Long term prediction of flood occurrence

Cristina Aguilar[1]**, Alberto Montanari**[2]**, and María José Polo**[3]

[1]Fluvial dynamics and hydrology research group, Andalusian Institute of Earth System Research, University of Granada, Granada, Spain
[2]Department DICAM, University of Bologna, Bologna, Italy
[3]Fluvial dynamics and hydrology research group, Andalusian Institute of Earth System Research, University of Cordoba, Cordoba, Spain

Correspondence to: Cristina Aguilar (caguilar@ugr.es)

Abstract. How long a river remembers its past is still an open question. Perturbations occurring in large catchments may impact the flow regime for several weeks and months, therefore providing a physical explanation for the occasional tendency of floods to occur in clusters. The research question explored in this paper may be stated as follows: can higher than usual river discharges in the low flow season be associated to a higher probability of floods in the subsequent high flow season? The physical explanation for such association may be related to the presence of higher soil moisture storage at the beginning of the high flow season, which may induce lower infiltration rates and therefore higher river runoff. Another possible explanation is persistence of climate, due to presence of long-term properties in atmospheric circulation. We focus on the Po River at Pontelagoscuro, whose catchment area amounts to $71\,000\,\mathrm{km}^2$. We look at the stochastic connection between average river flows in the pre-flood season and the peak flows in the flood season by using a bivariate probability distribution. We found that the shape of the flood frequency distribution is significantly impacted by the river flow regime in the low flow season. The proposed technique, which can be classified as a data assimilation approach, may allow one to reduce the uncertainty associated to the estimation of the flood probability.

1 Introduction

Perturbations occurring in large catchments may impact the flow regime for several weeks and months, therefore providing a physical explanation for the occasional tendency of floods to occur in clusters (Montanari, 2012). In the Po river, for instance, it has been observed that some flood events have been preceded by long lasting average flows. The physical explanation for such association may be related to the presence of higher than usual soil moisture storage, which may induce lower infiltration rates and therefore higher river runoff. Another possible explanation is persistence of climate, due to presence of long-term properties in atmospheric circulation.

It is well known that river flows are affected by forms of persistence that are not fully understood yet (O'Connell et al., 2015). These are referred to as the "Hurst Phenomenon", or the "Hurst Effect". The Hurst Effect has been physically explained as an implication of the principle of maximum entropy (Koutsoyiannis et al., 2011; Koutsoyiannis, 2014) and implies the presence of long-term cycles over a multitude of time scales. Therefore, the presence of long memory is connected to the possible occurrence of long-term cycles that imply the persistence of high and extreme flows.

With the idea that extreme floods may be induced by long term stress, rather than a short sequence of extreme rainfall, this paper explores the following research question: can higher than usual river discharges in the low flow season be associated to a higher probability of floods in the subsequent high flow season? An application in the Po River is carried out in order to set up a methodology to update the uncertainty associated to the estimation of flood occurrence probability.
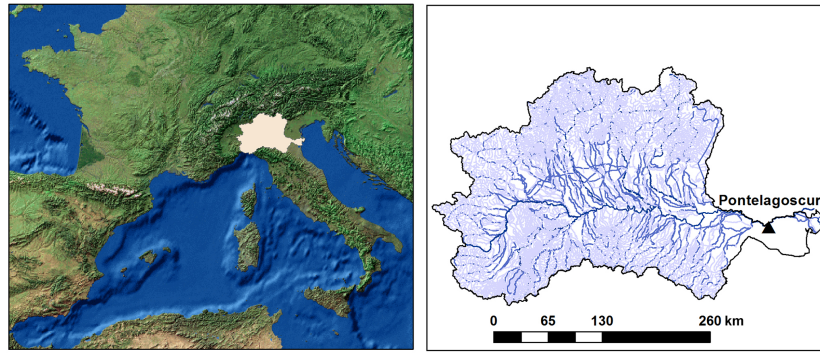
Figure 1. Po River Basin. Location (left), drainage network and closure section at Pontelagoscuro (right).

2 Study site and data sources

The Po River whose catchment has an area of about $71\,000\,\mathrm{km}^2$ is the longest river entirely flowing in the Italian Peninsula (Fig. 1). The average annual precipitation in the catchment is $78\,\mathrm{km}^3$, of which 60 % reaches the closure river cross-section at Pontelagoscuro where the mean annual flow is about $1470\,\mathrm{m}^3\,\mathrm{s}^{-1}$. An intense exploitation of water resources for irrigation, hydro-power production, civil and industrial use is found in the catchment. Even though the situation is currently sustainable on average, it might be problematic during drought periods (Montanari, 2012). The hydrological behavior of the Po River is described in detail in recent studies (Zanchettin et al., 2008; Montanari, 2012; Zampieri et al., 2015).

Daily discharge time series for the Po River Basin in Pontelagoscuro were analyzed in this study. The observation period of the complete series was 1920–2009. The discharge pattern shows a typical pluvial regime and thus a strong seasonality with two flood seasons in spring and autumn (Fig. 2).

3 Bivariate probability distribution fitting

In order to look at the stochastic connection between the average river flows in the pre-flood season and the peak flows in the flood season a bivariate probability distribution function is fitted to observed data sets. In what follows, random variables and their outcomes are identified with bold and unbold characters, respectively. The yearly random variables included in the analysis were:

 – Monthly mean flow in the pre-flood season, Q_{m}.

 – Peak flow in the flood season, Q_{p}.

First, the time series $Q_{\mathrm{m}}(t)$ and $Q_{\mathrm{p}}(t)$ with sample size n, where n is the number of years in the observation period, are extracted from the observed data sets. Then, the Normal Quantile Transform (NQT) is applied in order to make their marginal probability distributions Gaussian, therefore obtaining the normalized observations $NQ_{\mathrm{m}}(t)$ and $NQ_{\mathrm{p}}(t)$.
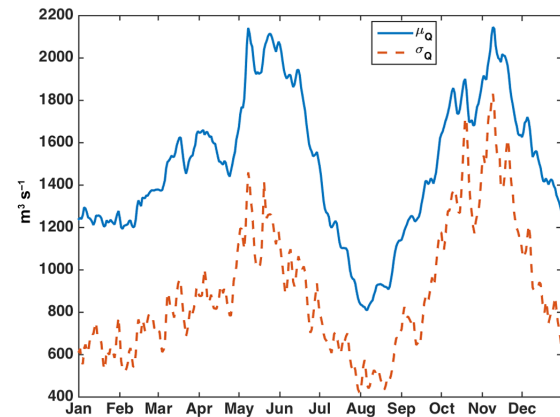

Figure 2. Daily mean value μ_{Q} ($\mathrm{m}^3\,\mathrm{s}^{-1}$) and daily standard deviation σ_{Q} ($\mathrm{m}^3\,\mathrm{s}^{-1}$) of the daily flows in the observation period (1920–2009) at Pontelagoscuro.

A detailed description of the application of the NQT in hydrological studies can be found in the literature (e.g. Moran, 1970; Montanari and Brath, 2004; Montanari, 2005; Montanari and Grossi, 2008; Bogner et al., 2012).

Finally, a bivariate Gaussian distribution function between both canonical Gaussian random variables is fitted. The parameters of the distribution are the mean $\mu(NQ_{\mathrm{m}}) = 0$ and $\mu(NQ_{\mathrm{p}}) = 0$ and the standard deviation $\sigma(NQ_{\mathrm{m}}) = 1$ and $\sigma(NQ_{\mathrm{p}}) = 1$ of the normalized series, and the Pearson's cross correlation coefficient between both normalized series, $\rho(NQ_{\mathrm{m}}, NQ_{\mathrm{p}})$. In the presence of dependence between NQ_{m} and NQ_{p}, the correlation coefficient will be significantly different from zero. The bivariate Gaussian distribution implies that, for an arbitrary (observed) $NQ_{\mathrm{m}}(t)$, the probability distribution of NQ_{p} is Gaussian, with parameters (Eqs. 1, 2):

$$\mu\left(NQ_{\mathrm{p}}\right) = \rho\left(NQ_{\mathrm{m}}, NQ_{\mathrm{p}}\right) \cdot NQ_{\mathrm{m}}(t), \tag{1}$$

$$\sigma\left(NQ_{\mathrm{p}}\right) = \left(1 - \rho^2\left(NQ_{\mathrm{m}}, NQ_{\mathrm{p}}\right)\right)^{0.5}. \tag{2}$$

Then, by taking the inverse of the NQT one can infer the updated probability distribution of Q_{p} conditioned to the ob-
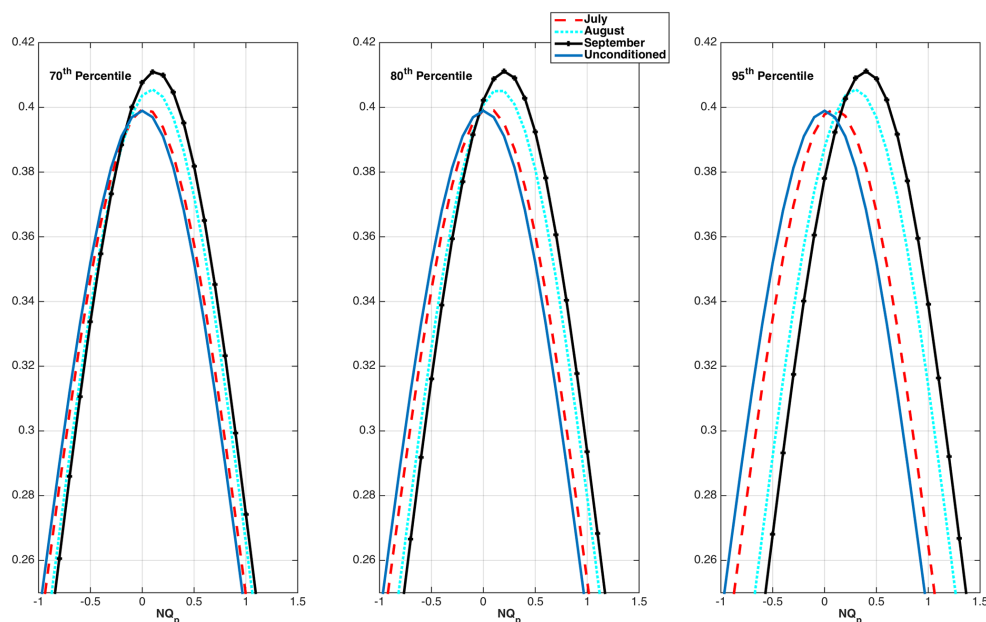
Figure 3. Probability distribution functions of the normalized dependent variable (NQ$_p$) conditioned to the occurrence of the 70th, 80th and 95th percentiles of the normalized variables in the pre-flood season.

served outcome $Q_m(t)$. Therefore, once the parameters of the distribution are computed, the probability distribution of the peak flow can be updated after observing the average flow in the considered low flow season.

The following two main assumptions are applied in this study. (1) The peak flows season covers the months of October and November in the Po River. Thus, the low flow season is assessed in the previous months to the peak flows season (July–September). Nevertheless, the methodology allows the user to select the seasons arbitrarily so that it can be applied to any other study site or hydrological regime. (2) For the sake of comparison, peak flows can be adequately modeled through the EV1 distribution.

In order to infer the actual impact of the dependence between peak flows and average flow in the low flow season, the unconditioned flood frequency distribution and the updated distributions inferred for several levels higher-than-average values of mean flow (e.g. 70, 80, and 95 % quantiles) in the pre-flood season were compared.

4   Results

The correlation coefficient between NQ$_p$ and NQ$_m$ was calculated by considering different observation periods for Q_m. In detail, we assumed that Q_m is given by the monthly mean flow in each of the 9 months preceding the high flow season (from September to January). Table 1 shows the decrease in the correlation coefficient as the considered low flow period moves backward, as one would expect.

Table 1. Correlation coefficient between NQ$_p$ and NQ$_m$ for varying low flow season in the abscissa.

Month	$\rho(\mathrm{NQ_m}, \mathrm{NQ_p})$
September	0.24
August	0.18
July	0.06
June	0.02
May	−0.06
April	−0.13
March	−0.18
February	−0.04
January	−0.07

The effect of the identified dependence on peak flow estimation, for an assigned return period, is shown in Fig. 3 for three different levels of mean flow (70, 80, and 95 % quantiles) in the considered pre-flood season. The probability distribution functions (pdf) of the normalized observed variable, NQ$_p$, with mean zero and standard deviation 1 is also displayed for the sake of comparison and denoted as unconditioned in Fig. 3. We can appreciate that the higher the cross correlation value, the lower the variability in the distribution of the normalized dependent variable and the higher the mean value. For example, when estimating the probability distribution of NQ$_p$ conditioned to the occurrence of the 95th quantile value in the normalized mean flow in September, the pdf is centered around a mean vaue of 0.4 and presents a standard deviation of 0.97. In contrast, if one attempts to estimate the probability distribution of NQ$_p$ conditioned to the occurrence
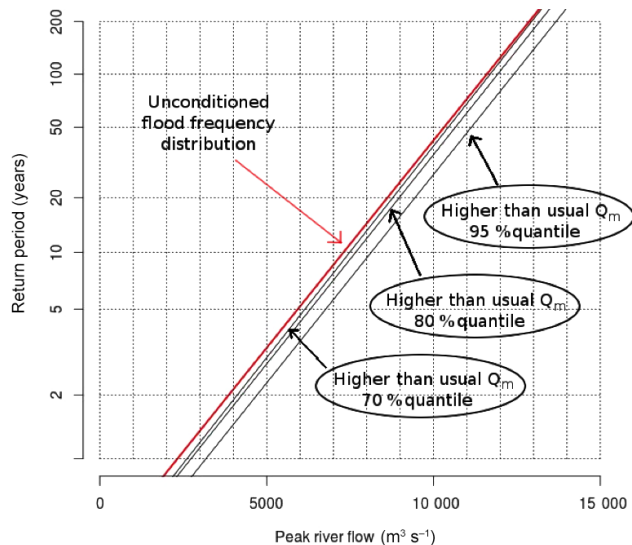
Figure 4. Peak flows in the flood season (October–November) vs return period modeled through the EV1 distribution function.

of the 95th quantile of the normalized mean flow in June, no significant change is found in the estimate with respect to the unconditioned distribution. In fact, the resulting probability density function (pdf) for NQ_p is centered around a mean value of 0.03 with a standard deviation of 0.998. In what follows, September was selected as the pre-flood season in the study site.

Once a pre-flood season was identified it is possible to update the flood frequency distribution after the observation of $Q_m(t)$. Figure 4 shows the comparison between the unconditioned flood frequency distribution and the simulated updated distributions when the flow in September is higher than usual (70, 80, and 95 % quantile). For example, the unconditioned expected flood for a return period of 200 years, $12\,507\,\text{m}^3\,\text{s}^{-1}$, increases up to $13\,790\,\text{m}^3\,\text{s}^{-1}$ when the mean flow in September corresponds to its 95 % quantile).

5 Conclusions

We found that the peak flow of the Po River is dependent on the average flow of the pre-flood season. Thus, we conclude that it is possible to update the flood frequency distribution basing on discharge observations during the low flow season. To this end, we use a bivariate Gaussian distribution function to model the above dependence. The methodology herein proposed can be applied to any other study site once the flood season is identified and the parameters of the bivariate distribution confirm the presence of the above stochastic dependence.

Several possible physical explanation can be postulated for the sensitivity of the peak flow to the mean discharge in the preceding low flow season, such as the impact of the catchment storage or soil moisture, which in turn impact the for-

mation of net rainfall, and the existence of memory in the weather. Current research is focusing on gaining a better understanding of the processes leading to the formation of the flood flows and in particular the related weather dynamics. Furthermore, we are carrying out experiments on several other rivers in the attempt to relate the above dependence to catchment properties.

References

Bogner, K., Pappenberger, F., and Cloke, H. L.: Technical Note: The normal quantile transformation and its application in a flood forecasting system, Hydrol. Earth Syst. Sci., 16, 1085–1094, doi:10.5194/hess-16-1085-2012, 2012.

Koutsoyiannis, D.: Entropy: from thermodynamics to hydrology, Entropy, 16, 1287–1314, doi:10.3390/e16031287, 2014.

Koutsoyiannis, D., Paschalis, A., and Theodoratos, N.: Two-dimensional Hurst–Kolmogorov process and its application to rainfall fields, J. Hydrol., 398, 91–100, doi:10.1016/j.jhydrol.2010.12.012, 2011.

Montanari, A.: Deseasonalisation of hydrological time series through the normal quantile transform, J. Hydrol., 313, 274–282, doi:10.1016/j.jhydrol.2005.03.008, 2005.

Montanari, A.: Hydrology of the Po River: looking for changing patterns in river discharge, Hydrol. Earth Syst. Sci., 16, 3739–3747, doi:10.5194/hess-16-3739-2012, 2012.

Montanari, A. and Brath, A.: A stochastic approach for assessing the uncertainty of rainfall-runoff simulations, Water Resour. Res., 40, W01106, doi:10.1029/2003WR002540, 2004.

Montanari, A. and Grossi, G.: Estimating the uncertainty of hydrological forecasts: A statistical approach, Water Resour. Res., 44, W00B08, doi:10.1029/2008WR006897, 2008.

Moran, P.: Simulation and Evaluation of Complex Water Systems Operations, Water Resour. Res., 6, 1737–1742, doi:10.1029/WR006i006p01737, 1970.

O'Connell, P. E., Koutsoyiannis, D., Lins, H. F., Markonis, Y., Montanari, A., and Cohn, T.: The scientific legacy of Harold Edwin Hurst (1880–1978), Hydrol. Sci. J., doi:10.1080/02626667.2015.1125998, online first, 2015.

Zampieri, M., Scoccimarro, E., Gualdi, S., and Navarra, A.: Observed shift towards earlier spring discharge in the main Alpine rivers, Sci. Total Environ., 503–504, 222–232, doi:10.1016/j.scitotenv.2014.06.036, 2015.

Zanchettin, D., Traverso, P., and Tomasino, M.: Po River discharge: a preliminary analysis of a 200-year time series, Climatic Change, 88, 411–433, doi:10.1007/s10584-008-9395-z, 2008.