Application of self-organising maps and multi-layer perceptron-artificial neural networks for streamflow and water level forecasting in data-poor catchments: the case of the Lower Shire floodplain, Malawi

F. D. Mwale, A. J. Adeloye and R. Rustum

ABSTRACT

With a paradigm shift from flood protection to flood risk management that emphasises learning to live with the floods, flood forecasting and warning have received more attention in recent times. However, for developing countries, the lack of adequate and good quality data to support traditional hydrological modelling for flood forecasting and warning poses a big challenge. While there has been increasing attention worldwide towards data-driven models, their application in developing countries has been limited. A combination of self-organising maps (SOM) and multi-layer perceptron artificial neural networks (MLP-ANN) is applied to the Lower Shire floodplain of Malawi for flow and water level forecasting. The SOM was used to extract features from the raw data, which then formed the basis of infilling the gap-riddled data to provide more complete and much longer records that enhanced predictions. The MLP-ANN was used for the forecasting, using alternately the SOM features and the infilled raw data. Very satisfactory forecasts were obtained with the latter for up to 2-day lead time, with both the Nash–Sutcliffe index and coefficient of correlation being in excess of 0.9. When SOM features were used, however, the lead time for very satisfactory forecasts increased to 5 days.

Key words | data-poor catchments, forecasting, multi-layer perceptron artificial neural networks, self-organising maps, streamflow, water level

F. D. Mwale (corresponding author)
A. J. Adeloye
School of Built Environment,
Heriot Watt University,
Riccarton,
Edinburgh EH14 4AS,
UK
E-mail: fdm4@hw.ac.uk

R. Rustum
School of Built Environment,
Heriot Watt University,
Dubai Campus,
United Arab Emirates

INTRODUCTION

The increasing severity of floods and the increase in economic damage caused by floods have resulted in efforts being directed at finding ways for more effective flood management. Consequently, there has been a change to the way floods are dealt with that combines the traditional structural measures with the new notion that people must understand and manage floods by learning to live with them (Kundzewicz & Menzel 2005; Samuels et al. 2006; WMO/GWP 2009). This new thinking frontier is exemplified in several national policy frameworks and projects, amongst them: the EU Flood Directive (EU Directive 2007), the EU Interreg Rhine–Meuse Activities (IRMA-SPONGE project) and the

Associated Program on Flood Risk Management (WMO/GWP 2009). While both measures are important in addressing flood risk, however, the non-structural measures have been described as more environmentally, economically and socially sustainable (Kundzewicz & Menzel 2005; Van Ogtrop et al. 2005; Lumbroso et al. 2008). This is of particular relevance for developing countries that face numerous challenges including competing socio-economic developmental demands for the meagre resources available that may preclude the huge investments often needed for developing effective structural measures of flood control (Lumbroso et al. 2008; Plate 2009). The lack of capacity in

these countries to maintain and operate flood management infrastructures (Van Ogtrop *et al.* 2005) poses another challenge. Thus as rightly observed by Plate (2009) sometimes non-structural measures can be the only viable options available for developing countries. The practical implementation of non-structural measures is evidenced in Mozambique where an array of tools (source book on sustainable flood risk management strategies, posters, a manual and card games), aimed at raising awareness and preparedness, were developed and successfully piloted in three rural communities of Carre, Languene and Matidze (Lumbroso *et al.* 2008).

Flood forecasting and early warning is one non-structural measure that has gained prominence in flood risk management. For example, early warning is integral to the five key pillars of the Hyogo Framework for Action (HFA) 2005–2015 (UNISDR 2005), a global blueprint for disaster reduction. The HFA recognises early warning as a measure that is cost effective and sustainable in the protection of lives, property and livelihoods. Hydrological modelling forms an important aspect of flood forecasting and early warning, yet data availability, reliability and adequacy in most developing countries are a big challenge when attempting to apply traditional conceptual and physically based models. This paper reports on the use of artificial intelligence (AI) tools, particularly the self-organising map (SOM) and the multi-layer perceptron feed-forward artificial neural networks (MLP-ANN) for forecasting flow and water level in the Lower Shire River floodplain of Malawi. In the next section, a summary review of hydrological modelling with regard to flow forecasting is presented. This is followed by a presentation of the essential aspects of the two AI modelling approaches adopted for the current research: the SOM and MLP-ANN. Next, the methodology is presented, followed by results and discussion sections. The final section contains the conclusions of the study.

HYDROLOGICAL MODELS

Hydrological models for forecasting fall into three groups: blackbox, conceptual and physically based. Blackbox models transform inputs to outputs using e.g. regression without reference to the internal physical processes controlling the transformation. Physically based models at the other extreme try to replicate the natural system using the basic mathematical representation of the flow at a point based on the principles of conservation of mass, momentum and energy. Conceptual models are those relying on a simple arrangement of a relatively small number of interlinked conceptual elements (the most common elements being storage elements), each representing a segment of the land phase of the hydrological cycle (Jain 1993).

According to O'Connor (2005) and Toth & Brath (2007), physically based models have found little application in flood forecasting because of their inherent complexity, long computational times, problems related to scale issues on the application of physical laws, extensive data demands and hence cost implications, all of which tend to limit them to impact studies in smaller catchments (Todini 1996; Moore & Bell 2001). Therefore, blackbox and conceptual models are the most common. Nonetheless, physically based models in flood forecasting include the TOPKAPI model which has been applied in the upper Xixian catchment in China (Liu *et al.* 2005) and the MIKE SHE model, applied in the Manoa–Palolo stream system in Hawaii (Sahoo *et al.* 2006). LISFLOOD, another physically based model, has been at the centre of the European Flood Alert System (Thielen *et al.* 2009).

Conceptual models in flood forecasting include the Nedbør-Afstrømnings-Model, which has been used as the rainfall-runoff model in the MIKE 11 model applied for operational flood forecasting system of the Jamuneswari river system of Bangladesh (Rahman *et al.* 2012); Soil Moisture Accounting and Routing, whose applicability was assessed in the Kilombero River basin of Tanzania for river flow forecasting (Yawson *et al.* 2005); and the Hydrologiska Byråns Vattenbalansavdelning (HBV) model which has been tested in the Sava River basin in Slovenia (Primožič *et al.* 2008). Simple blackbox models of the linear type, that is, simple linear model, linear perturbation model (LPM), linear varying gain factor model, stage regressions, auto regressive (AR), auto regressive moving average models (ARMA) and the ARMA model with an exogenous variable (ARMAX), have also found wide application. They have been applied for example to the

Kilombero River basin of Tanzania (Yawson *et al.* 2005) and the Kiewa and Ovens basins of the Murray Darling Basin, Australia for the operational software, Source (Dutta *et al.* 2012). In comparison to blackbox models, conceptual models tend to represent the hydrological processes effectively. However, their data requirements are high and on many occasions, these data are unavailable or expensive and time consuming to collect (Chau *et al.* 2005). Singh (2005) for example observed that in India, despite the availability of several conceptual models, they were hardly used due to data scarcity and problems of model calibration. According to the Ministry of Irrigation and Water Development (MoIWD) report in Malawi (MoIWD, undated), an attempt by the Malawi Government in 1998 to implement the HBV model for flood forecasting in the Lower Shire floodplain failed due to, among other factors, scarcity of data. In addition to high data demands, calibration of conceptual models is also not an easy task. Furthermore, some prior knowledge is required of physical processes taking place in the catchments. While the traditional linear blackbox models offer simplicity, they reduce the complex non-linear system to a linear system and may thus inadequately represent the system.

These challenges have resulted in increased attention towards data-driven models with ANN being the widely used data-driven technique in water resources management. ANNs have been used in water resources and environmental management for rainfall forecasting (Hung *et al.* 2009), rainfall-runoff modelling (Wu & Chau 2011), flow forecasting (Aquil *et al.* 2007; Wu *et al.* 2009), water level forecasting (Alvisi *et al.* 2006), reservoir parameters prediction (Adeloye & De Munari 2006), watershed water supply estimation (Iliadis & Maris 2007), groundwater level prediction (Nayak *et al.* 2006), sediment load estimation (Cigizoglu 2004), water quality parameter prediction (Baxter *et al.* 2001; Panda *et al.* 2004; Diamantopoulou *et al.* 2005; Rustum & Adeloye 2012), land use/land cover classification (Yuan *et al.* 2009), land use and land cover change prediction (Ito & Murata 2009), and land use/land cover change impact on water resources (Isik *et al.* 2013).

Despite the successful and increasing application of ANN for various studies, their application in data-poor African catchments remains quite sparse. Their application

is reported on small data sets in Bvumbwe, Mindawi I, Mindawo II and Mphezo areas of Malawi for the determination of sediment load transfer under different agricultural and land management practices (Abrahart & White 2001). The small data sets posed a challenge since ANNs characterise systems response from data presented (Toth & Brath 2007) through training and small data sets will lead to inadequate learning due to reduced degrees of freedom available for parameter estimation (Khamis *et al.* 2006). Therefore controlled random noise was added to the training set to increase data size and thus reduce the uncertainties in the parameter estimation. Mazvimavi (2003) used MLP-ANN to predict flow characteristics from catchment characteristics in selected catchments of Zimbabwe. Mazvimavi estimated mean annual flow, baseflow index, annual hydrograph, mean monthly flows and flow duration curves using mean annual precipitation, lithology, slope, mean annual potential evaporation, land cover and drainage density as inputs. The performance of ANN over linear regression was mixed. Due to data availability problems, Mazvimazi's work is limited to flow stations with a minimum of 10 years of data. Ilunga & Stephenson (2005) used ANN for infilling mean monthly seasonal flows on artificially created gaps in the Orange River drainage system of South Africa. ANN performance was high. Similarly, MLP-ANN yielded very satisfactory results when applied for real time flood forecasting to the Omo River in southern Ethiopia (Ayalew *et al.* 2007).

More applications of ANN are reported in the Blue Nile River basin of Ethiopia where Saliha *et al.* (2011) used a SOM for the identification of homogeneous regions in the basin for the subsequent prediction of daily stream flow in ungauged basins. Recently, Adeloye & Rustum (2012) used SOM to augment and extend the flow series at gauging stations in the Osun basin of Nigeria and then used this to derive flow data at an ungauged station downstream.

It is important, therefore, to recognise that African catchments pose a unique challenge in the sense that most are either ungauged or if gauged, have poor quality data characterised by gaps and short durations (Mazvimavi 2003; Saliha *et al.* 2011; Adeloye & Rustum 2012). Thus while the competence of ANN has been widely assessed in flow prediction, the Lower Shire floodplain of Malawi like

other African catchments presents a unique situation in terms of data quality and availability.

ARTIFICIAL NEURAL NETWORKS

Multi-layer perceptron artificial neural networks

Artificial neural networks are computational techniques that mimic the neural biological system of the human brain. There are various types of ANN but the most widely used ANN in water resources and more exclusively so in predictions and forecasting is the feedforward MLP-ANN (Maier & Dandy 2000). MLP-ANN consists of nodes arranged in parallel and connected to other nodes through weight vectors. They consist of an input layer, one or more hidden layers and an output layer. The use of one hidden layer in hydrological studies involving ANN is not atypical and has been considered complex enough to simulate non-linear patterns exhibited in hydrological problems (Demuth *et al.* 2009). Knowledge is acquired through training and the backpropagation training algorithm tends to be the most widely used in water resources problems. It is a computationally efficient mechanism for weight adjustments in feedforward networks (Maier & Dandy 2000, Maier *et al.* 2010). In the forward pass, the information at a particular node is a weighted sum of signals from the previous nodes plus a bias term calculated through a predetermined activation function. In the backward pass, weights are adjusted according to the error value between the network outputs and the targets. Non-linear activation functions such as the log or tan sigmoid transfer functions are normally used in the hidden layer as they are known to capture both linear and non-linear function (Demuth *et al.* 2009). In contrast, the output layer may contain either linear or non-linear activation functions. Training continues until a predetermined level of error is reached or when no further changes are observed in the error. The number of nodes in the input layer and output layers are problem-dependent being equal to the independent and dependent variables, respectively. The number of neurons in the hidden layer is subject to several propositions but the trial-and-error approach remains widely used (Maier *et al.* 2010). A comprehensive description of ANN can be found in several excellent publications

including Haykin (1999), Maier & Dandy (2000) and Dawson & Wilby (2001).

Self-organising maps

SOM are a competitive, unsupervised form of artificial neural networks pioneered by Kohonen *et al.* (1996). The SOM consists of two layers: the multi-dimensional input layer and the competitive or output layer as illustrated in Figure 1. Both of these layers are fully interconnected. The output layer consists of M neurons arranged in a two-dimensional grid of nodes. Each node or neuron i ($i = 1, 2, \ldots, M$) is represented by an n-dimensional weight or reference or code vector $W_{i=}[w_{i1}, \ldots, w_{in}]$, where n is the dimension of each input vector, that is, the maximum number of variables in the input vector. In other words, each neuron in the output layer of the SOM contains exactly the same set of variables contained in the input vectors and thus, unlike the MLP-ANNs, variables in the SOM are not partitioned into input or output variables. This is why the training of the SOM is often referred to as unsupervised in that, unlike the MLP-ANN, the SOM does not have any output exemplar which it attempts to match by 'supervision'.

A fuller description of the training algorithm for the SOM is provided by Adeloye & Rustum (2012) but, essentially, this involves identifying the output neuron that best matches the presented input vector and updating the identified output neuron vector so that it is closer to the input
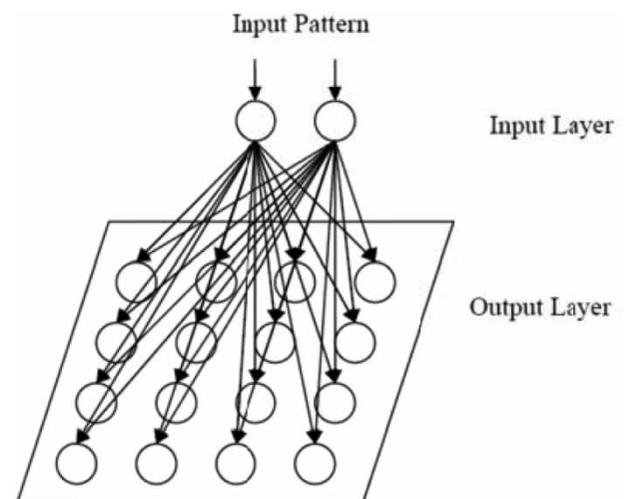

Figure 1 | The architecture of SOM.

vector. The identification is based on the Euclidian distance (Equation (1)) with the identified output neuron being the one with the minimum distance from the input vector and is known as the best matching unit (BMU) (Liukkonen *et al.* 2013)

$$D_i = \sqrt{\sum_{j=1}^{n} m_j (x_j - w_{ij})^2}; i = 1, 2, \ldots, M \qquad (1)$$

where D_i is the Euclidian distance between the input vector and the code vector i; x_j is the jth element of the current input vector; w_{ij} is the jth element of the code vector i; n is the dimension of the input vector; and m_j is the so called 'mask' which is used to include in ($m_j = 1$), or exclude from ($m_j = 0$), the calculation of the Euclidian distance, the contribution of a given element x_j of the input vector. This is very useful where the input vector contains missing elements because all that needs to be done is to set the mask (m_j) to zero for such elements.

At the end of the training, each of the BMUs will represent a weighted average of all the input vectors assimilated in them. These weighted averages are known as the features, which because they are complete and noise-free can form the basis of predicting the missing values of any input vector later presented to the trained map. In this way, SOM can be viewed as a tool for data compression and data filling.

METHODOLOGY

Study area and data

The Lower Shire floodplain lies at the southern tip of the Shire River basin of Malawi and cuts across two administrative districts of Chikwawa and Nsanje (Figure 2). The basin is drained by the Shire River, the only outlet of the lake, and its many tributaries.

The Lower Shire floodplain is the most vulnerable area to flooding and thus the most affected in Malawi. Its biophysical vulnerability arises from heavy rainfall in the upper and middle catchments of the Shire River and from the Ruo catchment on its main tributary. Although the floodplain rainfall is only around 600 mm annually, the upper

and middle Shire River catchments receive annual rainfall of around 900 mm and the Ruo catchment has average annual rainfall in excess of 2,000 mm (Shela *et al.* 2008). According to Shela *et al.*, flow measured at Chiromo on the Shire River averages 480 m³/s annually and 54 m³/s at Sinoya gauge station on the Ruo River. However, a list of major flooding events documented by Shela *et al.* (2008) shows that values as high as 1,480 and 5,400 m³/s, respectively, have been recorded. In terms of relief, the flood plain is low-lying at 107 masl at Chikwawa District Centre to about 61 masl at Nsanje (SVADD 1975). Flooding is further exacerbated by land degradation in the upper catchments. The population is predominantly rural and has the highest levels of poverty in the country (National Statistical Office 2012). Livelihoods are intricately linked to natural resources, especially water, which is why most of the settlements are along river courses. Due to resource constraints, in common with most developing countries, the catchment has very limited investment in structural measures for flood control. Thus the importance of flood forecasting cannot be over-emphasised. Previous attempts to forecast floods with the HBV conceptual model failed due to lack of data (MoIWD undated).

Whilst flooding occurs along many parts of the Shire River, the interest of this study is flooding that occurs around Chiromo and surrounding areas. This flooding is the most severe and has been largely attributed to the backflow effect at the confluence (Shela *et al.* 2008). The flows in the Ruo can be so high they create a reverse flow in the main river resulting in flooding around this area. Therefore, stream flow and water level are modelled for Chiromo gauging station using Chikwawa station upstream and Sinoya station, the most downstream on the Ruo. Chiromo gauge station is the most downstream station on the Shire River before the confluence with Ruo (Figure 2). Although the Shire River basin extends into Mozambique, flooding for the Lower Shire floodplain is a result of heavy rains in the upper and middle catchments of the Shire River below Lake Malawi and from the Ruo catchment. Therefore only the Shire River basin inside Malawi is considered.

Data used are daily rainfall collected by the Department of Climate Change and Meteorological Services, and daily stream flow and water level data collected by the MoIWD. Rainfall data are collected using manual rainfall gauges.
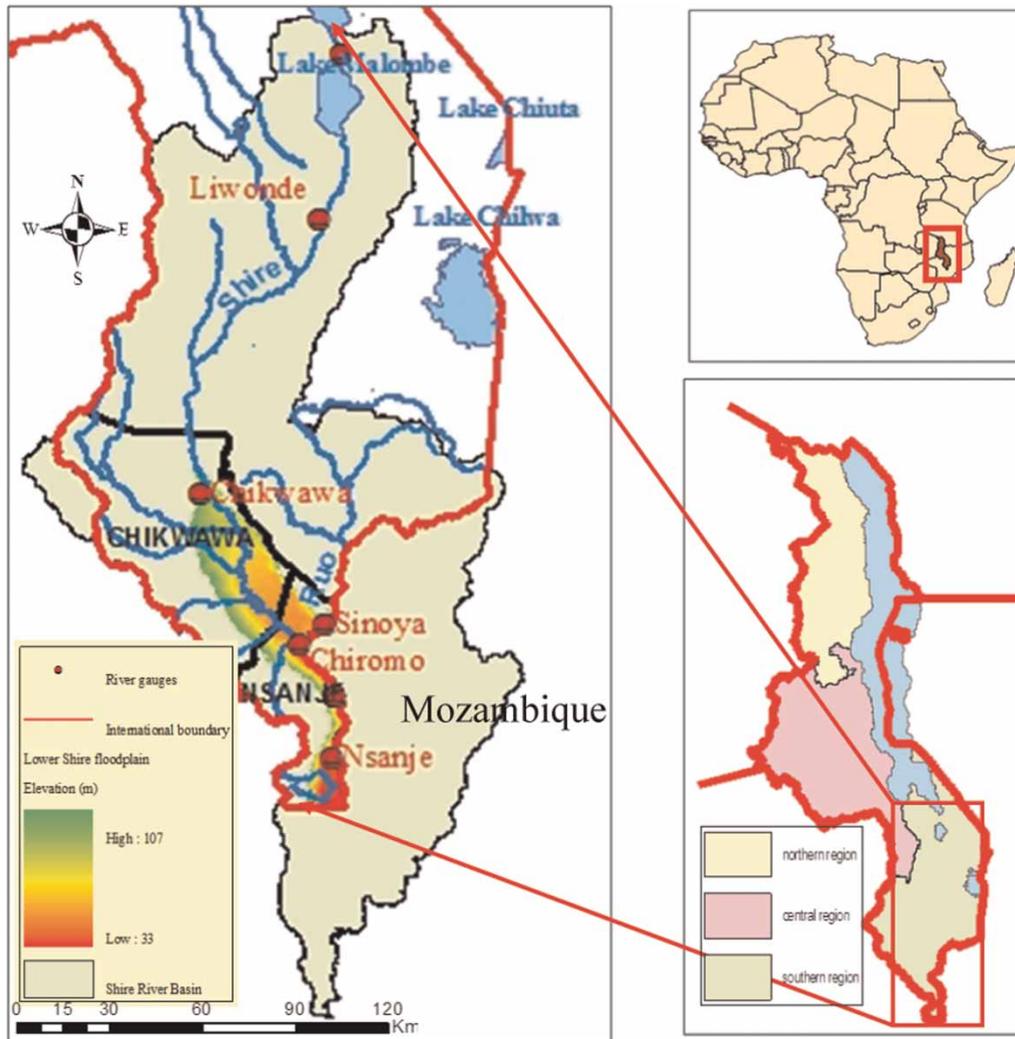
Figure 2 | Geographical position of Malawi, the Shire River basin and the floodplain.

Similarly, water levels are manually collected using staff gauges. The water levels are then converted to flow using rating curves but only when the former are available. Otherwise, a gap in water levels leads to a gap in flow data. The historical periods of data in the catchment are presented in Table 1, from which it is clear that the periods are non-uniform. Several of the records are missing and, although this is not evident in Table 1, some of the missing periods do overlap thus making the use of traditional infilling approaches such as regression impossible. Flow data at Sinoya, a potential station in the forecasting exercise, are the shortest. Thus the SOM was first used to infill and extend data for flow prediction.

A lack of both flow and water level data across all gauging stations exists between September 2008 and December 2008. Therefore although data on rainfall go as far as 2010, for the purpose of SOM application, the period January 1978–August 2008 with a substantial overlap was used for analysis. A hydrological year in Malawi starts in November and ends in October. Thus for the purpose of prediction with MLP-ANN, data from November 1978–August 2008 was used giving a total of 10,891 daily data records. Although only Chikwawa, Chiromo and Sinoya are the focus of the forecasting exercise, data infilling was extended to other gauge stations in the basin for the broader objective of flood risk assessment.

Table 1 | Record length of data used

Station	Record length	Proportion of missing data (%)	Rainfall cluster
Flow			
Mangochi	1956–2008	7.0	
Liwonde	1948–2008	2.8	
Chikwawa	1977–1998	11.7	
Chiromo	1953–1998	5.7	
Sinoya	1980–1990	23.0	
Water level			
Mangochi	1953–1996	0.3	
Liwonde	1980–2010	2.1	
Chikwawa	1980–2003	10.0	
Chiromo	1970–2009	14.4	
Sinoya	1962–2002	2.4	
Tengani	1970–2006	23.4	
Nsanje	1960–2003	41.5	
Rainfall			
Nsanje	1973–2009	8.4	
Makhanga	1953–2010	3.2	
Ngabu	1981–2010	0.0	Cluster 1
Chikwawa	1979–2010	10.7	
Nchalo	1981–2010	1.6	
Rainfall			
Neno	1980–2009	14.0	
Mwanza	1935–2006	17.1	
Mimosa	1958–2010	2.5	
Thyolo	1962–2010	0.2	Cluster 2
Bvumbwe	1953–2010	0.6	
Chileka	1961–2010	0.7	
Chichiri	1981–2010	0.0	
Makoka	1981–2010	0.7	
Rainfall			
Chingale	1952–2010	15.1	
Balaka	1981–2010	1.4	Cluster 3
Mangochi	1961–2010	0.02	

ANN application

Infilling with SOM

The SOM tool box developed at the Laboratory of Information and Computer Science at Helsinki University

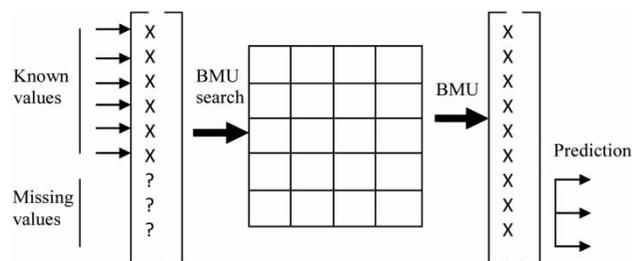**Figure 3** | Prediction of missing components of the input vector using SOM. *Source*: Rustum & Adeloye (2007).

of Technology (http://www.cis.hut.fi/projects/somtoolbox) was used in MATLAB environment by Mathworks Inc. Data are trained in four separate groups: flow and water level together, and three clusters of rainfall. The need to have separate models in the way outlined was informed by the poor performance that resulted when all were trained together (see Mwale *et al.* 2012). Kalteth & Berndtsson (2007) also observed that SOM prediction accuracy increases with the homogeneity of the training set and the three rainfall clusters (1, 2 and 3) used in this work are the homogeneous regions of southern Malawi as identified by Ngongondo *et al.* (2011). Table 1 lists the rainfall stations and the cluster they belong to. As seen in Table 1, there are five, eight and three stations in clusters 1, 2 and 3, respectively.

Once the SOM has been trained, it is used for infilling as illustrated in Figure 3, which has been adapted from Rustum & Adeloye (2007). As evident in Figure 3, there can be more than one variable needing to be predicted in a single input vector. In Figure 3 for example, there are three variables of the input vector that need to be predicted. This ability to simultaneously predict multiple variables is what makes the SOM a much more versatile tool than classical regression. The multivariate prediction using the SOM thus proceeds as follows:

(i) Decide on the variables needing prediction in the input vector. These will be variables that are unavailable because they are missing (e.g. missing river stage and discharge) resulting in a depleted input vector. In the schematic of Figure 3, the input vector has three values missing, which are represented by '?'.

(ii) Determine the Euclidian distance, D, of the depleted vector from each of the nodes of the output layer of the trained SOM using Equation (1). In doing this,

the mask, m_j, of each of the unavailable values will be set to zero while m_j will be set to unity for all the other values in the input vector.

(iii) Examine all the Ds for the minimum and hence isolate the SOM's BMU for the depleted input vector. While the input vector in step (i) above is depleted, that is, has values missing, the BMU identified here is a node of a trained SOM and hence has the full complement of values.

(iv) Replace the missing values of the input vector by their corresponding values in the BMU identified in step (iii) above.

For the details of this infilling exercise in this catchment, refer to Mwale *et al.* (2012).

Predicting with MLP-ANN

Determining the right inputs to an ANN model is an important exercise as not all inputs are equally informative: some inputs may be correlated, others may be noisy or may actually have no relationship to the output being modelled (Bowden *et al.* 2005). In flow, water level or rainfall runoff modelling, normally model inputs tend to be flow, stage or rainfall and their corresponding lagged inputs. Therefore these were the variables used. It was postulated that water levels at Chiromo and their corresponding flows are correlated with levels and flow respectively at Chikwawa, a station upstream of Chiromo and, levels and flows at Sinoya on Ruo River. Catchment rainfall becomes another variable as flooding in this catchment arises from rainfall. Input rainfall is average catchment rainfall calculated from all stations in the three clusters using the Thiessen Polygon method.

The numbers of lagged inputs to include were determined through autocorrelation (acf), partial correlation (pacf) and cross-correlation (ccf) functions as suggested by Sudheer *et al.* (2002). The acf for Chiromo levels and flows (Figures 4(a) and (c)) shows dominant autoregressive processes in both. The pacf (Figures 4(b) and (d)) nonetheless revealed that only six lags of water levels are statistically significant at 95% confidence level. Similarly, only 5-day lags on flow series are statistically significant. Therefore, from

Chiromo, a total number of 11 lags were used. As for exogenous variables, that is, flows and levels at Chikwawa and Sinoya and catchment rainfall, their lagged inputs were identified through ccf with Chiromo station flow and water level series and the appropriate number of lags identified as those up to the maximum cross-correlation (Figure 5). This approach has been used by Solomantine & Dulal (2003), Teschl & Randeu (2006) and Wu & Chau (2011). Lags corresponding to maximum correlation have been equated to travel times to the point of interest in the catchment. Figure 5 reveals that the maximum cross-correlations for both Chikwawa and Sinoya with Chiromo levels (Figure 5(a)) and Chiromo flows (Figure 5(b)) occur at zero lag. Thus only the current day's flow and water level from both Chikwawa and Sinoya were incorporated. As for rainfall, the maximum cross-correlation with Chiromo flows (Figure 5(b)) corresponded to a maximum of 18 days suggesting a concentration time of about 18 days. This is in agreement with results from a test carried out by the MoIWD in 1981 (Figure 6) which suggest a travel time of 17 days between Liwonde, 87 km from the river source, and Chiromo. The long concentration time is attributed to the Elephant marsh between Chikwawa and Chiromo.

Nonetheless, flooding at Chiromo is caused by Ruo River and therefore the concentration time of Ruo catchment is more relevant than that from the headwater catchments. The average slope of the Ruo catchment, derived with GIS, is 9.64%. The river has a total length of 130 km. An estimate of concentration time of the Ruo catchment at the confluence with Shire River with the Kirpich equation (Kirpich 1940) yields a value of 6.9 h validating the general view that flooding from Ruo is flash flooding. Since data used is available at a daily time step, no lag was applied to the rainfall and the day's rainfall was thus used. Based on this analysis therefore, the total number of inputs to the ANN model for flow and level forecasting at Chiromo was determined as 18 (a day's flow at Chiromo and flow on each of previous 5 days: sub-total = 6; a day's water level at Chiromo and the water level on each of previous 6 days: sub-total = 7; a day's flow and water level at both Chikwawa and Sinoya: sub-total = 4; and a day's average catchment rainfall: sub-total = 1). The model had two outputs: the water level and flow at Chiromo. Thus
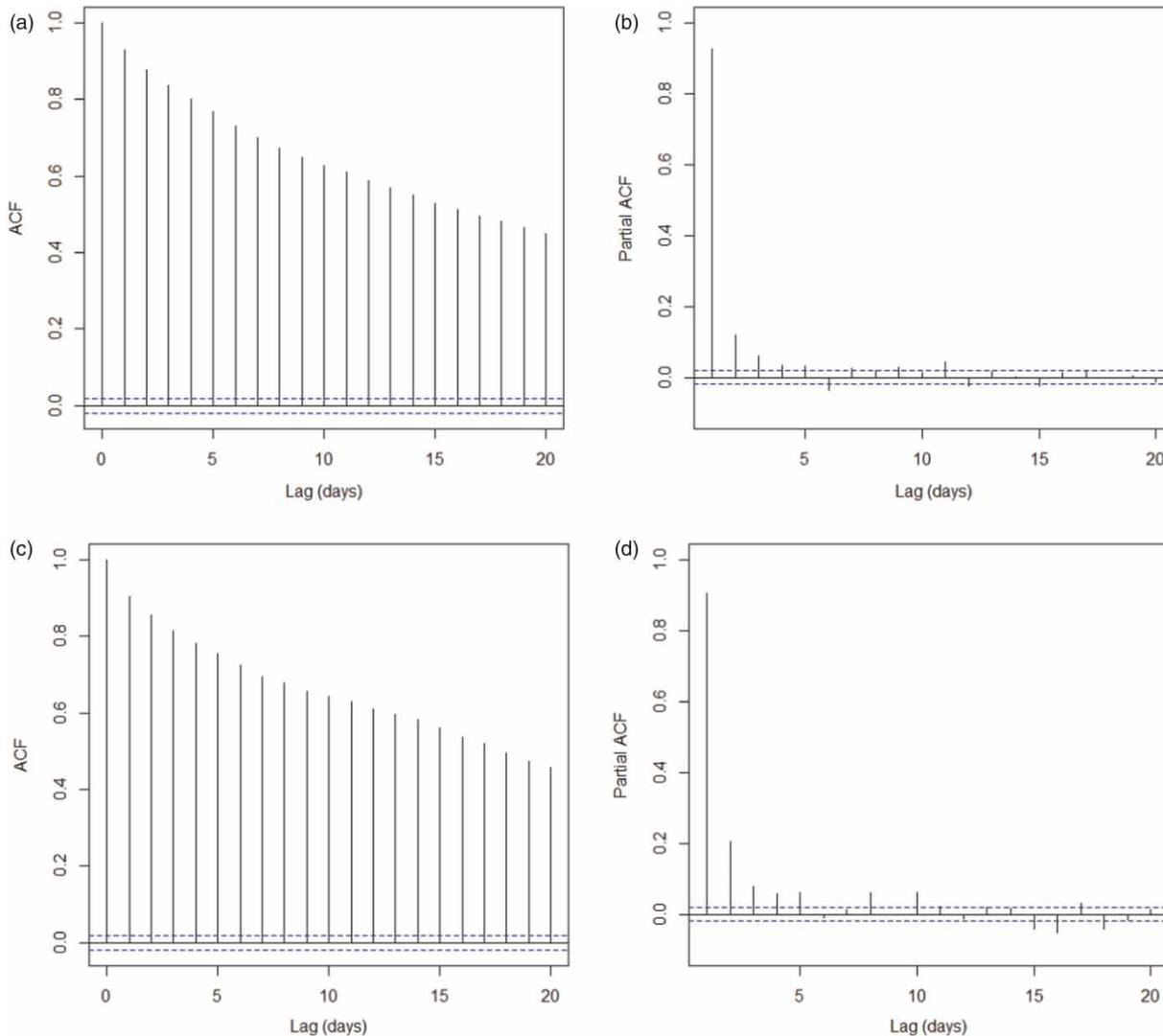
Figure 4 | ACF on (a) Chiromo water levels with the dashed lines representing the 95% confidence bound; and (c) Chiromo flows; PACF on (b) Chiromo water levels; and (d) Chiromo flows.

mathematically, the model is defined by the following equation.

Model 1:

$$(CRL_{t+1}, CRF_{t+1}) = f(CRL_t, CRL_{t-1}........CRL_{t-6}, CRF, \\ CRF_{t-1,}.......CRF_{t-5}, CKL_t, CKF_t, SL_t, SF_t, R_t) \quad (2)$$

where CR = Chiromo, CK = Chikwawa, S = Sinoya; L, F and R are water level, flow and rainfall, respectively. t is the point in time at forecasting and $t-1$, $t-2$, etc., are the lags. In search of a parsimonious model, further investigations were carried out. PACF on both flow and water level data at Chiromo (Figures 4(b) and (d)) show that

over 90% of the variance in both variables are explained by their 1-day lag, which is not unexpected. Therefore a second model (Model 2) with only one lag from flow and one lag from water level from the Chiromo series was investigated. Inputs from Chikwawa and Sinoya were excluded. Rainfall was retained as prediction accuracy has been found to be better when rainfall is included other than when flow is modelled solely on flow variable (Toth & Brath 2007; Wu & Chau 2011).

Model 2:

$$(CRL_{t+1}, CRF_{t+1}) = f(CRL_t, CRL_{t-1,}, CRF_t, CRF_{t-1}, R_t) \quad (3)$$
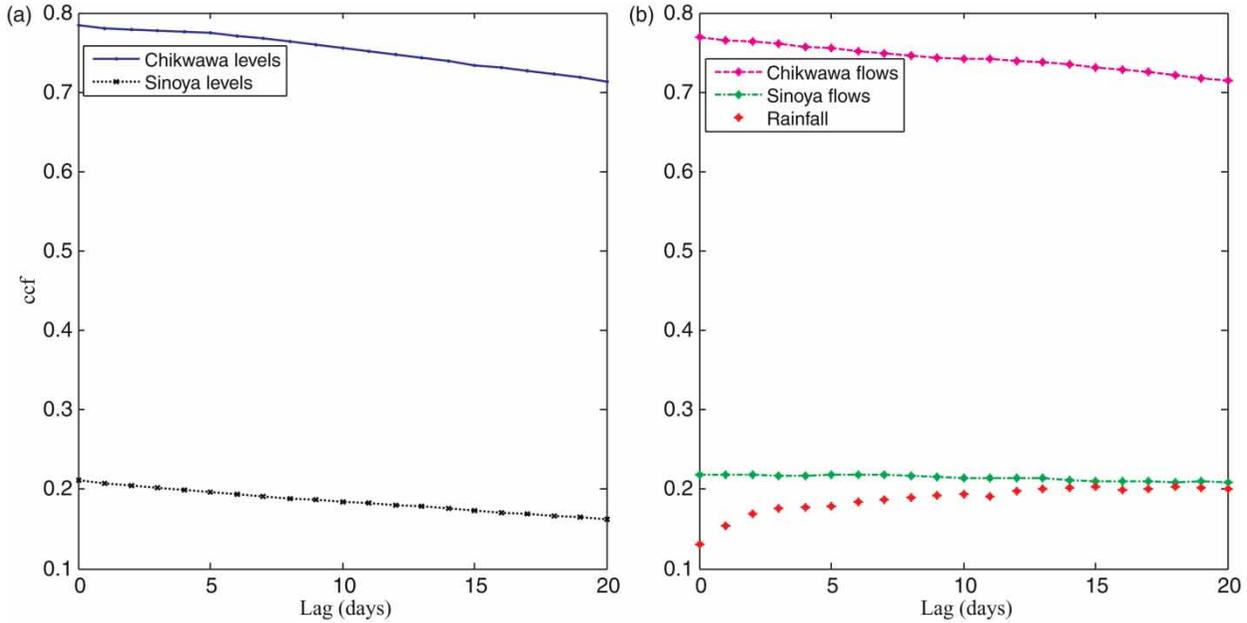
Figure 5 │ Cross correlations of (a) Chikwawa and Sinoya levels with Chiromo water levels and (b) Chikwawa and Sinoya flows, and rainfall with Chiromo flows.
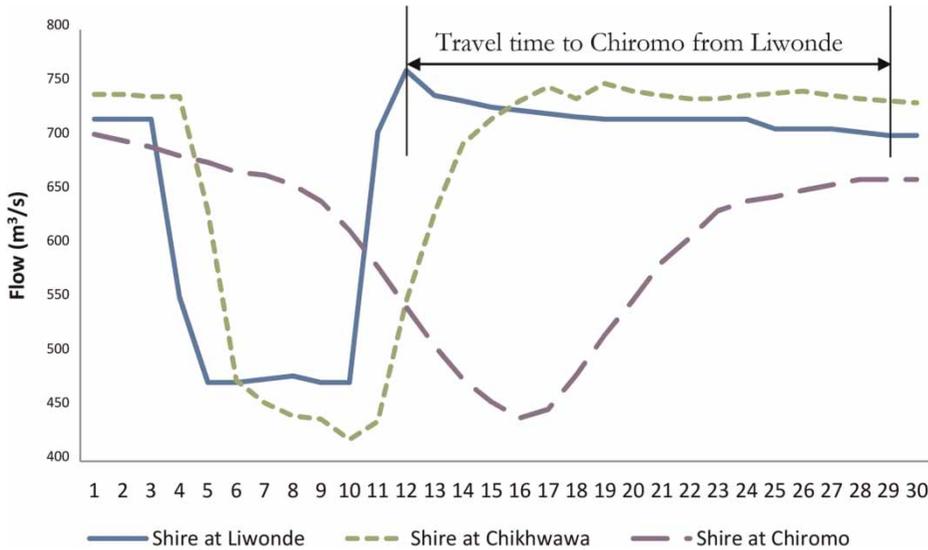

Figure 6 │ Hydrographs of Shire River flows at Liwonde, Chikhwawa and Chiromo following a test run at Liwonde barrage upstream in November 1981. *Source*: MoIWD.

On the same basis of parsimony, it was further investigated whether Chiromo levels and flows can be modelled with just 'today's' variables. This model is referred to as Model 3.

Model 3:

$$(CRL_{t+1}, CRF_{t+1}) = f(CRL_t, CRF_t, R_t) \qquad (4)$$

As justified earlier, the MLP feedforward neural network with three layers was used. The tan-sigmoid transfer function was used in the hidden layer and a linear transfer function to the output layer. A major criticism of ANN has been their failure to generalise beyond the calibration range and it has been suggested that the use of a linear function in the output layer circumvents this problem to a certain extent (Maier & Dandy 2000; Solomantine & Dulal 2003).

Data were partitioned in the ratio of 60, 20 and 20% for training, validation and testing, respectively. With a total number of 10,891 daily data records, 6,535 samples were used for training, 2,178 for validation and 2,178 for testing. To prevent over-fitting, the traditional early stopping criterion was used. In this technique, errors on both the training and validation sets are monitored during network training and training is stopped when the validation error starts increasing. Random sampling was applied to ensure representative data were used for training. The Levenberg–Marquardt backpropagation training algorithm was used to train the models. Studies have shown that the Levenberg–Marquardt algorithm performs better than the standard backpropagation algorithms. It is more accurate, faster and more reliable (Adeloye & De Munari 2006; Aquil *et al.* 2007; Okkan 2011).

Model performance was assessed using the Nash–Sutcliffe index, NS (Equation (5)), the correlation coefficient, R (Equation (6)) and the mean squared relative error, MSRE (Equation (7)). In these equations, n is the number of observations, $\hat{Q}_i$ is the predicted value, Q_i represents observed flow and $\overline{Q_i}$ denotes the mean of the observed values. To identify the best architecture for each of the three models, the number of neurons in the hidden layer was progressively increased from 2 to 15 at a time step of one hidden neuron and the best architecture in terms of the hidden neurons was picked based on best performance on test data. The NS is different from the R in that, unlike R, it quantifies the proportion of the observed variance that is explained by the model but in doing this, it does not assume *a priori* a linear relationship between the inputs and outputs. For a simple regression model, NS is the same as the square of R. Additionally, following recommendations against relying on a single measure of accuracy (Legates & McCabe 1999; Dawson & Wilby 2001; Jain & Sudheer 2008), MSRE was also used. The relative error measure was preferred to the absolute error because it removes any distortion that could result due to differences in units of output variables (Maier *et al.* 2010). Finally, time series plots and scatter plots have been produced to confirm the relative efficacies of the different models

$$NS = 1 - \frac{\sum_{i=1}^{n} \left(Q_i - \hat{Q}_i \right)^2}{\sum_{i=1}^{n} \left(Q_i - \overline{Q_i} \right)^2} \qquad (5)$$

$$R = \frac{n \sum Q_i \hat{Q}_i - \sum Q_i \sum \hat{Q}_i}{\sqrt{\left[n \sum Q_i^2 - \left(\sum Q_i \right)^2 \right] \left[n \sum \hat{Q}_i^2 - \left(\sum \hat{Q}_i \right)^2 \right]}} \qquad (6)$$

$$MSRE = \frac{\sum_{i=1}^{n} \dfrac{\left(Q_i - \hat{Q}_i \right)^2}{Q_i^2}}{n} \qquad (7)$$

RESULTS AND DISCUSSION

Ahead forecasting

Although not shown here for lack of space, it was observed that the MSRE, unlike NS and R, was sensitive to the number of hidden neurons. Consequently, the selection of best MLP-ANN structure in terms of hidden neurons was based on MSRE. The selection for the three models, based on average MSRE values on flow and water are shown in Table 2. Figure 7 shows corresponding scatter plots and

Table 2 | Performance of the three models in the testing phase

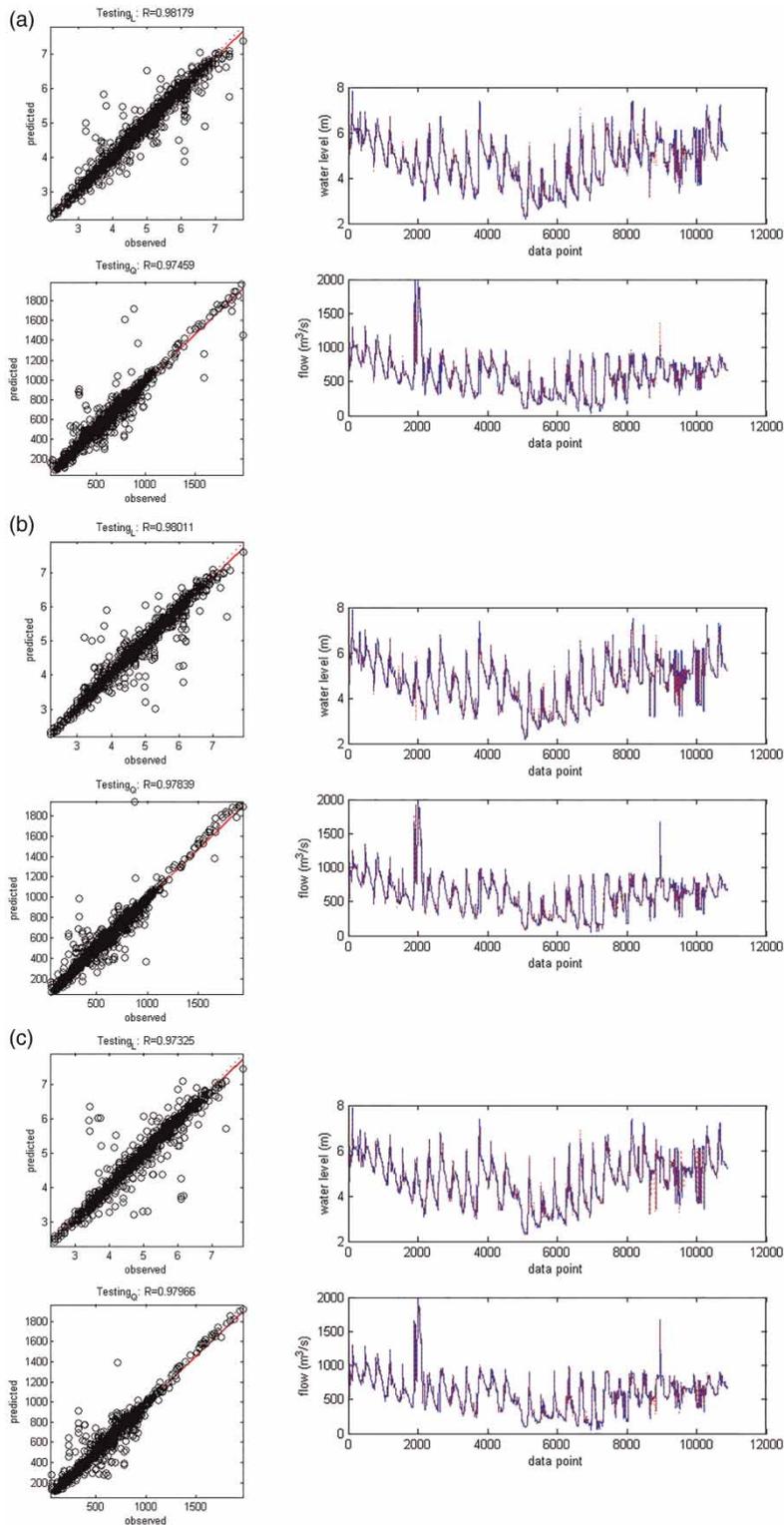| Model	Best architecture	Water level			Flow			Average values		
		MSRE	R	NS	MSRE	R	NS	MSRE	R	NS
1	18-8-2	0.0018	0.9818	0.9638	0.0262	0.9746	0.9498	0.0140	0.9782	0.9568
2	5-7-2	0.0019	0.9801	0.9604	0.0218	0.9784	0.9572	0.0119	0.9793	0.9588
3	3-3-2	0.0029	0.9732	0.9472	0.0356	0.9797	0.9596	0.0193	0.9765	0.9534

Figure 7 | Scatter plot and time series comparison of selected ANN structures. (a) Model 1; (b) Model 2; (c) Model 3.

time series. As shown in Table 2, all three models selected exhibit high performance in modelling both flow and water levels at Chiromo with respect to all the three statistics. According to Shamseldin (1997), an NS value of 0.9 and above is very satisfactory whilst a value between 0.8 and 0.9 presents a fairly good model. Anything below 0.8 characterises an unsatisfactory model. *R* values in excess of 0.9 reflect a very high efficacy of the model. Moriasi *et al.* (2007) rated as *very good* a model with a relative error of $< \pm 10\%$, 'good' for ± 10–15%, 'satisfactory' for ± 15–25% and 'unsatisfactory' for $> \pm 25\%$. In all the three models in Table 2, *R* and *NS* are above 0.9 and the maximum value of the average MSRE is below 2%. Similarly, high and low flows in the time series plots (time series plots are based on the whole data set) are well reproduced in all models. A comparison of the models (Table 2) shows that based on average values, model 2 performs better than the other two. Nonetheless, the differences are small, an aspect also reflected in the scatter and time series plots (Figure 7). Therefore, on the basis of parsimony, model 3 presents the best model and is chosen. Model 3 shows that remarkably good forecasts of the following day's water level and flow can be obtained at Chiromo with a parsimonious MLP-ANN model using only three inputs on a particular day: flow, water level and catchment rainfall. Forecasts of water levels are generally better compared with flows. This could be related to the quality of data.

When feature data (SOM predicted values) other than raw data were used for training, the ANN prediction accuracy was significantly improved (Figure 8). The best model identified in the previous section (i.e. with three inputs and two outputs) was used on features. The number of hidden neurons was systematically increased from 2 to 15 at a step of one. With MSRE (Figure 7), results are mixed; in some cases ANN trained on SOM features is better, in other cases ANN on raw data proves superior. Nevertheless, both *R* and *NS* show the modelling skills of ANN are enhanced when trained on feature data, supporting findings by Rustum & Adeloye (2012). This is important for data-poor areas where data are characterised by noise. By using SOM to first pre-process the data to obtain the error-free features as explained in a previous section, which are then used for forcing the MLP-ANN, results can be enhanced.
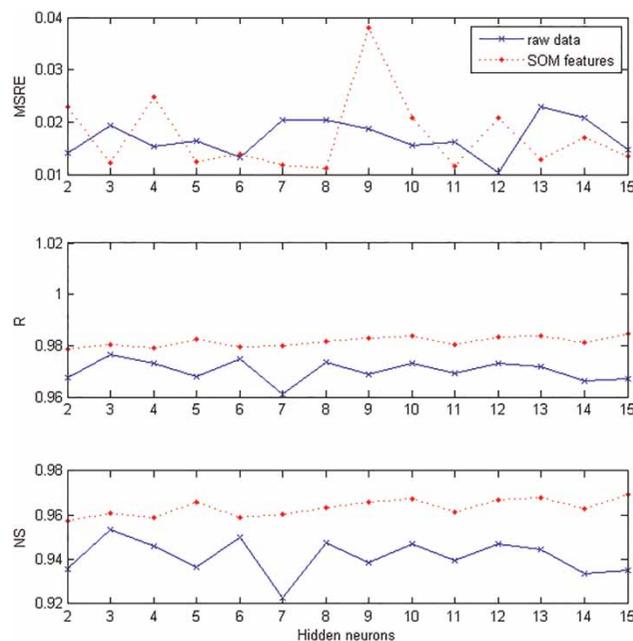

Figure 8 | ANN performance when trained on raw data and SOM features.

Flow and water level forecasting for 2- to 5-day leads

While a day's forecast is good, a longer lead prediction would be far more important for flood warning. The results on the 1-day ahead forecasts presented in the previous section confirmed that Model 3 with just the day's level, flow and rainfall will give satisfactory forecasts. This is an indication that for longer lead times, there is no need to consider longer memory models. Consequently, further ANNs were trained to forecasts 2–5 days ahead levels and flows at Chiromo using the day's measurement of two of the same inputs employed for model 3 for the 1-day lead forecasts. The input variable excluded was the catchment rainfall on the day and this was because, although the inclusion of rainfall as an input has been found to improve flow predictions (Toth & Brath 2007; Wu & Chau 2011), Wu & Chau (2011) observed that including rainfall in the input space only improved the results for lead times within the concentration time of a catchment. Otherwise, the impact of rainfall is insignificant. Since the time of concentration of Ruo catchment at Chiromo is less than a day, the day's rainfall was omitted and a series of 2-3-2 ANN models was trained on both raw data and the SOM features as before for 2–5 days' lead. The forecasts are mathematically

expressed as follows:

2-day lead: $(CRL_{t+2}, CRF_{t+2}) = f(CRL_t, CRF_t)$ (8)

3-day lead: $(CRL_{t+3}, CRF_{t+3}) = f(CRL_t, CRF_t)$ (9)

4-day lead: $(CRL_{t+4}, CRF_{t+4}) = f(CRL_t, CRF_t)$ (10)

5-day lead: $(CRL_{t+5}, CRF_{t+5}) = f(CRL_t, CRF_t)$ (11)

It is clear from Table 3 that in general model performance decreases with increasing lead time. The MSRE on both data sets increases with increasing lead time. Similarly, R and NS decrease with increasing lead. This behaviour is not unexpected. According to Solomantine & Dulal (2003), as the lead time increases, the inputs used are the remotest and, therefore, do not contain the recent information. It is also clear from the table that there is a significant improvement in prediction accuracy when ANN is trained on feature data. MSRE values are significantly reduced on SOM features. Similarly, R and NS values are higher when ANN is trained on feature data. Furthermore, Table 3

Table 3 | Comparison of raw and SOM-trained MLP-ANN for forecasting

Lead time (days)	Water level		Flow		Average	
	Raw	SOM	Raw	SOM	Raw	SOM
	MSRE					
2	0.0039	0.0029	0.0286	0.0223	0.0163	0.0126
3	0.0055	0.0030	0.0443	0.0275	0.0249	0.0153
4	0.0068	0.0038	0.0542	0.0334	0.0305	0.0186
5	0.0079	0.0044	0.0675	0.0394	0.0377	0.0219
	NS					
2	0.9214	0.9523	0.9244	0.9577	0.9229	0.9550
3	0.8873	0.9507	0.8723	0.9494	0.8798	0.9501
4	0.8779	0.9329	0.8723	0.9268	0.8751	0.9299
5	0.8393	0.9176	0.8469	0.9099	0.8431	0.9138
	R					
2	0.9599	0.9759	0.9615	0.9787	0.9607	0.9773
3	0.9420	0.9750	0.9343	0.9744	0.9382	0.9747
4	0.9370	0.9660	0.9363	0.9627	0.9367	0.9644
5	0.9162	0.9580	0.9203	0.9540	0.9183	0.9560

reveals that with ANN trained on raw data, remarkably good forecasts can be obtained up to 2 days ahead (based on the worst indicator, NS). Nonetheless, 3-, 4- and 5-day forecasts will still be fairly satisfactory. In contrast, if ANN is trained on features, a longer lead forecast is achievable with very satisfactory results. In this case, very satisfactory results for short-term forecasts of up to 5 days are attained.

CONCLUSION

The availability and quality of data in developing countries poses a big challenge for hydrological modelling. However, data-driven models such as ANN can help with such problems. While most hydrological studies have stalled due to data demands posed by traditional models, coupling SOM and MLP-ANN offers an alternative for these data-scarce countries. SOM is quite robust to missing data and can therefore infill for large gaps, something that would be impossible with traditional infilling methods, thus presenting a relatively long series needed for hydrological modelling. With most of the data characterised by noise, SOM filtered data present an alternative way of enhancing results. In this study, flows and water levels at Chiromo are reproduced to a high accuracy with a parsimonious MLP-ANN model. When SOM features are used, not only are forecasts very good for short lead times, the results are enhanced for longer lead times.

REFERENCES

Abrahart, R. J. & White, S. M. 2001 Modelling sediment transfer in Malawi: comparing backpropagation neural network solutions against a multiple linear regression benchmark using small data sets. *Physics and Chemistry of the Earth (B)* **26**, 19–24.

Adeloye, A. J. & De Munari, A. 2006 Artificial neural network based generalized storage-yield-reliability models using Levenberg–Marquardt algorithm. *Journal of Hydrology* **362**, 215–230.

Adeloye, A. J. & Rustum, R. 2012 Self-organising map rainfall-runoff multivariate modelling for runoff reconstruction in inadequately gauged basins. *Hydrology Research* **43**, 603–617.

Alvisi, S., Mascellani, G., Franchini, M. & Bardossy, A. 2006 Water level forecasting through fuzzy logic and artificial

neural network approaches. *Hydrology and Earth System Sciences* 10, 1–17.

Aquil, M., Kita, I., Yano, A. & Nishiyama, S. 2007 Neural networks for real time catchment flow modeling and prediction. *Water Resources Management* 21, 1782–1796.

Ayalew, L., Möller, D. P. F. & Reik, G. 2007 Using Artificial Neural Networks (ANN) for real time flood forecasting, the Omo River case in Southern Ethiopia. *Proceedings of the 2007 Summer Computer Simulation Conference*, Society for Computer Simulation International, San Diego, USA.

Baxter, C. W., Zhang, Q., Stanley, S. J., Shariff, R., Tupas, R. T. & Stark, H. L. 2001 Drinking water quality and treatment: the use of artificial neural networks. *Canadian Journal of Civil Engineering* 28, 26–35.

Bowden, G. J., Dandy, G. C. & Maier, H. R. 2005 Input determination for neural network models in water resources applications. Part 1–background and methodology. *Journal of Hydrology* 301, 75–92.

Chau, K. W., Wu, C. L. & Li, Y. S. 2005 Comparison of several forecasting models in the Yangtse River. *Journal of Hydrologic Engineering* 10, 485–491.

Cigizoglu, H. K. 2004 Estimation and forecasting of daily suspended sediment data by multi layer perceptrons. *Advances in Water Resources* 27, 185–195.

Dawson, C. W. & Wilby, R. L. 2001 Hydrological modelling using artificial neural networks. *Progress in Physical Geography* 25, 80–108.

Demuth, H., Beale, M. & Hagan, M. 2009 Neural Network Toolbox™ 6 User's Guide.

Diamantopoulou, M. J., Antonopoulos, V. Z. & Papamichail, D. M. 2005 The use of neural network technique for the prediction of water quality parameters of Axios River in Northern Greece. *European Water* 11/12, 55–62.

Dutta, D., Welsh, W. D., Vaze, J., Kim, S. S. H. & Nicholls, D. 2012 A comparative evaluation of short-term streamflow forecasting using time series analysis and rainfall-runoff models in ewater source. *Water Resources Management* 26, 4397–4415.

EU Directive 2007 Directive 2007/60/EC of the European parliament and of the council. *Official Journal of the European Union* L288, 27–34.

Haykin, S. 1999 *Neural Networks: A Comprehensive Foundation*. Prenticee Hall, New Jersey.

Hung, N. Q., Babel, M. S., Weesakul, S. & Tripathi, N. K. 2009 An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Sciences* 13, 1413–1425.

Iliadis, L. S. & Maris, F. 2007 An Artificial Neural Network model for mountainous water-resources management: the case of Cyprus mountainous watersheds. *Environmental Modelling & Software* 22, 1066–1072.

Ilunga, M. & Stephenson, D. 2005 Infilling streamflow data using feed-forward Back-Propagation (BP) artificial neural networks: application of standard BP and Pseudo Mac Laurin Power Series BP techniques. *Water SA* 31, 171–176.

Isik, S., Kalin, L., Schoonover, J. E., Srivastava, P. & Lockaby, B. G. 2013 Modeling effects of changing land use/cover on daily streamflow: an artificial neural network and curve number based hybrid approach. *Journal of Hydrology* 485, 103–112.

Ito, F. & Murata, A. 2009 Artificial neural network model estimating land use change in the Southwestern Part of Nagareyama City, Chiba Prefecture. In: *New Frontiers in Urban Analysis* (Y. Asami, Y. Sadahiro & T. Ishikawa, eds). CRC Press, Boca Raton, pp. 65–79.

Jain, S. K. 1993 Calibration of conceptual models for rainfall-runoff simulation. *Hydrological Sciences Journal* 38, 431–441.

Jain, S. K. & Sudheer, K. P. 2008 Fitting of hydrological models: a close look at the Nash–Sutcliffe Index. *Journal of Hydrologic Engineering* 13, 981–986.

Kalteth, A. M. & Berndtsson, R. 2007 Interpolating monthly precipitation by Self Organizing Map (SOM) and Multilayer Perceptron (MLP). *Hydrological Sciences* 52, 305–317.

Khamis, A., Ismail, Z., Haron, K. & Mohammed, A. T. 2006 Neural network model for oil palm yield modelling. *Journal of Applied Sciences* 6, 391–399.

Kirpich, Z. P. 1940 Time of concentration from small agricultural watersheds. *Civil Engineering* 10, 362.

Kohonen, T., Oja, E., Simula, O., Visa, A. & Kangas, J. 1996 Engineering application of the self organizing map. *IEEC Proceedings* 84, 1358–1384.

Kundzewicz, Z. W. & Menzel, L. 2005 Natural flood reduction strategies – a challenge. *International Journal of River Basin Management* 3, 125–131.

Legates, D. R. & Mccabe, G. J. 1999 Evaluating the use of goodness-of fit measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35, 233–241.

Liu, Z., Martina, M. L. & Todini, E. 2005 Flood forecasting using a fully distributed model: application of the TOPKAPI model to the upper Xixian catchment. *Hydrology and Earth System Sciences* 9, 347–364.

Liukkonen, M., Laakso, I. & Hiltunen, Y. 2013 Advanced monitoring platform for industrial wastewater treatment: multivariate approach using the self-organizing map. *Environmental Modelling and Software* 48, 193–201.

Lumbroso, D., Ramsbottom, D. & Spaliveiro, M. 2008 Sustainable flood risk management strategies to reduce rural communities' vulnerability to flooding in Mozambique. *Journal of Flood Risk Management* 1, 34–42.

Maier, H. R. & Dandy, G. C. 2000 Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling and Software* 25, 101–124.

Maier, H. R., Jain, A., Dandy, G. C. & Sudheer, D. 2010 Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. *Environmental Modelling and Software* 891–909.

Mazvimavi, D. 2003 Estimation of Flow Characteristics of Ungauged Catchments: Case Study in Zimbabwe. PhD Thesis, Wageningen University.

MOIWD (undated). Flood Forecasting and Warning Sytem for Lower Shire Valley Project Status Report. Lilongwe, Malawi.

Moore, R. J. & Bell, V. A. 2001 *Comparison of Rainfall-Runoff Models for Flood Forecasting. Part 1: Literature Review of Models*. Institute of Hydrology, Bristol

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transaction of the ASABE* **50**, 885–900.

Mwale, F. D., Adeloye, A. J. & Rustum, R. 2012 Infilling of missing rainfall and streamflow data in the Shire River basin, Malawi – a self organizing map approach. *Journal of Physics and Chemistry of the Earth* **50–52**, 34–43.

National Statistical Office 2012 Integrated Household Survey 2010–2011 – Household Socio-economic Characteristics Report, Zomba, Malawi.

Nayak, P. C., Rao, S. Y. R. & Sudheer, K. P. 2006 Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management* **20**, 77–90.

Ngongondo, C. S. X., Chong, Y. u., Tallaksen, L. M., Alemaw, B. & Chirwa, T. 2011 regional frequency analysis of rainfall extremes in Southern Malawi using the index rainfall and L-moments approaches. *Stochaistic Environmental Research and Risk Assessment* **25**, 939–955.

O'Connor, K. M. 2005 River flow forecasting. In: *River Basin Modelling for Flood Risk Mitigation* (D. W. Knight & A. Y. Shamseldin, eds). Taylor & Francis, London.

Okkan, U. 2011 Application of Levenberg-Marquardt optimization algorithm based multilayer neural networks for hydrological time series modeling. *An International Journal of Optimization and Control: Theories and Applications* **1**, 53–63.

Panda, S. S., Garg, V. & Chaubey, I. 2004 Artificial neural networks application in lake water quality estimation using satellite imagery. *Journal of Environmental Informatics* **4**, 65–74.

Plate, E. J. 2009 Classification of hydrological models for flood management. *Hydrology and Earth System Sciences* **13**, 1939–1951.

Primožič, M., Kobold, M. & Brilly, M. 2008 The implementation of the HBV model on the Sava River Basin. In: *XXIVth Conference of the Danubian Countries, Earth and Environmental Science.* http://dx.doi.org/10.1088/1755-1307/4/1/012004.

Rahman, M., Goel, N. & Arya, D. 2012 Development of the Jamuneswari flood forecasting system: case study in Bangladesh. *Journal of Hydrologic Engineering* **17**, 1123–1140.

Rustum, R. & Adeloye, A. J. 2007 Replacing outliers and missing values from activated sludge data using Kohonen self-organizing map. *Journal of Environmental Engineering* **133**, 909–916.

Rustum, R. & Adeloye, A. J. 2012 Improved modelling of wastewater treatment primary clarifier using hybrid ANNs.

International Journal of Computer Science and Artificial Intelligence **2**, 14–22.

Sahoo, G. B., Ray, C. & De Carlo, E. H. 2006 Calibration and validation of a physically distributed hydrological model, MIKE SHE, to predict streamflow at high frequency in a flashy mountainous Hawaii stream. *Journal of Hydrology* **327**, 94–109.

Saliha, A. H., Awulachew, S. B., Cullmann, J. & Horlacher, H. 2011 Estimation of flow in ungauged catchments by coupling a hydrological model and neural networks: case study. *Hydrology Research* **42**, 386–400.

Samuels, P., Klijn, F. & Dijkman, J. 2006 An analysis of the current practice of policies on river flood risk management in different countries. *Irrigation and Drainage* **55**, 141–150.

Shamseldin, A. Y. 1997 Application of a Neural Network Technique to Rainfall–Runoff Modelling. *Journal of Hydrology* **199**, 272–294.

Shela, O. N., Thompson, G., Jere, P. & Annandale, G. 2008 *Analysis of Lower Shire Floods & a Flood Risk Reduction and Recovery Programme Proposal for the Lower Shire Valley*, Department of Disaster Management Affairs, Lilongwe, Malawi.

Singh, R. D. 2005 *Real-Time Flood Forecasting – Indian Experiences*. International G-WADI Modelling Workshop, Roorkee, India.

Solomantine, D. P. & Dulal, K. N. 2003 Model Trees as an alternative to neural networks in rainfall – runoff modelling. *Hydrological Sciences Journal* **48**, 399–411.

Sudheer, K. P., Gosain, A. K. & Ramasastri, K. S. 2002 A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes* **16**, 1325–1330.

SVADD, Shire Valley Agricultural Development Project 1975 *An Atlas of the Lower Shire Valley, Malawi*. Department of Surveys, Blantyre.

Teschl, R. & Randeu, W. L. 2006 A neural network model for short term river prediction. *Natural Hazards and Earth Systems Sciences* **6**, 629–635.

Thielen, J., Bartholmes, J., Ramos, M. H. & De Roo, A. 2009 the European flood alert system – part 1: concept and development. *Hydrology and Earth System Sciences* **13**, 125–140.

Todini, E. 1996 The ARNO rainfall-runoff model. *Journal of Hydrology* **175**, 339–382.

Toth, E. & Brath, A. 2007 Multistep ahead streamflow forecasting: role of calibration data in conceptual and neural network modeling. *Water Resources Research*, **43**.

UNISDR 2005 *Hyogo Framework for Action 2005–2015: Building the Resilience of Nations and Communities to Disasters*. Geneva.

Van Ogtrop, F. F., Hoekstra, A. Y. & Van Der Meulen, F. 2005 Flood management in the Lower Incomati River Basin, Mozambique: two alternatives. *Journal of the American Water Resources Association* **41**, 607–619.

WMO/GWP 2009 Integrated Flood Risk Management Concept Paper. Geneva.

Wu, C. L. & Chau, K. W. 2011 Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology* **399**, 394–409.

Wu, C. L., Chau, K. W. & Li, Y. S. 2009 Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology* **372**, 80–93.

Yawson, D. K., Kongo, V. M. & Kachroo, R. K. 2005 Application of linear and non-linear techniques in riverflow forecasting in Kilombero River Basin, Tanzania. *Hydrologic Sciences* **50**, 783–796.

Yuan, H., Van Der Wiele, C. F. & Khorram, S. 2009 An automated artificial neural network system for land use/land cover classification from Landsat TM Imagery. *Remote Sensing* **1**, 243–265.