# Artificial neural network based real-time river flow prediction

**Mohammad T. Dastorani & Nigel G. Wright**

School of Civil Engineering, University of Nottingham, Nottingham NG7 2RD, UK

**Abstract** The potential of artificial neural network models for simulating the hydrologic behaviour of catchments is presented in this paper. The main purpose has been the modelling of river flow in a multi-gauging station catchment and real time prediction of peak flow downstream. The study area covers the Upper Derwent River catchment located in River Trent basin. The river flow has been predicted using upstream measured data. Three types of ANN were used for this application: Multi-layer perceptron, Recurrent and Lagged time recurrent neural networks. Data of different lengths (1 month, 6 months and 3 years) have been used, and flow with 3, 6, 9 and 12 hours lead-time has been predicted. In general, although the ANN shows a good capability to model river flow and predict downstream discharge by using only upstream flow data, however the type of ANN as well as the characteristics of the training data were found very important factors affecting the efficiency of the results.

## 1 Introduction

Computer models of some of the principal hydrological and hydraulic components of river floods have been successfully developed in recent years. Flood forecasting models, even in large catchments where the lead-time is long, are of most value when they operate in real-time. The lead-time is most important when flood forecasting is used to operate a flood warning system. For flow forecasters in small river basins, the achievement of an adequate lead-time for forecasts is more complicated as it is strongly dependent on the accuracy of the weather forecasts for a reasonable period. In a very small catchment, a good forecast can be made on the basis of the intensity of the rainfall. However, for larger catchments the flow characteristics will strongly influence the ability to make a good forecast. For these kinds of catchments, the effects of catchment storage, backwater and tidal effects are seen to be important in the process of river flow forecasting.

Reliable prediction of flow discharge and its variability along a river is an essential part of surface water planning projects. In addition, prediction of discharge in a particular point using upstream conditions helps design storage zones or control measures. 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 to process rainfall-runoff relationships (Hsu, Gupta and Sorooshian 1995, Minns and Hall 1996, Dawson and Wilby 1998, Tokar and Johnson 1999,); flow predictions (Karunanithi and Grenney 1994); and stage-discharge relationships (Bhattacharya and Solomatine 2000). In this work it was decided to investigate the application of ANN to model river flow discharge variation between gauging stations and real-time prediction of flow at downstream points using upstream discharge. This would be

Hydroinformatics 2002: Proceedings of the Fifth International Conference on Hydroinformatics, Cardiff, UK
 ISBN 1 84339 021 3 (set)

a useful step to prevent damage by predicting flood flow in flood plains, where the risk of flooding is high, using the measured flow data from upstream gauging stations. Three different types of ANN have been applied to evaluate the applicability of the technique for this particular application.

## 2 Real-time forecasting process

Improvements in river peak flow forecasting have resulted partly from the global increase in stream gauging stations and partly from accelerating advances in the technology of data collection as well as in computer-based data handling and telecommunication systems. The three main parameters in flood forecasting are accuracy, reliability, and timeliness. In flood warning applications there is likely to be more emphasis on timings and the reproduction of distinctive shapes on the rising limb and crest segment of the hydrograph. Prediction of flow by the method presented in this research is carried out for the next time step using the measured data at previous time steps. If Q represents the discharge at time t:

$$
\begin{aligned}
Q(t+1) = f(&Q_{(t)}, Q_{(t-1)}, \ldots, Q_{(t-n)} + q_{(t)}(1), q_{(t-1)}(1), \ldots q_{(t-n)}(1) \\
&+ q_{(t)}(2), q_{(t-1)}(2), \ldots q_{(t-n)}(2) \\
&+ q_{(t)}(3), q_{(t-1)}(3), \ldots q_{(t-n)}(3) + \mathrm{e}_{(t)})
\end{aligned}
$$

Where:

$f$ is an unknown non-linear mapping function.

$e_{(t)}$ is an unknown mapping error.

$n$ is the number of past inputs contributing to predict flow discharge at the next time-step.

$Q_{(t+1)}$ is river flow at Whatstandwell for next time step which is predicted using upstream flow at time $t$, $t-1$, ..., $t-n$, and $q_{(t)}(1)$, $q_{(t)}(2)$ and $q_{(t)}(3)$ are corresponding river flow at upstream three stations.

30-minutes flow data and four different lead times were used in this research, and predictions are as follows:

Q with 3 hours lead time = $f(Q_{(t-7)}, Q_{(t-8)}, \ldots, Q_{(t-n)} + q_{(t-7)}(1), q_{(t-8)}(1), \ldots q_{(t-n)}(1) + q_{(t-7)}(2), q_{(t-8)}(2), \ldots q_{(t-n)}(2) + q_{(t-7)}(3), q_{(t-8)}(3), \ldots q_{(t-n)}(3) + e_{(t)})$

Q with 6 hours lead time = $f(Q_{(t-13)}, Q_{(t-14)}, \ldots, Q_{(t-n)} + q_{(t-13)}(1), q_{(t-14)}(1), \ldots q_{(t-n)}(1) + q_{(t-13)}(2), q_{(t-14)}(2), \ldots q_{(t-n)}(2) + q_{(t-13)}(3), q_{(t-14)}(3), \ldots q_{(t-n)}(3) + e_{(t)})$

Q with 9 hours lead time = $f(Q_{(t-19)}, Q_{(t-20)}, \ldots, Q_{(t-n)} + q_{(t-19)}(1), q_{(t-20)}(1), \ldots q_{(t-n)}(1) + q_{(t-19)}(2), q_{(t-20)}(2), \ldots q_{(t-n)}(2) + q_{(t-19)}(3), q_{(t-20)}(3), \ldots q_{(t-n)}(3) + e_{(t)})$

Q with 12 hours lead time = $f(Q_{(t-25)}, Q_{(t-26)}, \ldots, Q_{(t-n)} + q_{(t-25)}(1), q_{(t-26)}(1), \ldots q_{(t-n)}(1) + q_{(t-25)}(2), q_{(t-26)}(2), \ldots q_{(t-n)}(2) + q_{(t-25)}(3), q_{(t-26)}(3), \ldots q_{(t-n)}(3) + e_{(t)})$

## 3 Study site and data manipulation

The study area covers the upper Derwent River catchment located in the River Trent basin. River flow has been predicted using upstream gauging stations data. Three gauging stations called Matlock, Chatsworth and Mythambdge have been selected with distances of about 10, 25 and 50 kilometres from the subject site of Whatstandwell gauging station respectively. Figure 1 shows an index plan of the upper Derwent river catchment and related discharge measuring stations.

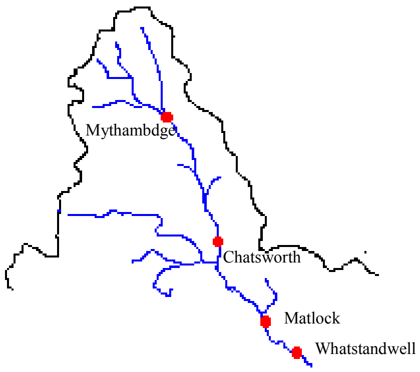


Figure 1. Plan of upper Derwent river catchment showing discharge stations

Measured flow data for related gauging stations were collected from the Environment Agency. In this research flow data at 30 minute intervals during January 1999, the first six months of the year 2000, and years 1998, 1999 and 2000 were used as input in different tests. For each experiment the data were split up to three parts; one for model training, one for cross validation (to prevent model over training), and another for testing the performance of the model. To prepare the data for ANN, and to determine the optimum time delay, correlation analysis was carried out. The correlation between discharges at Whatstandwell (the forecast station) and the other three upstream stations facilitated the calculation of the lag times. The results show that maximum correlation coefficients of Matlock, Chatsworth and Mythambdge with Whatstandwell are obtained with lag times of 1.5, 3.5 and 5.5 hours respectively.

## 4 Artificial neural network models

Neural networks must be trained with an appropriate set of typical input/output pairs of data called the training set (in supervised learning). In addition selection of an appropriate network structure for the problem to be solved is crucially important. Three types of ANN were used for this application called Multi-layer perceptron, Recurrent, and Time lag recurrent neural networks.

### 4.1 Multi-layer perceptron neural network

In this network a connection is allowed from a node in layer i only to nodes in layer i+1, and not vice versa. An advantage of MLP in terms of mapping abilities is its capability to approximate arbitrary functions. This is an important point in the study of non-linear dynamics, and solving other function mapping problems. Different types of transfer and output functions for hidden and output layers as well as different numbers of hidden layers were used to find the best structure of MLP for this application. Finally it was found that the tangent hyperbolic function was the most compatible one for the hidden layer. However, for the output layer the sigmoid function was the most suitable. The use of one hidden layer was most suitable for this model.

### 4.2 Recurrent neural networks

This type of network can be divided into fully and partially recurrent. Although recurrent networks are more powerful than feedforward networks, they are difficult to train and their properties are not well understood. The training of a recurrent network is much more sensitive to divergence. To construct the best architecture for the network many variations were implemented and the results considered. The number of hidden layers, number of processing elements in hidden layers, type of transfer and output functions and type of learning rule and its parameters have been considered and evaluated. After using different types of transfer and output functions for hidden and output layers, it was finally found that a tangent hyperbolic function was the best one for the hidden layer. However, for the output layer the sigmoid function has been the best one for all tests. Between the dynamic processing elements of Gamma, Laguarre and Time delay, the Laguarre and Time delay gave better results. The number of hidden layers differed from one in tests with a shorter length of input data to two for a longer length of input data. For the tests of this research, a partially recurrent network showed better adaptation than a fully recurrent one

### 4.3 Time lag recurrent neural networks

This type of network contains locally recurrent layers with a single adaptable weight. It usually suits temporal problems with short temporal dependency however it does not seem appropriate for more difficult temporal problems. For this type of neural network model it was found that tangent hyperbolic function and in few cases sigmoid function was the best one for hidden layer. However, for output layer the sigmoid function was the best one for all tests. Between the dynamic processing elements of Gamma, Laguarre and Time delay, the Gamma was found to be the most compatible for this application using this kind of neural network. Networks with only one hidden layer presented better performance.

## 5 Results

Flow at Whatstandwell has been predicted in four different experiments referred to as test 1 to 4. The length of the flow data input was one month, six months and three years in tests 1, 2 and 3. In test 4 the number of input patterns to the ANN models has been investigated. In this test several simulations have been used to predict flow at the subject site. These used the three ANNs and data from one, two and three upstream gauging stations. Figure 2 shows the predicted hydrographs with different lead-times for each type of neural network for one month of input data (test 1). Space does not allow for presentation of the result hydrographs for tests 2, 3 and 4. However, to consider the accuracy of the results and compare applicability of different types of ANN used in this research, mean squared error (rmse) and coefficient of efficiency ($R^2$) values for different length of input data and lead-times of the first three tests have been shown in table 1. As the table shows the Multi-layer perceptron network (a static network) did not perform well particularly for longer prediction horizons. For simulations using a long period of input data the Time lag recurrent network presented poor results although its outputs for simulations with shorter length of record is quite good and almost comparable to the Recurrent network. The Recurrent network has given the best results for all prediction horizons as well as against all three input data (record lengths).

Using different numbers of input patterns for the model (test 4) by presenting data from different numbers of upstream gauging sites did not indicate considerable difference between the results.

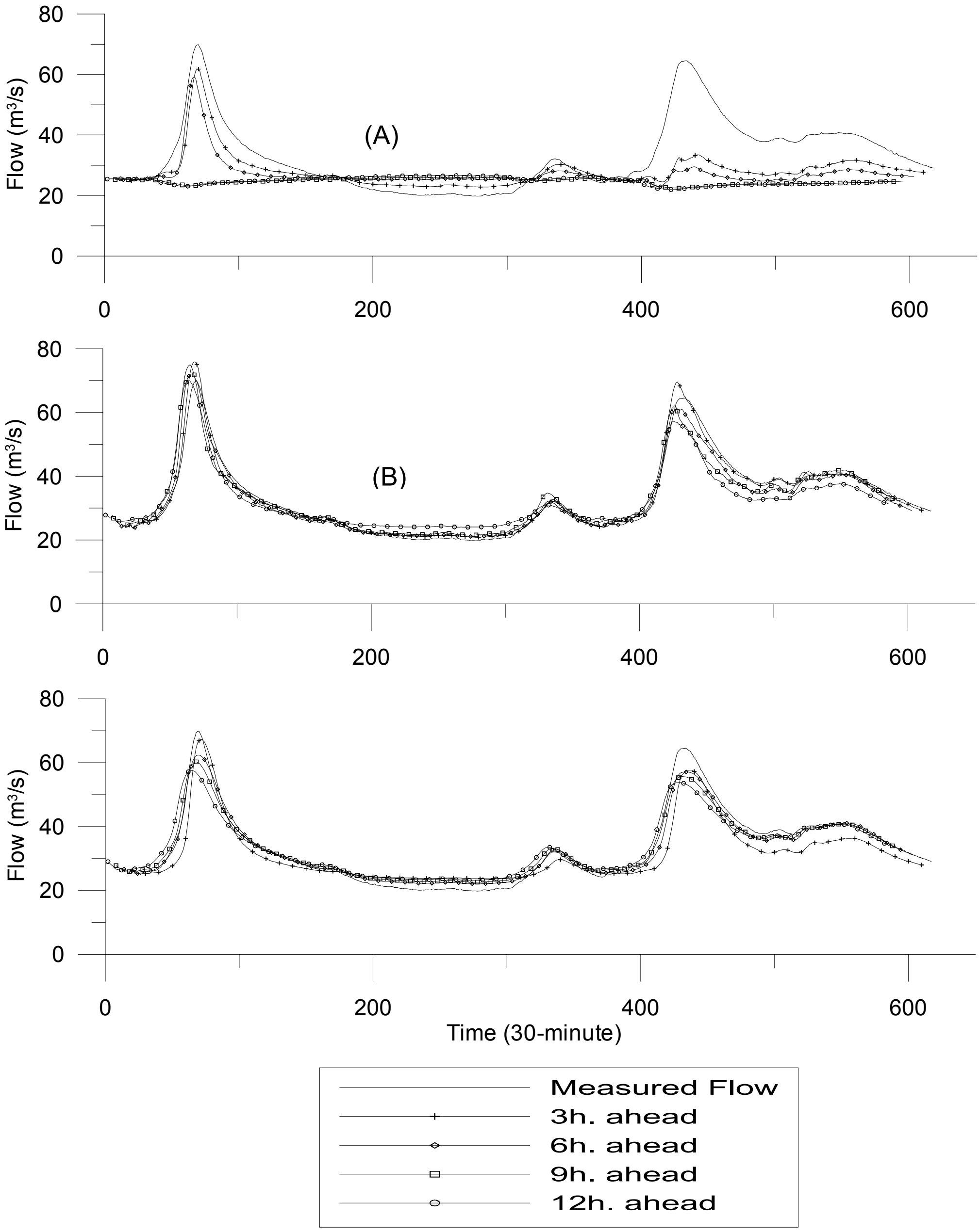


Figure 1 Results of the predicted flow by different types of neural networks and for different prediction lead-times: (A) Multi-layer perceptron, (B) Recurrent, (C) Time lag recurrent.

Table1 The values of rmse and $R^2$ of the predicted flow with different type of the neural networks, input data length and lead time in training and testing phases (for tests 1 to 3).

| Length of input data | Pred. Lead time | Root mean square error (rmse) | | | | | | Coefficient of efficiency ($R^2$) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Training | | | Testing | | | Training | | | Testing | | |
| | | MLP | Rec. | T.l. rec | MLP | Rec. | T.l. rec | MLP | Rec. | T.l. rec | MLP | Rec. | T.l. rec |
| 1 mo. | 3 h | 1.727 | 2.125 | 5.374 | 9.484 | 1.646 | 2.982 | 0.988 | 0.982 | 0.886 | 0.341 | 0.980 | 0.931 |
| | 6 h | 3.480 | 3.012 | 4.604 | 11.485 | 1.818 | 3.006 | 0.952 | 0.964 | 0.916 | 0.043 | 0.976 | 0.934 |
| | 9 h | 4.379 | 3.625 | 4.932 | 15.505 | 2.185 | 3.243 | 0.924 | 0.948 | 0.903 | -0.723 | 0.966 | 0.925 |
| | 12 h | 6.164 | 4.899 | 5.706 | 15.766 | 4.313 | 3.897 | 0.848 | 0.904 | 0.870 | -0.759 | 0.868 | 0.893 |
| 6 mo. | 3 h | 2.567 | 2.315 | 3.878 | 4.551 | 2.168 | 3.153 | 0.959 | 0.966 | 0.905 | 0.731 | 0.939 | 0.875 |
| | 6 h | 3.842 | 3.267 | 5.290 | 6.772 | 2.185 | 3.344 | 0.907 | 0.933 | 0.824 | 0.406 | 0.935 | 0.858 |
| | 9 h | 4.418 | 4.234 | 3.690 | 7.472 | 3.159 | 3.388 | 0.877 | 0.887 | 0.914 | 0.277 | 0.871 | 0.853 |
| | 12 h | 6.562 | 5.207 | 4.769 | 11.230 | 3.182 | 4.298 | 0.730 | 0.830 | 0.857 | -0.632 | 0.869 | 0.761 |
| 3 yr | 3 h | 3.352 | 5.664 | 12.171 | 9.298 | 10.456 | 12.647 | 0.978 | 0.876 | 0.599 | 0.978 | 0.820 | 0.621 |
| | 6 h | 6.237 | 6.990 | 14.976 | 14.958 | 10.783 | 13.084 | 0.923 | 0.862 | 0.559 | 0.489 | 0.793 | 0.609 |
| | 9 h | 8.689 | 9.813 | 18.329 | 21.367 | 11.539 | 18.736 | 0.851 | 0.811 | 0.190 | -0.043 | 0.696 | 0.182 |
| | 12 h | 10.727 | 11.637 | 21.504 | 24.342 | 11.734 | 19.930 | 0.774 | 0.734 | 0.091 | -0.353 | 0.686 | 0.093 |

## 6 Conclusion

Results obtained from this study show good performance for the dynamic types of the neural networks (Recurrent and Time lag recurrent). The Multi-layer perceptron network did not show good performance particularly for longer prediction horizons. It cannot cope well with hydrological time series that usually contain local features that do not have a fixed position in time. For simulations using a long period of input data the Time lag recurrent network presented poor results although its outputs for simulations with shorter length of record were quite good and almost comparable to recurrent network. Using data from different numbers of upstream gauging sites to the model (test 4) did not indicate considerable difference between the results.

## Reference

Bhattacharya B. and D. P. Solomatine, *Application of artificial neural network in stage-discharge relationship*, Proc. 4th International Conference on Hydroinformatics, Iowa city, USA, July 2000.

Dawson C. W., R. Wilby, *An artificial neural network approach to rainfall-runoff modelling,* J. of Hydrological Sciences, 43 (1), February 1998.

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.

Karunanithi N., W. J. Grenney, D. Whitley, and K. Bovee, *Neural networks for flow prediction,* J. Computing in Civil Engineering, Volume 8, Number 2, 201-220, April 1994.

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.

Tokar S. A. and P. A. Johnson, *Rainfall-runoff modeling using artificial neural networks*, J. Hydrologic Engineering, vol.4 No. 3, Jul 1999