Hydrology and Earth System Sciences, 9(4), *313–321* (2005) © EGU

Hydrology & Earth System Sciences

Simulation of flood flow in a river system using artificial neural networks

Rajesh Raj Shrestha[1], Stephan Theobald[2] and Franz Nestmann[2]

[1]Department of Hydrological Modelling, UFZ - Centre for Environmental Research Leipzig-Halle, Brückstrasse 3a, 39114 Magdeburg, Germany
[2]Institute for Water Resources Management, Hydraulic and Rural Engineering, University of Karlsruhe, D-76128 Karlsruhe, Germany

Email for corresponding author: rajesh.shrestha@ufz.de

Abstract

Artificial neural networks (ANNs) provide a quick and flexible means of developing flood flow simulation models. An important criterion for the wider applicability of the ANNs is the ability to generalise the events outside the range of training data sets. With respect to flood flow simulation, the ability to extrapolate beyond the range of calibrated data sets is of crucial importance. This study explores methods for improving generalisation of the ANNs using three different flood events data sets from the Neckar River in Germany. An ANN-based model is formulated to simulate flows at certain locations in the river reach, based on the flows at upstream locations. Network training data sets consist of time series of flows from observation stations. Simulated flows from a one-dimensional hydrodynamic numerical model are integrated for network training and validation, at a river section where no measurements are available. Network structures with different activation functions are considered for improving generalisation. The training algorithm involved backpropagation with the Levenberg-Marquardt approximation. The ability of the trained networks to extrapolate is assessed using flow data beyond the range of the training data sets. The results of this study indicate that the ANN in a suitable configuration can extend forecasting capability to a certain extent beyond the range of calibrated data sets.

Keywords: artificial neural networks, activation function, backpropagation, hydrodynamic numerical model, multilayer perceptron, Neckar River.

Introduction

The ability to simulate river flows quickly and accurately is of crucial importance in flood forecasting operations. Hydrodynamic models provide a sound physical basis for this purpose and have the capability to simulate a wide range of flow situations. However, these models require accurate river geometric data, which may not be available in many locations. It is also not possible to integrate observed data directly at desired locations to improve the model results. In this respect, artificial neural network (ANN) provides a quick and flexible approach for data integration and model development.

The use of ANN-based models for the simulation of flood flows has been gaining popularity in recent years. Several researchers have demonstrated the application for rainfall-runoff modelling and streamflow simulation (Thirumalaiah and Deo, 1998; Dawson and Wilby, 1999; Imrie *et al*., 2000; Solas *et al*., 2000; Dolling and Veras, 2002; Shamseldin *et al.*, 2002; Shrestha, 2003). A recent review can be found in the ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000). Typical applications involve the training of two- to three-layer networks using suitable network architectures like multilayer perceptron, radial basis networks or recurrent networks. The performances of the ANNs in river flow prediction have been found to be comparable with other data driven modelling approaches (Lekkas *et al*., 2001; Sivakumar *et al*., 2002).

An important criterion in application of ANN for flood flow simulation is to predict flows beyond the range of calibrated data sets. The network may perform very well for the training data set, but may be unable to generalise flows beyond the range of training data sets. Minns (1996) applied ANNs to both real and theoretical catchments, and found that the peak flows were considerably underestimated. Thirumalaiah and Deo (1998) used ANNs for river stage

forecasting and found that although lower water levels were predicted fairly accurately, higher water levels were underestimated. Solas *et al.* (2000) used average, dry and wet years' mean annual precipitation in a rainfall-runoff modelling application of ANNs, and observed that high flows were overestimated for the wet years.

The reasons why the ANNs underestimate or overestimate extreme flows may lie in the network structure used and range of training data sets. Thirumalaiah and Deo (1998) suggested that this could be due to a smaller number of training patterns for higher water levels. Minns (1996) emphasised the need to ensure that the training data sets actually contain all conceivable events.

Several methodologies have been proposed to improve the forecasting capability outside the range of training data sets. In the ANN applications, the inputs data are generally normalised in the range such as [0 – 1.0] or [–1.0 – 1.0]. To accommodate the data beyond the training range, alternative normalisation ranges have been suggested. The scaling of training data in a range such as [0.1 – 0.9] or [0.2 – 0.8] compared to the range [0.0 – 1.0] has been reported to be an effective means of improving generalisation (Imrie *et al*., 2000; Dawson *et al.,* 2002). This can accommodate validation and test data sets in excess of training data sets. The upper and the lower limits of activation functions such as [0 – 1.0] or [–1.0 – 1.0] also provide limiting amplitude to the data sets and affect the generalisation capability of the ANNs. Imrie *et al*. (2000) investigated the effects of different output activation functions for improving generalisation using cascade correlation network building strategy. Shamseldin *et al.* (2002) examined the significance of different non-linear activation functions for the hidden and the output layers in the context of overall performance of the multilayer feedforward networks. An alternative approach is the application of a non-linear activation function at the hidden and the linear function at the output layers. The application of linear activation in the output layers enables the network to take any range of values (Demuth and Beale, 2000).

This study explores methods for improving generalisation with different activation functions at the hidden layers using multilayer feedforward networks. A case study from the Neckar River in Germany demonstrates the application using historical flood data sets.

Artificial Neural Networks

Artificial neural networks (ANNs) are massively parallel distributed processors with a natural propensity for storing knowledge and making it available for use (Haykin, 1994). Typical neural networks consist of layers of neurons with

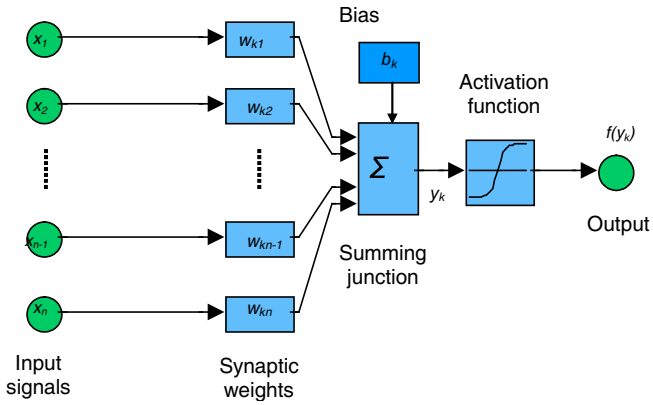

Fig. 1. *Structure of a neuron of an Artificial Neural Network*

weights providing inter-neuron connection for storage of knowledge. A summing junction acts as an adder for summing input signals weighted by respective weights. An activation function provides limiting amplitude of the neuron output, typically between [0.0 – 1.0] or [–1.0 – 1.0]. The structure of a typical neuron of an ANN is shown in Fig. 1.

The ANNs may also consist of a number of hidden layers between the input and output layers. Such networks are commonly known as multilayer feedforward networks or multilayer perceptrons (MLPs). The MLPs are one of the most widely used types of neural networks, which can be trained in a supervised manner to solve highly nonlinear problems. The capability of the MLPs in modelling dynamic systems can be enhanced using time delays. These provide a sequence of input vectors that occur in a certain time order. For the ANNs to predict a wide range of flow situations, it is important that the networks are able to generalise different ranges of data sets. A network with too few neurons may not approximate different flow situations. A network that is too complex may fit the noise, not just the signal, causing over-fitting. This may lead to a network that performs very well for the training data set, but becomes unable to generalise to a new situation.

It is customary to use validation and test data sets to assess the capability of the trained ANNs. Validation sets are separate sets of data, to be used during the training process to monitor the generalisation capability. Normally, the errors of the validation data sets decrease during the initial training iterations but begin to rise when the network begins to over-fit the training data. When the validation errors increase for the number of iterations in a criterion specified by 'early stopping', the training process is stopped. The test sets are independent sets of data, not used in training or validation, and are to be used to evaluate the network performance.

ACTIVATION FUNCTIONS

The activation function transforms signals y_k at a neuron using some function $f(y_k)$. Different activation functions investigated in this study are given below.

i. Linear function calculates the neuron's output by simply summing all the input y_k and bias b passed to it. This may be modified by a factor a to provide different limiting amplitudes. This function is defined as

$$f(y_k) = ay_k \text{ for all } y_k \quad (1)$$

ii. Sigmoidal function produces an output in the range 0 to +1. This function has the following form

$$f(y_k) = \frac{1}{1+\exp(-y_k)} \quad (2)$$

iii. Hyperbolic tangent function is mathematically equivalent to $tanh(y_k)$ and produces an output in the range -1 to +1. This function is given by

$$f(y_k) = \frac{2}{1+\exp(-2y_k)} - 1 \quad (3)$$

iv. Hyperbolic tangent + linear function can be used to combine non-linearity of the hyperbolic tangent function with the linear function using the weighing factor ±. This function is of the form

$$f(y_k) = \left(\frac{2}{1+\exp(-2y_k)} - 1\right)\alpha + (1-\alpha)y_k \quad (4)$$

Figure 2 shows these four different activation functions for the data range –2 to +2. From the figures it can be seen that the sigmoidal function provides the lowest limiting amplitude and the linear function the highest limiting amplitude. The limited amplitude range provided by the functional output of the sigmoidal and hyperbolic tangent functions produces a 'squashing effect' to the input signals. The sigmoidal and hyperbolic tangent functions are most commonly used only at the hidden layers. If the output layers use sigmoidal or hyperbolic tangent functions, the outputs are restricted to a small range of values. The application of the linear function at the output layer makes it possible for the network to take any value.

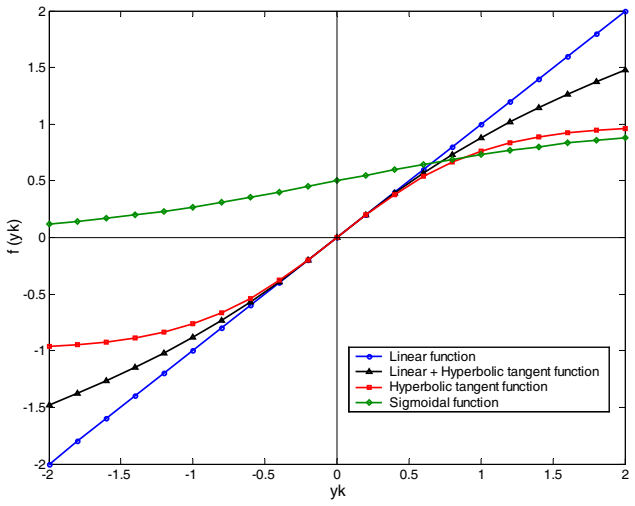

Fig. 2. Activation functions

Study area and data

The Neckar is a major tributary of the Rhine and flows through the region of Stuttgart, Heidelberg and Mannheim in south-west Germany. The study area consists of a reach of about 100 km from Lauffen to Heidelberg (Fig. 3) with a catchment area of 13 787 km^2 at the Heidelberg station. Time series of flow and water level data at one hour intervals are available from the gauging stations located at Lauffen, Rockenau and Heidelberg and water level time series are available from Gundelsheim station for the 1988, 1990 and 1993 flood events. Flow time series from the major tributaries Kocher, Jagst and Elz and the smaller tributaries Schwarzbach, Elsenz and Itter are also available for the same years. There is also a functioning one-dimensional hydrodynamic numeric (HN) model of the study area, with cross-sections at 100 m interval. Only the simulated flows from the HN model at Gundelsheim station, where no flow

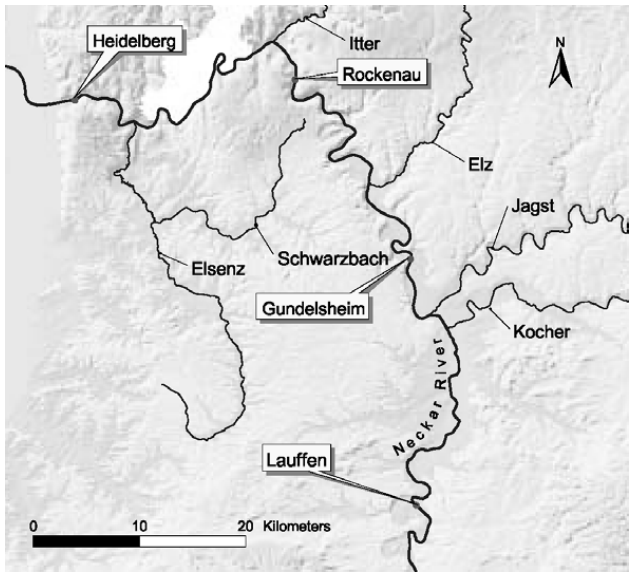

Fig. 3. Study Area

Table 1. Statistical characteristics of the flow data at Rockenau and Heidelberg gauging stations

Data set	Meanflow ($m^3 s^{-1}$)	Maximum flow ($m^3 s^{-1}$)	Minimum flow ($m^3 s^{-1}$)	Standard deviation
ROCKENAU				
Training set (1988)	958	1930	341	399
Validation set (1990)	583	2225	144	576
Test set (1993)	713	2680	232	579
HEIDELBERG				
Training set (1988)	1060	1945	374	401
Validation set (1990)	666	2299	224	604
Test set (1993)	807	2706	275	597

data are available, are included for model training and validation. The available flow data were divided in such a way that the maximum flows of validation and test data sets exceed the range of training data sets. The flow time series from the 1988 flood event was used as training set, and flood event data from 1990 and 1993 were used as validation and test data sets. The statistical characteristics of the flow data at the prediction stations Rockenau and Heidelberg are summarised in Table 1.

A cross correlation analysis was performed on the time series flows to identify a suitable lag time from upstream to downstream points. The cross correlation analyses of the time series water level data from Lauffen, Gundelsheim and Rockenau stations yielded the suitable lag time for Gundelsheim and Rockenau with respect to the upstream stations. Similarly, the analysis between the flow data from Rockenau and Heidelberg gave the lag time for the Heidelberg station. The lag times for the tributaries' inflows were calculated based on their distances. The lag times with respect to forecast stations Gundelsheim, Rockenau and Heidelberg for the upstream stations in the Neckar river and the tributaries are given in Table 2.

Table 2. Lag time with respect to forecast stations

Forecast station	*Upstream stations*	*Lag time (hrs)*
Gundelsheim (Neckar)	Lauffen (Neckar)	4
	Stein (Kocher)	3
	Untergriesheim (Jagst)	2
Rockenau (Neckar)	Gundelsheim (Neckar)	3
	Mosbach (Elz)	2
Heidelberg (Neckar)	Rockenau (Neckar)	2
	Eschelbronn (Schwarzbach)	2
	Eberbach (Itter)	2
	Meckesheim (Elsenz)	2

The forecast horizons at Gundelsheim, Rockenau and Heidelberg with respect to the upstream flows are each 2 hours. However, since the contribution from the tributaries Elz, Schwarzbach, Elsenz and Itter are quite small compared to flows in the rivers, their flows can be assumed to be constant for the duration of forecast to increase the forecast horizon. This increases the forecast horizon at Rockenau station to 5 hours and Heidelberg station to 7 hours. The forecast horizon can be increased further by integrating with external models such as the rainfall-runoff models.

River flow prediction

The selection of appropriate input and output data sets is an important consideration in the ANN modelling. A number of experiments was performed with the division of river reach into different ANN blocks. In the first experiment the ANN was used to predict flows at the gauging station Rockenau, based on the upstream flows from Lauffen (Neckar), and the tributaries Kocher, Jagst and Elz. However, although the network functioned quite well for the training and validation data sets, it did not perform well for the test data sets.

As a functioning HN model is available for the study area, it was decided to integrate the HN simulated results from Gundelsheim for the ANN training. It was observed that the integration of HN model results from Gundelsheim improved the performance of river flow prediction at Rockenau. Accordingly, the river reach was divided into three 'sub-reaches' represented by independently trained ANN blocks. The observed flows at Lauffen, Rockenau and Heidelberg stations, together with the tributaries and HN model results from Gundelsheim, were integrated for the ANN training. The inflows and the desired outflows for each of the ANN blocks are summarised in Table 3. Outputs of the best performing networks from previous blocks were used as inputs to next blocks.

Table 3. Network inputs and desired outputs

Network block	*River sub-reach*	*Network input*	*Desired output*
ANN block 1	Lauffen – Gundelsheim	Measured flows from gauging stations at Lauffen, and tributaries Jagst and Kocher	Simulated flows from the HN model at Gundelsheim
ANN block 2	Gundelsheim – Rockenau	Simulated flows from the ANN block 1 at Gundelsheim, and measured flow from gauging station at the tributary Enz	Measured flows from gauging stations at Rockenau
ANN block 3	Rockenau – Heidelberg	Simulated flows from the ANN block 2 at Rockenau, and measured flows from gauging stations at the tributaries Schwarzbach, Elsenz, and Itter	Measured flows from gauging stations at Heidelberg

The input data sets for the ANN trainings were scaled in the range [0.2 – 0.8]. Each of the ANN blocks was trained with one input layer, two hidden layers and one output layer. The number of neurons in the networks was kept to a minimum of eight in the first hidden layer, four in the second hidden layer and one in the output layer. There was no significant improvement of the model performance with the increase in number of neurons. Each of the networks consisted of a linear activation function in the output layer. The networks were trained with four different activation functions in the first hidden layer, namely linear, sigmoidal, hyperbolic tangent and a hyperbolic tangent + linear function. The weighing factor ± for the hyperbolic tangent + linear function was varied between 0.4 and 0.8 during training process.

The networks were trained using the procedures from MATLAB neural network toolbox. This involved network designing using text files containing MATLAB code (M-files). The training is done using a backpropagation algorithm with Bayesian regularisation of the Levenberg-Marquardt approximation. The early stopping criteria provided by the validation data sets were used to prevent overtraining. The test data sets were used independently for the evaluation of the model performance. The best performing network blocks in each of the sub-reaches (with the hyperbolic tangent and linear function in the first hidden layer for the first and third blocks and hyperbolic tangent function in the first and second hidden layers for the second block) were combined in the Simulink environment and an ANN simulation model was formulated. The combined ANN simulation model is schematised in Fig. 4.

Error Measurement

Table 4 shows the statistics of the error comparison used in this study. The error measurement process consists of analysis of errors between observed and calculated values. The overall performance of trained networks can be judged with respect to criteria such as the coefficient of efficiency (CE) and coefficient of determination (R^2). These coefficients are independent of the scale of data used and are useful in assessing the goodness of fit of the model (Dawson *et al.,* 2002; Dawson and Wilby, 1999). The root mean square error (RMSE) evaluates the error independent of sample size and can give useful insights into amplitude errors. The difference in peak flow between observed and calculated flows was considered to assess the prediction capability of the trained networks beyond the calibrated range.

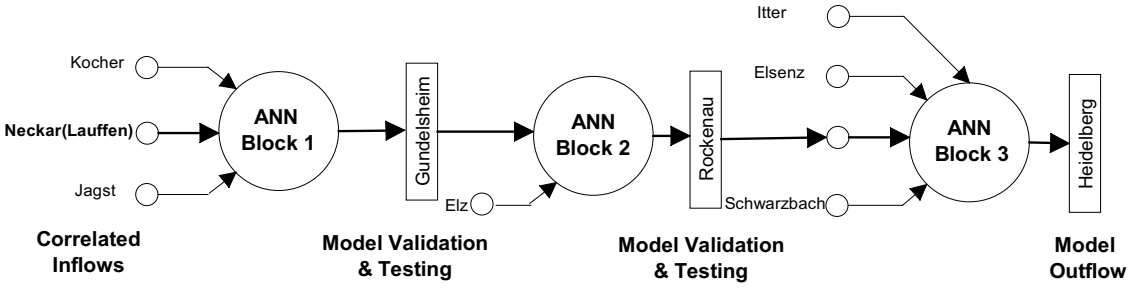

Fig. 4. *ANN simulation model*

Table 4. Error measurement formula

Error measurement	*Name*	*Formula*
Coefficient of efficiency	*CE*	$1-\frac{\sum_{i=1}^{n}(Q_{obs}-Q_{cal})^2}{\sum_{i=1}^{n}(Q_{obs}-Q_{av})^2}$
Coefficient of determination	R^2	$\frac{\left[\sum_{i=1}^{n}(Q_{obs}-Q_{av})(Q_{cal}-Q_{cal\_av})\right]^2}{\left[\sum_{i=1}^{n}(Q_{obs}-Q_{av})\right]\left[\sum_{i=1}^{n}(Q_{cal}-Q_{cal\_av})\right]}$
Root mean square errors	*RMSE*	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{obs}-Q_{cal})^2}$
Difference in peak flow	*DPF*	$Q_{obs}(max)-Q_{cal}(max)$

Where n is the number of observations, Q_{obs} and Q_{cal} are the observed and calculated values respectively, and Q_{cal_av} are the mean of the observed and calculated values. $Q_{obs}(max)$ and $Q_{cal}(max)$ are the maximum of observed and calculated values.

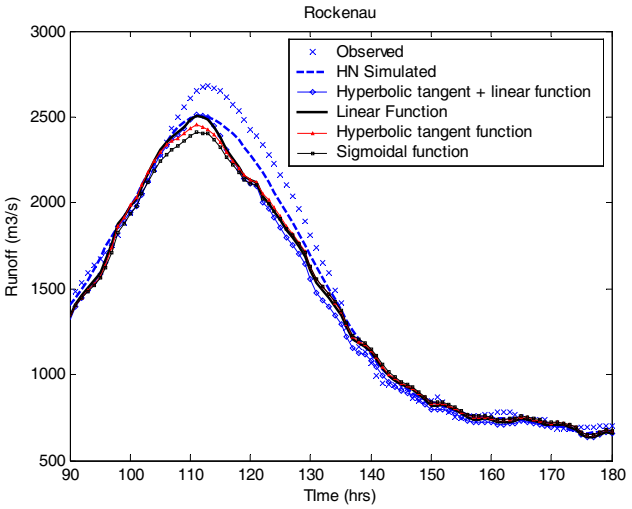

Fig. 5. *Comparison of model performance of different activation functions at Rockenau using upstream flows from Lauffen and tributaries*

Results

PERFORMANCE OF THE ROCKENAU ANN MODELS

The performance of the ANN models for the test data sets (1993 flood event) for the prediction of flow at Rockenau station are summarised in Table 5. The ANNs were trained with the upstream flows from Lauffen (Neckar) and the tributaries Kocher, Jagst and Elz. All the ANN models using different activation functions underestimated the flow at Rockenau station for the test data sets. The ANN with the activation function with a lower limiting amplitude such as a sigmoidal and hyperbolic tangent function in the hidden layer had a higher level of underprediction. Figure 5 shows the partial results of the model corresponding to the 1993 flood event, for the period 21.12.1993 to 25.12.1993.

PERFORMANCE OF THE SUB-REACH MODELS

The performances of different activation functions for the test data sets are summarised in Tables 6, 7 and 8. In the ANN block 1 between Lauffen and Gundelsheim, the network could easily approximate the training data sets. For training and test data sets, there was a tendency to decreasing error when activation functions with higher limiting amplitude were used. The application of the sigmoidal and hyperbolic tangent functions underestimated the peak flows whereas the linear function overestimated the peaks. The application of the combination of hyperbolic tangent and linear function ($\alpha = 0.5$) at the first hidden layer gave the best performance in terms of CE, R^2 and RMSE. The difference in peak flow was also found to be least using this activation function.

In the ANN block 2 between Gundelsheim and Rockenau, the performance of the networks was similar to the ANN block 1. The application of activation functions with higher limiting amplitude produced better results. However, there was an overall trend of underestimation of peaks, even with the linear activation function. Only with the application of hyperbolic tangent activation functions at the first and second hidden layers was the network able to predict peak flows with reasonable accuracy.

Table 5. Network performance in the reach between Lauffen and Rockenau

Activation functions	*Coefficient of efficiency CE*	*Coefficient of determination* R^2	*Root mean square error RMSE* ($m^3 s^{-1}$)	*Difference in peak flow* ($m^3 s^{-1}$)
Sigmoidal	0.9498	0.9771	141	273
Hyperbolic tangent	0.9565	0.9777	132	230
Hyperbolic tangent + linear	0.9534	0.9771	142	176
Linear	0.9535	0.9810	137	169

Table 6. Network performance in the reach between Lauffen and Gundelsheim

Activation functions	*Coefficient of efficiency CE*	*Coefficient of determination R^2*	*Root mean square error RMSE ($m^3 s^{-1}$)*	*Difference in peak flow ($m^3 s^{-1}$)*
Sigmoidal	0.9847	0.9907	72	118
Hyperbolic tangent	0.9887	0.9922	62	54
Hyperbolic tangent + linear	0.9918	0.9937	56	-36
Linear	0.9895	0.9921	66	-176

Table 7. Network performance in the reach between Gundelsheim and Rockenau

Activation functions	*Coefficient of efficiency CE*	*Coefficient of determination R^2*	*Root mean square error) RMSE ($m^3 s^{-1}$)*	*Difference in peak flow ($m^3 s^{-1}$)*
Sigmoidal	0.9698	0.9804	116	165
Hyperbolic tangent	0.9634	0.9820	121	246
Hyperbolic tangent + linear	0.9639	0.9801	119	205
Linear	0.9650	0.9787	122	154
Hyperbolic tangent in first & second hidden layers	0.9806	0.9877	95	-47

Table 8. Network performance in the reach between Rockenau and Heidelberg

Activation functions	*Coefficient of efficiency CE*	*Coefficient of determination R^2*	*Root mean square error RMSE ($m^3 s^{-1}$)*	*Difference in peak flow ($m^3 s^{-1}$)*
Sigmoidal	0.9563	0.9671	136	215
Hyperbolic tangent	0.9566	0.9817	135	163
Hyperbolic tangent + linear	0.9581	0.9752	142	62
Linear	0.9649	0.9790	109	-126

The statistical performances of the model from Gundelsheim to Rockenau are compared with the Rockenau ANN model, which do not integrate the HN model results from Gundelsheim (Table 5 and 7). The statistical performance in terms of CE, R^2, RMSE and difference in peak flow of the Gundelsheim to Rockenau ANN models using all activation functions are found to be superior.

The performance ANN block 3 between Rockenau and Heidelberg also improved with the application of activation functions with higher limiting amplitude. There was also a general trend of underestimation of peaks. The difference in peak flow prediction was also found to be lower using activation functions with higher limiting amplitude. The hyperbolic tangent + linear function ($\alpha = 0.7$) gave the best performance in terms of *CE*, and difference in peak flow.

Partial results of the model showing peak discharges for the test data sets in the period between 21.12.1993 and 25.12.1993 are shown in Figs. 6, 7 and 8. In Fig. 6, the

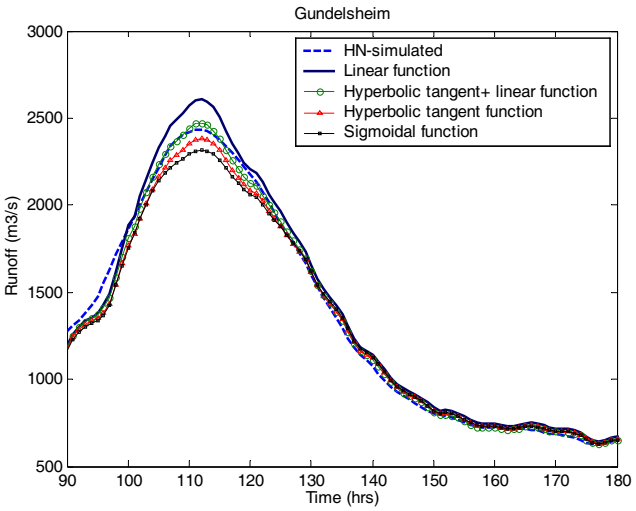

Fig. 6. *Comparison of model performance of different activation functions at Gundelsheim (Block 1)*

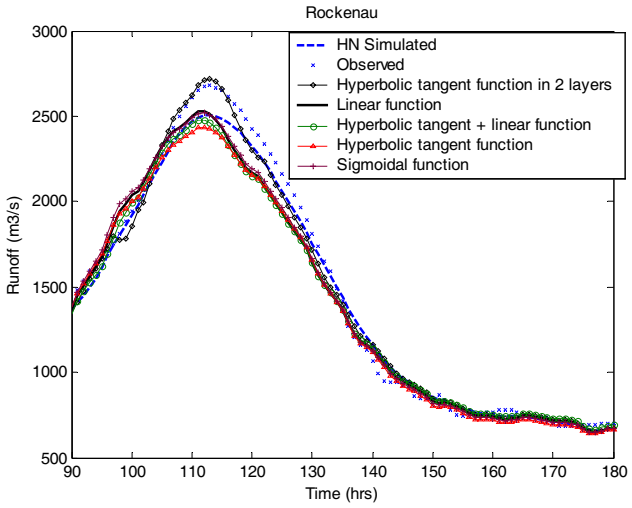

Fig. 7. *Comparison of model performance of different activation functions at Rockenau (Block 2)*

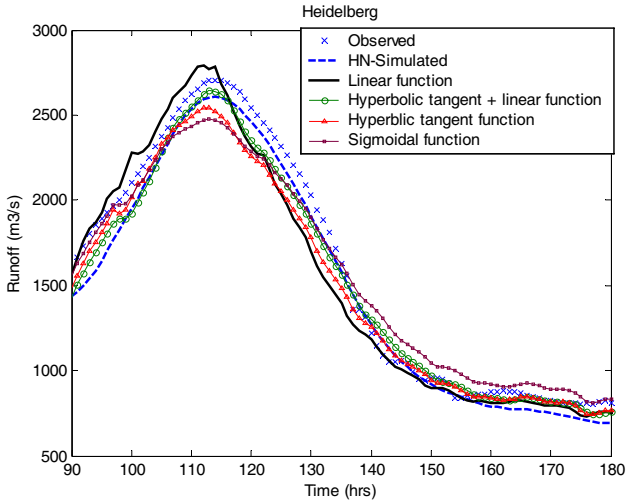

Fig. 8. *Comparison of model performance of different activation functions at Heidelberg (Block 3)*

ANNs results are compared with the simulated flows from hydrodynamic numerical (HN) model. Figures 7 and 8 show the comparison with both HN model results and measured flow series.

It is to be noted that the statistical performance of the ANN model for block 1 is not consistent with blocks 2 and 3 as the results of block 1 are compared with the HN model simulations and blocks 2 and 3 with the measured data.

ANN MODELS FOR EXTREME FLOWS

To test the ability of ANNs to predict extreme events, simulations were made with all the upstream inflows multiplied by a factor 1.5. The ANN simulation model containing the best performing ANN blocks in the previous section was used. Simulations using the HN models were also made by multiplying the upstream flows by 1.5. The outputs of the HN model and ANN simulated results were compared with each other. Figures 8, 9 and 10 show the comparison of the results at Gundelsheim, Rockenau and Heidelberg respectively with 1993 flows multiplied by 1.5. The comparison of results of the two models indicated a good match for the Gundelsheim station and an under- or overestimation for the Rockenau and Heidelberg stations. The ANN outputs showed underestimation at Gundelsheim (ANN block 1) and Heidelberg (ANN block 3), and

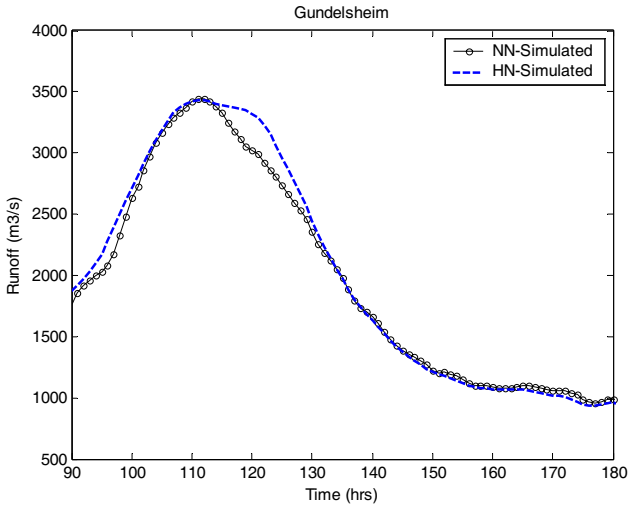

Fig. 9. *Comparison of model performance at Gundelsheim using 1993 flows multiplied by 1.5 (Hyperbolic tangent + linear activation function)*

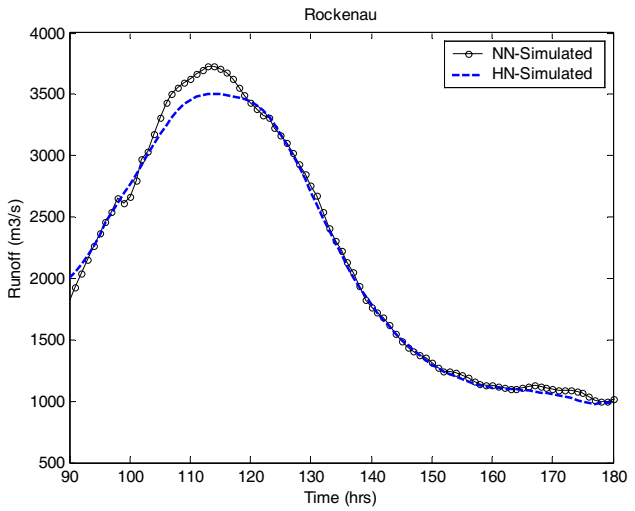

Fig. 10. *Comparison of model performance at Rockenau using 1993 flows multiplied by 1.5 (Hyperbolic tangent activation function in two hidden layers)*

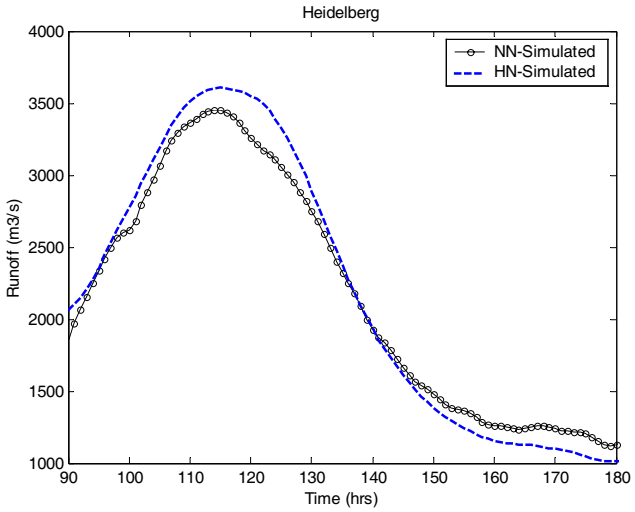

Fig. 11. Comparison of model performance at Heidelberg using 1993 flows multiplied by 1.5 (Hyperbolic tangent + linear activation function)

overestimation at Rockenau (ANN block 2). It is interesting to note here that ANN blocks 1 and 3 have a combination of linear and hyperbolic transfer functions at the first hidden layers and ANN block 2 has hyperbolic transfer functions at the first and second hidden layers.

The result highlighted a random behaviour of the ANNs when used to predict extreme flow events. It might function quite well in some cases and not so well in other cases. Hence caution needs to be exercised in the use of ANNs for forecasting extreme events. It is also important to specify the forecasting range of the trained ANNs.

Conclusions

This paper has presented an ANN-based approach for the simulation of flood flows in the Neckar River in Germany. The river reach was divided into three 'sub-reaches' for the ANN trainings. The effects of different activation functions at the first hidden layer of multilayer perceptron (MLP) neural networks were evaluated for predicting flows beyond a calibrated range. This evaluation was made in terms of test data sets with higher peaks, above the range of training data sets. Four different activation functions, the sigmoidal, hyperbolic tangent, linear, and a combination of hyperbolic tangent and linear functions were investigated in this study.

The results of this study indicate that the ANNs provide an efficient means of flood flow forecasting. Compared to a HN model, which requires a lot of cross-sectional data, the ANN model can be quickly trained to forecast flows at specific sections in the river reach. The division of river reach into 'sub-reaches' facilitates the integration of data from observations and numerical model results. This also provides guidance to the network training and enhances the overall model performance. The assessment of results of the trained networks shows that a combination of hyperbolic tangent and linear transfer functions at the first hidden layer generally produced the best performance. This function has higher limiting amplitude and also imparts non-linearity to the networks. Hence, the ANNs in a suitable configuration can extend the forecasting capability to a certain extent beyond the range of calibrated data sets. It is, however, important to exercise caution in using ANNs for extreme flood events.

References

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000. Artificial neural networks in hydrology II: Hydrologic applications. *J. Hydrolog. Eng.-ASCE*, **5**, 124–137.

Dawson, C.W. and Wilby, R.B., 1999. A comparison of artificial neural networks for flow forecasting. *Hydrol. Earth Syst. Sci*., **3**, 529–540.

Dawson, C.W., Harpham C., Wilby, R.B. and Chen, Y., 2002. Evaluation of artificial neural network techniques for flow forecasting in River Yangtze, China. *Hydrol. Earth Syst. Sci*., **6**, 619–626.

Demuth, H. and Beale M., 2000. *Neural network toolbox user's guide*. The MathWorks Inc. Online documentation: http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/

Dolling, O.R. and Veras, E.A., 2002. Artificial neural networks for streamflow prediction. *J. Hydraul. Res.,* **40**, 547–554.

Haykin, S., 1994. *Neural networks a comprehensive foundation* (1st Edition). Macmillan College Publishing Company, Inc., New York.

Imrie, C. E., Durucan, S. and Korre, A., 2000. River flow prediction using artificial neural networks: generalisation beyond calibration range. *J. Hydrol.*, **233**, 138–153.

Lekkas, D.F., Imrie, C.E. and Lees M.J., 2001. Improved nonlinear transfer function and neural network methods for flow routing for real-time flood forecasting. *J. Hydroinformatics*, **3**, 153–164.

Minns, A.W., 1996. *Extended rainfall-runoff modelling using artificial neural networks*, A. Müller (Ed.), Proc. Hydroinformatics '96, Balkema, Rotterdam, the Netherlands. 207–213.

Shamseldin, A.Y., Nasr, A.E. and O'Connor, K.M., 2002. Comparison of different forms of the multi-layer feed-forward neural network method used for river flow forecasting. *Hydrol. Earth Syst. Sci*., **6**, 671–684.

Shrestha, R.R., 2003. *Flood routing using artificial neural networks*. Proc. XXX IAHR Congress, JF Kennedy Student Paper Competition, Thessaloniki, Greece.

Sivakumar, B., Jayawardena, A.W. and Fernando, T.M.K.G., 2002. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *J. Hydrol.*, **265**, 225–245.

Solas, J.D., Markus, M. and Tokar, A.S., 2000. Streamflow forecasting based on artificial neural networks. In: *Artificial neural networks in hydrology*, R.S. Govindaraju and A. Ramachandra Rao (Eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands. 23–51.

Thirumalaiah, K. and Deo, M.C., 1998. River stage forecasting using artificial neural networks. *J. Hydrolog. Eng.-ASCE*, **3**, 26–31.