SCIENTIFIC BRIEFING
HYDROLOGICAL PROCESSES
Hydrol. Process. **17**, 677–686 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.5103

Improving peak flow estimates in artificial neural network river flow models

K. P. Sudheer*
P. C. Nayak
and K. S. Ramasastri

National Institute of Hydrology, Deltaic Regional Centre, Kakinada (KPS, PCN) and National Institute of Hydrology, Roorkee (KSR), India

*Correspondence to:
K. P. Sudheer, National Institute of Hydrology, Deltaic Regional Centre, Kakinada (KPS, PCN) and National Institute of Hydrology, Roorkee (KSR), India. E-mail: sudheer_nih@rediffmail.com

Abstract

In this paper, the concern of accuracy in peak estimation by the artificial neural network (ANN) river flow models is discussed and a suitable statistical procedure to get better estimates from these models is presented. The possible cause for underestimation of peak flow values has been attributed to the local variations in the function being mapped due to varying skewness in the data series, and theoretical considerations of the network functioning confirm this. It is envisaged that an appropriate data transformation will reduce the local variations in the function being mapped, and thus any ANN model built on the transformed series should perform better. This heuristic is illustrated and confirmed by many case studies and the results suggest that the model performance is significantly improved by data transformation. The model built on transformed data outperforms the model built on raw data in terms of various statistical performance indices. The peak estimates are improved significantly by data transformation. Copyright © 2003 John Wiley & Sons, Ltd.

Introduction

The application of artificial neural networks (ANNs) to various aspects of hydrological modelling has undergone much investigation in recent years. This interest has been motivated by the complex nature of hydrological systems and the ability of ANNs to model non-linear relationships. ANNs are essentially semi-parametric regression estimators and well suited for hydrological modelling, as they can approximate virtually any (measurable) function up to an arbitrary degree of accuracy (Hornik *et al*., 1989). A significant advantage of the ANN approach in system modelling is that one need not have a well-defined process for algorithmically converting an input to an output. Rather, all that is needed for most networks is a collection of representative examples of the desired mapping. The ANN then adapts itself to reproduce the desired output when presented with training sample input. The emergence of neural network technology has provided many promising results in the field of hydrology and water resources simulation. A comprehensive review of the application of ANN to hydrology can be found in ASCE Task Committee (2000a,b) and in Maier and Dandy (2000).

Despite the huge amount of network theory and the importance of neural networks in applied work, many researchers have indicated that ANN models are unable to predict extreme values in the river flow

Received 24 September 2002
Accepted 26 November 2002

Copyright © 2003 John Wiley & Sons, Ltd.

(Minns and Hall, 1996; Dawson and Wilby, 1998; Compolo *et al*., 1999). Imrie *et al*. (2000) argue that there may be a number of reasons why ANN models are unable to predict extreme values, and a variety of remedies have been proposed (e.g. Karunanithi *et al*., 1994; Hsu *et al*., 1995; See *et al*., 1997). Imrie *et al*. (2000) attribute these extreme values to the climatic variations observed over the past few years that have given rise to record-breaking flood and drought conditions. This paper addresses this concern of extreme value prediction in ANN-based river flow modelling through discussion of the possible causes, and proposes a statistical approach to effectively eliminate this trepidation. The proposed methodology has been illustrated by presenting applications of the procedure to many Indian river basins.

Artificial Neural Networks

An ANN attempts to mimic, in a very simplified way, the human mental and neural structure and functions (Hsieh, 1993). It can be characterized as massively parallel interconnections of simple neurons that function as a collective system. The network topology consists of a set of nodes (neurons) connected by links and usually organized in a number of layers. Each node in a layer receives and processes weighted input from the previous layer and transmits its output to nodes in the following layer through links. Each link is assigned a weight, which is a numerical estimate of the connection strength. The weighted summation of inputs to a node is converted to an output according to a transfer function (typically a sigmoid function). Most ANNs have three layers or more: an input layer, which is used to present data to the network; an output layer, which is used to produce an appropriate response to the given input; and one or more intermediate layers, which are used to act as a collection of feature detectors.

There are no fixed rules for developing an ANN, even though a general framework can be followed based on previous successful applications in engineering. The goal of an ANN is to generalize a relationship of the form:

$$Y^m = f(X^n) \tag{1}$$

where X^n is an n-dimensional input vector consisting of variables $x_1, \ldots, x_i, \ldots, x_n$; Y^m is an m-dimensional output vector consisting of the resulting variables of interest $y_1, \ldots, y_i, \ldots, y_m$. In the river flow modelling, values of x_i may include runoff and any other exogenous variables at various time lags, and the value of y_i is generally the flow during subsequent periods.

The ability of neural networks to process information is obtained through a learning process, which is the adaptation of link weights so that the network can produce an approximate output. In general, the learning process of an ANN will reward a correct response of the system to input by increasing the strength of the current matrix of nodal weights. Therefore, the likelihood of producing similar output when the same inputs are entered in the future will increase. An incorrect response from the system is discouraged by adjusting the nodal weights so that the system will respond differently when it encounters similar inputs in the future (Hsieh, 1993). The learning process, in general, is a non-linear optimization of an error function. During learning, iterative techniques are used to minimize the error function and the iteration is terminated when the error function value reaches a predefined goal, thus completing the learning process. The working of a three-layer ANN can be described mathematically as follows.

Let the weight vector between the input and hidden layer of the ANN be denoted by $\boldsymbol{w}_h$, then the output of the ith node (hidden) in the ANN is given by:

$$y_{hi} = \sum_{i=1}^{n} x_i w_{hi} + b_{hi} = \mathbf{x}^{\mathrm{T}}\mathbf{w}_h + \mathbf{b}_h \tag{2}$$

and the network output is given by:

$$y_i = f(\mathbf{x}^{\mathrm{T}}\mathbf{w} + \mathbf{b}_h) + \mathbf{b}_o \tag{3}$$

where $\boldsymbol{b}_h$, $\boldsymbol{b}_o$ are the bias vectors of the hidden and output layer, respectively, and f is the transfer function. Determination of an appropriate network architecture is one of the most important, but also one of the most difficult, tasks in the model building process. Unless carefully designed an ANN model can lead to over-parameterization, resulting in an unnecessarily large network.

Extreme Value Prediction: Issues and Remedies

The accuracy of the flood peaks predicted by ANN models is one of the major concerns for the hydrologic community. Although no study has discussed

Copyright © 2003 John Wiley & Sons, Ltd.

the possible causes of this problem, it could be illustrated by theoretical considerations of the non-linear transformation that occurs in the hidden layer nodes of the ANN. Consider the commonly used sigmoid transfer function, which is defined for any variable u as:

$$f(u) = \frac{1}{1 + \exp(-u)} \quad (4)$$

where $u = \mathbf{x}^{\mathrm{T}}\mathbf{w}$, in which $\mathbf{x}$ is the input vector and $\mathbf{w}$ is the weight vector to the hidden node considered. This function is bounded and monotonically increasing, tending to 0 as $\mathbf{x}^{\mathrm{T}}\mathbf{w} \to -\infty$, and approaches 1 as the linearly combined input tends to $+\infty$. For an n-dimensional input space, the output of a sigmoid in the hidden layer is constant (c) along the ($n-1$)-dimensional hyperplanes given by $\mathbf{w}_0 + \mathbf{x}^{\mathrm{T}}\mathbf{w} = c$, where $\mathbf{w}_0$ is the bias for the node (Brown and Harris, 1994). Thus the nodes, which are composed of an adaptive linear combiner and a sigmoidal-type transfer function, are termed ridge functions (Mason and Parks, 1992), as the output is constant along hyperplanes in their input space. Hence, if the desired function can be concisely decomposed into similar ridge functions, multi-layer perceptrons (MLPs) may be suitable models. In contrast, if the modelling functions have local variations, the output of all the hidden layer nodes is generally non-zero, and the resulting optimization problem (weights optimization) can be very complex, and an optimal solution is not guaranteed. This observation illustrates that the MLPs, which are commonly employed in hydrological modelling, are generally unsuitable for modelling functions that have local variations. In river flow modelling, the extreme values contribute to the local variations in the time series function, and consequently the MLPs developed for the flow series fail to capture these extreme values.

Hence it is apparent that for MLPs to provide an accurate estimate of the peak flows, it is necessary to remove the local variations caused by these high flows from the time series function being mapped. Data transformations are often used to simplify the structure of the data so that they follow a convenient statistical model (Carrol and Ruppert, 1988). The objective of transformation may be to induce a simple systematic relationship between response and predictor variables in regression; to stabilize a variance, that is, to induce a constant variance in a group of populations or in the residuals after the regression analysis; and/or to induce a particular type of distribution or to remove extreme skewness. The extreme values in the time series data can be considered as outliers that are poorly fit by the model, though the model fits the bulk of the data. When a transformation is induced, outliers are difficult to identify. For instance, if the original data are lognormally distributed, apparent outliers in the long right tail may be seen to be conforming after the normalizing transformation. Conversely, the smallest observations may be quite close to the bulk of the original data but outlying after the log transformation, implying the transformed series is a smooth function without local variations or outliers.

Until recently, there has been little interaction between the neural network and statistical communities, and ANN and statistical models have developed virtually independently. However, many studies dealing with various engineering applications indicate that ANN models are not significantly different from a number of statistical models. For instance, Hill *et al.* (1994) have suggested that certain ANN models are equivalent to time series models of the autoregressive moving average (ARMA) type. Connor *et al.* (1994) have shown that feed forward neural networks are a special case of non-linear autoregressive (NAR) models. Chon and Cohen (1997) have demonstrated the equivalence of ARMA and non-linear ARMA models with feed forward ANNs utilizing polynomial transfer functions. Maier and Dandy (2000) present equivalent statistical models for neural network models. Despite the similarities, the rules governing the traditional statistical models are seldom considered in ANN model building. However, some studies indicate that consideration of statistical principles in the ANN model building process may improve model performance (e.g. Cheng and Titterington, 1994; Ripley, 1994; Sarle, 1994). Consequently, it is vital to adopt a systematic approach in the development of ANN models, taking into account factors such as data pre-processing, the determination of adequate model inputs and a suitable architecture, parameter estimation and model validation. Maier and Dandy (1997) and Sudheer *et al.* (2002) have discussed guidelines for deciding appropriate inputs to ANN models based on the statistical properties of data.

In most traditional statistical models, the data have to be normally distributed before the model

Copyright © 2003 John Wiley & Sons, Ltd.

coefficients can be estimated efficiently. If the data is not in the normal domain, suitable transformations to normality have to be found. Burke and Ignizio (1992) argue that the probability distribution of the input data does not have to be known *a priori* in ANN modelling. However, more recently it has been pointed out that as the mean squared error function is generally used to optimize the weight vector in ANN models, the data need to be normally distributed in order to obtain optimal results (Fortin *et al.*, 1997). Similarly, the issue of stationarity has rarely been considered in ANN model development. However, there are good reasons why the removal of deterministic components in the data (i.e. trends, variance, seasonal and cyclic components) should be considered. It is generally accepted that ANNs cannot extrapolate beyond the range of the training data. Consequently, it is unlikely that ANNs can account for trends and heteroscedasticity in the data. One way to deal with this problem is to remove any deterministic components in the data prior to input to ANN. Normalization of data is warranted since most frequency curves of the hydrological variables are asymmetrically distributed, or are bound by zero (i.e. they are positive valued variables). Further, a highly skewed data series induces local variations in the function being mapped, and normalization of the data series brings the skewness to near zero, thus reducing the effect of local variations.

Hence it seems that if the guidelines to statistical model building are followed in the ANN approach, the performance of the MLP may be improved significantly. This issue has not so far been addressed or confirmed by empirical trials, where the model fits were the same regardless of whether raw or transformed data were used. The following case studies illustrate this heuristic by comparing the performance of independently developed ANN models using both raw and transformed data.

Case Studies

The concern of flood peak (extreme value) prediction in ANN river flow modelling has been addressed by developing ANN river flow models for the Baitarani river basin in the eastern part of India. The data series consists of 24 years (1972–1995) of daily values of runoff. The basin has a drainage area of 8570 km^2. Apart from this application, the approach has been tested for a few more basins and is summarized in the results and discussions section.

Methodology

A plot of the historic flow series of Baitarani river basin is presented in Figure 1 and it is apparent that

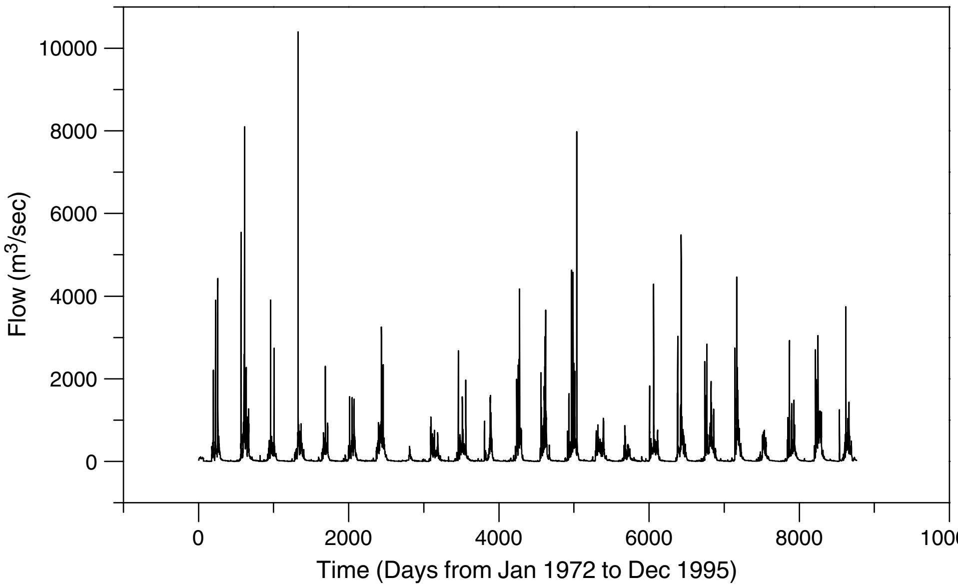

Figure 1. Historic flow series for the Baitarani River, India

Copyright © 2003 John Wiley & Sons, Ltd.

the data series is highly skewed and exhibits heteroscedasticity. It is clear from Figure 1 that the time series function contains high local variations. The skewness coefficient of the historic data considered in the study is 9·45, and confirms the local variations in the function. Therefore a transformation to reduce this skewness closer to zero is carried out. Different transformations are made to the original flow series and the Wilson–Hilferty transformation (Wilson and Hilferty, 1931) is found to fit well to the given data on the basis of chi-square statistics for different distributions. The procedure for a Wilson–Hilferty transformation is as follows (Salas *et al.*, 1985).

Firstly, normalization by log-transformation is applied to the original flows, $x_{v,T}$, via:

$$w_{v,T} = \log(x_{v,T}) \tag{5}$$

where x_T is the mean flow for day T and v denotes the year. Standardization is obtained by:

$$y_{v,T} = \frac{w_{v,T} - w_T}{\sigma_T} \tag{6}$$

where the resulting $y_{v,T}$ will have zero mean and unit variance, and w_T and σ_T are the mean and standard deviation of $w_{v,T}$.

Finally, the modified Wilson–Hilferty transformation (Kirby, 1972), which is valid for any value of skewness coefficient that preserves the first three moments, is given by:

$$z_{v,T} = \frac{6}{\gamma_T}\left\{\left[\frac{\gamma_T y'_{v,T}}{2} + 1\right]^{1/3} - 1\right\} + \frac{\gamma_T}{6} \tag{7}$$

where $y'_{v,T}$ satisfies the expression:

$$y'_{v,T} = \begin{cases} \max[y_{v,T}, -2/\gamma_T] & \text{if} \quad \gamma_T > 0 \\ \min[y_{v,T}, -2/\gamma_T] & \text{if} \quad \gamma_T < 0 \end{cases} \tag{8}$$

where $y_{v,T}$ is the log-transformed standardized variable and γ_T is the coefficient of skewness of the original series.

The transformed data series is presented in Figure 2, and the plot illustrates that an appropriate transformation reduces the local variations in the time series function. The transformed series showed a mean of −0·0034, standard deviation of 1·0217 and a skewness of −0·14. Figure 2 suggests that the transformed series is more or less a smooth function without any outliers.

However, it may be noted that the Wilson–Hilferty transformation used in this study need not fit to all series effectively as it depends on the statistical properties of the series. Therefore, for any data series the modeller has to find an appropriate transformation

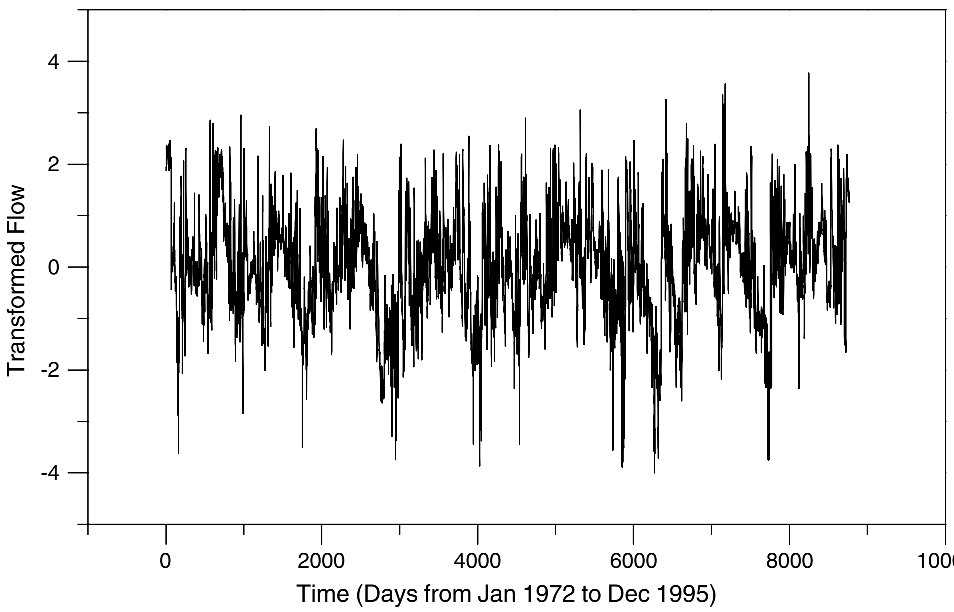

Figure 2. Wilson–Hilferty transformed flow series for Baitarani River, India

Copyright © 2003 John Wiley & Sons, Ltd.

based on goodness-of-fit statistics, so that the effect of local variations in the function is reduced.

A standard back-propagation algorithm is employed for training, and the hidden layer neurons are optimized by trial and error. Two models are developed, one using the raw data (ANN-RD) and the other using transformed data (ANN-TD). The input vector was identified using the procedure suggested by Sudheer *et al.* (2002), and consists of runoff values at two time lags (two antecedent values). The models are trained using data for 18 years (1972–1989) and validated on the rest of the data (1990–1995). The resulting hydrographs from both the models are analysed statistically using various indices employed for performance analysis of models. The goodness-of-fit statistics considered are the root mean square error (RMSE) between the computed and observed runoff, the coefficient of correlation (CORR) and the model efficiency (EFF).

Results and Discussion

The values of performance indices for the optimal neural network models for Baitarani basin are presented in Table I. The RMSE statistic is a measure of residual variance and is indicative of the model's ability to predict high flows (Hsu *et al.*, 1995). Considering the magnitude of the peak flow during the period of study (10 339 $m^3 s^{-1}$, see Figure 1), the ANN models were able to compute the high flows with reasonable accuracy as can be evidenced by the low RMSE values (see Table I). Comparing ANN-RD and ANN-TD models, the relatively low RMSE values suggest that the ANN-TD model estimates the high flows relatively better. It is worth noting that the ANN-RD model has poor efficiency during training, suggesting a large amount of unexplained variance for the model. The efficiency statistic is an indicator of the accuracy of model predictions away from the mean, and the results suggest that the ANN-RD model's prediction away from the mean would not be accurate. However, the efficiency of the ANN-RD model improved during validation, indicating good generalization properties.

It may be noted that in the total training data set corresponding to the period 1972–1989, the number of patterns containing high flows is significantly less than the number of patterns containing low and medium flows (see Figure 1), and therefore the

Table I. Comparison of performance of ANN models developed on raw data and transformed data

	Training		Validation	
	ANN-RD	ANN-TD	ANN-RD	ANN-TD
RMSE	0·0263	0·0181	0·0411	0·0401
Efficiency (%)	53·62	70·38	81·93	81·32
Correlation	0·7323	0·839	0·9051	0·9020

parameters estimated based on this training set have a bias towards the low and medium flows in the case of the ANN-RD model, and thus exhibit poor efficiency. The validation period 1990–1995 consists of relatively low and medium flows (see Figure 1), and consequently the model exhibited good performance during validation. On the other hand, the ANN-TD model shows improved performance over the ANN-RD model in terms of efficiency during calibration as well as validation (see Table I). The explained variance for the ANN-RD model is found to be 73·22%, which improved to 90·51% for the ANN-TD model. It is worth noting that unlike the ANN-RD model there is no significant change in performance during training and validation for the ANN-TD model, implying the transformation reduces the effect of high flows in data series. This reinforces the earlier discussed heuristic that an appropriate transformation can smooth the data series by removing any outlier.

Table II depicts the percentage error in annual peak flow estimates for the validation years for both models. While Table II is self-explanatory, it indicates that the annual peak flow estimates are improved by the ANN-TD model. However, it is worth mentioning that the ANN models tend to underestimate the peak flow even after data transformation, but the magnitude of underestimation is significantly reduced for ANN-TD. This observation implies that the local variations

Table II. Percentage error in peak estimation for ANN models

Year	ANN-RD	ANN-TD
1990	−44·85	−10·12
1991	−48·51	−34·15
1992	10·41	−3·01
1993	−65·01	−26·20
1994	−77·15	−2·75
1995	−58·67	−37·83

Copyright © 2003 John Wiley & Sons, Ltd.

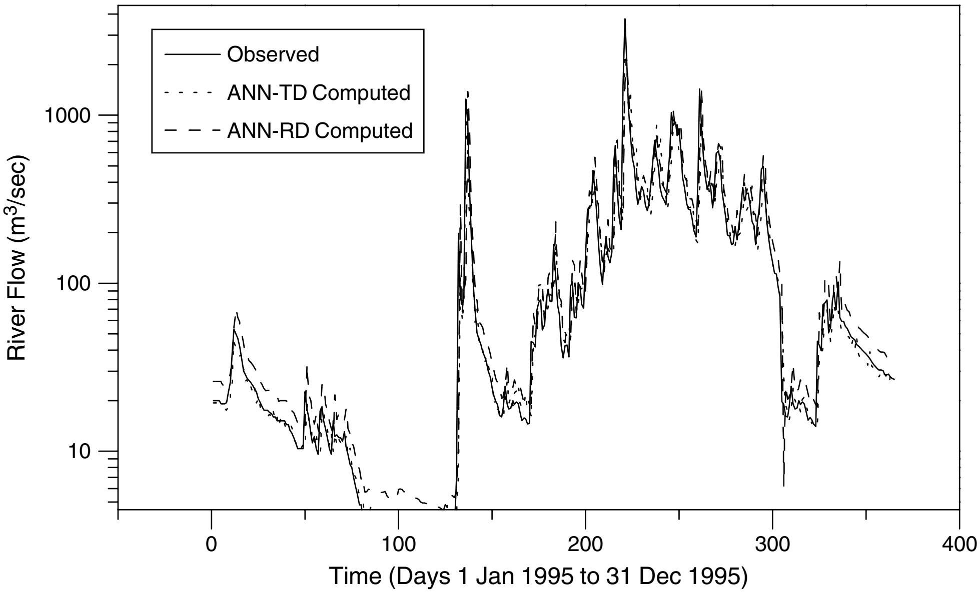

Figure 3. Computed and recorded hydrograph for the year 1995

in the function being mapped are not fully removed by transformation, but are considerably reduced. The observed and computed hydrographs (by both models) for the validation year 1995 are presented in Figure 3 and the effectiveness of data transformation is clear. It is observed that all ranges of flow have been mapped well by ANN-TD, while the low flows are overestimated by ANN-RD, implying an improvement in performance of ANN by data transformation.

An analysis to assess the potential of each of the models to preserve the statistical properties of the historic flow series reveals that the flow series computed by the ANN-TD model reproduces the first three statistical moments (i.e. mean, standard deviation and skewness) better than that computed by the ANN-RD model. The values of the first three moments for the historic and computed flow series for the validation years are presented in Table III for comparison. It can be observed from Table III that if the data series is less skewed (e.g. during the year 1992), implying reduced local variations in the time series function, the models preserve the moments satisfactorily irrespective of data transformation being performed or not. Also, as the skewness increases, ANN-RD fails in preserving moments compared to ANN-TD.

It is observed that data transformation does not affect the model parsimony as the final architecture for both models is the same (two hidden neurons). Similarly, there is no significant change in the computational time required (about 350 epochs) for training the models, implying better model performance with no additional cost by data transformation.

Table III. Statistical moments of the observed and computed flow series during validation

	Year	Historical	ANN-RD	ANN-TD
Mean	1990	204·34	211·85	190·49
	1991	193·61	179·03	186·00
	1992	86·97	92·26	87·45
	1993	143·13	153·13	136·75
	1994	239·26	230·14	245·74
	1995	152·04	158·59	150·00
Standard Deviation	1990	341·95	313·88	331·58
	1991	459·93	343·05	447·69
	1992	144·06	155·68	143·69
	1993	270·36	249·62	272·57
	1994	430·40	371·08	464·90
	1995	312·55	260·20	332·00
Skewness	1990	3·45	2·92	3·63
	1991	4·74	3·34	5·24
	1992	2·43	2·22	2·38
	1993	4·67	3·59	4·37
	1994	2·82	2·25	3·18
	1995	5·97	3·89	5·44

Copyright © 2003 John Wiley & Sons, Ltd.

To evaluate the statistical significance of the proposed methodology, the method has been tested for modelling river flows of a few more basins, Tel, Uttei, Vamsadhara and Choladipuzha. The details of the basins, and other methodological particulars, are presented in Table IV. These case studies provide a preliminary investigation of the proposed approach on basins of varied size (in terms of basin area). It can be observed from Table IV that the skewness of the original series is considerably reduced by appropriate data transformation. It may be noted that the transformation function varies from basin to basin, as it depends on the statistical properties of the data series.

The results pertaining to these case studies are summarized in Table V, and all the results are in direct agreement with the advantages of the proposed methodology. The explained variance of ANN-TD is much better in all cases compared to ANN-RD, implying a general improvement of ANN models when developed on transformed data. The magnitude of error in peak flow estimation is considerably reduced by data transformation in all cases. From Table V, it is apparent that there exists a definite correlation between the skewness of the data series and the percentage error in peak flow estimation, as observed in the case of Baitarani basin also. This observation reinforces the heuristic that the underestimation of peak flows in ANN models is caused by local variations due to high skewness in the data series. The results of these case studies establish the statistical significance of the proposed approach. The foregoing discussions clearly demonstrate that the performance of an MLP can be significantly improved by an appropriate transformation to the historical data prior to model building.

Summary and Conclusions

This paper discusses the issue of peak flow prediction by ANN river flow models, and presents an appropriate statistical procedure to address the problem. From the aforementioned case studies and discussions it can be concluded that the proposed procedure would lead to better estimates of the peak flows by ANN models. Since the proposed methodology is based on the information contained in the data series itself, and is based on clear statistical properties as decision rules, the approach becomes more explicit and can be adopted for any basin. The specific advantages of the approach are: (1) it can improve the overall performance of the ANN models; (2) it preserves the statistical properties of the original series; (3) it is simpler and quicker to use, since there is no need for large pre-processing of the data; (4) it does not affect the model parsimony, as well as no additional computational burden being introduced. The results of applications of the proposed approach strengthen the potential of the procedure, and show the statistical significance of the approach. However, further empirical studies to investigate the effect of data transformation on other activation functions used in ANN may be required to reinforce this conclusion. In addition, the comparative performance of other training algorithms on transformed data should also be investigated.

Acknowledgement

We wish to express our sincere gratitude to Professor Hoshin V. Gupta, University of Arizona, Tucson whose exuberant suggestions and remarks have greatly helped us to improve the quality of the manuscript.

Table IV. Basin and methodological details of other case studies

Sl. no.	Basin	Drainage area (km^2)	Data length (years)	Skewness of original data series	Transformation procedure employed	Skewness of transformed data series
1	Tel (sub-basin of Mahanadi)	19 600	1973–1994	6·954	Wilson–Hilferty	0·1181
2	Uttei (sub-basin of Mahanadi)	456	1976–1994	31·530	Log normal	0·0764
3	Vamsadhara	10 830	1982–1988	8·750	Square root	0·5010
4	Choladipuzha	98	1995–1999	2·350	Square root	0·1449

Copyright © 2003 John Wiley & Sons, Ltd.

Table V. Performance indices of ANN models for different basins

Model	Indices	Tel		Uttei		Vamsadhara		Choladipuzha	
		Training (1973–89)	Validation (1990–94)	Training (1976–90)	Validation (1991–94)	Training (1982–86)	Validation (1987–88)	Training (1995–97)	Validation (1998–99)
ANN-TD	RMSE	0·053	0·040	0·043	0·032	0·079	0·056	0·107	0·108
	Efficiency (%)	88·300	88·890	95·600	88·530	88·070	84·340	83·660	77·930
	Correlation	0·940	0·944	0·978	0·942	0·939	0·925	0·915	0·883
	Explained variance (%)	93·960	95·020	97·780	98·530	93·840	93·880	91·460	91·600
	Percentage error in peak flow	−3·280	−2·940	−2·180	−2·280	−10·760	−4·354	−6·254	−6·101
ANN-RD	RMSE	0·033	0·043	0·009	0·011	0·034	0·008	0·046	0·060
	Efficiency (%)	67·020	69·010	83·560	15·190	51·110	74·330	86·710	82·710
	Correlation	0·819	0·831	0·914	0·413	0·715	0·878	0·931	0·913
	Explained variance (%)	81·860	80·490	91·410	54·780	71·290	73·200	93·090	83·100
	Percentage error in peak flow	−32·750	−36·130	−10·004	−53·460	−48·016	−17·166	−24·080	−32·560

Copyright © 2003 John Wiley & Sons, Ltd.

References

ASCE Task Committee. 2000a. Artificial neural networks in hydrology—I: Preliminary concepts. *Journal of Hydrologic Engineering, ASCE* **5**(2): 115–123.

ASCE Task Committee. 2000b. Artificial neural networks in hydrology—II: Hydrologic applications. *Journal of Hydrologic Engineering, ASCE* **5**(2): 124–137.

Brown M, Harris C. 1994. *Neuro-fuzzy Adaptive Modeling and Control*. Prentice Hall: New York.

Burke LI, Ignizio JP. 1992. Neural networks and operation research: an overview. *Computer and Operation Research* **19**(3/4): 179–189.

Carrol RJ, Ruppert D. 1988. *Transformation and Weighting in Regression*. Chapman and Hall: New York.

Cheng B, Titterington DM. 1994. Neural networks: A review from a statistical perspective. *Statistical Sciences* **9**(1): 2–54.

Chon KH, Cohen RJ. 1997. Linear and nonlinear ARMA model parameter estimation using an artificial neural network. *IEEE Transactions on Biomedical Engineering* **44**(3): 168–174.

Compolo M, Andreussi P, Soladati A. 1999. River flood forecasting with a neural networks model. *Water Resources Research* **35**: 1191–1197.

Connor JT, Martin RD, Atlas LE. 1994. Recurrent neural networks and robust time series prediction. *IEEE Transactions on Neural Networks* **5**(2): 240–254.

Dawson CW, Wilby R. 1998. An artificial neural network approach to rainfall runoff modeling. *Hydrological Science* **43**(1): 47–66.

Fortin V, Quarda TBMJ, Bobee B. 1997. Comments on 'The use of artificial neural networks for the prediction of water quality parameters' by Maier HR, Dandy GC. *Water Resources Research* **33**(10): 2423–2424.

Hill T, Marquez L, O'Connor M, Remus W. 1994. Artificial neural network models for forecasting and decision making. *International Journal of Forecasting* **10**: 5–15.

Hornik K, Stinchcombe M, White H. 1989. Multilayer feed forward networks are universal approximators. *Neural Networks* **2**: 359–366.

Hsieh C. 1993. Some potential applications of artificial neural networks in financial management. *Journal of Systems Management* **April**: 12–15.

Hsu K, Gupta VH, Sorooshian S. 1995. Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research* **31**(10): 2517–2530.

Imrie CE, Durucan S, Korre A. 2000. River flow prediction using artificial neural networks: generalization beyond the calibration range. *Journal of Hydrology* **233**: 138–153.

Karunanithi N, Grenney WJ, Whitley D, Bovee K. 1994. Neural networks for river flow prediction. *Journal of Computing in Civil Engineering, ASCE* **8**(2): 201–220.

Kirby W. 1972. Computer oriented Wilson–Hilferty transformation that preserves the first three moments and the lower bound of the Pearson type 3 distribution. *Water Resources Research* **8**(5): 1251–1254.

Maier HR, Dandy GC. 1997. Determining inputs for neural network models of multivariate time series. *Microcomputers in Civil Engineering* **12**: 353–368.

Maier HR, Dandy GC. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and application. *Environmental Modeling & Software* **15**: 101–124.

Manson JC, Parks PC. 1992. Selection of neural network structures—some approximation theory guidelines. In *Neural Networks for Control Systems*, Warwicks K, Irwin GW, Hunt KJ (eds). IEE Control Engineering Series 46, Peter Peregrinus: Stevenage; 151–180.

Minns AW, Hall MJ. 1996. Artificial neural networks as rainfall-runoff models. *Hydrological Sciences Journal* **41**(3): 399–417.

Ripley BD. 1994. Neural networks and related methods of classification. *Journal of the Royal Statistical Society, Series B* **56**(3): 409–456.

Salas JD, Tabios QV, Bartolini P. 1985. Approaches to multivariate modeling of water resources time series. *Water Resources Bulletin* **21**(4): 683–708.

Sarle WS. 1994. Neural networks and statistical models. In *Proceedings of the Nineteenth Annual SAS Users Group International Conference*. SAS Institute: Cary, NC; 1538–1550.

See L, Corne S, Dougherty M, Openshow S. 1997. Some initial experiments with the neural network models of flood forecasting on the river Ouse. In *Proceedings of the 2nd International Conference on GeoComputations*, University of Otago, Dunedin, New Zealand, 26–29 August.

Sudheer KP, Gosain AK, Ramasastri KS. 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes* **16**: 1325–1330.

Wilson EB, Hilferty MM. 1931. The distribution of chi-square. *Proceedings of the National Academy of Science* **17**: 684–688.

Copyright © 2003 John Wiley & Sons, Ltd.