Flow estimation for ungauged catchments using a neural network method

Mohammad T. Dastorani[1] & Nigel G. Wright[2]

This research focused on the application of artificial neural networks for flood prediction in ungauged catchments. Catchment descriptors were used as input data and the index flood was the output of the model. Different types and numbers of catchment descriptors (17 descriptors and more than 1000 catchments) were used to choose those that gave the best relationship with the hydrological behaviour and flood magnitude. ANN models with different architectures were developed and applied to training and validation sets of data to find the best type of ANN for this application. Selection of pooling groups of catchments either randomly or according to geographical proximity did not produce desirable results. Therefore hydrologically similar catchments were clustered using the WINFAP-FEH Software before entering descriptors into the ANN model. This improved the accuracy of predicted floods.

Introduction

Floods mean many things to many people depending on the profession they are engaged in. To the civil engineers who are responsible for designing flood protection measures it means how to build up a design flood based on the knowledge of hydrology of the catchment area. This is required to plan engineering structures such as storage reservoirs, schedule of operations etc. Further as the flood wave passes through a river it is necessary to know how the stage varies with respect to time and distance for the design of river engineering works as well as for establishment and operation of flood warning systems by the civil authorities. To do this, the most important factor is to predict flood discharge magnitude accurately. Flood prediction is a complicated situation specially for the ungauged catchments considered here as there is no flow data to use.

The technique of artificial neural networks has been found to be a powerful tool for solving different problems in a variety of applications ranging from pattern recognition to system optimisation. A neural network is an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neurone. It is based on learning the process from inputs to outputs using a training data set, and it mimics these for a new set of inputs to reach corresponding outputs. Recently, it has been applied to problems of water resources engineering, river flow modelling, and catchment rainfall-runoff processing. The technique of artificial neural networks is widely used as an efficient tool in different areas of water related research activities. It has been used in different water research areas such as evaluation of rainfall-runoff relationship (Hsu, Gupta and Sorooshian 1995, Minns and Hall 1996).

The present study focuses on the application of artificial neural networks for river flood prediction of ungauged catchments using catchment descriptors. In addition, the identification of appropriate factors affecting flood flows of ungauged catchments (for use as inputs to the models) as well as the suitable architecture of neural networks for this particular application have been considered.

[1]School of Civil Engineering, University of Nottingham, Nottingham NG7 2RD, UK.
Email: evxmtd@nottingham.ac.uk
[2]School of Civil Engineering, University of Nottingham, Nottingham NG7 2RD, UK.
Email: nigel.wright@nottingham.ac.uk

Artificial neural networks

A neural network consists of a large number of simple processing elements that are variously called neurones, units, cells, or nodes. Each neurone is connected to other neurones by means of direct communication links, each with an associated weight. The weights represent information being used by the network to solve a problem. Neural networks operate on the principle of learning from a training set. They must be trained with a set of typical input/output pairs of data called the training set. The final weight vector of a successfully trained neural network represents its knowledge about the problem. In general, it is assumed that the network does not have any a priori knowledge about the problem before it is trained. At the beginning of training the network weights are usually initialised with a set of random values. There are a variety of neural network models and learning procedures. Two classes of neural networks that are usually used for prediction applications are feed-forward networks and recurrent networks. The neural network approach is a black-box approach and the user need not know much about the flow process. The ANN model structure is suited ideally for the modelling of highly non-linear relationship between inputs and outputs. In this research the software The NeuroSolutions neural networks environment produced by NeuroDimension, Inc. (NeuroSolutions, 2001) was used to construct neural network models for this study. ANN models with different architectures have been constructed and applied to training and validation sets of data to find the best ANN for this application. Different values for the parameters of learning rate, number of PEs, number of hidden layers, type of activation and output functions were tested for each architecture as well as for each set of data. Finally, it has been found that the Multi-Layer Perceptron (MLP) network with three layers, and Tangent hyperbolic function in hidden layer and Sigmoid function in output layer is the most accurate network for this purpose. Figure 1 shows a simple architecture of a typical 3-layer feedforward neural network, which has been used in this research. An advantage of MLP in terms of mapping abilities is its capability of approximating arbitrary functions. This is an important point in the study of nonlinear dynamics, and solving other function mapping problems.

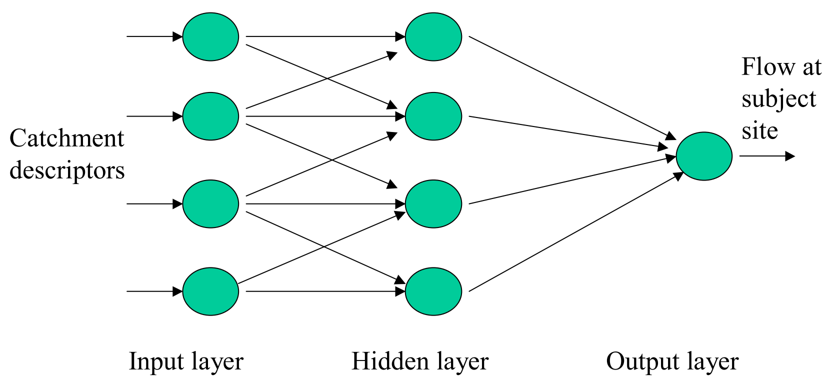

Fig. 1 A simple architecture of MLP neural network used in this research.

This type of the neural networks is normally trained with backpropagation algorithm. The backpropagation rule, propagates the errors through the network and allows adoption of the hidden processing element. It works with error correction learning to update the weights that from the system response at processing element i at iteration n $(y_i(n))$, and the desired response $(d_i(n))$ for a given input pattern an instantaneous error $(e_i(n))$ is defined by:

$$e_i(n) = d_i(n) - y_i(n) \qquad 1$$

Using the theory of gradient descent learning, the weights are adopted by correcting the value using:

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_i(n) x_i(n) \qquad 2$$

Where $w_{ij}(n+1)$ is the weight of processing elements i to j for the iteration $n+1$, $w_{ij}(n)$ the value of the same weight for iteration n, $\delta_i(n)$ is the local error computed directly from $e_i(n)$, and the constant η is the step size. Momentum learning, which has been used for the MLP neural network in this research, is an improved version of gradient descent in which a memory term is used to speed up and stabilize convergence. In this type of learning, weights are updated by following equation.

$$w_{ij}(n+1) = w_{ij}(n) + \eta\delta_i(n)x_i(n) + \alpha\left(w_{ij}(n) - w_{ij}(n-1)\right) \qquad 3$$

α is the momentum and normally is set between 0.1 and 0.9. The final weight vector of the successfully trained network, which represents its knowledge about the problem is used to apply to a new set of data to evaluate the performance of the model. A further parameter to be evaluated in this study was the number of processing elements (PEs) for the hidden layer of the neural network models. Results obtained from a simple test showing the quality of the results (R^2 for predicted and measured values) obtained with different numbers of PEs (10, 14, 50, 100 and 200) for a pooling group of catchments with seven inputs. this result showed that the variation of the accuracy of the outputs for different PEs is not considerable, although there is a decrease when the number of the PEs passes 100.

Each set of data was split into three parts for training, cross validation (to prevent over training), and testing (to test the model performance by using a trained network for a new set of data) purposes. In each pooling group, 60% of the catchments was used for training, 10% for cross validation, and 30% for testing of the model performance.

Identification of appropriate model inputs

As mentioned earlier, catchment descriptors have been used as the inputs to train and test artificial neural network models. However, some of the descriptors have more influence on runoff generation and flooding process than others. Identification of most important catchment characteristics, which have the strong relationship with runoff, is an important step to accurate prediction of the flow in an ungauged catchment. Although selection of the descriptors which have the most influence to runoff is the most important point, the number of descriptors used in the model is also important. Decreasing the inputs to make the modelling as simple as possible while keeping the accuracy of the prediction in an acceptable level, is a big advantage in terms of time and effort. It also affects usefulness of the model where there is limitation or additional cost for preparation or extraction of the required descriptors. To identify the most relevant descriptors to use as input of the neural network models, in this study the technique of neural network itself has been used. A multi-layer preceptron neural network was used with different types and number of inputs and the results were compared.

The number of input patterns is important in neural network modelling, as it has considerable influence on the ability of the models to process the task and go from input to output. A very small number of inputs may cause insufficient recognition by the network of the nature of the problem in order to map the relationship between inputs and outputs. A very large number of inputs may also lead to a complexity of the relationship and consequent poor performance of the modelling. From the hydrological point of view the optimum state is to use the minimum number of inputs whilst keeping the results to a desirable or acceptable level of accuracy. This was done in this case. Of course in addition to the number of descriptors the type of descriptors is very important, as usually runoff has a stronger relationship with some of the descriptors of the drainage catchment than others. Identification of these descriptors and using them will help to improve the accuracy of the outputs. Several simulations were carried out using different numbers and types of descriptors, and the results were compared to identify the most important number and types of the catchment descriptors to use as input to ANN models and predict river flood of the ungauged catcments. From figure 2, which shows correlation coefficient of the outputs from the ANN simulations using different number of inputs with the related measured values, it is clear that models with 5, 6, 7, 8, and 9 inputs represent the more accurate results. To get more confidence in the choice of the right number of inputs, the results of the tests has also been considered using root mean square error (RMSE), and the results confirm what is seen in figure 2. By comparing the the results of these tests and the desire to have as few inputs as possible, the number of 7 descriptors has been selected for further simulations. These are as follows:

- AREA = Catchment drainage area using an IHTDM-derived boundary (km^2).
- BFIHOST = Bas Flow Index derived using the HOST classification.
- SPRHOST= Standard Percentage runoff derived using the HOST classification.
- FARL= Index of Flood Attenuation attributable to Reservoirs and Lakes.
- SAAR= Standard period (1961-1990) Average Annual Rainfall (mm).
- SMDBAR= Mean SMD for the period 1941-70 calculated from MORECS monthend values (mm).
- PROPWET= Proportion of time when SMD was ≤6 mm during 1961-90.

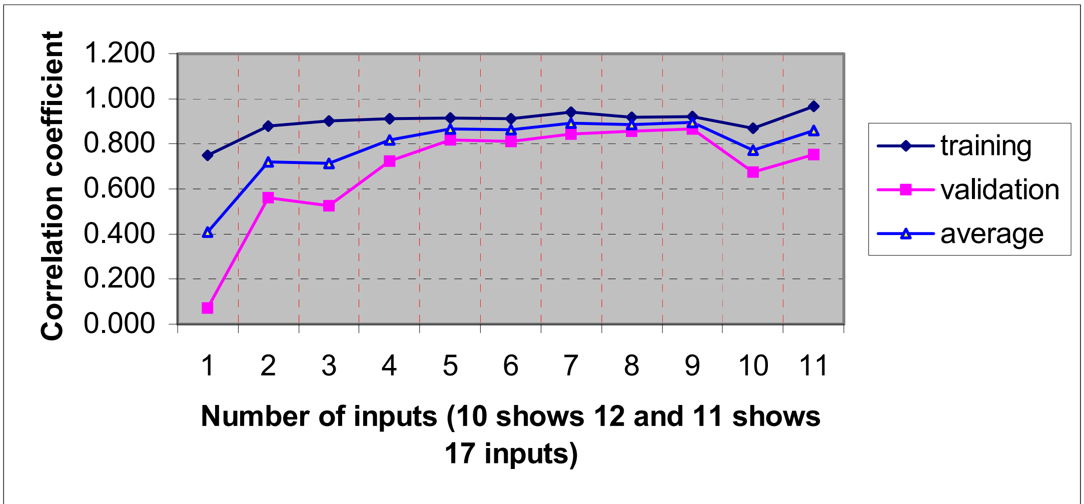

Figure 2 Squared correlation coefficient (R) between measured flow and predicted flow by ANN using different number of catchment descriptors in training and validation phases, and the average value

Randomly and geographically selected group of catchments

Two groups of catchments have been considered in this step. The first one, whose members have been selected randomly from all over the United Kingdom, contains 70 catchments. After selection of the members of the group, all descriptors of the catchments have been extracted and prepared to enter into the model. The second group, which contains 52 catchmens was formed according to an initial consideration of similarity in terms of drainage area and geographical location. For the second group, after specification of the catchments their descriptors have been extracted and used to train and test the models. All catchments in this group have an area less than 100 km^2. In each group, 60%, 30% and 10% of the members have been used for training, testing and cross validation purposes respectively. Figures 3 and 4 show the results obtained from the tests using the first and second groups in testing phases of the simulation.

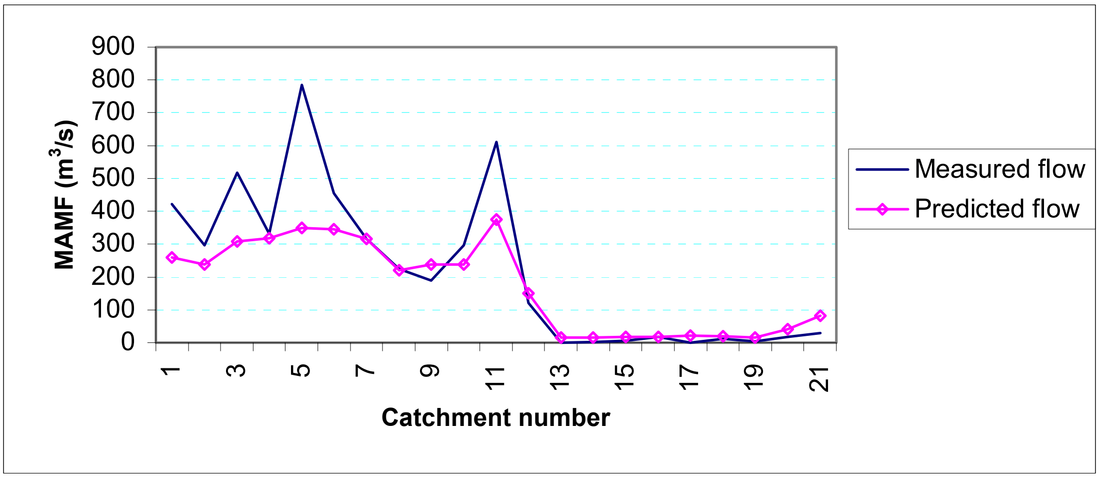

Figure 3 Results obtained from ANN model for a group of 70 random catchments against the actual values - testing phase

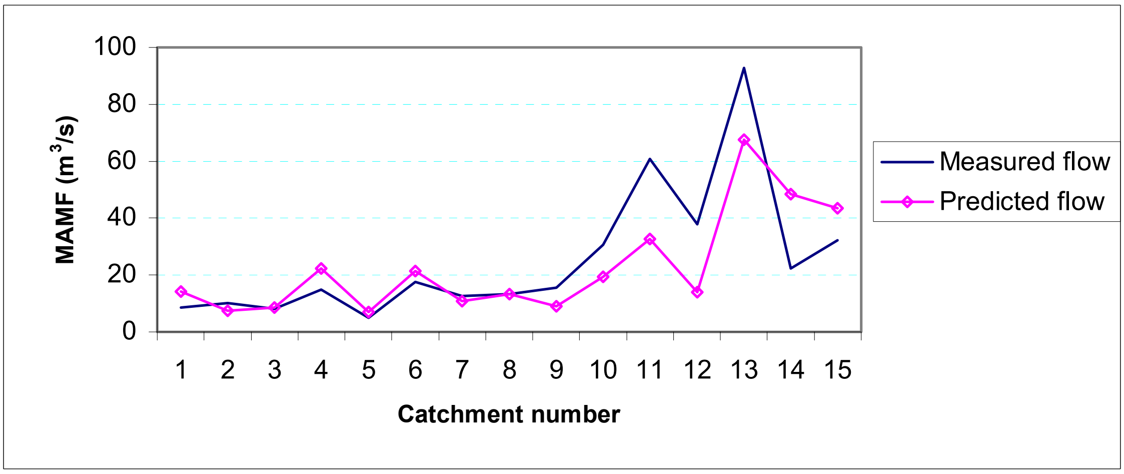

Figure 4 Results obtained from ANN model for a group of 52 selected catchments against the actual values - testing phase

As can be seen from the figures, the accuracy of the results is not satisfactory, and also there is no considerable difference between two results. Although in the second test catchments were selected from same geographical location and the same category of drainage area (all under 100 km^2), the results show there is no important improvement in output accuracy over the first test that has been selected randomly. The coefficient correlation between predicted and measured floods (R) for the first test is 0.98 and 0.87 for training and testing phases respectively. For the second test where catchments have been selected according to the area and geographical location R, is 0.91 and 0.79 for training and testing phases respectively. Although in the training phase the predicted flow data and measured flow values show a high correlation coefficient, they do not show good performance in the testing phase, which is more important in terms of the evaluation of model performance. This shows that the model cannot learn the process in training phase properly to be able to reach appropriate weights and apply them to the new set of data in testing phase.

Prediction results for pooling groups

The WINFAP-FEH was used to form pooling groups. This procedure identifies the catchments based on hydrological similarities. The ANN was trained again based on these pooling sets (60% of each pooling group) of catchments and was tested by 30% of the catchments in each pooling group. After employing the WINFAP-FEH and forming pooling-groups, the accuracy of the results of the neural networks was improved by about 14% and 7% for the groups containing 52 and 70 catchments respectively. The outputs of the model where closer to the measured values especially in testing phase. This shows that employing of the WINFAP-FEH in this stage has helped artificial neural networks to produce more practical outputs. Figure 5 shows the results from a pooling group of 52 catchments for the subject site of Tay in testing phase. For this test R was 0.96 and 0.93 for training and testing phases respectively. For another pooling group (formed for the subject site of Thurso with 70 catchments), the results have been shown in figure 6 for testing phase. For this test R was 0.98 and 0.94 for training and testing phases respectively.

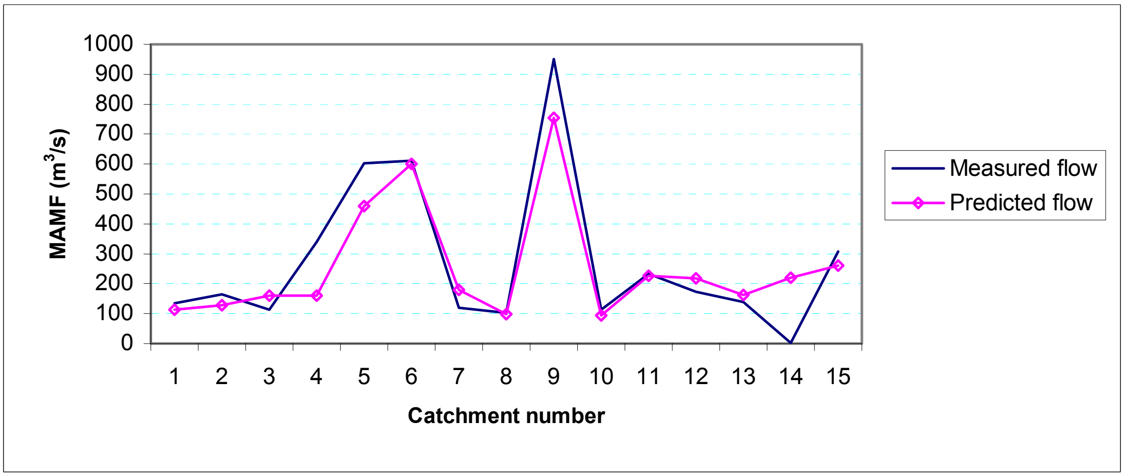

Figure 5 Results obtained from ANN model for a pooling group of 52 catchments (for Tay subject site) in testing phase.

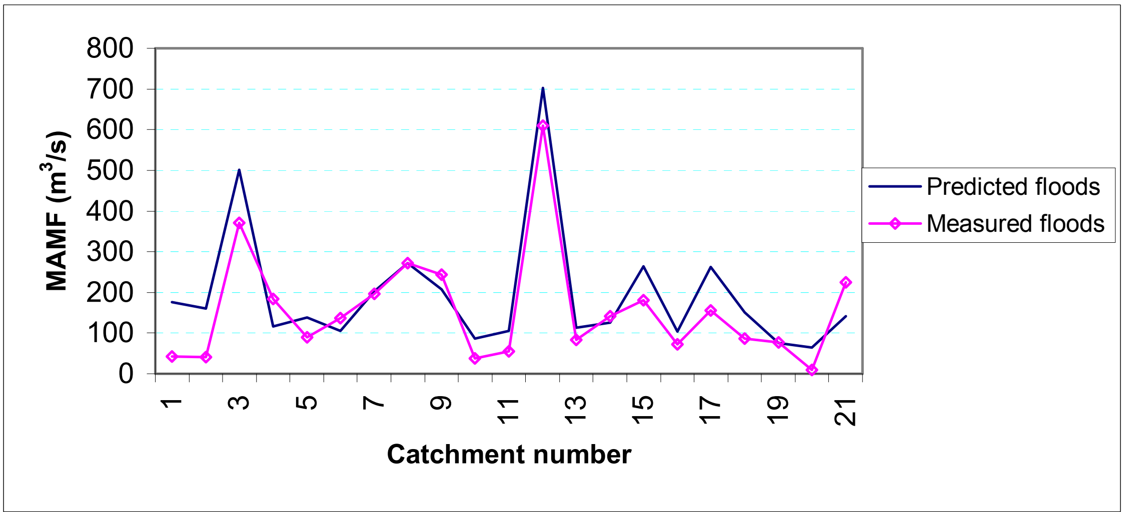

Figure 6 Results obtained from ANN model for a pooling group of 70 catchments (for Thorso subject site) in testing phases

Relationship between accuracy and group homogeneity

It was decided to analyse the relationship between two appropriate factors showing the accuracy of the model results (agreement between the predicted values and the actual values) and homogeneity of the pooling groups. The purpose was to consider the effects of homogeneity of the formed groups on the outputs of the ANN models. To meet this purpose, more pooling groups for subject sites located in different parts of the UK were formed and used to predict the flow. R^2 (squared of the correlation coefficient between predicted results and the measured values) was plotted against heterogeneity factor of the pooling groups, H_2. Evaluation of these two parameters for several different pooling groups show that an increased heterogeneity factor decreases the accuracy of the results of ANN predicted flows. First the number of pooling groups was increased to 6 and these two parameters were considered. Results have been plotted in figures 7. In terms of the pooling groups homogeneity evaluation, FEH suggest that a pooling group is homogenous when $H_2<2$, it is considered heterogeneous when $4>H_2>2$, and it is very heterogeneous when $H_2 >4$. Looking at these figures show that for six pooling groups except one of them (number 1) there is a good agreement between predicted and measured flow for the pooling groups when heterogeneity is smaller and especially when it is less than 1. Heterogeneity factor (H_2) of the pooling groups was also plotted against the root mean square error (rmse) of the results obtained from ANN models using these pooling groups. This graph also confirmed the clear effect of forming hydologically similar groups of catchments in performance of the technique of artificial neural networks in ungauged catchment flood prediction. The number of pooling groups was again extended to 12 and the results between H_2 and R^2 indicated almost the same conditions as for 6.

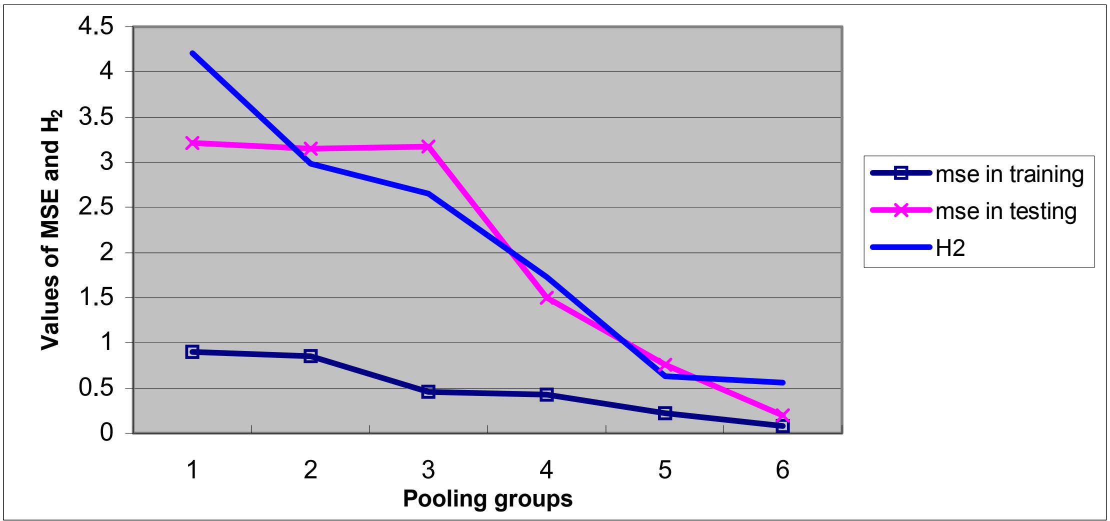

Figure 5.17 Heterogeneity factor (H_2) of WINFAP-FEH and coefficient correlation between predicted and measured peak flow (R^2) in testing phase for 6 different pooling groups.

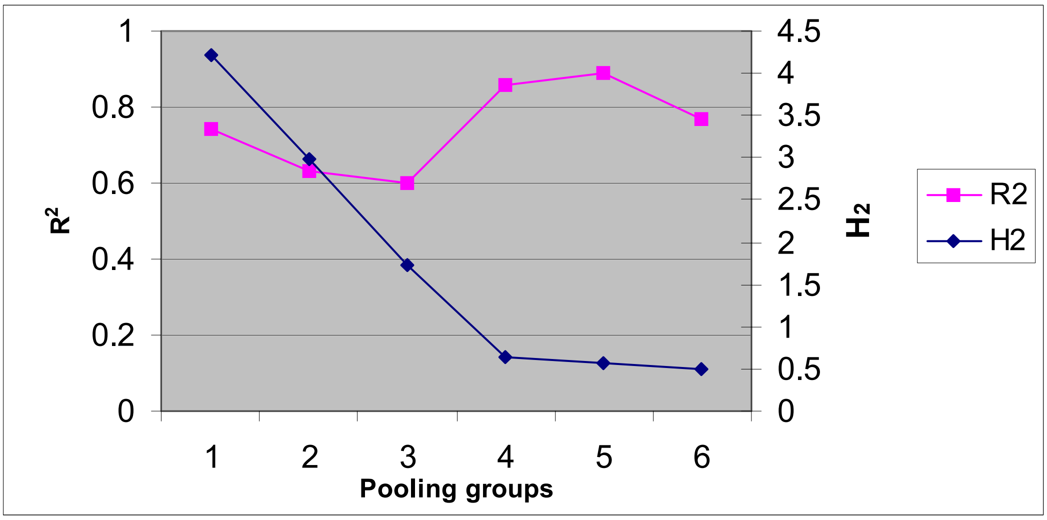

Figure 5.18 Relationship between the H_2 (pooling group heterogeneity factor) and the rmse (root mean square error between predicted and measured flow) in training and testing phases of ANN for the six pooling groups formed by WINFAP-FEH.

Comparison of the results to FEH method

The Flood Estimation Handbook (FEH) presented a method to estimate annual maximum flood (QMED) for rural catchments (URBEX<0.025) using catchment descriptors (see FEH volume 3 for more information). To evaluate the efficiency of the ANN based procedure presented in this research, the results of the testing phase data (catchments that have not contributed in training phase) were compared to the results obtained by FEH for those catchments. The testing phase results of four pooling groups formed in this research (Thurso, Derwent, Tay and Urel) were used for this comparison (only catchments with urbanisation factor less than 0.025 which has been mentioned in FEH). Root mean square error (rmse) and coefficient of efficiency (R^2) between the measured flow and predicted flow by ANN (presented in this research) and FEH were calculated to evaluate the accuracy of the estimations. Table 1 shows these values for the results of each pooling group.

Table 1 Root mean square error (rmse) and coefficient of efficiency (R^2) between the measured flow and predicted flow by ANN (presented in this research) and FEH

Group	Root mean square error (rmse)		Coefficient of efficiency (R^2)	
	FEH method	ANN method	FEH method	ANN method
Derwent	102.6608	87.97555	0.744576	0.812424
Thurso	65.18505	53.60212	0.765902	0.841705
Urel	137.0868	117.6087	0.5015	0.633096
Tay	145.3701	100.7791	0.658049	0.835656
Total together	147.9979	90.10736	0.474627	0.80525

Discussion

The initial results obtained from the ANN model of ungauged catchment flood prediction were not desirable. These results were the outputs of the model that used descriptors from randomly or geographically formed groups of catchments. However, after employing the WINFAP-FEH and forming pooling-groups, the accuracy of the results of the neural networks was improved considerably. This is due to the homogeneity of the catchments selected in pooling groups. In clustering the catchments by this method hydrological factors are considered and hydrological responses through the members of these pooling groups are more similar than the catchments clustered randomly or according to geographical proximity. Using descriptors extracted from hydrologically similar catchments in the ANN model gives more relevant representation of the whole drainage area to the model. It is to be expected that the ANN model using these input data can produce more reliable results. By looking at the outputs in this stage it is clear that the outputs of the model were closer to the measured values especially in the testing phase, which shows the ability of the model to cope with new sets of data.

Another finding of this research was the specification of the catchment descriptors that have the highest influence on flood magnitude. This is a very important point as the type of inputs is one of the most important factors in the success of this type of investigation. Several types and number of catchment descriptors can be used as inputs and there is no given rule representing the catchment characteristics that have the strongest relationship with the discharge. The neural network technique was again used to do this. Several simulations were carried out using different numbers and type of descriptors, and the results were compared to identify the most important number and types of the catchment descriptors to use as input to ANN models. After a correlation coefficient analysis for the outputs of the ANN simulations with different number of inputs, it became clear that models with 5, 6, 7, 8, and 9 inputs represent the more accurate results. Calculation of the values of root mean square error (rmse) for the results of the tests also confirmed what was seen earlier in correlation coefficient analysis. Finally by taking this finding into account and also attempting to have as small as possible number of inputs, the number of 7 descriptors was found to be the most appropriate number of inputs for further simulations. In addition to the number of inputs the type of inputs was nother point to be cleared. Several tests were carried out by selection and use of different types of the descriptors for each identity test the most relevant types of descriptors were specified for each identity test. As a result the descriptors of AREA, SAAR, BFIHOST, SPRHOST, FARL, SMDBAR, PROPWET were found to be the most relevant inputs to the ANN model to predict flow of ungauged catchments. By looking at these descriptors it becomes clear that the results taken from these tests are reliable. These descriptors are physically the most important factors in terms of catchment hydrological analysis. These descriptors represent characteristics such as drainage area, rainfall, river base flow index, lakes and reservoirs, and catchment soil property and moisture. These characteristics are very important in runoff analysis and some of them are also the main parameters used in empirical methods of flood prediction. Identification of the type and number of catchment descriptors as inputs of the model showed its maximum effect when the WINFAP-FEH pooling groups were used to the model.

The relationship between accuracy of the results and homogeneity of the pooling groups was evaluated. This was carried out to find effects of homogeneity of the formed groups on the outputs of the ANN model. More pooling groups were created to complete this evaluation, and heterogeneity factor (H_2) was considered against the closeness of the predicted peak flow to measured values (R^2) for several different pooling groups. The result was that an increased heterogeneity factor decreases the accuracy of the results of ANN predicted flows, and normally a good agreement between predicted and measured flow is seen when heterogeneity is less than 1.

Consideration of the pooling group homogeneity factor and the closeness of the predicted results to the measured values indicates that the forming of the pooling groups is efficient only when the groups have a high enough level of homogeneity. In this study for the groups with H_2 less than 1, the neural network model produces predictions close enough to the measured values. It can be said that ANN seems to be an appropriate tool to model river flow and predict peak flows for unguaged catchments or catchments with a short record period of data when a suitable method is used to identify hydrologically similar catchments in the region. This research once again indicates that selection of the catchments to form pooling groups based on geographical proximity is not an efficient way in most cases. Homogeneity of the catchments must be considered based on hydrological parameters, which show similarity in reaction to precipitation, runoff generation and hydrological responses of the catchments.

By comparing the results of the procedure presented in this research to those produced by the FEH method it was shown that the accuracy of the outputs presented by this method is higher than the FEH method. For all four pooling groups considered R^2 for ANN method results is higher than the FEH method. The difference in R^2 is between 7 to 18 percent for the results in different pooling groups (table 1). This comparison shows that ANN can be an appropriate alternative to produce more practical predictions in ungauged catchments.

Conclusions

In this paper the application of an artificial neural network for ungauged sites was evaluated. In addition to the type of the compatible neural networks for this problem, the type and number of catchment descriptors were found important to get an acceptable result. By choosing the right type and number of basin characteristics as inputs, and compatible type of the neural network model as well as suitable model parameters, this technique gives an efficient tool to solve the problem of sites where the lack of data limits the efficiency of other modelling tools. In terms of flow prediction for these sites, correct selection of the members of the groups seems necessary. This should be done using a sophisticated method, which can select the catchments according to their similarity in response to the hydrological events, as geographical proximity does not seem so efficient in terms of catchment group homogeneity. The results obtained from this research were compared to those of the method presented by FEH for ungauged sites. Values of root mean square error and coefficient of efficiency calculated for the results of these two methods showed that the results of this research are more accurate than those produced by FEH method.

References

Burn, D.H., Evaluation of regional flood frequency analysis with a region of influence approach, Water resources research, 26 (10), 1990.

Dastorani, M. T. and N. G. Wright*, Application of artificial neural networks for ungauged catchments flood prediction*, Floodplain Management Association conference, 13-16 March 2001, San Diego, USA.

Hsu K., H. V. Gupta, and S. Sorooshian, *Artificial neural network modeling of the rainfall-runoff process*, J. of Water resources reseach, Vol. 31, No. 10, pp. 2517-2530, October 1995.

Institute of Hydrology, *Flood Estimation Handbook, Volumes 1-5*, 1999.

Institute of Hydrology, *FEH CD-ROM*, 1999.

Institute of Hydrology, *FEH Software*, 1999.

Minns A. W. and M. J. Hall, *Artificial neural networks as rainfall-runoff models*, J. of Hydrological Sciences, 41 (3) June 1996.

NeuroDimensions*, NeuroSolutions*, www.nd.com, 2001.

Wright N. G. and M. T. Dastorani, *Effects of river basin classification on Artificial Neural Networks based ungauged catchment flood prediction*, in the Proceedings of the 3rd International Symposium on Environmental Hydraulics, December 5-8, 2001, Phoenix, USA.