PREDICTION OF NANOPLANKTON CHANGE IN IZMIR BAY (AEGEAN SEA-TURKEY) WITH TIME BY NEURAL NETWORKS

F.S. Sunlu [1], I. Demir [2], G. Onkal Engin [3] *, U. Sunlu [1], B. Buyukisik [1], T. Koray [1], T.M. Sever [1], S. Kukrer [1], A. Aydin [1]
[1] Ege University, Faculty of Fisheries, Dept. of Hydrobiology, 35100 Bornova, Izmir, Turkey
[2] University of Georgia, Warnell School of Forest and Natural Resources, Athens, GA 30602, USA
[3] Gebze Institute of Technology, Environmental Eng. Dept., 41400, Gebze, Kocaeli, Turkey - guleda@gmail.com

Abstract
In this study, neural network analysis was used to construct prediction models of nanoplankton population change with nutrients and other environmentally important parameters. The results indicated that, with a data set of 52 weeks, it is possible to predict nanoplankton change.
Keywords : Aegean Sea, Coastal Waters, Eutrophication, Phytoplankton.

Introduction
Coastal areas are subject to diverse anthropogenic influence including industrial development, domestic wastes, maritime transport and agricultural activities. Neural networks, a statistical tool for data analysis, can be used to extract the relations between variables depending on the predictive variables used by means of a mechanism called training or learning. After a training process, Neural Networks are able to give estimations by using the relationship developed during the learning phase. The aim of this research was to predict of temporal changes of nanoplankton variation and physico-chemical environmental parameters using Neural Networks.

Materials and Methods
Experimental: In this study, physico-chemical environmental parameters, nutrients and some general biological parameters were measured weekly during an one year period. All these parameters were measured at different depths of 3 selected sampling stations which are located in middle and inner part of Izmir Bay.
Neural networks: Artificial Neural Networks (ANNs) are information processing algorithms that are inspired by the way biological nervous systems work, such as learning from past experience, making generalizations from similar situations and producing decisions out of incomplete knowledge of the states [1]. Neural networks are generally used in two main application areas, function approximation and pattern classification [2]. A feed-forward ANN, trained with the back-propagation algorithm, was implemented using the software package Neuroshell 2 (Ward Systems Group Inc., 2000).
Data division and pre-processing: 468 data was collected from selected stations, however, 18 points from PON and 18 points from nanoplankton data were missing. Missing data points replaced with the values from best matching weight vector for corresponding variable after training process using Self Organizing Map (SOM). Self Organizing Map (SOM) is an unsupervised neural network method which has properties of both vector quantization and vector projection algorithms. The obtained data were smoothened by Simple Moving Average (5 weeks) technique.
The 464 available data sets were divided into three groups, of which the 63 were used for testing, a set of 53 for evaluation and the remaining 348 data were used for training. The training set was used for adjusting the connection weights, whereas the testing set was used for the determination of network geometry and model parameters. Finally, the validation set was used for testing the optimality and generalization ability of the model developed [3]. The data were normalized between 0.0 and 1.0 before the NN analysis, corresponding to the limits of the transfer function, in order to improve training characteristics.
Network: The best results were obtained with a network consisting of two hidden layers, with having 14 and 4 neurons, respectively. Hyperbolic tangent transfer function used at hidden layers. Every 4^{th} data for testing and evaluation set extracted from the data.

Results and Discussion
During training, a number of different descriptive correlation values, namely the R Squared and Minimum Average Error (MAE) were observed for both training and testing sets. Using these values, monitoring of the local and the global minimum values for correlation values could be possible. Therefore, training was continued until overtraining was observed. Correlation values were calculated for evaluation data set after training.

The best network geometry was chosen according to highest correlation and lowest MAE error value of the testing/evaluation sets. Overall results are given in Table 1.

Tab. 1. Overall results. * percent of data within specific error range.

	Training		Testing		Evaluation	
	DO	NP	DO	NP	DO	NP
Corr. Coeff. R	0.975	0.946	0.959	0.891	0.962	0.916
MAE	0.23	73375	0.29	95672	0.29	86363
% within 5*	82.76	25.00	69.84	12.70	69.81	22.64
% within 5 to 10	14.37	19.54	26.98	20.64	30.19	20.76
% within 10 to 20	2.87	22.99	3.18	19.05	0	24.53
% within 20 to 30	0	15.23	0	23.81	0	15.09
% over 30	0	17.24	0	23.81	0	16.98

The prediction models demonstrated good abilities to model the training data. The denormalized actual and simulated DO and nanoplankton values are presented in Figure 1. As seen, the actual and predicted values were in good agreement. The performance of neural networks can be seen from training and evaluation data sets. As can be seen, the proposed network exhibited a good performance overall.

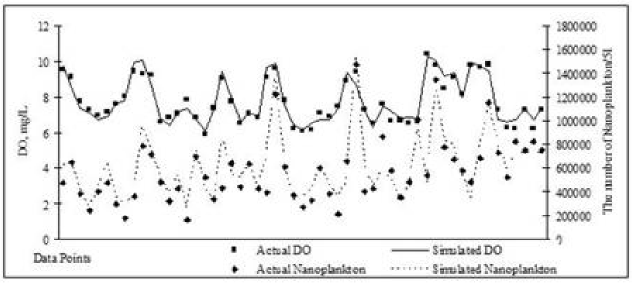

Fig. 1. Comparison between actual and predicted values of the NN model.

Conclusions
ANNs were applied to the obtained data. These analyses were carried out so that less analyses in laboratory conditions are needed in future to see the effect of nutrients on the nanoplankton population.

Acknowledgements
The authors would like to thank to Turkish Scientific and Technical Research Council for financial supports.

References
1 - McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115-133.
2 - Kosko, B., 1992. A Dynamical Systems Approach to Machine Intelligence. Prentice-Hall, Englewood Cliffs, NJ, 449p.
3 - Maier,H.R.,Dandy, G.C.,2000.Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues& applications. *Environmental Modelling & Software* 15(1),101-124.