High return period annual maximum reservoir water level quantiles estimation using synthetic generated flood events

F. Francés, R. García-Bartual and G. Bussi
Universitat Politècnica de València, Research Institute of Water and Environmental Engineering, Valencia, Spain

ABSTRACT: One of the most important variables for dam designing is the annual maximum reservoir water level (X) for different return periods. Traditionally, these quantiles have been estimated using the concept of "Design Flood”. The exceedance probability of each Design Flood is assumed to be the exceedance probability of the total accumulated precipitation of a corresponding “Design Storm". One of the main problems of a deterministic "Design Storm" is the poor representation of the space-time structure of real storms, resulting in a significant additional uncertainty in estimating flood related quantiles.

Our proposed innovative methodology to solve such problems is based on the construction of a trivariate statistical model between: i) the annual maximum daily precipitation (P), as representative of the total storm; ii) the basin initial soil moisture (H), which is relevant for the runoff production specially in arid and semi-arid climates; and iii) the variable of interest (X).

1 INTRODUCTION

One of the most important variables for dam designing is the annual maximum reservoir water level (X) for different return periods. Traditionally, these quantiles have been estimated using the concept of "Design Flood”, which is a deterministic single hydrograph for each return period. The exceedance probability of each Design Flood is assumed to be the exceedance probability of the total accumulated precipitation of a corresponding “Design Storm". This traditional methodology has two main drawbacks: First, it makes no statistical sense assigning an exceedance probability to a multivariate process, such as a hyetograph or hydrograph. Secondly, the main problem of a deterministic "Design Storm" is the poor representation of the space-time structure of real storms, resulting in a significant additional uncertainty in estimating flood related quantiles.

Our proposed innovative methodology to solve such problems is based on the construction of a trivariate statistical model between: i) the annual maximum daily precipitation (P), as representative of the total storm; ii) the basin initial soil moisture (H), which is relevant for the runoff production specially in arid and semi-arid climates; and iii) the variable of interest (in this work, X).

In the presented dam design case study, the marginal distribution of P was computed by means of a regional statistical analysis, in order to reduce the uncertainty for high return period quantiles. A set of 368 synthetic storms were generated using a space-time stochastic rainfall model based on the theory of point processes. This model is event-oriented and specifically conceived for representation of the main features and structural properties in extreme convective Mediterranean storms.

The distributed hydrological model TETIS was used to properly transform the synthetic storms into dam reservoir input hydrographs and the corresponding synthetic X. Also, the discrete marginal probabilities of 3 representative states of H were estimated by continuous hydrological simulation of the available 57 years of daily rainfall and temperature records.

Finally, by means of basic concepts of probability, it was possible to estimate the empirical probability distribution of the 368x3 annual maximum reservoir water levels, as it can be seen in the presented case study.

2 REGIONAL ANALYSIS OF ANNUAL MAXIMUM DAILY PRECIPITATION

To reduce quantile estimation uncertainty in the frequency of annual maximum precipitation is advisable to use daily data (longer records and higher number of stations) within a regional framework. In this case study, it was check the regional homogeneity using a coefficient of variation test. The total number of “equivalent years” of the regional series was of 1054, much more than the longest local record of 48 years. For the statistical model parameter estimation by Maximum Likelihood, it was used the free software developed within our research group called AFINS.

Two regionalization methods were used: the index-variable method, for the GEV, SQRT-ETmax and TCEV distribution functions, and the local Gumbel regionalization with the TCEV distribution function (Rossi et al., 1984). These four models obtained excellent and similar results. Figure 1 shows the local estimated quantiles for one of the longest raingauge station (44 complete years) using the 4 regional models and the plotting positions obtained from the local data.

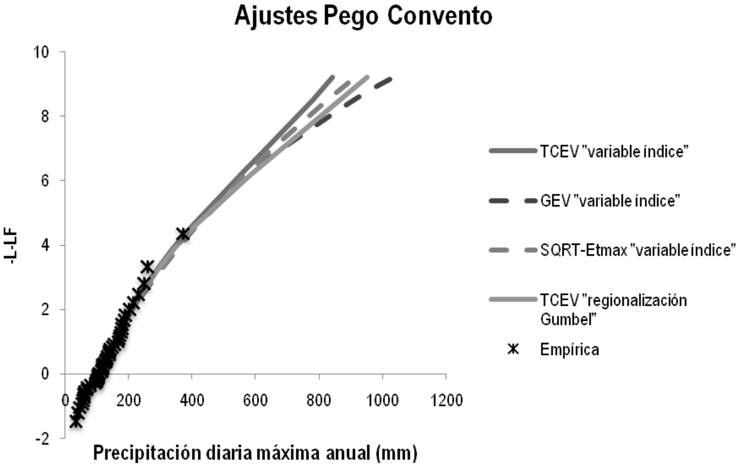

Figure 1. Local quantiles of the raingauge “Pego Convento” for the 4 fitted regional statistical models and the data plotting positions.

At the end, the TCEV with Gumbel regionalization was selected, because the physical meaning of the TCEV basic assumption of two different populations (ordinary and extraordinary events), and higher number of extraordinary events detected by the Gumbel regionalization, compared with the index-variable method. This is in accordance with previous studies for the Western Mediterranean floods (Francés, 1998).

3 GENERATION OF SYNTHETIC STORMS

3.1 *Rainfall stochastic model*

Climatic inputs derived from a space-time stochastic rainfall model have been used in this research. Such alternative is becoming more popular in recent years due to the quality rainfall data available from automatic high-resolution pluviometer networks. In addition to this, meteorological radar information provides a better knowledge of the structural patterns and its characteristic scales found in the rainfall fields after the occurrence of intense rainstorms. These sources of data facilitate an adequate parameter estimation of models, and explains the expansion of stochastic and fractal rainfall models for applied purposes (Wheater et al. ,2005; Qin, 2011; Vandenberghe et al., 2010; Calenda et al., 2005). The multidimensional definition of the rainfall field, including the evolution of the storm and a realistic representation of rainfall intensities fluctuations in space and time is a key aspect for an adequate hydrological modeling of high return period floods.

The stochastic model after (Salson and García-Bartual, 2003) was conceived to adequately represent space-time patterns and internal structures of the rainfall fields typically found during the occurrence of intense Mediterranean storms of convective nature. This model is based on previous mathematical developments of point processes applied to rainfall modeling, presented in the 80's (Rodriguez-Iturbe and Eagleson, 1987; Sivapalan and Wood, 1987). The model adapts such kind of formulations to the specific properties of the Mediterranean storms. Other approaches of stochastic modeling also based on point process theory have been used in Europe (Burton et al., 2008; Willems, 2001).

The application of this stochastic model allows definition of a wide range of climatic scenarios consisting on space-time rainfall events, each of which consisting on a rainfall intensity field with 10-minutes time level of aggregation and spatial resolution of 1x1 km^2. The associated return period is estimated for each event after a detailed analysis of the files, by means of a procedure presented herein. For this particular hydrological application, a total of 368 climatic scenarios were generated.

3.2 *Parameter estimation*

The parameters of the model are estimated by the method of moments, using some relevant empirical statistics obtained from rainfall intensity records. These include the spatial correlation function, the time autocorrelation function for different time levels of aggregation, the mean normalized function, which represents the evolution in time of the ratio between cumulative rainfall at time "t" and the total cumulative rainfall at the end of the storm and both. Finally, both the mean and variance of the total cumulative depth are used.

3.3 *Estimation of the return period for a multidimensional storm*

The concept of non-exceedance probability, and therefore the concept of return period, is inevitably linked to a given random variable and its statistical distribution. From this point of view, a conceptual problem arises when trying to assign a return period value to a given event (i.e., rainstorm) which is basically defined in terms of a family of higher order matrix, each of them containing rainfall intensities over a geographical grid of points.

To tackle this problem, the practical method proposed is based on the estimation of a unique equivalent value of rainfall, obtained from the storm multidimensional data. This value is named as the daily equivalent rainfall, PD eq, being a single representative value of cumulative rainfall to be later introduced in the extreme-value-distribution derived from the previous statistical regional analysis.

PD eq can be estimated as [24 Ix /ARF], being the return period $T=[1-F(PD\ eq)]^{-1}$, where ARF is the areal reduction factor for the hydrological catchment, and "x" is time interval in hours so that $Ix \approx PD/24$, being Ix (mm/h) the annual maximum of average rainfall intensity for an interval of duration "x" hours and PD the corresponding annual maximum of daily rainfall.

According to the hydrological technical guidelines in Spain, a value of x=28 h is mostly used, with ARF=[1- (log A)/15], if no specific studies are available in the region. "A" represents the catchment area in km2. For the case study presented herein (Marina Alta in the Valencian Community-Spain), both aspects were investigated. After analysis of historical records in selected raingauges, a value of x= 29h 10min was obtained for the region, very close to the general recommendation of 28 hours.

Additionally, and after areal analysis of extreme events occurred in the Valencian Community, a value of $ARF=[1+\lambda A\gamma]^{-1}$ with $\lambda=0.00783$ and $\gamma=0.65676$ was obtained.

For a given multidimensional storm, either generated from a stochastic rainfall model or defined by actual rainfall records from the raingauge network or radar, the associated value of Ix should be obtained. To do so, the areal averaged hyetograph over the catchment is first estimated. Figure 2 below shows the case of synthetic storm #106, over the Girona upper river catchment in the Valencian Community.

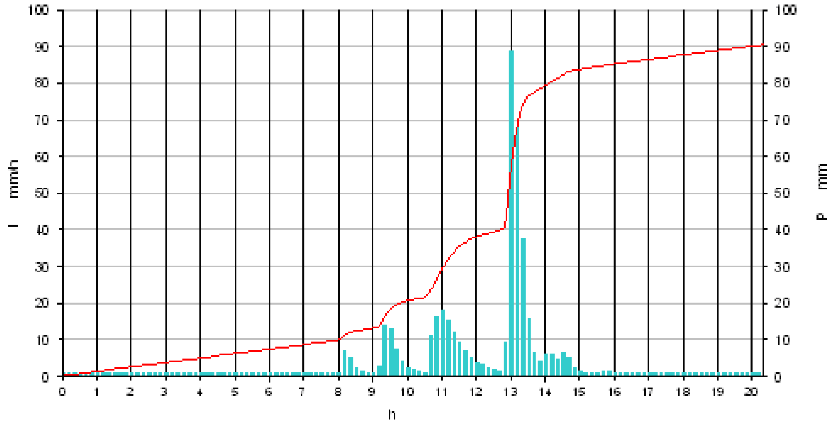

Figure 2. Areal averaged - synthetic rainfall event # 106. Upper Girona river catchment.

4 HYDROLOGICAL MODELING WITH TETIS

4.1 *Event model implementation*

The TETIS model is a distributed and conceptual hydrological model with physically based parameters (Francés et al. 2002, Francés et al. 2007). It has been calibrated and validated at flood event scale (Δt = 10 minutes) in the actual Isbert dam, using the recorded events by the Flood Warning System of the Jucar Basin Authority since 1989. The calibration and validation was satisfactory, with Nash-Sutcliffe eficiency indexes greater than 0.8, in calibration (Fig. 3) and in most of the temporal validations.

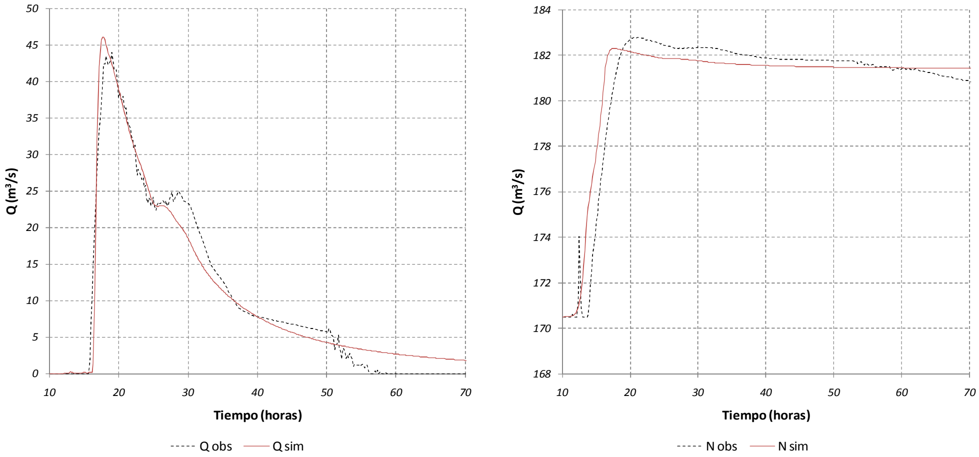

Figure 3. Actual Isbert dam (Girona River) observed and simulated reservoir output hydrograph (left) and reservoir water level (right) for the flood event of April 2003.

4.2 *Initial soil moisture analysis*

The objective of the hydrological modeling is the simulation of a set of independent synthetic storms, without a definition of the initial condition of the state variables, in particular, the initial soil moisture condition. For this reason, the TETIS model has been also calibrated and validated at daily temporal discretization. In this case, it was used the flow gauge station of "Gallinera" due to its better daily record.

Once the daily model has been adjusted, it has been used to simulate the soil moisture since 1943 to nowadays, using the daily precipitation records of the Spanish National Weather Service within the Gallinera basin. From these results, three different initial conditions for dry, medium and wet soil (10, 40 and 80% of soil moisture respectively) has been considered and it has been assigned an occurrence probability.

4.3 *Synthetic events simulation*

Finally, at least with respect the hydrological modeling, TETIS model has been used to simulate the rainfall-runoff transformation for the 368 available synthetic storms, using the three different initial soil moisture conditions computed previously. At the end, it was produced 1104 input hydrographs to the new dam site.

5 PROBABILITY ESTIMATION

The synthetic storm generator doesn't assign a particular date at each storm, because it is a single storm generator (not a continuous precipitation generator). So, it is needed an specific statistical model to obtain non-exceedance probability of the random variable of interest, in this paper the annual maximum reservoir water level. We propose the use of a trivariate model of the annual maximum daily precipitation, the initial soil moisture condition (a discrete random variable in this work) and the annual maximum reservoir water level. What it is a priori known are the marginal distributions of the first two and their cross-correlation.

From elementary concepts of probability, it is possible the estimation of the empirical marginal distribution function (or plotting positions) of the random variable of interest, for the sample synthetically generated. The final equation is:

$$F_X(a) \approx \sum_{j=1}^{3}\left\{ Pj \sum_{i=0}^{\infty} \frac{n_{ij}(a)}{N_{ij}} \left[F(R_{i+1}) - F(R_i) \right] \right\}$$

where Pj corresponds to the probability initial soil moisture is in the state j; $n_{ij}(a)$ is the number of observations less or equal to a, within the interval i, which covers the annual maximum daily preciptation range $[R_i, R_{i+1}]$ and N_i is the total number of observations within the precipitation interval i. In Figure 4 is represented the result for the case study.

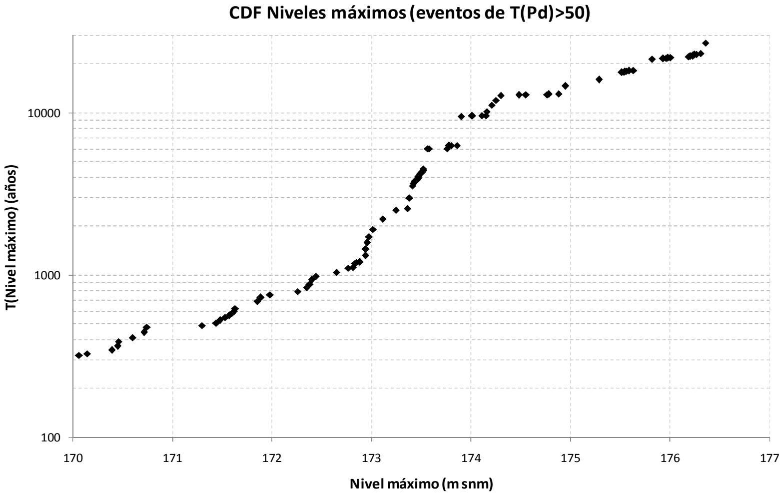

Figure 4. Plotting positions of the annual maximum reservoir water level for the new Girona River dam.

6 CONCLUSIONS

One of the most common objectives of hydrological studies is the flood frequency analysis at a given section of a river. This analysis is used in the design of many hydraulic works (spillways of dams, diversion tunnels, channeling, etc..), transverse drainage systems of linear infrastructures (bridges and culverts), flood hazard and risk mapping, and so on. The analysis of the frequency of floods has the main target of getting the relationship between flood flows (or quantiles) and their corresponding non-exceedance probability (or return period), although in most cases you can just get the flows quantiles for a few number of return periods.

In many of the above described situations, the socioeconomic implications are enormous, so it is enforceable to work as accurately as possible: an underestimation error leads to an increase of the actual risk, while if the error is and overestimation, which is causing is an unnecessary increase in the infrastructure cost.

The traditional methodology based on the concept of the Flood Design have enough misconceptions to that recommend its abandonment, to the extent that is desirable the maximum precision in the calculations. In this article we have shown that it is possible to develop methodologies that address these misconceptions.

7 REFERENCES

Burton, A., Kilsby, C.G., Fowler, H.J., Cowpertwait, P.S.P. and O'Connell, P.E.,2008. RainSim: A spatial temporal stochastic rainfall modelling system. *Environmental Modelling and Software,* 23, 1356-1369.

Calenda, G., Gorgucci, E., Napolitano, F., Novella, A. and E.Volpi, 2005. Multifractal analysis of radar rainfall fields over the area of Rome. *Advances in Geosciences*, 2, 293-299.

Francés, F., 1998. Using the TCEV distribution function with systematic and non-systematic data in a regional flood frequency analysis. *Stochastic Hydrology and Hydraulics*, 12 (4), 267-283.

Francés, F., I. Vélez, J. Vélez and M. Puricelli, 2002. Distributed Modeling of Large Basins for a Real Time Flood Forecasting System in Spain. In: *Second Federal Interagency Hydrologic Modeling Conference*, Las Vegas, USA.

Francés, F. I. Vélez and J. Vélez, 2007. Split-parameter structure for the automatic calibration of distributed hydrological models. *Journal of Hydrology*, 332, 226-240.

Qin, J.,2011. A high- resolution hierarchical model for space-time rainfall. *Phd Thesis. Univ. of Newcastle.*

Rodriguez-Iturbe, I. and Eagleson, P.S.,1987. Mathematical models of rainstorm events in space and time. *Water Resources Research, 23 (1) pp. 181-190.*

Rossi, F., Fiorentino, M., and Versace, P., 1984. Two-Component Extreme Value Distribution for Flood Frequency Analysis. *Water Resour. Res.,* 20 (7), 847-856.

Salsón, S. and R. García-Bartual, 2003. A space-time rainfall generator for highly convective Mediterranean rainstorms. *Natural Hazards and Earth System Sciences*, 3, 103-114.

Sivapalan, M. and Wood, E.F.,1987. A multidimensional model of non stationary space-time rainfall at the catchment scale. *Water Resources Research, 23 (7) pp. 1289-1299.*

Wheater, H.S., Chandler, R.E., Onof, C.J., Isham, V.S., Bellone, E., Yang, C., Lekkas, D., Lourmas, G. and M.L.Segond, 2005. Spatial-temporal rainfall modelling for flood risk estimation. *Stoch Environ Res Risk Assess* 19. 403-416.

Vandenberghe, S., Verhoest, N.E.C., Buyse, E. and B.De Baets, 2010. A stochastic design rainfall generator based on copulas and mass curves. *Hydrol. Earth Syst. Sci.,* 14, 2429-2442.

Willems, P., 2001. A spatial rainfall generator for small spatial scales. *Journal of Hydrology*, 252: 126-144.