AN ARTIFICIAL NEURAL NETWORK'S DYNAMICS FOR REPRESENTING THE RAINFALL – RUNOFF EXTREME PROCESS

D. Panagoulia (1), N.Maratos (2) and A. Grammatikogiannis (1)

(1) N.T.U.A., 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) N.T.U.A., Faculty of Electrical and Computer Engineering, Department of Signals, Control and Robotics, 9 Heroon Polytechniou, 15780 Zographou, Athens, Greece.

ABSTRACT

An algorithm combining a nonlinear simplex optimization method with linear least squares (LLSSIM), is presented to show the potential of a three-layer feed forward ANN models to represent the high and low rainfall-runoff relationship over a mountainous catchment. The output "rain plus melt" from the snowmelt model of the US National Weather Service (US NWS) applied on the mountainous Mesochora catchment in Central Greece was used as input to ANN model. The RMSE and % VE statistics were computed separately for flows above and below the 15-year mean daily flows. The nonlinear ANN model appears to provide a better representation of the rainfall-runoff extremes than the conceptual runoff model of the US NWS applied over the same catchment.

Η ΔΥΝΑΜΙΚΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ ΝΑ ΑΝΑΠΑΡΙΣΤΑ ΤΗ ΔΙΑΔΙΚΑΣΙΑ ΒΡΟΧΗΣ-ΑΠΟΡΡΟΗΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Δ. Παναγούλια (1), Ν. Μαράτος (2) και Α. Γραμματικογιάννης

(1) Ε.Μ.Π., Σχολή Πολιτικών Μηχανικών, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Εργων, Ηρ. Πολυτεχνείου 5, 15780 Ζωγράφου, Αθήνα, Ελλάδα, dpanag@central.ntua.gr/ τηλ: +3-2107722858.
(2) Ε.Μ.Π., Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ, Τομέας Σημάτων, Ελέγχου και Ρομποτικής, Ηρ. Πολυτεχνείου 9, 15780 Ζωγράφου, Αθήνα, Ελλάδα.

ΠΕΡΙΛΗΨΗ

Παρουσιάζεται ένας αλγόριθμος συνδυασμού της μη γραμμικής μεθόδου βελτιστοποίησης simplex και των ελαχίστων τετραγώνων (LLSSIM), για να δείξει τη δυναμική των νευρωνικών δικτύων με προσοτροφοδότηση τριών στρωμάτων να αναπαριστούν τις υψηλές και χαμηλές τιμές απορροής σε ορεινή λεκάνη. Το νερό βροχής και τήξης χιονιού που προέκυψε από την εφαρμογή του μοντέλου τήξης χιονιού της ΕΜΥ των ΗΠΑ στη λεκάνη της Μεσοχώρας αποτέλεσε δεδομένο εισόδου στο νευρωνικό μοντέλο. Υπολογίσθηκαν τα στατιστικά RMSE και %VE για ροές πάνω και κάτω της μέσης ημερήσιας απορροής. Το μη γραμμικό νευρωνικό μοντέλο έδειξε ότι αναπαριστά καλύτερα τις ακραίες συνθήκες βροχής-απορροής από το εννοιολογικό μοντέλο απορροής της ΕΜΥ των ΗΠΑ που εφαρμόσθηκε στην ίδια λεκάνη.

1. INTRODUCTION

The climatic variations observed over the past few years have given rise to record-breaking high and low flow events, which stresses the need for a flexible model being able to capture such extreme values. The application of artificial neural networks (ANNs) to various aspects of hydrological modeling has already gained a great interest since the ANNs are able to model nonlinear relationships and perform well with respect to conventional models. However, ANNs in capturing rainfall-runoff high and low values are restricted to the research environment [1], or they forecast poorly at high-and-low levels overestimating or underestimating the flows [2, 3, 4].

Perhaps, there are some reasons for inability of ANNs to predict extreme values, which could be met by some remedies. Some authors [e.g. 2, 3] attributed the underestimation of peak flows to a lack of information provided to the network, such as the antecedent rainfall. Some others [e.g. 4, 5] suggested to include more high-flow patterns in the training data sets, or to use log transformations in flow values to reduce the gap between the high and low flow conditions.

Model generalization in itself is another important consideration in predicting of extreme values. Care has to be taken during training so that the ANN does not become over-fitted to the training data and hence it captures only those relationships that are representative of the catchment [6]. A standard procedure is cross-validation, whereby another set of data is used to monitor the generality of the model during training [7, 8].

This study demonstrates the ability of the ANN approach in developing an effective rainfall-runoff extreme process without the intention to substitute a physically based conceptual model. The algorithm for training three-layer feed forward ANNs adopted here was a combination of linear least squares and nonlinear simplex optimization (LLSSIM) [5] since it has been found to perform best with respect to input-output function approximations [6]. The flows above and below the mean daily flow were considered as high and low values respectively. The ANN model was calibrated (trained) over low-flow values (dry years data) and it was validated over high and medium flow values, in order to be achieved a better generalization with respect to extreme events predicting.

The study is organized as follows. First, a brief overview of the architecture of three-layer feed forward ANN models and the LLSSIM training algorithm is described. Next, the LLSSIM methodology is used to develop an ANN model for the mountainous Mesochora catchment in Central Greece. Finally, the performance of this model with respect to extreme values is compared to the soil moisture accounting (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 NWS applied over the Mesochora catchment.

2. ARTIFICIAL NEURAL NETWORKS (ANNs) AND THE LLSSIM TRAINING ALGORITHM

Several ANN structures, e.g. feed forward networks, self-organizing feature maps, radial basis networks etc., have been proposed in the literature and their properties extensively studied. In this study a feed forward ANN [9] is used since such networks have been found to have very good performance when used as input-output approximations. A typical three-layer feed forward ANN is shown in Figure 1 [5]. The first layer (input layer) is connected directly to the input variables $x_i(p)$ and to hidden layer nodes, the second layer (hidden layer) is connected to both the input and output layer nodes while the third layer (output layer) is connected to hidden layer nodes and to the output variables $z_k(p)$.

The node j receives incoming signals x_i of each node from the previous layer. Within node j, a linear combination of the incoming signals x_i is formed, with weights w_{ji}, to produce the effective incoming signal s_j:

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

Still within node j, the signal s_j is passed through a nonlinear activation function (sigmoid function) to produce the output y_j of the node. The sigmoid function is a smooth monotonically increasing and bounded function. In this study, the logistic function,

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

is used as a sigmoid function.

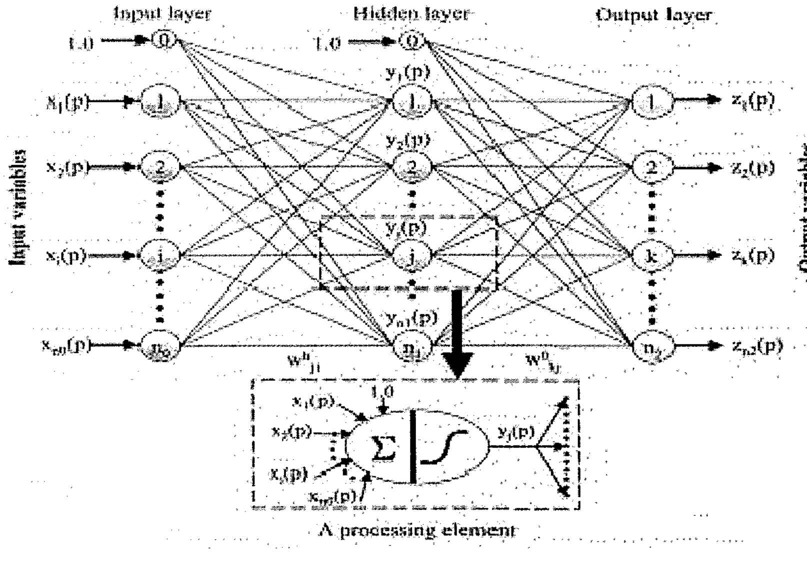


Figure 1. Typical three layer feed forward neural network [5].

Thus the outputs $y_j(p)$ and $z_k(p)$ of hidden layer node j and output node k respectively, can be expressed as follows:

$$y_j(p) = f(\sum_{i=0}^{n_0} w_{ji}^{h} x_i(p)) \tag{3}$$

$$z_k(p) = f(\sum_{j=0}^{n_h} w_{kj}^{o} y_i(p)) = f\left(\sum_{j=0}^{n_h} w_{kj}^{o} f(\sum_{i=0}^{n_0} w_{ji}^{h} x_i(p))\right) \tag{4}$$

where $x_i(p)$ is the input of input node i, w_{ji}^{h} is the connection weight from input node i to hidden node j, w_{kj}^{o} is the connection weight from hidden node j to output node k, n_0 is the number of input nodes and n_h the number of hidden layer nodes.

ANNs of this type can be trained to approximate a given input-output relationship. The objective of network training is to choose optimal connection weights w_{ji}^{h} and w_{kj}^{o} so that network predicted outputs $z_k(p)$ are a best fit to the measured output values, say $t_k(p)$. In our case, where the rainfall-runoff relationship is to be approximated, a single output value $t(p)$ (runoff) is measured, hence the approximating ANNs must have a single output node with output value $z(p)$ given by (4) with the index k dropped. In order to train the network, rainfall-runoff measurements are organized in m sets of input-output patterns (indexed by p), each having n_0 inputs $[x_1(p), x_2(p), \cdots, x_{n_0}(p)]^T$ and a single target value (runoff) $t(p)$. Training the ANN is equivalent to minimizing the following error function:

$$F(w^h, w^o) = \frac{1}{2}\sum_{p=1}^{m}(t(p) - z(p))^2 = \frac{1}{2}\sum_{p=1}^{m}\left[t(p) - f(\sum_{j=0}^{n_h} w_j^o y_i(p))\right]^2 = \\ = \frac{1}{2}\sum_{p=1}^{m}\left[t(p) - f\left(\sum_{j=0}^{n_h} w_j^o f(\sum_{i=0}^{n_0} w_{ji}^h x_i(p))\right)\right]^2 \tag{5}$$

with respect to the weights w_{ji}^h and w_j^o.

Many methods have been proposed in the literature for training ANNs. In this study we make use of the efficient LLSSIM network training algorithm, proposed by [5]. With respect to the input-to-hidden layer weights w_{ji}^h, the optimization is performed using the nonlinear simplex algorithm of [10].

In order to obtain explicit optimal values for hidden-to-output weights the LLSSIM method defines an approximately equivalent problem of linear least squares. It concerns: (a) the target values $t(p)$ are transformed backwards through the logistic function of the output node, i.e. $\tilde{t}(p) = \ln\left(\frac{t(p)}{1 - t(p)}\right)$ are obtained from $t(p)$, and (b) a new error function $F_1(w^o)$ is defined in terms of the transformed target values $\tilde{t}(p)$:

$$F_1(w^o) = \frac{1}{2}\sum_{p=1}^{m}\left[\tilde{t}(p) - \sum_{j=0}^{n_h} w_j^o y_j(p)\right]^2 \tag{6}$$

Now, minimization of $F_1(w^o)$ with respect to hidden-to-output weights w_j^o is a problem of linear least squares whose optimal solution $\hat{w}^o = [\hat{w}_1^o, \hat{w}_2^o, \cdots, \hat{w}_{n_h}^o]^T$ is explicitly obtained by solving the following set of linear equations:

$$\sum_{p=1}^{m} y_i(p) \sum_{j=0}^{n_h} y_j(p)\hat{w}_j^0 = \sum_{p=1}^{m} \tilde{t}(p) y_i(p), \qquad i = 1, \cdots, n_h \tag{7}$$

Of course, both the hidden node outputs $y_j(p)$ and the solution $\hat{w}^o$ of (7) depend on the input-to-hidden layer weights w^h, therefore we write $\hat{w}^o(w^h)$ for the optimal solution of (6).

The nonlinear part of the LLSSIM method can now be derived by replacing the optimal hidden-to-output weights $\hat{w}^o(w^h)$, obtained from (7), into (5):

$$\tilde{F}(w^h) = F(w^h, \hat{w}^o(w^h)) = \frac{1}{2}\sum_{p=1}^{m}\left[t(p) - f\left(\sum_{j=0}^{n_h} \hat{w}_j^o(w^h) f(\sum_{i=0}^{n_0} w_{ji}^h x_i(p))\right)\right]^2 \tag{8}$$

The error function $\tilde{F}(w^h)$ now depends only on input-to-hidden layer weights w^h and the resulting nonlinear least squares problem is solved by the nonlinear simplex method by [10].

3. 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 extreme events over a medium-sized mountainous catchment. The two model structures are (1) three-layer feed forward ANN model, and (2) the SMA model 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. The first five years of data (1972 to 1976) including the mostly low flows were used for model identification, while the remaining ten years (1977 to 1986) were used for ANN model validation including medium and high flows. Whole the 15-year period was used for SMA and ANN0 model calibration. Identification and calibration periods are long enough to extract representative results of catchment low-and-high flows behavior, in contrast to [5] study in which the identification period was only one wet year, considered as a disadvantage by the same authors.

4. ANN MODEL IDENTIFICATION

A first step in building an ANN model for the rainfall-runoff relationship is data normalization. Since sigmoid functions are used in ANN models, the hydrologic data should be normalized to the range $[0,1]$. Next, it is assumed that, at time t, the runoff output z(t) is related to past rainfall inputs r(t-j), $j=1,\cdots,n_a$ and to past runoff outputs $\tilde{t}(t-j)$, $j=1,\cdots,n_b$, i.e.,

$$\tilde{t}(t)=g\big(r(t-1),r(t-2),\cdots,r(t-n_a),\tilde{t}(t-1),\tilde{t}(t-2),\cdots,\tilde{t}(t-n_b)\big)+e(t) \tag{9}$$

Where, the unknown nonlinear mapping g(.) is to be approximated by the ANN model by minimizing the unknown mapping error e(t). Here n_a and n_b are the numbers of past inputs and outputs, respectively, contributing to the current output. Thus at time t the input vector $x(p)=[r(t-1),r(t-2),\cdots,r(t-n_a),\tilde{t}(t-1),\tilde{t}(t-2),\cdots,\tilde{t}(t-n_b)]^T$ is presented to the input layer nodes of the ANN. This input vector together with the current output (runoff) measurement $\tilde{t}(p)=\tilde{t}(t)$ constitutes the p^{th} input-output training pattern for the ANN model.

The notation $ANN(n_a,n_b,n_h)$ is used in the sequel to denote the above model structure, where $n_a+n_b=n_0$ is the number of nodes in the input layer, n_h is the number of nodes in the hidden layer, and the output layer consists of a single node. Thus an ANN model is identified by selecting values for n_a, n_b and n_h, and by using the LLSSIM algorithm in order to estimate network weights w_{ji}^h and w_j^o which minimize the prediction error function $F(w^h,w^o)$ of equation (5).

In this study several combinations of n_a, n_b, n_h were examined in order to make an assessment of the suitability of ANN models to represent the rainfall-runoff extreme process in mountainous catchment. Selected results of this exploration are given in Tables 1and 2 below.

5. 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 [11, 12]. The inputs to the SMA model are precipitation (SAA model snowmelt) 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.

6. RESULTS AND DISCUSSION

The statistical performance of the identified ANN models for the five-year calibration period and the overall 10-year validation period, on a daily basis, is summarized in Table 1. The results are presented and discussed below.

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

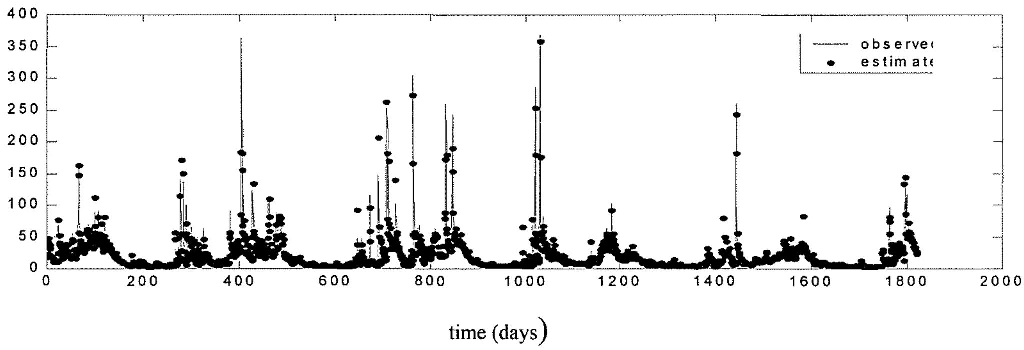
No	Model	RMSE		%VE		%MF		CORR	
		Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
1	ANN1(2,3,3)	16.785	18.170	-2.267	-2.019	-67.300	-71.100	0.8165	0.8411
2	ANN2(5,4,3)	13.190	16.277	-1.410	-1.790	-23.800	-53.700	0.8895	0.8730
3	ANN3(3,5,4)	12.440	16.185	-1.230	-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

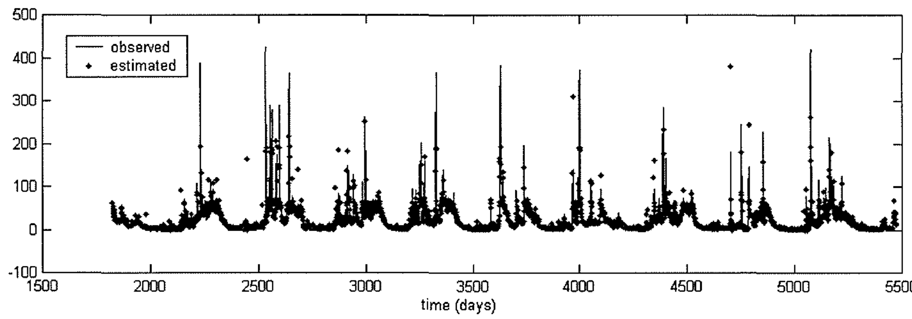
RMSE denotes root-mean-square error; %VE, percent volume error; %MF, percent error of maximum flow; and CORR, correlation statistic

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 (in mm) and observed one for the calibration and validation period respectively. The ANN3 (3,5,4) model tends to fit the low-flows quite well, while the high-flow performance is not extremely good.

In the following paragraphs, the ANN3 (3,5,4) model is compared with SMA, with respect to low and high flows. In order to examine these results in more detail, Table 2 and Figures 3a-3d present the RMSE and %VE statistics computed separately for flows above and below the 15-year mean daily flow ($24.41 m^3/s$). This provides an indication of performance on high– and low– flow events. Figure 3a presents the $RMSE_A$ statistic (root-mean-square error for above-mean flows). Clearly, the ANN model performs best for dry and wet years (whole the15-year period). Figure 3b presents to $RMSE_B$ statistic (root-mean-square error for below-mean flows). The ANN model performs best for all of the years, except for 1983, that both models present the same behavior.





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

TABLE 2. Statistics for Fifteen Years Calibration Study and Above-Below Mean Flows

						Above mean flow		Below mean flow	
No	Model	RMSE	% VE	% MF	CORR.	$RMSE_A$	$\%VE_A$	$RMSE_B$	$\%VE_B$
1	Hydrological	22.455	2.071	-77.752	0.749	37.313	-4.412	8.828	21.121
2	ANN0	13.77	-1.112	-41.96	0.911	23.521	-3.737	3.839	6.641

Figure 3c presents the $\%VE_A$ statistic (percent volume error for above-mean flows). The ANN model provides consistently low bias, while the SMA performance is poor. Figure 3d presents the $\%VE_B$ statistic (percent volume error for below-mean flows). The SMA model performance is very poor. We could point out that the poor performance of the SMA model does not reflect so much on the ability of the model as it does on the calibration procedure used. On the other hand, an 'expert' would not allow the model error bias at various flow levels to deteriorate so much while minimizing the error variance.

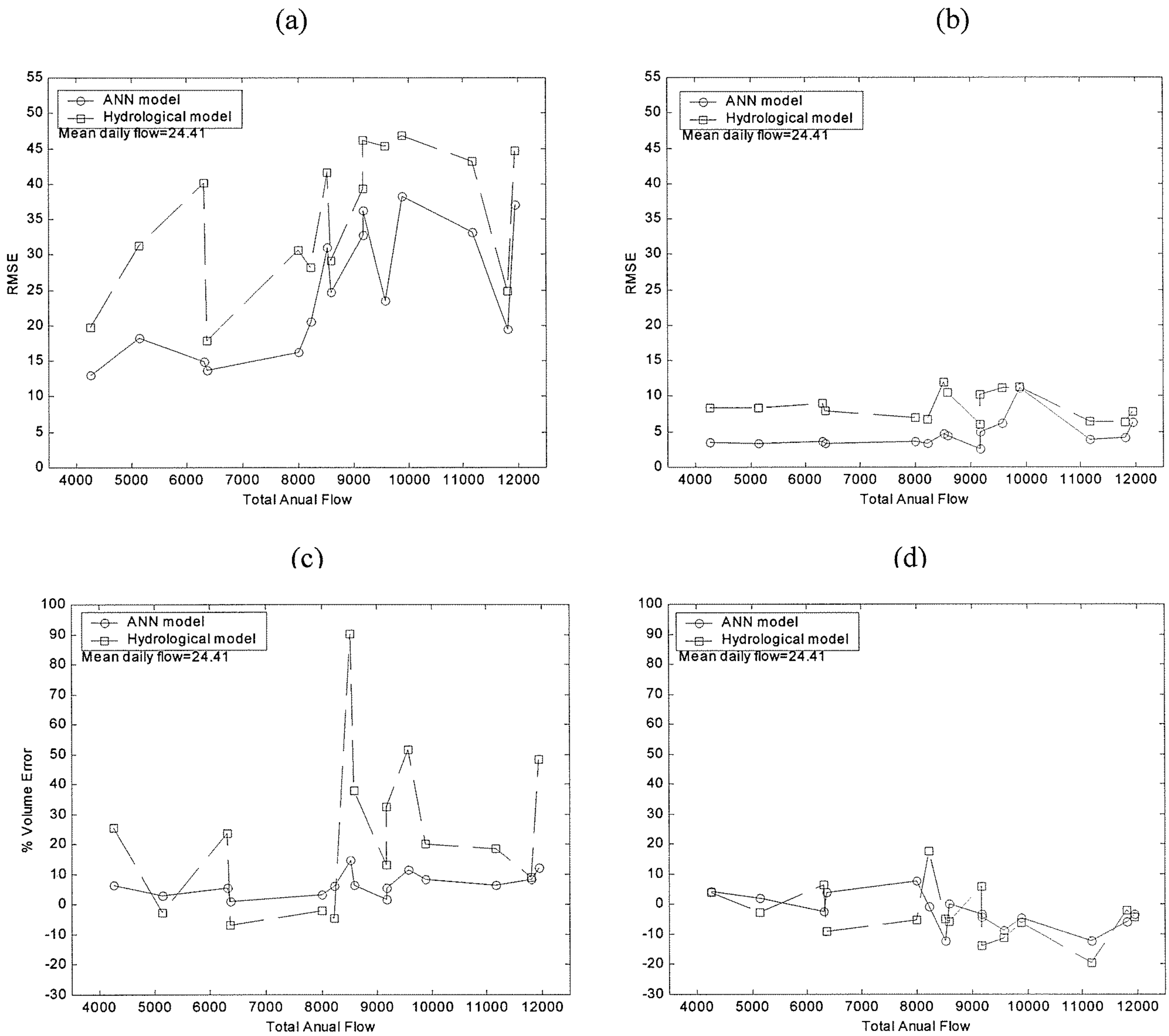


Figure 3 Annual performance statistics of above –and- below mean flow for each model plotted against total annual flow for each data year: (a) root-mean-square error for above-mean daily flow, (b) root-mean-square error for below-mean daily flow, (c) percent volume error for above-mean daily flow, and (d) percent volume error for below-mean daily flow.

7. SUMMARY AND CONCLUSIONS

The potential of artificial neural network models for simulating the hydrological behavior of extreme values over a mountainous catchment has been presented in this study. An efficient procedure (LLSSIM, [5]) 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 extreme values 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 of extreme flows in situations that do not require modeling of the internal structure of the catchment.

REFERENCES

1. Minns, A. W., and Hall, M. J., (1996) Artificial neural networks as rainfall-runoff models, **Journal of Hydrological Sci., 41, 339-417**.

2. Dawdson, C. D., and Wilby, R., (1998) An artificial neural network approach to rainfall-runoff modeling, **Journal of Hydrological Sci., 43, 14-66**.

3. Campolo, M., Andreussi, P., Soldati, A. (1999) River flood forecasting with a neural network model. **Water Resources Research, vol. 35, no. 10, p. 1191-1197**.

4. Karunanithi , N., Grenney, W. J., Whitely, D., and Bovee, K., (1994) Neural networks for river flow prediction, **Journal Computing Civil Enging. 8, 201-219**.

5. Hsu, K., H.V.Gupta, and Soroosh Sorooshian, (1995) Artificial neural network modeling of the rainfall-runoff process, **Water Resources Research, vol. 31, no. 10, p. 2517-2530**.

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. Panagoulia, D., and N. Maratos (2003a) Artificial neural network modeling of the rainfall-runoff extreme events. **Proceedings of 8th International Conference on Environmental Science and Technology**, Editor: I.P. Lekkas, Lemnos Island, Greece 8-10 September, p 48-56.

8. Panagoulia, D., and N. Maratos (2003b) Rainfall-runoff process in mountainous catchment with artificial neural network. **In Proceedings of XXX IAHR congress**, Theme D: hydroinformatics and Advanced Data Technology in Engineering Practice, p. 39-46, Editors: J. Ganoulis & P. Prinos, August 2003. AUTh, Thessaloniki, Greece.

9. Rumelhart, D. E., E. Hinton, and J. Williams (1986) Learning internal representation by error propagation, in **Parallel Distributed Processing**, vol. I. Pp. 318-362, MIT press, Cambridge, Mass.

10. Nelder, A. J., and R. Mead, (1965) A simplex method for function minimization, **Comput. J., 7(4), 308-313**.

11. Burnash, R. J. E., R. L. Ferral, and R. A. McGuire, (1973) A generalized streamflow simulation system, **Reort. 220**, Jt. Fed-State River Forecast. Cent., Sacramento, Calif.

12. Panagoulia, D., (1992) Hydrological modeling of a medium-sized mountainous catchment from incomplete meteorological data, **Hydrological Sciences Journal 137(1-4): 279-310**.