Flash flood prediction in large dams using neural networks

J.C. Múnera Estrada [1] and R. García Bartual [2]

[1,2]Departamento de Ingeniería Hidráulica y Medio Ambiente, Universidad Politécnica de Valencia - Spain

[1] juamues1@doctor.upv.es ; [2] rgarciab@hma.upv.es

Introduction

A flow forecasting methodology is presented as a support tool for real time flood prediction in large dams. The practical and efficient use of hydrological real-time measurements is necessary to operate early warning systems for flood disasters prevention in catchments regulated with reservoirs. In this case, the optimal dam operation during flood scenarios should reduce the downstream risks and achieve a compromise between the structural security and the objectives of the water resources system management.

A dam operation during a flood event requires to take appropriate management strategies depending on the flood magnitude and the initial freeboard at the reservoir. The most important flow prediction difficulties arise from the inherent stochastic character of peak rainfall intensities, their strong spatial and temporal variability, and the highly nonlinear response of arid and semiarid catchments resulting from a high sensitivity to the soil moisture initial conditions and the dominant flow mechanisms.

The efficient integration of a flow forecast model in a real-time prediction system should include combined techniques of data pre-processing and completion, assimilation of information and implementation of real time filters depending on the system characteristics.

This work explores the capability of flood forecast algorithms based on artificial neural networks (ANN) techniques and their integration in a real time prediction tool developed that has been named PCTR, which is the Spanish acronym for "Real Time Flood Forecasting".

Methodology and case study

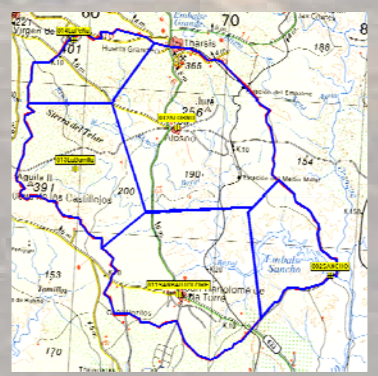


Figure 1. The Meca river catchment

The proposed forecasting methodology has been tested in the Meca River catchment (Huelva, Spain), regulated by El Sancho dam (figure 1).

A hydrological data network of 5 telemetered rain gauges (Δt=10 min) and 3 high precision water level sensors are operating in the catchment and reservoir, with real-time data transmission to a central database, making the data available at the dam control site.

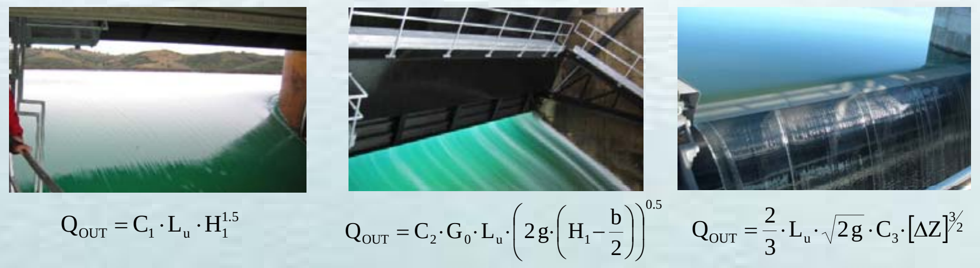
The applied methodology includes an "on line" time series reconstruction of historical output flows derived from the hydraulics of the gates (figure 2), while the inflows estimation is made by means of mass balance equation in the reservoir.

The few-hours ahead inflows are predicted with an ANN model using as input variables a sequence of current and past average hourly inflows and rainfalls in the catchment, each one with different time delays. Moreover, it is included the immediate future quantitative precipitation forecast (QPF) from an outside model.

The mass balance equation:

$$\overline{Q}_{IN} = \frac{\Delta V}{\Delta t} + \overline{Q}_{OUT}$$

$$\Delta V = V(Z_{t+\Delta t}) - V(Z_t)$$



$$Q_{OUT} = C_1 \cdot L_u \cdot H_1^{1.5} \qquad Q_{OUT} = C_2 \cdot G_0 \cdot L_u \cdot \left(2g \cdot \left(H_1 - \frac{b}{2}\right)\right)^{0.5} \qquad Q_{OUT} = \frac{2}{3} \cdot L_u \cdot \sqrt{2g} \cdot C_3 \cdot [\Delta Z]^{3/2}$$

Figure 2. Three cases of weir flow and the corresponding discharge equations

In the above equation: $\overline{Q}_{IN}$ is the mean inflow, ΔV is the storage variation, and $\overline{Q}_{OUT}$ is the output reservoir discharge in the Δt interval. Figure 3, shows the high sensitivity of the $\overline{Q}_{IN}$ estimation as a function of the level sensor precision **e** and the Δt.

The Q_{OUT} estimation sensitivity problem:

$$E(\overline{Q_{IN}}) = (\overline{Q_{IN}})_{\max} - (\overline{Q_{IN}})_{\min}$$

$$E(\overline{Q_{IN}}) = \frac{V(Z_{t+\Delta t}+e) - V(Z_t - e)}{\Delta t} - \frac{V(Z_{t+\Delta t}-e) - V(Z_t + e)}{\Delta t}$$

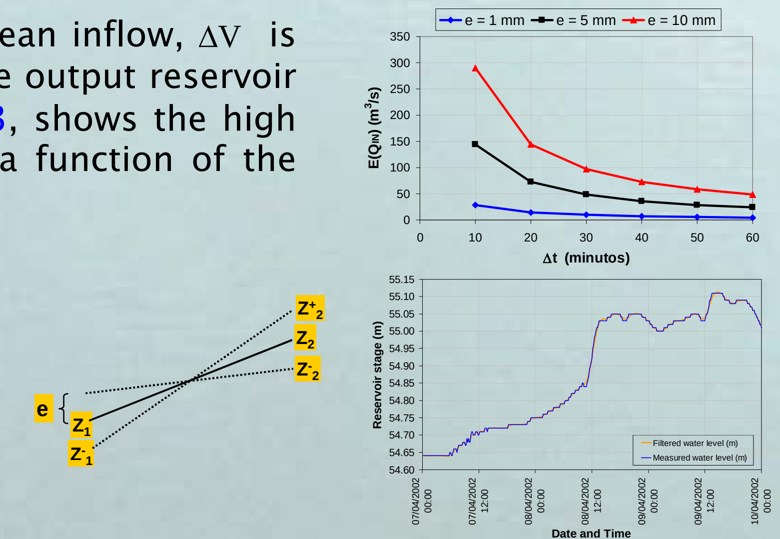


Figure 3. Flow input estimation sensitivity for different sensor precision (e) and the Δt compute time interval.

The prediction models

The ANN models have been trained and validated from 12 flood events estimated off-line (figure 4). A cross correlation analysis between precipitation data and inflows was previously performed for several historical events (figure 5).

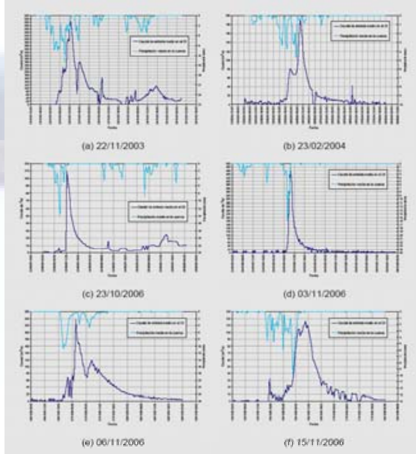


Figure 4. Some recorded flood events used in the ANN model training and validation process

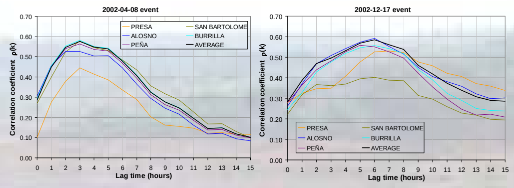


Figure 5. Cross correlation analysis for some events

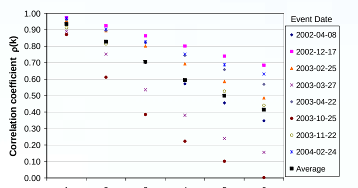


Figure 6. Autocorrelation analysis for some events

Optimal time lags were found to be in the range of 2 to 6 hours, depending on the event.

Additionally, an event based autocorrelation analysis (figure 6) shows an average correlation coefficient near to 0.50 for a 5-hours lag time, suggesting a reasonable prediction horizon.

Several ANN models architectures have been evaluated and compared. All of them have a very simple architecture based on the conventional Three Layer Feed Forward Perceptron, with a variable number of nodes in the hidden layer and one single node in the output layer producing the next-hour flow value.

A potential based transformation is applied to the original input and output variables for the ANN models.

ANN Input variables preprocessing potential functions:

$$p = \left(\frac{P}{P_M}\right)^{\beta} \qquad q = \left(\frac{Q}{Q_M}\right)^{\alpha}$$

For the following time steps, a serial-propagated neural networks structure scheme is used following the strategy suggested by F. Chang J. et al (2007), see figure 8. The evaluated architectures of ANN models are shown in figure 9.

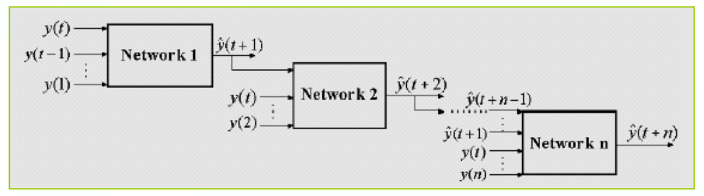


Figure 8. Serial propagation neural networks structure (from F. Chang J. et al, 2007)

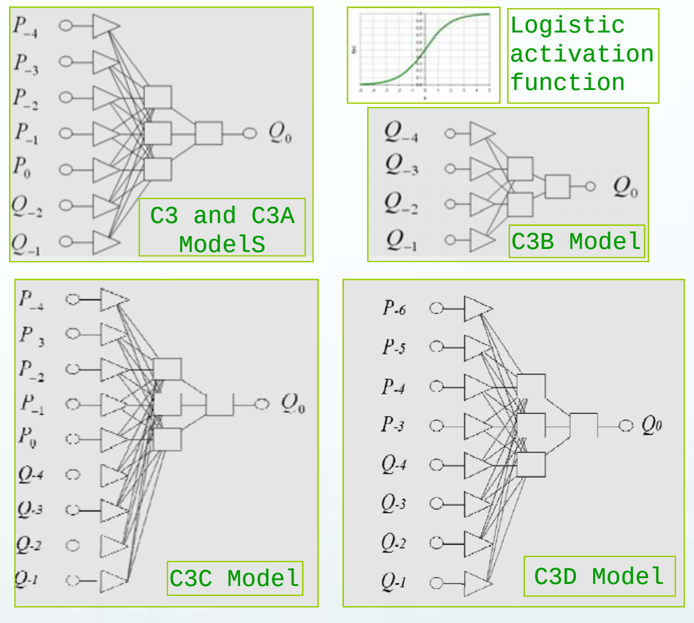


Figure 9. Several ANN architectures evaluated: C3, C3A,C3B, C3C and C3D nets

Training and validation of ANN models

The ANN models have been compared using the Root Mean Square Error (RMSE) and the Nash-Sutcliffe efficiency (NSE) statistical indices. The prediction horizon has been set to 3 hours, although results show that it could be extended a few extra hours if the precipitation forecasts were reliable enough. Initially, it has been evaluated in the training and validation processes the statistical performance of 3 different ANN models : C3, C3A - C3B and C3C. Between them, the C3C model showed a better performance in the validation data set. See figure 10.

Statistical indices for model performance assessment:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(\hat{Q}_i - Q_i)^2}{\sum_{i=1}^{n}(Q_i - \overline{Q}_i)^2} \qquad RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{n}}$$

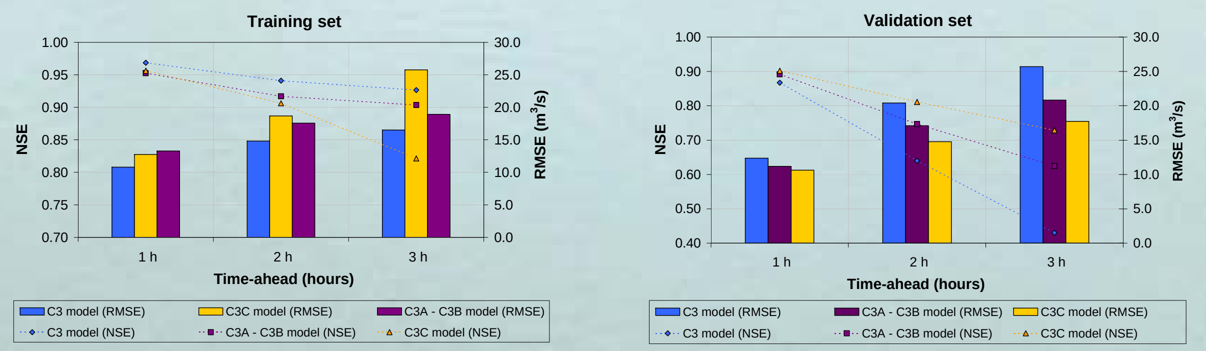


Figure 10. NSE and RMSE statistical indices obtained for 3 ANN models. Left: training data set. Right: validation data set.

Figure 11 shows the forecast response of the chosen C3C model for two recorded flood events used in the training and validation sets respectively. Additionally, the dispersion diagrams for the totality of patrons used in the training and validation data sets are shown too.

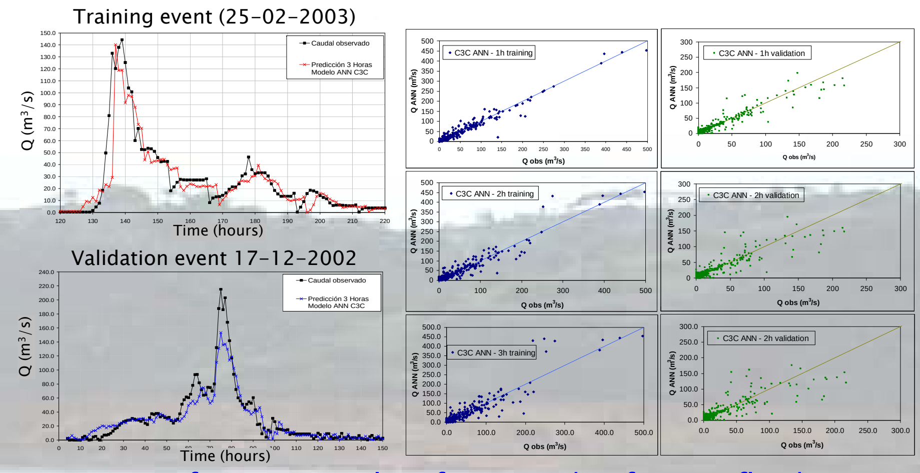


Figure 11. Left: Two examples of estimated vs forecast flood events from training set (up) and validation set (down) Right: Comparison of estimated and forecast flows for 1h to 3h time-ahead in both series.

Analysis of QPF uncertainty in C3C model

Above, it has been shown that C3C ANN model has good prediction capabilities. Nevertheless, the proposed model has a high dependence on the QPF precipitation forecast. Usually, QPF have errors greater than 50 %, and therefore, it has been analyzed what would be the effect of a systematic error of 50 % in QPF over the C3C forecast model performance. The results show a worsening of the NSE and RMSE indices, especially in the case of QPF overestimation. For this reason, it has been proposed a fourth ANN model (C3D) with greater precipitation time delays and thus eliminating this source of uncertainty. See figure 12.

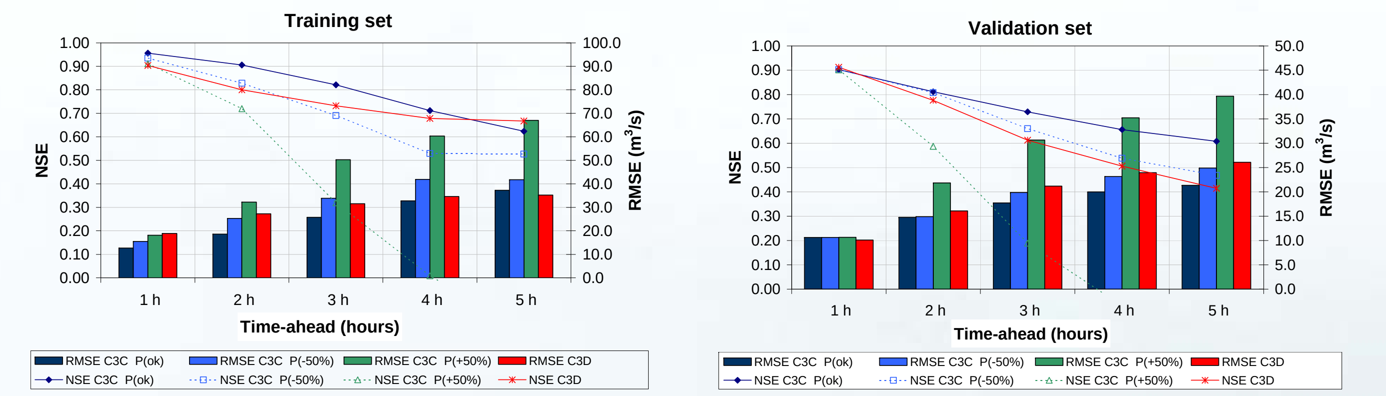


Figure 12. NSE and RMSE statistical indices obtained for C3C model with an hypothetical perfect QPF, the same model with an assumed error of ±50 % and the proposed C3D model without QPF as input variable. Left: training data set. Right: validation data set.

Conclusions

Quality of predictions has been found to be strongly affected by reliability of rainfall predictions, in particular when it is overestimated, while not so much when it is underestimated. To reduce such sensitivity, a new model (C3D) was proposed eliminating completely the predicted rainfalls in the input variables set. Although results are slightly poorer than in C3C model, the NSE index reveals a satisfactory performance in the validation set (near 0.80 for 2 hours and 0.60 for 3 hours).

The robustness and simplicity of ANN schemes makes them particularly appropriate in real-time systems, as they can easily be integrated and programmed, handling well the presence of possible errors and uncertainties in data.

On the other hand, this models are computationally very efficient, and over all, they are easily updated without changing the general conception and operation of the real-time decision making support tool.

References

Beven, K., Romanowicz R., Pappenberger, F., Young, P., and Werner, M., 2005. The Uncertainty Cascade in Flood Forecasting. International conference on innovation FloodRelief. Bergen-Tromsø, Norway. 9p.

Campolo, M., Soldati, A., Andreussi, P., 2003. Artificial neural network approach to flood forecasting in the River Arno. Hydrological Sciences-Journal-des Sciences Hydrologiques, 48(3). 381-398.

Chang F. J., Chiang, Y. M., Chang, L. C., 2007. Multi-step-ahead neural networks for flood forecasting. Hydrological Sciences-Journal-des Sciences Hydrologiques, 52(1). 114-130.

Corani G. and Guariso G., 2005. Coupling Fuzzy Modeling and Neural Networks for River Flood Prediction. IEEE Transactions on systems, man, and cybernetics—Part C: Applications and reviews, 35 (3). 382-390.

García Bartual, R., Múnera, J. C. 2007. A Decision Support Tool for Flash Flood Control in Large Dams. (Poster). EGU General Assembly 2007. EGU2007-A-11012. Viena.

Hsu, K., Gupta, H., Sorooshain, S., 1995. Artificial neural network modeling of the rainfall-runoff process. Water Resources Research 31 (10), 2517-2530.

Nash, J. E.y Sutcliffe, J. V., 1970. River flow forecasting through conceptual models. I. A discussion of principles. Journal of Hydrology, Vol. 10, p. 282-290.

Zhang, B., Govindaraju R. S., 2003. Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. Journal of Hydrology 273. 18-34.

European geosciences union – EGU 2009 – Vienna, April 20-24. Poster Session HS10.5 11456