

Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting

E. Toth

University of Bologna, Faculty of Engineering, viale Risorgimento 2, 40136 Bologna, Italy

Received: 14 January 2009 – Published in Hydrol. Earth Syst. Sci. Discuss.: 23 February 2009

Revised: 17 July 2009 – Accepted: 20 July 2009 – Published: 3 September 2009

Abstract. This paper presents the application of a modular approach for real-time streamflow forecasting that uses different system-theoretic rainfall-runoff models according to the situation characterising the forecast instant. For each forecast instant, a specific model is applied, parameterised on the basis of the data of the similar hydrological and meteorological conditions observed in the past. In particular, the hydro-meteorological conditions are here classified with a clustering technique based on Self-Organising Maps (SOM) and, in correspondence of each specific case, different feed-forward artificial neural networks issue the streamflow forecasts one to six hours ahead, for a mid-sized case study watershed. The SOM method allows a consistent identification of the different parts of the hydrograph, representing current and near-future hydrological conditions, on the basis of the most relevant information available in the forecast instant, that is, the last values of streamflow and areal-averaged rainfall. The results show that an adequate distinction of the hydro-meteorological conditions characterising the basin, hence including additional knowledge on the forthcoming dominant hydrological processes, may considerably improve the rainfall-runoff modelling performance.

1 Introduction

Metric (or system-theoretic) and hybrid metric-conceptual (see Wheater et al., 1993) models have always represented a natural candidate for online forecasting of the rainfall-runoff transformation (WMO, 1992; Young, 2002), since the real-time framework gives more importance to the simplicity and robustness of the model implementation rather than to an accurate description of the various internal sub-processes.

System-theoretic models are *data-driven* models, since they are based primarily on observations, and seek to characterise the system response from extensive records of past input and output variables. They are, therefore, particularly sensitive to the set of data used for their calibration, which must be suitable for inferring an adequate input-output relationship. On the other hand, also the use of physically-based approaches cannot, yet, overcome the need to calibrate at least a part of the model parameters, so that the significance of calibration data is crucial in any kind of rainfall-runoff transformation model.

The significance of the data belonging to a particular period, and therefore the reliability of a model parameterised on that data set, are strictly linked to the hydrological processes taking place in the period. Such processes are in fact strongly variable in time: the physical phenomena governing the streamflow generation at the beginning of a storm are certainly extremely different from those dominating the falling limb of the same flood hydrograph, and even further from those responsible for the low flows.

This constation is at the basis of the formulation of event-based models, that are explicitly aimed at modelling only the processes that are dominant during flood events. But the same consideration guides the calibration procedure of continuously-simulating models where the hydrologist has to choose which part of the observed hydrograph is most important to fit, either implicitly, through the visual agreement in manual calibration, or explicitly, through the choice of the objective function(s). Changing the objective functions it is in fact possible to emphasise different kind of errors, giving them more weight in the calibration phase, for example with functions distinguishing high and low flows (Coulibaly et al., 2001; de Vos and Rientjes, 2007), or with peak, time-to-peak or volume errors in case of flood events. In order to adequately reproduce more aspects of the observed hydrograph, multi-objective calibration algorithms have been successfully developed in the recent years (e.g., Gupta et al.,



Correspondence to: E. Toth
(elena.toth@unibo.it)

1998; Madsen, 2000; Vrugt et al., 2003; Tang et al., 2006; de Vos and Rientjes, 2007), with the aim of helping the hydrologists to choose an optimum (even if always subjective) trade-off.

A different line will be followed in this study, consisting in the implementation of multiple models, that is, a modular approach diversifying the rainfall-runoff models on the basis of the specific hydro-meteorological situation presenting itself in each forecast instant.

The hydrological and meteorological conditions characterising the instant in which the forecast is issued determine in fact which hydrological processes will be dominant in the following period. The future evolution of the streamflow values is therefore simulated with a different model for each forecast instant, chosen in function of the hydro-meteorological situation and parameterised on the basis of the evolution of the similar situations observed in the past. This approach is particularly suitable for system-theoretic, data-driven models: in this work, multi-layer Artificial Neural Networks (ANN) will be used, where there is no explicit a priori representation of the known physical processes and the models are set up exclusively on the basis of the available data.

The identification of the different hydro-meteorological conditions corresponding to each forecast instant will be done with a classification technique based on the use of Self-Organising Maps (SOMs, Kohonen, 1982, 2001). The SOMs were originally used principally for signal recognition, organization of large collections of data and information processing, but they are now acknowledged as a powerful clustering technique (Mangiameli et al., 1996; Astel et al., 2007) and have been recently used also in a variety of water resources studies (see Kalteh et al., 2008, for an exhaustive review). The main advantages of the SOM clustering algorithm are that it is non-linear and it has an ability to preserve the topological structure of the data (ASCE Task Committee, 2000a), thus allowing also an evaluation of the affinity between the clusters, as explained in the following.

In the present work, a modular approach is applied on a mid-sized watershed, the Sieve River, for issuing multi-step ahead streamflow forecasts referring to long continuous periods (thus including a variety of flow conditions) of calibration and validation data. A SOM is first used for clustering the vectors characterising each forecast instant: such vectors are formed not only by the antecedent streamflow, but also by past precipitation values, given the importance of meteorological forcing in the evolution of future flows. Secondly, rainfall-runoff models based on multi-layer Artificial Neural Networks are parameterised accordingly to each specific hydro-meteorological condition, for issuing multi-step ahead forecasts.

2 The Sieve River case study

The case study herein considered is referred to the Sieve River basin, a first tributary of the Arno River, located on the Apennines Mountains in Tuscany, North-Central Italy. The Sieve River basin is elongated in shape and the drainage area is around 830 km² at the outlet section of Fornacina, where the time of concentration is approximately 10 h. The watershed is morphologically characterised by moderate to strong relief in the upper and lower sections and by a gently rolling plain in the central part. Except in the valleys, dedicated to agriculture, the terrain is forested and mountainous. The fact that Mediterranean water is warmer than Atlantic water throughout the year and the presence of island barriers in the Mediterranean serve as preconditions for strong cyclogenesis causing most rainfall over the Sieve River between late Fall and early Spring, November being the wettest month. The summer months, especially July, are the driest, owing to the dominance of the Azores high-pressure cell.

At the closure section, hourly discharge observations [m³/s] were collected between 1 January 1992 and 31 December 1996. For the same observation period, hourly rainfall depths [mm] at 12 raingauges are available, thus allowing the computation of the average areal precipitation over the watershed with an inverse squared distance weighting of the raingauges observations. The calibration procedures described in the following are based on the continuous data belonging to the first three hydrological years of the observation period, from 1 September 1992 to 31 August 1995. The last 16 months, from 1 September 1995 to 31 December 1996, are used for validation purposes. The main statistics of the streamflow values for the calibration and validation periods are shown in Table 1. The mean and the percentiles are similar, but the variability of the calibration streamflow data is more pronounced than that of the validation data: in particular at the beginning of the calibration period, in autumn 1992, occurred the major events (high but not exceptional: two peaks with a return period of about 5 years) of the observation period. This is not a drawback for split-sample calibration experiments, but quite the opposite: in fact, the calibration period must have enough information contents, including a wide range of hydrological conditions and in particular it is useful that it includes the highest output values, due to the difficulties ANNs may experience in extrapolation (see, e.g., De Vos and Rientjes, 2008).

For each time instant, six forecasts will be issued, corresponding to a lead-time varying from one to six hours.

Table 1. Statistics of the streamflow observation data sets.

	Number of data	Mean (m ³ /s)	Coeff. of variation	75% percentile (m ³ /s)	95% percentile (m ³ /s)
Calibration period	26 280	14.84	2.69	13.78	47.38
Validation period	11 688	12.11	2.26	13.73	45.72

Forecasting performance measures

The performances of the streamflow forecasting models will be evaluated by the Nash-Sutcliffe efficiency,

$$E_L = 1 - \frac{\sum_{t=1, N} [Q_{\text{obs}}(t+L) - Q_{\text{sim}}(t+L)]^2}{\sum_{t=1, N} [Q_{\text{obs}}(t+L) - \mu_{\text{obs}}]^2}, \quad L=1 \div 6, \quad (1)$$

and through an error measure in the same units of the simulated variable (as suggested also by Legates and McCabe, 1999), namely the mean absolute error,

$$\text{MAE}_L = \frac{\sum_{t=1, N} |Q_{\text{obs}}(t+L) - Q_{\text{sim}}(t+L)|}{N} \quad (2)$$

where t is the forecast instant, Q_{obs} and Q_{sim} are the observed and simulated streamflow, respectively, μ_{obs} is the mean value of Q_{obs} , N is the total number of forecast instants and L is the lead-time, varying from one to six hours in the present study.

The efficiency coefficient varies in the range $]-\infty, 1]$, where 1 indicates a perfect agreement and negative values mean that the forecast is worse than assuming future occurrences equal to the mean value μ_{obs} . The meaningful value of zero provides a convenient reference point to compare the model with the predictive abilities of the observed mean, but the efficiency coefficient, like all the squared measures, tends to inflate the highest errors, that generally correspond to the highest flows. The MAE, on the contrary, gives the same weight to all errors and it is more significant for comparing the forecasting performances over average and low flow regimes.

As an additional benchmark, the forecasting models will be compared also with a naïve persistent model, where future streamflow is supposed to be equal to the last observed value over all the lead-times:

$$Q_{\text{pers}}(t+L) = Q_{\text{obs}}(t), \quad \forall L. \quad (3)$$

3 Artificial neural networks for streamflow forecasting

The appeal of the use of Artificial Neural Networks (ANNs) as hydrological models lies mainly in their capability to flexibly and rapidly reproduce the highly non-linear nature of the

relationship between input and output variables, and it is certainly worthy considering ANN models as powerful tools for real-time short-term runoff forecasts.

An extensive review of the potentiality of ANNs in hydrological modeling was given, for example, by the ASCE Task Committee (2000b) and by Maier and Dandy (2000). In the majority of the applications of river flow prediction, the networks are fed by both past flows and past precipitation observations: extremely encouraging results have been obtained in literature on both real and synthetic rainfall-runoff data (among the many others, in the recent years: Cameron et al., 2002; Solomatine and Dulal, 2003; Jain et al., 2004; Khan and Coulibaly, 2006; Shamseldin et al., 2007; Srivastav et al., 2007). Despite the importance of calibration information in a data-driven technique, little attention has been paid, so far, to the influence that the calibration period has on the forecasting performances of ANN rainfall-runoff modeling. Even if it is acknowledged that the choice of the training set has a fundamental weight (see, for instance, Minns and Hall, 1996; Campolo et al., 1999), only a few studies have presented, so far, an analysis of the impact of the use of different training data sets on ANN performances in validation (e.g., Dawson and Wilby, 1998; Anctil et al., 2004; Toth and Brath, 2007). In the proposed approach, different calibration data sets are identified, to be used specifically for modelling the future evolution of similar data.

ANNs distribute computations to processing units called neurons, grouped in layers and densely interconnected. In the supervised feed-forward multilayer networks, three different layer types can be distinguished: an input layer, connecting the input information to the network (and not carrying out any computation), one or more hidden layers, acting as intermediate computational layers, and an output layer, producing the final output.

In correspondence of a computational node J , each one of the N_j entering values (I_i) is multiplied by a connection weight (w_{ij}). Such products are then all summed with a neuron-specific parameter, called bias (b_j), used to scale the sum of products into a useful range. The computational node finally applies an activation function (f) to the above sum producing the node output (O_J):

$$O_J = f\left(\sum_{i=1}^{N_j} w_{ij} I_i + b_j\right). \quad (4)$$

The ANNs applied in the present work have only one hidden layer: tan-sigmoidal activation functions were chosen for the hidden layer and linear transfer functions for the output layer.

Weights and biases are determined by means of the quasi-Newton Levenberg-Marquardt BackPropagation optimisation procedure (Hagan and Menhaj, 1994), minimising a learning function expressing the closeness between observations and ANN outputs, in the present case the mean squared error. To mitigate overfitting and to improve generalization, a Bayesian regularization of the learning function (Foresee and Hagan, 1997; Antil et al., 2004) was applied.

For each lead-time, a distinct mono-output network will be implemented: the output of each network, $Q_{\text{sim}}(t+L)$, is the streamflow forecast issued, in the forecast instant t , for each lead-time L .

The input data consist of the most relevant information that is generally available in a real-time flow forecasting system, namely, past rainfall and streamflow observations.

The optimal number of input nodes (corresponding to past streamflow and mean areal precipitation values) and of hidden nodes to be included in the network is strongly case-dependent. The number of input nodes may be obtained either with a model-free approach, using statistical measures of dependence (such as correlation or mutual information) to determine the strength of the relationship between candidate model inputs and the model output, prior to model specification and calibration (e.g., Solomatine and Dulal, 2003; Bowden et al., 2005; Fernando et al., 2009), or with a model based approach, that analyses the performance of models that are calibrated with different inputs, for choosing the most appropriate input vector.

In the present work, a model-based approach was used for identifying the dimension of both the input and the hidden layers: the investigation of the performances of several combinations of input and hidden layers dimensions was performed (through a trial-and-error procedure based on a “forward selection method”, consisting in beginning by selecting a small number of neurons and then increasing it) in past researches on the same study watershed (partly reported in Toth and Brath, 2007) and will not be described here for sake of brevity. The architecture providing the best trade-off between parsimony and forecasting performances was the one feeding to the input layer four streamflow and three precipitation values preceding the forecast instant t , $Q_{\text{obs}}(t-3)$, $Q_{\text{obs}}(t-2)$, $Q_{\text{obs}}(t-1)$, $Q_{\text{obs}}(t)$, $P(t-2)$, $P(t-1)$, $P(t)$, with three nodes in the hidden layer and one output node $Q_{\text{sim}}(t+L)$. It was examined the possibility to implement a different architecture for each network, that is for each lead-time L , but the validation results showed, for each L , an analogous behaviour when varying the dimension of the layers.

4 Multi-network modeling

Extremely different methods for combining the river flow forecasts issued by a set of different rainfall-runoff models have been recently proposed in the literature, for example by Shamseldin et al. (1997, 2002, 2007), Abrahart and See (2002), Georgakakos et al. (2004), Solomatine and Siek (2006). This work, in particular, presents an implementation of multiple, alternative models, that is, a modular approach that uses different, specialised rainfall-runoff models, chosen on the basis of the specific hydro-meteorological situation presenting itself in each forecast instant.

Modular neural networks (or *multi-network* models) for streamflow forecasting have been successfully applied in the hydrological literature in the most recent years: interesting applications, considering different input variables and different methods for identifying the model appropriate to each case, have been presented with the objective of forecasting future streamflow at extremely variable time-scales (from hourly to monthly). Furundzic (1998) was the first to propose a multi-network approach with decomposition of the modelling domain in a study on the relevancy of input variables. Zhang and Govindaraju (2000) introduced a modular architecture where different modules within the network were trained to learn subsets of the input space in an expert fashion: a gating network was used to mediate the responses of all the experts and the model was applied for forecasting monthly runoff values. A hybrid structure of Artificial Neural Networks, SORB, was proposed by Moradkhani et al. (2004): the architecture employed consisted of a Self-Organising Map (SOM) as an unsupervised training scheme for data clustering, which correspondingly provided the parameters required for the Gaussian functions in a Radial Basis Function (RBF) neural network. Such scheme was inspired by the Self Organizing Linear Output mapping (SOLO) proposed by Hsu et al. (2002): SOLO classifies the input information using a SOM and then maps the inputs into the outputs using multivariate linear regression. Parasuraman and Elshorbagy (2007) clustered the data set in two groups with a K-means algorithm before applying two different networks for forecasting monthly runoff values, obtaining a better reproduction of the dynamics of high flows. Gopakumar et al. (2007) used Self-Organising Maps (SOMs) for identifying a seasonal pattern classifying the monthly rainfall and runoff values: subsequently an ANN was developed for daily flow forecasting using only the data of the identified rainy season. A pioneer work that proposed clustering algorithms for grouping high-resolution streamflow data (at hourly time scale), thus explicitly decomposing the hydrograph in separate parts, for ANN multi-network modelling, is that by Abrahart and See (2000): they implemented two separate ANN models, specifically developed for two rising limbs clusters. In their work, as also in the one by Wang et al. (2006), the classification was based on past river flow only. However, information on the recent

precipitation depths is extremely valuable for the identification of the streamflow evolution: in the period immediately following the forecast, a rising limb, for example, will keep increasing or will reach the peak and begin to decrease depending if the rainfall is continuing or if it has already stopped. Jain and Srinivasulu (2006) used both rainfall and flow values for decomposing the flow hydrograph and then forecasting one-step ahead daily streamflow with a multi-network approach: the decomposition was performed with methods based on physical concepts and with a small SOM network, which classified the flows in low, medium and high ranges. Corzo and Solomatine (2007) applied a modular architecture based on the distinction of baseflow and excess flow obtained with i) a K-means clustering algorithm, ii) a semi-empirical constant slope method or iii) filtering algorithms of the hydrographs (where i) and ii) are again based on past flows only).

The principal difference with the above cited works is that the objective of this study is to forecast the future hourly streamflow not only one-step ahead but for increasing lead-times: to do so it is crucial to identify the conditions of each forecast instant not only in terms of past streamflow data but also of past rainfall data, given the importance of meteorological forcing in the evolution of future, farther flows. It is therefore applied a classification algorithm that is based on both past streamflow and rainfall values, rather than an algorithm that performs a separation of the hydrograph in different rising and falling limbs, based on past streamflow data alone. The present work will thoroughly explore the potential of SOMs for identifying the different meteorological and hydrological conditions of each forecasting instant, and therefore the future dominant hydrological processes, for improving streamflow forecasts over lead-times from one to six hours.

5 Classification of hydro-meteorological conditions

There are no predefined classes of the conditions characterising the watershed in each forecast instant: a clustering algorithm is here used as an *unsupervised* classifier, where the task is to learn a classification from the data. Such partitioning will be based on the most relevant available information, that is, past rainfall and flow observations, assuming that such variables are able to characterise both the current situation and its near-future evolution. It is important to underline that the combination of rainfall and streamflow observations prior to the forecast contains valuable information on the state of saturation of the basin and hence on its capability to respond to recent and current rainfall perturbation. The vector chosen for representing each forecast instant is therefore the same that will be provided in input to the multilayer feedforward ANNs modelling the future streamflow values.

The classification is based on the use of a SOM (Self Organised Map), which organises the data according to their similarity.

5.1 Self Organising Maps

Self Organising Maps (SOMs), or Kohonen networks (Kohonen, 1982, 2001), are artificial neural networks of the unsupervised type: as opposite to supervised networks (like the multilayer networks introduced in Sect. 3 for rainfall-runoff modelling) there is no known user-defined target that the output vector should reproduce: the desired solutions are not given and the network learns to cluster the input data by recognizing different patterns. Unsupervised networks may be viewed as classifiers, where the classes are the clusters that are discovered in the calibration data and new data, such as those of the validation set, may be successively assigned to the same classes.

A SOM is formed by only two layers of nodes: the input layer contains a node for each of the n variables characterising the unit to classify and the output layer is an array, generally two-dimensional for the convenience of visual understanding, whose nodes are connected, by weighted connections, to the input layer. Each input vector “activates” only one output node (the *winning* node, that will represent its class), using the Kohonen competitive learning rule.

Initially the weights are randomly assigned. When the n -dimensional input vector (\underline{x}) is sent through the network, each neuron of the network computes a distance measure: a Euclidean distance was here chosen, as in the majority of SOM applications, between the weight (\underline{W}) and the input:

$$\|\underline{x} - \underline{W}\| = \sqrt{\sum_{i=1}^n (x_i - W_i)^2}. \quad (5)$$

The neuron responding maximally to the given input vector, that is the weight vector having the minimum distance from the input vector, is chosen to be the winning neuron. The winning neuron and its neighbouring neurons are allowed to learn by changing the weights at each training iteration t , in a manner to further reduce the distance between the weights and the input vector:

$$\underline{W}(t+1) = \underline{W}(t) + \alpha(t)h_{lm}(\underline{x} - \underline{W}(t)), \quad (6)$$

where α is the learning rate, $\in[0, 1]$, l and m are the positions of the winning and its neighbouring output nodes and h_{lm} is the neighbourhood shape, that reduces the adjustment for increasing distance:

$$h_{lm} = \exp\left(-\frac{\|l-m\|^2}{2\sigma(t)^2}\right), \quad (7)$$

where $\|l-m\|$ is the lateral distance between l and m on the output grid and σ is the width of the topological neighbourhood.

The weights of the SOM nodes are adjusted, through the learning process, on the vectors of the calibration set. In the learning process all the calibration input vectors are processed through the SOM incrementally, one after the other, re-iteratively: for each sample input vector \underline{x} , the weights of the winner node and of the nodes in its neighbourhood are changed closer to \underline{x} .

Lateral interaction between neighbouring output nodes ensures that learning is a topology-preserving process in which the network adapts to respond in different locations of the output layer for inputs that differ, while similar input patterns activate units that are close together. In this way, a SOM produces a topologically ordered output that displays the similarity between the samples presented to it (Foody, 1999).

At the end of the learning phase, the SOM is used (without changing the weights any more) to classify the calibration vectors: the trained network identifies which output node to activate in correspondence of each input vector and all the input vectors that activate the same node belong to the same class. In exactly the same way, the tuned SOM may be used to associate any new vector, such as those of the validation set, to one of the units of the SOM output layer, thus attributing the new data to the clusters identified before.

5.2 SOM-clustering of the hydro-meteorological conditions

The use of a SOM in the proposed research activity entails the association of each of the input variables defining the current hydro-meteorological condition, $Q_{\text{obs}}(t-3)$, $Q_{\text{obs}}(t-2)$, $Q_{\text{obs}}(t-1)$, $Q_{\text{obs}}(t)$, $P(t-2)$, $P(t-1)$, $P(t)$, to an input node. In the classification phase, such values are standardised to have mean equal to zero and variance equal to one, in order to give them the same importance in the distance measure.

There is not a predefined number of possible conditions and it was chosen to have an output layer formed by three rows by three columns, for a total of nine nodes, each one corresponding to a class, believing that such number is sufficient for representing a variety of hydro-meteorological conditions without preventing their following interpretation. The output layer topology is hexagonal, rather than rectangular, so that diagonal neighbours have the same distance as horizontal and vertical neighbours, as suggested by Kohonen himself (Kohonen, 2001) and by several works on SOM clustering (e.g., Van der Voort et al., 1996; Hsu and Halgamuge, 2003; Shirazi and Menhaj, 2005). The trained network will indicate, for any input vector, the class of the matching forecast instant, along with the affinity with other classes.

The SOM was initially applied to the calibration set, that is, to the first three hydrological years of the observation period, from 1 September 1992 to 31 August 1995. The vectors characterising each one of the instants of such period, for a total of 26 280 records, were iteratively given in input to the

SOM: at the end of the tuning phase, these vectors were classified in nine homogeneous groups, formed by all the vectors resulting assigned to the same node on the output layer.

The hexagonal output layer is shown in Fig. 1, using markers that have similar colour and/or shape for the neighbouring nodes. The figure displays also a part of the observed hydrograph where, at each time t , representing each instant in which a forecast (or better, six forecasts for the varying lead-times) will be issued, the flow value is indicated by a marker having the colour and shape of the class to which the forecast instant is assigned. It is therefore possible to visualise which parts of the hydrograph are associated to the different classes. It should be noted that the hydro-meteorological condition, that is the class, of each forecast instant is the same, independent of the lead-time that will be successively considered for the forecast.

It may be observed in Fig. 1 that classes 1 and 2 (whose nodes are adjacent on the output layer) correspond to the rising limbs (beginning of the rising for class 1, values closer to the peak for class 2), whereas nodes 3 and 6 (contiguous as well, even if diagonally, on the hexagonal map) correspond to the maximum flow values, respectively around the peak and at the beginning of the falling limb. Nodes 7 and 8, even if it is less evident in the hydrograph zoom reported in the figure, are associated to recession low flows. The hydro-meteorological conditions corresponding to the remaining nodes (4, 5 and 9) are instead intermediate between the previously described classes and less easily identifiable.

The nature of the various classes pictured in Fig. 1 is recognisable also by analysing their size and the mean values of the different variables in each class, reported in Table 2.

The table highlights that classes 1 and 2 are characterised by the highest precipitation values, as expectable along rising limbs, while the highest streamflow values are associated to nodes 3 and 6. Minimum streamflow values and precipitation practically null are associated to nodes 7 and 8 and it may also be noted that such conditions are largely dominant in terms of class occupancy.

Overall, the SOM seems to be able to clearly recognize the different conditions, distinguishing the parts of the hydrograph not only in terms of the flow value observed in the forecast instant and in the previous ones, but taking into account also the recent meteorological forcing: such distinction may be advantageous for discerning the near-future trend of the hydrograph evolution.

The first column of Table 2 highlights, nonetheless, that some of the classes, in particular those characterised by the highest precipitation and streamflow values, may be not numerous enough to allow a proper calibration of the rainfall-runoff models. In the modular approach, in fact, a different model is parameterised on the only data belonging to each one of these classes: the informative content of small-sized calibration sets may result insufficient for a satisfactory characterisation of the input-output relationship in the calibration

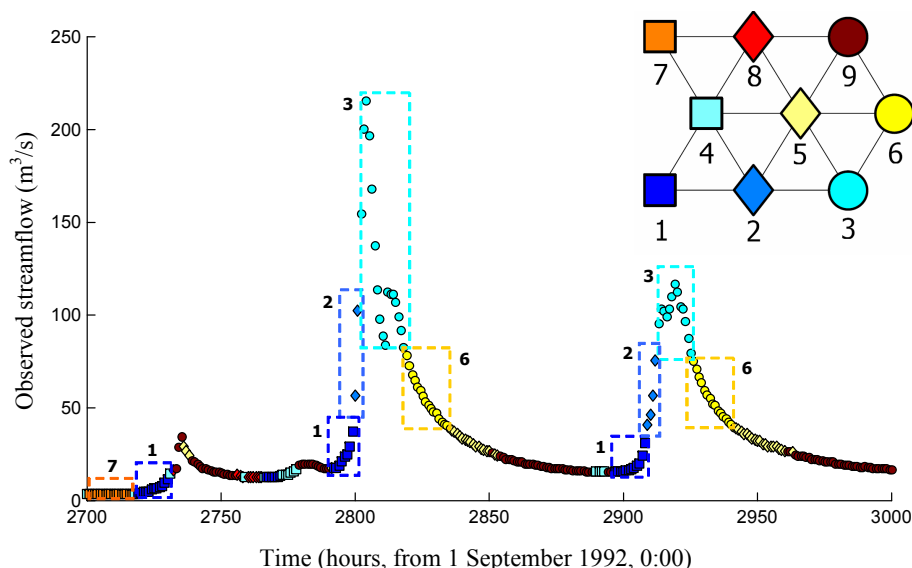


Fig. 1. Markers associated to the SOM output layer nodes (upper right-hand corner) and part of the observed hydrograph: the streamflow value relative to each forecast instant is indicated with the marker of the corresponding class. The elements belonging to classes 1, 2, 3, 6, and 7 are put in evidence by the dotted boxes of the matching colours.

Table 2. Size of the nine classes obtained with the SOM and mean values, for each class, of the variables forming the input vectors of the calibration set.

Class	Class size	Mean value						
		Q_t (m ³ /s)	Q_{t-1} (m ³ /s)	Q_{t-2} (m ³ /s)	Q_{t-3} (m ³ /s)	P_t (mm/h)	P_{t-1} (mm/h)	P_{t-2} (mm/h)
1	1141	9.11	8.50	8.14	7.92	1.84	1.81	1.62
2	329	46.40	41.36	38.52	37.02	1.19	1.54	1.64
3	602	215.13	216.42	215.78	213.80	0.73	0.80	0.89
4	1295	5.87	5.79	5.74	5.69	0.28	0.23	0.29
5	1403	30.68	30.77	30.90	31.10	0.13	0.11	0.12
6	940	54.86	55.97	57.04	58.07	0.07	0.06	0.07
7	12 166	2.13	2.14	2.15	2.15	0.00	0.00	0.00
8	5462	9.09	9.12	9.16	9.20	0.01	0.00	0.01
9	2942	19.38	19.47	19.58	19.72	0.02	0.01	0.02

phase (as will be confirmed by the forecasting results on validation data described in Sect. 6).

To overcome this problem, the opportunity to form wider classes of observations (but always homogeneous from a hydrological point of view) was tested, so to ensure a greater size of the data sets used in the calibration procedure.

Partitioning of the hydro-meteorological conditions in wider classes

The SOM classification offers a straightforward solution for the identification of similar classes, which may be joined to form broader, homogeneous groups of data. In fact, as said in Sect. 5.1, input vectors belonging to similar classes ac-

tivate nodes that are adjoining on the output layer: in this way, nodes that are nearby may be considered representative of akin classes. Once identified, on the output map, an association of similar, adjacent nodes, it is therefore identified a new, wider class, formed by all the elements that activate each one of the neighbour nodes. One such possible association may be based on the fusion of the following classes: 1 and 2 (rising limbs), 3 and 6 (flows close to the peak and beginning of falling limb), 7 and 8 (null precipitation and low flows) and the union of the remaining classes, corresponding to intermediate situations.

Table 3. Class size and mean values of the variables forming the input vectors of the calibration set for the two 4-class partitionings.

Class	Class size	Mean value						
		Q_t (m ³ /s)	Q_{t-1} (m ³ /s)	Q_{t-2} (m ³ /s)	Q_{t-3} (m ³ /s)	P_t (mm/h)	P_{t-1} (mm/h)	P_{t-2} (mm/h)
1st 4-class partitioning (reasoned association of the original nine classes)								
A (7+8)	17 628	4.29	4.30	4.32	4.33	0.00	0.00	0.00
B (4+5+9)	5640	19.09	19.14	19.22	19.33	0.11	0.09	0.11
C (1+2)	1470	17.46	15.85	14.94	14.43	1.69	1.75	1.62
D (3+6)	1542	117.43	118.61	119.01	118.87	0.33	0.35	0.39
2nd 4-class partitioning (4-node SOM)								
I	17 120	3.8	3.8	3.8	3.8	0.01	0.01	0.01
II	4850	20.2	20.3	20.5	20.6	0.02	0.02	0.02
III	1463	8.5	8.3	8.2	8.1	0.54	0.46	0.49
IV	2847	75.5	75.3	75.1	74.8	0.93	0.98	0.95

The four classes resulting from this reasoned association of the original classes may be compared with those that are obtainable setting up a new SOM, with only four nodes in the output layer, thus getting a second partitioning of the data in four classes.

The properties of these two 4-class partitionings of the calibration data are reported in Table 3.

The classes identified automatically by the 4-node SOM and those obtained by the reasoned associations (here named with letters) of the nodes of the original 9-node SOM do not coincide: in particular, the 4-node SOM does not seem able to clearly identify the cluster of the rising limbs, characterised by the highest rainfall (classes 1 and 2 of the 9-node SOM, joined in one class, named class C, in the reasoned associated classes) and to distinguish it from the data that are around the peak and at the beginning of falling limbs (classes 3 and 6 of the 9-node SOM, joined in class D). The approach for obtaining wider classes that is based on the topological properties of the original SOM appears therefore more suitable for the preservation of the hydrological distinctiveness of the classes.

6 Rainfall-runoff modelling

Preliminary to the design of the modular approaches, in order to have a term of comparison for the multi-network results, one traditional, global rainfall-runoff ANN model is implemented, trained on all the data belonging to the calibration period. As a matter of fact, as described in Sect. 3, six different mono-output feed-forward networks, with seven nodes in the input layer and three hidden nodes, were implemented for forecasting the future streamflow from one to six hours ahead, $Q_{\text{sim}}(t+L)$.

Having identified, in Sect. 5, the nature of the different hydro-meteorological conditions and the corresponding classes of forecast instants, it is then possible to build the modular rainfall-runoff systems.

The first modular approach is built on the basis of the 9-class partitioning: nine different rainfall-runoff ANN models are implemented, each one formed by six mono-output networks for the varying lead-times. Every model is parameterised through a training procedure that uses exclusively the input-output vectors, of the calibration period, belonging to the same class. In this way, a different model is built for each class, to be used for each particular hydro-meteorological condition.

In the validation phase, streamflow forecasts are issued in correspondence of every hour belonging to the last 16 months of the observation period, whose data was not used in any way in the tuning of the SOM, nor in the parameterisation of the rainfall-runoff models. In the modular approach, the tuned SOM already used to classify the calibration data is first used to associate every forecast instant of the validation period to one of the identified nine classes. The rainfall-runoff module representing that class is then chosen for issuing the streamflow forecasts.

The goodness-of-fit measures of the validation forecasts that are presented in Fig. 2 indicate, as expected, a remarkable improvement for both the global model (red bars) and the 9-class modular one (green bars) in comparison with the simple persistent model (blue bars). It is, on the other hand, evident that the use of the 9-module model allows an improvement of the MAE index, but it entails a deterioration of the efficiency coefficients, if compared to the global model.

It may be hypothesised that this deterioration is related to the low occupancy of some of the nine classes of the original partitioning: in fact, as said in Sect. 5.2, the parameterisation of the rainfall-runoff models may be inadequate because

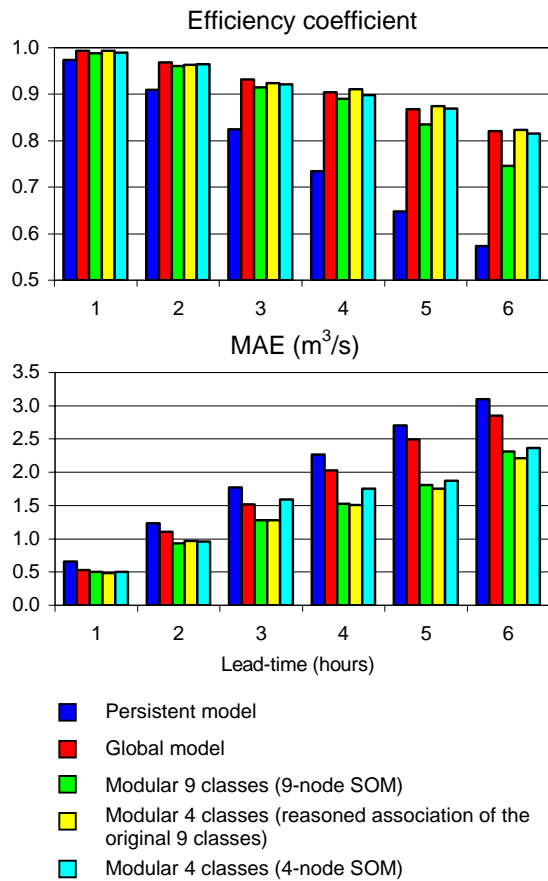


Fig. 2. Performance measures of streamflow forecasts for the validation data set.

of the insufficient informative content of the calibration data. This inadequacy is likely to affect the classes different from 7 and 8, which, in addition to be the most numerous, are also those associated to the lowest streamflow values: it follows that less reliable performances may be expected in the prediction of the higher flows. Since the efficiency coefficient amplifies the highest errors, which generally coincide with the highest flows, this would justify the deterioration of such coefficient for the 9-class modular model.

Two additional modular approaches were then implemented, based on the 4-class partitionings that were identified in Sect. 5.2, whose classes are more numerous.

The second modular system is based on the four classes obtained from the association (on the basis of their similarity) of the original nine classes. Four different rainfall-runoff network models are calibrated using all and only the data belonging to each one of the four classes of hydro-meteorological conditions. Figure 3 shows the 1 to 6 h ahead forecasts issued by this approach in correspondence of different forecast instants (blue diamonds) for three validation events: the behaviour, even if somewhere fluctuating (as expected since the forecasts are issued by independent models), is not too unrealistic.

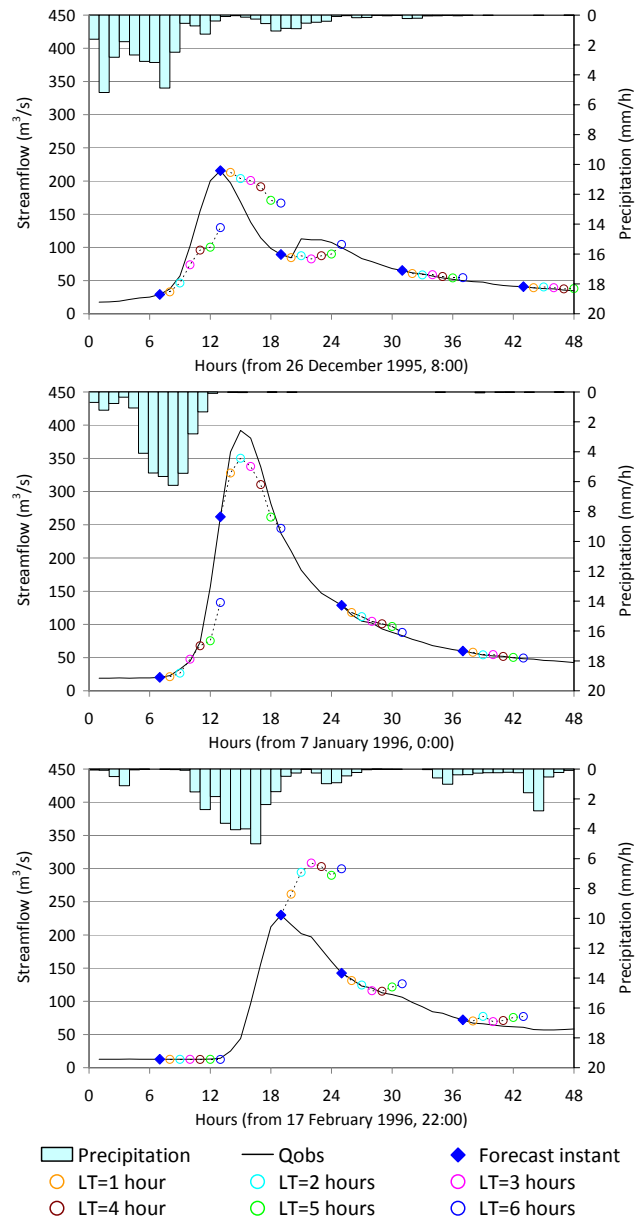


Fig. 3. Observed hydrographs (Q_{obs}) for three validation events and, in correspondence of different forecast instants, the forecasts (for lead-times $LT=1\div 6$ h) issued by the modular approach based on the reasoned association of the original 9 classes.

In analogous way, the third, and last, forecasting modular approach was implemented on the basis of the classes automatically identified by the 4-node SOM.

The goodness-of-fit indexes of the forecasts obtained with these second and third modular approaches over the validation data are shown by the yellow and cyan bars in Fig. 2. As far as the comparison with the global model is concerned, the MAE obtained with both the 4-module approaches are always substantially lower. The efficiency coefficients are higher for the forecasts issued over the longest time-horizons,

but are analogous, or even slightly worse, for the shortest lead-times. This is probably due to the fact that for short lead-times, due to the response time of the watershed, there is less influence of the most recent rainfall values: such values will instead control the evolution of the phenomena over longer time horizons, especially for the highest flow values. Therefore, for short lead-times, also the global model may allow satisfactory efficiencies, whereas, for longer time horizons, differentiating the hydro-meteorological conditions becomes crucial.

The forecasts issued by the 4-class modular approach whose classes are formed by the reasoned associations of the nine original classes are always better, especially as far as the MAE index is concerned, than those based on the 4-node SOM. This is to be ascribed to the fact that, as said in Sect. 5.2, the groups obtained from the similarity of the original classes seem able to better preserve the distinctive features of the hydro-meteorological conditions.

Overall, the modular approach based on the four, wider classes obtained on the basis of the affinity among the characterising hydro-meteorological conditions appears the best performing one, especially for the longest lead-times.

7 Conclusions

The SOM method has proved to be an instrument suitable for an objective, automatic classification of the hydro-meteorological conditions of the watershed: its use allowed in fact a satisfactory identification of the different parts of the hydrograph representing current and near-future hydrological conditions, on the basis of the most relevant information available in the forecast instant, that is recent streamflow and rainfall observations.

As far as the real-time rainfall-runoff modelling is concerned, the performances of the first modular approach, based on nine classes of hydro-meteorological situations, appear penalised by the low occupancy of some of the classes. The reduced informative content of not sufficiently numerous classes may in fact prevent an adequate characterisation of the input-output relationship in the calibration phase. Broader classes were therefore formed, through an association of the clusters representing similar hydro-meteorological conditions, exploiting the property of the SOM, unique among the other clustering techniques, to provide indications on the similarity between the classes. The new modular system, differentiating the rainfall-runoff models according to classes that are wider but still preserve the hydrological distinctiveness of the hydro-meteorological conditions, allowed a remarkable improvement of the performances in validation, in comparison to both the 9-class modular approach and to the global one. Such finding highlights the important influence, on the streamflow forecasts, of the number and properties of the classes that are identified by the SOM: additional research on this aspect will be the topic of future work.

Overall, the results show that an adequate distinction of the hydro-meteorological conditions that characterise the basin at the forecast instant, thus including additional knowledge on the forthcoming hydrological processes, may considerably improve the rainfall-runoff modelling performance.

Acknowledgements. The author wishes to thank E. Zehe, D. Solomatine, U. Ehret and an Anonymous Referee for their constructive and helpful comments, which certainly enhanced the quality and readability of the paper.

Edited by: E Zehe

References

- Abrahart, R. J. and See, L.: Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments, *Hydrol. Process.*, 14(11), 2157–2172, 2000.
- Abrahart, R. J. and See, L.: Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments, *Hydrol. Earth Syst. Sci.*, 6, 655–670, 2002, <http://www.hydrol-earth-syst-sci.net/6/655/2002/>.
- Anctil, F., Perrin, Ch., and Andreassian V.: Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models, *Environ. Modell. Softw.*, 19, 357–368, 2004.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology: Artificial neural networks in hydrology. I: preliminary concepts, *J. Hydrol. Eng.*, 5, 115–123, 2000a.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology: Artificial neural networks in hydrology II: hydrologic applications, *J. Hydrol. Eng.*, 5, 124–137, 2000b.
- Astel, A., Tsakovski, S., Barbieri, P., and Simeonov, V.: Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets, *Water Res.*, 41(19), 4566–4578, 2007.
- Bowden, G. J., Dandy G. C., and Maier, H. R.: Input determination for neural network models in water resources applications. Part 1 – background and methodology, *J. Hydrol.*, 301, 75–92, 2005.
- Cameron, D., Kneale, P., and See, L.: An evaluation of a traditional and a neural net modelling approach to flood forecasting for an upland catchment, *Hydrol. Process.*, 16, 1033–1046, 2002.
- Campolo, M., Andreussi, P., and Soldati, A.: River flood forecasting with neural network model, *Water Resour. Res.*, 35(4), 1191–1197, 1999.
- Corzo, G. A. and Solomatine, D. P.: Baseflow separation techniques for modular artificial neural networks modelling in flow forecasting, *Hydrolog. Sci. J.*, 52(3), 491–507, 2007.
- Coulibaly, P., Bobée, B., and Anctil, F.: Improving extreme hydrologic events forecasting using a new criterion for artificial neural network selection, *Hydrol. Process.*, 15, 1533–1536, 2001.
- Dawson, C. W. and Wilby, R.: An artificial neural network approach to rainfall-runoff modeling, *Hydrolog. Sci. J.*, 43(1), 47–65, 1998.
- de Vos, N. J. and Rientjes, T. H. M.: Multi-objective performance comparison of an artificial neural network and a conceptual rainfall-runoff model, *Hydrolog. Sci. J.*, 52(3), 397–413, doi:10.1623/hysj.52.3.397, 2007.

- de Vos, N. J. and Rientjes, T. H. M.: Multiobjective training of artificial neural networks for rainfall-runoff modeling, *Water Resour. Res.*, 44, W08434, doi:10.1029/2007WR006734, 2008.
- Fernando, T. M. K. G., Maier, H. R., and Dandy, G. C.: Selection of input variables for data driven models: An average shifted histogram partial mutual information estimator approach, *J. Hydrol.*, 367, 165–176, 2009.
- Foody, G.: Applications of the self-organising feature map neural network in community data analysis, *Ecol. Model.*, 120, 97–107, 1999.
- Foresee, F. D. and Hagan, M. T.: Gauss-Newton approximation to Bayesian learning, *IEEE IJCNN*, New York, 3, 1930–1935, 1997.
- Furundzic, D.: Application example of neural networks for time series analysis: rainfall-runoff modelling, *Signal Process.*, 64, 383–396, 1998.
- Georgakakos, K. P., Seo, D., Gupta, H., Schaake, J., and Butts, M. B.: Towards the characterization of streamflow simulation uncertainty through multimodel ensembles, *J. Hydrol.*, 298, 222–241, 2004.
- Gopakumar R., Takara, K., and James E. J.: Hydrologic Data Exploration and River Flow Forecasting of a Humid Tropical River Basin Using Artificial Neural Networks, *Water Resour. Manag.*, 21(11), 1915–1940, 2007.
- Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34, 751–764, 1998.
- Hagan, M. T. and Menhaj, M.: Training feedforward networks with the Marquardt algorithm, *IEEE T. Neural Networ.*, 5(6), 989–993, 1994.
- Hsu, A. L. and Halgamuge, S. K.: An unsupervised hierarchical dynamic selforganizing approach to cancer class discovery and marker gene identification in microarray data (supplementary information on “Dynamic SOM with hexagonal structure for data mining”), *Bioinformatics*, 19(16), 2131–2140, 2003.
- Hsu, K., Gupta, H. V., Gao, X., Sorooshian, S., and Imam, B.: Self-organizing linear output map (SOLO): An artificial neural network suitable for hydrologic modeling and analysis, *Water Resour. Res.*, 38(12), 1302, doi:10.1029/2001WR000795, 2002.
- Jain, A., Sudheer, K. P., and Srinivasulu, S.: Identification of physical processes inherent in artificial neural network rainfall runoff models, *Hydrol. Process.*, 18, 571–581, 2004.
- Jain, A. and Srinivasulu, S.: Integrated approach to model decomposed flow hydrograph using artificial neural network and conceptual techniques, *J. Hydrol.*, 317, 291–306, 2006.
- Kalteh, A. M., Hjorth, P., and Berndtsson, R.: Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application, *Environ. Modell. Softw.* 23, 835–845, 2008.
- Khan, M. S. and Coulibaly, P.: Bayesian neural network for rainfall-runoff modeling, *Water Resour. Res.*, 42(7), W07409, doi:10.1029/2005WR003971, 2006.
- Kohonen, T.: Self-organized formation of topologically correct feature maps, *Biol. Cybern.*, 43, 59–69, 1982.
- Kohonen, T.: *Self-Organizing Maps* (third edn.), Springer, Berlin, Germany, 2001.
- Legates, D.R. and McCabe, G.J.: Evaluating the use of ‘goodness-of-fit’ measures in hydrologic and hydroclimatic model validation, *Water Resour. Res.*, 35, 233–241, 1999.
- Madsen, H.: Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *J. Hydrol.*, 235, 267–288, 2000.
- Maier, H. and Dandy, G.: Neural networks for the prediction and forecasting of water resources variables: A review of modeling issues and applications, *Environ. Modell. Softw.*, 15(1), 101–104, 2000.
- Mangiameli, P., Chen, S. K., and West, D.: A comparison of SOM neural network and hierarchical clustering methods, *Eur. J. Oper. Res.*, 93, 402–417, 1996.
- Minns, A. W. and Hall, M. J.: Artificial neural networks as rainfall runoff models, *Hydrolog. Sci. J.*, 41, 399–417, 1996.
- Moradkhani, H., Hsu, K., Gupta, H. V., and Sorooshian, S.: Improved streamflow forecasting using self-organizing radial basis function artificial neural networks, *J. Hydrol.*, 295, 246–262, 2004.
- Parasuraman, K. and Elshorbagy, A.: Cluster-based hydrologic prediction using genetic-algorithm trained neural networks, *J. Hydrol. Eng.*, 12(1), 52–62, 2007.
- Shamseldin, A. Y., Nasr, A. E., and O’Connor, K. M.: Comparison of different forms of the Multi-layer Feed-Forward Neural Network method used for river flow forecasting, *Hydrol. Earth Syst. Sci.*, 6, 671–684, 2002, <http://www.hydrol-earth-syst-sci.net/6/671/2002/>.
- Shamseldin, A. Y., O’Connor, K. M., and Liang, G. C.: Methods for combining the outputs of different rainfall-runoff models, *J. Hydrol.*, 197, 203–229, 1997.
- Shamseldin, A. Y., O’Connor, K. M., and Nasr, A. E.: A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall-runoff models, *Hydrolog. Sci. J.*, 52(5), 896–916, 2007. natural phenomena, *Neural Networks*, 19, 215–224, 2006.
- Shirazi, J. and Menhaj, M. B.: A SOM Based 2500-Isolated-Farsi-Word Speech Recognizer, in: *ICANN 2005, LNCS 3696*, edited by: Duch, W., Kacprzyk, J., Oja, E., and Zadrony, S., Springer-Verlag Berlin Heidelberg, 589–595, 2005.
- Solomatine, D. P. and Siek, M. B.: Modular learning models in forecasting
- Solomatine, D. P. and Dulal, K. N.: Model trees as an alternative to neural networks in rainfall-runoff modelling, *Hydrolog. Sci. J.*, 48(3), 399–411, 2003.
- Srivastav, R. K., Sudheer, K. P., and Chaubey, I.: A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models, *Water Resour. Res.*, 23(10), W10407, doi:10.1029/2006WR005352, 2007.
- Tang, Y., Reed, P., and Wagener, T.: How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration?, *Hydrol. Earth Syst. Sci.*, 10, 289–307, 2006, <http://www.hydrol-earth-syst-sci.net/10/289/2006/>.
- Toth, E. and Brath, A.: Multistep ahead streamflow forecasting: Role of calibration data in conceptual and neural network modeling, *Water Resour. Res.*, 43, W11405, doi:10.1029/2006WR005383, 2007.
- Van der Voort, M., Dougherty, M., and Watson, S.: Combining Kohonen maps with ARIMA time series models to forecast traffic flow, *Transport. Res.*, C4(5), 307–318, 1996.
- Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39,

- 1–19, 2003.
- Wang, W., Gelder, P., Vrijling, J. K., and Ma, J.: Forecasting daily streamflow using hybrid ANN models, *J. Hydrol.*, 324(1), 383–399, 2006.
- Wheater, H. S., Jakeman, A. J., and Beven, K. J.: Progress and directions in rainfall-runoff modelling, in: *Modelling change in environmental systems*, edited by: Jakeman, A. J., Beck, M. B., and McAleer, M. J., Wiley, Chichester, 101–132, 1993.
- World Meteorological Organization: Simulated real-time intercomparison of hydrological models, WMO Publications, Geneva, Swiss, 1992.
- Young, P. C.: Advances in real-time flood forecasting, *Philos. T. R. Soc. Lond.*, 360, 1433–1450, 2002.
- Zhang, B. and Govindaraju, S.: Prediction of Watershed Runoff using Bayesian Concepts and Modular Neural Networks, *Water Resour. Res.*, 36(3), 753–762, 2000.