# Advances in shot noise modeling of daily streamflows

Claps P., Giordano A. and Laio F.

Dipartimento di Idraulica Trasporti e Infrastrutture Civili, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy.

**Abstract**

Univariate shot noise models for streamflow generation at short time scales are examined in detail, to reconsider the verification of the basic hypotheses behind the models, the problem of objectively evaluating their performances, and the importance of model parsimony. The classical approach to model estimation is shown to produce some inconsistencies in the inverse evaluation of the model input, in particular regarding the assumed independence and Poissonianity of the pulses; an alternative procedure for pulses identification is proposed, which enables the mentioned hypotheses to be respected. To evaluate model performances, two indices are proposed, respectively based on the comparison of real and generated flow duration curves ($I_1$) and annual maxima statistics ($I_2$). A method for explicitly accounting for the dependence of $I_1$ and $I_2$ on the number of model parameters is described. An application to 7 daily streamflow time series in northern Italy demonstrates the validity of the proposed procedure for the identification of the input and the usefulness of the performance indices in discerning among competing models.

**keywords**: streamflow simulation; daily runoff; independence; Poissonianity; model parsimony.

# 1 Introduction and scope of the paper

Synthetic streamflow data at different time resolutions (hourly to yearly) are often required for planning and management of water resources systems (see e.g. Salas, 1993). Filtered point processes are commonly used to generate synthetic data at short time scales (hourly to weekly), basically for their better aptitude to reproduce the presence of peaks and recessions that appear in daily discharge data. In particular, filtered Poisson processes (also named Shot Noise processes) have received considerable attention in the hydrologic literature (Bernier (1970), Lawrence and Kottegoda (1977), Weiss (1977), Koch (1985), Murrone et al. (1997), Xu et al. (2002)). Most efforts in this field have been devoted to the definition of the (deterministic) system response function, or to the parameter estimation methods, while other significant problems have been somewhat neglected. Among these, the verification of the model basic hypotheses, such as the independence of consecutive pulses in the marked point process and the Poissonianity of the distribution of their occurrences, the problem of objectively evaluating model performances, and the role of the number of model parameters on the model overall efficiency.

The first point (verification of model hypotheses) is strictly connected to the method used to estimate the pulse sequence representing the model input, which is meant to be the effective rainfall sequence. A typical approach is to assign a pulse to any rise in the discharge time series (Battaglia, 1986; Murrone et al., 1997; Xu et al., 2002); however, this procedure tends to produce pulse sequences with a very large average number of events per year, $\lambda$, and the basic hypotheses behind the Poisson model (independence and Poissonianity) tend not to be respected. A new procedure, borrowed from statistical hydrology, is proposed here to cope with this problem. The effective rainfall events are selected with a modified Peak Over Threshold approach (Claps and Laio, 2003) in a way that guarantees the respect of the mentioned hypotheses (see Section 3).

In the second part of the paper (Section 4) the problem of properly evaluating model performances is considered (see for example Eshete and Vandewiele, 1992), with the scope of providing simple tools which enable one to compare different models, or applications of the same model at different sites. The focus is on the definition of objective indices of performance, for example based on the distance between real and generated flow duration curves. A point of specific interest is the analysis of the impact of the number of parameters on model efficiency. For synthetic daily streamflow simulation this number can be rather large when seasonally varying parameters are employed (e.g., 67 parameters in the Treiber and Plate (1977) model; 108 parameters in the Aksoy and Bayazit (2000) model). So far, the little attention devoted to the use of such largely parameterized models was justified by the fact that daily samples are rather large in size. However, it is shown here that the temporal dependence of daily discharges strongly reduces the information content of available samples, increasing the relevance of having parsimonious models. A correction term to the mentioned indices of performance is thus introduced to explicitly account for the number of model parameters. An application to 7 time series from basins in North-Western Italy demonstrates the practical usefulness of the proposed techniques (Section 5).

## 2 Shot noise models

A shot noise process is a filtered sequence of independent and instantaneous pulses that can be represented as (e.g., Yue and Hashino, 1999)

$$q(t) = \sum_{i=1}^{N_t} u_i h(t - \tau_i) \tag{1}$$

where $t$ is time, $N_t$ is the total number of occurrences up to time $t$, $u_i$ is the intensity of the $i$-th pulse, $h(t - \tau_i)$ is the response function, and $\tau_i$ is the time of occurrence of the $i$-th pulse. The random times $\{\tau_i\}$ form a Poisson sequence, i.e. the number of occurrences $N_t$ follows a Poisson distribution with rate $\lambda$. The presence of the summation term in (1) comes from the principle of superimposition of effects, which reveals the underlying assumption of linearity of the input-output relation. The daily averaged streamflow, $q_t = \int_{t-1}^{t} q(t)\mathrm{d}t$, has the same representation as the continuous time variable $q(t)$ in (1), the only variation being the response function, that assumes the form $h_d(t - \tau_i) = \int_{t-1}^{t} h(s - \tau_i)\mathrm{d}s$ (Weiss, 1977; Yue and Hashino, 1999).

Application of shot-noise models to daily streamflow generation requires the choice of the probability distribution $F_U(u)$ of the random marks $u_i$, and the definition of the shape of the response function $h$. Since the focus of this paper is on other model details, we make very simple assumptions regarding this two points: following Weiss (1977), $F_U(u)$ is assumed to be exponential with mean value $\alpha$ ($f_U(u) = \frac{1}{\alpha}e^{-u/\alpha}$). This corresponds to using a Poisson-Exponential (PE) model to simulate the effective rainfall input. The assumed system response function is

$$h(t) = c_0\delta(0) + \frac{c_1}{k_1}e^{-t/k_1} + \frac{c_2}{k_2}e^{-t/k_2}, \tag{2}$$

where $\delta(0)$ is the Dirac delta function, $c_0$, $c_1$, $c_2$ are nondimensional constants subject to the constraint $c_0 + c_1 + c_2 = 1$, and $k_1$ and $k_2$ (with $k_2 > k_1$) are storage coefficients (having dimension of time). As pointed out by Murrone et al. (1997), equation (2) corresponds to a conceptual model of the watershed where the effective rainfall is partitioned into a fraction $c_0$, which directly reaches the basin outlet with sub-daily response time, and other components $c_1$ and $c_2$ that flow through two distinct linear reservoirs with storage coefficient $k_1$ and $k_2$ before reaching the outlet. We refer to the paper by Murrone et al. (1997) for a thorough discussion of the advantages and drawbacks of the use of this response function.

Two possible strategies can be followed for the estimation of the model parameters: Weiss (1977) proposes to calculate parameter values according to an analogous of the method of moments, preserving the first two moments and the lag-1 autocorrelation of the observed streamflow sequence. When more than 3 parameters need to be estimated, the modeled and observed moments and autocorrelation are equated also for the discharges aggregated at the monthly scale. In alternative, the model estimation procedure can proceed in four steps: (i) the effective rainfall events occurrences are identified in correspondence to the days when the streamflow increases; the magnitude of these events is initially taken equal to the amount of discharge increment (Battaglia, 1986; Murrone et al., 1997), and the parameters of the PE model ($\lambda$ and $\alpha$) are estimated using the method of moments. For convenience of exposition, this procedure will be called the "Discharge Increments Pulses", or DIP, approach to effective rainfall identification. (ii) Once the effective rainfall sequence is reconstructed, the parameters of the response function $h(t)$ are found by minimizing the sum of quadratic distances between observed and reconstructed data. (iii) A new pulse series is inversely estimated through deconvolution between the actual discharge time series and the estimated response function. (iv) The steps (ii) and (iii) are repeated until convergence (see Murrone et al., 1997). This approach seems to be more attractive than the classical approach of Weiss (1977), since it enables a first evaluation of the model through the comparison of the reconstructed and observed time series. Moreover, the method provides a means to clearly separate the estimation of the stochastic component (effective rainfall) from that of the deterministic one (response function).

# 3 Effective rainfall identification

Although frequently adopted, the mentioned DIP approach presents some drawbacks and difficulties of application: first, the presence of measurement noise can produce small rises in discharge that should not be mistaken for effective rainfall events to avoid distortions in the modeling results. In fact, the DIP procedure tends to produce pulse sequences with very large $\lambda$ values (see Figure

1a), sometimes larger than the average number of rainy days in a year, which clearly contrasts with the physical interpretation of pulses as effective rainfall events. Second, the basic hypotheses that the pulses are Poisson-distributed and that they are mutually independent are often not respected by the estimated sequences. For example, if the independence hypothesis of subsequent peaks is tested by employing Kendall's $\tau$ test at the 5% significance level (see Claps and Laio (2003) for details), the hypothesis is rejected in most of the cases. Analogously, when the dispersion index test of Cunnane (1979) is used to test for Poissonianity, the Poisson hypothesis at the 5% level is often rejected (see Section 5).

A method to derive a more appropriate pulse sequence is therefore necessary. Following Claps and Laio (2003) the pulses can be identified by following a filtered peak over threshold (FPOT) procedure, summarized as follows. (i) The peak events are found in correspondence to all of the local maxima of the daily discharge time series. (ii) A sequence of filtered peaks (FP) is obtained by subtracting from each peak the discharge measured at the first minimum preceding the event. A similar approach to the selection of pulse intensities was used by Pegram (1980). (iii) A threshold filter is applied to the FP sequence to retain only the significant peaks. (iv) The appropriate threshold $s$ that filters out noisy peaks is selected by testing the independence of the peaks in the sample (Kendall's $\tau$ test) and the distribution of occurrences (Cunnane (1979) test for the Poisson distribution). The threshold $s$ is gradually increased until the two tests are jointly met. The convergence towards independence for large

$s$ values can be attributed to the increase of the distance between subsequent peaks. Analogously, when the threshold $s$ is increased the numbers of crossings of $s$ in disjoint time intervals tend to become independent random variables, and this guarantees the asymptotic convergence towards Poissonianity. A formal demonstration of the asymptotic convergence towards independence and Poissonianity can be found in Cramer and Leadbetter (1967, pp. 256-271) for the case of a Gaussian stochastic process.

The adoption of the FPOT approach allows one to avoid the deconvolution step in the peaks identification procedure, as used by Murrone et al. (1997), with substantial advantages in terms of simplicity and robustness of the procedure. Moreover, the method allows one to obtain a pulse sequence which automatically meets the independence and Poissonianity requirements. However, the number of selected peaks is reduced to 5-20 per year (Figure 1b), which probably underestimates the actual number of effective rainfall events. This result is derived on a larger number of runoff series analyzed by Claps and Laio (2003) and can be a clue of the inadequacy of the Poisson independent model in the correct reproduction of the effective rainfall behavior. More complicate models have been proposed which account for the mutual dependence of peaks (e.g. Markov-chain models: see for example Vandewiele and Dom (1989), Aksoy and Bayazit (2000), Xu et al. (2001), Aksoy (2003), Aksoy (2004)) and for the clustering of rainy days (Neymann-Scott models or similar, e.g Cowpertwait and O'Connell, 1992). However, the increased complication of such models can hardly be properly supported, mainly due to the shortness of the available time series (see also the discussion in Section 4.2). This makes the simpler, yet possibly erroneous, Poisson independent model a good choice in many cases.

A last consideration is needed with regard to the problem of seasonality: a periodic behavior is often detected in riverflow time series, and its origin is traced back to the presence of seasonality in the climatic forcing. Following these lines, the response function in shot noise models is often kept constant throughout the year, while the parameters of the PE models for effective rainfall ($\alpha$ and $\lambda$) are allowed to vary from season to season. In the simpler approach different values of $\lambda_\tau$ and $\alpha_\tau$ are estimated in periods of the year wherein they are supposed to remain constant (see Murrone et al. (1997) for a detailed description of this and other approaches). When $\lambda$ varies in time, the basic hypothesis becomes that the pulse sequence is a non-homogeneous Poisson process; the Kendall's $\tau$ test is still appropriate in this case, as also is the Cunnane (1979) test, since the number of events in any time window $t_w$ is Poisson-distributed even when $\lambda$ varies inside $t_w$ (Ross, 1996, p. 79). In contrast, when the average event intensity $\alpha$ varies in time, the sample should be split in seasons wherein $\alpha$ is supposed to remain constant before applying a goodness-of-fit test for the intensities distribution. This produces a major loss of significance of the test, due to the reduction of the sample sizes. For this reason the test for the distribution of pulse intensities, which was present in the original FPOT procedure as a further constraint to the choice of the threshold $s$ (Claps and Laio, 2003), is not considered here.

# 4 Model evaluation

It is commonly accepted that the value of a stochastic model for data generation can be defined from its ability to reproduce the statistical features of the observed record in the generated time series. Typically, this ability is evaluated by comparing the moments of the observed and simulated series at different aggregation scales (e.g., Weiss, 1977) or, in some cases, the full probability density functions (e.g., Vandewiele and Dom, 1989). Two problems affect this approach to model verification, and are considered in the following subsections: the first relies on the lack of objective statistics for the comparison, the value of the model being often judged by a subjective visual analysis of real and generated data. The second problem relies on the fact that model parsimony is seldom considered explicitly.

## 4.1 Indices of performance

Some confusion in the assessment of model performances derives from the use of both the reconstructed and the generated time series as possible terms of comparison with the real observations. The reconstructed time series is obtained from the convolution of the estimated input pulse sequence with the system response function. Reconstructed and observed series are easily compared using any measure of reciprocal distance between the two signals (see Figure 2a-b). If this comparison can be significant to validate the model hypotheses regarding the form of the response function, it has in contrast no value for judging the real ability of the model to generate time series similar to the observed one.

The real value of the model must rather be judged by comparing observed and generated sequences. The time correspondence of the peaks of the two series is obviously lost in this case, and the distance between contemporary values becomes meaningless.

In this latter case, a better option is to consider the reciprocal distance between the observed and synthetic cumulative distribution functions, represented as flow duration curves. In order to build the flow duration curves, the observed discharge dataset of size $n$ is sorted in ascending order, and an empirical frequency of occurrence $F_{(i)} = \frac{i}{n+1}$ is assigned to the $i$-th order statistic in the sample, $q_{o(i)} = q_o(F_{(i)})$. The same procedure is followed for the generated sample, whose size $N$ is much larger than $n$, obtaining $F_{(j)} = \frac{j}{N+1}$ and $q_{g(j)} = q_g(F_{(j)})$. The flow duration curves are given by the $F_{(i)}$ and $F_{(j)}$ values plotted versus $q_{o(i)}$ and $q_{g(j)}$; the curves are not shown in Figure 2 because the resulting real and generated graphs are nearly indistinguishable. To evaluate the distance between these curves, a value of generated discharge with the same frequency of occurrence as $q_{o(i)}$, namely $q_g(F_{(i)})$ need to be found. $q_g(F_{(i)})$ is the empirical quantile corresponding to $F_{(i)}$, i.e. the $j$-th order statistic in the generated sample, with $j = \frac{N+1}{n+1}i$ ($j$ can be conveniently approximated to the closer integer because of the large sample size $N$). The mean squared distance between the two flow duration curves (”model error variance”) is then evaluated as

$$s^2 = \frac{\sum_{i=1}^{n} \left[q_o(F_{(i)}) - q_g(F_{(i)})\right]^2}{n}. \quad (3)$$

In order to facilitate the comparison between different applications, the value in (3) is rescaled by the variance $\sigma^2$ of the observed discharges, and an index of performance similar to the coefficient of determination of linear regression models is proposed as a measure of model adequacy:

$$I_1 = 1 - \frac{s^2}{\sigma^2}. \tag{4}$$

The closer $I_1$ to its limit value 1, the more adequate is the model to represent the flow duration curve. Other measures of model adequacy are easily defined by considering other characteristics of the observed and generated sequences. For example, to test the correct reproduction of the annual maxima (AM) statistics (see Figure 2e-f), a specific index can be defined as

$$I_2 = 1 - \frac{s_{AM}^2}{\sigma_{AM}^2}. \tag{5}$$

where $s_{AM}^2 = \frac{\sum_{j=1}^{k} \left[q_o^{AM}(F_{(j)}) - q_g^{AM}(F_{(j)})\right]^2}{k}$ is the mean squared distance between the empirical frequency curves of the $k$ observed AM, $q_o^{AM}$, and the corresponding curves for generated data, $q_g^{AM}$, and $\sigma_{AM}^2$ is the variance of the observed AM sample.

## 4.2 Model parsimony

A major problem with the indices of performance approach is that the number of parameters is not explicitly accounted for in model evaluation. This implies a tendency to favor models with a large number of adjustable parameters, without considering the sampling errors associated with parameter estimates (Vogel and Stedinger, 1988). As a limit example, when non-parametric techniques are employed (e.g. Sharma et al., 1997; Tarboton et al., 1998), one can obtain a perfect equivalence of the frequency distribution of the real and generated series. The indices $I_1$ and $I_2$ would equal 1 in this case, but the generated samples are little more than duplicates of the real time series, and they tend to mechanically reproduce sampling errors. Speaking in terms of extreme-value frequency analysis, a non-parametric method can be considered as the analogous of a polynomial passing through all of the points in the sample, while a good data generation model is the analogous of a probability distribution fitted to the data. When the latter is correctly chosen, it has clear advantages towards the former, for example the possibility to provide adequate extrapolations. Moreover, the smaller is the number of parameters, the more robust is the model; the above considerations should help clarify the need to explicitly account for model parsimony.

As a first guess, one could think of simply correcting equation (3) with a similar approach as that followed in multivariate regression analysis, i.e. by multiplying the error variance $s^2$ by a term $\frac{n}{n-p}$, where $p$ is the number of adjustable parameters. In non-parametric models, $p$ is instead the number of constraints imposed to the kernel probability density estimator, which approximates $n$ when the bandwidth tends to zero (Sharma et al., 1997). With a sample size $n$ of the order of $10^3 - 10^4$ (365 multiplied by the number of years of observations), the role of the number of parameters would seem negligible at a first sight, $p$ being of the order of 10-100. However, the daily discharge time series are strongly autocorrelated, and this causes a dramatic drop of the information content in a sample of given size. This loss of information was accounted for in an ad hoc manner by Eshete and Vandewiele (1992), by constructing the frequency curves using a subsample of observations consisting of one value each two months, which corresponds to having a sample size reduced of a factor 6/365. A manner of facing this problem with a more systematic approach is by defining an "equivalent" sample size, $n_e$, as the size of a sample of independent data having the same marginal distribution and providing the same amount of information as the original autocorrelated sample. $n_e$, rather than $n$, will then be considered to correct $s^2$ for the presence of adjustable parameters. A similar approach was followed, in a different context, by Yevjevich (1972, pp. 45-46).

In order to determine $n_e$, we make the hypothesis that the amount of information of a given sample can be measured in terms of the standard error of estimation of the sample moments: the smaller the standard error, the larger the information provided by the sample. Define the $r$-th moment as $y_r = \frac{1}{n}\sum_{i=1}^{n} q_i^r$ (non central moments are considered for simplicity of exposition), and consider the case when all the random variables $q_i$, $i = 1, .., n$ are identically distributed, with mean $\mu$, variance $\sigma^2$, and covariance $\mathrm{cov}(q_i, q_j) = \sigma^2\rho(i-j)$, where $\rho(\ )$ is the autocorrelation coefficient ($\rho(0) = 1$).

Standard sampling error theory, based on the Taylor expansion of the non-linear function $q_i^r$ around the mean, provides the relation (e.g., Kendall and

Stuart, 1977, pp. 246-247)

$$\mathrm{var}(y_r) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial y_r}{\partial q_i}\bigg|_{q_i=\mu} \frac{\partial y_r}{\partial q_j}\bigg|_{q_j=\mu} \mathrm{cov}(q_i, q_j) = \left[\frac{r\mu^{r-1}}{n}\right]^2 \sigma^2 \sum_{i=1}^{n}\sum_{j=1}^{n} \rho(i-j). \tag{6}$$

Considerations of symmetry of $\rho(k)$, analogous to those invoked by Vanmarcke (1983, p. 116), enable one to simplify (6) as

$$\mathrm{var}(y_r) = \frac{r^2\mu^{2r-2}}{n^2}\sigma^2 \left(2\sum_{i=1}^{n}(n-i)\rho(i) + n\right) \tag{7}$$

Equation (7) with $r = 1$ provides the exact variance for the sample mean $y_1$ (e.g. Hosking, 1996), while for higher order moments Equation (7) is an approximation valid to the order of $n^{-2}$. When the $q_i$ are independent, the approximate standard error in (7) reads

$$\mathrm{var}_{\mathrm{ind}}(y_r) = \frac{r^2\mu^{2r-2}}{n}\sigma^2. \tag{8}$$

The reduction of information content of a correlated sample with respect to an independent sample of the same size can be determined by taking the ratio of the standard errors of estimation of the sample moments,

$$r_{inf} = \frac{\mathrm{var}_{\mathrm{ind}}(y_r)}{\mathrm{var}(y_r)} \tag{9}$$

The equivalent sample size is finally determined as $n_e = r_{inf} \cdot n$. Straightforward calculations allow one to find

$$n_e = \frac{n^2}{n + 2\sum_{i=1}^{n}(n-i)\rho(i)}. \tag{10}$$

Equation (10) is rather general, since it neither depends on the order $r$ of the considered moment, nor on the mean and variance of the random variables, $q_i$. However, Equation (10) depends on an initial hypothesis that the random variables $q_i$ are identically distributed, an hypothesis that is not verified for periodic time series. In this case, the actual value of $r_{inf}$ would depend on the order of the considered sample moment. The computation of $r_{inf}$ is thus related to the standard error of estimation of the sample mean $y_1$ only (not to that of all moments $y_r$). Considering an annual periodicity (which is typical for geophysical time series), equation (9) then reads,

$$r_{ann} = \frac{\mathrm{var}_{\mathrm{ind}}(y_1)}{\mathrm{var}(y_1)} = \frac{\sum_{i=1}^{365} \mathrm{var}(q_i)}{\sum_{i=1}^{365} \sum_{j=1}^{365} \mathrm{cov}(q_i, q_j)} \tag{11}$$

where $\mathrm{var}(q_i)$ is the variance of the discharge in the $i$-th day of the year, which is variable with $i$ for a periodic process, and $\mathrm{cov}(q_i, q_j)$ is the covariance between the discharge in the $i$-th and in the $j$-th days of the year, i.e the covariance calculated "over realization" in the terminology of Mitosek (2000). The reduction of information content of a dependent sample can be considerable for riverflows, with $r_{ann}$ coefficients of the order of 0.02-0.1 (see Table 3 in Section 5).

In conclusion, the error variance $s^2$ in (3) is corrected as $s^2 \frac{n_e}{n_e - p}$, and the index $I_1$ becomes

$$I_1(p) = 1 - \frac{s^2}{\sigma^2} \frac{n_e}{n_e - p}. \tag{12}$$

Analogously, the effect of the number $p$ of adjustable parameters on the second index yields

$$I_2(p) = 1 - \frac{s_{AM}^2}{\sigma_{AM}^2} \frac{n_e}{n_e - p}. \tag{13}$$

# 5 Application

7 time series of daily runoff are considered in the present analysis, relative to drainage basins located in the North-West of Italy. The basins analyzed cover a variety of climatic and geologic features, but the morphology and climate of the Alps influence the majority of them. We report in Table 1 some characteristic features of the drainage basins and of the daily runoff time series. Note the marked differences in drainage areas $A$ and average elevations $h_m$. Record length in years, $n_y$, also shows a great variability.

Table 2 reports the estimated parameters for the shot noise model when the FPOT (values in italic) or DIP approaches are adopted to obtain an initial estimate of the effective rainfall, and the procedure described in Section 2 is followed for determining the coefficients $c_0$, $c_1$, $c_2$, $k_1$ and $k_2$ of the response function. We refer to Giordano (2004) for an extended graphical presentation of the modeling results. The main differences between the two procedures are in the estimates of $\alpha$ and $\lambda$ (last two columns), with the DIP model producing pulse sequences with more frequent and smaller events. The reasons behind these differences are fully explained in Section 3. The results in Table 2 refer to parameters estimated on an annual time basis, but analogous results are found when $\alpha$ and $\lambda$ are allowed to vary on a monthly scale (see Figures 2c-d). The differences in $\alpha$ and $\lambda$ values also affect the storage coefficients $k_1$ and $k_2$, that tend to be larger when the FPOT approach is adopted, probably as a consequence of the reduced number of peaks.

As for the evaluation of the models, the independence and Poissonianity hypotheses deserve a first comment: when applied to the DIP pulse sequences, both hypotheses are rejected at the 5% significance level for all of the 7 basins. The independence condition in particular is not even met when the significance level is decreased to 1% or lower. In contrast, both tests are automatically passed when the FPOT approach is followed. Regarding the performance indices, Table 3 reports a summary of the results for the 7 considered river basins. The first two columns show the information reduction coefficient $r_{ann}$ in (11) and the equivalent sample size $n_e$, that can be as low as 140 days despite the fact that the available sample size is never smaller than 12 years, or 4380 data. The remaining columns in Table 3 report the indices of performance $I_1$ and $I_2$, calculated in three different cases: first (columns 3-4) the observed and reconstructed data are compared: these indices have no real value in the judgement of the quality of the models, but they are reported for the sake of comparison of the DIP and FPOT approaches (see below). The other columns refer to the comparison of observed and generated data, either with annual parameters (columns 5-6), in which case $p$=6, or with seasonally varying $\alpha$ and $\lambda$ values (13 "months" of 28 days each, for a total of 30 parameters: results in columns 7-8). Again, the values obtained with the FPOT procedure are reported in italic.

The $I_1$ index shows a smaller variability than the $I_2$, but for both the general tendency is for the DIP approach to produce better results in the reconstruction phase, while the FPOT procedure is generally better in generation. The reason behind this behavior is that the increased number of pulses in the DIP approach brings the reconstructed curve closer to the observed one, but then the mentioned problems with the modeling hypotheses affect the results in the generation phase. In fact, one generates an effective rainfall sequence with independent Poisson distributed pulses, but estimates $\alpha$ and $\lambda$ from a pulse sequence that does not have these characteristics, introducing a distortion in the model. An additional comment regards the use of the seasonal modeling of effective rainfall, which produces a general improvement of modeling results despite the increased number of parameters. Finally, it is worth mentioning that the reproduction of the annual maxima statistics with the FPOT approach is satisfactory even when seasonality is not accounted for, which represents a substantial advantage towards other commonly adopted models (see Katz et al. (2002) for a discussion).

## 6 Conclusions and future developments

The rationale behind daily streamflow generation by shot noise models is discussed here, with reference to some crucial aspects of model specification and verification. In particular, the identification of the effective rainfall sequence and the evaluation of model performances in relation to parsimony have been examined, also by means of model application to a set of 7 stations. The obtained results have indicated the necessity, in future applications, to provide greater attention to the verification of statistical assumptions and to the objective presentation of model results. More in detail, issues that require additional investigations are: (i) the use of an exponential (or other) distribution for pulse intensities (see Section 2) needs to be verified using appropriate goodness-of-fit tests. Preliminary analyses in this direction have shown that the exponential distribution is seldom adequate, while two-parameter distributions (e.g. Generalized Pareto as in Claps and Laio (2003) or Gamma as in Murrone et al. (1997) or Xu et al. (2002)) provide a better representation, even if at the expense of an increase of the number of parameters. (ii) In order to verify the model performances, other statistics similar to the indices $I_1$ and $I_2$ could be proposed, possibly based on the comparison of the autocorrelation structure of the observed and generated time series, or on other characteristics more strictly connected to the final use of the models (statistics of flow volumes, annual minima, ...). For a fair comparison between different model structures (shot-noise versus Markov chain models, seasonally varying versus constant parameters, ...), special attention should be provided to the number of model parameters, for example following the procedure described in Section 4. (iii) The choice of the system response function should be derived from a better comprehension of the relations between the model structure and the physical mechanisms in the rainfall-runoff transformation. Appropriate tools to validate this choice should be developed, for example based on tests of linearity versus non-linearity, or on specific data-based verifications of the adopted shape, possibly starting from well established methods to determine the empirical response function (e.g. Bruen and Dooge, 1984).


## Acknowledgements

The financial support of the National Research Council (CNR-GNDCI grant 03.00022.GN42) and of the Regional Environmental Protection Agency (ARPA Piemonte, Interreg 3A project) are gratefully acknowledged.

# References


Aksoy H. Markov chain-based modeling techniques for stochastic generation of daily intermittent streamflows. Adv Water Resour 2003;26:663-671.

Aksoy H, Bayazit M. A model for daily flows of intermittent streams, Hydrol Processes 2000;14:1725-44.

Aksoy H, Pulse generation models for daily intermittent streamflows, Hydrol Sci J 2004;49(3):399-411.

Battaglia F. Stochastic models for the representation and prediction of riverflows, Dept. of Statistics Papers, No. 3, Univ. "La Sapienza", Rome, Italy, 1986. (in Italian)

Bernier J. Inventaire des modeles de processus stochastiques applicables a la description des debits journaliers des rivieres. Revue Inst Intern Stat 1970;38(1):49-61.

Bruen M, Dooge JCI. An efficient and robust method for estimating unit hydrograph ordinates. J Hydrol 1984;70:1-24.

Claps P, Laio F. Can continuous streamflow data support flood frequency analysis? An alternative to the Partial Duration Series approach. Water Resour Res 2003;39(8):1216, doi:10.1029/2002WR001868.

Cowpertwait PSP, O'Connell PE. A Neymann-Scott shot noise model for the generation of daily streamflow time series, in Advances in theoretical hydrology, ed. JP O'Kane, Elsevier, Amsterdam, pp.75-94, 1992.

Cramer H, Leadbetter MR, Stationary and related stochastic processes, John Wiley and sons, New York, USA, 1967.

Cunnane C. A note on the Poisson assumption in partial duration series models. Water Resour Res 1979;15(2):489-494.

Eshete Z, Vandewiele GL. Comparison of non-Gaussian multicomponent and periodic autoregressive models for river flow. Stochastic Hydrol Hydraul 1992;6:223-38.

Giordano A. An application of shot noise models to daily riverflow time series in North-Western Italy, DITIC working paper 2-2004, Polytechnic of Turin, Turin, Italy, 2004. Available at: http://www.idrologia.polito.it/∼giordano (in Italian).

Hosking JRM. Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. J Econometrics 1996;73: 261-84.

Katz RW, Parlange MB, Naveau P. Statistics of extremes in hydrology Adv Water Resour 2002;25:1287-1304.

Kendall MG, Stuart A. The advanced theory of statistics. Vol 1. Distribution theory, 4th edition, Griffin and Co., London, 1977.

Koch RW. A stochastic streamflow model based on physical principles. Water Resour Res 1985;21(4):545-53.

Lawrance AJ, Kottegoda NT. Stochastic modelling of riverflow time series. J Royal Stat Soc A 1977;140(1):1-47.

Mitosek HT, On stochastic properties of daily river flow processes. J Hydrol 2000; 228:188-205.

Murrone F, Rossi F, Claps P. Conceptually-based shot noise modeling of streamflows at short time intervals. Stoch Hydrol Hydraul 1997;11:483-510.

Pegram GGS. A continuous streamflow model. J Hydrol 1980;47: 65-89.

Ross S. Stochastic processes, 2nd edition, John Wiley and sons, New York, USA, 1996.

Salas JD. Analysis and modeling of hydrologic time series, in Handbook of Hydrology, edited by D.R. Maidment, McGraw-Hill, New York, USA, 1993.

Sharma A, Tarboton D, Lall U. Streamflow simulation: a nonparametric approach. Water Resour Res 1997;33(2):291-308.

Tarboton DG, Sharma A, Lall U. Disaggregation procedures for stochastic hydrology based on nonparametric density estimation. Water Resour Res 1998;34(1):107-19.

Treiber B, Plate EJ. A stochastic model for the simulation of daily flows. Hydrol Sci Bull 1977;22:175-192.

Vandewiele GL, Dom A. A non-Gaussian multicomponent model for river flow. Water Resour Res 1989;25(3):397-404.

Vanmarke E. Random fields, The MIT press, Cambridge (MA), USA, 1983.

Vogel RM, Stedinger JR. The value of stochastic streamflow models in overyear reservoir design applications. Water Resour Res 1988;24(9):1483-90.

Weiss G. Shot noise models for the generation of synthetic streamflow data. Water Resour Res 1977;13(1):101-8.

Xu ZX, Schumann A, Brass C, Li J, Ito K. Chain-dependent Markov correlation pulse model for daily streamflow generation. Adv Water Resour 2001;24:179-85.

Xu ZX, Schultz GA, Schumann A. A conceptually based stochastic point process model for daily stream-flow generation. Hydrol Processes 2002;16:3003-17.

Yevjevich V. Stochastic processes in hydrology, Water Resources Publications, Littleton (CO), USA, 1972.

Yue S, Hashino M. A stochastic model for deriving the basic statistics of J-Averaged streamflow. Water Resour Res 1999;35(10):3127-37.

## Captions

Table 1. Characteristic features of the 7 drainage basins considered in the application. Drainage area $A$, average elevation $h_m$, and record length $n_y$ are reported for each station.

Table 2. Estimated parameters of the system response function ($c_0$, $c_1$, $c_2$, $k_1$ and $k_2$) and of the effective rainfall PE model ($\alpha$ and $\lambda$) for the 7 considered riverflow time series, when the DIP of the FPOT (values in italic) procedures are used for effective rainfall identification.

Table 3. Performance indices $I_1$ (Equation (12)) and $I_2$ (Equation (13)) for the 7 considered riverflow time series, when the DIP of the FPOT (values in italic) procedures are used for effective rainfall identification. The first two columns contain the annual information reduction coefficient, from Equation (11), and the effective sample size.

Figure 1. Comparison of estimated effective rainfall sequences with the DIP (a) and FPOT (b) approaches. The example is relative to river Tanaro at Nucetto (see Table 1).

Figure 2. Shot noise model performances for river Tanaro at Nucetto (see Table 1): observed and reconstructed time series (a-b), generated data with annual (ann.) or monthly parameters (seas.) (c-d), and annual maxima statistics (e-f). Left panels refer to the DIP approach to effective rainfall estimation, right panels to the FPOT approach.

Table 1: *Claps et al.* [2004]

| Name | $A$ [Km$^2$] | $h_m$ [m a.s.l] | $n_y$ [years] |
|---|---|---|---|
| Ayasse at Champorcher | 42 | 2392 | 22 |
| Borbera at Baracche | 202 | 880 | 14 |
| Bormida at Cassine | 1483 | 493 | 12 |
| Chisone at S. Martino | 580 | 1751 | 26 |
| Orco at Pont Canavese | 617 | 1930 | 29 |
| Scrivia at Serravalle | 605 | 695 | 14 |
| Tanaro at Nucetto | 375 | 1227 | 29 |

Table 2: *Claps et al* [2004]

| | $c_0$ | $c_1$ | $c_2$ | $k_1$ [d] | $k_2$ [d] | $\alpha$ [mm] | $\lambda$ [1/year] |
|---|---|---|---|---|---|---|---|
| Ayasse at Champorcher | 0.06 | 0.08 | 0.86 | 1.2 | 19.7 | 25.9 | 48.4 |
| | *0.05* | *0.17* | *0.78* | *7.0* | *32.5* | *87.0* | *14.4* |
| Borbera at Baracche | 0.10 | 0.20 | 0.70 | 2.4 | 53.6 | 49.9 | 16.6 |
| | *0.08* | *0.19* | *0.73* | *3.3* | *76.3* | *108.6* | *7.6* |
| Bormida at Cassine | 0.12 | 0.48 | 0.40 | 2.2 | 55.8 | 31.2 | 16.6 |
| | *0.12* | *0.38* | *0.50* | *2.5* | *89.1* | *63.1* | *8.1* |
| Chisone at S. Martino | 0.06 | 0.28 | 0.66 | 6.3 | 220.0 | 40.1 | 17.1 |
| | *0.04* | *0.14* | *0.82* | *5.1* | *126.0* | *143.0* | *4.8* |
| Orco at Pont Canavese | 0.10 | 0.18 | 0.72 | 2.1 | 61.9 | 37.1 | 26.6 |
| | *0.06* | *0.13* | *0.81* | *3.1* | *135.0* | *130.5* | *7.5* |
| Scrivia at Serravalle | 0.12 | 0.34 | 0.54 | 2.0 | 43.7 | 27.2 | 28.9 |
| | *0.12* | *0.28* | *0.60* | *2.7* | *56.9* | *52.2* | *15.0* |
| Tanaro at Nucetto | 0.10 | 0.25 | 0.65 | 2.0 | 135.3 | 43.0 | 20.4 |
| | *0.09* | *0.21* | *0.70* | *2.9* | *205.7* | *93.9* | *9.3* |

Table 3: *Claps et al* [2004]

| | | | reconstructed | | gen. annual | | gen. seasonal | |
|---|---|---|---|---|---|---|---|---|
| | $r_{ann}$ | $n_e$ [d] | $I_1$ | $I_2$ | $I_1$ | $I_2$ | $I_1$ | $I_2$ |
| Ayasse at Champorcher | 0.027 | 220 | 1.00 | 0.95 | 0.74 | -1.23 | 0.95 | -0.07 |
| | | | *0.99* | *0.92* | *0.95* | *0.82* | *0.99* | *0.91* |
| Borbera at Baracche | 0.027 | 140 | 0.97 | 0.95 | 0.85 | 0.37 | 0.87 | 0.56 |
| | | | *0.97* | *0.99* | *0.93* | *0.71* | *0.94* | *0.79* |
| Bormida at Cassine | 0.090 | 393 | 0.98 | 0.94 | 0.87 | 0.40 | 0.90 | 0.55 |
| | | | *0.95* | *0.97* | *0.96* | *0.92* | *0.96* | *0.94* |
| Chisone at S. Martino | 0.030 | 282 | 0.94 | 0.95 | 0.80 | 0.32 | 0.85 | 0.48 |
| | | | *0.97* | *0.99* | *0.97* | *0.83* | *0.97* | *0.88* |
| Orco at Pont Canavese | 0.042 | 445 | 1.00 | 0.99 | 0.94 | 0.49 | 0.98 | 0.83 |
| | | | *0.98* | *0.98* | *0.99* | *0.90* | *0.99* | *0.95* |
| Scrivia at Serravalle | 0.059 | 300 | 0.99 | 0.88 | 0.92 | -0.32 | 0.94 | -0.02 |
| | | | *0.98* | *0.98* | *0.96* | *0.92* | *0.97* | *0.92* |
| Tanaro at Nucetto | 0.045 | 475 | 0.96 | 0.93 | 0.88 | 0.24 | 0.93 | 0.57 |
| | | | *0.94* | *0.99* | *0.95* | *0.82* | *0.95* | *0.87* |

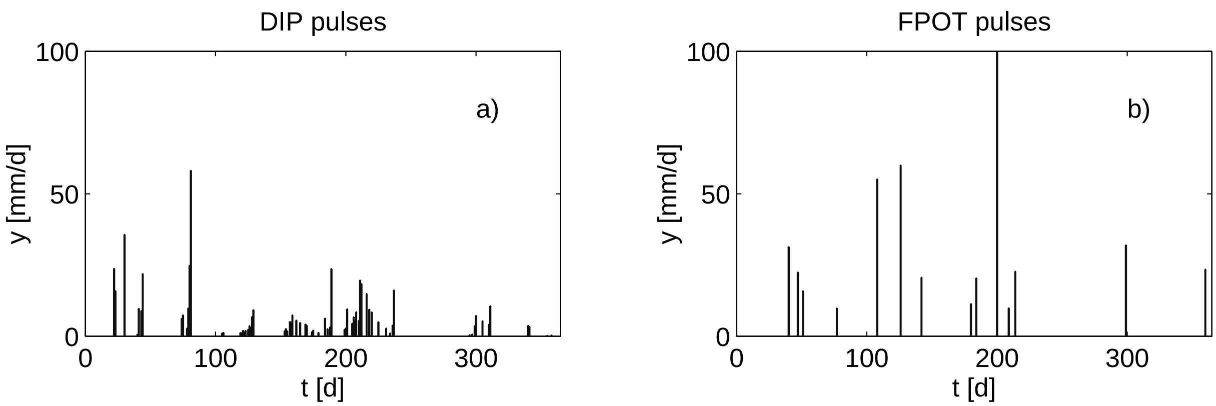


Figure 1: Claps et al. [2004]

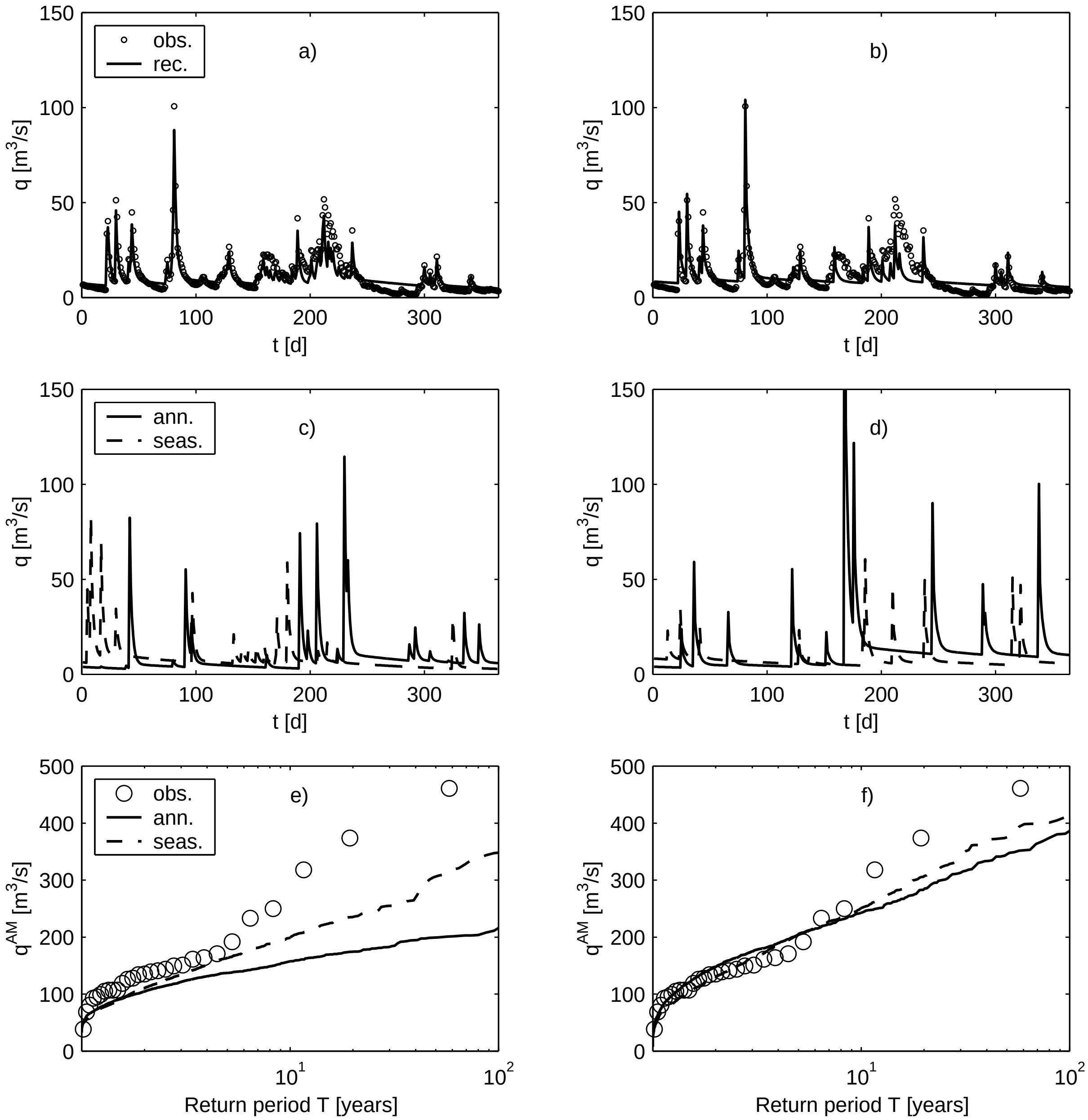


Figure 2: Claps et al. [2004]