

Neural network modelling of non-linear hydrological relationships

R. J. Abrahart¹ and L. M. See²

¹School of Geography, University of Nottingham, Nottingham NG7 2RD, UK

²School of Geography, University of Leeds, Leeds, LS2 9JT, UK

Received: 8 January 2007 – Published in Hydrol. Earth Syst. Sci. Discuss.: 22 February 2007

Revised: 26 July 2007 – Accepted: 10 September 2007 – Published: 20 September 2007

Abstract. Two recent studies have suggested that neural network modelling offers no worthwhile improvements in comparison to the application of weighted linear transfer functions for capturing the non-linear nature of hydrological relationships. The potential of an artificial neural network to perform simple non-linear hydrological transformations under controlled conditions is examined in this paper. Eight neural network models were developed: four full or partial emulations of a recognised non-linear hydrological rainfall-runoff model; four solutions developed on an identical set of inputs and a calculated runoff coefficient output. The use of different input combinations enabled the competencies of solutions developed on a reduced number of parameters to be assessed. The selected hydrological model had a limited number of inputs and contained no temporal component. The modelling process was based on a set of random inputs that had a uniform distribution and spanned a modest range of possibilities. The initial cloning operations permitted a direct comparison to be performed with the equation-based relationship. It also provided more general information about the power of a neural network to replicate mathematical equations and model modest non-linear relationships. The second group of experiments explored a different relationship that is of hydrological interest; the target surface contained a stronger set of non-linear properties and was more challenging. Linear modelling comparisons were performed against traditional least squares multiple linear regression solutions developed on identical datasets. The reported results demonstrate that neural networks are capable of modelling non-linear hydrological processes and are therefore appropriate tools for hydrological modelling.

1 Introduction

The last decade has witnessed a virtual explosion of neural network (NN) modelling activities throughout the hydrological sciences. It is readily apparent from the increasing number of published case studies that the development of data-driven solutions based on the use of neural tools or smart technologies is being trialled and tested in most sectors of hydrological modelling and hydraulic engineering. Numerous extended descriptions exist and for detailed summaries the interested reader is referred to the following papers: ASCE (2000a, b); Maier and Dandy (2000); Dawson & Wilby (2001) and edited volumes: Govindaraju and Rao (2000); Abrahart et al. (2004). Neural technologies continue to make enormous strides in their struggle to become established as recognized tools that offer efficient and effective solutions for modelling and analysing the behaviour of complex dynamical systems. Time series forecasting has been a particular focus of interest and superior performing models have been reported in a diverse set of fields that include rainfall-runoff modelling (ASCE, 2000a, b; Dawson and Wilby, 2001; Birikundavy et al., 2002; Campolo et al., 2003; Huang et al., 2004; Riad et al., 2004; Hettiarachchi et al., 2005; Senthil Kumar et al., 2005) and sediment prediction (Abrahart and White, 2001; Nagy et al., 2002; Yitian and Gu, 2003; Kisi, 2004; Bhattacharya et al., 2005; Kisi, 2005). Moreover, for flood forecasting purposes, neural solutions offer practical advantages related to operational costs and socio-economic resources that would be of interest in developing countries, e.g. rapid development; rapid execution; parsimonious requirements; open source code (Shamseldin, 2007). Two recent catchment studies have nevertheless questioned the use of such tools for non-linear hydrological modelling purposes. Gaume and Gosset (2003) and Han et al. (2007) concluded: (1) that for short term forecasting purposes neural solutions offered no real advantages over traditional linear transfer functions; (2) that the demands

Correspondence to: R. J. Abrahart
(bob.abrahart@nottingham.ac.uk)

and complexities involved in the development of neural solutions made them difficult to use and therefore “uncompetitive” (Han et al., 2007, p.227); (3) that there is still much to be done to improve our understanding about the uncertain nature and hydrological characteristics of neural solutions “before [such mechanisms] could be used as a practical tool in real-time operations” (Han et al., 2007, p.228); and (4) that the potential merit of putting further resources into the development of black box computational intelligence methodologies such as feedforward neural networks remains questionable since “the quest for a universal model requiring no hydrological expertise might well be hopeless” (Gaumeand Gosset, 2003, p.705).

The main objective of this paper is to evaluate the ability of a NN model to capture non-linear effects within the non-linearity range employed in traditional rainfall-runoff models of the type that are widely used in operational flood forecasting systems, e.g. black box and conceptual rainfall-runoff models. There are various sources of non-linearities in the rainfall-runoff transformation process, e.g. antecedent conditions are a substantial cause of such effects. The present paper will examine hydrological non-linearities attributed to spatial variation of maximum soil moisture storage capacities as implemented in the Xinjiang Rainfall-Runoff Model (Zhao et al., 1980). No specific river records were involved; so no general principles, derived from a single, perhaps atypical case, would be produced, i.e. arising out of the particularities of an observed dataset or related to an individual catchment. The modelling procedures that are presented in this paper can also be used to capture other sources of non-linearity in that particular model as well as hydrological non-linearities that are simulated in a complex physical-based model. This paper considers the potential impact of non-linear effects over a modest range of varying hydrological parameters. The reported selections are considered to be sufficient for the purposes of our argument; using input parameters that encompassed “extended ranges” or “limiting cases” for the purposes of completeness will not alter the main conclusions of our paper for the water sector. If modelling typical hydrological ranges and minor non-linear properties can produce incontrovertible outcomes, applying identical procedures to more pronounced non-linear relationships, will produce nothing other than a set of similar or stronger confirmations. Exploring the extent to which the full range of complex non-linear relationships that might be contained in the original conceptual model might however be a useful computer science endeavour and could be tested at a later point.

This paper will also examine the modelling assumptions and individual interpretations that were reported in the two earlier critical studies: Gaume and Gosset (2003) reported that a NN model can sometimes produce results that are similar to a weighted linear transfer function (WLTF) model; Han et al. (2007) reported that a NN model offered no advantages for short-term forecasts over a weighted linear transfer function (WLTF) model. The need to perform non-linear

transformations is a fundamental aspect of most hydrological modelling applications and real-time forecasting operations. The need to question technological and methodological approaches that fail to encapsulate such properties is paramount. Neural solutions purport to model simple linear, complex non-linear, or multifaceted hybrid relationships so the potential reasons for such apparent shortcomings must be clarified. It is not sufficient for such matters to be left unanswered. The current paper will address a straightforward matter: NN capabilities to discover and reproduce non-linear hydrological relationships. The need for more experiments of a similar nature that attempt to settle fundamental issues within the hydrological sciences is also championed. It is suggested that a complacent misbelief has developed; enthusiasts and proponents of neural solutions might well consider a set of recorded findings that document the power of such technologies to model non-linear relationships as tantamount to a “confirmation of established science” or about “preaching to the converted”. The main issues can perhaps be related to a common trigger; most inventions and discoveries experience initial rapid development that is often accompanied by exaggerated claims about what can or cannot be achieved. However, following an initial series of reported successes, a detailed set of methodological underpinnings is required to support the development of subsequent applications. NN hydrological modelling has now reached this stage; there are large gaps in our knowledge and substantial issues still need to be resolved. Han et al. (2007, p.223) commented that the large number of unsolved questions continues to hinder the application of such tools amongst practising hydrologists. Neural solutions also encounter “institutional barriers”: Zhang et al. (2004, p.iv) noted that “. . . it is relatively new and vastly different from conventional approaches. As such, it is often summarily dismissed by experimenters who do not wish to apply new approaches, even in cases where use of the technology has been proven”. The extent of such rejection is perhaps best exemplified by considering the limited number of operational applications that have been implemented. It is argued that continued research and development in this field will provide a stronger understanding and appreciation of the hydrological modelling opportunities that are on offer such that good scholarship and greater awareness might encourage the wider acceptance of neural solutions. The debate about the potential benefits of computational intelligence methodologies in the hydrological sciences is ongoing and should be encouraged in the scientific press. This paper is an attempt to engage in one particular aspect of that debate.

Modelling involved two groups of controlled experiments. It was based on an ideal non-linear hydrological modelling problem set in a data-rich environment. For each initial experiment that comprised a full set of inputs, further models were developed on a reduced set of inputs, to explore the relative impact of each input, and to assess the overall competencies of neural solutions for resolving difficult situations

under imperfect conditions, i.e. a missing parameter. Eight neural network models were developed; four full and partial emulations of the Xinanjiang Rainfall-Runoff Model; four further solutions developed on an identical set of inputs and a calculated runoff coefficient output. The aim of an emulator, in the most general sense, is to duplicate the functions of one system with a different system, so that the second system appears to behave like the first system. However, unlike a simulator, it does not attempt to precisely model the state of the device being emulated; it only attempts to reproduce its behaviour. Full and partial emulation supported two outcomes; it permitted a direct comparison to be performed against the computed mathematical relationship; it also provided more general information on the power of neural solutions to model non-linear relationships. The second set of experiments is not an “emulation process”. Four neural solutions were developed on identical input datasets but the output is different: it is a calculated “runoff coefficient”. Loke et al. (1997) modelled this dimensionless empirical parameter with a NN. It describes the proportion of total rainfall in a storm event that is converted into runoff. It is a traditional catchment parameter and was part of the “Rational Method” (or Lloyd-Davies Method) first published in 1851 (Mulvaney, 1851; Beven, 2001). It measures “overall responsiveness” and not actual discharge: in spatial terms it relates to different or changing land use and hydrologic soil groups; in non-spatial terms it depends on antecedent conditions and rainfall. The motivation for modelling such outputs is twofold: the relationship is implicit as opposed to explicit and will need to be discovered; the target surface is more challenging and presented a stronger non-linear relationship. The conceptual model was formulated as a single equation that had a small number of input variables and no temporal component. The mathematical relationship was transparent and the nature of the relationship is clearly non-linear. The inputs to the rainfall-runoff equation were produced in a statistical random pattern generator and spanned a range of different hydrological scenarios. The eight experiments and their reported results offer unambiguous proof that neural solutions can capture non-linear hydrological relationships. It is thus logical to conclude that there is potential merit in continuing to undertake substantive research on different topics in this field of hydrological modelling.

2 Xinanjiang Rainfall-Runoff Model

The Xinanjiang Rainfall-Runoff Model (named after the river to which it was first applied) was developed in 1973 and first published in 1980 (Zhao et al., 1980; Zhao, 1992). It is a conceptual rainfall-runoff forecasting tool that was designed for humid and semi-humid regions and is based on the concept of runoff formation on repletion of storage, i.e. runoff is not produced until the soil moisture content of the aeration zone reaches field storage capacity and there-

after runoff equals rainfall excess without further loss. It can also be employed in a semi-distributed manner: the catchment is divided into a number of sub-catchments to which the model is thereafter applied using different parameter sets and/or input information, i.e. a “probability soil moisture distributed model”. The model has been applied with success to large areas including all of the agricultural, pastoral and forested lands (except for the loess) of China (Zhao and Liu, 1995, p.230). This model has a small number of parameters, its structure and components have strong physical meaning, and such factors in combination make it a popular tool for hydrological modelling. The basic model has experienced numerous internal modifications, e.g. use of a double parabolic curve to describe complex soil moisture storage distributions (Jayawardena and Zhou, 2000); incorporation of two different mechanisms for the production of runoff (Hu et al., 2005). The model has also been used in large scale land-surface hydrological applications (e.g. the Variable Infiltration Capacity Model; Wood et al., 1992) and coupled to mesoscale precipitation forecasts where it produced encouraging flood simulation outputs (Lin et al., 2006). The non-complicated nature of this model continues to make it a popular choice for hydrological experimentation, e.g. for testing intelligent calibration procedures (Cheng et al., 2002; 2006) or for distributed modelling purposes (Su et al., 2003; Chen et al., 2007). The model has also been incorporated into nationwide forecasting methodologies, e.g. the USA National Weather Service River Forecasting System (M. Kane, Riverside Technology Inc., personal communication). There are strong similarities between this model and the Probability Distributed Model developed at the Centre for Ecology and Hydrology (Moore, 2007; see also Moore, 1985; Senbeta et al., 1999). The ARNO Rainfall-Runoff Model (Todini, 1996) is derived from the Xinanjiang Model and has been incorporated into a climate model (Dümenil and Todini, 1992); the NUARNO Model is based on the ARNO Model and forms an integrated part of the UK NERC-ESRC Land Use Programme Decision Support System where it is used to predict the direction and magnitude of the hydrological response that results from proposed changes in land-use (Adams et al., 1995, p.56–58).

In its simplest form the model comprises a single equation:

$$R = P - (Wm - Wo) + Wm \left[\left(1 - \frac{Wo}{Wm} \right)^{\frac{1}{1+b}} - \frac{P}{(1+b)Wm} \right]^{1+b} \quad (1)$$

in which R = runoff; P = effective precipitation; Wm = maximum field storage capacity; Wo = initial field storage capacity; and b is an exponent that represents “non-uniform spatial distribution”, i.e. the non-uniform distribution of surface conditions including factors such as topography, geology, soil type and vegetation coverage.

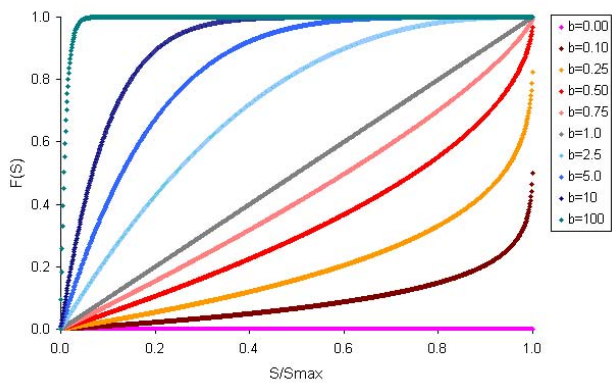


Fig. 1. Field storage distribution for different settings of b parameter.

The model was implemented according to the following rules:

$$VOL = \left[\left(1 - \frac{W_o}{W_m} \right)^{\frac{1}{1+b}} - \frac{P}{(1+b)W_m} \right]^{1+b} \quad (2)$$

If $VOL \leq 0$ $R = P - (W_m - W_o)$ (3)

If $VOL > 0$ $R = P - (W_m - W_o) + W_m (VOL)^{1+b}$ (4)

If $W_o = W_m$ $R = P$ (5)

Dividing by W_m permits Eq. (1) to be converted into a non-dimensional equivalent:

$$\frac{R}{W_m} = \frac{P}{W_m} - \left(1 - \frac{W_o}{W_m} \right) + \left[\left(1 - \frac{W_o}{W_m} \right)^{\frac{1}{1+b}} - \frac{P}{(1+b)W_m} \right]^{1+b} \quad (6)$$

Defining $C_1 = \frac{P}{W_m}$; $C_2 = \frac{W_o}{W_m}$; $C_3 = \frac{R}{W_m}$ permits Eq. (2) to be expressed as:

$$C_3 = C_1 + C_2 - 1 + \left[\left(1 - C_2 \right)^{\frac{1}{1+b}} - \frac{C_1}{(1+b)} \right]^{1+b} \quad (7)$$

The “runoff coefficient” was calculated on inputs and outputs for Eq. (1) as:

$$C_4 = \frac{R}{P} \quad (8)$$

Each model was developed and tested on a dataset of random input variables comprising 5000 records split into two equal groups: 2500 cases for training purposes; 2500 cases for split sample testing operations. To permit meaningful assessment of non-linear hydrological responses that are attributed to maximum soil moisture storage capacities a random number between “half-full” and “maximum” was assigned to initial soil water [W_o]. It also avoided the need to

model exceptional responses related to drier antecedent conditions. The input values for effective precipitation [P], maximum soil water storage [W_m] and the curve fitting exponent [b] were random samples computed in MINITAB. The random sampling operation was used to create uniform distributions that had fixed ranges:

- Effective precipitation (mm h^{-1}) between 0 and 50
- Maximum soil water storage (mm) between 50 and 100
- b (dimensionless) between 0.1 and 0.5

This selection of input ranges denotes the extent to which our reported findings are valid. Each selection can be explained and placed in its historical hydrological context. Minimums were fixed at or near to their lowest permitted levels; maximums were set at sufficient magnitudes to ensure that the final dataset contained a broad assortment of different input scenarios and a reasonable set of output responses. It was also considered prudent to avoid the use of extreme values and excessive ranges in an attempt to provide results that were considered to be more representative of the “general case”. Todini (1996, p.359) reports that maximum rainfall intensities are in most cases $< 100 \text{ mm h}^{-1}$. To avoid extremes the inputs for effective precipitation were set to range from zero to half this level. He also reported that average catchment soil water storage was around 50–300 mm (p.357). His lower limit was accepted as the established minimum level; setting the upper limit, however, is more difficult. This input variable reflects a complex interaction between the soil layer itself and the prevailing climatic conditions, such that in most cases maximum soil water storage in humid regions is lower than in arid regions, i.e. ranging from 80 mm in South China to 170 mm or more in North China (Zhao and Liu, 1995, p.225). To reduce potential disparities the upper limit for soil moisture storage was set at less than or equal to twice the rainfall input driver; the established level being near to that recorded in more arid regions; near to half that that recorded in more humid regions; whilst at the same time being reduced to one third of its highest potential value in an effort to contain the scope of the modelling effort.

The dimensionless curve fitting exponent b can have a wide range of values. Field storage capacity is assumed to vary spatially and this variation can be described in terms of a probability distribution function $F(S)$ which yields the proportion of the catchment area that has field storage capacity values less than or equal to S . The Xinanjiang Rainfall-Runoff Model uses the Pareto Probability Distribution Function according to:

$$F(S) = 1 - \left[1 - \frac{S}{S_{\max}} \right]^b \quad \text{for } 0 \leq S \leq S_{\max} \quad (9)$$

Figure 1 reveals the impact that this parameter has on controlling spatial variability where S_{\max} is the maximum value

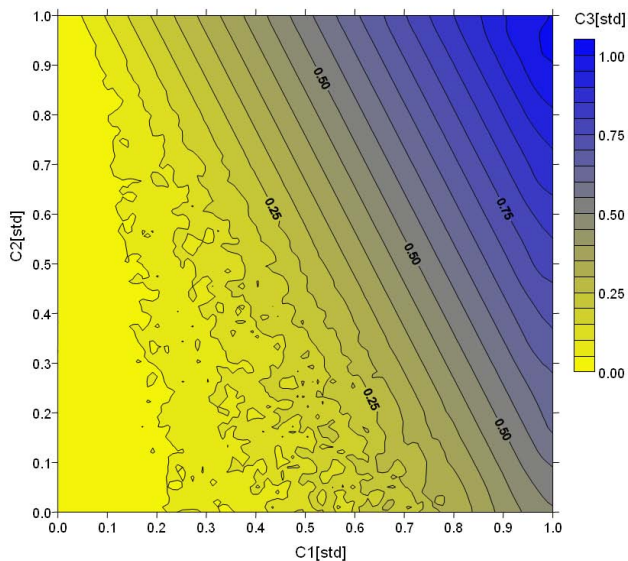


Fig. 2. Target surface of equation outputs for Experiment A4.

of the field storage capacity. The shape of the curve changes from concave to convex if the value of b exceeds 1. There are special cases when $b=0$ (constant storage), $b=1$ (uniform distribution) and $b=\infty$ (constant storage). Experience suggests that for traditional basins, i.e. 10 km^2 , $b=0.1$; whereas for basins measured in thousands of square kilometres, $b=0.4$ or higher, perhaps even extending to 1 or 2 or more (Zhao and Liu, 1995). Todini (1996) in his initial considerations for the ARNO Model is more conservative stating that b should perhaps be in the order of 0.1 to 0.01 (p.357); since large values for the soil moisture curve shape parameter, i.e. $b > 1$ can produce a “peaky response” on drier soils (p.358), i.e. flows that rise and fall in the manner of a direct and “near-linear” or “near-instantaneous” response to rainfall events. Todini (2002; p.707) in a subsequent paper nevertheless suggests a useful permitted range of 0.1 to 0.5. This range has been adopted as a good hydrological compromise but it is nevertheless recognised that the following methodological limitations will result: (i) insufficient material will be created to support the capture of a comprehensive profile of non-linear events; (ii) the final product will not include mixed cases of concave and convex curves; and that (iii) the special limiting cases are not considered.

Linear scaling was applied to the input drivers and output responses with each variable being standardised to a fixed range [0–1]. To minimise the number of numerical conversions required all inputs and outputs are henceforth reported in terms of standardised units: $C_1(\text{std})$, $C_2(\text{std})$, $C_3(\text{std})$ and $C_4(\text{std})$. Target plots for paired inputs are provided in Figs. 2 and 3. Each plot reveals the complex nature of the non-linear relationships that are encapsulated in the conceptual model and such figures hold the key for interpreting our results. Figure 2 depicts the complex nature of the $C_3(\text{std})$ “target sur-

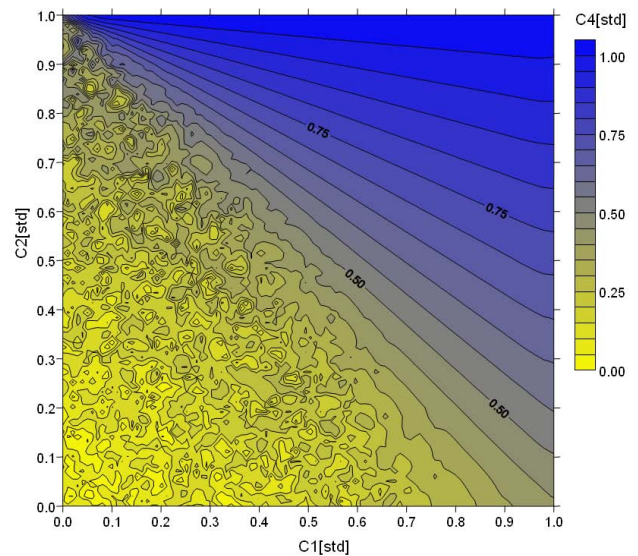


Fig. 3. Target surface of equation outputs for Experiment B4.

face”. This surface reveals a strong trend running from lower left to upper right. The predictand is confirmed to be a modest non linear output in two respects: (i) surface flexing is apparent, towards the lower left corner; (ii) rotational curvature occurs, pivoting about the upper left corner. The lowest levels of $C_1(\text{std})$ produced the lowest levels of $C_3(\text{std})$ and a set of shallower slopes; increased levels of $C_1(\text{std})$ produced a rougher surface, with a speckled pattern, and a stronger gradient. This trend towards stronger gradients develops into a well defined smoother and steeper inclined planar surface that runs from mid to higher values of $C_3(\text{std})$. Figure 3 depicts the complex nature of the $C_4(\text{std})$ “target surface”. This surface reveals a similar trend running from lower left to upper right. However, on this occasion, the predictand is confirmed to possess intensified non linear properties: (i) surface flexing is more pronounced, producing stronger diversities across a broader range of outputs; (ii) rotational curvature is more pronounced, having marked pivoting about the upper left corner. The lower levels of $C_1(\text{std})$ produced a rougher surface and shallower slopes; increased levels of $C_1(\text{std})$ produced a well defined smoother and steeper inclined curvilinear surface that runs from mid to higher values of $C_3(\text{std})$. This twofold separation of the output surface for both predictands arises out of Eqs. (3) and (4). The lower section is attributed to Eq. (4). This is a complex non-linear multi-valued function in which b is an important controlling parameter. However, if effective precipitation is sufficient to augment the storage deficit, the impact of this controlling parameter is eliminated and runoff is a simple function of $C_1(\text{std})$ and $C_2(\text{std})$. This results in the production of a smooth linear or curvilinear surface and corresponds to the second part of the diagram (upper right). The two target surfaces nevertheless exhibit different properties such that a quantitative

Table 1. Experiments classified and numbered according to input and output variables.

Group	Experiment	Input(s)			Output
A	A1	$C_1(\text{std})$	$C_2(\text{std})$	$b(\text{std})$	$C_3(\text{std})$
	A2		$C_2(\text{std})$	$b(\text{std})$	$C_3(\text{std})$
	A3	$C_1(\text{std})$		$b(\text{std})$	$C_3(\text{std})$
	A4	$C_1(\text{std})$	$C_2(\text{std})$		$C_3(\text{std})$
B	B1	$C_1(\text{std})$	$C_2(\text{std})$	$b(\text{std})$	$C_4(\text{std})$
	B2		$C_2(\text{std})$	$b(\text{std})$	$C_4(\text{std})$
	B3	$C_1(\text{std})$		$b(\text{std})$	$C_4(\text{std})$
	B4	$C_1(\text{std})$	$C_2(\text{std})$		$C_4(\text{std})$

$$C_1 = P/Wm; C_2 = Wo/Wm; C_3 = R/Wm; C_4 = R/P$$

comparison of numerical fit is awkward. Each group of models will be modelling a different sort of relationship; in comparison to $C_4(\text{std})$, $C_3(\text{std})$ is less demanding in terms of gradients and rotational factors.

Seven different performance statistics were used for comparing the output results. HydroTest¹ was used to perform the required calculations; a description of the relevant metrics and their associated equations can be found on that website and in its companion paper (Dawson et al., 2007). Model performance was assessed on the basis of two absolute statistics, two relative statistics and three dimensionless indices: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Relative Error (MARE), Mean Squared Relative Error (MSRE), Coefficient of Efficiency (CofE), Coefficient of Determination (RSqr) and Index of Agreement (IoAd).

3 Neural network experiments

Two groups of experiments were performed: Group A comprised four experiments to predict $C_3(\text{std})$; Group B comprised four experiments to predict $C_4(\text{std})$. Further details are provided in Table 1. The eight experiments were designed to explore the potential power of different input drivers under two different modelling scenarios and to assess the specific impact of excluding the most influential non-linear input driver from the model development process. Abrahart and See (2007) reported the differential impact of reducing and conflating selected input drivers. Modelling with and without the b parameter was, as expected, observed to be important in terms of modelling non-linear relationships. This parameter accounts for non-uniform distributions. The power of individual solutions to provide non-linear outputs without the benefit of an important non-linear input was therefore being tested. Each experiment in this revised paper investigates the individual loss of a single input variable; modelling

Table 2. MLIN parameter coefficients for the eight reported experiments.

Group A	Exp. A1	Exp. A2	Exp. A3	Exp. A4
Intercept	-0.1992	0.0565	-0.0637	-0.1812
$C_1(\text{std})$	0.7041	–	0.6965	0.7041
$C_2(\text{std})$	0.2686	0.2573	–	0.2670
$b(\text{std})$	0.0345	0.0355	0.0219	–
Group B	Exp. B1	Exp. B2	Exp. B3	Exp. B4
Intercept	-0.1834	0.0034	0.1744	-0.1040
$C_1(\text{std})$	0.5143	–	0.4943	0.5146
$C_2(\text{std})$	0.7089	0.7007	–	0.7021
$b(\text{std})$	0.1524	0.1530	0.1190	–

with the two remaining inputs permits a comprehensive set of paired input combinations to be tested. No conflation of input variables is performed. This methodological improvement avoids the subjective process of selecting particular inputs to be removed or conflated and opens up several opportunities: using a full set of inputs will offer incontrovertible evidence on the power of NN modelling capabilities; using a partial set of inputs will permit the relative influences of lost drivers to be established; the NN outputs can in each case be compared and contrasted to a set of linear modelling counterparts.

TNNS (Trajan Neural Network Simulator²) was used to develop and implement the eight reported solutions. NN feedforward models were trained using “backpropagation with momentum”; each run produced a “Back Propagation Neural Network” (BPNN). Each processing unit was connected to all processing units in the adjacent layers and a full set of initial connections was maintained throughout. No transfer function is applied in the input layer and unaltered inputs are passed as direct outputs to units in the next layer. Each processing unit in the hidden layer and output layer contained a logistic transfer function, i.e. sigmoid curve. Models incorporating sigmoid transfer functions can support improved generalisation and superior learning characteristics. The resultant models can also offer higher accuracies but this is not always the case (e.g. Kim et al., 2001). The use of this function is widespread and in most cases it is the “default option”. Thus, for most modelling operations, it provides an ideal starting point. Moreover, if other options are to be considered, it should be used as a standard against which the alternatives can be compared.

Each weighted connection and processing unit bias was assigned an initial random setting in the range of ± 1 . Training material was presented in random order and the training programme was stopped at 10 000 epochs. Training parameters were set for automatic adjustment over the period; the learning rate was set to decrease from 0.8 to 0.2; the momentum

¹<http://www.hydrotest.org.uk>

²<http://www.trajan-software.demon.co.uk/>

Table 3. HydroTest evaluation statistics for Group A.

	Exp.	Equal Treatment		Favour High Flows		Comparison with Norms			
		MAE	MARE	RMSE	MSRE	CofE	RSqr	IoAd	
BPNN	Training	A1	0.0038	0.2181	0.0052	5.8353	0.9992	0.9992	0.9998
		A2	0.1279	6.6003	0.1723	3669.8383	0.1370	0.1826	0.4756
		A3	0.0696	0.6805	0.0979	18.2474	0.7215	0.7281	0.9088
		A4	0.0105	0.3252	0.0156	5.3870	0.9929	0.9929	0.9982
	Testing	A1	0.0037	0.1608	0.0049	2.7567	0.9992	0.9992	0.9998
		A2	0.1228	5.5770	0.1651	2245.1113	0.1404	0.1745	0.4796
		A3	0.0677	0.6208	0.0964	12.7876	0.7071	0.7111	0.9050
		A4	0.0104	0.2798	0.0156	3.1716	0.9923	0.9923	0.9981
MLIN	Training	A1	0.0505	4.3081	0.0619	3118.3660	0.8888	0.8888	0.9697
		A2	0.1313	7.9846	0.1689	4844.5527	0.1708	0.1708	0.5334
		A3	0.0777	2.3705	0.1009	425.0863	0.7042	0.7042	0.9067
		A4	0.0506	4.3005	0.0627	3507.8157	0.8859	0.8859	0.9689
	Testing	A1	0.0499	4.4901	0.0608	4283.5170	0.8834	0.8846	0.9693
		A2	0.1264	6.7539	0.1615	2837.9453	0.1776	0.1785	0.5474
		A3	0.0759	2.3315	0.0978	520.2014	0.6986	0.6994	0.9074
		A4	0.0500	4.4191	0.0614	4493.1153	0.8810	0.8821	0.9685

factor to decrease from 0.4 to 0.1. This use of fixed stopping conditions does not prevent overfitting. However, overfitting was not considered to be a danger as a result of (i) the smooth nature of the mathematical relationship that was being replicated; (ii) the comprehensive nature of the numerical datasets upon which the models were developed and tested; and (iii) the achievement of similar performance statistics on split sample datasets. Error plots of the development process also indicated that the model was not overfitted.

Each BPNN model contained six hidden units positioned in one hidden layer. Abrahart and See (2007) in their original discussion paper reported findings related to a much larger architectural configuration, i.e. two hidden layers that had twelve hidden units in each layer. The full extent to which such reported outputs differ is not addressed in the present paper. However, where comparisons can be made, the latest results depict a similar set of patterns and substantiate the earlier reported arguments. Abrahart and See (2007) made no attempt to develop (i) an optimal or (ii) a minimal architectural configuration. This paper adopts a similar standpoint. Instead, it was considered desirable to have a sufficient number of potential parameters at the start of the learning exercise that could, if required, be used to support the modelling process. Too few a number would put constraints on the capacity of the model to capture the full relationship. The earlier model was one that sought to maintain and not restrict important hydrological relationships. It is accepted that complex models might produce superior output scores but the real purpose of optimisation is not to develop an optimal model. It is to improve the quality of the solution. The need to search for a parsimonious solution equates to a different form of optimisation and one in which important properties, such as

robustness, might be lost; indeed, a parsimonious solution is a solution with as few parameters as possible for a given quality of model, but the nature of what does or does not constitute an important set of operational qualities remains a fundamental water management issue.

Traditional least squares multiple linear regression (MLIN) models were also fitted to the eight training datasets and each model was thereafter applied to its relevant test dataset. This use of identical predictors and predictands facilitated an unequivocal comparison to be performed against linear modelling solutions. The parameter coefficients for each linear modelling counterpart are provided in Table 2.

4 Results

HydroTest BPNN and MLIN performance statistics for the eight reported experiments are provided in Tables 3 and 4. Figures 4 and 5 provide Group A scatterplots of actual against predicted values for Experiments A1-4; Figs. 6 and 7 provide companion scatterplots for Group B. Training and testing output scatterplots are provided for each experiment. BPNN outputs relate to the trained product that was created at end of the learning and development process, i.e. after 10 000 epochs. The superior performance of the non-linear BPNN solutions, in contrast to the poor relative performance of their linear MLIN counterparts, is apparent in cases where a clear relationship exists. Several experiments in the current paper correspond to modelling operations reported in our earlier "discussion paper" (Abrahart and See, 2007). Each NN model in the previous set of experiments was developed on a much larger architectural setup; such earlier advantages

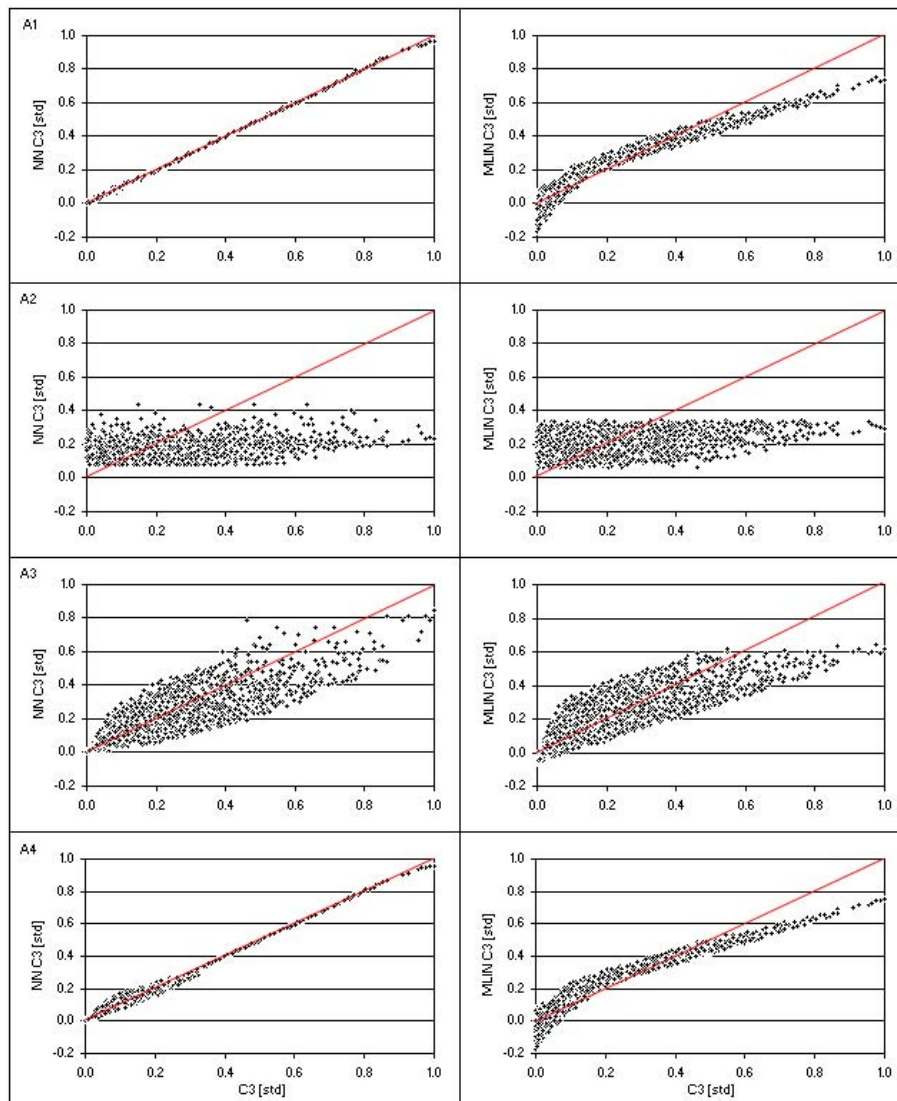


Fig. 4. Training dataset output scatterplots for Experiments A1-4.

produced predicted outputs that corresponded to the expected output values in a tighter manner. The corresponding models developed on different configurations nevertheless created a similar set of overall results that contained matching non-linear approximations to the hydrological processes concerned.

4.1 Group A

Figures 4 and 5 and the statistics in Table 3 confirm that the two split sample datasets produced similar outputs in each experiment which is indicative of unbiased solutions and minimal potential overfitting. Experiments A1 and A4 and their associated findings equate to procedures and results obtained during the first and second experiments in our preceding “discussion paper” on the use of different input drivers for model emulation purposes (Abrahart and See, 2007).

Experiment A1 comprised a direct emulation process. No input drivers were missing. The resultant scatterplots depict strong agreement between expected and predicted runoff values. The two latest scatterplots thus demonstrate that simple as well as complex neural solutions are able to emulate this particular hydrological model and perform non-linear processing operations in an efficacious manner. MLIN counterpart scatterplots reveal clear patterns of error: a curvilinear profile has resulted that is “twisted” around the “line of perfect agreement”. The higher and lower level values are underpredicted whilst the central region values are overpredicted. Moreover, the lower output values exhibit the greatest levels of spread, and this level of spread is reduced in a progressive manner as output values are increased. The nature of such problems indicates that the linear solution has failed to capture important non-linearities. This visual interpretation

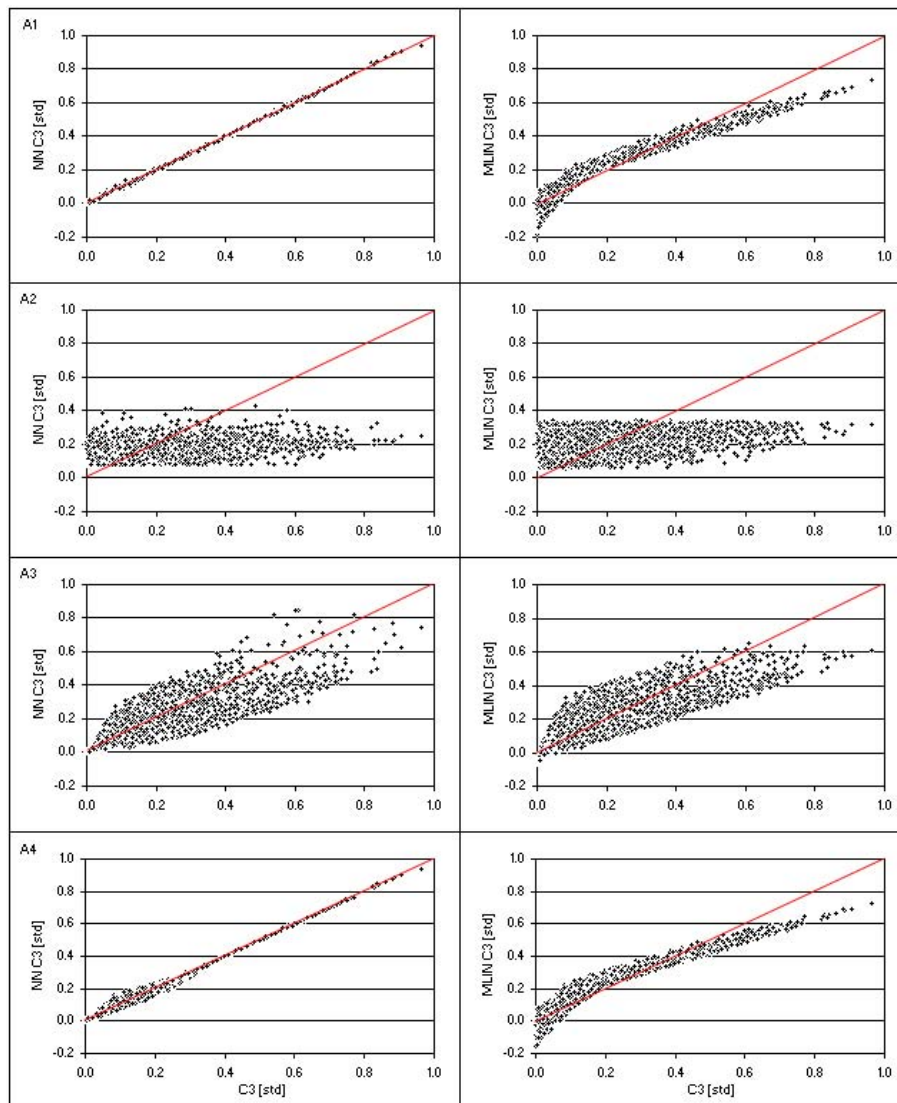


Fig. 5. Testing dataset output scatterplots for Experiments A1-4.

of events is supported in an assessment of the evaluation statistics. MLIN, in comparison to BPNN, exhibits poorer overall modelling capabilities on most reported evaluation metrics. MLIN also produced erroneous negative outputs.

Experiment A2 involved dropping the precipitation input driver $C_1(\text{std})$. This input was the main driver in a deterministic process without which the original model would be illogical and nonsensical. The resultant scatterplots depict a horizontal band that tapers at higher magnitudes; such plots in combination with the output evaluation statistics confirm that no direct relationship exists between the two remaining input drivers and the output predictand. This experiment has also confirmed that neither of the two approaches is able to detect or express a spurious relationship, i.e. no bogus solutions were developed. Of interest is the fact that neither model fitted the lowest outputs; MLIN also resulted in a sharp near linear upper margin.

Experiment A3 involved dropping the soil moisture condition input driver $C_2(\text{std})$. This input was a major volumetric control; the revised model still retains some physical sense but will operate in a diminished manner. The resultant scatterplots depict a modest relationship between the input drivers and output predictand. NN output provides a scattering of points around a central trend. The largest amount of dispersal is in the central region. This tails off in a progressive manner towards each end of the permitted output, but with a marked scattering of upper outliers. Higher values are nevertheless underpredicted reflecting issues related to the missing soil water component. MLIN output also provides a scattering of points and underpredicted higher values. The greatest amount of scattering is in the central region and tails off towards each end of the scatterplot. There is, however, an overall difference; the higher values are more underpredicted and instead of revealing equal convexities above and

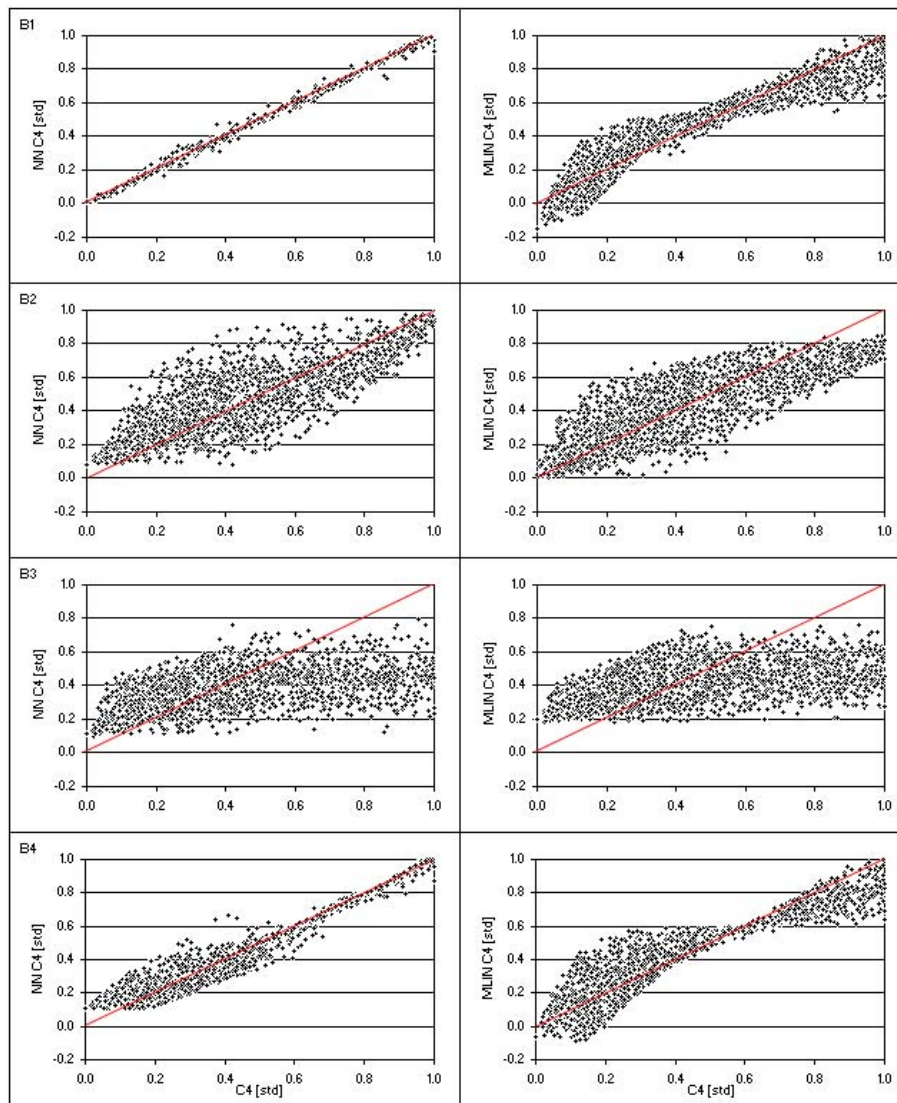


Fig. 6. Training dataset output scatterplots for Experiments B1-4.

beneath some central trend, there is a sharp near linear lower margin to the scattering of points. MLIN also produced erroneous negative outputs. Experiment A4 involved dropping the curve fitting exponent input driver $b(\text{std})$. This input provided a meaningful non-linear parameter that represents an important control for lower magnitude events; the revised model will nevertheless continue to produce a rational output. The resultant scatterplots depict strong agreement between expected and predicted runoff values. BPNN outputs exhibit a clear trend. Model outputs are superior to Experiment 3 but somewhat less impressive in comparison to Experiment 1 that used a full set of input drivers. The extent of the impairment is however restricted to a modest scattering of points in the lower sections of the scatterplot. This omission of a principal non-linear component created a substantial decrease in the level and content of information that was avail-

able for the construction of a model but the end result is nevertheless considered to offer a reasonable non-linear approximation to the original mathematics. MLIN scatterplots reveal a similar sequence of errors to that found in Experiment 1. This “no change situation” would suggest that the curve fitting exponent input driver is not being put to good use in the linear model for Experiment 1. Table 2 confirms the result; negligible weightings are assigned to this variable in each linear model. Figure 2 provides the target surface outputs for Experiment 4. Figure 8 provides the BPNN predicted output equivalent and confirms the non-linear nature of the neural approximation that was developed in response to this particular modelling scenario. The neural solution was also required to implement a certain degree of “simplification” or “averaging” in regions of conflict. The construction process is thus observed to have (i) produced a modelled surface that

Table 4. HydroTest evaluation statistics for Group B.

	Exp.	Equal Treatment		Favour High Flows		Comparison with Norms			
		MAE	MARE	RMSE	MSRE	CofE	RSqr	IoAd	
BPNN	Training	B1	0.0083	0.0365	0.0119	0.0419	0.9978	0.9979	0.9995
		B2	0.1058	0.4289	0.1354	1.9835	0.7179	0.7285	0.9155
		B3	0.1819	0.6768	0.2279	3.9597	0.2003	0.2146	0.5723
		B4	0.0417	0.2795	0.0583	3.1354	0.9477	0.9477	0.9864
	Testing	B1	0.0082	0.0356	0.0111	0.0051	0.9981	0.9982	0.9995
		B2	0.1049	0.4313	0.1330	0.5831	0.7284	0.7400	0.9188
		B3	0.1790	0.6730	0.2281	1.3223	0.2013	0.2121	0.5666
		B4	0.0410	0.2720	0.0566	0.5564	0.9508	0.9508	0.9872
MLIN	Training	B1	0.0671	0.3245	0.0857	5.4655	0.8869	0.8869	0.9692
		B2	0.1165	0.4112	0.1433	0.4970	0.6840	0.6840	0.8993
		B3	0.1879	0.8007	0.2271	10.9126	0.2060	0.2060	0.5711
		B4	0.0725	0.3444	0.0963	1.4138	0.8572	0.8572	0.9603
	Testing	B1	0.0681	0.3180	0.0863	0.6065	0.8856	0.8857	0.9688
		B2	0.1148	0.4182	0.1415	0.5144	0.6927	0.6928	0.9021
		B3	0.1856	0.7908	0.2269	2.4023	0.2095	0.2098	0.5673
		B4	0.0724	0.3356	0.0957	0.3583	0.8593	0.8594	0.9605

exhibits a recognisable trend; (ii) contains two sorts of non-linear modelling in terms of changing slope and pivoting; (iii) maintained important upper and lower region relationships; and (iv) resolved potential conflicts in the lower regions of the output surface.

4.2 Group B

Similar to Group A, Figs. 6 and 7 and the statistics in Table 4 confirm that the two split sample datasets produced similar outputs with no obvious signs of overfitting. Experiment B4 and its associated findings equate to procedures and results obtained during the final experiment in our preceding “discussion paper” on the use of different input drivers for model emulation purposes (Abrahart and See, 2007). The differences between the reported findings for Experiments A1-4 and B1-4 can be related to the fact that the nature of the non-linear modelling exercise has changed.

Experiment B1 used an identical set of inputs to Experiment A1 and a calculated runoff coefficient output $C_4(\text{std})$. This modified output equates to a different standardisation of the original runoff output and is dependent upon the physical characteristics of a watershed. The resultant scatterplots depict strong agreement between expected and predicted runoff coefficient outputs. The level of fit, however, was not as tight as that produced under “direct emulation” conditions in Experiment A1. The two latest scatterplots nevertheless demonstrate that neural solutions can emulate this traditional hydrological relationship and perform more challenging non-linear processing operations in an efficacious manner. MLIN counterpart scatterplots reveal clear patterns of error; but the nature of such errors is somewhat different to that found in Ex-

periment A1. The output in both cases nevertheless equates to a band that is “pivoted” around a central point on the “line of perfect agreement”. The distribution of values spreads out from this point in both directions such that the maximum spread of values occurs in the uppermost and lowermost corners. Higher level values tend to be underpredicted. Lower level values straddle the line of perfect agreement. MLIN, in comparison to BPNN, exhibits poorer overall modelling capabilities on most reported evaluation metrics. MLIN also produced erroneous negative outputs.

Experiment B2 (pursuant to A2) involved dropping the precipitation input driver $C_1(\text{std})$; but on this occasion a meaningful relationship is identified for the runoff coefficient in direct contrast to the poor result achieved for discharge under A2. This input is an important discriminating factor; it is required to resolve non-linear ambiguities related to the soil moisture status and curve fitting exponent drivers $C_2(\text{std})$ and $b(\text{std})$. Models that possess no rainfall input driver are instead required to discover a far broader relationship developed on physical catchment conditions. The simplified model that results will fail to distinguish unique output responses in the central ranges; potential conflicts and inconsistencies related to different non-linear effects will be less pronounced at either end of the output extent. The resultant scatterplots depict a modest relationship between the input drivers and output predictand. The upper values tend to be underpredicted. The lower values are overpredicted and no values occur in the lowest 10%. The greatest amount of scattering is in the central region and tails off towards each end of the scatterplot. MLIN output also provides a similar scattering of points. The greatest amount of scattering is in the central region and tails off towards each end of the

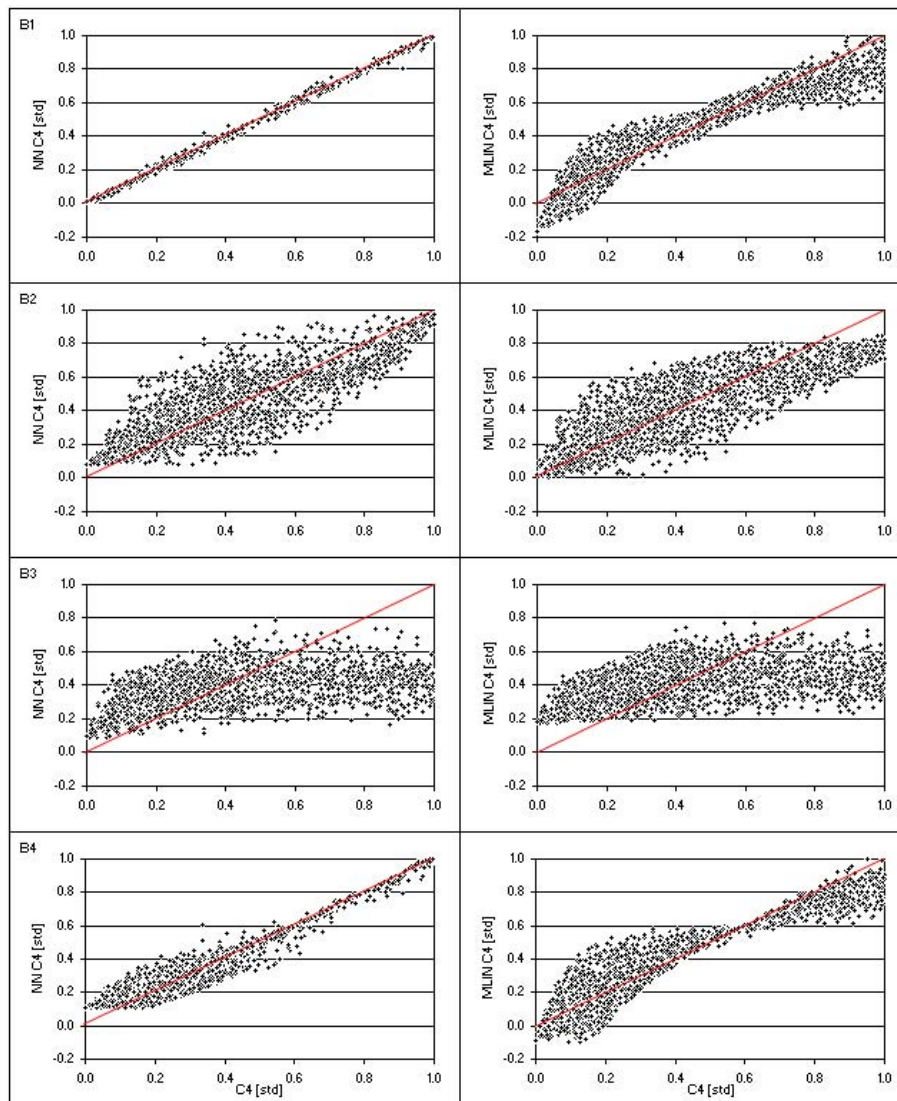


Fig. 7. Testing dataset output scatterplots for Experiments B1-4.

scatterplot. The highest values exhibit a marked underprediction; the lowest values, in contrast, exhibit no erroneous cut-off point and continue downwards to the zero level. There is, however, an overall difference; the higher values exhibit greater underprediction and the level of central spread is less.

Experiment B3 (pursuant to A3) involved dropping the soil moisture condition input driver C_2 (std); but on this occasion no meaningful relationship is identified for the runoff coefficient in direct contrast to the modest result achieved for discharge under A3. The state of the catchment in terms of antecedent conditions is a major physical factor in catchment response; drier or wetter conditions represent a fundamental control over the proportion of rainfall that is converted into runoff and without this input driver the model would be expected to produce inadequate results. The resultant scatterplots depict a curved “spreading out of responses” that

evolves into a horizontal band at higher magnitudes; such plots in combination with the output evaluation statistics confirm that no direct relationship exists between the two remaining input drivers and the output predictand. MLIN output also provides a similar scattering of points with under-predicted higher values and over-predicted lower values. The upshot of this failed experiment is to support the conclusions for A2; neither model is able to detect or express a spurious relationship and such outcomes could be important in developing a level of trust. Of interest is the fact that neither model fitted the lowest outputs; plus no model had a sharp near-linear upper or lower margin.

Experiment B4 (pursuant to A4) involved dropping the curve fitting exponent input driver b (std). The missing parameter is a measure of spatial inequalities in response; so removing it would be expected to produce some measure of

disorder in sensitive regions of the output space. The resultant scatterplots depict a reasonable level of agreement between expected and predicted runoff coefficient outputs. The level of fit, however, was not as tight as that produced under “direct emulation” conditions in Experiment A4. Higher magnitudes depict a stronger relationship, lower magnitudes present a broader spread of estimated outputs, and a clear cut-off point exists below which the neural solution does not produce output predictions. This particular situation represents a different form of generalisation that occurs under conditions of insufficient or conflicting observations in the dataset. It can be equated to “pit-filling” in the lower regions of the target output. Indeed, in different regions of the solution space, a crude generalisation is the best that can be done under such circumstances and serves to confirm that when there is a firm relationship to be modelled, the neural solution will extract it, whereas in other cases it will attempt to fit a broad higher level approximation. Model outputs are superior to Experiment B2, but somewhat less impressive in comparison to Experiment B1, that used a full set of input drivers. Most of this impairment is however restricted to a modest scattering of points in the lower sections of the scatterplot. This omission of a principal non-linear component created a substantial decrease in the level and content of information that was available for the construction of a model but the end result is nevertheless considered to offer a reasonable non-linear approximation to the original numbers. MLIN scatterplots reveal a similar sequence of errors to that found in Experiment B1. This “no change situation” once more suggests that the curve fitting exponent input driver is not being put to good use c.f. Experiment A4. Table 2 confirms matters with small weights being assigned to this variable in each linear model. The target surface outputs for Experiment B4 in Fig. 3 and its BPNN predicted output equivalent in Fig. 9 provide similar findings to that described for Experiment A4.

5 Discussion

NN can be evaluated and compared on a number of different factors. The numerical and graphical assessment procedures that are reported in this paper confirm that neural solutions can be used to model different groups of outputs related to conceptual model input and output datasets. Models developed on a full set of inputs produced excellent results. Singh et al. (2007) recommend that for operational purposes solutions should also be found to be consistent with the underlying physical processes. Their input parameters were varied over the full range and the output response verified so as to match up with the one that might perhaps be expected from the established physics of the underlying process(es). No such input variations were performed in the current paper; models were instead developed on different paired input combinations and assessment of changes in output related to a missing physical input performed. No useful runoff model

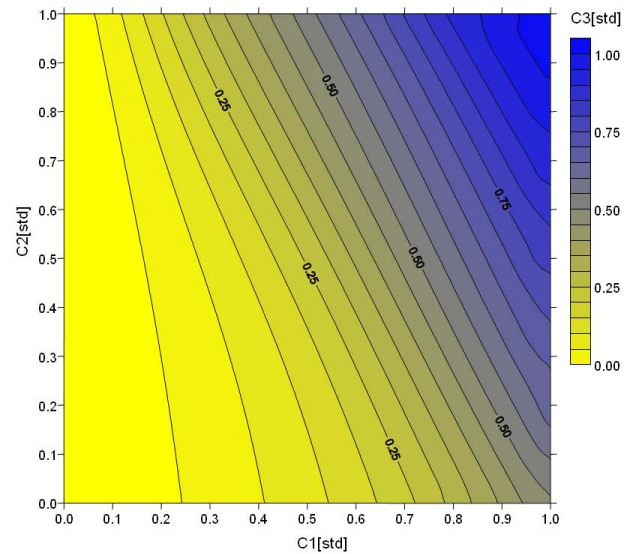


Fig. 8. BPNN predicted output surface for Experiment A4.

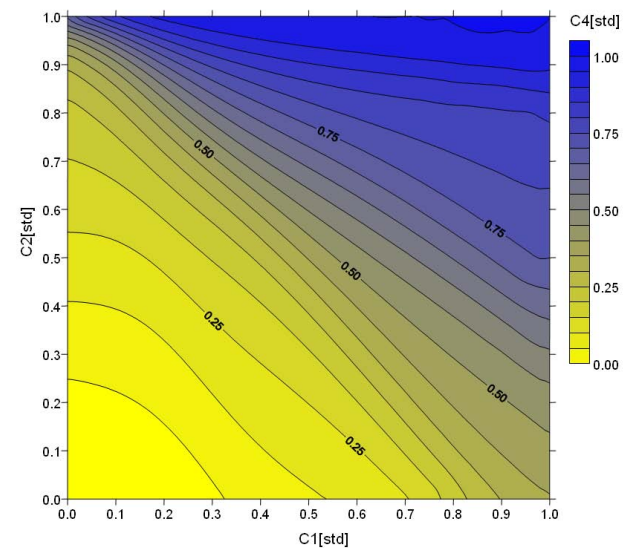


Fig. 9. BPNN predicted output surface for Experiment B4.

could be produced without an effective precipitation input that could be converted into a discharge output; no useful runoff coefficient model could be produced without a soil moisture condition input that reflected antecedent catchment conditions. The b parameter is an important spatial control at lower levels of runoff and will be more influential in terms of reflecting catchment conditions as opposed to direct discharge outputs. BPNN and MLIN discrepancies for the different groups and experiments reflected the neglected non-linear components. Experiments 1 and 4 have marked non-linear modelling requirements; Experiments A3 and B2 have partial non-linear modelling requirements; Experiments A2 and B3 have no real non-linear modelling requirements.

The exact relationship and overall intent surrounding the use of linear and non-linear hydrological modelling methodologies should be clear. The controlled experiments that are reported in this paper have established that a NN will produce an appropriate non-linear solution if it is presented with an appropriate non-linear situation to model. If the problem is linear, or near-linear, it is axiomatic that WLTF models and trained NN models will produce a similar set of results; the tools are performing as expected. The anticipated “superior performance” of a NN in relation to a WLTF will, as a result, appear to be limited. It is also possible to obtain better generalisation using a simple linear model, in the case of a function that contains mild non-linearities, if the datasets are too small, or contain too much noise, since such factors will prevent the NN from accurately estimating the non-linearities (Sarle, 1997). The two earlier critical studies can therefore be interpreted in a different light. Linear models should be used as standards against which neural solutions are tested so as to indicate the degree to which the presented relationship that needs to be modelled is linear and therefore requires the application of a linear modelling solution. Moreover, if required, following the adoption of a linear modelling solution, it is also the case that non-linear tools could thereafter be used for the identification of neglected non-linearities (Curry and Morgan, 2003). The question of which tool would be more appropriate in a linear or near-linear modelling situation is not a matter for scientific contest; it is a practical issue that equates to picking the “right tool for the right job”.

Modelling involves a number of assumptions but neither the non-linear nature of hydrological processes, nor the importance of modelling from an operational point of view, in which observed measurements are required for decision making purposes, is being questioned in this paper. The two earlier critical studies considered the nature of the relationship that existed between NN and WLTF models developed on hydrological datasets. The eight controlled experiments that are reported in this paper permit the previous studies to be viewed in context. The earlier reported interpretations and conclusions would appear to have been based on the following underlying premises:

1. The rainfall-runoff relationship is a recognised non-linear catchment process;
2. Measurements of observed rainfall and runoff can be used to develop a non-linear catchment response dataset;
3. The measured datasets and the manner in which such datasets are used have encapsulated the non-linear catchment response in a suitable format for subsequent identification and extraction using machine learning algorithms.

This list does not however contain an explicit method of testing for the presence of non-linear relationships in each

dataset. The simplest test would be to develop a linear model on the selected dataset, as a measure of the extent to which a linear or near-linear relationship exists, and thereafter select the most appropriate tool for subsequent modelling operations.

The exact manner in which a specific problem has been formulated is also important. The two earlier critical studies did not consider the extent to which the required solution called for the production of a simple model, with limited non-linear modelling capabilities, and no real need for the incorporation of complex dynamics. The solutions incorporated the last observed record at their point of forecast and as such the modelling operation might simply amount to calculating the change in discharge, which in most cases will be a near-linear operation, as opposed to something more physical. Further particulars on the dominant effect of including the last observed record at the point of forecast and reported findings related to countering or suppressing such factors using a constraint-based method can be found in Abrahart et al. (2007). It is also axiomatic in such cases that the use of smaller forecasting horizons will lead towards the development of linear or near-linear solutions. Han et al. (2007) reported modelling experiments performed over different forecasting horizons which provides further insight into this question. Indeed, as the forecasting horizon was increased, the dominant effect of the last recorded input was reduced such that the requirement for producing a non-linear solution is more apparent. If a marked non-linear relationship exists it should be obvious that a non-linear solution will be required to model it – as demonstrated in this paper. However, if a theoretical non-linear relationship appears to have been captured in an acceptable manner using a linear or near-linear model, then the exact reason(s) for this unexpected result should be questioned. It is, moreover, insufficient to assert that such findings can be attributed to potential failings, either in a specific dataset, or to the tool that has been applied to model it, without first having undertaken detailed investigations that are able to confirm or disprove such matters in the manner of “hypothesis testing”.

Hydrological modelling requires consistent measures of merit and trust. Hillel (1986: p.42) advocated that hydrological modelling solutions should be: “parsimonious” – each model should contain a minimum number of parameters that can be measured in the field; “modest” – the scope and purpose to which a specific model can be applied must not be overstated; “accurate” – the correctness of the forecast or prediction need not be better than the correctness of the input measurements; and “testable” – the limits within which the model outputs are valid can be defined. This paper and its predecessors have focused on one aspect of merit and trust: the production of more accurate outputs. However, other qualities and issues are also important with respect to practical operational implementations, and mechanistic properties such as “robustness” and “graceful degradation” will not in all cases have an optimal relationship with model output

accuracies and so must be treated as independent properties that impart a set of constraints. To provide a robust solution each model must exhibit a constant or stable behaviour and be insensitive to potential uncertainties in the construction and parameterisation process, e.g. problems related to measurements that cannot be obtained with sufficient accuracies or are not constant over long(er) periods. To be reliable and trusted an operational model must also exhibit the properties of “graceful degradation”; a gradual and progressive reduction in overall performance such that the model continues to operate and function in a normal manner, but provides a reduced level of service, as opposed to taking incorrect actions or suffering a total collapse of processing activities. Environmental modelling investigations into the changing nature of NN outputs related to the provision of degraded inputs are reported for hydrological forecasting in Abrahart et al. (2001) and for sediment transfer in Abrahart and White (2001). For more detailed discussion on the requirements constraint issue the interested reader is referred to Alippi (2002).

It is important in closing to highlight several issues that could hinder or prevent a direct comparison of the reported findings with the two earlier critical studies, not on fundamental questions about modelling non-linear relationships, but in terms of the different processes that are being represented.

- Major differences exist between the diverse types of non-linear hydrological process involved in: (i) a model that converts rainfall into storm runoff without considering time distribution factors or other essential components, e.g. infiltration, evapotranspiration, sub-surface flow; and (ii) the earlier reported studies sought to model discharge hydrographs in a time domain (with higher lead times in some cases). However, the purpose of the current paper is not to compare different models or the modelling of different processes, but to confirm that a NN can model simple non-linear hydrological relationships. Two case studies that had no temporal component were used to provide a return to “first principles” and such investigations offer an indispensable starting point for subsequent studies. The next step in this proof of concept testing and reporting exercise would be to inspect one or more time series forecasting models in a similar manner and thus add further findings to the debate.
- The two earlier critical studies had different transfer functions in their output units: Gaume and Gosset (2003) “linear”; Han et al. (2007) “positive linear”. There are two obvious advantages related to the use of a linear transfer function in the output unit: (i) to restrict potential impacts and distortions associated with upper limit and lower limit saturation effects; and (ii) to address potential deficiencies and ceilings associated with undershoots or requirements to extrapolate beyond the

range of the training dataset. The potential impact of using linear as opposed to non-linear transfer functions in the output unit is the subject of ongoing studies.

- The model outputs in the current paper were compared to traditional least squares multiple linear regression models developed on the same datasets. No direct comparison in such cases is intended between the earlier WLTF and later MLIN models; the two linear numerical approaches are nevertheless of a similar statistical type, providing modelling norms and benchmarking standards, irrespective of the fact that the individual mechanisms and processes involved were rather different.

6 Conclusions

The skill of computational mechanisms to model important non-linear processes is addressed in this paper. Two important issues have been raised. To what extent: (i) can a NN model perform non-linear hydrological modelling operations given a suitable challenge and an unproblematic set of information-rich observations, i.e. to help overcome issues related to poor content and noise; and (ii) are most reported NN hydrological modelling applications about the production of near-linear as opposed to non-linear solutions. This paper has shed some light on the first question based on a selective set of modest non-linear modelling operations.

It is possible to infer from the success of the reported experiments that NN hydrological modelling solutions, under appropriate conditions, can be used to produce reliable non-linear transformations, and to represent catchment processes at different levels of interest, i.e. generalisation can be performed on different scales of measurement. The selection of purposeful inputs can moreover lead to the development of dissimilar models that reflect different physical process constraints. More challenging situations and pertinent explorations related to development of improved processing methodologies are the subject of ongoing experiments, as are attempts to discover the limits of what can or cannot be achieved with a NN.

The power of a neural solution to provide acceptable modelling operations based on a reduced set of inputs and to explore alternative relationships has been demonstrated. The power to omit one or more problematic variables is of particular importance in the case of scarce or difficult to obtain datasets, and, in addition, has clear cost-benefit implications. The power to develop insights into the inner mechanisms of a conceptual model through input removal procedures offers numerous interesting possibilities. Multifaceted plots of difficult-to-observe process-based relationships could perhaps be used for the assessment of more complex models, or of their internal components.

Minimum effort was devoted to design and development issues. Each end product, nevertheless, was considered to be

acceptable but not necessarily optimal. For operational purposes such solutions might be sufficient. The act of building a NN model was also discovered to be a rather robust operation that required limited expert involvement. The solutions were quick to create and simple to test which makes them ideal tools for bootstrapping, sensitivity analysis and rapid prototyping implementations. Further emulation exercises of a similar nature are to be encouraged.

Acknowledgements. The valuable insights and useful suggestions of the guest editor and two anonymous reviewers on our initial discussion paper were both helpful and very much appreciated. Their input has resulted in the development of a stronger and more coherent structure. It has also promoted a clearer account of important issues.

Edited by: E. Toth

References

- Abrahart, R. J.: Neural networks and the problem of accumulated error: an embedded solution that offers new opportunities for modelling and testing, in: *Hydroinformatics'98: Proceedings Third International Conference on Hydroinformatics*, edited by: Babovic, V. and Larsen, C. L., Copenhagen, Denmark, 24–26 August 1998, A.A. Balkema Publishers, Rotterdam, The Netherlands, 2, 725–731, 1998.
- Abrahart, R. J., Heppenstall, A. J., and See, L. M.: Timing error correction procedure applied to neural network rainfall-runoff modelling, *Hydrol. Sci. J.*, 52(3), 414–431, 2007.
- Abrahart, R. J., Kneale, P. E. and See, L. M. (Eds.): *Neural Networks for Hydrological Modelling*, A. A. Balkema Publishers, Rotterdam, The Netherlands, 2004.
- Abrahart, R. J. and See, L. M.: Neural network emulation of a rainfall-runoff model, *Hydrol. Earth Syst. Sci. Discuss.*, 4, 287–327, 2007, <http://www.hydrol-earth-syst-sci-discuss.net/4/287/2007/>.
- Abrahart, R. J., See, L., and Kneale, P. E.: Investigating the role of saliency analysis with a neural network rainfall-runoff model, *Computers and Geosciences*, 27(8), 921–928, 2001.
- Abrahart, R. J. and White, S.: Modelling Sediment Transfer in Malawi: Comparing Backpropagation Neural Network Solutions Against a Multiple Linear Regression Benchmark Using Small Data Sets, *Phys. Chem. Earth (B)*, 26(1), 19–24, 2001.
- Alippi, C.: Selecting Accurate, Robust, and Minimal Feedforward Neural Networks, *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, 49(12), 1799–1810, 2002.
- Adams, R., Dunn, S. M., Lunn, R., Mackay, R., and O'Callaghan, J. R.: Assessing the Performance of the NELUP Hydrological Models for River Basin Planning, *J. Environ. Planning Manage.*, 38(1), 53–76, 1995.
- ASCE: Artificial Neural Networks in Hydrology. I. Preliminary Concepts, *J. Hydrol. Eng.*, 5(2), 115–123, 2000a.
- ASCE: Artificial Neural Networks in Hydrology. II. Hydrologic Applications, *J. Hydrol. Eng.*, 5(2), 124–137, 2000b.
- Bhattacharya, B., Price, R. K., and Solomatine, D. P.: Data-driven modelling in the context of sediment transport, *Phys. Chem. Earth*, 30(4–5), 297–302, 2005.
- Beven, K. J.: *Rainfall-Runoff Modelling: The Primer*, John Wiley & Sons Ltd., Chichester, England, 2001.
- Birikundavy, S., Labib, R., Trung, H. T., and Rousselle, J.: Performance of Neural Networks in Daily Streamflow Forecasting, *J. Hydrol. Eng.*, 7(5), 392–398, 2002.
- Campolo, M., Soldati, A., and Andreussi, P.: Artificial neural network approach to flood forecasting in the River Arno, *Hydrol. Sci. J.*, 48(3), 381–398, 2003.
- Chen, X., Chen, Y. D., and Xu, C.: A distributed monthly hydrological model for integrating spatial variations of basin topography and rainfall, *Hydrol. Processes*, 21(2), 242–252, 2007.
- Cheng, C. T., Ou, C. P., and Chau, K. W.: Combining a fuzzy optimal model with a genetic algorithm to solve multi-objective rainfall-runoff model calibration, *J. Hydrol.*, 268(1–4), 72–86, 2002.
- Cheng, C.-T., Zhao, M.-Y., Chau, K. W., and Wu, X.-Y.: Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure, *J. Hydrol.*, 316(1–4), 129–140, 2006.
- Curry, B. and Morgan, P. H.: Neural networks, linear functions and neglected non-linearity, *Computational Management Science*, 1, 15–29, 2003.
- Dawson, C. W., Abrahart, R. J., and See, L. M.: HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts, *Environ. Modell. Software*, 22(4), 1034–1052, 2007.
- Dawson, C. W. and Wilby, R. L.: Hydrological modelling using artificial neural networks, *Progress in Physical Geography*, 25(1), 80–108, 2001.
- Dümenil, L. and Todini, E.: A rainfall-runoff scheme for use in the Hamburg climate model, in: *Advances in Theoretical Hydrology – A tribute to James Dooge*, edited by: O'Kane, J. P., Elsevier, Amsterdam, The Netherlands, 129–157, 1992.
- Gaume, E. and Gosset, R.: Over-parameterisation, a major obstacle to the use of artificial neural networks in hydrology?, *Hydrol. Earth Syst. Sci.*, 7(5), 693–706, 2003.
- Govindaraju, R. S. and Rao, A. R. (Eds.): *Artificial Neural Networks in Hydrology*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
- Han, D., Kwong, T., and Li, S.: Uncertainties in real-time flood forecasting with neural networks, *Hydrol. Processes*, 21(2), 223–228, 2007.
- Hettiarachchi, P., Hall, M. J., and Minns, A. W.: The extrapolation of artificial neural networks for the modelling of rainfall-runoff relationships, *J. Hydroinformatics*, 7(4), 291–296, 2005.
- Hillel, D.: Modeling in soil physics: A critical review. In: *Future Developments in Soil Science Research*, Soil Society of America, Madison, Wisconsin, USA, 35–42, 1986.
- Hu, C., Guo, S., Xiong, L., and Dingzhi Peng, D.: A modified Xinanjiang model and its application in northern China, *Nordic Hydrology*, 36(2), 175–192, 2005.
- Hsu, C. Y., Lee, H. Y., and Lin, Y. L.: Estimation of sediment yields in a watershed area using ANN method, in: *River Basin Management II (Transactions of the Wessex Institute, Ecology and the Environment Vol. 60, Progress in Water Resources Vol. 7)*, edited by: Brebbia, C. A., WIT Press, Southampton, UK, 461–470, 2003.
- Huang, W., Xu, B., and Chad-Hilton, A.: Forecasting flows in Apalachicola River using neural networks, *Hydrol. Processes*, 18(13), 2545–2564, 2004.

- Jayawardena, A. W. and Zhou, M. C.: A modified spatial soil moisture storage capacity distribution curve for the Xinanjiang Model, *J. Hydrol.*, 227(1–4), 93–113, 2000.
- Kim, B., Choi, W., and Kim, H.: Using neural networks with a linear output neuron to model plasmaetch processes, *Proceedings IEEE International Symposium on Industrial Electronics*, 1, 441–445, 2001, doi:10.1109/ISIE.2001.931830
- Kisi, O.: Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation, *Hydrol. Sci. J.*, 49(3), 1025–1040, 2004.
- Kisi, O.: Suspended sediment estimation using neuro-fuzzy and neural network approaches, *Hydrol. Sci. J.*, 50(4), 683–696, 2005.
- Lin, C.A., Wen, L., Lu, G., Wu, Z., Zhang, J., Yang, Y., Zhu, Y., and Tong, L.: Atmospheric-hydrological modeling of severe precipitation and floods in the Huaihe River Basin, China, *J. Hydrol.*, 330(1–2), 249–259, 2006.
- Liong, S. Y. and Chan, W. T.: Runoff Volume Estimates with Neural Networks, in: *Neural Networks and Combinatorial Optimization in Civil and Structural Engineering*, edited by: Topping, B. H. V. and Khan, A. I., *Proceedings Third International Conference on the Application of Artificial Intelligence to Civil and Structural Engineering*, Edinburgh, UK, 17–19 August 1993, Civil-Comp Press, Stirling, UK, 67–70, 1993.
- Loke, E., Warnnaars, E. A., Jacobsen, P., Nelen, F., and Almeida, M. D.: Artificial neural networks as a tool in urban storm drainage, *Water Sci. Technol.*, 36(8–9), 101–109, 1997.
- Maier, H. R. and Dandy, G. C.: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, *Environ. Modell. Software*, 15(1), 101–124, 2000.
- Moore, R. J.: The probability-distributed principle and runoff production at point and basin scales, *Hydrol. Sci. J.*, 30(2), 273–297, 1985.
- Moore, R. J.: The PDM rainfall-runoff model, *Hydrol. Earth Syst. Sci.*, 11(1), 483–499, 2007.
- Mulvaney, T. J.: On the use of self-registering rain and flood gauges in making observations of the relations of rainfall and flood discharges in a given catchment, *Proceedings of the Institution of Civil Engineers in Ireland*, 4, 19–31, 1851.
- Nagy, H. M., Watanabe, K., and Hirano, M.: Prediction of Sediment Load Concentration in Rivers using Artificial Neural Network Model, *J. Hydraul. Eng.*, 128(3), 588–595, 2002.
- Rao, Z. and Alvarruiz, F.: Use of an artificial neural network to capture the domain knowledge of a conventional hydraulic simulation model, *J. Hydroinformatics*, 9(1), 15–24, 2007.
- Riad, S., Mania, J., Bouchaou, L., and Najjar, Y.: Predicting catchment flow in a semi-arid region via an artificial neural network technique, *Hydrol. Processes*, 18(13), 2387–2393, 2004.
- Rogers, L. L. and Dowla, F. U.: Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling, *Water Resour. Res.*, 30(2), 457–481, 1994.
- Sarle, W. S. (Ed.): *Neural Network FAQ, Part 3 of 7: Generalisation* [periodic posting to the Usenet newsgroup comp.ai.neural-nets]. (<ftp://ftp.sas.com/pub/neural/FAQ.html>) 1997 [Last accessed: 8 November 2006]
- Senbeta, D. A., Shamseldin, A. Y., and O'Connor, K. M.: Modification of the probability-distributed interacting storage capacity model, *J. Hydrol.*, 224(3), 149–168, 1999.
- Senthil Kumar, A. R., Sudheer, K. P., Jain, S. K., and Agarwal, P. K.: Rainfall-runoff modelling using artificial neural networks: comparison of network types, *Hydrol. Processes*, 19(3), 1277–1291, 2005.
- Shamseldin, A. Y.: Artificial neural network model for flood forecasting in a developing country, *J. Hydroinformatics*, in press, 2007.
- Singh, A. K., Deo, M. C., and Kumar, V. S.: Prediction of littoral drift with artificial neural networks, *Hydrol. Earth Syst. Sci. Discuss.*, 4(4), 2497–2519, 2007.
- Solomatine, D. P. and Avila Torres, L. A.: Neural network approximation of a hydrodynamics model in optimizing reservoir operation, in: *Hydroinformatics 96: Proceedings Second International Conference on Hydroinformatics*, edited by: Müller, A., Zurich, Switzerland, 9–13 September 1996, A.A. Balkema Publishers, Rotterdam, The Netherlands, 1, 201–206, 1996.
- Su, B., Kazama, S., Lu, M., and Sawamoto, M.: Development of a distributed hydrological model and its application to soil erosion simulation in a forested catchment during storm period, *Hydrol. Processes*, 17(14), 2811–2823, 2003.
- Todini, E.: The ARNO rainfall-runoff model, *J. Hydrol.*, 175(1–4), 339–382, 1996.
- Todini, E.: The ARNO model, in: *Mathematical Models of Large Watershed Hydrology*, edited by: Singh, V. P. and Frevert, D. K., Water Resources Publications LLC, Highlands Ranch, CO, USA, 687–716, 2002.
- Toth, E., Brath, A., and Montanari, A.: Comparison of short-term rainfall prediction models for real-time flood forecasting, *J. Hydrol.*, 239(1–4), 132–147, 2000.
- Varoonchotikul, P.: *Flood forecasting using artificial neural networks*, A.A. Balkema Publishers, Rotterdam, The Netherlands, 2003.
- Wilby, R. L., Abrahart, R. J., and Dawson, C. W.: Detection of conceptual model rainfall-runoff processes inside an artificial neural network, *Hydrol. Sci. J.*, 48(2), 163–181, 2003.
- Wood, E. F., Lettenmaier, D. P., and Zartarian, V. G.: A land-surface hydrology parameterisation with subgrid variability for general circulation models, *J. Geophys. Res.*, 97(D3), 2717–2728, 1992.
- Yitian, L. and Gu, R. R.: Modeling Flow and Sediment Transport in a River System Using an Artificial Neural Network, *Environ. Manage.*, 31(1), 122–134, 2003.
- Zhang, Q., Smith, D. W., and Baxter, C. W.: Introduction, *J. Environ. Eng. Sci.*, 3, S1, iii–iv, 2004.
- Zhao, R.-J.: The Xinanjiang model applied in China, *J. Hydrol.*, 135(1–4), 371–381, 1992.
- Zhao, R.-J. and Liu, X.-R.: The Xinanjiang Model, in: *Computer Models of Watershed Hydrology*, edited by: Singh, V. P., Water Resources Publications, Fort Collins, USA, 215–232, 1995.
- Zhao, R.-J., Zhang, Y.-L., Fang, L.-R., Liu, X.-R., and Zhang, Q.-S.: The Xinanjiang model, In: *Hydrological Forecasting: Proceedings of the Symposium on the Application of Recent Developments in Hydrological Forecasting to the Operation of Water Resource Systems*, Oxford, April 1980, International Association of Hydrological Sciences Press, Wallingford, UK, IAHS Publication No. 129, 351–356, 1980.