Fuzzy modelling of basin saturation state and neural networks for flood forecasting

G. Corani [a], G. Guariso[a]

[a]Dipartimento di Elettronica ed Informazione
Politecnico di Milano
e-mail:corani(guariso)@elet.polimi.it

Abstract: Over the last decade, neural networks-based flood forecast systems have been increasingly used in hydrological studies. Usually, input data of the network are composed by past measurements of flows and rainfalls, without providing a description of the saturation state of the basin, which in contrast plays a key role in the rainfall-runoff process. Here, we adopt a fuzzy approach in order to provide a description of the basin saturation state; the basin state is classified as belonging with different degrees of membership to different saturation classes, starting from the analysis of the cumulated rainfall information. A different neural predictor is specialized to mimick the rainfall-runoff relationship which pertains to each different saturation class. The forecast is obtained weighting the outputs of the specialized neural predictors, the weights being given by the current memberships of the basin state to the different saturation classes. The framework has been tested on an Italian catchment where it overperforms a classical neural networks.

Keywords: feedforward neural networks, fuzzy sets, flood forecasting

1 INTRODUCTION

An efficient flood alarm system may significantly improve public safety, and mitigate economical damages caused by inundations. Flood forecasting is undoubtedly a challenging field of operational hydrology, and a huge literature has been developed in years; in particular, the rainfall-runoff relationship has been recognized to be non linear. Although conceptual models allow a deep understanding of the hydrological processes, their calibration requires to collect a great amount of information regarding the physical properties of the watershed under study, which may be expensive and very time consuming. Since flood warning systems do not aim at providing an explicit knowledge of the rainfall-runoff process, black box models are largely used besides the traditional physically-based ones. Over the last decade, artificial neural networks (ANN) have been increasingly used in hydrological forecasting (see, for instances Maier and Dandy [2000], where tens of paper on the topic are quoted). Furthermore, their computational speed in simulating and forecasting is very welcomed in real time operations.

A well known issue is that the catchment response to rainfall impulses may strongly change depending on the saturation state of the basin. Piecewise approaches (Kachroo and Natale [1992]) have been proposed to represent the effects of basin saturation on the rainfall runoff relationship and usually antecedent precipitations are used as a proxy for the saturation state. A great drawback of piecewise approaches is however the abrupt model switching which occurs across the fixed thresholds.

In this work, we adopt a proxy constituted by the cumulated rainfall measured in the days preceding the prediction time. Starting from such an information, the basin state is evaluated in a fuzzy manner, i.e. generating a set of memberships (having unitary sum) with reference to the different saturation classes.

Assuming that each saturation class results in a different non linear rainfall-runoff relationship, a different neural network is trained to mimick the rainfall-runoff dynamic which takes place on a specific saturation class. In order to be specialized to a certain state of the basin, the neural network is trained minimizing a weighted least squares objective function, thus taking into greater consideration the rainfall-runoff observations taken in correspondence of a basin state of interest. Using a Takagi Sugeno approach (Takagi and Sugeno [1985]), the returned forecast is obtained weighting the outputs of the different neural networks, the weights being given by the current memberships of the basin state. Therefore, a smoothly increasing higher confidence is given to a certain model on the final prediction as the basin conditions becomes closer to that used for its calibration.

The paper is organized as follows: we first describe the methodological framework, i.e. the neural network model of the rainfall runoff relationship, the fuzzy representation of the basin state, and the calibration procedure adopted. Then, we compare the forecast accuracy of the system with a classical neural network approach for a river basin located in the Northern Italy.

2 NEURAL NETWORK MODELLING OF THE RAINFALL RUNOFF RELATIONSHIP

The proposed forecast system is based, according to a typical hydrological approach, on the availability of rain gauges distributed within or near the watershed. The forecasts are issued after the arrival of the rainfall events; since rainfall-runoff response times are of the order of few hours for small and medium sized basins (i.e. under or about $1000\,km^2$), this is a natural bound on the forecast lead times. In a recent work Kim and Barros [2001], lead times have been successfully increased up to 24 hours even on small basins, acquiring data from radar, radiosondes and satellite, thus monitoring on a wide area the synoptic evolution of the atmospheric conditions. However, such an advanced and expensive instrumentation is rarely available: in fact, most catchments are simply gauged with much cheaper hydrometers and rain-gauges, as is in the case study presented in the paper.

The basic modelling unit of our framework is a feed forward neural network with one hidden layer, having the water level $y(t+k)$ as target variable to be predicted, *k* being the forecast horizon. We train the networks as *direct predictors*, i.e. to return the *k*-steps ahead prediction avoiding the intermediate steps which characterize recursive forecast approaches; in fact, under mild assumptions, direct predictors can be demonstrated to provide higher performances than recursive (Atiya et al. [1999]). The variables input set comprises some *autoregressive* terms (i.e., past water level measurements), and several rainfall variables, associated with the available rain gauges. Rainfall inputs are delayed by time lags in order to take into account the travel times took by the rainfall to join the river running on the land surface.

The proposed modelling framework uses *h* different neural networks, each targeted to a different basin saturation class.

3 BASIN SATURATION ISSUES

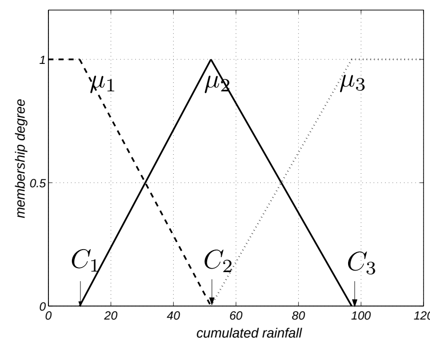


Figure 1: Sample of membership functions used to fuzzify the basin saturation state. Here three saturation classes are considered.

Let us denote as $R(\Psi, t)$ the rainfall cumulated on the whole basin area over the time window $[t-\Psi,\, t]$; it can be estimated via an interpolation method, such as Thiessen polygons, starting from the measures taken at the individual gauges. The information is obtained elaborating rainfall measures which are already available, and does not involve any significant computational overload.

We first identify a set of centroids $[C_1,\, C_2,\, \ldots C_h]$ on the time series of $R(\Psi, t)$ by means of the C-means fuzzy clustering algorithm (Bezdek [1981]). C_1, indicator of a low cumulated rainfall condition, is thought to represent a dry basin condition, while on the other hand C_2,... C_h correspond to increasing levels of the basin saturation; a basin saturation class is therefore associated to each centroid of cumulated rainfall. At this point, fuzzy logic is used to map each point of the cumulated rainfall domain to a set of scalar values μ_j $([\sum_{j=1}^{j=h} \mu_j] = 1)$, called *memberships*. Depending on the memberships, a cumulated rainfall value corresponds to several basin saturation classes at different extents. Memberships are obtained as output of the so-called membership functions; in particular, we use triangular (intermediate classes) and trapezoidal (first and last class) functions. A sample is provided in Figure 1.

4 DATA PRE-PROCESSING

According to the split sample approach (Bishop [1995]), we divide the available data into three different subsets, ensuring as far as possible the similarity of statistical properties between them:

- a training set $\mathcal{T}$, with cardinality N, used in the parameters identification;
- a validation set $\mathcal{V}$, with cardinality M, used to implement *early stopping* (Bishop [1995]): at each iteration the training objective function is evaluated on both training and validation set, and the training is stopped once the objective function on the validation set begins to increase. The union of $\mathcal{T}$ and $\mathcal{V}$ corresponds to the whole calibration set, in that the two sets are jointly exploited in the predictor choice;
- a testing set $\mathcal{S}$, used to assess the model performances on previously unused data, and to get an estimate of its generalization capabilities.

Data have been standardized in order to ensure the robustness of the training algorithm and to have a faster convergence.

5 SPECIALIZED NEURAL NETWORKS IDENTIFICATION

Neural networks are usually trained according to a least squares criterion, and thus the parameter estimate tends to represent the average situation in the dataset. However, we are interested here in specializing the neural network on data which correspond to a certain basin saturation condition, and this can be accomplished exploiting a weighted least squares criterion during the training. In particular, the weighted least squares objective function for the network specialized on the basin saturation class *j* can be written as:

$$\phi_j(\theta) = \left[\sum_{t \in \mathcal{T}} \mu_j(t) [y(t+k) - \hat{y}_j(t+k,\theta)]^2 \right] \quad (1)$$

where $\hat{y}_j(t+k)$ denotes the *k*-step ahead forecast computed by neural network *j*. We remark that the network will not consider those hydrological situations which are too far from its own saturation class, since if $\mu_j(t) = 0$, the error at time *t* does not affect the objective function of the j-th predictor. A regularization term (*weight decay*) proportional to the norm of the weights is also added to the objective function to improve the generalization capability of the model. The function to be minimized on the training data is finally defined as:

$$\Phi_j(\theta) = \phi_j(\theta) + D\,\|\theta\| \quad (2)$$

where θ is the parameters vector, and D is the weight decay coefficient. We implemented the weighted least squares Levemberg Marquardt algorithm (Press et al. [1988]) starting from the standard least squares implementation provided in the Neural Network Based System Identification Toolbox for Matlab (Norgaard et al. [2000])[1].

For each basin saturation class, the network architecture showing the best generalization (i.e., the lowest value of the unregularized criterion $\phi_j(\theta)$ on the validation set) is selected via trial and error. Each network architecture is trained 20 times in order to avoid local minima estimates.

5.1 Forecast computation

In order to compute the forecast, all the *h* neural networks are simulated; their outputs are then combined in a fuzzy manner in order to obtain the final prediction. In particular, their outputs are linearly weighted, the weights being given from the current memberships of the fuzzified basin state:

$$\hat{y}(t+k) = \sum_{j=1}^{h} \mu_j(t)\hat{y}_j(t+k) \quad (3)$$

Such an approach allows to smoothly give higher confidence to a certain model on the final prediction as the basin conditions become closer to the corresponding saturation class.

6 RESULTS

Several indicators of forecast effectiveness have been proposed besides the classical measures of the root mean square error ($RMSE = \frac{1}{N}\sqrt{\sum[y-\hat{y}]^2}$) and correlation ρ between pre-

[1] Available at
`www.iau.dtu.dk/research/control/nnsysid.html`

dicted and real flows. For instance, the model efficiency R^2, defined as:

$$R^2 = \frac{F_0 - F}{F_0} \tag{4}$$

where F_0 is the variance of water levels $\left(\frac{1}{N}\sum[y-\overline{y}]^2\right)$ and F is the mean square error $\left(\frac{1}{N}\sum[y-\hat{y}]^2\right)$, is widely used. The "prediction" of the average value, which can be considered as the prediction available even in the worst case, has then an efficiency of 0. An efficiency value of 90% indicates a very satisfactory model performance while a value in the range $80-90\%$ indicates a fairly good model (Shamseldin et al. [1997]).

Besides these indicators, which assess the mean performances on the whole dataset, the application claims for a specialized investigation of high flows situations. We define a "high flows error rate" hf, which gives an idea of the relative error on the flows data exceeding the average value $\overline{y}$ plus twice the square deviation σ_y:

$$hf = \frac{1}{K}\sum_{y>(\overline{y}+2\sigma_y)}\left|\frac{y-\hat{y}}{y}\right| \tag{5}$$

where K is the number of flows data exceeding $\overline{y}+2\sigma_y$ (about $2-5\%$ of the time series in our case study).

Figure 2: Map of the Olona catchment

We applied the approach presented above to the case of the river Olona (Figure 2), located in Lombardia, Northern Italy. Its average flow is about $2.5\,m^3/sec$, while the maximum expected flow over a time period of 10 years is about $100\,m^3/sec$. The area at the considered closing section (Castellanza) sizes about $200\,km^2$, and the basin is divided into two parts: the upper one, mountainous and weakly anthropized, and the lower one, flat and strongly urbanized. Besides Castellanza hydrometer, which measures water levels, three rain gauges are available on the basin. They are located in:

- Arcisate, in the mountainous part of the basin;
- Varese, at the beginning of the urbanized area;
- Vedano, in the urbanized area.

Data refer to 13 events (with an overall length of about 1100 hourly steps) occurred within the period 1999-2001; training, validation and testing sets contain respectively about 540, 400, 140 patterns. Due to the lack of the water level/flow relationship, we adopted the Castellanza water level as variable to be predicted. We will evaluate the model performances on a three hours forecast horizon, judged as suitable by Civil Protection technicians. Two experiments have been performed, using as saturation proxy the cumulated rainfall computed on 2 and 5 days.

In order to provide a term of comparison for the performances of the fuzzy framework, we identified also a traditional feed forward neural network predictor, having hyperbolic tangent activation function in the hidden layer neurons and linear activation function in the output layer. The optimal network architecture has been found via trial and error, training 20 times, by using early stopping, each candidate architecture.

6.1 Forecast performances

We present the resuts obtained using two saturation classes, since using three or more classes does not lead in our experiments to a performances improvement. Similar findings are reported also in (Xiong et al. [2001]), where a fuzzy combination of hydrological models has been tested. Centroids have been identified at 25 and 100 mm on the 2-days cumulated rainfall, and at 31 and 114 mm on the 5-days cumulated rainfall.

In Table 1 we show the optimal number of neurons in the hidden layer found for each neural network. It is worth noticing that the networks trained on

Model	hidden nodes
ann	3
fuzzy $(\Psi = 5days)$	
network 1 (“dry basin“)	3
network 2 (“wet basin“)	9
fuzzy $(\Psi = 2days)$	
network 1 (“dry basin“)	3
network 2 (“wet basin“)	9

Table 1: Models architectures found via trial and error

Model	**testing**			
	R^2	ρ	*RMSE*	*hf*
ann	.853	.933	.030	.294
fuzzy $(\Psi = 5days)$	.879	.948	.027	.286
fuzzy $(\Psi = 2days)$	.857	.941	.029	.319

Table 2: 3-hours ahead forecast performances.

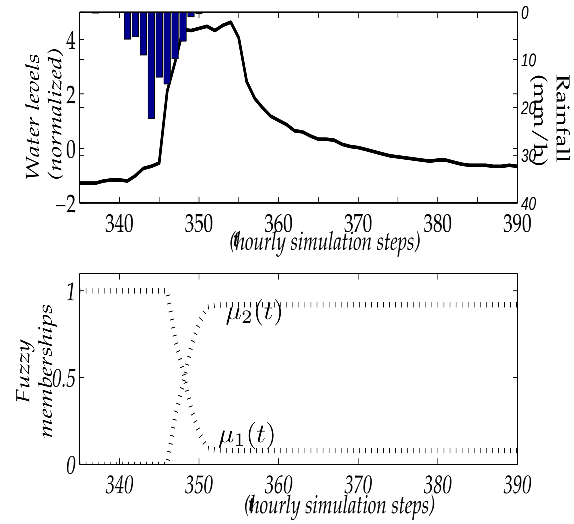
the second saturation class contain much more hidden nodes than those trained on dryer conditions. It appears thus that the rainfall runoff relationship requires an higher degree of model complexity as the basin tends towards saturation. For the classical ANN however, a low number of neurons has been found; an explanation of such results may be perhaps the significant prevalence of memberships pertaining to the first saturation class in the time series.

Table 2 shows the models performances over the testing set, where models run against previously unseen data. Remarkably, the framework based on the 5-days cumulated rainfall allows the highest performances on all the indicators. Advantages are more evident on average indicators, rather than on the peaks indicator.

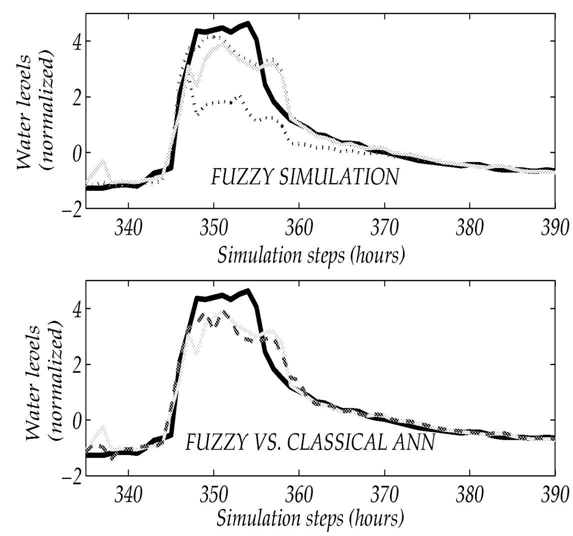
Figure 3 presents a simulation sample taken from the testing set, using the 5-days cumulated rainfall information.

7 CONCLUSIONS

The proposed frameworks fuzzifies the cumulated rainfall information in order to evaluate the basin saturation state. Specialized neural networks predictors are set up, in order to mimick the rainfall-runoff relationship which holds over the different basin saturation classes. To this end, a weighted least squares variant of the Levember Marquardt



(a) Upper plot: instant flows and rainfall on the basin (reversed *y* axis). Lower figure: fuzzy memberships



(b) Upper plot: fuzzy simulation. The bold line represents the observed values, the dotted lines the predictions of the specialized models, and the light solid line the final fuzzy forecasts. Lower plot: observed values (bold), classical ann predictions (dashed) and fuzzy predictions (solid, light).

Figure 3: 3-hours prediction samples (testing set). As μ_2 increases because of rainfall, the prediction becomes closer to that of the second predictor.

training algorithm has been implemented. The forecast is obtained combining the output of the different specialized neural predictors, according to the current basin state. Prediction performances improve in our case study by a 1-10% depending on the indicator considered, with reference to a classical neural network approach.

The complexity added in the model is not negligible, since the proposed framework is significantly more parametrized than a unique neural network. However, such a complexity appears to effectively increase the generalization of the model. With reference to the computational overload, we note that it involves in practice just the model identification step, which is however performed only once. During the operational use of the forecasting system, there is no real speed difference between simulating a neural network or a few of them. Remarkably, the cumulated rainfall additional input comes at no cost and does not require the installation of any new equipment. Finally, the case study considered here may be considered unfavorable since a large portion of the catchment is anthropized and therefore almost impervious, i.e. not sensitive to saturation issues. One can realistically expect larger improvements when dealing with less anthropized basins.

References

Atiya, A., S. El-Shoura, I. Shaheen, and M. El-Sherif. A comparison between neural-network forecasting techniques case study: River flow forecasting,. *IEEE Trans. on Neural Networks,*, 10(2):402–409, 1999.

Bezdek, J. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.

Bishop, C. *Neural networks for pattern recognition*. Oxford University Press, 1995.

Kachroo, R. and L. Natale. Non linear modelling of the rainfall runoff tranformation. *J. Hydrology*, 133:41–97, 1992.

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

Maier, H. and C. Dandy. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Mod. & Soft.*, 15:101–124, 2000.

Norgaard, M., O. Ravn, N. Poulsen, and L. Hansen. *Neural networks for modelling and control of dynamic systems*. Springer-Verlag, London, 2000.

Press, W., B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical recipes in C*. Cambridge University Press, 1988.

Shamseldin, A., K. O'Connor, and K. Liang. Methods for combining the outputs of different rainfall-runoff models. *J. Hydrology*, 197:203–229, 1997.

Takagi, T. and M. Sugeno. Fuzzy identification of systems and its application to modelling and control. *IEEE Trans. Syst. Man. Cybern.*, (15):116–132, 1985.

Xiong, L., A. Shamseldin, and K. O'Connor. A non linear combination of the forecasts of rainfall runoff models by the first order Takagi Sugeno fuzzy system. *J. Hydrology*, 245:196–217, 2001.