European Water 11/12: 55-62, 2005.

The Use of a Neural Network Technique for the Prediction of Water Quality Parameters of Axios River in Northern Greece

M. J. Diamantopoulou[1], V. Z. Antonopoulos[2] and D. M. Papamichail[2]
[1] *School of Forestry and Natural Environment, Aristotle University, 54124 Thessaloniki, Greece, e-mail: papamich@agro.auth.gr*
[2] *School of Agriculture, Aristotle University, 54124 Thessaloniki, Greece, e-mail: papamich@agro.auth.gr*

Abstract: Axios River is one of the most important transboundary rivers between the Greek and the neighbour country FYROM in the Balkan area. In this paper, Artificial Neural Networks (ANNs) were used to derive and to develop models for prediction the monthly values of some water quality parameters of the river Axios at a station located at Axioupolis site of Greece near the Greece - FYROM borders by using the monthly values of the other existing water quality parameters as input variables. The monthly data of twelve water quality parameters and the discharge, for the time period 1980-1994 were selected for this analysis. The results demonstrate the ability of the appropriate Neural Network models for the prediction of water quality parameters. This provides a very useful tool for filling the missing values of time series of water quality parameters that is a very serious problem in most of the Greek monitoring stations.

Key words: Neural Networks, Water quality parameters, Missing values, Axios river

1. INTRODUCTION

The water quality and quantity of water resources worldwide is a subject of ongoing concern. The assessment of long–term water quality changes is also a challenging problem. During the last decades, a gradual accumulation of reliable long-term water quality data have been monitored for many rivers in the world (Helsel and Hirsch, 1992; Antonopoulos et al., 2001).

Many spatial and temporal techniques require that the monitoring be equally spaced over time or space. Missing values of time series is a very serious problem in most of the Greek monitoring stations where the monitoring of the water quality parameters is usually on a monthly basis. A number of estimation techniques can be used to estimate the missing values (Holder, 1985; Helsel and Hirsch, 1992; Mitsiou et al., 1999; Skoulikidis, 2002). Some of these methods include linear interpolation, polynomial fitting, or nonparametric methods.

The monthly missing values of the water quality parameters can be estimated, taking into consideration the relationship between predicted variable with monthly missing value and predictors which are the monthly values of the other existing water quality parameters, from multiple linear and nonlinear regression models. Because of the bias is detected on least squares assumptions and the high values of errors, a neural network technique can be used. In recent years, Artificial Neural Networks (ANNs) have become extremely popular for prediction and forecasting in a number of areas, including finance, power generation, medicine, water resources and environmental science (Maier and Dandy, 2000; Diamantopoulou, 2005). Shamseldin et al. (1997) used ANNs for combining the outputs of different rainfall-runoff models in order to improve the estimation and prediction of streamflow. Mantoglou and Kourakos (2002) used also the ANNs for modelling rainfall-runoff process.

According the new European Water Framework Directive 2000/60 (EU, 2000), the achievement of a good ecology, quantitative and qualitative status of surface waters has been obligatory. The river's water management plans should be development on the watershed basis. This should contain a summary of all significant pressures and impacts of human activities on the aquatic environment. The watershed of transboundary rivers and lakes are shared between different countries. In these cases the cooperation among the member of state are needed and especially for rivers of which part of basin is under non-member countries. Greece, a member of European Union, share five rivers (Evros, Nestos, Strymon, Axios and Aoos) and three lakes (Megali and Mikri Prespa and Doirani) with the non-member neighbour countries. Axios is one of the most important transboundary river between Greece and FYROM.

In this paper, Artificial Neural Networks (ANNs) were used to derive and to develop models to predict the monthly values of some water quality parameters of the transboundary river Axios at a station located near the Greece - FYROM borders. The purpose of the above models is to fill the monthly missing values of some water quality parameters at this station of river.

2. MATERIALS AND METHODS

2.1 Artificial Neural Networks Methodology

The basic structure of an Artificial Neural Network (ANN), usually, consists of three distinctive layers, the input layer, where the data are introduced to the ANN, the hidden layer or layers, where data are processed, and the output layer, where the results of ANN are produced. The structure and operation of ANNs is discussed by a number of authors (Fausett, 1994; Haykin, 1994; Dowla and Rogers, 1995; Patterson, 1996; Gurney, 1999).

The ANNs are designed by putting weights between neurons, by using a transfer function that controls the generation of the output in a neuron, and using adjustable laws that define the relative importance of weights for input to a neuron. In the training, the ANN defines the importance of the weights and adjusts them through an iterative procedure.

In this study, the training of ANNs was achieved by the cascade correlation algorithm (Fahlman and Lebiere, 1990) which is a supervised algorithm in the multilayer feed-forward ANNs. The Cascade part refers to the architecture and its mode of construction entails adding hidden units once at a time and always connecting all the previous units to the current unit. The Correlation part refers to the way hidden units were trained by trying to maximize the correlation between output of the hidden unit and the desired output of the network across the training data. The information is processed in the forward direction from the input layer to the hidden layer or layers to the output layer. Kalman's learning rule (Kalman, 1960; Brown and Hwang, 1992; Grewal and Andrews, 1993; Demuth and Beale, 2001) was used to modify the ANN weights.

The training procedure of ANNs is composed of a forward pass (Fig. 2). In this pass, the input data are multiplied by the initial weights and then, are summed to yield the net to each neuron. The net of a neuron is passed through a transfer function to produce the output of the neuron. Since, the adjustment of the network weights is accomplished with the derivative of the transfer function, continuous transfer functions are desirable. In this study, the hyperbolic-tangent function was used (Fausett, 1994):

$$f(s) = \tanh(s) = \frac{e^{s} - e^{-s}}{e^{s} + e^{-s}} = \frac{1 - e^{-2s}}{1 + e^{-2s}} \tag{1}$$

where: $s_i = \sum_{i=1}^{n} w_i x_i$, in which w_i are weights and x_i are input values.

The error between the output of the ANN and the target value of the output was computed, as well. In order to achieve an estimation of the monthly missing values of the time series of one water quality parameter, the monthly values of the other existing water quality parameters are introduced as input parameters into ANNs. In this sense, the input layer of ANNs consists of a number input neurons and one output neuron, which is the water quality parameter with missing values (Fig. 2). During the training and verification period of ANNs, the simulated monthly values of the six water quality parameters are compared with the corresponding measured monthly values to identify the simulation errors. The geometry of ANNs, which determines the number of connection weights and how these are arranged, depends on the number of hidden layers and the number of the hidden nodes in these layers. In the developed ANNs, one hidden layer is used and the number of the hidden nodes is optimized by maximizing the correlation between output of the hidden unit and the desired output of the network across the training data. However, the final network architecture and geometry are tested to avoid over-fitting as suggested by Maier and Dandy (2000).

2.2 Site Description and data

The Axios River, is a transboundary river of Greece and FYROM. Its length is about 376 km, approximately 76 km of which are in Greek territory. The area of the catchment is approximately 22,450 km^2, 2,300 km^2 (10 %) of which are in Greece (Figure 1). The delta of Axios river is one of the most important Greek conservation areas and one of the 11 Ramsar sites in Greece. The water of river irrigates about 30.000 ha in Thessaloniki plain of Greece. The Greek Ministry of Agriculture measures water flow and takes water samples once every month at the Axioupolis station of the Axios River (HMA-Hellenic Ministry of Agriculture, 1997). This station is located a few kilometers downstream the Greek-FYROM border, and have been included in the European Community's water quality monitoring program, since 1982 (EC, 1994).

Among the variables measured by the Department of Irrigation and Water Protection of the Greek Ministry of Agriculture are water temperature (T), pH, specific conductivity (EC_w), dissolved oxygen (DO), bicarbonates (HCO_3^-), sulfates (SO_4^{2-}), chlorides (Cl^-), sodium (Na^+), magnesium (Mg^{2+}), calcium (Ca^{2+}), nitrates (NO_3^-), ammonia (NH_4^+), total phosphorus (TP) and discharge (Q). The available monitoring data of the aforementioned variables for the analysis of this paper are on a monthly basis for the period 1980-1994. Discharge is a parameter with many missing values, especially since 1989. The values of discharge since 1994 were impossible to be completed. Nitrate and total phosphorus have neither been measured since October 1994. The missing monthly values of the original time series were completed using the methods proposed by Holder (1985).

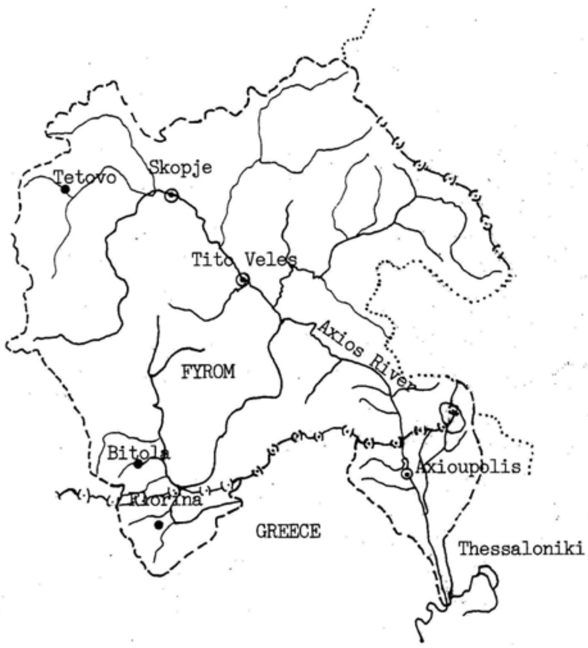

Figure 1. River Axios, its catchment boundaries and Axioupolis water quality sampling station

3. RESULTS AND DISCUSSION

ANNs were used to derive and to develop models to fill the monthly missing values of water quality parameters at Axios River. The monthly data of twelve water quality parameters (NO_3^-, EC_w, DO, T, SO_4^{2-}, Na^+, Mg^{2+}, Ca^{2+}, TP, pH, Cl^-, HCO_3^-) and the discharge (Q), at the Axioupolis station, for the time period 1980-1994 were selected for this analysis. The statistical measures of time series of monthly water quality variables are given in Table 1. For the six of water quality parameters (nitrates, specific conductivity, dissolved oxygen, sodium, magnesium, and calcium) the appropriate ANNs were developed.

Table 1. Statistical parameters of the time series of monthly values of water quality parameters and discharge of Axios River at the Axioupolis station for the period 1980-1994

Parameter	Mean	Max value	Min value	Variance
Q, m^3/sec	101.08	443.6	4.50	9597.8
T, ^oC	13.48	30.0	-2.0	54.33
EC_w, μmhos/cm	477.25	870.0	260.0	16671.5
pH	7.70	8.620	6.90	0.120
DO, mg/l	9.89	13.50	2.20	3.35
Na^+, me/l	0.701	3.20	0.30	0.101
Mg^{2+}, me/l	1.252	3.80	0.10	0.251
Ca^{2+}, me/l	3.299	6.00	1.80	0.845
NO_3^-, mg/l	7.14	18.25	0.018	12.725
TP, mg/l	0.625	4.359	0.00	0.292
SO_4^{2-}, me/l	1.504	5.70	0.10	1.026
Cl^-, me/l,	0.413	0.10	1.10	0.041
HCO_3^-, me/l	3.298	4.60	2.20	0.252

For neural network models construction, monthly data randomly partitioned into training (90% of all data) and test (the remaining 10% of all data) data sets, were used. The cascade correlation algorithm achieved the training of neural networks, which is a feed-forward and supervised algorithm. Kalman's learning rule was used to modify the artificial neural networks weights. The networks are designed by putting weights between neurons, by using the hyperbolic-tangent function of training of the form of Eq. 1. The number of nodes in the hidden layer was determined based on the maximum value of coefficient of correlation. Different networks structures tested in order to determine the optimum number of hidden layers and the number of nodes in each. The architecture of the neural networks for the nitrates is shown in Fig. 2.

The best solution given by the neural network for the NO_3^- (Fig. 2) composed of one input layer with nine input variables, one hidden layer with nineteen nodes and one output layer with one output variable with value of coefficient of correlation equal to 0.9382.

The correlation coefficient (R), the mean absolute error (MAE), the root mean square error (RMSE), the (%) of the mean, between the output of the hidden unit and the desired output of the networks and the input variables for the nitrates, the specific conductivity, the dissolved oxygen, the sodium, the magnesium and the calcium for total data set, training and test set are given in Table 2.

For the ANN models construction for the six water quality parameters (NO_3^-, ECw, DO, Na^+, Mg^{2+} and Ca^{2+}) chosen to be used as outputs, many combinations of all measured variables tested as inputs. The best combinations with the better prediction results for each of the six variables used as outputs are given in Table 2.

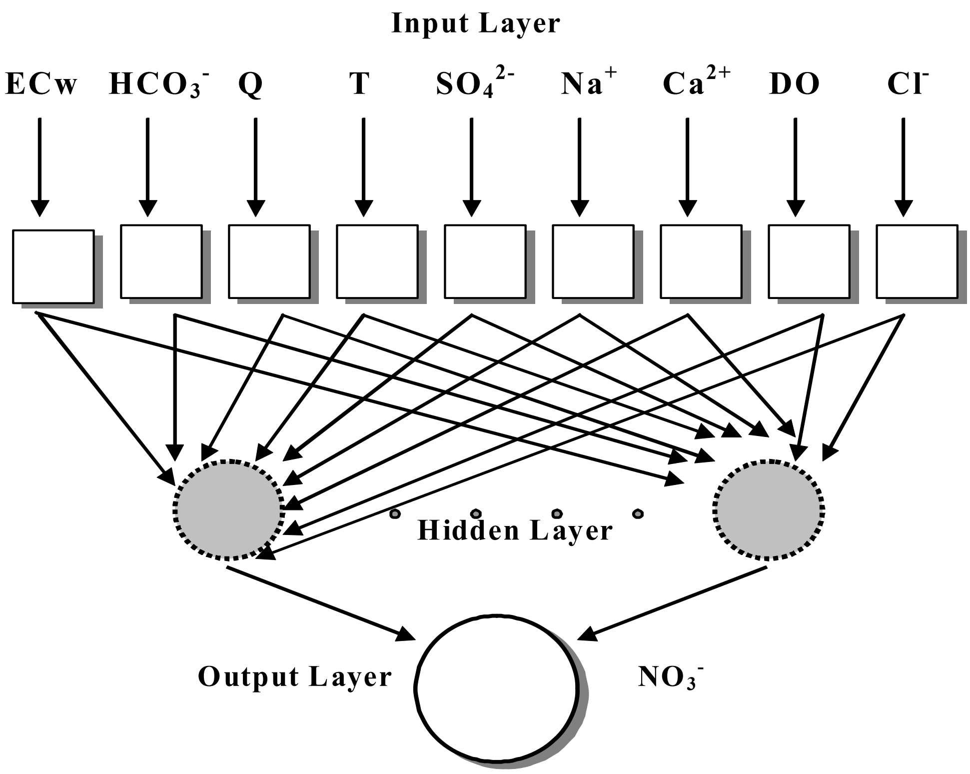

Figure 2. Artificial Neural Network architecture for the Nitrates (NO3-), at the Axioupolis station of Axios river

Table 2. Correlation coefficient (R), mean absolute error (MAE), root mean square error (RMSE), the (%) of the mean and the input variables of the ANNs, for the NO_3^-, ECw, DO, Na^+, Mg^{2+} and Ca^{2+} for total, training and test set

	NO_3^- / ANN: 9-19-1/0.9382			ECw / ANN: 10-4-1/0.9744		
Data set	R	MAE	RMSE	R	MAE	RMSE
Total	0.9278	0.962	1.333 (18.67%)	0.9755	18.915	28.377 (5.95%)
Train	0.9382	0.886	1.246 (17.44%)	0.9744	19.119	29.060 (6.09%)
Test	0.8420	1.634	1.941 (27.17%)	0.9863	17.117	21.415 (4.49%)
Inputs	Q, T, ECw, DO, Na^+, Ca^{2+}, SO_4^{2-}, Cl^-, HCO3$^-$			Q, T, pH, Na^+, Mg^{2+},Ca^{2+}, TP, SO_4^{2-}, Cl^-, HCO3$^-$		
	DO / ANN: 9-14-1/0.9188			Na^+ / ANN: 11-24-1/0.9331		
Data set	R	MAE	RMSE	R	MAE	RMSE
Total	0.9103	0.552	0.759 (7.67%)	0.9353	0.035	0.117 (16.65%)
Train	0.9188	0.531	0.735 (7.43%)	0.9331	0.033	0.121 (17.25%)
Test	0.8741	0.736	0.941 (9.52%)	0.9678	0.053	0.069 (9.84%)
Inputs	Q, T, pH, ECw, Na^+, NO_3^-, TP, SO_4^{2-}, Cl^-			Q, T, pH, ECw, DO, Ca^{2+}, NO_3^-, TP, SO_4^{2-}, Cl^-, HCO3$^-$		
	Mg^{2+} / ANN: 8-10-1/0.9640			Ca^{2+} / ANN: 5-10-1/0.9454		
Data set	R	MAE	RMSE	R	MAE	RMSE
Total	0.9660	0.065	0.135 (10.84%)	0.9481	0.223	0.294 (8.90%)
Train	0.9640	0.067	0.145 (11.22%)	0.9454	0.226	0.301 (9.11%)
Test	0.9890	0.046	0.0821 (6.56%)	0.9718	0.193	0.225 (6.81%)
Inputs	pH, ECw, Na^+, Ca^{2+}, NO_3^-, SO_4^{2-}, Cl^-, HCO3$^-$			ECw, Na^+, TP, NO_3^-, SO_4^{2-}		

Monthly values of nitrates, dissolved oxygen, specific conductivity, magnesium, calcium and sodium estimated by the Neural Network models versus the corresponding monthly measured values are shown in Figures 3(a) and 3(b), 3(c) and 3(d), 3(e) and 3(f), 4(a) and 4(b), 4(c) and 4(d), 4(e) and 4(f), for the training and the test data set, respectively.

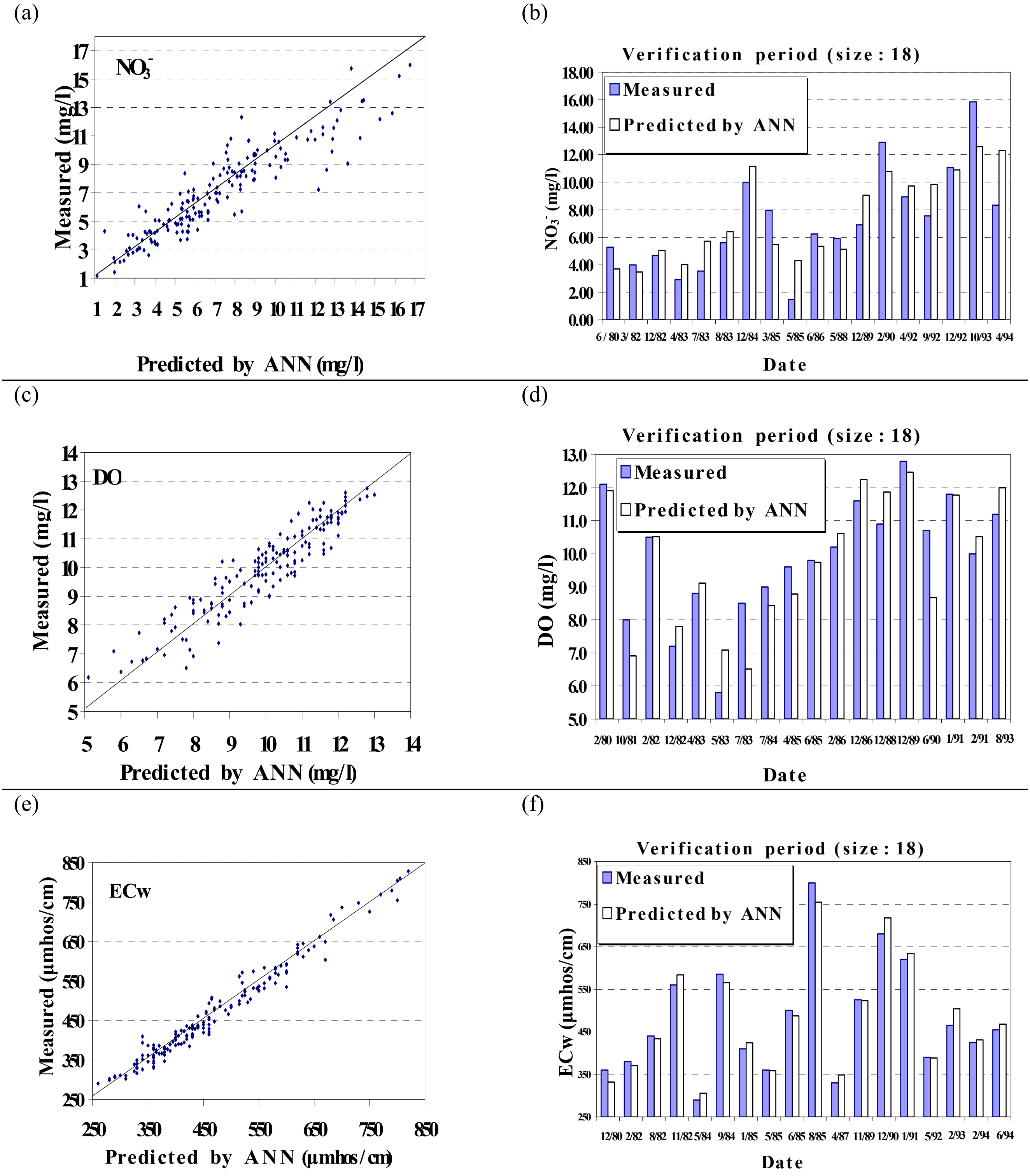

Figure 3. Monthly values of nitrates (NO_3^-), dissolved oxygen (DO) and specific conductivity (ECw) estimates by the Neural Networks versus the corresponding measured values for the training and the test data set

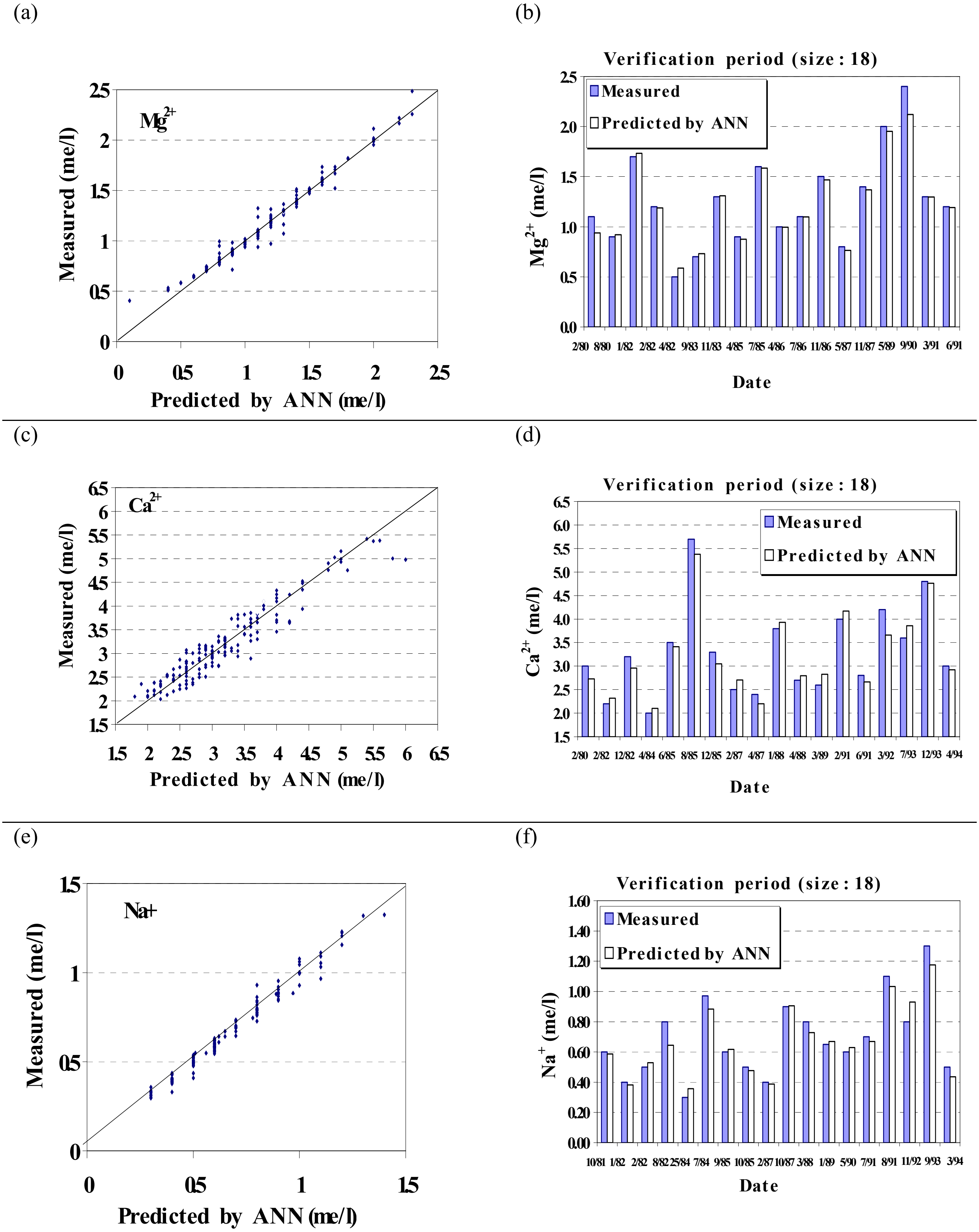

Figure 4. Monthly values of magnesium (Mg^{2+}), calcium (Ca^{2+}) and sodium (Na^{+}), estimates by the Neural Networks versus the corresponding measured values for the training and the test data set

Table 2 and Figures 3 and 4 clearly demonstrate the ability of the Neural Network models to predict very well the monthly values of the NO_3^-, DO, ECw, Mg^{2+}, Ca^{2+} and Na^+ at the Axioupolis station.

4. CONCLUSIONS

In this paper, ANNs were developed for the prediction the monthly values of the six water quality parameters nitrates (NO_3^-), specific conductivity (ECw), dissolved oxygen (DO), sodium (Na^+), calcium (Ca^{2+}) and magnesium (Mg^{2+}) at the Axioupolis station, of Axios River, Greece. The monthly data of these and six other water quality parameters and discharge (Q), at the Axioupolis station, for the time period 1980-1994 were selected for this analysis. The training of neural networks was achieved by the cascade correlation algorithm which is a feed-forward and supervised algorithm. Kalman's learning rule was used to modify the artificial neural networks weights. The networks are designed by putting weights between neurons, by using the hyperbolic-tangent function of training. The number of nodes in the hidden layer was determined based on the maximum value of coefficient of correlation. The results for the training and the test data sets were satisfactory. Consequently, the Neural Network models can be used for the prediction of water quality parameters and allow the filling of the missing values of time series of water quality parameters that is a very serious problem in most of the Greek monitoring stations.

REFERENCES

Antonopoulos, V.Z., Papamichail, D.M., Mitsiou, K.A.: 2001, 'Statistical and trend analysis of water quality and quantity data for the Strymon River in Greece', Hydrology and Earth System Sciences, 5(4), 679-691.

Brown, R.G. and Hwang, P. Y. C.: 1992, 'Introduction to Random Signals and Applied Kalman Filtering' 2nd Ed., Wiley, New York.

Demuth, H. and Beale, M.: 2001, 'Neural Network Toolbox. For Use with Matlab®', User's Guide, V4.

Diamantopoulou, M.J.: 2005, 'Artificial neural networks as an alternative tool in pine bark volume estimation', Comput. Electron. Agric. 48, 235-244.

Dowla, U.F. and Rogers, L.: 1995, 'Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks', Massachusetts, USA.

EU: 2000, 'Directive of the European Parliament and the European Council for the creation of a framework for measures of the union in the area of water policy', Official Journal, L327, 22/12/2000, P.0001.

European Community: 1994, Water Quality of Surface Waters-Common procedure for exchange of information, General Direction XI, Environment, nuclear safety and public protection.

Fahlman, S.E. and Lebiere, C.: 1990, 'The Cascade Correlation Learning Architecture', in D.S. Touretsky (ed.), Advances in Neural Information Processing Systems 2, San Mateo, CA, Morgan Kaufmann, pp. 524-532.

Fausett, L.: 1994. 'Fundamentals of Neural Networks Architectures. Algorithms and Applications', Prentice Hall, USA.

Grewal, M. S. and Andrews, A.P.: 1993, 'Kalman Filtering Theory and Practice', Upper Saddle River, Prentice Hall, NJ, USA.

Gurney, K.: 1999, 'An Introduction to neural networks', UCL Press, UK

Haykin, S.: 1994, 'Neural Networks: A Comprehensive Foundation', Macmillan, New York, USA.

Helsel, D.R. and Hirsch, R.M.: 1992, 'Statistical Methods in Water Resources', Elsevier Science Publishers B.V., The Netherlands.

HMA (Helenic Ministry of Agr.): 1997. Water quality and discharge monitoring data for the Greek Water Resources, Department of Irrigation and Water Protection of Greek Ministry of Agriculture, Greece.

Holder, R.L.: 1985, 'Multiple regression in Hydrology', Institute of Hydrology, Wallingford, Oxfordshire, UK.

Kalman, R.E.: 1960. A New Approach to Linear Filtering and Prediction Problems. Transaction of the ASME. Journal of Basic Engineering, pp. 35-45.

Maier, H. R. and Dandy, G. C.: 2000, 'Neural Networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications',. Environmental Modeling & Software, 15, 101-124.

Mantoglou, A. and Kourakos, G.: 2002, 'Modeling the rainfall-runoff process using neural networks with emphasis on optimization of network configuration using genetic algorithms', Proc. of Inter. Conf., Water Resources Management in the ERA of Transition, European Water Resources Association, Athens, pp.229-242.

Mitsiou, K., Antonopoulos, V., Papamichail, D. : 1999. Statistical analysis of water quality parameters time series of Strymon river. Hydrotechnika, 9, 59-74.

Patterson, D.: 1996, 'Artificial Neural Networks', Prentice Hall, Singapure.

Shamsheldin, A. Y., O'Connor, K. M. and Liang, G. C.: 1997, 'Methods for combining the outputs of different rainfall-runoff models', J. Hydrol. 197, 203-229.

Skoulikidis, N.T.: 2002. Hydrochemical character and spatiotemporal variations in a heavily modified river of western Greece. Environmental Geology, 43, 814-824.