Artificial neural network approach to flood forecasting in the River Arno

MARINA CAMPOLO*

Centro Interdipartimentale di Fluidodinamica e Idraulica, Università di Udine, I-33100 Udine, Italy

ALFREDO SOLDATI

Centro Interdipartimentale di Fluidodinamica e Idraulica and Dipartimento di Energetica e Macchine, Università di Udine, Via delle Scienze 208, I-33100 Udine, Italy
soldati@uniud.it

PAOLO ANDREUSSI

Dipartimento di Chimica e Chimica Industriale, Università di Pisa and *Centro TEA, Consorzio Pisa Ricerche, Piazza Mazzini 1, I-56100 Pisa, Italy*

Abstract The basin of the River Arno is a flood-prone area where flooding events have caused damage valued at more than 100 billion euro in the last 40 years. At present, the occurrence of an event similar to the 1966 flood of Firenze (Florence) would result in damage costing over 15.5 billion euro. Therefore, the use of flood forecasting and early warning systems is mandatory to reduce the economic losses and the risk for people. In this work, a flood forecasting model is presented that exploits the real-time information available for the basin (rainfall data, hydrometric data and information on dam operation) to predict the water-level evolution. The model is based on artificial neural networks, which were successfully used in previous works to predict floods in an unregulated basin and to predict water-level evolution in the Arno basin under low flow conditions. Accurate predictions are obtained using a two-year data set and a special treatment of input data; which allows a balance to be found between the spatial and temporal resolution of rainfall information and the model complexity. The prediction of water-level evolution remains accurate within a forecast time ahead of 6 h, which is the minimum time lag for the river to respond to dam releases under saturated conditions of the basin. The predicted flow rate percentage error ranges from 7 to 15% from the 1-h ahead to 6-h ahead predictions, and the accuracy of prediction increases for each time ahead of prediction, as the flow rate increases, suggesting that the model is particularly suited for flood forecasting purposes.

Key words flood forecasting; artificial neural network; system response identification; nonlinear modelling; rainfall–runoff; River Arno, Italy

Une approche à base de réseau de neurones artificiels pour la prévision des crues du fleuve Arno

Résumé Le bassin du fleuve Arno est une zone sujette au phénomène des inondations, où le coût des dégâts dus aux inondations durant les 40 dernières années se chiffre à plus de 100 milliards d'euros. De nos jours, un événement aussi grave que l'inondation de 1966 à Florence produirait des dommages pour plus de 15.5 milliards d'euros. C'est pourquoi l'utilisation de systèmes de prévision et d'annonce précoce de crues s'avère nécessaire en vue d'une réduction des pertes économiques et du risque pour les personnes. Nous présentons dans ce travail un modèle de prévision de crues exploitant les informations en temps réel disponibles pour le bassin (données de précipitations, enregistrements hydrométriques et opérations de barrage), dans le but de prévoir l'évolution du niveau de l'eau. Le modèle est basé sur des réseaux de neurones artificiels qui ont été employés avec succès dans des travaux développés précédemment pour prévoir les crues dans un bassin non-régulé et pour prévoir

* *Currently at:* Instituto Pluridisciplinar, Universidad Complutense de Madrid, Paseo Juan XXIII, ES-28040 Madrid, Spain.

Open for discussion until 1 December 2003

l'évolution du niveau de l'eau dans le bassin de l'Arno en conditions d'écoulement réduit. Des prévisions précises sont obtenues à partir d'un jeu de données de deux ans et d'un traitement spécial des données d'entrée, qui permet de trouver un équilibre entre la résolution spatiale et temporelle des informations de précipitation et la complexité du modèle. La prévision de l'évolution du niveau de l'eau reste précise pour une durée de six heures, ce qui est le délai minimum pour que le fleuve réagisse à l'ouverture des barrages quand le bassin est en condition de saturation. L'erreur de prévision du débit est comprise entre 7 et 15% pour les anticipations de une à six heures. D'autre part, la précision des prévisions augmente quand le débit croît, ce qui laisse penser que le modèle est particulièrement approprié pour être utilisé en prévision de crues.

Mots clefs prévision de crues; réseau de neurones artificiels; identification de la réponse du système; modélisation non-linéaire; pluie–débit; fleuve Arno, Italie

INTRODUCTION

Flooding of the River Arno has caused serious damage to the city of Firenze (Florence), Italy, which is the first large city the river crosses as it leaves the mountain region. To reduce damage from flood events, a net of telemetered raingauges and river gauges has been installed and is operating in the basin, making real-time data available at the Ufficio Idrografico and Mareografico of Pisa; a meteorological radar has been installed in the area of Pisa to allow quantitative prediction of rainfall; and several studies have been performed to identify the areas mostly exposed to the risk of floods. For the successful implementation of policies for flood mitigation and risk reduction, information from these different sources should be integrated and used to obtain an accurate and timely prediction of the river-level rise.

Rain falling on the mountains and reservoir operations are the two inputs that cause the river-level rise at the basin closing section. The state of saturation of the basin is an additional variable, which determines the velocity of this process. These variables are generally used in traditional hydrological models to evaluate the amount of water available in the basin. In the Arno basin, this information is essential in order to set up a basin-scale water management system with the following objectives: (a) flood control and risk mitigation, (b) water supply for municipal and industrial uses, (c) water quality control, and (d) power production optimization.

An extensive review of traditional physically-based models may be found in WMO (1992). Following a less traditional system approach, the basin may be considered as a complex, nonlinear system, and the modelling of the rainfall–runoff process can be stated as a system-response identification problem. Predictability of future behaviour is a consequence of the correct identification of the system transfer function. Among the different approaches proposed and discussed in the literature, the phase-state reconstruction method, based on univariate (Porporato & Ridolfi, 1997; Liu *et al.*, 1998; Sivakumar *et al.*, 2001a,b), or multivariate nonlinear prediction (Porporato & Ridolfi, 2001), and artificial neural networks (Hsu *et al.*, 1995; Smith & Eli, 1995; Minns & Hall, 1996; Shamseldin, 1997; Dawson & Wilby, 1998; Fernando & Jayawardena, 1998; Thirumalaiah & Deo, 1998; Campolo *et al.*, 1999a,b; Imrie *et al.*, 2000; Maier & Dandy, 2000; Hu *et al.*, 2001; Xiong & O'Connor, 2002; Rajurkar *et al.*, 2002) are receiving the greatest attention. The interest in these models is due to their ability to produce reasonably accurate results in a very short computation time, exploiting a reduced set of data that may be readily available in real time. For these reasons, they are optimal candidates for an easy and successful integration into flood forecasting and early warning systems (Kim & Barros, 2001).

The object of this work is to present and discuss an artificial neural network-based model developed for the real-time forecasting of floods in the River Arno. A similar model was successfully used in a previous work to forecast floods in a different basin (Campolo *et al.*, 1999a) and to forecast river flow rate in the Arno basin during low-flow periods (Campolo *et al.*, 1999b). These models may be eventually used as input for river basin management (Campolo *et al.*, 2002).

In the previous flood forecasting application (Campolo *et al.*, 1999a,b), an un-regulated basin, smaller (1950 km^2) than the Arno basin (4000 km^2) was considered. In the present work, the larger basin makes it necessary to evaluate carefully the effect of the spatial distribution of rainfall, at least on a sub-basin scale. Therefore, in order to adjust the need for distributed rainfall information against the growing complexity of the neural network model, a special procedure is proposed for the aggregation in time and space of rainfall data. In the previous application on the Arno basin the water-level evolution during low flow periods was predicted. In those conditions, mechanisms determining runoff are completely different, the variation in time of water level is smooth and predictions may be issued on a daily basis. In this work, the water-level evolution is predicted hourly for medium–high flow periods, when the water level of the river exceeds a threshold value such that heavy rainfall can produce floods.

Experimental data made available for the work include rainfall, hydrometric and power production data collected in the basin during 1992 and 1993. Due to the kind of data exploited for prediction, it was found that model performances remain satisfactory up to 6 h ahead, which is too short for operational flood forecasting. The availability of new data, mainly radar data, should allow the time of prediction of the model to be extended to operationally useful values.

SITE AND DATA

Figure 1 shows the River Arno basin, Italy. In this study, only the upper part of the basin, closed by the section of Nave di Rosano, a few kilometres upstream of Firenze, is considered. This part covers about 4000 km^2 and can be divided into five sub-basins

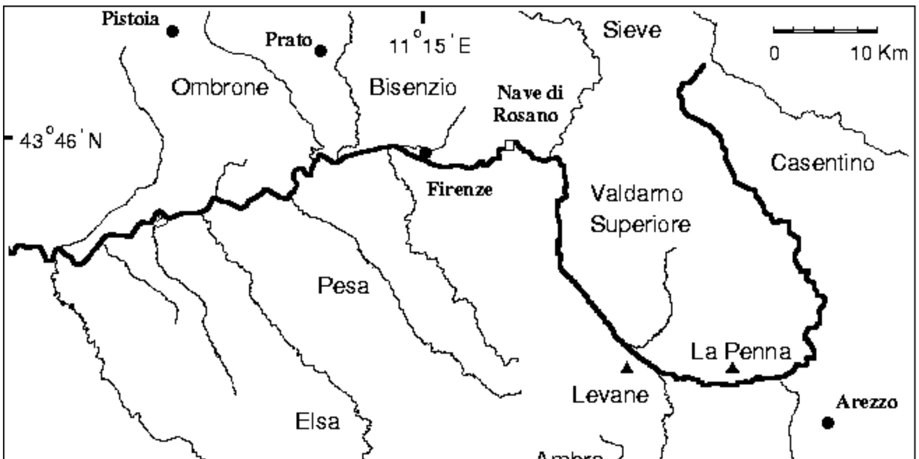

Fig. 1 Map of the River Arno basin. (Nave di Rosano closes the upper part of the basin; two dams contribute to flow regulation.)

(Casentino, valley of Valdarno Superiore and sub-basins of the tributaries Chiana, Ambra and Sieve). The sub-basins are characterized by differences in topography, climate and vegetation. Slopes range from 20% in the hilly and mountainous zones to 3% in the plain areas around the main river valleys. This complex topography determines a high variability in the distribution of rainfall and mean temperature. The change in climate generated by the altitudinal gradients and distances from the sea determines high variability in vegetation types and land cover. As a consequence, the response of each sub-basin to a rainfall event must be considered separately, since lag time and soil permeability to rainfall are different.

A network of raingauges and river gauges collects rainfall and water-level data with a time interval of 15 min and these are automatically available at the Ufficio Idrografico and Mareografico of Pisa; meteorological data (temperature, relative air moisture) are available at the same gauging stations. In the present study, a set of 31 raingauges, located in the upper part of the basin, and the gauge at Nave di Rosano were selected. A stage–discharge relationship available at Nave di Rosano allows the water level to be converted into the corresponding flow rate.

The flow rate of the River Arno is highly variable during the year and is regulated by two artificial water reservoirs, La Penna and Levane, holding 13 and 3×10^6 m^3, and located 40 and 30 km upstream of Firenze, respectively. At the section of Nave di Rosano the yearly averaged flow rate is about 50 m^3 s^{-1} (a minimum flow rate of 0.560 m^3 s^{-1} was recorded on 29 August 1958, and a maximum flow rate of 3540 m^3 s^{-1} was recorded on 4 November 1966).

There are two main contributions to the river flow rate at Nave di Rosano: rainfall production and storage regulation. Snowmelt contribution to runoff may be significant only during spring. The La Penna Reservoir does not contribute directly to flow rate during high flow periods since its large volume allows water storage for flood mitigation. The Levane Reservoir modifies the flow rate in the Arno, discharging the water through turbines for power production. Data on the hydroelectric power production of the Levane Dam with a time interval of 1 h were provided by the National Power Agency (ENEL). Hydropower production is directly related to the water discharged from dam by the relationship:

$$P = k \cdot \frac{Q}{H} \tag{1}$$

where P is the power, Q is the flow rate and H is the height of the waterfall, being the difference between the level of water in the reservoir and that of the tail water.

The level of the reservoir is not continuously recorded and the discharged flow rate may be estimated only with some uncertainty. Nevertheless, since artificial neural network models may use multivariate time series for flood prediction, it was decided to use the raw power data as input to the model.

METHODOLOGY

Artificial neural networks (ANN) are nonlinear, multi-dimensional interpolating functions. The functional form of the ANN model developed herein is extrapolated using a best-fit procedure from pairs of input–output data, each one representing an example of the transformation to be modelled. Transformation of the input array of data, $\overline{I} = \{x_i, i = 1, N_I\}$, consisting of N_I components including rainfall, water level and power data, into the output array, $\overline{O} = \{o_k, k = 1, N_O\}$, consisting of N_O components representing the water-level evolution in the next hours, is obtained by combining a sequence of simple computations performed in parallel by a number of elementary processing units, or nodes. Components of the input array, x_i, are weighted, summed up and then transformed by each node using a nonlinear activation function, *f*:

$$h_j = f\left(\sum_{1}^{N_I} w_{i,j} \cdot x_i - \sigma_j\right) \qquad j = 1, N_H \tag{2}$$

The result of this operation is an array, $H = \{h_j, j = 1, N_H\}$, where the number of components N_H is equal to the number of processing internal nodes that can be processed by a subsequent layer of nodes in the same way as the input array:

$$o_k^c = f\left(\sum_{j=1}^{N_H} w_{j,k} \cdot h_j - \sigma_k\right) \qquad k = 1, N_O \tag{3}$$

The array calculated by the last layer of nodes, $O_c = \{o_k^c, k = 1, N_O\}$, is the output array, $\overline{O}$. The weighting parameters, $w_{i,j}$ and $w_{j,k}$, the thresholds, σ_j and σ_k, and the number of processing units of internal layers, N_H, are the adjustable parameters of the model which are modified during the model building phase—training (see Haykin, 1995, for details) and finally fixed in order to make the calculated output, $\overline{O}_c$, as close as possible to the target, $\overline{O}$, thus minimizing the error of the model. Generally, and in this work, the correction of the weights is made proportional to the actual error of the model (standard back-propagation error algorithm). When the error of the model decreases below a specified threshold, the building phase is over and the model can be used to compute transformation of new series of input data.

In this study, a standard methodology based on feedforward neural networks, trained with the standard back-propagation algorithm was used (see Haykin, 1995, for details). Feedforward neural networks are widely used for hydrological applications because they are simple, accurate and allow high processing speeds. Other configurations, for example recurrent neural networks, do not provide clear practical advantage (Maier & Dandy, 2000). The neural model was implemented and calibrated using the software Stuttgart Neural Network Simulator (SNNS), developed by the University of Stuttgart, Germany, and available as free software from the Internet (SNNS, 1995).

Analytical methods and hydrological expertise helped to determine the inputs for the ANN model in order to reduce the network size and training time (Maier & Dandy, 1996, 1997).

For the present application, the efficient use of input data is a critical issue. Given the basin topography and climate variability, the distributed contribution of rainfall to runoff needs to be accounted for. Nevertheless, the use of distributed rainfall information adds to the complexity of the model, to the training time and, in the absence of extensive data sets, to the possibility of overfitting. Therefore, during the model building phase, the input data were carefully analysed to achieve an optimum balance between input data information and model complexity.

DATA ANALYSIS AND MODEL SET-UP

Performance of neural networks, as any data-driven model, is extremely sensitive to the data used during model set-up and calibration: (a) data must be adequate and specific for the modelling task, (b) input data must be selected according to their relevance for the modelling, and (c) input data must be as limited as possible to reduce the training time and the possibility of overfitting.

With these basic rules in mind, it was first decided to make the database problem-specific by removing very low flow periods that may lead to model performance degradation. In these conditions, the probability of flood is negligible and the mechanisms controlling runoff production are considerably different. Campolo *et al.* (1999b) chose a threshold value equal to 1.5 m relative to the position of the recording gauge to identify low flow conditions and predict the daily water-level evolution. In this work, since the object is to predict hourly water-level variations, which may produce floods, a different threshold value was chosen for the water level (–0.25 m) and for prediction only those periods were chosen during which the water level is higher than the threshold. This lower threshold value allows floods generated by heavy rainfall falling on the basin initially in unsaturated conditions to be included in the data set.

Second, to determine the input of the neural model, the data available to establish a causality relationship and time-lags between input and output data were analysed. The sampling rate of available data is one hour, as fixed by power data on dam regulation. Figure 2 shows a typical time series of water-level variation following a rainfall event and reservoir regulations. The main peak in the water-level time series is due to the rainfall, whereas the number of subsequent peaks are due to dam releases that may sensibly modify the falling limb of the hydrograph. Using cross-correlation between segments of power data and water level and between rainfall and water level, the lag time necessary for the system to respond to these different perturbations was evaluated, providing valuable information about the rainfall–runoff dynamics. The lag was identified as the value producing a peak in the cross-correlation diagram. As expected, the lag is influenced, first, by the distance of the raingauge or the power production site from the gauge and, second, by the state of the basin, i.e. the lag is smaller in a saturated basin. In the Arno basin, maximum and minimum lags, corresponding to unsaturated and saturated conditions, are of the order of 16–8 h for the rainfall and 10–6 h for the power data, respectively. From this analysis, the segment of rainfall and power data relevant for water-level prediction on a hourly basis was identified. Yet, the use of relevant hourly data for each of the 31 raingauges poses a crucial problem of high number of inputs. Then, distributed rainfall information (in space and time) needs to be used to account for rainfall variability and its effect in floods production, and the number of inputs needs to be kept small because the data set is small.

Therefore, it was decided to group the data based on the results of the correlation analysis as follows. First, groups of raingauges were identified for which the time lag was similar. These data, corresponding to raingauges belonging to the same sub-basin, were averaged on an hourly basis. Sub-basin averaged rainfall data given as input to the neural model for a time period equal to the maximum time lag should represent the forcing input for water-level variation (input from I5 to I39 in Fig. 3). Next, when the time lag becomes larger than the maximum correlation value, it was found that the contribution of rainfall from each raingauge to flow rate variation could not be

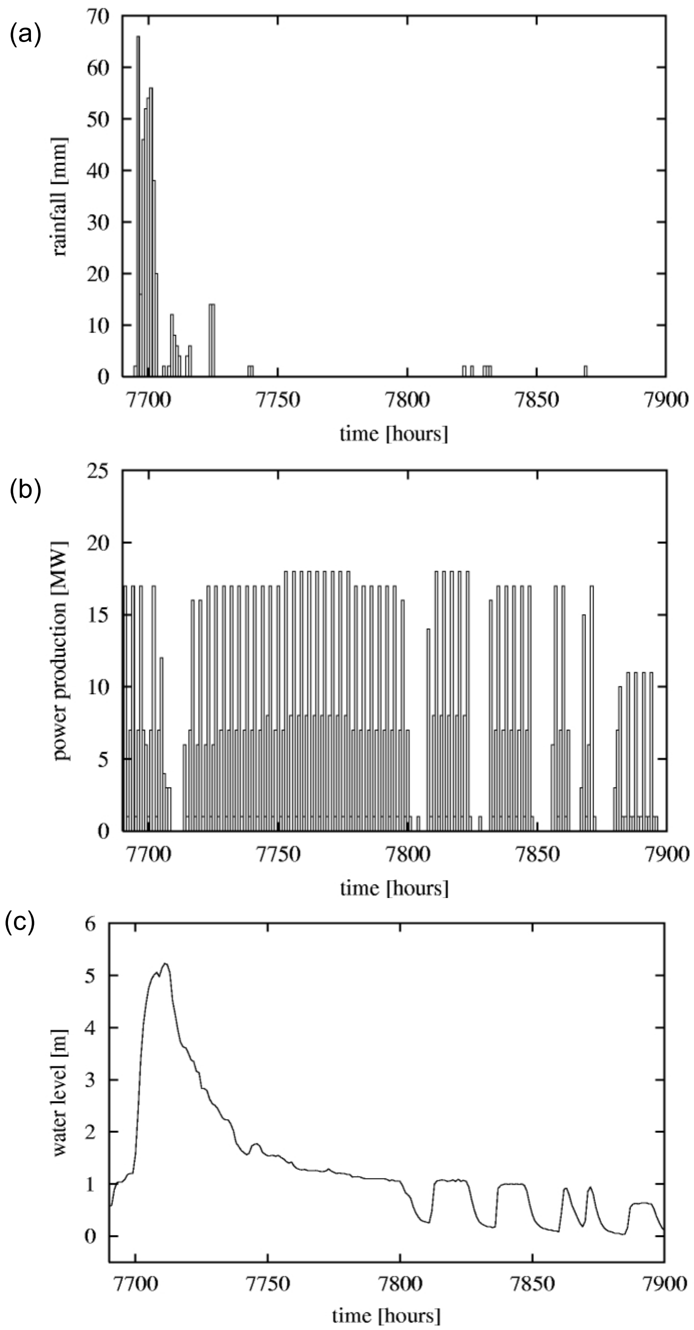

Fig. 2 Time series of (a) rainfall, (b) power production, and (c) water-level variation at Nave di Rosano. Rainfall and power production are forcing functions determining system response.

identified. Yet, these rainfall data may give useful information on the state of saturation of the basin. Therefore, it was decided to aggregate over the basin-scale rainfall values delayed more than the maximum lag. Furthermore, analysing the trend of basin-averaged rainfall data only smooth variations over time were found, suggesting that the input dimension for the model may be further reduced. It was decided to describe soil moisture variation using rainfall cumulated over a fixed time period and the time evolution of this variable. The trend of rainfall cumulated over a period of k hours (with k in the range 2–8 h) was considered and the possibility of describing the time evolution of this function using a reduced sampling rate was evaluated. It was found that using $k = 4$ and a four-step approximation for the trend, an optimal compromise could be found between accurate description of soil moisture variation and number of inputs. This procedure is in line with previous analysis by Nalbantis (2000), who found that the importance of rainfall resolution decreases moving back in time.

As previously observed by Minns & Hall (1996), information about the forcing inputs (rainfall and power production) alone is not sufficient to compute the flow rate, since the state of the basin plays an important role in determining flow rate behaviour. Even though soil moisture content is partially accounted for in this study using basin-averaged rainfall information, it was decided to also use the gauge recording at certain time intervals before the time of prediction as additional input. As reported by Porporato & Ridolfi (1997), the high level of autocorrelation for subsequent values in the time series indicates that the water level is the best water-level predictor.

The correlation analysis also gave a final suggestion as to the maximum time advance for accurate prediction. Only ground-truth data were used to make predictions and it is clear that accuracy of prediction decreases if these data become incomplete or inaccurate. Since the minimum time lag between the Levane Reservoir operations and water-level variation is 6 h when the basin is in saturated conditions, it will be difficult to extend the time advance of prediction beyond this period. The future scheduling for reservoir operations, which influences the water-level rise, should be available in order to extend the time advance of the forecast further into the future.

Figure 3 shows the structure of the input and output data adopted for the model. To predict the water level from time T, i.e. the next hour, to $T + 5$, i.e. 6 h ahead, the following data were used: (a) the rainfall cumulated over 4 h for the entire basin for time $T - 20$, $T - 16$, $T - 12$ and $T - 8$; (b) the rainfall averaged over each sub-basin from hour $T - 7$ to $T - 1$; (c) the power production data from hour $T - 9$ to $T - 1$; and (d) the water level from time $T - 9$ to $T - 1$. Power production data are given for the time period corresponding to the maximum lag time calculated from the correlation analysis and up to the present time in order to use all the information available to extend the time ahead of prediction. Based on the structure chosen for the input data, the time series were divided into two independent sets—calibration and validation.

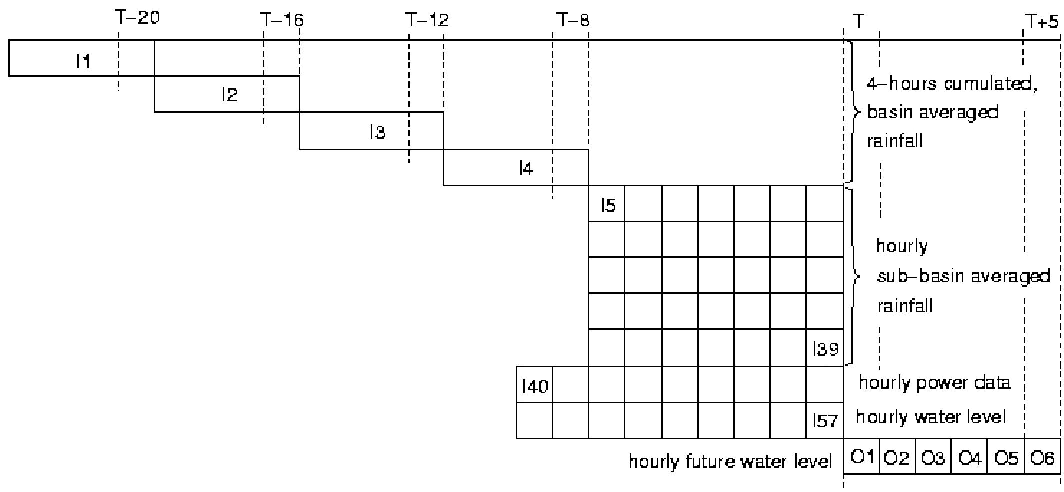

Fig. 3 Input data used for water-level prediction up to 6 h ahead.

Table 1 Division of data into training, testing and validation sets. Range of variation of water level in the data sets.

Data set	Period	Water-level range (m)
Testing	10–22 January 1992 7 February–7 March 1992 23 March–25 April 1992 30 April–5 May 1992 12–16 June 1992 5–10 July 1992	–0.25 to 5.3
Training	4 October 1992–1 January 1993 25 March–5 April 1993 10–23 April 1993	–0.25 to 6.72
Validation	1–4 October 1993 6–29 October 1993 3–25 November 1993 29 November–31 December 1993	–0.25 to 5.8

The data were grouped carefully in order to ensure that statistical properties of each subset—i.e. mean, variance, minimum and maximum value of water level—were most similar (see Table 1 for details). Using the same criteria, the calibration ensemble was further divided into two portions, the training set and the testing set. The training set was used to minimize the error of the model while the testing set was used to check generalization capability as calibration proceeds (see Bishop, 1994; Haykin, 1995 for details).

The procedure for data division was adopted to avoid the problem of extrapolation. The inability of neural networks to extrapolate beyond the range of data used for training is widely acknowledged (Maier & Dandy, 2000; Minns & Hall, 1996) and very few examples of neural network configurations with improved extrapolation ability are reported in the literature (Imrie *et al.*, 2000). This is a potential problem for flood forecasting models, since the probability of flooding events exceeding the range covered by the training data may be low and yet remains different from zero. The problem of extrapolation was tackled by standardizing all data with the help of linear scaling from the overall range of variation evaluated from available data to the range [0.1–0.9], the output of the activation function being bounded between 0 and 1. In particular, the water level was scaled from [–0.25 to 6.72] to [0.1 to 0.9]. The reduced range selected for the scaling allows output water levels within the range [–1.12 to 7.59] to be obtained. These values represent physically realistic boundaries for the present problem, since the lower value is below the zero flow gauge level and the larger value corresponds to a discharge in the river of about 3540 $m^3 s^{-1}$, which is a 200-year-flood (value recorded in 1966). Similarly, the rainfall time series was scaled at each raingauge assuming that greater rainfall may fall on the basin. Greater rainfall would result in the calculation of water levels higher than the maximum in the training set, and would allow prediction of water levels in the river that are higher than those used for training.

Based on the structure identified for the input–output transformation and on the division of data into training, testing and validation sets, a number of patterns—i.e. related pairs of input–output data—was generated, being 2416, 1874 and 1708 for the training, testing and validation sets, respectively.

A trial-and-error procedure was used to fix the number of hidden nodes and determine the final structure of the neural network model. Five different networks were trained with 20, 25, 30, 35 and 40 hidden nodes following the same fixed scheme (four epochs of 1000 iterations each, decreasing the learning rate from 0.2 to 0.05 by step 0.05 at each epoch). For each network configuration, the training procedure was repeated 10 times, starting from different random initialization values for the weights. The best performing among these nets was considered for further comparison with the other configurations. During training, the error in prediction over the entire time horizon was checked on the testing set to avoid overfitting. It was found that the net with 30 hidden nodes allowed optimal performances—i.e. minimum error in prediction—to be obtained.

The final structure of the net exploits 57 input nodes, 30 hidden nodes and six output nodes. As suggested in many previous works (e.g. Minns & Hall, 1996), the number of hidden nodes is roughly half of the number of input nodes. The total number of adjustable parameters is 1926 and that of available patterns (training set) is 2416. This gives a ratio of training sample to weight number of 1.25, which is a small value compared to many empirical relationships suggested in the literature (see Maier & Dandy, 2000, for discussion). Following Rogers & Dowla (1994), the number of weights should not exceed the number of training samples, whereas a ratio of training sample to weight number up to 30 is necessary to obtain a good generalization ability. To verify the generalization ability of the present model, a cross-validation approach (Bishop, 1995) was used. The calibration data were divided into 10 independent sets at random, and the network was trained again leaving out one set of data at a time for testing. No significant differences were found in network performances by training the net using this enhanced data set.

The possibility was also considered of reducing the number of inputs by testing the performance of an ANN model exploiting basin-averaged rainfall from hour $T-7$ to $T-1$ instead of rainfall averaged over each sub-basin. All the other settings (other input data, number of hidden nodes, training procedure) were left unchanged. This cheaper model was not able to reproduce the rainfall–runoff transformation accurately, and produced 1-h ahead prediction 20% in error for events generated by rainfall falling in single sub-basins.

RESULTS AND DISCUSSION

To verify the adequacy of information presented to the model and the existence of prediction limits, the distribution of the error was plotted against the prediction lead time.

Figure 4 compares river level predictions and measured data for the 1-h and 6-h forecasts on the calibration and validation sets. Dispersion of points increases from the 1-h to the 6-h ahead prediction but remains satisfactory. For the validation set, points corresponding to extreme values of water level are slightly overestimated or underestimated in the 6-h ahead prediction, while some points corresponding to medium-flow conditions are heavily underestimated.

Figure 5 shows two segments of the time series of the measured water level compared with the 1-h ahead and 6-h ahead prediction. The segments shown from the calibration and validation sets are representative of predictions obtained over the entire

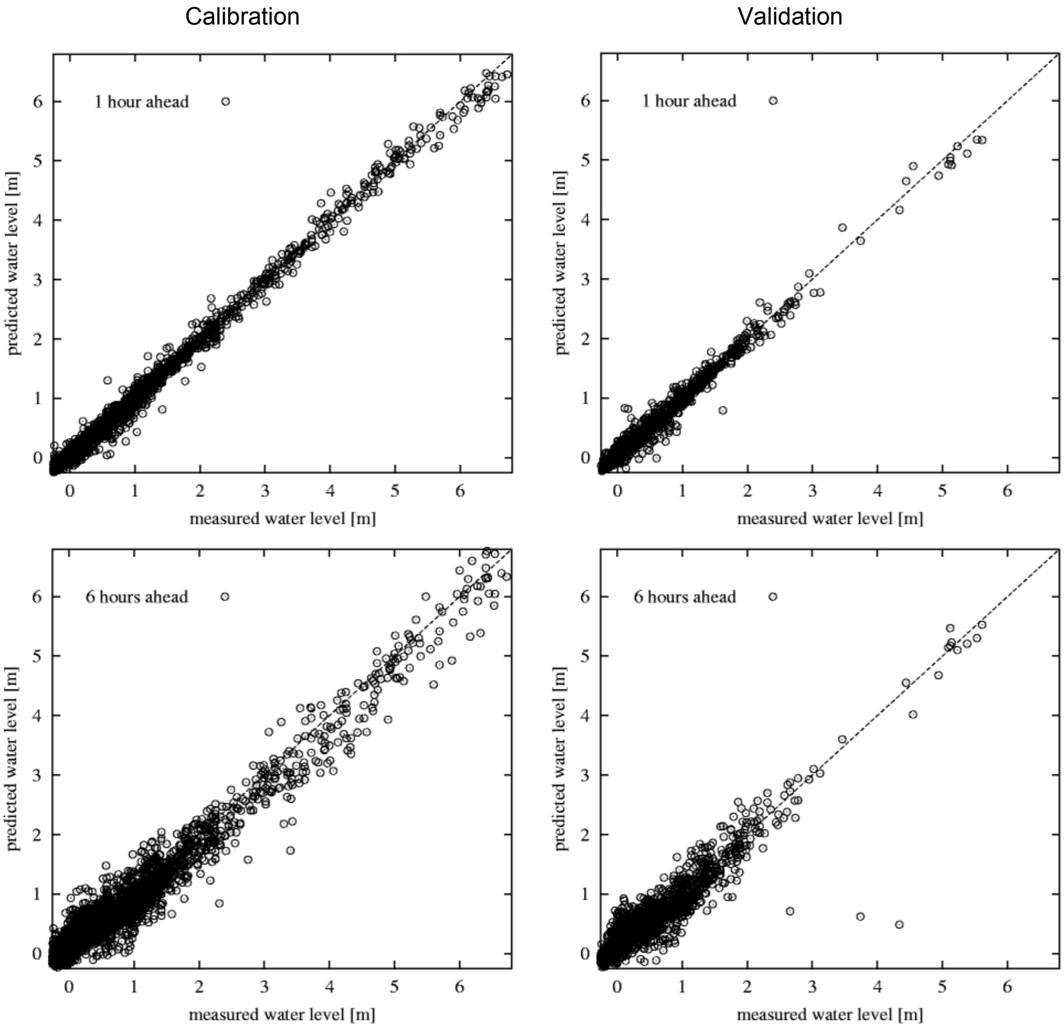

Fig. 4 Comparison between measured and predicted water level for (top) 1-h ahead and (bottom) 6-h ahead prediction on (left) calibration and (right) validation sets.

set of data. The 1-h ahead prediction reproduces the data accurately. In the 6-h ahead prediction, deviations are found for water-level variations due to dam operations (peaks on the recession limb) and for the main peak.

The overall accuracy of the model was evaluated by calculating *RMSE* and correlation coefficients. Since the water-level height is relative to the gauge position and statistics calculated on this variable can falsify the results, the stage–discharge relationship was first used to convert water level into flow rate; second, *RMSE* and correlation coefficient were calculated using these transformed values. Figure 6(a) and (b) shows the *RMSE* and the correlation coefficient calculated from the calibration and validation sets. The *RMSE* increases with time ahead of prediction for both sets. The value is smaller for the validation set than for the calibration set. The increasing trend within the time horizon becomes steeper for the 6-h ahead prediction. As initially suggested by the cross-correlation analysis, this proves that the given input information becomes progressively insufficient to compute water levels. These results were compared with performances of a neural model exploiting the same set of inputs and predicting the water-level evolution for the next 12 h. Figure 6(c) and (d) shows that

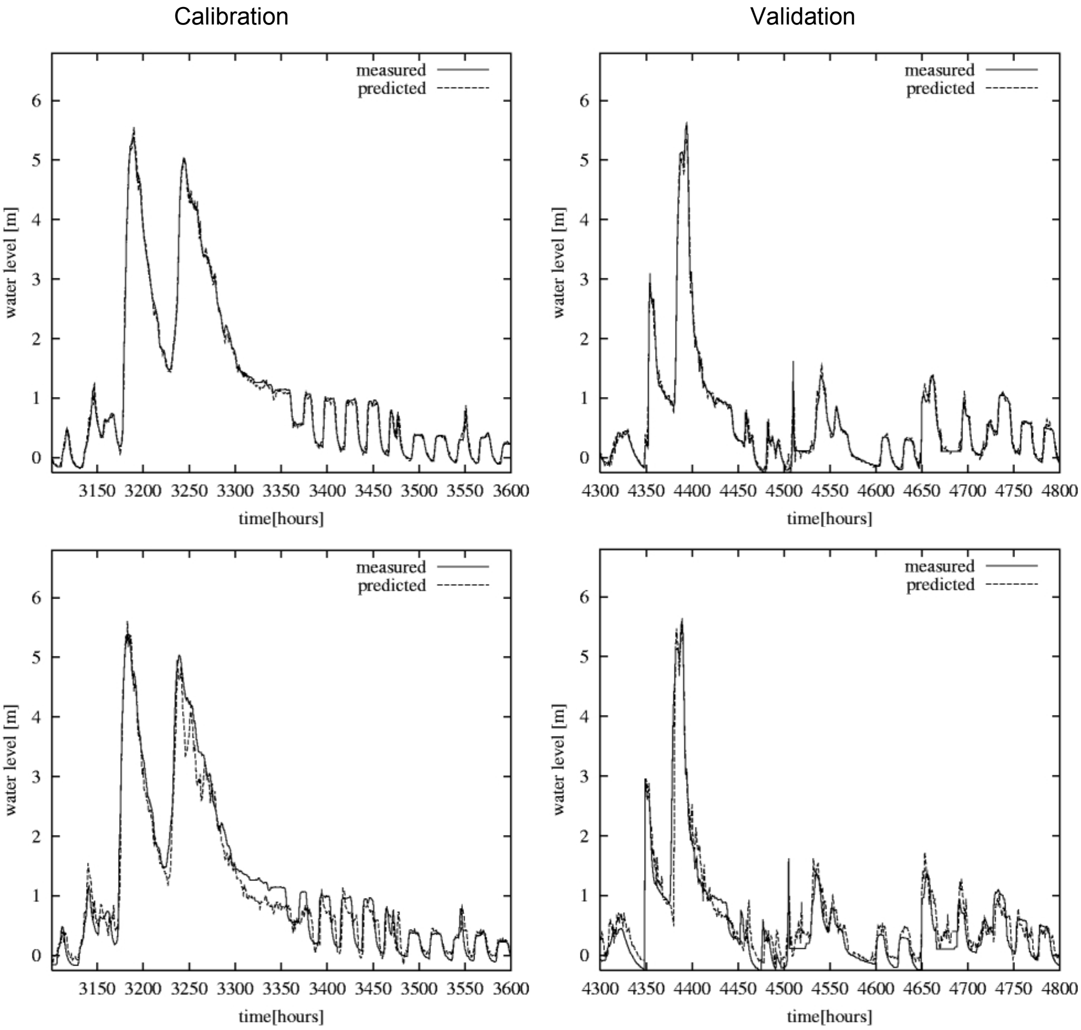

Fig. 5 Time series of predicted and measured water level for (top) 1-h ahead and (bottom) 6-h ahead prediction for (left) calibration and (right) validation sets.

RMSE values for $T = 1–6$ are larger than in the six-output model. Furthermore, a steep change is observed for $T > 6$. For the model presented in this work, the correlation coefficient between calculated and measured flow rate, shown in Fig. 6(b), remains about 99% for the calibration set and above 95% for the validation set for the entire time ahead of prediction. Lower values are found for the 12-h ahead prediction (Fig. 6(d)), suggesting that input information is adequate for water-level forecasts up to 6 h in advance.

The final purpose of the model is to predict water-level evolution when water-level rise may produce floods. To assess the performance of the model in the prediction of the larger values of discharge, the absolute and relative error were calculated for the increasing value of water level. Calculations are summarized in Fig. 7. Errors in discharge values for each value of water level, h, were calculated considering all the forecast values greater than h. Higher water levels and associated errors correspond to floods. For the 1-h ahead prediction (Fig. 7(a)), it was found that for each water level the absolute error in discharge increases with the stage. The relative percentage error in discharge is about 5% for the calibration set, and about 7%, on average, for the

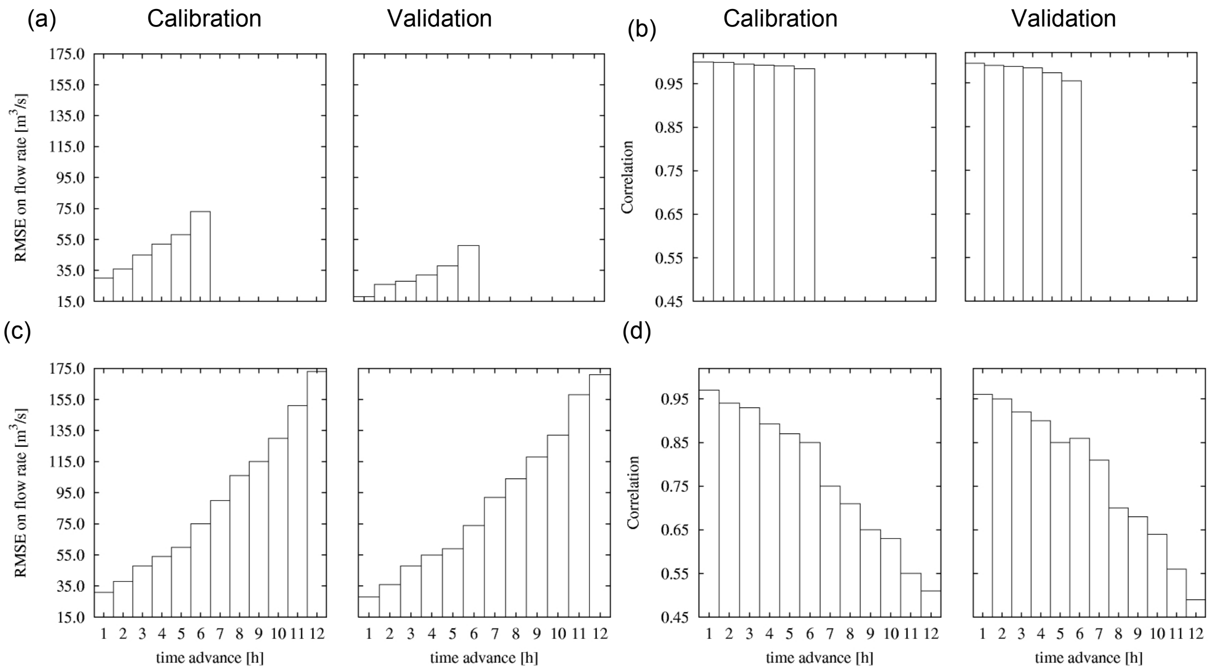

Fig. 6 RMSE and correlation coefficient for increasing time advance of prediction for ANN model predicting water-level evolution: (a) and (b) 6 h ahead; (c) and (d) 12 h ahead. Values for calibration and validation set are calculated on flow rates.

validation set. The error increases with the time ahead of prediction (see Fig. 7(b) and (c)—up to 15%, on average, for calibration and validation for the 6-h ahead forecast), but the increase is more pronounced for the lower levels than for the higher ones. This behaviour is regular for the calibration set and more irregular for the validation set. The reduced increase in the percentage error for $h \geq 5$ m up to the 6-h ahead prediction (about 7%) suggests that the model is suitable for flood forecasting, and less accurate for medium-flow predictions. The stability of model performance for increasing lead time of prediction when the predicted water level is very high deserves some physical explanation. In Fig. 8, two different peaks of discharge with the same maximum value are presented. As shown in the upper part of the figure, these are generated by two different time series of rainfall. The peak on the left is generated by rainfall of medium intensity lasting for a time period of 10 h and the similar peak on the right is generated by a shorter rainfall event of higher intensity. Predictions for the peaks in the two cases are shown in the lower part of Fig. 8: the peak (a) is correctly reproduced even for the larger lead time, whereas a shift can be observed in the prediction of the peak (b). The time of the rising limb is incorrect, even though the peak value is captured. This suggests that the main errors in prediction are found for intermediate values of water level because, in these conditions, the rising limb can be generated either by rainfalls of medium intensity and long duration (that are well reproduced) or by heavy rainfalls of shorter duration (that are not well reproduced). Considering the relative frequency of these events in the data set, it was found that more frequent events are better reproduced than less frequent events. On the other hand, peaks corresponding to larger values of discharge (and water level) are always generated by rainfalls that are heavy and of long duration, and this explains the reduced error.

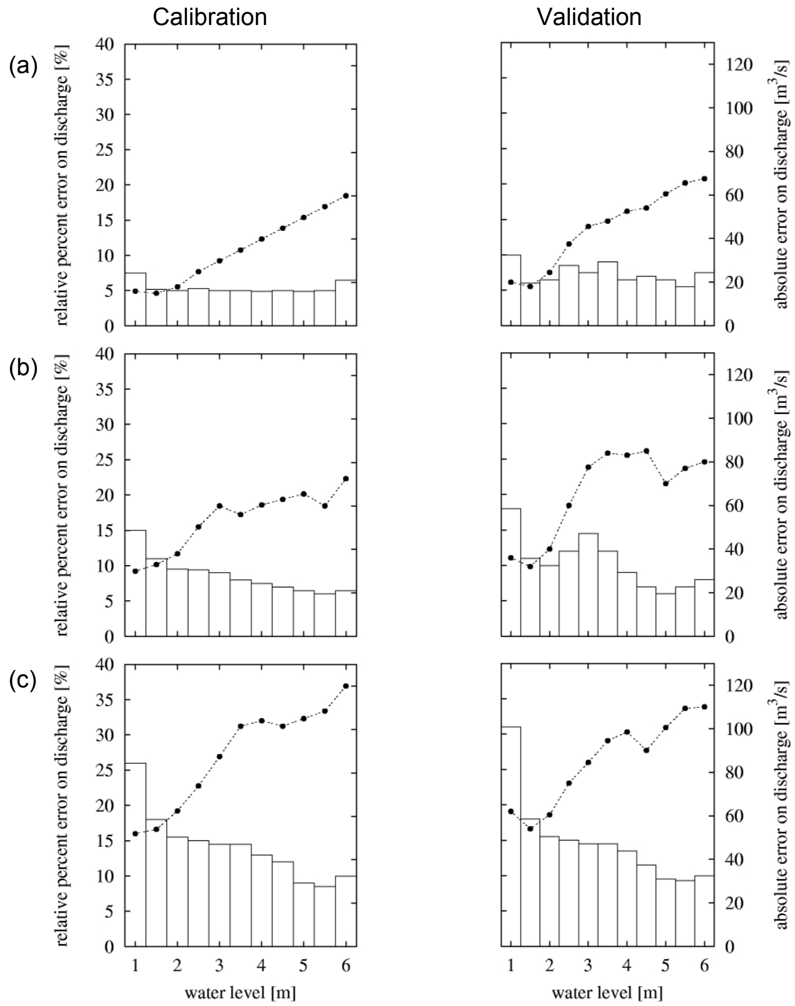

Fig. 7 Absolute and relative percentage error on discharge for increasing water-level: values calculated for calibration (left) and validation (right) for (a) 1-, (b) 3- and (c) 6-h ahead prediction.

Finally, the possibility was considered of improving network performances on the 6-h ahead prediction by using the same input data and a more specific network, with a single output. This is the configuration generally used in the literature to produce delayed forecasts (Maier & Dandy, 1996, Dawson & Wilby, 1998). A similar network can be obtained from the already trained model eliminating all the weights from the hidden layer to the lost output nodes. Nevertheless, the purpose here is to verify if embedding of the output time series into the model outputs increases network performances, as suggested by Cheng *et al.* (1995). To compare single-output and multiple-output models, a new neural network model was trained with the same number of hidden nodes as before (four epochs of 1000 iterations each, with a learning rate decreasing from 0.2 to 0.05). It was found that performance of this simplified network was slightly worse than of the multiple-output network, with *RMSE* of 90 and 65 $m^3 s^{-1}$ and correlation coefficients of 0.94 and 0.84 for the calibration and validation sets, respectively. Figure 9 shows the percentage error computed for the single-output net on the calibration and the validation sets, that can be compared with Fig. 7(c). The

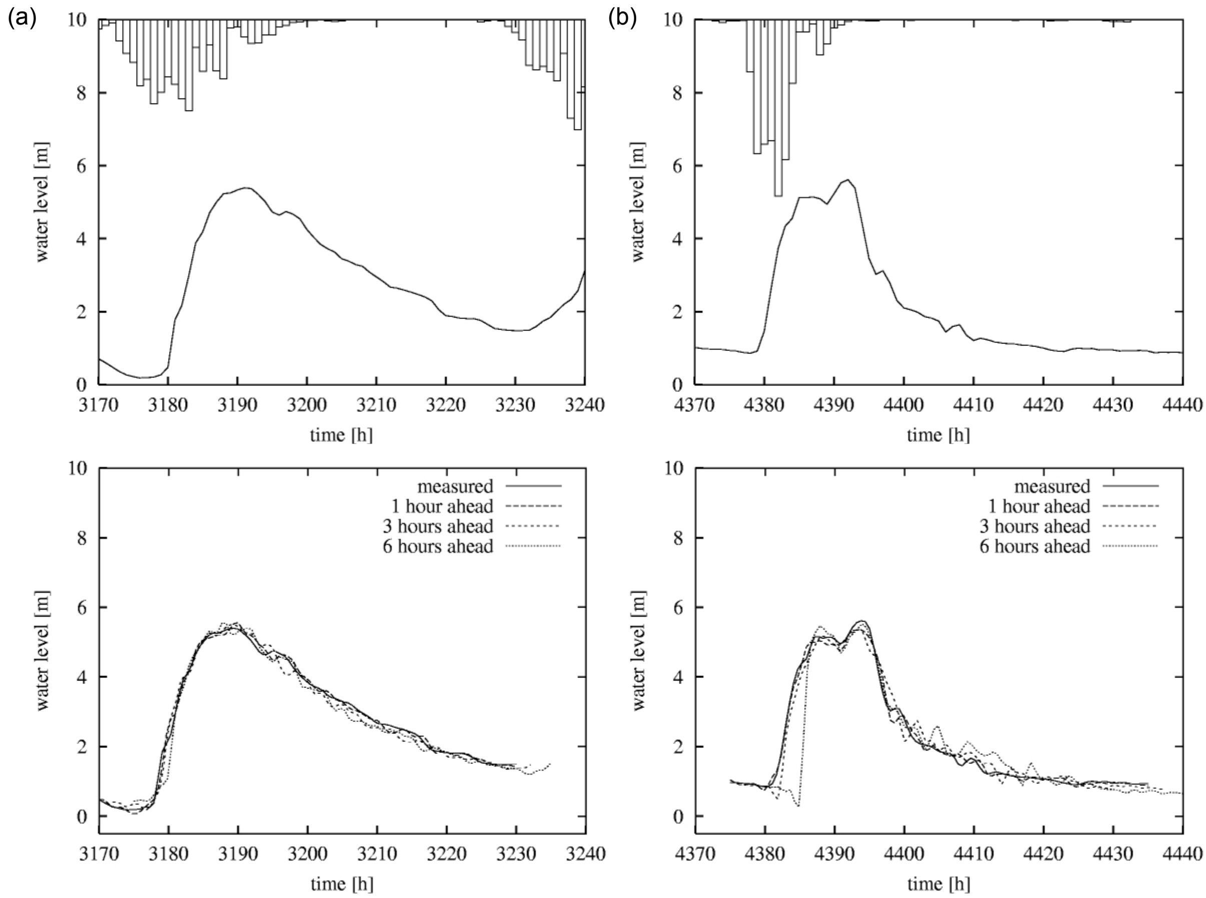

Fig. 8 Water-level variations produced by different time series of rainfall. (a) Peaks generated by long duration rainfall are reproduced correctly even for the larger time advance. (b) Peaks generated by rainfall of short duration are not correctly reproduced.

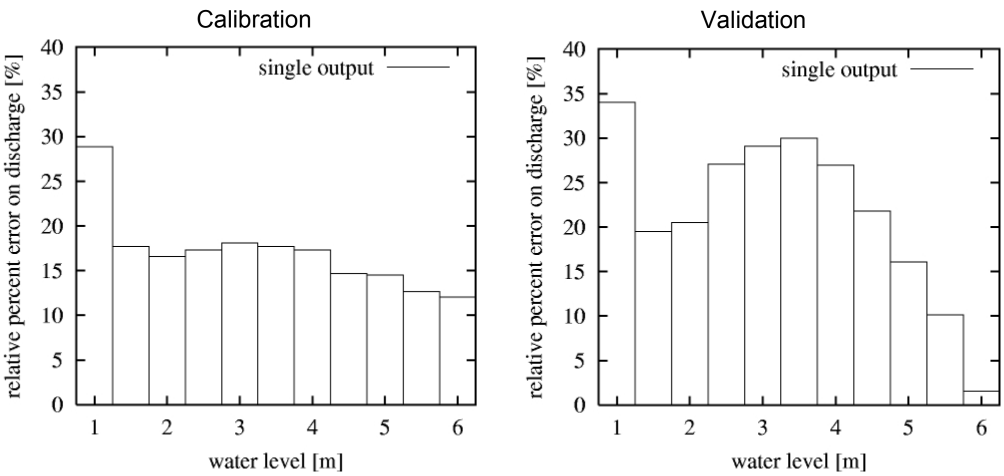

Fig. 9 Relative percentage error on discharge for increasing water level for single-output neural model. Performances are slightly poorer than for six-output model.

error is larger for each class of water level for both the calibration and the validation sets. A reduced value is found only for $h \geq 6$ m in the validation set.

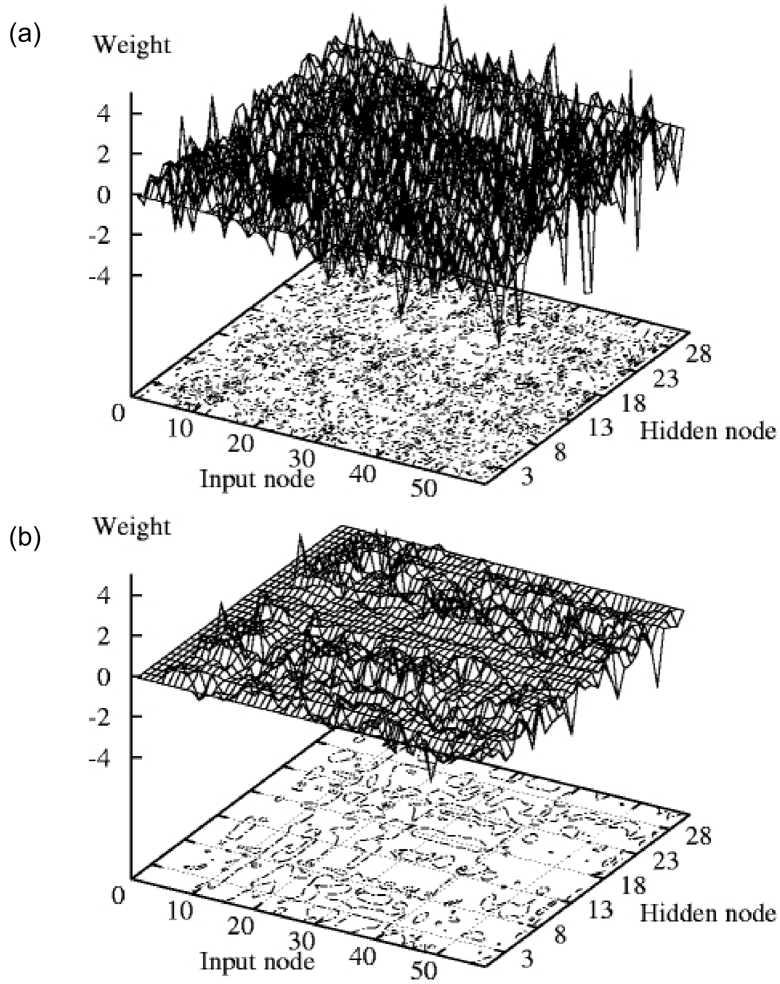

Fig. 10 Distribution of weights from input to hidden nodes for (a) the multiple output model, and (b) the single output model.

In order to understand the effect of the output selection (single value or segment of water-level time series) on model calibration, the map of the weights of the trained nets was studied. In Fig. 10 the values of the weights for each link between input and hidden nodes are plotted. The distribution of weights is extremely irregular for the multiple-output model, while the surface is only slightly deformed for the single-output model. Furthermore, a large number of weights is near to zero in the single-output case (Fig. 10(b)), indicating that input information from the corresponding input is neglected. Obviously, the transfer function to reproduce the water-level evolution for the next six hours is more complex than the one reproducing a single future value of water level and a less expensive model may be used to predict the water level 6 h in advance. Nevertheless, for flood forecasting applications the water-level evolution is as important as the peak value, and for this reason the authors prefer to use a multiple-output model.

CONCLUSIONS

A neural network model predicting water-level variation up to 6 h in advance is presented and discussed. The model, built using data collected on the basin of the

River Arno, has the advantages of both low cost and simplicity and can be easily integrated with automatic data acquisition systems into a real-time flood forecasting system. A methodology is described for the selection of model inputs based on analytical procedures (cross-correlation analysis) and hydrological expertise, rather than on extensive model sensitivity analysis to input data. This procedure allows the identification and calibration of the model to be speeded up because the relevant rainfall, power production, and water-level data to be used for prediction and the possible limit to the prediction lead time can be identified in advance. Performance degradation of the model for the 6-h ahead prediction confirms the limitations in the prediction time advance. Furthermore, it was found that the model predicts satisfactorily the evolution of water levels during floods. The percentage error in flow rate is less than 7% for the 6-h ahead prediction when the water level is higher than 5.0 m. Intermediate peaks are always predicted in magnitude, but a timing errors exists in the prediction when the peak is generated by heavy rainfall of short duration. Finally, performances were compared of the 6-h ahead forecast with the help of a multiple-output model, predicting hourly data for the entire time horizon, and a single-output model, making only the 6-h ahead forecast. It was found that predicting the entire time horizon gives more useful information, while the accuracy of predictions is not significantly improved by a single output model.

Acknowledgements The authors thank the Ufficio Idrografico e Mareografico of Pisa, Italy, for making available the data for this work, and the National Power Agency (ENEL, Divisione Produzione, Direzione Produzione Idroelettrica Alto Appennino), Rome, Italy for the power production data.

REFERENCES

Bishop, C. M. (1994) Neural network and their applications. *Rev. Sci. Instrum.* **65**, 1803–1832.

Bishop, C. M. (1995) *Neural Network for Pattern Recognition.* Oxford University Press, Oxford, UK.

Campolo, M., Andreussi, P. & Soldati, A. (1999a) River flood forecasting with a neural network model. *Water Resour. Res.* **35**, 1191–1197.

Campolo, M., Soldati, A. & Andreussi, P. (1999b) Forecasting river flow rate during low-flow periods using neural networks. *Water Resour. Res.* **35**, 3547–3552.

Campolo, M., Andreussi, P. & Soldati, A. (2002) Water quality control in the River Arno. *Water Res.* **36**, 2673–2680.

Cheng, Y., Karjala, T. W. & Himmelblau, D. M. (1995) Identification of non linear dynamic processes with unknown and variable dead time using an internal recurrent neural network. *Indust. Engng Chem. Res.* **34**, 1735–1742.

Dawson, C. W. & Wilby, R. (1998) An artificial neural network approach to rainfall–runoff modelling. *Hydrol. Sci. J.* **43**(1), 47–66.

Fernando, D. A. K. & Jayawardena, A. W. (1998) Runoff forecasting using RBF networks with OLS algorithm. *J. Hydrol. Engng* **3**, 203–209.

Haykin, S. (1995) *Neural Networks: A Comprehensive Foundation.* Macmillan College Publishing Company, New York, USA.

Hsu Kuo-Lin, Gupta, H. V. & Sorooshian, S. (1995) Artificial neural network modeling of the rainfall–runoff process. *Water Resour. Res.* **31**, 2517–2530.

Hu, T. S., Lam, K. C. & Ng, S. T. (2002) River flow time series prediction with a range-dependent neural network. *Hydrol. Sci. J.* **46**(5), 729–745.

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

Kim, G. & Barros, A. P. (2001) Quantitative flood forecasting using multisensor data and neural networks. *J. Hydrol.* **246**, 45–62.

Liu, Q., Islam, S., Rodriguez-Iturbe, I. & Le, Y. (1998) Phase space analysis of daily streamflow: characterization and prediction. *Adv. Water Resour.* **21**, 463–475.

Maier, H. R. & Dandy, G. C. (1996) The use of artificial neural networks for the prediction of water quality parameters. *Water Resour. Res.* **32**, 1013–1022.

Maier, H. R. & Dandy, G. C. (1997) Determining inputs for neural network models of multivariate time series. *Microcomputers in Civil Engng.* **12**, 353–368.

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. *Environ. Modell. Software* **15**, 101–124.

Minns, A. W. & Hall, M. J. (1996) Artificial neural networks as rainfall–runoff models. *Hydrol. Sci. J.* **41**(3), 399–417.

Nalbantis, I. (2000) Real-time flood forecasting with the use of inadequate data. *Hydrol. Sci. J.* **45**(2), 269–284.

Porporato, A. & Ridolfi, L. (1997) Nonlinear analysis of river flow time sequences. *Water Resour. Res.* **33**, 1353–1367.

Porporato, A. & Ridolfi, L. (2001) Multivariate non linear prediction of river flows. *J. Hydrol.* **248**, 109–122.

Rajurkar, M. P., Kothyari, U. C. & Chaube, U. C. (2002) Artificial neural networks for daily rainfall–runoff modelling. *Hydrol. Sci. J.* **47**(6), 865–877.

Rogers, L. L. & Dowla, F. U. (1994) Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. *Water Resour. Res.* **30**, 457–481.

Shamseldin, A. Y. (1997) Application of a neural network technique to rainfall–runoff modeling. *J. Hydrol.* **199**, 272–294.

Sivakumar, B., Berndtsson, R., Olsson, J. & Jinno, K. (2001a) Evidence of chaos in the rainfall–runoff process. *Hydrol. Sci. J.* **46**(1), 131–146.

Sivakumar, B., Berndtsson, R. & Persson, M. (2001) Monthly runoff prediction using phase space reconstruction. *Hydrol. Sci. J.* **46**(3), 377–387.

Smith, J. & Eli, R. N. (1995) Neural network models of rainfall–runoff process. *J. Water Resour. Plan. Manage.* **121**, 499–508.

SNNS (1995) *User Manual Version 4.0.* Report no. 6/95, University of Stuttgart, Institute for Parallel and Distributed High Performance Systems, Stuttgart, Germany.

Thirumalaiah, K. & Deo, M. C. (1998) River stage forecasting using artificial neural networks. *J. Hydrol. Engng* **3**, 26–32.

WMO (1992) *Simulated Real Time Intercomparison of Hydrological Models.* Operational Hydrology Report no. 38, World Meteorological Organization, Geneva, Switzerland.

Xiong, Lihua & O'Connor, K. M.(2002) Comparison of four updating models for real-time river flow forecasting. *Hydrol. Sci. J.* **47**(4), 621–639.

Received 8 April 2002; accepted 21 February 2003