Neural net aided detection of astronomical periodicities in geologic records

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

1 Dipartimento di Informatica ed Applicazioni, Università di Salerno
2 Osservatorio Astronomico di Capodimonte, Napoli
3 Istituto di Ricerca GEOMARE sud, CNR, Napoli
4 Dipartimento di Scienze della Terra, Università di Napoli "Federico II"
5 Dipartimento di Fisica Teorica, Università di Salerno

Abstract

Astronomically controlled variations in the climate induce cyclic trends in the sedimentary process and record (Milankovitch periodicity). 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 how to recognize periodicities in the signal. The application to two sequences of shallow water carbonate deposits from the Cretaceous of Southern Italy has shown this approach to be particularly effective, confirming the existence of Milankovitch-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

Periodic changes in climate, induced by astronomically induced variations in the distribution of solar energy over the Earth have been recognized as influencing the production of carbonate sediments, which are mostly formed by marine organisms and by their activity [1-4]. On the other hand, these periodic signals may be recorded in and have been extracted from a variety of other non-carbonate deposits originated in different sedimentary environments, from continental to deep marine [1,4].

The whole topic presents two incongruent aspects. The availability of detailed dynamical models of the Earth-Moon-Sun system, as well as of the Solar System as a whole, together with reliable algorithms, has led to very accurate predictions of the secular variations in the Earth's orbital parameters [5]; while the signature left by these periodic modulations in the stratigraphic record may appear *"extremely imperfect, unreliable, noisy and poorly time calibrated"* [3]. This lack of precision in the sedimentary records is the outcome of a complex process driven by a number of variables of which climate (together with sediment supply, global sea level change, subsidence, biologic activity and so on) is just one. An extensive review of these processes can be found in a recently published volume of proceedings [3], as well as in other recent publications [1,4,6]. In addition, we must consider that, if the time series analyzed are acquired with not very dense sampling, many of the high-frequency cycles may be lost in the rock record [3,4].

For these reasons high frequency ciclicity has been more and more studied in modern stratigraphy for its potential predictive value, in order to tie sedimentary processes to absolute time. In spite of the pioneeristic work of Croll and Gilbert [7,8] at the turn of the last century, it has been only in the past 15 years that an increasing interest has developed and the new term of *"cyclostratigraphy"* has become established, stressing the genetic connection of the orbital perturbations induced by the variability of Earth's thermal budget (Milankovitch cyclicity) with the climate, sedimentation and biosphere [2]. The vast majority of works in ciclostratigraphy have been devoted to the pelagic realm; that is, where the biostratigraphy gives the best time definition, allowing for the detected cyclicities to be framed into the Milankovitch periodicity [2-4]. More recently, a number of authors have focused their attention on shallow marine carbonates, which seem to offer the best conditions for the detection of the shortest period components (see, e.g. [9-13]) even though, for these types of data, the chance of gaps punctuating the record may be high (e.g. in the peritidal domain).

In the Cretaceous sequences of the Southern Apennines more than 700 m of well bedded carbonate platform deposits, spanning from Hauterivian to Coniacian, have been analyzed on a centimetric scale. In these shallow water sequences three main rock types alternate and the relative lithofacies associations qualify them as subtidal, tidal-supratidal and storm controlled deposits [9-12], [14]. The cycles are aggradationally stacked, shallow upwards and often capped by ephemeral emersion related features. Environmental changes expressed by texture or lithofacies variations, plotted at cm-scale on the aggradation axis show a hierarchy of 3 orders of cycles [15].

A package of spectral analysis based on analytical techniques, originally tailored for astronomical applications, has been used to search for possible periodicities [16]. This package includes preprocessing and spectral analysis of the signal; it deals with unevenly spaced data; it allows to discriminate between meaningful frequencies and spurious ones on a physical basis and, finally, reduces the effects of aliasing bias in the data.

The duration of cycles was calculated for about 450 m of well bedded sequences by comparing the Relative Ratio Sets (RRS) of the recurrence of sedimentary features (expressed in cm) with the RRS of orbital parameters (Precession, Obliquity, Eccentricity) calculated by Berger [17,5] for the Cretaceous (expressed in years). The two ratio sets show, for every sequence analyzed, a very good linear correlation (r >0.99) suggesting that Cretaceous carbonate platform strata of Southern Italy have an allocyclic organization (not inherent to the sediment deposition [11,12,14]), and that the relative time scale falls in the Milankovitch periodicities. Furthermore, spectral analysis of palaeomagnetic data from a 88 m thick core, drilled in lower Cretaceous shallow marine carbonate deposits from Mount Raggeto in Southern Italy, also demonstrates cyclic recurrence of the remanent magnetization values (declination, inclination, intensity) [18,19]. These periodicities are also linked to the orbital parameters with r >0.99.

We discuss here a new approach based on neural networks, to identify meaningful

Figura 1: Textural signal extracted from Monte Raggeto sequence

Figura 2: Lithofacies signal extracted from Monte Tobenna sequence

periodicities, a tool which has proved particularly effective in speech processing [20] and in pattern recognition [21]. In Sect. 2, the data are presented together with a short summary of the classical "power spectrum" approach to the search for periodicities in stratigraphic records. The general background of the neural network implementation and training is introduced in Sect. 3, data reduction and analysis are detailed in Sect. 4, while results and concluding remarks are presented in Sect 5. Finally, in the Appendix A we provide the reader with some details on the training of the neural network.

2 The data

The data used for the present analysis refer to two stratigraphic sequences observed along exposed open air quarries near Mount Raggeto ([15], [11] and Mount Tobenna [14], located near Caserta and Salerno, Southern Italy, respectively. The geological characterization of these sequences has been discussed at length elsewhere [15,14] and we refer to these papers for further details. We want, however, to stress a few points:

- The data were obtained by identifying/sampling the rocks directly from the outcrop at a centimetric scale and the lithofacies where determined by supplementary examination using thin sections and acetate peels;
- Both sequences refer to carbonate strata formed in shallow water environments, within the photic zone, at a depth never exceeding -on the average- a few or few tens of meters and, therefore, in environments very sensitive to global sea level oscillations (eustasy);
- The sedimentation rates in these environments are usually fairly high and, therefore, sampling at a centimetre scale provides a good time resolution.

3 Methodology

Neural nets are powerful tools in dealing with:
i) function approximation; it has been shown that they are universal function approximators [22], [23];
ii) classification and clustering: a neural net can learn from examples to classify input patterns in a supervised (cf. [24], [25]) or unsupervised (cf. [26], [27] manner.
In particular their use is recommended when:

- there is no good algorithm to solve the problem;
- input data are incomplete, noisy, and not directly understandable;
- applications are data intensive, i.e. more data then computations;

Table 1: Codes used for the Mount Raggeto and Mount Tobenna deposits.

Mount Raggeto Textures	Mount Tobenna Lithofacies	Codes
bindstone	"Stromatolitic" and "loferitic" boundstones (criptalgal laminites)	0
mudstone	Aeolissaccus sp. and Thaumatoporella sp., wack and wack-packstone, locally with burrows	0.1
wackestone	Large Miliolid pack and pack-wack with Thaumatoporella sp. and Ostracods	0.2
packstone	Wack, wack-pack and pack thin levels with criptalgal laminites	0.3
grainstone	Miliolid wack and wack-pack with benthonic foraminifers, Thaumatoporella sp., Aelisaccus sp., Dyciclina sp., Ostracods and micritized grains	0.4
rudstone	Oncoid pack, pack-wack and grains with Dyciclina sp., benthonic foraminifers and small planktonics	0.5
floatstone	Foraminifers wack-pack and pack with small planktonics	0.6
undetermined	Foraminifer wack and mud-wack, with small planktonics and with bioturbations	0.7
	Radiolitid floatstone with bioclastic matrix	0.8
clay levels	clay levels	0.9

- the procedure needs to be repeated many times;
- many examples for "training" are available.

All these requirements are fulfilled by our data thus making neural nets an ideal tool to deal with periodicity recognition in stratigraphic signals.

3.1 The neural net

A neural net is a computational structure made by many processing elements (units) – the neurons– operating in parallel. These neurons are generally organized into clusters or layers. They are grouped in "input", "output" and "hidden" (i.e. those units which are neither input nor output) layers. Three fundamental elements characterize any neural nets:

- The net structure or topology, i.e. the way the layers are linked;
- The activation function, which represents the answer of a neuron to the input stimuli;
- The learning algorithm.

In the present case, the adopted net model is the well known "Multilayer Perceptron" [25] shown in Fig. 3.1. It consists of one input layer, one output layer and one or two hidden layers. Each neuron of a given layer is connected to all the neurons of the next one.

Our model is syncronous: at each time every neuron receives as input the weighted sum of the input patterns and/or of the other neuron outputs, as shown in the following equation:

$$O_k = f[\sum_n W_{kn} O_n - B_k] \tag{1}$$

Figura 3: The topological structure of the Multilayer Perceptron Neural network used for our experiment

Figura 4: The Sigmoidal function

Where W_{kn} is the weight associated to the link from neuron n to neuron k, O_n is the output of neuron n or of the n-th input and B_k is the neuron threshold, generally called bias. Thus the neuron output is a continuous and derivable function of its net input, with values comprised in the $[0, 1]$ range. This function f is not linear for the hidden units. For our experiment we choose the sigmoidal function, shown in Fig. 3.1, which looks like:

$$f(x) = 1/(1 + e^{-x}) \tag{2}$$

The training procedure was the so called "back propagation" which works as follows: the first pattern is presented to the input neurons and then the net gives its output. If it is not equal to the desired output pattern, we compute the difference (error) between these two values and change the weights in order to minimize it. Then we propagate the information to the previous layer, changing the weights. For a detailed explanation of the algorithm see Appendix A. We repeat these operations for each input pattern until we minimize the mean square error of the system. Given the p-th pattern in input, the error E_p is:

$$E_p = 1/2 \times \sum_j (t_{pj} - O_{pj})^2 \tag{3}$$

where t_{pj} is the p^{-th} desired output value and O_{pj} is the output of the corresponding neuron.

3.2 Application to geologic record

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 biologic [26] as well as in stratigraphic signals [1,3,4]. Several authors have suggested that these signals might be linked to the secular variations of the Earth's orbit [5,3,17]. 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 the Mesozoic, by Berger and his collaborators on the basis of N-body simulations of the solar system [5,17], 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 strata considered. 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.

4 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 [11,12] and leads to a rough estimate of S_{est}.

The Scargle algorithm has exactly the required characteristics. To be more precise:

- it deals with unevenly spaced data; this characteristic is important because the rebinning of unevenly sampled data into equally spaced bins and following computation of a conventional periodogram may alter the spectrum and the significance of a periodic signal;
- it allows discrimination between meaningful frequencies and spurious ones on a physical basis; this is made by evaluating the false alarm probability [16], that is an estimate of the significance of the height of a peak in the power spectrum; this is possible only if the periodogram has an exponential probability distribution function [16], achieved by normalizing the periodogram with respect to σ^2, i. e. the variance of the data;
- it reduces the effects of aliasing bias in the data.

To be more detailed, let $X(t_i)$, with $i = 1, 2, \ldots, N_0$, be the time series data, sampled at discrete intervals, and $P_X(\omega)$ the corresponding power spectrum at frequency ω. The equation to calculate $P_X(\omega)$ is:

$$P_X(\omega) = \frac{1}{2}\left[\frac{\left[\sum_j X_j cos\omega(t_j-\tau)\right]^2}{\sum_j cos^2\omega(t_j-\tau)} + \frac{\left[\sum_j X_j sin\omega(t_j-\tau)\right]^2}{\sum_j sin^2\omega(t_j-\tau)}\right] \tag{4}$$

where τ is defined as follows:

$$tan(2\omega\tau) = \frac{\left(\sum_j sin2\omega t_j\right)}{\left(\sum_j cos2\omega t_j\right)} \tag{5}$$

and represents a term that makes the periodogram invariant to a shift in the zero point of the time scale.

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

The stratigraphic record obtained for the exposed sequences in Mount Tobenna and Mount Raggeto was processed according to the previous section, leading to the identification of the spatial periodicities shown in Figs. 5 and 6 and listed in Table 2 and shown in.

Figura 5: Periodogram of the Mount Tobenna lythofacies. The signal has been corrected for spectral leakage. The horizontal solved line gives the 90% confidence level.

Figura 6: Periodogram of the Mount Raggeto texture. The signal has been corrected for spectral leakage. The horizontal solved line gives the 90% confidence level.

4.1.1 Pre-processing

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 Milankovitch-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 Milankovitch periodicities;
- additive noise;
- the same length of the real stratigraphic sequence;
- a similar square-like shape;
- the same sampling rate as the real data set;

Due to the different domains of the real (space) and simulated (time) series it was first necessary to find the conversion factor (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 Milankovitch frequency, as in [5,17], thus obtaining an estimate of α. This value gave the sampling rate for the synthetic signals, which were therefore degraded by adding randomly generated additive noise. Fig. 7 shows an example of synthetic signal, containing all six Milankovic' frequencies, created by our procedure.

4.1.2 Creation of training and test sets

The net model used for these experiments is not shift invariant and, on the other hand, it is impossible to submit the whole signal to the net at the same time. In order to by-pass this problem, we subdivided every simulated data set into smaller subsequences long enough to contain all the information. Every subsequence is slightly overlapping the next in order to simulate a sort of continuity in submitting the signals to the net. For each sub-sequence, a Fourier transform allows identification of the six amplitude values which represent the pattern input; so every signal is submitted to the network as a pattern set. We give the

Figura 7: Synthetic signal obtained for the Mount Tobenna texture record by combining the six Milankovic' cycles and assuming $S_{est} = 371\ cm \times kyz^{-1}$

Table 2: 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 trains
Tobenna	12-6-6-6	77.000	0.013117
Raggeto	12-6-6-6	95.000	0.009991

Table 3: Test phase. The names F_i, $i = 1, \ldots 6$ 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

net two consecutive overlapped sequences in order to obtain the best experimental results. In this way we obtained the 12 input values of the single input pattern. Each time series is composed of 10 and 13 consecutive overlapped sequences for the Mount Tobenna and Mount Raggeto experiments, respectively (depending on the length of the related original stratigraphic signal); we obtained 9 and 12 input patterns for each synthetic signal related to the Tobenna and Raggeto experiments, respectively. We have 64 (2^6) different synthetic signals, containing all the combinations of the 6 Milankovitch periodicities plus an additive pseudo-random noise. The training set is composed of 30 synthetic signals, chosen from the 64 available, in such a way that samples of the 6 fundamental periodicities and their combinations are present. For each time series the output error is evaluated as the mean value of the output of the 9 or 12 consecutive input patterns. The error is back-propagated after each submission of all 30 synthetic signals to the net.

5 Results

We used the pre-processing strategies, described in Section 4.1.1, to evaluate the sedimentation rate and the sampling rate for the sequences related to Mount Tobenna and Mount Raggeto. Once the sampling rate was fixed, a complete series of synthetic signals covering all possible combinations of the six main Milankovitch frequencies was produced.

Table 4: Test phase. Level of confidence for positive detections for the six Milankovic' frequencies. A value higher than 0.50 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

We the apply the procedure described in Section 4.1.2 to create the training and test set. After the training phase (see also Appendix A) we used a subset of synthetic signals not used in the training phase in order to evaluate the network performance. With these signals we obtained result with 100% of correct output detection. (with an output value greater than 0.9 for the output nodes related to the Milankovitch cycles present into the signals). Finally, we used the Tobenna and Raggeto sequences in order to evaluate the network behavior with real signals. The results are illustrated in Tables 3, 4 and 5. We used a threshold on the output equal to 0.5 for the detection. Therefore this procedure allowed us to recognize the Milankovitch periodicities present in the stratigraphic signals studied. These results are confirmed by spectral analysis developed with the Scargle algorithm and the evidence from the outcrop (thickness of the cycle) [14,11]. The neural approach here illustrated is completely automatic and permits the user to detect Milankovitch periodicities without complex interpretations of the power spectrum.

6 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. In fact, instead of searching blindly 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/no base by training the net to recognize signals of a given frequency. The time and effort required for 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 Cretaceous sequences, formed by well bedded carbonate rocks of shallow water origin. The data processed were obtained with centimetric accuracy along two exposed sections of Mount Raggeto and Mount Tobenna, which had been studied previously by the authors with traditional PST's [11], [14]. The results of the present study confirm the existence of Milankovitch type signals in both sequences and in particular:

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

In forthcoming papers we shall discuss the application of wavelets to the pre-processing of the data as well as the effect of other net architectures on the accuracy of the data.

Appendix A. Neural Net Learning Procedure

In this section we discuss in more detail the neural network learning procedure. First, the error back propagation algorithm (EBP) is presented and then we discuss the learning procedure, explaining the meaning of its main parameters. The algorithm is shown below.

As already stated, it is necessary that the units have non linear threshold functions that are continuously differentiable, such as, for instance, the sigmoidal function (Eq. 2):

$$O_k = f[\sum_n w_{kn} O_n - B_k] \tag{6}$$

Algorithm

1. Let us assign small random values to weight and thresholds and a fixed small value to the error confidence;
2. input patterns $X_p = x_{p0}, ...x_{pn-1}$, where p is the index of the p^{-th} pattern and n is the number of the input nodes;
3. for the first hidden layer calculate the actual output $O_{pj} = f[\sum i_{0n-1} w_{ji} x_i - B_j]$ and pass it as input to the next layer and so on for the next layers. f is the activation function. At the output layer calculate the actual error by a comparison between the actual net output and relative corrected output;
4. after an entire cycle of input pattern presentations, calculate the whole error function: $E = 1/2 \times \sum_j (t_{pj} - O_{pj})^2$ on all input patterns p.
5. Adapt weights starting from the output layer and going backwards. Calculate: $W_{ji}(new) = W_{ji}(old) + \eta \times \delta_{pj} O_{pj} + \alpha \times (W_{ji}(new) - W_{ji}(old))$, where η and α are constants called, respectively, learning rate and momentum and: $\delta_{pj} = f'(O_{pj})(t_{pj} - O_{pj})$ for output units $\delta_{pj} = f'(O_{pj})(\sum_k W_{kj} \times \delta_{pk})$ for hidden units, are the error terms for pattern p on node j.
6. If $E > error\ confidence$ then *go to step* 2 else *stop.*

The meaning of the hidden units into the net can be summarized as feature detectors. It can be seen as a recording of the inputs so that a neural net can learn the mapping of input patterns to output patterns. This recording or internal representation, in terms of the number of hidden units, is critical to the behavior of the network. In fact, to form internal representations of any input pattern, and to produce the correct response of the output units, requires enough hidden units. The weight updating phase in the above algorithm, well known in literature as "generalized delta rule", provides a method for teaching multilayer perceptron networks, producing the necessary internal representation of the hidden nodes. It must be recalled that the learning procedure is not guaranteed to produce convergence. It is possible for the network to fall into a so called *local minimum* of the error function E. The correct learning situation is, instead, the falling of the error function E to a so called *absolute minimum.* In fact, the error function and its related network weight can be depicted as an energy surface that is a rippling landscape of hills and valleys, wells and mountains, with points of minimum energy corresponding to the wells and maximum energy found on the peaks [24].

In this context the generalized delta rule aims to minimize the error function E by adjusting the weights so that they correspond to those at which the energy surface is the lowest. It does it by a method known as *Gradiend Descent*, where the energy function is calculated and changes are made in the steepest downward direction. This is guaranteed to find a solution in cases where the energy landscape is simple. Each possible solution is represented as a hollow in the landscape. This "basins of attractions" represent the solution to the values of the weights that produce the correct output. Three layers of neurons can therefore form arbitrarily complex shapes, and are capable of separating any classes, as it is stated by the well known Kolmogorov Theorem [24]. One of the main criticisms against the multilayer perceptron is that it requires many presentations of the input pattern set and a repetition of the corresponding calculations, before the network is able to settle into a stable solution. The method of gradient descent is slow to converge due to the complexity of the error surface. The addition of the two terms, learning rate and momentum, often speeds up convergence. The η term is a measure of the degree of influence of the updating weights formula of the error term, whereas the α term determines the influence of the past history of weights changes into the same formula. Determining the optimal choices for all these parameters must be achieved on the basis of a trial and error approach.

Acknowledgements

This work was partly sponsored by the institute Geomare Sud, CNR (National Research Council), Naples; by the Italian Space Agency (ASI) and the Osservatorio Astronomico di Capodimonte (OAC), Naples; and the Department of Earth Sciences, University of Naples. Financial support is also acknowledged from the Ministry of University and Scientific Technical Research (MURST 40%), through the Salerno University (Unit of the INFM). Finally, we thank the Caianiello International Institute for Advanced Scientific Studies (IIASS), Vietri sul Mare, Salerno, for supporting this research through a CNR grant.

References

(1) G. Einsele, W. Ricken, A. Seilacher, Cycles and events in stratigraphy, 955 pp., Springer-Verlag, Berlin, 1991.

(2) A.G. Fischer, P.L. de Boer, I. Premoli Silva,