SEA SURFACE-WATER TEMPERATURE AND ISOTOPIC RECONSTRUCTIONS FROM NANNOPLANKTON DATA USING ARTIFICIAL NEURAL NETWORKS

Marco Pozzi, Björn A. Malmgren and Simonetta Monechi

Marco Pozzi and Simonetta Monechi. Dipartimento di Scienze della Terra, Università di Firenze, via La Pira 4, 50121 Firenze, Italy.

Björn A. Malmgren. Department of Earth Sciences-Marine Geology, Earth Sciences Centre, University of Göteborg, Box 460, SE-405 30 Göteborg, Sweden.

ABSTRACT

Artificial neural networks (ANN) are computer systems that differ from other pattern recognition methods in their ability to 'learn' one or more target variables from a set of input variables. These systems learn by self-adjusting a set of parameters to minimize the error between the desired output and network output.

To explore the potential of artificial neural networks for predictions of paleo-oceanographic parameters from relative abundances of calcareous nannoplankton species we analysed observations taken from the literature for (1) the prediction of sea surface-water temperature (SST) in samples from offshore southern California, and (2) the prediction of oxygen isotopic values in a Quaternary core from the eastern Mediterranean.

We employed a back propagation (BP) neural network to assess the ability of the network to predict SST and oxygen isotopic values. Each of the data sets was divided into five random training and test sets to assess the stability of the error rate estimates. For the California Bight samples we obtained an average Root-Mean-Square-Error of Prediction (RMSEP) in the test sets of 0.68, implying that an unknown SST can be predicted with a precision of ± 0.68°C. In the Mediterranean samples the average RMSEP in the test sets was 0.64; hence an unknown oxygen isotope value can be predicted with a precision of ± 0.64 $\delta^{18}O$‰ vs. PDB.

ANN techniques can be seen as a complementary tool to more conventional approaches to paleoceanographic data analysis. Such techniques hold great potential for making predictions of various types of variables from paleontological data.

INTRODUCTION

The large amounts of nannoplankton data collected worldwide require sophisticated and powerful quantitative methods of analysis for the extraction of the optimum amount of information on species

distributions through which paleoceanographic reconstructions might be based. One of the analytical methods now available is the artificial neural network (ANN) approach for data analysis and forecasting.

Artificial neural networks are computer systems that have the ability to 'learn' a set of output, or target, variables from a set of input variables. Neural networks have been employed in many disciplines for problems of prediction, classification, or control of various processes. This remarkable success can be attributed to a few key factors.

Neural networks are numerically sophisticated modelling techniques, capable of modelling extremely complex functions. For many years, linear modelling has been the most commonly used modelling technique in paleo-oceanography because linear models have well-known optimization strategies. In cases where the linear approximation was not valid (which was frequently the case in nature) the models suffered accordingly. Neural networks analyse the dimensionality problem. This distinguishes them from attempts at modelling nonlinear functions with large numbers of variables;

Neural networks learn by example. The neural network user gathers representative data, and then invokes training algorithms to automatically learn the structure of the data. In addition, neural computers have the ability to learn from experience in order to improve their performance, and can adapt their behaviour to new and changing environments. The level of user knowledge needed to successfully apply neural networks is much lower than would be the case using more traditional statistical methods;

Neural networks tend to be more reliable and versatile than conventional data analysis and modelling methodologies. They have the ability to cope well with incomplete (e.g., small samples), 'fuzzy' data, and can deal with previously unspecified or unencountered situations. Neural networks are very tolerant to analytic faults. This contrasts with conventional systems, where the failure of one component of the analysis usually means the failure of the entire analytic system.

In the earth sciences, neural networks have been applied to problems of well-log interpretation (Baldwin et al. 1989; Baldwin et al. 1990; Rogers et al. 1992), for the identification of linear features in satellite imagery (Penn et al. 1993); for geophysical inversion problems (Raiche 1991), for the correlation of volcanic ash layers (Malmgren and Nordlund 1996); and for the establishment of present-day climatic zonation in Puerto Rico (Malmgren and Winter 1999). Malmgren and Nordlund (1997) applied a BP neural network in an attempt to predict modern sea surface-water temperatures (SST) from relative abundances of planktonic foraminifer species in the southern Indian Ocean. That study showed the BP technique to be able to reproduce the SST data more faithfully than conventional techniques such as the Imbrie-Kipp Transfer Functions (Imbrie and Kipp 1971) and Modern Analog Technique (Hutson 1979). These results indicated that late Quaternary summer and winter SST's may be predicted with a precision of ±0.7-0.8°C using the trained BP network.

With data gleaned from the literature we here make a first attempt at testing the applicability of ANN to reconstruct SST and stable isotope data from relative abundance data of calcareous nannoplankton species.

MATERIAL AND METHODS

The Nannoplankton Data Sets

We used two datasets from published studies. The first dataset concerns seasonal changes in coccolithophore cell den-

Table 1. Percent abundances of the coccolithophores in the California Bight samples analysed using SEM.

Samples	E. huxleyi %	G. oceanica %	U. sibogae %	R. longistylis %	* Syracos. spp.	SST ×C
10/11/91	70.21	1.06	0	7.45	0	18.73
12/16/91	47.12	0.52	0	0	2.61	14.34
1/29/92	36.94	2.24	4.85	23.13	6.34	14.99
3/16/92	77.63	1.97	1.32	2.96	5.92	15.13
3/23/92	59.53	4.65	2.79	4.19	5.14	15.71
4/17/92	57.77	0.49	17.96	4.85	7.78	17.9
5/1/92	62.87	0	7.92	0.99	6.94	18.29
5/11/92	58.94	1.45	8.21	0	5.31	18.5
5/15/92	40.69	0	10.39	3.03	3.04	18.44
6/5/92	63.89	0.46	0.01	0.46	3.69	17.16
6/12/92	63.29	0.97	1.93	3.86	3.87	18.18
7/2/92	43.48	2.17	2.17	0	0	17.89
7/10/92	80.21	2.08	2.08	0	2.08	19.93
7/20/92	62.5	3.29	9.21	0	1.32	21.33

*Syracosphaera spp. consists of: **S. anthos, S. binodata, S. corolla, S. corrugis, S. epigrosa, S. histrica, S. nana, S. nodosa. S. orbiculos, S. pirus, S. prolongata, S. pulchra, S. rotula,** S. sp. 1, S. sp.2.

sity and species composition, and the relationship between species composition and SST in the San Pedro Basin, Southern California Bight (Ziveri et al. 1995b). The second dataset utilizes data from a paleoceanographic study based on nannofossil assemblages from the eastern Mediterranean (Castradori 1992, 1993).

The Californian dataset consists of nannoplankton samples collected at 14 stations and SST measurements simultaneously obtained during several expeditions between October 11, 1991 and July 20, 1992. Several studies in coastal upwelling environments (Winter 1985; Mitchell-Innes and Winter 1987; Klejine et al. 1989; Giraudeau et al. 1993; Ziveri et al. 1995a; and Thunell et al. 1996) have shown that coccolithophores can be important contributors to the total phytoplankton population as well as that their

distributions are related to variation in nutrient compositions and SST. This region, influenced by the El Niño-Southern Oscillation (ENSO), is marked by the following oceanographic changes from nearshore to deeper waters: deepening of the thermocline, warming of the surface-water mixed layer, reduced coastal upwelling, enhanced onshore transport of low salinity waters (Ziveri et al. 1995b). Changing coccolithophore assemblages—and more complex relationships between cell density, abundance variations, and SST—reflect these changes.

From the total association we considered, the most abundant and continuous species (Table 1)—**Emiliania huxleyi**—was found to be the dominant species accounting for approximately 60% of the assemblages (with a maximum of 80% in July 1992). High percentages of **E. hux-**

leyi indicate a fairly well-stratified water column (Brand 1994); blooms of **E. huxleyi** sometimes follow large blooms of diatoms (Probyn 1993). **Umbilicosphaera sibogae** was the second most-abundant species; this species has an ecological preference for warm oligotrophic waters (Okada and McIntyre 1979; Giraudeau 1992). Other abundant species recognized in this area include*:* **Rhabdosphaera longistylis**, which is abundant in temperate waters (McIntyre et al. 1970; Gaarder, 1970) with no specific sensitivity to high nutrient levels (Brand 1994); **Gephyrocapsa oceanica**, which is adapted to warm and nutrient-rich coastal waters (MacIntyre et al. 1970; Okada and Honjo 1975; Okada and McIntyre 1977), and **Syracosphaera** spp., mainly represented by **Syracosphaera pulchra***,* which has a preference for warm temperate waters with low nutrients (McIntyre et al. 1970; Roth and Coulbourn 1982; Pujos 1992).

The second data set is derived from a quantitative analysis of calcareous nannofossils in four eastern Mediterranean cores. We selected one core (Ban82-15PC), located near the Erodoto Abyssal Plain (32°42' N, 26°44' E), for which both oxygen isotope and nannofossil data were available (Parisi 1987). This core spans the last 500,000 years and contains seven sapropel layers. The average sedimentation rate is 2.22 cm/k.y. The oxygen isotope values range from –4.5‰ to 2.6‰ with a maximum glacial-interglacial change of approximately 6.7‰ at Termination II; the highest enrichment lies at the Pleistocene-Holocene boundary (Parisi 1987). The nannoplankton assemblage alters from dominance of E. huxleyi in the upper part (**Emiliania huxleyi** Acme Zone) to dominance of small **Gephyrocapsa** and **G. oceanica** in the lower part. Other species, like **Helicosphaera carteri***,* **Syracosphaera** spp.*,* **Gephyrocapsa**

caribbeanica*,* were also recognized in all samples but with lower relative abundances. On the whole, the nannofossil species abundance reflects a temperate oligotrophic water association. We selected the eight most abundant species (Table 2). Often this group of species represents 90% of the association.

THE NEURAL NETWORK STRUCTURE

Among the several ANN learning algorithms available, BP is the most popular (Figure 1A). The BP network consists of an interconnected series of layers, each containing a number of processing units called 'neurons' (Figure 1B). The basic steps in our application of the BP network are: (1) training of the network on the basis of a number of training sets, and (2) assessment of the performance of the network by computations of the error rates in the test sets (details are given below).

The main steps in the neural network procedure are as follows: the input signals (e.g., nannoplankton relative abundances), enter the network via the input layer; each neuron in the network processes the input data, with the resulting values steadily seeping through the network layer by layer, until a result is generated in the output layer. The output of the network is then compared with the actual output value. This results in an error value, representing the sum-squared difference between the actual and predicted input. In order to minimize this error value all the weights at each connection of the network are gradually adjusted in the direction of the steepest descent with respect to the error (the steepest-descent algorithm). This process involves working backwards from the output layer, through the hidden layer, and back to the input layer, until the specified error limit is reached. Fine-tuning the weights in this way has the effect of 'teaching' the network how to produce the

Table 2. Percent abundances of the nannofossils in Ban82-15PC core analysed using MOP (optical polarized microscope).

| Samples	G. oceanica	E. huxleyi	G. caribbeanica	small Gephiroc. spp.	H. carteri	Syrac. spp.	C. leptoporus	Rhabd. spp.	$\delta^{18}O$
10	0.7	79.7	1	1.3	0.5	5.7	0.3	2.9	-1.67
15	1	76.7	0.3	5.3	4	4.2	0	3.7	-3.02
22	1.7	69	1	4.7	0.7	4.5	0	3.8	-0.48
30	1.3	66.7	2.3	5.3	0.7	2.9	0.2	2.4	2.63
39	0.3	37	1.3	5.3	3	4.7	0	3.3	2.35
50	1	26.7	3	7.7	1.7	6.3	0	2.7	1.96
81	1	30	1	9.3	3	5.7	0	3	1.57
100	1.3	54.3	1.7	3.3	1.7	2.7	0	2	2.22
120	1	42	4.3	12	2.3	5.3	0	3	0.63
139	2.3	34.3	1	22	3.3	6.7	0.3	2.3	1.83
165	9	4.3	25	35.7	0.5	3.6	0.3	2.1	-0.65
185	14.7	4.3	4	55	0.2	3.3	0	3.1	-0.9
202	6.7	22	4.7	19.7	1.3	4.7	0.3	2.7	-1.04
219	13.7	5	0.7	40.3	9.7	4.7	0.3	5.7	-3.37
221	33.7	4.7	0.3	42.7	5	4.5	0.6	3.5	-0.22
223	42.3	3.7	2.3	22.3	2.3	8.5	0.3	4.7	-1.17
247	4.7	7.3	10	57.3	0.8	5.6	0.4	1.2	-1.97
270	16.7	4.3	1.7	37	17	7	1.7	9.7	0.27
289	0.7	4.7	3	29	2.3	4.7	0.7	7.7	-1.14
304	9.7	13	0.3	52.3	3	8.4	1	2.2	-1.42
314	19	9.7	3	43	7.9	3	7.3	1.3	-2.39
329	41.7	6.7	2.7	31.3	3.5	1.6	1.6	2.8	2.23
349	1	2.7	5	74.3	0.9	1.2	0.2	1.2	0.52
360	19.7	1.7	1.3	52.7	1.7	5.2	0.2	3.9	-3.57
365	12.3	3.7	4	51.3	13.2	2.6	3.7	0.9	0.22
371	7	4.3	0.3	61.7	8.3	5.3	4.8	1.9	1.34
380	4.3	1	1	64.3	1.5	2.6	1.5	2.3	-0.42
391	12	1	0.7	63	7.5	4.7	7.2	1.9	-0.05
398	29.3	1	0.3	51.7	9.5	4.2	1.9	1.1	-0.7

Table 2 (continued).

Samples	G. oceanica	E. huxleyi	G. caribbeanica	small Gephiroc. spp.	H. carteri	Syrac. spp.	C. leptoporus	Rhabd. spp.	$\delta^{18}O$
412	30.7	1	0.3	36.7	20.7	6.6	1.6	0.9	-0.41
439	20	4	8.7	32.7	3	4	0.3	1.7	1.05
460	24.7	1	26	19.3	8.4	5.5	0.6	0.3	-0.8
485	11.3	0.3	8	63.3	2.9	2.4	1.2	1.5	-0.9
510	1.3	0	7.7	61.3	0.6	2.1	0	0.3	-0.98
516	19.7	0	1	64	2.9	3.8	0.8	4	-3.07
522	5	0	0.7	76.7	9	2.8	0.9	3.2	-1.38
529	4.3	0	0.3	80	5.7	4	1.4	2.6	-1.92
550	6.3	0	0.3	88.3	0.4	2.2	0.2	0.7	-1.41
575	33.7	0	14	31.3	0.6	4.4	0.2	3.4	0.28
598	12.7	0	31.3	36.7	0	2.7	0	1.5	1.42
625	52.3	0	16.3	23.7	0.3	2.8	0.1	1.7	0.57
646	40	0	29.3	22.3	0.1	3.2	0.1	1.8	0.26
666	25	0	31	31.3	0.3	2.5	0.4	1.5	-1.35
672	67.7	0	4.3	13	1.8	6.9	0.2	3	-2.53
674	31.7	0	47	15.3	1.1	1.5	0.2	0.9	-1.44
693	53	0	7.7	19.7	0.2	3.7	0.2	2	-1.59
726	15.3	0	43.3	34	0.4	2	0.1	1.5	-1.38
746	58.7	0	7	34.3	0.1	2.8	0	2.4	-0.64

'desired' output for a particular input. In this way the network 'learns'.

The last three steps described above usually have to be repeated a number of times until the error value is minimized (in the ideal case this error is zero). These steps may potentially involve many thousand training passes. This iterative process is the kernel of the back propagation algorithm. Finally, when the network has converged (meaning, reached a preset error limit), it will ideally be able to produce the correct output for each input. Once the network has been trained, it can be used to predict the output signals used in the training phase from new input signals.

In the analysis of the dataset from the California Bight, we used five neurons in the input layer (corresponding to the five species of nannoplankton used as input variables). In the Mediterranean dataset, we used eight neurons in the input layer. In both cases, the number of neurons in the output layer is one (corresponding to the SST and oxygen isotope data to be predicted; Figure 2). The software used was the NeuroGenetic Optimizer (NGO), version 2.6, from BioComp Systems, Inc. (http://www.biocompsystems.com).

This program automatically attempts either one or two hidden layers to find the optimum network. The program also
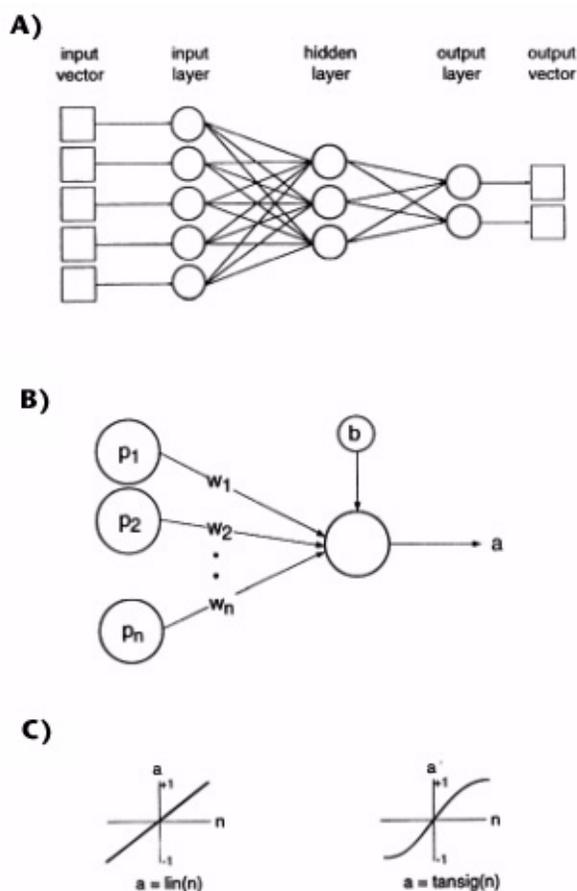
Figure 1. (A) Diagram showing the general architecture of a back propagation network. Each neuron in the hidden and output layers receives weighted signals from the neurons in the previous layer. (B) Diagram showing the elements of a single neuron in a BP network. In forward propagation, the incoming signals from the neurons of the previous layer (p) are multiplied by the weights of the connections (w) and summed. The bias (b) is then added, and the resulting sum is filtered through the transfer function (c), linear or sigmoidal, to produce the activity (a) of the neuron.

allows the number of network cycles to be specified prior to the start of the data processing. Each of these cycles utilizes a different network configuration. The cycles are divided into populations and generations that can both be varied. In addition, the program searches the best solution by varying the different number of neurons in each of the different layers. The NeuroGenetic Optimizer attempts different types of transfer functions within single neurons (linear, logarithmic and hyperbolic tangent) when performing genetic searches.

RESULTS

In the California Bight data 14 samples were available. Half of these samples have been used as the training set (seven observations). The remaining half constituted the test set (also seven observations). To make sure that the error rate is representative of the dataset, the network was run five times, each with 50% random training-set and 50% test-set members (the observations for the training sets are in italics in Table 1). The number of network configurations attempted was 600 (20 generations of 30 populations each). Best network configurations for the various partitions are shown in Table 3. The average Root-Mean-Square-Error of Prediction (RMSEP) in the test sets was 0.68, implying that an unknown SST can be predicted from the relative abundances of the five nannoplankton species used here with a precision of ± 0.68° C (Figure 3). Correlation coefficients between observed and predicted SSTs in the five test sets range between 0.80 and 0.98 (Figure 4).

For the Mediterranean data the number of samples was 48. We trained the network using 80% of the samples as the training set (39 observations) and the remaining 20% as the test set (nine observations). This subdivision of the training and test sets was performed randomly by the program. Again, five different random subdivisions of the original data set were run to generate an error estimate. Also in this case, the number of network configurations attempted was 600 (20 generations of 30 populations each). The configurations of the best networks are provided in Table 3. The average RMSEP in the test sets was 0.64, implying that an unknown oxygen isotope value can be predicted from the relative abundances of the selected nannofossil species in this
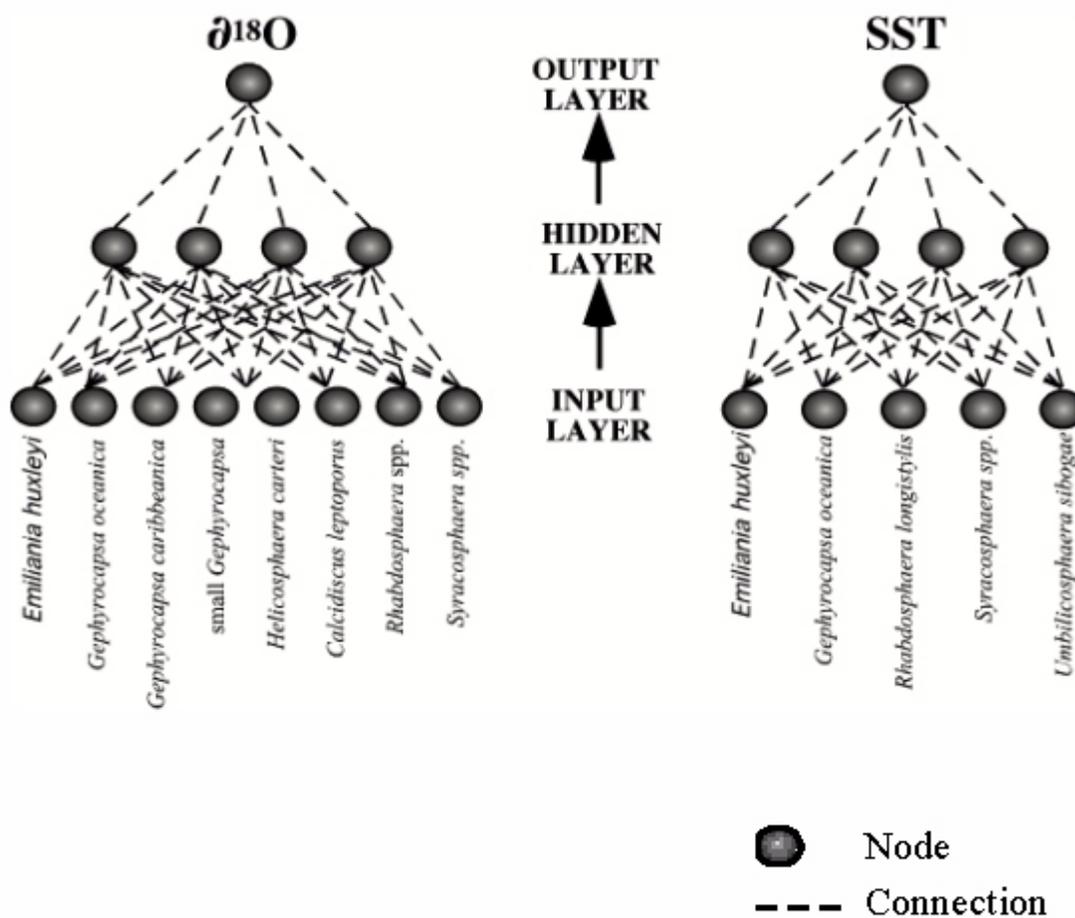
Figure 2. Diagram showing the network structure with five- and eight-input neurons, corresponding to the number of species of nannoplankton used as input signals in the Californian and Mediterranean studies, respectively. This is an example comprising one hidden layer with four neurons and an output layer with a single neuron, corresponding to the variable to be predicted.

core with a precision of ± 0.64 $\delta^{18}O‰$ PDB (Figure 5). Correlation coefficients between observed and predicted isotope data in the five test-sets range between 0.64 and 0.96 with an average value of 0.88 (Figure 6).

SUMMARY AND CONCLUSIONS

By means of the application of the ANN BP algorithm the input signal from calcareous nannoplankton assemblages was predicted in terms of two parameters, SST and $\delta^{18}O$. The same technique could be applied to predictions of any type of physio-chemical variables. The number of samples in both examples was small, but the error estimates obtained from the ANN suggest that even when using small samples highly reliable results can be obtained.

The ability to learn from the examples reveals the goal of the analysis architecture. The examples could also be represented by nonlinear and nonhomogeneous data. In the specific case of a nannoplankton assemblage, the examples could comprise absences of some species in some samples. In the case of the samples from the California Bight, the network has realized the best solution by choosing the samples in such a way as to have represented in the training session all the seasons comprising the time period between October 11, 1991 and

Table 3. Configurations of the best neural networks for each of the five different random partitions of the Californian and Mediterranean data sets. The output layer always contains a single neuron. 'Li' indicates linear transfer functions, 'Lo' logistic transfer functions and 'T' hyperbolic tangent transfer functions..

California data set			
Partition	First hidden layer	Second hidden layer	Output layer
1	12Lo, 3T, 13Li	2Lo, 156T	1Lo
2	9Lo, 12T, 8Li	5Lo, 12T, 4Li	1Lo
3	11Lo, 1Li	--	1Lo
4	7Lo, 8T	--	1Lo
5	4Lo, 5T, 6Li	--	1Lo
Mediterranean data set			
Partition	First hidden layer	Second hidden layer	Output layer
1	12Lo, 3T, 13Li	--	1Li
2	13Lo, 5T	--	1Lo
3	14Lo, 7T	--	1Lo
4	3Lo, 8T, 2Li	--	1Lo
5	31Lo	--	1Lo

July 20, 1992 (the italic word in Table 1). So, in this case, the training set includes the months of October, January, March, May, June, and July, whereas the test set consists of December, March, April, May, June, and July. This arrangement was made from the neural network automatically, without any information during the initialization phase. Thus, the ability of the network to learn from a small, random data set allows the organisation of the data in the correct order and with a correct interpretation.
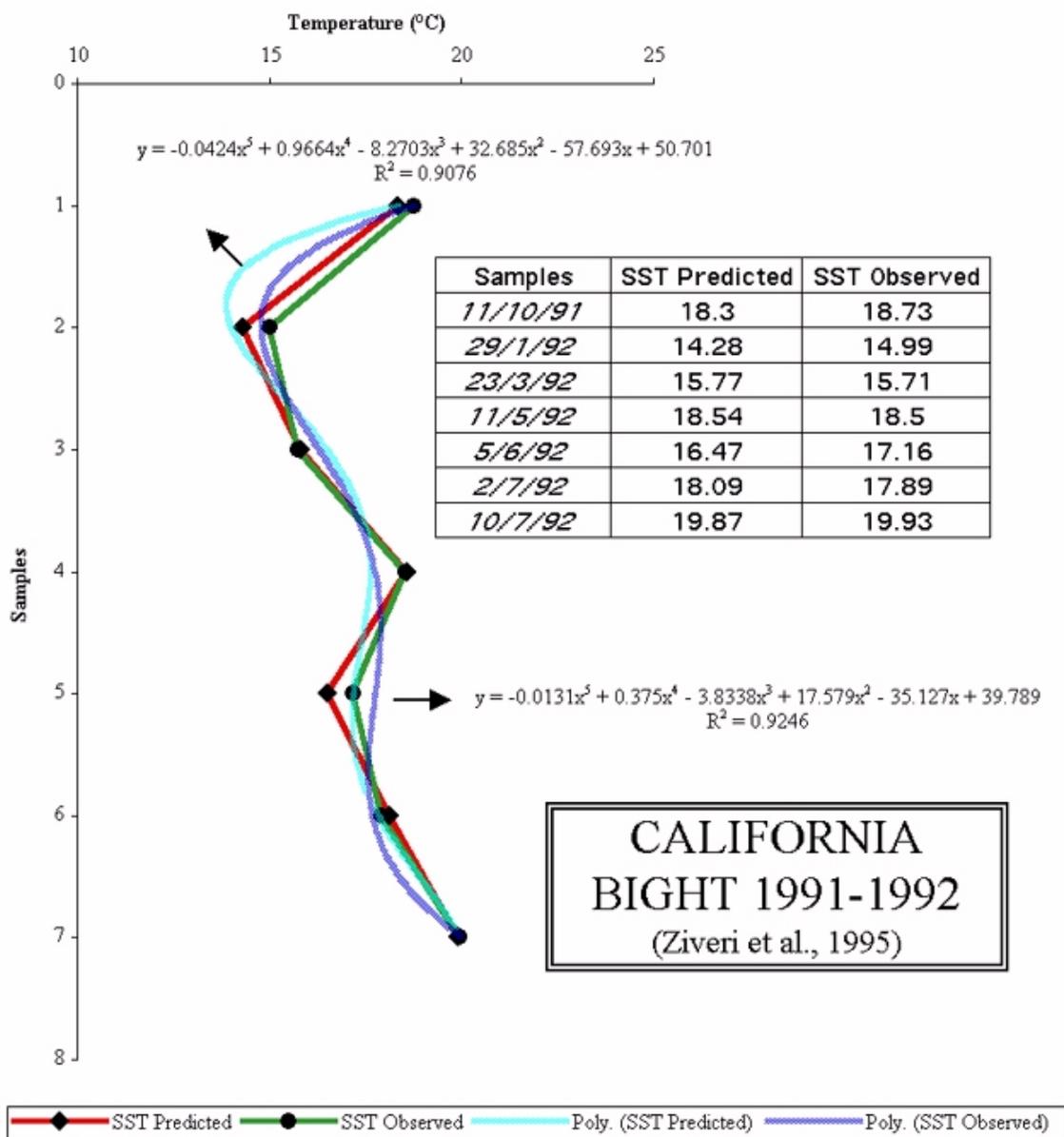
In the case of the Mediterranean Sea, the network has selected the best configuration based on training and test sets established by the program, and the final chart (Figure 5) shows a good relationship between predicted and observed values.

In particular, this chart documents a very similar pattern of change in observed and predicted $\delta^{18}O$ values with increasing core depth. This demonstrates the optimum choice of the neural network concerning the training test-set subdivision. In this case, there was a large number of 'zero' values in the distribution of E. huxleyi in the lower part of the core. However, this problem did not in any way harm the final result. The neural approach reveals a highly satisfactory result despite the occurrence of many 'zero' values as well as with consideration of the relatively few observations in this dataset.

The application of ANN to nannoplankton data could be potentially useful for studies of paleoproductivity and paleoclimate, paleobiogeographic patterns, and paleo-oceanography. For example, some areas of nannoplankton research where ANN might be used involve predictions of sea surface-water temperatures (SST), nutrient and micronutrient composition, species distribution and algal blooms from the analysis of modern calcareous nannoplankton distributions in parts of the World Ocean.

REFERENCES

Baldwin, J. L., Otte, D. N. and Wheatley, C. L. 1989. Computer emulation of human mental process: Application of neural network simulations to problems in well log interpretation. **Society of Petroleum Engineers,** Paper 19619:481-493.

Baldwin, J. L., Bateman, R. M. and Wheatley, C. L. 1990. Application of neural network to the problem of mineral identification from well logs. **Log Analyst,** 3:279-293.

Brand, L. E.1994. Physiological ecology of marine coccolithophores, p. 39-49. In Winter, A. and Siesser, W. (eds.), **Coccolithophores**. Cambridge University Press.

Castradori, D. 1992. **I nannofossili calcarei come strumento per lo studio biostratigrafico e paleoceanografico del Quaternario nel Mediterraneo orientale.** Unpublished Ph.D. Thesis, University of Milano, Italy.

Castradori, D. 1993. Calcareous nannofosils and the origin of Eastern Mediterranean sapropels. **Paleoceanography**, 8:459-471.

Temperature (°C)

$y = -0.0424x^5 + 0.9664x^4 - 8.2703x^3 + 32.685x^2 - 57.693x + 50.701$
$R^2 = 0.9076$

Samples	SST Predicted	SST Observed
11/10/91	18.3	18.73
29/1/92	14.28	14.99
23/3/92	15.77	15.71
11/5/92	18.54	18.5
5/6/92	16.47	17.16
2/7/92	18.09	17.89
10/7/92	19.87	19.93

$y = -0.0131x^5 + 0.375x^4 - 3.8338x^3 + 17.579x^2 - 35.127x + 39.789$
$R^2 = 0.9246$

CALIFORNIA
BIGHT 1991-1992
(Ziveri et al., 1995)

SST Predicted — SST Observed — Poly. (SST Predicted) — Poly. (SST Observed)

Figure 3. Relative abundances of nannoplankton species considered in this work and a chart showing the relationship between observed and predicted values of SST. Samples used as training sets are in italics. 'Poly.' in legend means: polynomial regression.

Gaarder, K. R. 1970. Three new taxa of coccolithinae. Nytt Magasin for Botanikk, 17:113-126.

Giraudeau, J. 1992. Coccolith paleotemperature and paleosalinity estimates in the Caribbean Sea for the Middle-Late Pleistocene (DSDP Leg 68-Hole 502B). **Memorie Scienze Geologiche,** Padova, 43:375-387.

Giraudeau, J., Monteiro, P. M. S. and Nikodemus, K. 1993. Distribution and malformation of living coccolithophores in the northern Benguela upwelling system off Namibia. **Marine Micropaleontology,** 22:93-110.

Hutson, W. H. 1979. The Agulhas Current during the late Pleistocene: Analysis of modern analogs. **Science**, 207:64-66.

Imbrie, J. and Kipp, N. G. 1971. A new micropaleontological method for quantitative paleoclimatology: Application to a late Pleistocene Caribbean core, p. 71-181. In Turekian, K.K. (ed.), **The Late Cenozoic Glacial Ages,** Yale University Press, New Haven.

Klejine, A., Kroon, D. and Zevernoom, W. 1989. Phytoplankton and foraminiferal frequencies in northern Indian Ocean and Red Sea surface waters. **Netherlands Journal of Sea Research,** 24:531-539.
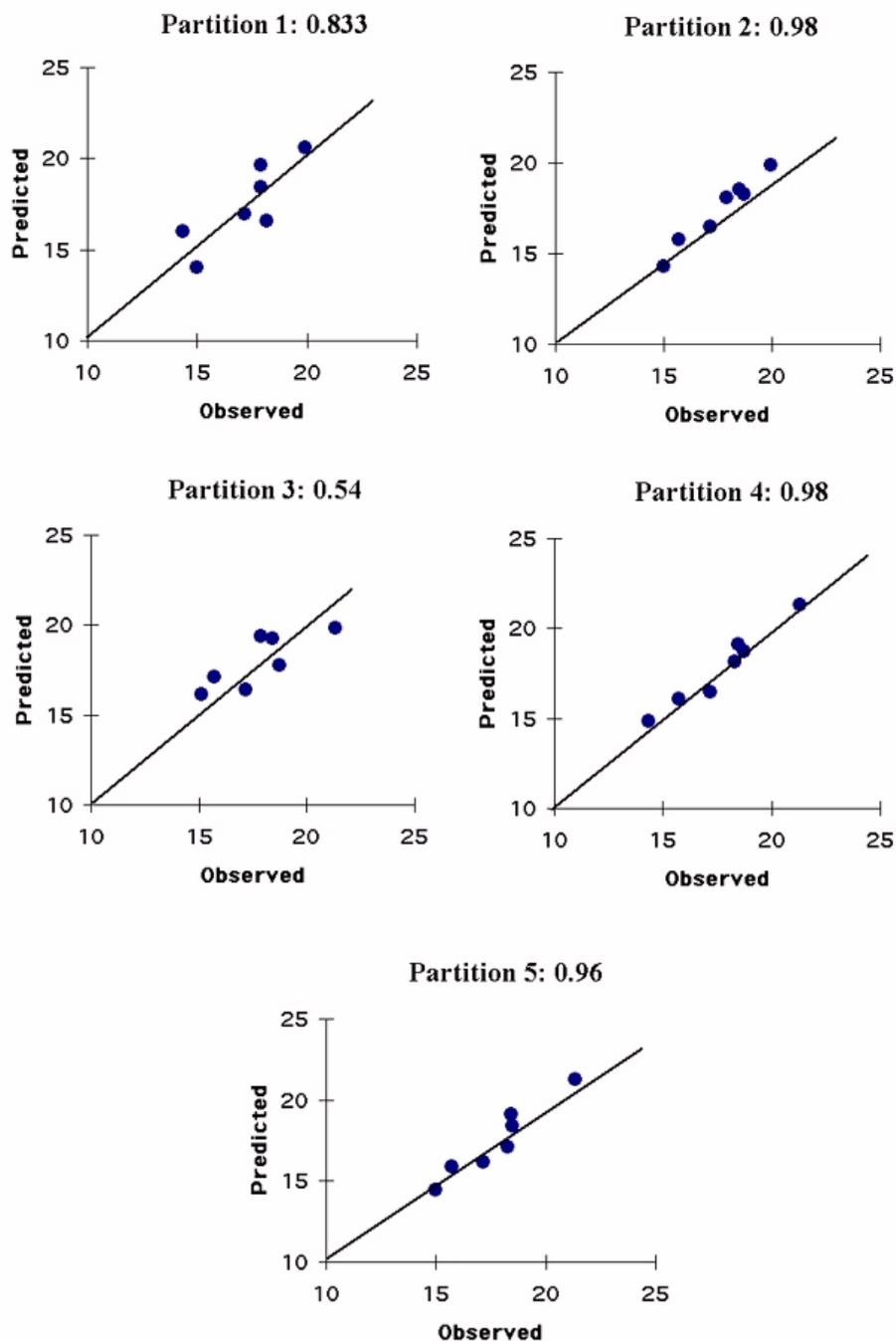
Figure 4. The correlation coefficients between observed and predicted SST's in the five test-sets of the Californian samples are as follows: Partition 1: 0.83; Partition 2: 0.98; Partition 3: 0.54; Partition 4: 0.98; and Partition 5: 0.96.

Malmgren, B. A. and Nordlund, U. 1996. Application of artificial neural networks to chemostratigraphy. **Paleoceanography**, 11:505-512.

Malmgren B. A. and Nordlund U. 1997. Application of artificial neural networks to paleoceanographic data. **Palaeogeography, Palaeoclimatology, Palaeoecology,** 136:359-373.

Malmgren, B. A. and Winter, A. 1999. Climate zonation in Puerto Rico based on principal components analysis and an artificial neural network. **Journal of Climate,** 12: 977-985.
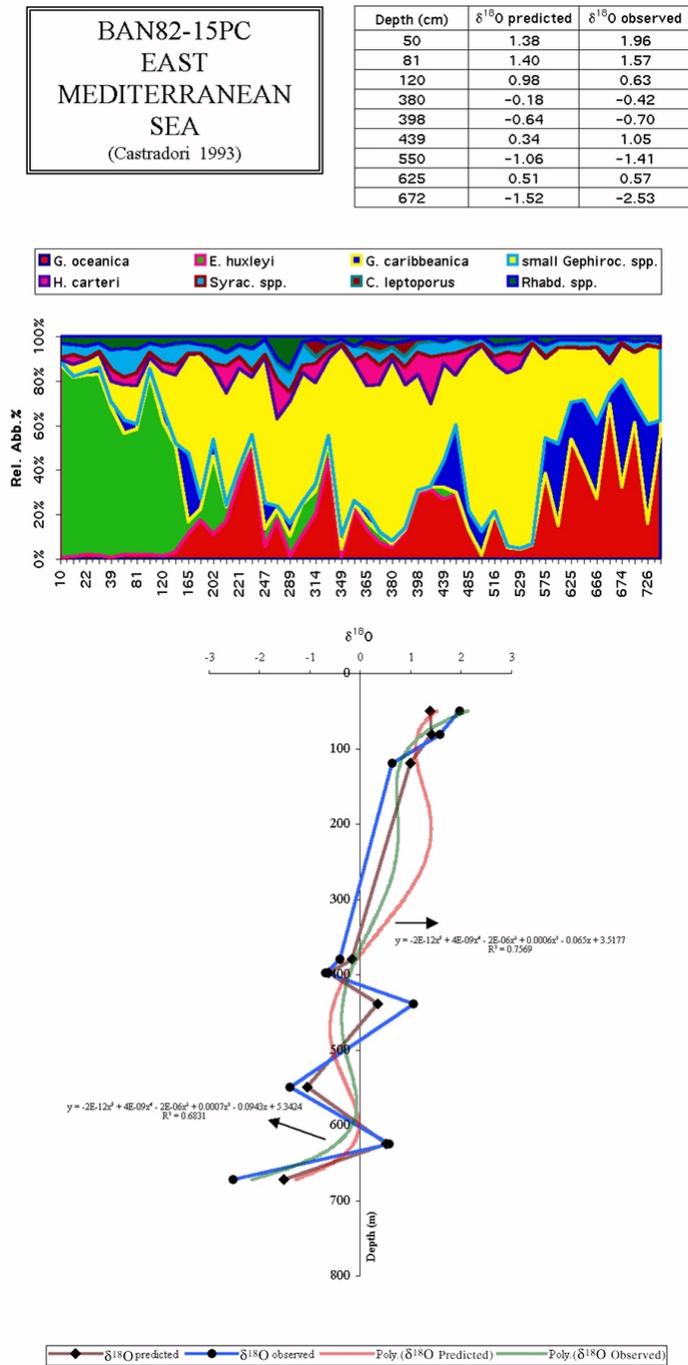
McIntyre, A., Bé, A. W. H. and Roche, M. B. 1970. Modern Pacific **Coccolithophorida: A paleontological thermometer. Transactions New York Academy of Sciences,** Ser. II, 32:720-731.

Mitchell-Innes, B. A. and Winter, A. 1987. Coccolitho- phores: A major phytoplankton component in mature upwelled waters off the Cape Peninsula, South Africa in March, 1983. **Marine Biology,** 95:25-30.

Okada, H. and McIntyre, A. 1979. Seasonal distribution of modern coccolithophores in the western North Atlantic Ocean. **Marine Biology,** 54:319-328.

Okada, H. and Honjo, S. 1975. Distribution of coccolitho- phores in Marginal Sea along the Western Pacific Ocean and in the Red Sea. **Marine Biology,** 31:271- 285.

Parisi, E. 1987. Carbon and oxygen isotope composition of Globogerinoides ruber in two deep-sea cores from the Levantine Basin (eastern Mediterranean). **Marine Geology,** 75:201-219.

Penn, B. S., Gordon, A. J. and Wendlandt, R. F. 1993. Using neural networks to locate edges and linear fea- tures in satellite images. **Computers & Geo- sciences,** 19:1545-1565.

Probyn, T. 1993. The inorganic nitrogen nutritions of the phytoplankton in the southern Benguela: New pro- ductions, phytoplankton size and implication for pelagic food webs. **South African Journal of Marine Science,** 12:411-420.

Pujos. A. 1992. Calcareous nannofossils of Plio-Pleis- tocene sediments from the northwestern margin of tropical Africa. In Summerhayes, C. P., Prell, W. L., and Emeis, K. C. (eds.), Upwelling Systems: Evolu- tion since the Early Miocene. **Geological Society of London Special Publication** 60:343-356.

Raiche, A., 1991. A pattern recognition approach to geo- physical inversion using neural nets. **Geophysical Journal International,** 105:629-648.

Rogers, S. J., Fang, J. H., Karr, C. L., and Stanley, D. A. 1992. Determination of lithology from well logs using a neural network. **American Association of Petro- leum Geologists Bull,** 76:731-739.

Roth, P. H. and Coulbourn, W.T. 1982. Floral and solu- tion pattern of coccoliths in surface sediments of the North Pacific. **Marine Micropaleontology,** 7:1-52.

Thunell, R., Pride, C., Ziveri, P., Muller-Karger, F., Sanc- etta, C. and Murray, D. 1996. Plankton response to physical forcing in the Gulf of California. **Journal of Plankton Research,** 18:2017-2026.

Winter, A. 1985. Distribution of living coccolithophores in the California Current system, Southern California Borderlands. **Marine Micropaleontology,** 9:385- 393.

Ziveri, P., Thunell, R. C. and Rio, D. 1995a. Export pro- duction of coccolithophores in an upwelling region: Results from San Pedro Basin, Southern California Borderlands**. Marine Micropaleontology,** 24:335- 358.

Ziveri, P., Thunell, R. C. and Rio, D. 1995b. Seasonal changes in coccolithophore densities in the Southern California Bight during 1991-1992. **Deep Sea Research I,** 42 :1881-1903.

ELECTRONIC REFERENCES

http://www.biocompsystems.com for Neurogenetics Optimizer software.

http://www.hh.se/staff/nicholas/NN_Links.html for a complete electronics links.

http://asgard.kent.edu/meso/neural/neural/ppframe.htm for an ANN summary.

http://www.interstate95.com/home/adaptive/index.html another useful group of electronic links.

http://gias720.dis.ulpgc.es/Links/links.html Vision Systems, Pattern Recognition, Machine Learning, Artificial Intelligence links.

Figure 5. Relative abundances of nannofossil species considered in this work and a diagram showing the relationship between observed and predicted $\delta^{18}O$. 'Poly.' in legend means: polynomial regression.

Depth (cm)	$\delta^{18}O$ predicted	$\delta^{18}O$ observed
50	1.38	1.96
81	1.40	1.57
120	0.98	0.63
380	−0.18	−0.42
398	−0.64	−0.70
439	0.34	1.05
550	−1.06	−1.41
625	0.51	0.57
672	−1.52	−2.53

Figure 6. The correlation coefficients between observed and predicted $\delta^{18}O$ in the five test-sets of the Mediterranea samples are as follows: Partition 1: 0.92; Partition 2: 0.96; Partition 3: 0.93; Partition 4: 0.94; and Partition 5: 0.64.