XXX IAHR CONGRESS

AUGUST 2003, AUTh, THESSALONIKI, GREECE

THEME D

HYDROINFORMATICS AND ADVANCED DATA TECHNOLOGY IN ENGINEERING PRACTICE

Also includes papers from the AGORA on "Hydraulic Instrumentation"

SERIES EDITORS *Ganoulis J. (AUTh, Greece)*
Prinos P. (AUTh, Greece)

THEME EDITORS *Korfiatis G. (Stevens Institute of Technology, USA)*
Christodoulou G. (NTUA, Greece)

RAINFALL-RUNOFF PROCESS IN MOUNTAINOUS CATCHMENT WITH ARTIFICIAL NEURAL NETWORK

Panagoulia D.[1] and **Maratos N.**[2]

[1]National Technical University of Athens, Faculty of Civil Engineering, Department of Water Resources-Hydraulic & Maritime Engineering, 5 Heroon Polytechniou, 15780 Zographou, Athens, Greece, dpanag@central.ntua.gr/ tel: +3-2107722858.

[2]National Technical University of Athens, Faculty of Electrical Engineering, Department of Signals, Control and Robotics, 9 Heroon Polytechniou, 15780 Zographou, Athens, Greece.

Abstract: An algorithm using a combination of linear least squares and multi-start simplex optimization (LLSSIM), originally proposed by *Hsu et al., 1995*, is used in order to show the mechanism and parameters of three-layer feed forward ANN models and the potential of such models for simulating and forecasting the nonlinear hydrological behavior of mountainous catchments. The output 'rain plus melt' from the snow accumulation and ablation model (SAA) of the US National Weather Service (US NWS) applied on a medium-sized mountainous catchment (the Mesochora catchment in Central Greece) was used as input to ANN model. The nonlinear ANN model approach is shown to provide a better representation of the rainfall-runoff relationship in medium and extreme conditions than the conceptual soil moisture accounting (SMA) model of the US NWS applied over the same catchment. Because the ANN approach presented here has not physically realistic components and parameters, it is by no means a substitute for conceptual catchment modeling. However, the ANN approach does provide a viable and effective alternative for developing input-output simulation and forecasting models in cases that do not require modeling of the internal dynamics of the catchment.

Keywords: Artificial neural network, linear least squares and multi-start simplex optimization, conceptual modeling, rainfall-runoff process, mountainous catchment.

1. INTRODUCTION

In recent years, artificial neural networks (ANNs) have been used successfully to model complex nonlinear input-output time series relationships in a wide variety of hydrological systems. Although deterministic conceptual models strive to account for all physical processes, their ability may be restricted by the need for catchment-specific data and the simplification involved in solving the governing equations. ANNs offer a relatively quick and flexible way of modelling, and as such applications are widely reported in hydrological literature (*Hsu et al., 1995; Raman and Sunilkumar, 1995; Campolo et al., 1999; Imrie et al., 2000*). However, the applicability of ANNs is restricted in situations where explicit knowledge of the internal hydrological sub-processes is not required.

This study demonstrates the applicability of the ANNs approach in developing effective nonlinear models of the rainfall-runoff process without the intention to substitute a physically based conceptual model. The identification of the ANN model presented here includes the structure and the model parameters (weights) that are calibrated (trained) through an interactive procedure exploring an objective function surface in search of an optimum.

The algorithm for training three-layer feed forward ANNs adopted here was a combination of linear least squares and multi-start simplex optimization (LLSSIM) (*Hsu et al., 1995*). This is more effective and efficient than the back propagation algorithms (BPA) because it reliably obtains the global or near-global solution of the problem with fewer function evaluations. Moreover, the LLSSIM algorithm incorporates a procedure for automatically identifying a parsimonious model structure through stepwise expansion of the network size. This results in relatively rapid model development requiring no user intervention.

This study is organized as follows. First, a brief overview of the fundamentals of ANN modeling is present. Next, the LLSSIM methodology for training the ANN model weights and identifying a parsimonious model structure is described. Finally, the LLSSIM methodology is used to develop an ANN model for the medium-sized mountainous Mesochora catchment in Central Greece. The performance of this model is compared to the SMA model used by the US NWS. Inputs to both ANN and SMA models were used the output 'rain plus melt' from the snow accumulation and ablation model (SAA) of the US National Weather Service (US NWS) applied over the Mesochora catchment.

2. ARTIFICIAL NEURAL NETWORKS (ANNs)

The most widely researched and used structures of ANNS are the multi-layer feed forward networks [*Rumelhart et al.*, 1986], which are adopted here due to their best performance with regard to input-output function approximation. A feed forward ANN can have many layers. A typical three-layer feed forward ANN is shown in Figure 1. Each node j receives incoming signals from every node i in the previous layer. Associated with each incoming signal (x_i) is a weight (w_{ji}). The effective incoming signal (S_j) to node j is the weighted sum of all the incoming signals:

$$s_j = \sum_{i=0}^{n_0} w_{ji} x_i \tag{1}$$

where x_0 and w_{j0} are called the bias ($x_0 = 1.0$) and the bias weights, respectively. The effective incoming signal, S_j, is passed through a non-linear activation function (sometimes called a transfer function or threshold function) to produce the outgoing signal (y_j) of the node. The characteristics of a sigmoid function are that it is bounded above and below, it is monotonically increasing, and it is continuous and differentiable everywhere [*Hecht-Bielsen*, 1990]. The sigmoid function most often used for ANNs is the logistic function:

$$y_j = f(s_j) = 1/[1 + \exp(-s_j)] \tag{2}$$

in which S_j can vary on the range $\pm\infty$, but y_j is bounded between 0 and 1.

In this study, the training of three-layer feed forward ANNs is based on the use of an effective and efficient network training algorithm, entitled LLSSIM (*Hsu et al.*, 1995). The LLSSIM algorithm uses a partition of the weight space to implement an optimal synthesis of two training strategies. The input-hidden layer weights are estimated using a multi-start version of the simplex nonlinear optimization algorithm (*Nelder and Mead*, 1965), while the hidden-output layer weights are estimated using optimal linear least squares estimation (LLS) (*Scalero and Tepedelenlioglu*, 1992). The algorithm takes advantage of this weight space partition to conduct the nonlinear portion of the search in a reduced dimensional space, resulting in an acceleration of the training process. The simplex search algorithm provides improved global search characteristics owing to the use of multiple starts initiated randomly in the search space and its ability do not be trapped by minor local optima. Identification of the structure of the ANN is done using a strategy of progressively adding nodes to the hidden layer until a structure appropriate to the complexity of the problem is achieved.

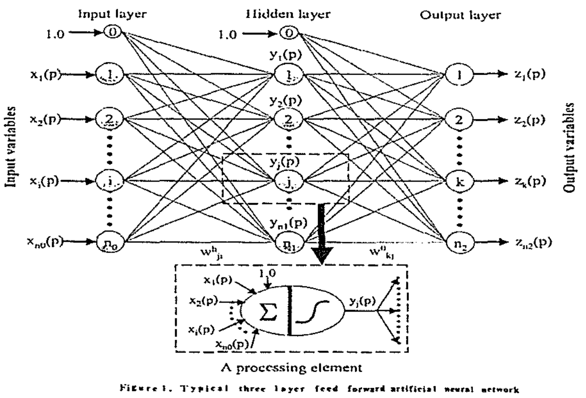


Figure 1. Typical three layer feed forward neural network

3. THE LLSSIM NETWORK TRAINING ALGORITHM

Let as, $t_k\,(p)$ target value of output node k of training pattern p; $x_i\,(p)$ input value of input node i of training pattern p; $y_j\,(p)$ output value of hidden node j of training pattern p; $z_k\,(p)$ output value of output node k of training pattern p; $s_j^h(p)$ weighted sum of inputs to hidden node j of pattern p; $s_k^0(p)$ weighted sum of hidden node outputs entering output node k of pattern p; w_{ji}^h connection weight from input node i to hidden node j; w_{kj}^0 connection weight from hidden node j to output node k.

Then, the LLSSINM weight-training strategy is formulated to the hidden-output weights such as:

$$\frac{\partial F_1}{\partial w_{kj}^0} = -\sum_{p=1}^{m} (\mathrm{TS}_k(p) - s_k^0(p))y_j(p) = 0 \tag{3}$$

The above equation can be rewritten as

$$\sum_{p=1}^{m} \mathrm{TS}_k(p)y_1(p) = \sum_{p=1}^{m}\sum_{l=0}^{n_1} w_{kl}^0 y_l(p)y_j(p)$$

$$= \sum_{p=1}^{m} y_j(p)\sum_{l=0}^{n_1} y_l(p)w_{kl}^0 \tag{4}$$

Define:

$$R_{jl} = \sum_{p=1}^{m}\sum_{l=0}^{n_1} y_j(p)y_l(p) \tag{5}$$

$$Q_j = \sum_{p=1}^{m} \mathrm{TS}_k(p)y_j(p) \tag{6}$$

The hidden-output weights, w_k^0, can now be found by solving the matrix equation:

$$w_k^0 = R^{-1}Q \tag{7}$$

where, $w_k^0 = \left[w_{k_0}^0, w_{k_1}^0, \ldots, w_{kn_1}^0\right]^T$ are the conditionally optimal hidden-output weights because their values depend on the values selected for the input-hidden weights.

4. RAINFALL-RUNOFF MODELING

This study compares the performance of two kinds of different model structures with respect to their ability to represent the rainfall-runoff process of a medium-sized mountainous catchment. The two model structures are (1) three-layer feed forward ANN model, and (2) the Soil moisture accounting model (SMA) of the US NWS. The first is nonlinear system theoretical model and the second is a conceptual model. Fifteen consecutive years of daily rainfall (rain plus melt generated from SAA model) and runoff data for the Mesochora catchment (633 km^2), were selected for model development and testing. Five years of data (1972 to 1976) are used for model identification, while the remaining ten years (1977 to 1986) are used for ANN model validation, while whole the period was used for SMA model calibration. Identification and calibration periods are enough long to extract representative results of catchment behavior, in contrast to *Hsu et al.*, (1995) study in which the identification period was only one year.

5. ANN MODEL IDENTIFICATION

The ANN model structure is ideally suited for highly nonlinear rainfall-runoff modeling. The runoff output $z(t)$ was assumed to be related to past inputs $x\ (t-j)$ and outputs $z\ (t-j)$ using a general nonlinear model structure:

$$Z(t) = g_{\text{non}}(z(t-1) - \dots, z(t-n_a), x(t-1), \dots, x(t-n_b)) + e(t) \quad (8)$$

where: $g_{\text{non}}(\ \ )$ is the unknown nonlinear mapping function, $e(t)$ is the unknown mapping error (to be minimized), and n_a and n_b are the (unknown) number of past inputs and outputs contributing to the present output. This model structure is represented by the notation ANN (n_a, n_b, n_h, n_0), where: $n_a + n_b$ is the number of nodes in the input layer, n_h is the number of nodes in the hidden layer, and n_0 is the number of nodes in the output layer ($n_0 = 1$ in our case). To identify an ANN model, values for n_a, n_h, and n_0 must be selected, and values for the network weights $w^{\eta}_{\xi\iota}$ and w^{0}_{kj} must be estimated so that the prediction error is minimized.

In this study, n_a and n_b were each varied over the range 2 to 5. For each of the combinations of n_a and n_b, the LLSSIM algorithm was used to estimate n_h and the values for the network weights using the calibration data. For each model, the calibration data was evaluated using three popular residual statistics: the root-mean-square error (RMSE), the A information criterion (AIC) [*Akaike*, 1974], and the B information criterion (BIC) [*Rissanen*, 1978]. The AIC and BIC are computed using the equations:

$$\text{AIC} = m \ln(\text{RMSE}) + 2\text{npar} \quad (9)$$

$$\text{BIC} = m \ln(\text{RMSE}) + \text{npar} \ln(m) \quad (10)$$

where, as defined earlier, m is the number of input-output patterns and npar is the number of parameters to be identified. These three are listed as models 1, 2, 3 and 4 in Table 1.

6. SMA MODEL IDENTIFICATION

The SMA model is a conceptual rainfall-runoff model that is one of the components of the National Weather Service River Forecast System (NWSRFS) used to convert precipitation inputs into streamflow outputs [*Burnash et al.*, 1973; *Peck*, 1976; *Kitanidis and Bras*, 1980a, b; *Panagoulia*, 1992a,b]. The inputs to the SMA model are precipitation and potential evapotranspiration. Precipitation is provided in the form of mean areal precipitation over elevation bands. The outputs from the model are estimated evapotranspiration and channel inflow; the latter is converted into streamflow by means of a unit hydrograph. The model was calibrated manually by using the15-year period that is a suitable length for calibration of conceptual models.

7. RESULTS AND DISCUSSION

The statistical performance of the identified ANN models for the five-year calibration period and the overall 15-year validation period, respectively, are summarized in Table 1. The results are presented and discussed below.

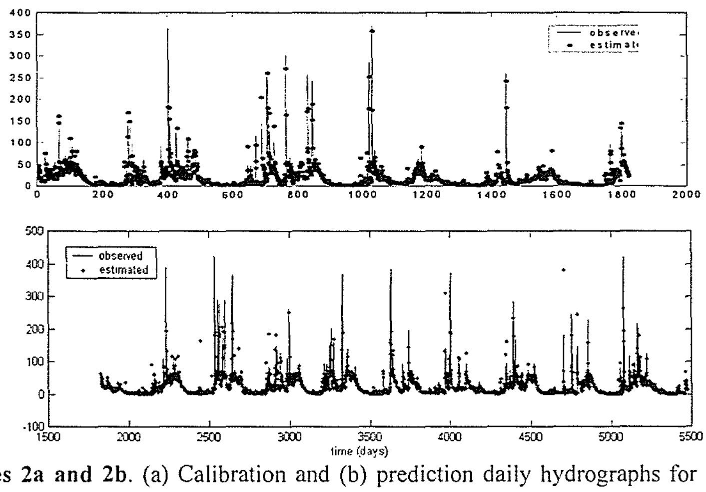
Table 1. Calibration and Validation Statistics for Five-Years Calibration Study

No	Model	RMSE		%VE		%MF		CORR	
		Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
1	ANN1(2,3,3)	16.785	18.170	-2.267	-2.019	-67.300	-71.100	0.8165	0.8411
2	ANN2(5,4,3)	13.190	16.277	-1.410	-1.790	-23.800	-53.700	0.8895	0.8730
3	ANN3(3,5,4)	12.440	16.185	-1.230	-1.480	-3.200	-57.210	0.9100	0.8750
4	ANN4(3,5,4)	11.820	17.600	-1.100	-1.840	-23.200	-61.600	0.9190	0.8500

RMSE denotes root-mean-square error; %VE, percent volume error; %MF, percent error of maximum flow; and CORR, correlation statistic

The RMSE statistic measures the residual variance with optimal value 0.0. The % VE statistic measures the percent error in volume (bias) under the observed and simulated hydrographs, summed over the data period; 0.0 is best, positive values indicate overestimation, and negative values indicate underestimation. The %MF statistic measures the percent error in matching the maximum (peak) flow of the data record; 0.0 is best; positive values indicate overestimation, and negative values indicate underestimation. The correlation (CORR) statistic measures the linear correlation between the observed and simulated flows with optimal value 1.0. On average, between calibration and validation periods, as well as among all the statistics the ANN3 (3,5,4) model performs best.

Figures 2a and 2b present the simulated by ANN3 (3,5,4) daily streamflow and observed one for the calibration and validation period respectively. The high-flows performance of ANN model is not so good. In the following paragraphs, the ANN3 (3,5,4) model is compared with SMA.



Figures 2a and 2b. (a) Calibration and (b) prediction daily hydrographs for ANN3 (3,5,4) model.

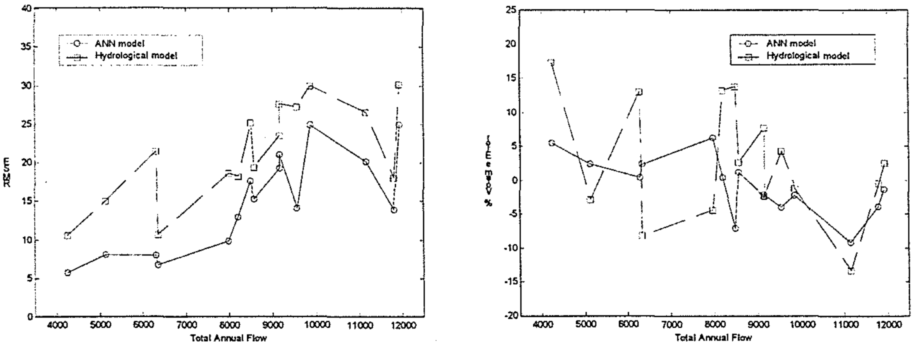


Figure 3. Annual performance statistics for each model plotted against total annual flow for each data year: (a) root mean square error, and (b) percent volume error.

Figures 3a and 3b present the RMSE and %VE statistics, for each model, computed separately for each of the fifteen years and presented as a function of the mean flow for that year. The RMSE statistic is presented in Figure 3a. There the lower mean flow years appear to have generally lower model RMSE. Clearly, the RMSE performance of the SMA is worse than that of the ANN on all years. The %VE statistic is shown in Figure 3b. The results indicate that the ANN model consistently has the smaller bias.

As a general notice, both models tend to fit the higher flows quite well (on a relative error bias), however the low-flow performance is not so good. Most likely, this is largely a result of the methods used for model identification; if the fitting criterion used to calibrate the models had been based on matching the logs of the flows, we would expect to see a more even distribution of error size across the entire flow range. In ongoing research, we are exploring the usefulness of using log flows for model identification.

Simulated versus observed flow plots of both models for the calibration and validation periods are given in Figures 4a and 4b for ANN3 (3,5,4) model and Figures 4c for Hydrological (SMA) model. Once again, we note that the models were calibrated to actual flows, while the data are presented using logs. Notice the tendency of the ANN and SMA models to underestimate in the very low flow range and to overestimate in the medium-flow range. The SMA model tends to have the largest deviations from the 1:1 line, while the ANN model shows the closest matching of simulated and observed over the entire flow range. As suggested before, this may indicate that the ANN model is implicitly doing a better job of representing the non-linearities inherent in partitioning precipitation into precipitation excess.

The residual autocorrelation functions for both models for the calibration and validation periods are presented in Figure 5. These plots indicate that the ANN 3(3,5,4) and SMA results in residuals are strongly systematic.

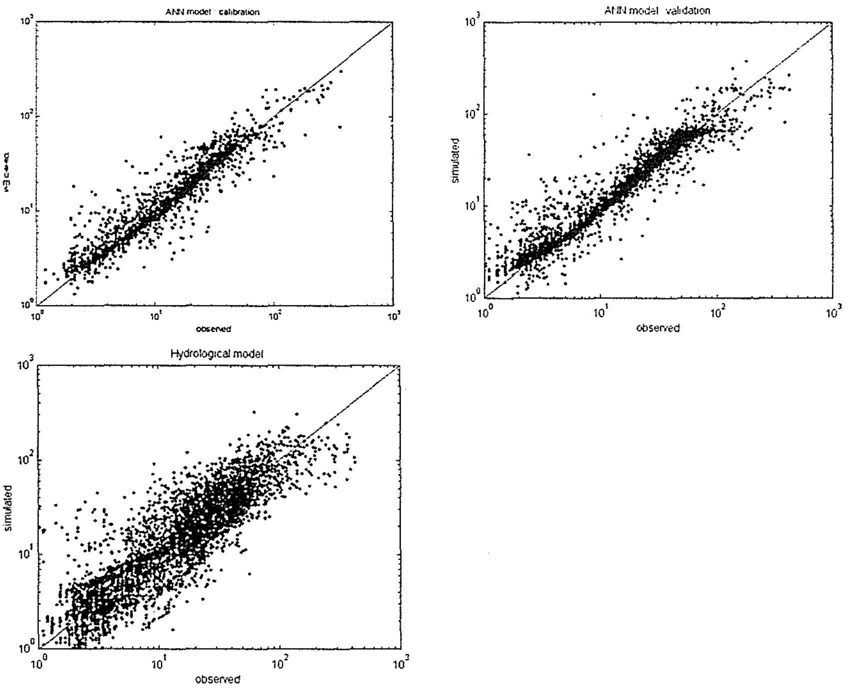


Figure 4. Scatterplots comparing simulated and observed flows for calibration and observed data.

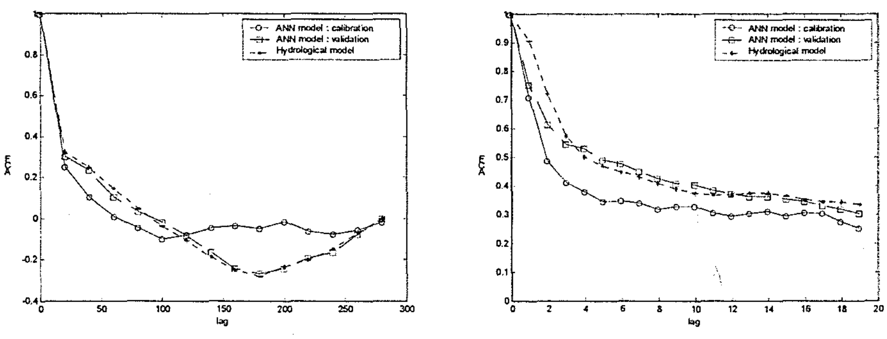


Figure 5. Autocorrelation function (ACF) of ANN 3(3,5,4), and SMA model.

8. SUMMARY AND CONCLUSIONS

The potential of artificial neural network models for simulating the hydrological behavior of mountainous catchment has been presented in this study. An efficient procedure (LLSSIM, *Hsu et al., 1995*) for estimating the weights (parameters) of a three-layer ANN was used. The nonlinear ANN model identified using the LLSSIM identification procedure seems to provide a better system theoretical representation of the rainfall-runoff relationship of the medium-sized mountainous Mesochora catchment, in Central Greece, than the conceptual SMA model.

Because the ANN approach presented here does not provide models that have physically realistic components and parameters, it is by no means a substitute for conceptual catchment modeling. However, the results suggest that the ANN approach may provide an alternative to SMA model for developing input-output simulation and forecasting models in situations that do not require modeling of the internal structure of the catchment.

REFERENCES

1. Akaike, H., A new look at the statistical model identification, *IEEE Trans. Automat. Control*, AC-19, 716-723, 1973
2. Burnash, R. J. E., R. L. Ferral, and R. A. McGuire, A generalized streamflow simulation system, *Reort. 220*, Jt. Fed-State River Forecast. Cent., Sacramento, Calif., 1973
3. Campolo, M., Andreussi, P., Soldati, A. River flood forecasting with a neural network model. *Water Resources Research, vol. 35, no. 10, p. 1191-1197,* 1999
4. Hecht-Bielsen, R. Neurocomputing, Addison-Wesley, Reading, Mass., 1990
5. Hsu, K., H.V.Gupta, and Soroosh Sorooshian, Artificial neural network modeling of the rainfall-runoff process, *Water Resources Research, vol. 31, no. 10, p. 2517-2530,* 1995
6. Imrie, C., E., Durucan, S., Korre, A. River flow prediction using neural networks: generalisation beyond the calibration range. *Journal of Hydrology 233 138-153,* 2000
7. Kitanidis, P.K., and Bras, R. L. Adapting filtering through detection of isolated transient errors in rainfall-runoff models. *Water Resources Research, vol. 16, no. 4, 740-748,* 1980a
8. Kitanidis, P.K., and Bras, R. L. Real-time forecasting with a conceptual hydrological model, 1. Analysis of uncertainty. *Water Resources Research, vol. 16, no. 6, 1025-1033,* 1980b
9. Nelder, A. J., and R. Mead, A simplex method for function minimization, *Comput. J.*, 7(4), 308-313, 1965
10. Panagoulia, D., G. Dimou, Linking space-time scale in hydrological modeling with respect to global climate change. Part 1. Models, model properties, and experimental design, *Journal of Hydrology 194 15-37,* 1997
11. Panagoulia, D., Hydrological modeling of a medium-sized mountainous catchment from incomplete meteorological data, *Hydrological Sciences Journal 137(1-4): 279-310,* 1992b
12. Panagoulia, D., Impacts of GISS-modeled climate changes on catchment hydrology. *Hydrological Sciences Journal 37(2): 141-163,* 1992a
13. Panagoulia, D., Hydrological response of a medium-sized mountainous catchment to climate changes. *Hydrological Sciences Journal 36(6): 525-547*, 1991
14. Peck, E. L., Catchment modeling and initial parameter estimation for National Weather Service River Forecast System, *NOAA Tech. Memo. NWS HYDRO-31*, Off. of Hydrol., Natl. Oceanic and Atm. Admin., Silver Spring, Md., *1976*
15. Raman, H., and Sunilkumar, R. Multivariate modeling of water resources time series using artificial neural networks, *Journal of Hydrology 40 145-163,* 1995
16. Rissanen, J., Modeling by short data description, *Automation*, 14, 465-471, 1978
17. Rumelhart, D. E., E. Hinton, and J. Williams, Learning internal representation by error propagation, in *Parallel Distributed Processing*, vol. 1, pp. 318-362, MIT press, Cambridge, Mass., 1986
18. Scalero, R.S. and N. Tepedelenlioglu, A fast new algorithm for training feed-forward neural networks, *IEEE Trans. Signal Process.*, 40(1), 202-210, 1992