Conceptually-Based Shot Noise Modelling of Streamflows at Short Time Interval

FEDERICO MURRONE, FABIO ROSSI

Dipartimento di Ingegneria Civile, Università di Salerno, Via Ponte Don Melillo, Fisciano, I-84084 Italy

PIERLUIGI CLAPS

Dipartimento di Ingegneria e Fisica dell'Ambiente, Università della Basilicata, Potenza, I-85100 Italy

ABSTRACT

A conceptual-stochastic approach to short time runoff data modelling is proposed, according to the aim of reproducing the hydrological aspects of the streamflow process and of preserving as much as possible the dynamics of the process itself. This latter task implies preservation of streamflow characteristics at higher scales of aggregation and, within a conceptual framework, involves compatibility with models proposed for the runoff process at those scales. At a daily time scale the watershed response to the effective rainfall is considered as deriving from the response of three linear reservoirs, respectively representing contributions to streamflows of large deep aquifers, with over-year response lag, of aquifers which run dry by the end of the dry season and of subsurface runoff. The surface runoff component is regarded as an uncorrelated point process. Considering the occurrences of effective rainfall events as generated by an independent Poisson process, the output of the linear system represents a conceptually-based multiple shot noise process. Model identification and parameter estimation are supported by information related to the aggregated runoff process, in agreement to the conceptual framework proposed, and this allows parameter parsimony, efficient estimation and effectiveness of the streamflow reproduction. Good performances emerged from the model application and testing made with reference to some daily runoff series from Italian basins.

Key words: Conceptual-stochastic models; Shot Noise; Streamflow simulation; Time aggregation.

1. Introduction

Many hydrologic applications, from the water resources management to the modelling of water quality, require synthetic time series of streamflows, as those generated by stochastic models. Depending on the particular problem at hand, reproduction of the streamflow process is desired at different resolution. In the past thirty years, several classes of stochastic models were proposed, which generally looked at each time scale of aggregation individually.

This paper focuses on the reproduction of short-time runoff data (say, daily to weekly) according to two main requirements: 1. a physically-consistent model structure, which has significant advantages when dealing with inadequate data; 2. the model capability to coexist with corresponding models suited for more aggregated scale, in a homogeneous framework for streamflow modelling.

Short time streamflows are characterised by the presence of the intermittent pattern of rain events and by the skewed nature of the hydrographs, with sudden discharge increases and slow recessions. These features prevent the use of ARMA-type of models [*Box and Jenkins*, 1976], which are successfully applied to monthly and annual data that – as required in the canonical ARMA framework – can be somewhat reduced to stationary continuous processes.

Models which explicitly consider the intermittent pattern of rain events date back to *Bernier* [1970], who introduced in streamflow modelling the *filtered Poisson process* [*e.g. Parzen*, 1962]. The structure of this process, usually referred to as *shot noise*, consists of a point process, that reproduces the occurrence of effective rainfall events, which acts as the input of a system that is representative of the transformations operated by the watershed. Runoff is obtained by filtering the input through the system response function.

The first comprehensive work on shot noise models of runoff is due to *Weiss* [1973, 1977], who introduced a model in which effective rainfall events are reproduced through a Poisson process of occurrences coupled to exponentially-distributed intensities. Response of the watershed consists of two components, one representing the base flow and the other the direct runoff. Both are assumed to have an exponential response to precipitation. A two-component shot noise process results from this scheme, and the method of moments is proposed for parameter estimation.

A bivariate version of Weiss' model was proposed by *Koch* [1985], with more strict correspondence between parameters of the physical system and model parameters. Another, more recent, improved variant of Weiss' model is due to *Cowpertwait and O'Connell* [1992], who proposed the Neyman-Scott model to reproduce the effective rainfall process.

Between *Weiss* [1973] and *Cowpertwait and O'Connell* [1992], several other approaches, often based on the shot noise formulation, were proposed for short-time runoff modelling. In a rough classification of the literature on the subject one can distinguish: *a*. models in which the response function derives, as in *Weiss* [1973], from a linear conceptual scheme of the watershed [*e.g. Pegram*, 1980; *Hino and Hasebe,* 1981; *Vandewiele and Dom,* 1989]; *b*. non-linear or non-parametric models [*Treiber and Plate,* 1977; *Yakowitz,* 1979]. Some of these models use Markov processes as input [*Treiber and Plate,* 1977; *Yakowitz*, 1979; *Vandewiele and Dom,* 1989] and often the input process is reconstructed by inverse estimation [*Treiber and Plate,* 1977; *Hino and Hasebe,* 1981; 1984; *Battaglia,* 1986; *Kron et al.*, 1990; *Wang and Vandewiele,* 1994] as opposed to Weiss' approach, in which parameters of the input model are directly estimated from runoff through the moments method.

As emerges from analysis of the literature, much work was made in the attempt to reproduce the peculiar statistical features of short-time streamflows. Most of the models proposed, however, are based on the assumption that enough data is available for the process under study. Therefore, much work is still needed toward the achievement of sufficient model performances in condition of scarcity or lack of data, as well as in combination with corresponding models to be used for generation of data at more aggregated scales. The latter point becomes important when simulation of runoff data in a water resources system is required at different scales, due to different characteristics of the system elements or to different goals of the planning.

When both short-time and aggregated data are required, one is practically forced to use different models, because models for, say, daily runoff pay little or no attention to statistics referred to higher scales.

As an attempt to give an answer to the above-mentioned points, the building of a shot noise model was conceived in this paper under the framework of models with stochastic input and deterministic conceptual watershed response. The substantial step forward with respect to previous similar models is in the consideration that, to comply with the process dynamics at all of the scales of interest, the distinction in the watershed response of only two components (one fast and one slow) is inadequate. To provide more detail in the watershed response, information is transferred from aggregated data both in model identification and in parameter estimation, according to the approach proposed by *Claps et al.* [1993]. This scheme provides parsimonious and efficient models, with parameters that can be validated and related to physical variables in view of their determination in ungauged stations.

The main features of the structure of the proposed model are presented below.

2. Model Structure

2.1. Stochastic model of the effective rainfall

The task of selecting a stochastic process representative of the *effective rainfall* presents serious difficulties, particularly if short aggregation interval are considered, due to the fact that the natural process is unobserved. At the stage of the selection of a broad class of stochastic models, it can be sufficient to assume that effective rainfall retains most of the stochastic characters of total rainfall. This assumption, however, needs to be substantiated properly at the application stage, and this at least requires that the series of the effective rainfall must be inversely estimated. This point will be discussed in the next section.

In our approach to model building, the choice has gone toward the class of point processes, which are physically-based and can be expressed in functional form. The alternative class of Markovian processes is not as attractive, because they are discrete processes which do not preserve their structure in the aggregation.

Among point processes we have considered the classical marked Poisson white noise processes [e.g., *Todorovic and Yevjevich*, 1969; *Eagleson*, 1978] and the ones reproducing the rainfall events as a sequence of storms made of clusters of rain cells, such as the Neyman-Scott or the Bartlett-Lewis processes [e.g., *Kavvas and Delleur*, 1981; *Rodriguez-Iturbe et al.*, 1984; 1987]. The occurrence of rainfall events or, in cluster processes, the occurrence of storms, are usually modelled by a Poisson point process. The choice of the particular model to use is mainly driven by the scale of aggregation considered, because serial correlation decreases with the aggregation.

For aggregation scales up to 1-2 days, rainfall data present significant autocorrelation, which is generally reproduced correctly by cluster-type models (see, e.g., *Bo et al.*, 1994). However, it is not guaranteed that the same degree of correlation exists on the effective rainfall, which is an estimated process with a structure that depends in part by the configuration of the watershed response model. As a matter of fact, application of cluster-type models to the effective rainfall turned out to be unsatisfactory even on daily data. In particular, *Cowpertwait and O'Connell* [1992] attempted to directly estimate the four parameters of a Neyman-Scott model with exponential instantaneous pulses within a shot-noise model of daily streamflows and we applied the same model directly to our data obtained by inverse estimation. In both cases it was not possible to detect the "within storm" cellular structure of the effective rainfall, which represents the peculiar feature of this kind of models.

Given this outcome, it was judged that the two-parameter Poisson white noise model with exponential pulses (PWNE) represents a reasonable choice for scales of aggregation greater than one day, even because the effective rainfall process is only the input to a complex system that induces significant transformations on it, so that it is not guaranteed that disregarding some autocorrelation in the input will have a significant impact on the features of the generated streamflow series.

In conclusion, considering also that the method of selection of the aggregation scale proposed in this paper, discussed in the next section, does not necessarily require the analysis of data at the daily scale, it appears that the use of models more complex than the PWNE can only be proposed in a more refined framework, that allows efficient parameter estimation and justifies adequately the reduction in parsimony due to the greater number of parameters.

2.2. Components of the watershed response

Effective rainfall reaches the basin outlet through three main runoff components: *baseflow* (return flow from groundwater), *subsurface flow* (i.e. interflow, representing rapid flow through pipes, macropores and seepage zones in the soil) and saturated overland flow (*surface flow*). The two latter components form the *direct runoff* (or quickflow).

Separation of the effective rainfall into the above components is non-linear, because the relative weight between direct and groundwater runoff depends on the infiltration capacity, that is function of the soil moisture state and of the intensity of rainfall. This non linearity is not considered here, essentially because it is thought that for the aims of the analysis undertaken it is more important to test how well the proposed framework is able to reproduce main runoff process features at different frequencies.

Regarding the baseflow, *Claps et al.* [1993] showed that two different types of groundwater components can be clearly identified from analysis of runoff series. In particular, *Claps et al.* [1993] clarified that even in small basins annual data can be autocorrelated due to the presence of a groundwater contribution with very slow response time. In large basins, the hydrogeologic scheme could be more complex, but it is convenient, following the logic of conceptualisation, to assume in all cases that when annual time series are autocorrelated that is due to the presence of one large, deep, aquifer.

In monthly data autocorrelation is present, regardless of seasonality, even when annual runoff is uncorrelated. Thus, there exist an independent groundwater element which introduces memory in runoff data with a delay time of the order of few months. This over-month groundwater component is due to the presence of aquifers which run dry within the dry season, such as, for instance, overflow springs.

As regards direct runoff, if the basin is sufficiently large, even on daily data one can recognise the presence of both the subsurface and the surface runoff components.

Therefore, from this analysis it follows that one cannot exclude that all four of the above components can be identified from the streamflow dynamics on a short time scale. Consequently, the most general watershed scheme proposed includes four conceptual elements in parallel, which reduce to three if the over-year groundwater is absent.

The basin response function *h(t)* is assumed linear and its form derives from the linear combination of the individual responses of its components:

$$h(t) = u_0 + \frac{c_1}{k_1} e^{-t/k_1} + \frac{c_2}{k_2} e^{-t/k_2} + \frac{c_3}{k_3} e^{-t/k_3}, \qquad t \ge 0 \tag{1}$$

In (1) the impulse responses $u_i(t) = c_i / k_i e^{-t/k_i}$ of linear reservoirs represent the individual IUH of the groundwater and subsurface components, where the subscript *i* increases with the component lag time. So, k_3 is the storage coefficient of the *over-year* groundwater component.

Coefficients c_i, which are considered constant, distribute the input volume among system elements, according to the continuity condition $\sum_i c_i = 1$.

The response function of the surface component is indicated in (1) as $u_0(t)$ because, depending on the basin size and the time scale considered, it can be taken as a more or less complex function, becoming a Dirac delta function if basin surface response lag is enough smaller than the time interval of data aggregation.

Being the output of a linear system, runoff $X(t)$ is derived from the convolution integral

$$X(t) = \int_0^{+\infty} h(\tau)\, dI(t-\tau) \tag{2}$$

where $dI(\tau)$ is the input process. Considering the effective rainfall as a process of instantaneous pulses following a compound (or marked) Poisson process [*e.g. Snyder*, 1975] and assuming that pulses are uncorrelated and the distribution of their intensities is independent from the process of occurrences, the runoff process (2) represents a *filtered Poisson process*.

Convolution integral (2) gives the representation of a continuous shot noise process, while a discretized form is needed in the applications. The scheme adopted for discretizing the proposed model is shown in the Appendix.

3. Model Building

3.1. Identification of the basin response

Based on the above considerations, identification of model structure only consists in evaluating the presence of an over-year groundwater component. One may also want to assess if the daily resolution is sufficient to reproduce the surface runoff with some delay using, say, an exponential IUH.

The identification procedure should start with a check on the presence of autocorrelation in annual data, which is a statistical information to preserve. However, in the series we have analysed, autocorrelation has seldom a clear statistical significance, mainly due to the limited amount of data available. *Claps et al.* [1993] discussed this aspect with regard to identification of the stochastic model of annual flows, and suggested that the definitive test for the existence of the over-year groundwater component is represented by the comparison of estimates of parameters k_3 and c_3 with two hydrological indexes, called DFI and SFI, which have the same meaning.

The DFI (deep flow index) was defined as the ratio between the mean of annual minima of the mean monthly discharge and the average total discharge. The spring flow index (SFI) represents the ratio between the average of annual minima of the mean daily discharge and the average total discharge. Since the SFI is more sensitive to measurement errors, it is suggested to use it only as a check index.

If the above evaluations exclude the presence of an over-year component, the conceptual model of watershed will include only three linear elements, accounting for the over-month groundwater, the subsurface and the surface runoff.

In the presentation of the basin response function, the surface runoff component was left as a generic IUH. Considering scales of interest ranging from one to seven days, the form of the surface IUH can result significantly smoothed by aggregation, depending on its characteristic lag time which, in turn, depends on the basin size. Given the aims of this paper, which include the testing of some conceptual hypotheses, it was thought that identification of the surface response function at the maximum detail available would have reduced the attention on the role and the effects determined by the others components.

Therefore, the minimum scale of analysis was not fixed, but determined as T such that for the aggregation scale of T days the surface IUH is certainly reduced to a Dirac function. In other words, T is enough greater than the surface runoff lag. On the viewpoint of technical application, this way to proceed does not produce particular disadvantages, because the resolution needed to reproduce the runoff dynamics decreases with the basin size.

The value of the reference scale T could be set by using an empirical relation for the evaluation of the basin mean lag time as a function of area [*e.g. Murrone et al.*, 1992]. However, to make objective as much as possible the selection of T we made the following considerations: when T is small with respect to the surface time lag the surface hydrographs spread over a number of intervals, introducing autocorrelation which could be incorrectly attributed to the subsurface flow component effect. Conversely, when *T* is large, the variance of the component decreases, since high-frequency features of the flow are smoothed. Therefore, the reference scale must be the one that maximises the variance σ^2_s of the surface flow component, as evaluated once model parameters are estimated. Additional comments on this procedure are given in the application section.

3.2. Parameter Estimation

The procedure for parameters estimation of the shot noise models depends on the approach followed for identification and estimation of the model of effective rainfall. For the latter task there are essentially two alternatives.

In the first one, common to most of the shot noise models in literature [*Weiss,* 1973, 1977; *O'Connell,* 1977; *O'Connell and Jones,* 1979; *Cowpertwait and O'Connell,* 1992], the form of the underlying input process is pre-determined and its parameters are estimated through the method of moments applied to the streamflow statistics. This procedure does not give the possibility of verifying the hypothesis made on the input process and does not allow one to evaluate the influence of the effects of the watershed transformations on the estimation of the input parameters.

The alternative approach overcomes the above problems, since the input series is entirely reconstructed by inverse estimation. On this series, parameters of the desired stochastic model are then estimated. This procedure was followed, with different techniques, by *Treiber and Plate* [1977], *Hino and Hasebe* [1981, 1984], *Battaglia* [1986], *Wang and Vandewiele,* [1994], among others.

According to this latter procedure, which is the one adopted here, estimation of parameters of a shot noise model arise from the following steps: *(a)* identification of the pulse occurrences; *(b)* evaluation of the pulse intensities; *(c)* estimation of parameters of the system response function.

Since the correct inverse estimation of the input series requires the estimation of the vector $\boldsymbol{\theta}=\{c_0, c_1, k_1, c_2, k_2, c_3, k_3\}$ of parameters of the response function h', which, in turn, is conditioned by the determination of the pulse series, the estimation procedure must be iterative. The steps required in the estimation procedure are:

(***a***) determination of trial occurrences and values of pulses;

(***b***) estimation of θ by means of minimisation of the sum of quadratic errors:

$$\min_{\theta} \; SQ(\theta) = \min_{\theta} \; \sum_{t=1}^{N} \left\{ x_t - \sum_{s=1}^{t} Y'_{t-s+1} \, h'_s(\theta) \right\}^2 \tag{3}$$

where x_t is the data observed at time t, N is the number of observations and Y'_t is the cumulative input in the interval t (see Appendix);

(***c***) inverse estimation of a new pulse series through deconvolution.

After step *(c)* the sum of quadratic errors SQ(θ) resulting from the new pulse series is evaluated. The procedure is assumed to converge when the improvement obtained in SQ(θ) with the new Y'_t is lower than the 5% of the previous SQ(θ) value.

Some details regarding the actual application of the iterative Least Squares procedure are worth adding.

In the step *(a)*, a net rainfall occurrence is assumed in each time interval t presenting a discharge increase. To account for errors in the discharge measurements, a threshold value L is considered. When the condition $x_{t+1} \geq x_t + L$ is met, a trial value of the net rainfall amount is assumed as $Y'_t = x_{t+1} - x_t$, following *Battaglia* [1986]. The choice of the threshold level is critical, because as L increases the number of the detected events decreases, while its decrease originates many events with small intensities, which can be heavily affected by errors in streamflow data. These errors, which are found at high frequencies, produce a great number of pulses of very low value that alter the correlation structure of the reconstructed process. Then, the choice of L must be the result of a trade-off between the number of events detected and the significance of their values.

In step *(b)*, convergence of the procedure is ensured since the minimisation is applied to a positive-definite quantity. The minimisation problem is subject to some constraints, defined by the relations:

$$k_i > 0\ ;\quad c_i > 0\ ,\quad i = 1, 2, 3; \qquad \sum_{i=0}^{3} c_i = 1 \tag{4}$$

determined by the conceptual meaning of parameters. Step (*b*) is carried out through a Nelder-Mead simplex algorithm [*Press et al.*, 1988], which is convenient when the objective function is of a highly nonlinear type. The procedure starts from an initial condition $\boldsymbol{\theta}^o$ belonging to the feasible region of parameters. Constraints (4) are not directly implemented in the minimisation algorithm but are checked on each solution vector. It is also checked that in each optimisation phase the solution vector belongs to the feasible region of the problem.

In step (*c*) constraints are defined by the condition of non-negativity of pulses, i.e. $Y'_t \geq 0$. This constraint cannot be easily implemented in the deconvolution algorithm and is imposed on the series after deconvolution, by removing the negative values and their occurrences. In this way the estimated occurrences are implicitly tested. Actually, because of the influence of errors, a negative pulse either indicates a very low intensity summed to a great negative error, or a negative error alone. In both cases the occurrence of a rainfall event is highly unlikely.

3.3. Parameter Estimation on Different Time Scales

From a theoretical viewpoint the model proposed should be able to estimate all the response parameters at the daily time scale. However, this resulted practically impossible for the storage and the recharge coefficients of the over-year groundwater component. The main reason for this outcome is that the high time lag represented by k_3 makes this runoff component indistinguishable, at the daily (or *T*-day) scale, from an additive constant. To preserve the long-term correlation structure, a separate estimation of the parameters k_3 and c_3 is then needed.

The estimation scheme proposed here is the same used for the monthly scale by *Claps et al*. [1993], who suggested to aggregate the data on the annual scale in order to mask the undesired information (*i.e.* the seasonality) and maximise the visibility of the effect to capture (*i.e.* the interannual correlation). In the cited paper, for estimation of k_3 and c_3 the authors refer to the corresponding annual conceptually-based stochastic model, which is an ARMA(1,1).

As a support to this procedure, *Claps and Murrone* [1994] reported, in even simpler systems, positive effects of data aggregation with regard to parameter estimation of ARMA models. Aggregation was found effective when the response time of the conceptual element is much greater than the reference scale and when the weight of the component (in this case c_3) is relatively small.

The need to refer to a different class of stochastic models when data are aggregated requires to be further commented. In the framework of streamflow models with stochastic input and deterministic basin response, the increase in the aggregation scale modifies both the input process and the response function. As the aggregation interval becomes significantly greater than the time lag of a response component, the effects determined by this component are no longer apparent in the data. In the conceptual system, that 'fast' response is then associated with the component representing the zero-lag translation (diversion). This quite natural system modification does not change the meaning of parameters and would not require in itself a modification of the nature of the streamflow model. However, the structure of the input process also changes significantly, going from a periodic - approximately independent - point process on the daily scale, to a continuous quasi-gaussian process on the annual scale.

These drastic changes in the input process require equivalent modifications in the form of the streamflow model. The latter goes, according to the scheme initially proposed by *Claps et al*. [1993], from the ARMA(1,1) model on the annual scale, to the PIR-ARMA(2,2) (ARMA with periodic independent residual) on the monthly scale, up to the shot noise model presented here, for a scale of T days.

Once estimated k_3 and c_3 as described above, a constrained estimation is possible for the remaining parameters, starting on the monthly scale with parameters k_2 and c_2 [*Claps et al.*, 1993]. Finally, in the shot noise framework, parameters k_1 and c_1 can be estimated (with c_0 resulting by continuity).

The separate estimation of the parameter k_3 gives the necessary support for setting the memory horizon q of the shot noise model, which is needed to make finite the infinite memory mixed exponential response function. Since k_3 is the mean lag time of the slowest runoff component, q can be fixed as $q = n\ k_3$, with n such that q is the base time underlying 99% of the area of the exponential recession.

4. Multiple Shot Noise Model Application and Testing

The multiple shot noise model proposed was applied to 8 time series of daily flows, recorded in 7 watersheds located in the Apennine region of central southern Italy (Table 1). Two sub-series were considered for the Tiber River since the streamflow record is interrupted. For all of the rivers no significant diversions or regulations were reported during the observation period and the influence of the snowmelt runoff can be neglected. The basins under study are all characterised by the climate and the geology of Apennine mountains. For the whole region it is possible to distinguish two main climatic seasons: a wet season, during autumn and winter, and a dry season, during spring and summer. When analysing the streamflow process on the annual and monthly scale it was made reference to the hydrologic year, starting at the end of the dry season, conventionally at October the 1st.

4.1. Model Identification and Parameter Estimation

As discussed in the previous section, model identification essentially coincides with the evaluation of the presence of an over-year groundwater component. Existence of this component on the basins under study was preliminary checked using the DFI and SFI. Based on the values computed for the above indexes, reported in Table 2, in the seven basins only one "impermeable" watershed (group 1) was clearly identified, presenting very low runoff during the dry season.

Estimation of parameters k_3 and c_3 through the ARMA(1,1) model (Table 3) confirmed the indications emerged from the use of the hydrological indexes. The use of the Matlab™ package for ARMA parameter estimation made it possible to slightly revise the results of the basin classification (on the corresponding stations) obtained by *Claps et al.* [1993], by substantiating the discrimination suggested by the DFI.

TABLE 1. Characteristics of the basins and of the daily runoff series analysed

Series #	Gauging Station	Continuous Recording Period	Area (km^2)
1	ALENTO at Casalvelino	1.1.58 - 31.12.72	284
2	CALORE IRPINO at Montella	28.12.44 - 31.12.70	123
3	TAMMARO at Pago Veiano	1.1.58 - 31.12.70	555
4	SACCO at Ceccano	1.1.59 - 31.12.70	922
5	GIOVENCO at Pescina	1.1.60 - 31.12.70	139
6*a*	TIBER at Rome	8.7.20 - 31.5.42	16545
6b	"	1.1.47 - 31.5.69	"
7	NERA at Torre Orsina	1.1.47 - 31.5.69	1445

According to the conceptual framework for estimation of parameters on different scales of aggregation, coefficients c_2 and k_2 should be estimated on the monthly scale. It is worth specifying that for all series, except series #3, coefficients c_2 and k_2 are to be estimated through a PIR-ARMA(2,2) model, due to the presence of the over-year groundwater component, while the model for the monthly data of station # 3 is a PIR-ARMA(1,1).

TABLE 2. Watershed classification based on the flow indexes. Group1 includes impermeable basins. Group 2 includes basins with significant over-year groundwater runoff.

Group	Series #	Gauging Station	Spring Flow Index (%)	Deep Flow Index (%)
1	3	TAMMARO at Pago Veiano	0	2.09
2	1	ALENTO at Casalvelino	1.0	8.8
	2	CALORE IRPINO at Montella	0.4	19.7
	4	SACCO at Ceccano	2.3	9.7
	5	GIOVENCO at Pescina	34.4	55.2
	6*a*	TIBER at Rome	39.6	50.4
	6*b*	"	28.9	50.1
	7	NERA at Torre Orsina	40.1	74.3

However, application of the shot noise model showed that c_2 and k_2 can be efficiently estimated even on the T-day scale. This last result let us suppose that parameters c_3 and k_3 cannot be "seen" on the daily scale because the frequency of the response of that component is simply too low to be evident in the frequency spectrum computed with daily data. In particular, as discussed in the next section, estimates obtained with the shot noise proved to give a better picture of c_2 than the one obtained on monthly data. It was then decided for c_2 and k_2 to give credit to the shot noise estimates (Table 3).

TABLE 3. Final values of shot noise model parameters for analysed series and correspondent values of standard errors (in italic)

Series #	1	2	3	4	5	6*a*	6*b*	7
T (days)	3	2	3	3	2	4	4	5
c_0	0.340	0.155	0.272	0.319	0.106	0.098	0.101	0.025
c_1	0.281	0.194	0.236	0.242	0.112	0.158	0.136	0.018
St. Err.	*0.010*	*0.004*	*0.013*	*0.010*	*0.004*	*0.007*	*0.011*	*0.005*
k_1 *(days)*	3.105	2.672	2.912	2.706	3.274	4.076	5.552	5.401
St. Err.	*0.241*	*0.098*	*0.337*	*0.257*	*0.194*	*0.357*	*0.631*	*3.190*
c_2	0.297	0.546	0.492	0.261	0.191	0.225	0.243	0.247
St. Err.	*0.011*	*0.004*	*0.015*	*0.010*	*0.004*	*0.008*	*0.012*	*0.005*
k_2 *(days)*	60.4	72.2	35.6	53.9	56.3	43.7	40.2	109.3
St. Err.	*5.03*	*1.42*	*1.98*	*4.64*	*2.78*	*2.74*	*2.72*	*6.13*
c_3	0.082	0.106	-	0.178	0.590	0.520	0.520	0.710
k_3 *(days)*	551.4	228.5	-	507	1073	1233	1233	1533

Estimation of the subsurface component parameters and selection of the reference scale T are performed in the same time, since the final c_1 and k_1 are selected as the estimates made on the scale T that complies with the hypothesis $u_0(t) = \delta(0)$. Reference intervals found for all of the analysed series are reported in Table 3 together with the estimates of c_1 and k_1. The coefficient c_0 results from the volume continuity condition.

The procedure for selection of T through maximisation of the surface runoff variance to the total runoff variance ratio $\sigma^2{}_s/\sigma^2{}_t$ is exemplified for the station #3 in Table 4. Aggregation scales ranged from 1 to 7 days. Observation of these results, which are common to all other series, allows one to notice that the subsurface runoff storage coefficient k_1 is quite sensitive with respect to aggregation, while the corresponding recharge coefficient c_1, as well as the other parameters, resulted quite stable with the increase of T.

The variability of k_1 with aggregation was not unexpected, since the aggregation on multiple days needed for the correct estimation of surface component, produces also a loss of resolution on the subsurface component, having a delay only slightly greater than the one of the surface component. This is perhaps the reason why the maximum of the ratio $\sigma^2{}_s/\sigma^2{}_t$ is not very evident. On the other hand, it is worth remarking that the sensitivity shown by k_1 resulted not particularly important, as compared with the stability of the parameter c_1, in terms of the reproduction of the process features.

With respect to the selection of the reference scale, it is to mention that the presence of residual correlation in the effective rainfall series is of great support, because it indicates that some effects introduced by the lag of the surface runoff component are transferred to the estimated input. This means that the surface lag is not yet enough smaller than T .

TABLE 4. Estimation over different aggregation scales. Series # 3: Tammaro at Pago Veiano (the values correspondent to the reference scale T are bold)

aggregation (days)	c_0	c_1	k_1 (days)	c_2	k_2 (days)	$\sigma^2{}_s/\sigma^2{}_t$
1	0.2175	0.2501	1.4476	0.5324	32.9918	0.2771
2	0.2406	0.2868	2.1295	0.4726	33.8132	0.2460
3	**0.2721**	**0.2359**	**2.9121**	**0.4920**	**35.5948**	**0.2846**
4	0.2564	0.2713	3.3299	0.4723	31.9288	0.2252
5	0.2049	0.2914	1.9579	0.5037	30.6993	0.1276
6	0.2948	0.1506	2.3006	0.5546	34.7493	0.2725
7	0.2462	0.2210	4.7390	0.5328	32.1816	0.1820

4.2. Effective Rainfall Model Estimation

The Poisson white noise exponential model of precipitation (PWNE) was fitted on the input series obtained after estimation of the shot noise parameters. In the PWNE model the number N of rainfall events in the interval Δt follows a Poisson distribution, with parameter λ, and an exponential distribution, with mean $1/\beta$, is assumed for the intensity of the instantaneous pulse, The input intensity can also follow different continuous distributions, such as a gamma or a mixture of two negative exponential functions.

The probability density function of the cumulated effective rainfall Y over the duration Δt is given by [*e.g. Eagleson*, 1978]:

$$f_Y(y) \; = \; e^{-\lambda\,\Delta t}\left[\delta(y)+\sqrt{\frac{\lambda\,\Delta t\,\beta}{y}}\;\; I_1\!\left(2\sqrt{\lambda\,\Delta t\,\beta\,y}\right)\, e^{-\beta y}\right] \tag{5}$$

where $\delta(.)$ is the Dirac delta function and $I_1(.)$ is the first order modified Bessel function.

Moment estimates of parameters are given by

$$\hat{\lambda} = \frac{2\; m_Y^2}{s_Y^2\;\Delta t}; \qquad \hat{\beta} = \frac{2\; m_Y}{s_Y^2} \tag{6}$$

where m and s are the sample mean and standard deviation, while Maximum Likelihood (ML) estimates are given by the following relations [*Sirangelo and Versace*, 1990]

$$\frac{1}{Nj}\sum_{j=N_{j0}+1}^{N_j}\frac{I_0\!\left(2\hat{\beta}\sqrt{m_Y\,y_j}\right)}{I_1\!\left(2\hat{\beta}\sqrt{m_Y\,y_j}\right)}\sqrt{y_j} = \sqrt{m_Y} \tag{7}$$

$$\hat{\lambda} = \frac{m_Y\;\hat{\beta}}{\Delta t} \tag{8}$$

where variables y_j indicate the N_j historic values of Y, rearranged such that if N_{j0} is the number of zero values, the first N_{j0} values of y_j are equal to zero and $y_j > 0$ for $N_{j0}+1 \le \; j \le \; N_j$. $I_0(.)$ is the modified Bessel function of zero order.

Parameters of the PWNE model were estimated on seasons ranging from 14 to 33 days according to the scale of aggregation of the reconstructed input. Following the suggestions by *Yevjevich and Harmancioglu* [1989], the seasonal variability of each parameter p was described by a truncated Fourier series:

$$p(t) \; = \; A_0+\sum_{i=1}^{n}\left[A_i\;\; sin\!\left(\frac{2\pi i}{T}\,t+\Phi_i\right)\right] \tag{9}$$

in which coefficients A_i and Φ_i respectively represent the amplitudes and the phases of the series, and A_0 is the mean value of the parameter. In the case under study, based also on the results of the application made by *Sirangelo* [1994] on central-southern Italy series, two harmonics were considered sufficient.

Estimation of coefficients A_i and Φ_i, as well as A_0, were performed both with the moments and with the maximum likelihood methods, and are reported in Table 7.

5. Model Testing

5.1. Hydrological Validation of Parameters Estimates

In the procedure suggested by *Claps et al*. (1993) validation of parameters of the over-year groundwater was considered as part of the identification stage and was based on the indirect evaluation of the coefficient c_3 through the DFI and the SFI. In the previous section, effectiveness of the above hydrological indexes in the validation of the conceptual parameter c_3 was shown once again.

As regards parameters of the over-month groundwater component, given that the estimates of c_2 and k_2 obtained on the T-day scale with the shot noise model look quite efficient, it was natural to compare them with the estimates obtained through the PIR-ARMA model applied to monthly data. For this comparison, the interval T was set at the reference value, even though, as can be seen in Table 4, the parameters under investigation are almost insensitive to aggregation from 1 to 7 days.

Table 5 reports the estimates of parameters c_2 and k_2 (and c_0) obtained with both models, along with the shot noise estimates of parameters c_1 and k_1 of the subsurface runoff component, reported again for reference. The results obtained require some comments.

TABLE 5. Comparison between PIR-ARMA and shot noise model parameters

Series #	Model	c_0	c_1	k_1 *(days)*	c_2	k_2 *(days)*
1	SN	0.340	0.281	3.11	0.297	60.4
	PIR-ARMA	0.186			0.732	58.5
2	SN	0.155	0.194	2.67	0.546	72.2
	PIR-ARMA	0.155			0.740	60.7
3	SN	0.272	0.236	2.91	0.492	35.6
	PIR-ARMA	0.210			0.790	55.2
4	SN	0.319	0.242	2.71	0.261	53.9
	PIR-ARMA	0.145			0.677	90.1
5	SN	0.106	0.112	3.27	0.191	56.3
	PIR-ARMA	0.133			0.277	66.9
6*a*	SN	0.097	0.158	4.08	0.225	43.7
	PIR-ARMA	0.117			0.363	49.8
6*b*	SN	0.101	0.136	5.55	0.243	40.2
	PIR-ARMA	0.117			0.363	49.8
7	SN	0.025	0.018	5.40	0.247	109.3
	PIR-ARMA	0.034			0.256	106.8

It can be first observed that, apart form station # 4, values obtained for k_2 from both models are substantially coinciding, particularly if considering the acceptable region of parameters shown in Table 6 and discussed in the next subsection. On the other hand, values obtained for c_2 with the PIR-ARMA models are systematically greater than those obtained from the shot noise model. On this side, *Claps et al.* [1993] attributed their relatively high values to the interference of the subsurface component on the over-month component. Results presented here confirms that conjecture, because the sum of c_1 and c_2 obtained through the shot noise model practically matches the values of c_2 estimated on the monthly scale. It is also worth remarking that c_0 was estimated quite well already on the monthly scale.

5.2. Statistical Testing of Parameters Estimates

Verification of the results of model estimation were made under different viewpoints. First, some tests were performed to ascertain the reliability of parameter estimates. The testing concerns the possible presence of multiple local minima of the objective function in the feasible region of parameters, which would not allow the numerical algorithm to find the absolute minimum. This aspect was verified using a procedure proposed by *Duan et al*. [1992], who suggest to search minima of the objective function, like the function defined by relation (3), starting from different points of the feasible region of the parameters. Application of this procedure on all of the series analysed confirmed the stability of the parameter estimates.

A further verification of this aspect was performed by exploring the objective function surface on the subspaces determined by couples of parameters. Patterns of this surface are shown in Fig. 1, with reference to four couples of parameters. In all of the plots the presence of only one region of attraction for the minimum can be recognised, and the region is also quite well defined.

Evaluation of efficiency of estimates can also be made by determining their variance (or standard errors). This determination can be achieved through the knowledge of the covariance matrix of estimates, that allows one to obtain the standard error (SE) of estimates using an asymptotical result given by *Bard* [1974]. Values of the standard errors computed for all of the parameter estimates are given in Table 3 for all series. Standard errors of k_3 and c_3 were not reported since from the ARMA estimation result only the SE of the AR and MA parameters.

An additional test was performed regarding the robustness of the numerical scheme or, in other words, to verify the capability of the estimation technique to attain the “true” parameter values regardless of possible distortions due to small fluctuations in the data. To this end, a non-standard jackknife resampling procedure, proposed by *Künsch* [1989] for ergodic stationary time series, was used. The mean and the standard deviation of the estimated parameters were computed, for each series, resampling 200 sub-series of 1825 observations, corresponding to a shifting window of 5 years [*Murrone*, 1994]. To encompass the whole series the shift was adapted to the series length.

For all of the series, parameter estimates fell inside the acceptable region, defined as having the centre on the jackknife mean and with half-band width equal to twice the jackknife standard deviation. In almost all cases estimates fell even inside the one-standard-deviation interval (Table 6).

5.3. Generation of Synthetic Streamflow Series

To verify the efficiency of the model in terms of reproduction of the statistical characteristics of observed series, extensive simulation studies are required. In general one is interested to test if a number of significant characteristics of the observed series are correctly reproduced.

According to the aims of runoff modelling, flow sequences are mainly needed in water resources planning and management, so one is mainly interested in: partition of flows among different components, cumulative frequency function of total flow, minimum and maximum daily flows, runoff volumes over different duration, flow duration curves, run length statistics. Among the many possible statistical characteristics of the observed series, the ones tested were the first three moments, the flow duration curves and the maxima and minima of flows, useful to check the model behaviour in the tails of the distribution.

TABLE 6. Jackknife robust estimates of parameters (μ) and their standard deviations (σ) compared to final values of model parameters (in italic)

Series #		c_0	c_1	k_1 *(days)*	c_2	k_2 *(days)*
		0.340	*0.281*	*3.11*	*0.297*	*60.4*
1	μ	0.282	0.315	2.92	0.321	47.6
	σ	0.074	0.076	1.15	0.046	11.5
		0.155	*0.194*	*2.67*	*0.546*	*72.2*
2	μ	0.148	0.219	3.46	0.527	86.2
	σ	0.038	0.061	1.99	0.064	40.2
		0.272	*0.236*	*2.91*	*0.492*	*35.6*
3	μ	0.267	0.289	3.29	0.435	45.8
	σ	0.047	0.068	0.88	0.089	12.9
		0.319	*0.242*	*2.71*	*0.261*	*53.9*
4	μ	0.296	0.281	2.51	0.245	51.7
	σ	0.041	0.070	0.59	0.050	13.4
		0.106	*0.112*	*3.27*	*0.191*	*56.3*
5	μ	0.105	0.133	3.08	0.172	51.8
	σ	0.012	0.015	0.50	0.013	12.1
		0.097	*0.158*	*4.08*	*0.225*	*43.7*
6*a*	μ	0.092	0.180	4.17	0.208	48.5
	σ	0.040	0.025	1.61	0.034	13.9
		0.101	*0.136*	*5.55*	*0.243*	*40.2*
6*b*	μ	0.072	0.142	3.80	0.266	38.7
	σ	0.038	0.033	1.86	0.030	12.4
		0.025	*0.018*	*5.40*	*0.247*	*109.3*
7	μ	0.011	0.029	4.15	0.250	89.4
	σ	0.015	0.023	5.85	0.029	17.9

These comparisons were performed for the station # 1 using the PWNE model for the inputs, with ML estimates of parameters. The PWNE model was fitted on the input series inversely estimated on 3-day data, with model parameters set at the values reported in Table 3. Data generation was made on a daily scale using the seasonal approximation of input parameters given in Table 7. For each calendar day t, parameter values were obtained through relation (9). The inputs were then convoluted with the system response function, as in (15) to obtain the streamflow series.

Daily flows were generated with length equal to 20 times the observation period. To make the generated series independent on the initial conditions, 20 years of data were first generated and not considered in the subsequent analyses.

To give a general idea of the quality of the simulated runoff, a comparison of runoff patterns of observed and generated daily data are shown in Fig. 2 for two solar years with reference to station # 1. Table 8 shows the comparison of moments of the observed and generated flows, averaged on the different months. It can be seen that the first three moments are quite well preserved in simulation, both for the general and for the monthly values. Positive bias of the generated skewness is to be noticed, even though it resulted compatible with the sample variability, measured by the $\pm 2\sigma$ bands, reproduced in Fig. 3 for all three moments.

Reproduction of the flow duration curves and of maxima and minima of flow over different duration are shown in Fig. 4 and Fig. 5, respectively. Good performances of the model can be recognised with respect to the reproduction of these characteristics. It is to notice that preservation of the effects of the groundwater runoff component produces remarkably good results in the simulation of runoff during long dry periods, which is one of the points where previous shot noise models gave inadequate results [e.g. *Weiss*, 1977; *Battaglia*, 1986].

Maximum of flow over fixed duration were also very well reproduced. Comparisons are referred to the average yearly maximum (and minimum) flow of duration 1-30 days.

Observation of the series of the yearly maxima of daily flow in Gumbel probability paper (Fig. 6), also shows the good performance and adequacy of the PWNE model for the inputs with regard to the upper extremes.

TABLE 7. Fourier expansion of PWNE model parameters

Series #	Fourier Coefficients	Moment method λ (1/*days*)	Moment method 1/β (*mm*)	Maximum Likelihood method λ (1/*days*)	Maximum Likelihood method 1/β (*mm*)
1	A_0	0.0627	18.4652	0.1222	9.8799
	A_1	0.0366	19.3238	0.0568	11.3508
	Φ_1	0.5194	0.6652	1.0510	0.5246
	A_2	0.0098	-3.6871	0.0146	-2.9899
	Φ_2	0.6329	-2.8692	2.0364	-3.1582
2	A_0	0.0780	17.0055	0.1541	9.0492
	A_1	0.0482	12.2677	0.0634	9.1986
	Φ_1	0.7778	1.0153	1.5916	0.7400
	A_2	-0.0060	-2.9475	0.0237	-2.5246
	Φ_2	4.0475	-2.2568	3.5817	3.7131
3	A_0	0.0870	8.9844	0.1135	7.0678
	A_1	0.0781	7.6915	0.0739	7.2112
	Φ_1	0.4628	1.0713	1.1218	0.5598
	A_2	0.0113	-1.9819	0.0213	-1.6621
	Φ_2	6.1570	-1.0450	2.8850	-3.1367
4	A_0	0.0713	15.9804	0.1194	9.2217
	A_1	-0.0410	15.0478	0.0465	11.1499
	Φ_1	3.5851	1.4098	1.2894	1.0611
	A_2	0.0237	-3.3802	0.0359	-3.5134
	Φ_2	1.9072	-1.9942	2.8818	-2.5602
5	A_0	0.1573	5.4907	0.2321	2.9636
	A_1	-0.0444	4.8171	0.0668	3.3123
	Φ_1	1.9184	1.2622	2.1649	0.8149
	A_2	-0.0423	-1.4042	0.0282	-0.9944
	Φ_2	4.0489	1.7715	3.8458	-2.9002
6*a*	A_0	0.0893	14.2871	0.1118	11.5891
	A_1	-0.0341	-8.5488	-0.0328	-9.3243
	Φ_1	0.5707	0.9060	1.7152	0.5343
	A_2	0.0169	4.0876	0.0225	1.7111
	Φ_2	3.5682	1.7722	3.4993	2.2347
6*b*	A_0	0.0929	12.1795	0.1307	8.5534
	A_1	0.0282	7.3971	0.0368	8.0486
	Φ_1	0.8619	1.0213	1.9636	0.6803
	A_2	-0.0130	-1.9725	0.0160	-1.6221
	Φ_2	4.6728	1.9306	3.3149	-3.1410
7	A_0	0.0696	24.4386	0.0865	16.5933
	A_1	-0.0131	19.1208	0.0191	11.2012
	Φ_1	1.2360	0.2672	0.2020	-0.1480
	A_2	0.0134	-9.8283	0.0061	-1.9034
	Φ_2	4.0353	3.6219	0.6079	3.7185

TABLE 8. Statistics of historical and generated daily flows for each calendar month (averaged over the years) and for the whole period. Series # 1

Month	Mean (cms)		Standard deviation (cms)		Skewness	
	Hist.	Gen.	Hist.	Gen.	Hist.	Gen.
Jan	11.9935	10.9199	16.4726	15.0339	2.2738	3.0530
Feb	9.2241	9.9710	9.8953	12.6215	2.2370	3.0152
Mar	7.3553	7.4254	7.2697	8.9105	2.4190	3.2204
Apr	5.2805	5.0500	5.5319	4.7972	2.4582	2.9815
May	2.6469	3.1805	2.0212	2.5145	1.6508	3.1770
Jun	1.5716	2.0350	1.1604	1.2588	1.2716	2.8848
Jul	1.0335	1.3792	0.3167	0.6837	1.2813	2.8224
Aug	0.6509	0.9004	0.2051	0.2093	1.1545	1.9361
Sep	0.7878	0.7798	0.9144	0.3904	2.1076	3.1654
Oct	1.7465	1.9661	3.0342	2.9991	2.0680	3.3723
Nov	5.2338	5.5723	9.0731	8.4227	3.0213	2.9807
Dec	10.2225	9.7141	16.6532	13.7824	2.8460	2.8551
whole period	4.7905	4.8832	12.1982	10.1628	7.7051	7.0382

6. Concluding remarks

The building of a shot noise model for streamflows at short scales of aggregation is presented here according to a conceptually-based interpretation of the basin response to precipitation, in which linearity is assumed for the effective rainfall to runoff transformation.

Our basic aim was to build a model able to preserve the dynamics of the streamflow process over different scales of aggregation. The respect of this requirement allowed the model to comply with the structure of corresponding conceptually-based models proposed for the monthly and annual scales [*Claps et al.*, 1993]. Moreover, parameter estimation is part of this compatibility, since coefficients responsible of the long-term autocorrelation in runoff series are estimated, through the ARMA(1,1) model, on the annual scale. Preservation of the effects of the groundwater runoff component with over-year lag determines a marked improvement in the quality of the runoff series reproduced by the model, even because it allows to better discriminate the effects due to groundwater components and direct runoff.

To simplify the watershed response function, a reference scale of *T* days was indicated for parameter estimation, such that the response of the surface runoff component could be assumed as equivalent to a Dirac delta function. For the selection of this scale a procedure based on the maximisation of the surface runoff variance was proposed, and its application is simultaneous with the parameter estimation.

Model application to eight time series of central-southern Italy basins showed interesting results, supported by hydrological validations and statistical tests involving also a simulation analysis. This latter was performed using a Poisson white noise model with exponential pulses (PWNE) for reproduction of the effective rainfall. PWNE model parameters were estimated on the inverse estimate of the effective rainfall series. Comparison of observed and generated statistics and properties of the runoff series showed satisfying agreement between the two, with excellent reproduction of maxima of runoff for different duration.

Improvements in the model structure can be achieved, in the authors' opinion, essentially from refinements in the effective rainfall inverse estimation and in the related stochastic model, on which additional investigations are being carried out by the authors. Given the good overall performances of the model, it is thought that refining the modelling of effective rainfall is more important than introducing non linearity in the separation of effective rainfall in the groundwater and subsurface subsystems.

In addition, further investigations about the modelling of the surface runoff response are needed, starting from refinement of the method for selection of the reference scale when dealing with daily data. Finally, by the viewpoint of the conceptual interpretation of watershed response, hourly runoff data could allow us to improve the distinction of the subsurface and surface response, giving, in the meantime, the possibility to estimate the latter as an exponential or gamma function.

ACKNOWLEDGEMENTS

This work was supported in part by funds granted by the Italian National Research Council - Group for Prevention of Hydrogeological Disasters, grant 93.01431.42.

REFERENCES

Bard, Y., 1974; *Nonlinear Parameter Estimation*. Academic Press, San Diego, California.

Battaglia, F., 1986; Modelli stocastici per la rappresentazione e la previsione dei deflussi. *Dept. of Statistics Papers*, No. **3**, Univ. di Roma "La Sapienza", Dip. di Statistica, Probabilità e Statistica Applicata, Rome.

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

Bo, Z., S. Islam, and E. A. B. Elthair, 1994; Aggregation-disaggregation properties of a stochastic rainfall model. *Water Resour. Res.*, **30**(12), 3423-3435.

Box, G. E., and G. Jenkins, 1976; *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco, California, Revised Edition.

Claps, P., 1990; Modelli stocastici dei deflussi dei corsi d'acqua, Ph.D. dissertation, 279 pp., Università degli Studi di Napoli, Naples, Italy.

Claps, P., F. Rossi, and C. Vitale, 1993; Conceptual-stochastic modeling of seasonal runoff using Autoregressive Moving Average Models and different scales of aggregation. *Water Resour. Res.*, **29**(8), 2545-2559.

Claps, P. and F. Murrone, 1994; Optimal parameter estimation of conceptually-based streamflow models by time series aggregation, in K.W. Hipel *et al.* (eds.) *Stochastic and statistical methods in hydrology and environmental engineering*, vol.3, 421-434, Kluwer Acad., Dordrecht, The Netherlands.

Cowpertwait, P. S. P., and P. E. O'Connell, 1992; A Neyman-Scott shot noise model for the generation of daily streamflow time series, in *Advances in Theoretical Hydrology*, edited by J. P. O'Kane, Part A, Chap. 6. Elsevier, The Netherlands.

Duan, Q., S. Sorooshian, and V. K. Gupta, 1992; Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.*, **28**(4), 1015-1031.

Eagleson, P. S., 1978; Climate, soil, and vegetation: 2. The distribution of annual precipitation derived from observed storm sequences. *Water Resour. Res.*, **14**(5), 713-721.

Hino, M., and M. Hasebe, 1981; Analysis of hydrologic characteristics from runoff data - A hydrologic inverse problem. *Jour. of Hydrology*, **49**(3/4), 287-313.

Hino, M., and M. Hasebe, 1984; Identification and prediction of nonlinear hydrologic systems by the filter-separation autoregressive (AR) method: extension to hourly hydrologic data. *Jour. of Hydrology*, **68**: 181-210.

Kavvas, M. L., and J. W. Delleur, 1981; A stochastic cluster model of daily rainfall sequences. *Water Resour. Res.*, **17**(4), 1151-1160.

Koch, R. W., 1985; A stochastic streamflow model based on Physical principles. *Water Resour. Res.*, **10**(4), 737-744.

Kron, W., E. J. Plate, and J. Jhringer, 1990; A model for the generation of simultaneous daily discharges of two rivers at their point of confluence. *Stochastic Hydrology and Hydraulics*, (**4**), 255-276.

Künsch, H. R., 1989; The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, Vol. **17**, No.3, 1217-1241.

Murrone, F., 1994; Modelli stocastici puntuali dei deflussi fluviali, Ph.D. dissertation, 208 pp., Università degli Studi di Napoli "Federico II", Naples, Italy.

Murrone, F., F. Rossi, and P. Claps, 1992; A conceptually-based multiple shot noise model for daily streamflows, in *Stochastic Hydraulics '92*, edited by J.-T. Kuo and G.-F. Lin, pp. 857-864, NTU, Taipei, Taiwan R.O.C.

O'Connell, P. E., 1977; Shot noise models in synthetic hydrology, in *Mathematical Models for Surface Water Hydrology*, edited by T. A. Ciriani, U. Maione and J. R. Wallis, pp. 19-26, John Wiley, New York.

O'Connell, P. E., and D. A. Jones, 1979; Some experience with the development of models for the stochastic simulation of daily flows, in *Inputs for Risk Analysis in Water Resources*, edited by E. A. McBean et al., pp. 287-314, Water Resour. Publ., Fort Collins, CO.

Parzen, E., 1962; *Stochastic Processes*. Holden Day Inc., San Francisco, CA.

Pegram, G. G. S., 1980; A continuous streamflow model. *Jour. of Hydrol.*, **47,** 65-89.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1988; *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge Univ. Press, Cambridge, UK.

Rodriguez-Iturbe, I., V. K. Gupta, and E. Waymire, 1984; Scale considerations in the modelling of temporal rainfall.*Water Resour. Res.*, **20**(11), 1611-1619.

Rodriguez-Iturbe, I., D. L. Cox, and V. Isham, 1987; Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond.* A **410**, 269-288.

Sirangelo, B., 1994; Un modello poissoniano non omogeneo per l'analisi delle piogge cumulate su lunghi intervalli temporali. *Idrotecnica*, 3, 125-158.

Sirangelo, B., and P. Versace, 1990; Stima dei parametri ed intervalli di aggregazione nei modelli puntuali di precipitazione. *Proc. of the XXII Convegno di Idraulica e Costruzioni Idrauliche*, Cosenza, Italy, Vol. 5, 373-405.

Snyder, D. L., 1975; *Random Point Processes*. J. Wiley & Sons Inc., New York, NY.

Todorovic, P., and V. Yevjevich, 1969; Stochastic process of precipitation. *Hydrology Papers*, 35, Colorado State Univ., Fort Collins, Colorado.

Treiber, B., and E. J. Plate, 1977; A stochastic model for the simulation of daily flows. *Hydrological Sciences Bulletin*, XXII, 1 **3**/1977, 175-192.

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

Wang, Z.M., and G. L. Vandewiele, 1994; Stochastic simulation of streamflow with short time interval. in *Coping with Floods*, edited by G. Rossi, N. B. Harmancioglu and V. Yevjevich, pp. 257-269, NATO ASI Series E, Vol. 257. Kluwer, Dordrecht, The Netherlands.

Weiss, G., 1973; Shot noise models for synthetic generation of multisite daily streamflow data. *Proc. of the Symposium on the Design of Water Resources Projects with Inadequate Data*, Unesco, WMO, IAHS, Madrid, Vol. 2.

Weiss, G., 1977; Shot noise models for the generation of synthetic streamflow data. *Water Resour. Res.*, **13**(1), 101- 108.

Yakowitz, S. J., 1979; A nonparametric Markov model for daily river flow. *Water Resour. Res.*, 15(5), 1035-1043.

Yevjevich, V., and N. B. Harmancioglu, 1989; Description of periodic variation in parameters of hydrologic time series. *Water Resour. Res.*, 25(3), 421-428.

Appendix. Discretized representation of the shot noise model

In a linear system fed by a point process, if the input pulse at the instant τ_i has intensity Y_i, the system output, for $t > \tau_i$, is equal to $Y_i\, h(t-\tau_i)$. Given the linearity of the transformations operated by the watershed, the discharge at time t will be the sum of the outputs to the previous individual inputs [*Parzen*, 1962], whose number is given by the counting function of the arrival process, $N(t)$:

$$X(t) = \sum_{N(-\infty)}^{N(t)} Y_i\, h(t-\tau_i) \tag{10}$$

Relation (10) has theoretically infinite memory. The problem to overcome with regard to this formulation concerns the correct classification of the initial condition (t=0) and the possibility of actually reproducing the numerical scheme. To avoid the carry-over of "infinite" events influencing the output at any time, it is sufficient to limit the "active" history of events to a finite IUH base time q (the memory horizon, underlying 99% of the IUH area). Even then, the setting of the initial conditions with the usual method of the "warm up length" (e.g. removing the first n data influenced by the initial conditions) can be impracticable if the IUH base time is particularly large. Reasonable initial conditions can be set if long dry periods exist, since at their end the over-month and the subsurface components can be neglected compared to the over-year runoff component. In this case, the model expression becomes

$$X(t) = X_0\, e^{-t/k_3} + \sum_{N(0)}^{N(t)} Y_i\, h(t-\tau_i) \tag{11}$$

where X_0 is $X(t\text{=}0)$ and is entirely represented by the over-year component.

Process (11) is the representation of the continuous streamflow process, while time series are usually available in discretized form, such as the mean daily discharge analysed here. To represent process (11) integrated on a time interval Δt, with Δt equal to T days (T = 1, 2, ...), a time distribution of the rainfall intensity within the aggregation interval Δt must be assumed.

The rainfall volume Y'_t in each interval is computed as the integration of the instantaneous pulses Y_i occurred within Δt. The discretized output X_t will be then obtained through convolution between the input volume and the discretized response function $h_{\Delta t}$.

Considering a time scale Δt greater than the lag of the surface flow component, this latter can be considered as a random process with intensity proportional to the effective rainfall at any time. In this hypothesis, the surface runoff response $u_0(t)$ can be represented by a pure translation element, with zero lag. Its IUH is then represented by a Dirac delta function, $\delta(0)$.

The system unit hydrograph $h_{\Delta t}(\tau)$ is the response to a unit volume of effective rainfall of duration Δt and intensity $1/\Delta t$. From the expression (1) of the system IUH, where $u_0(t)$ is replaced by $\delta(0)$, if the within-interval time distribution of the input is assumed as uniform, the unit hydrograph is expressed as:

$$h_{\Delta t}(t) = \int_0^t \frac{1}{\Delta t} u(t-\tau) d\tau = \frac{1}{\Delta t}\left[c_0 \delta(0) + \int_0^t \sum_{i=1}^{3} \frac{c_i}{k_i} e^{-(t-\tau)/k_i} \, d\tau \right] =$$

$$= \frac{1}{\Delta t}\left[c_0 \delta(0) + \sum_{i=0}^{3} c_i (1 - e^{-t/k_i}) \right], \qquad t \le \Delta t \qquad (12)$$

The discharge is maximum when $t = \Delta t$ and then decreases following an exponential recession:

$$h_{\Delta t}(t) = \frac{1}{\Delta t} \sum_{i=1}^{3} c_i (1 - e^{-\Delta t/k_i}) e^{-(t-\Delta t)/k_i}, \qquad t > \Delta t \qquad (13)$$

Runoff at the end of the intervals $1\Delta t$, $2\Delta t$, ..., $s\Delta t$, is the integral of the two above expressions. From the former, the outflow volume h'_1 at the end of the first interval is obtained:

$$\begin{aligned} h'_1 &= c_0 \delta(0) + \int_0^{\Delta t} \frac{1}{\Delta t} \sum_{i=1}^{3} c_i \left(1 - e^{-\tau/k_i}\right) d\tau = \\ &= c_0 \delta(0) + \frac{1}{\Delta t} \sum_{i=1}^{3} c_i \left[k_i \left(e^{-\Delta t/k_i} - 1 \right) + \Delta t \right] \end{aligned} \qquad s = 1 \qquad (14)$$

The latter expression gives the volumes h'_s in the subsequent intervals due to the rainfall in the first interval:

$$h'_s = \int_{(s-1)\Delta t}^{s\Delta t} \frac{1}{\Delta t} \sum_{i=1}^{3} c_i \left(1 - e^{-\Delta t/k_i}\right) e^{-(\tau-\Delta t)/k_i} \, d\tau =$$

$$= \sum_{i=1}^{3} c_i \frac{k_i}{\Delta t} \left[e^{\Delta t/k_i} + e^{-\Delta t/k_i} - 2 \right] e^{-\Delta t\,(s-1)/k_i} , \qquad s > 1 \tag{15}$$

By aggregation of the input pulses Y_i one obtains a new intermittent process, $\{Y't\}$, defined as the total volume of pulses occurring in the generic interval $[(t-1)\Delta t,\ t\Delta t]$, denoted as interval t:

$$Y'_t = \sum_{N((t-1)\Delta t)}^{N(t\,\Delta t)} Y_i \tag{16}$$

Streamflow in the interval t due to an input occurred s intervals before is given by the product $Y'_{t-\ s+1}h'_s$. Then, the discretized form of the model (11) is given by

$$X_t = \int_{(t-1)\Delta t}^{t\Delta t} X(\tau)\, d\tau = k_3\, e^{-t\,\Delta t\,/\,k_3} \left(e^{\Delta t\,/k_3} - 1 \right) X_0 + \sum_{s=1}^{t} Y'_{t-s+1}\, h'_s \tag{17}$$

If the process $\{Y'_t\}$ and the response function h' are mutually independent and $\{Y'_t\}$ is serially uncorrelated, moments of the process (17) can be expressed as [*Murrone* 1994] :

$$E(X_t) = p\, E(Y') \sum_{s=1}^{t} h'_s ;$$

$$VAR(X_t) = \{p\, E(Y'^2) - p^2\, E^2(Y')\} \sum_{s=1}^{t} h'^2_s \tag{18}$$

$$COV(X_t,\, X_{t+k}) = \{p\, E(Y'^2) - p^2\, E^2(Y')\} \sum_{s=1}^{t} h'_s\, h'_{s+k}$$

where p represents the marginal probability $P(Y' > 0)$.

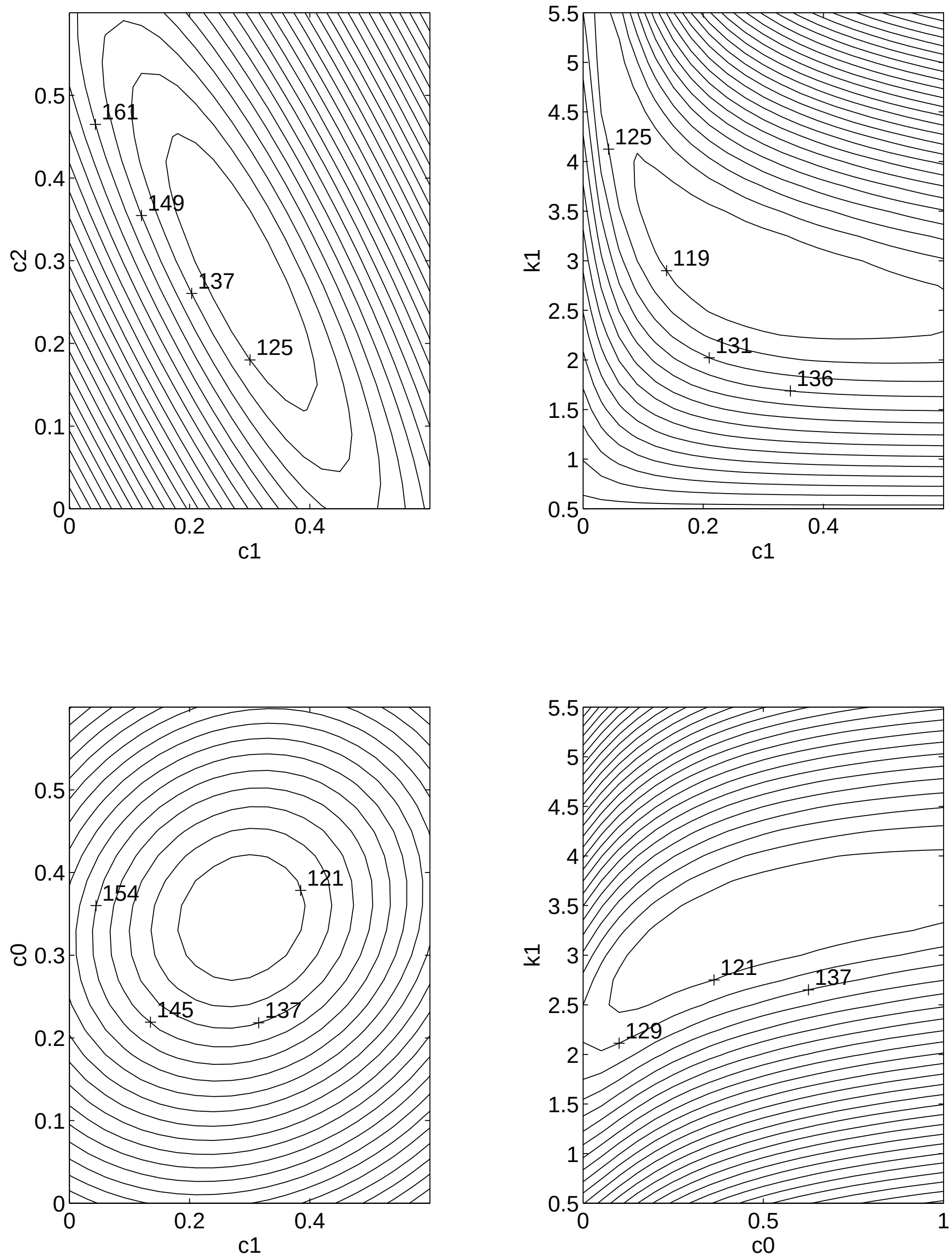

Figure 1. Series # 1: contour plot of the objective function of the model for four couples of parameters

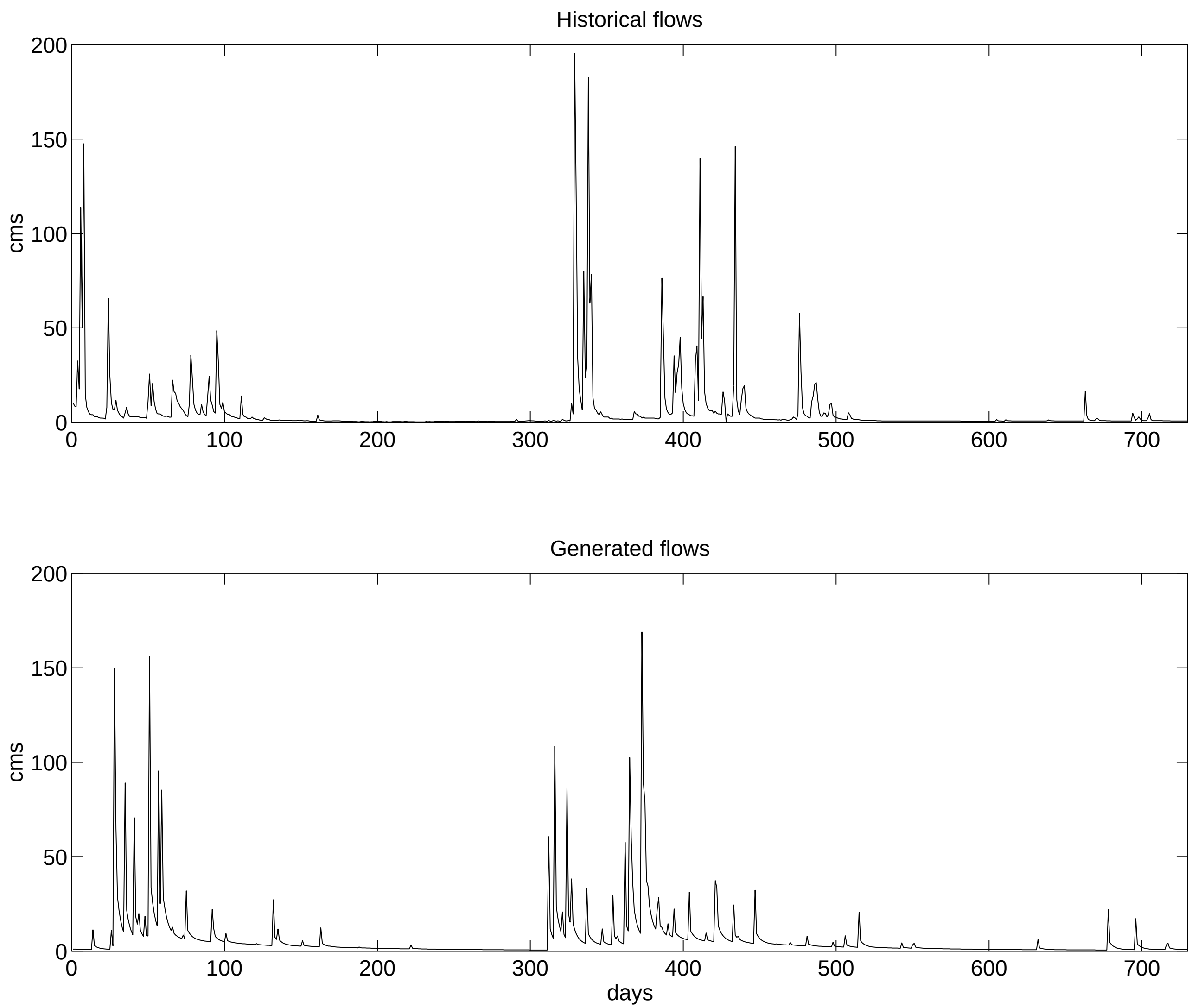

Figure 2. Series # 1: two years of historical and generated daily flows (Jan. 1 ÷ Dec. 31)

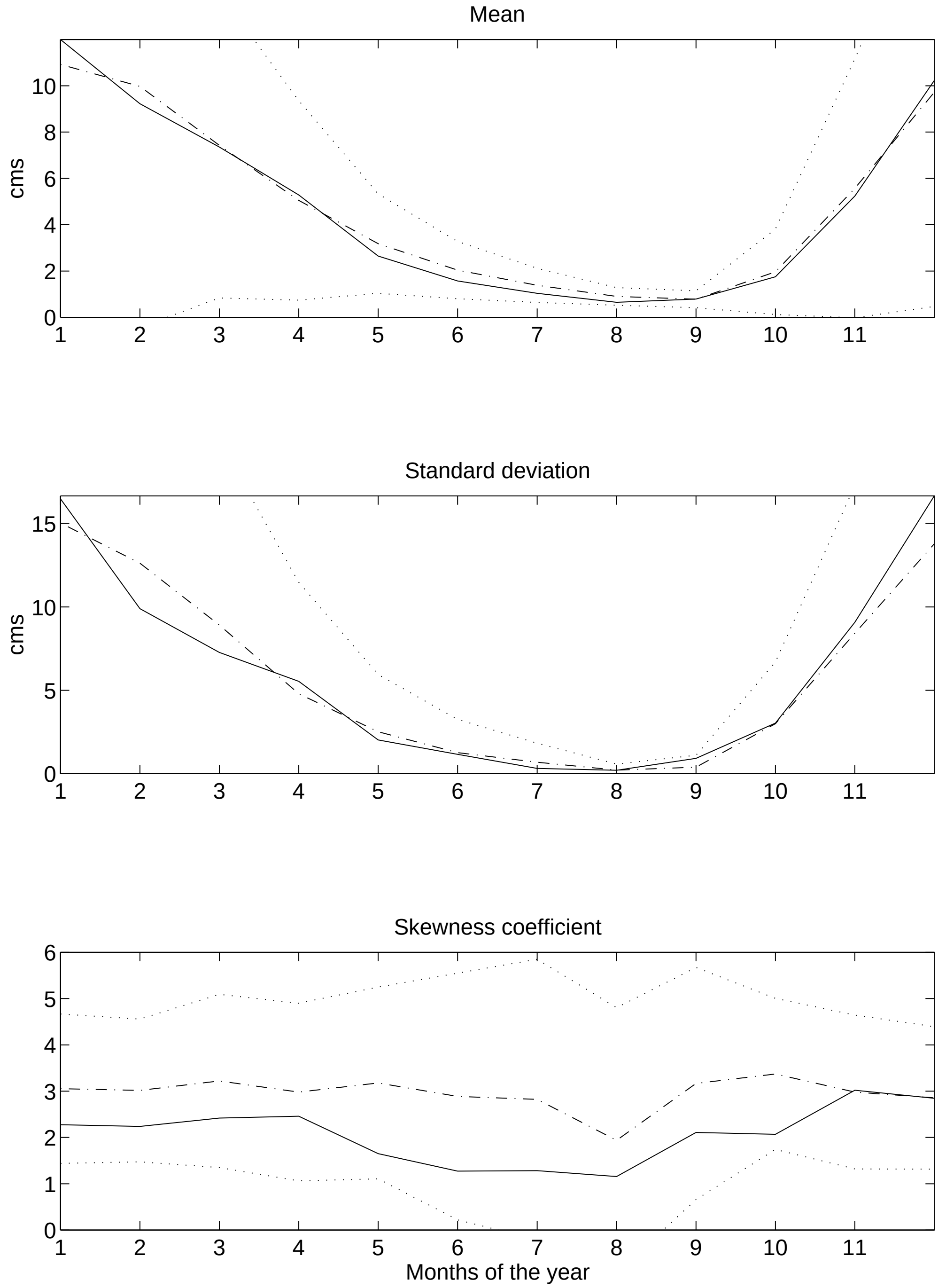

Figure 3. Series # 1: comparison of mean statistics of daily flows for each calendar month;

(—) historical flows;(-·-·) generated flows; (·····) 2-std. bands

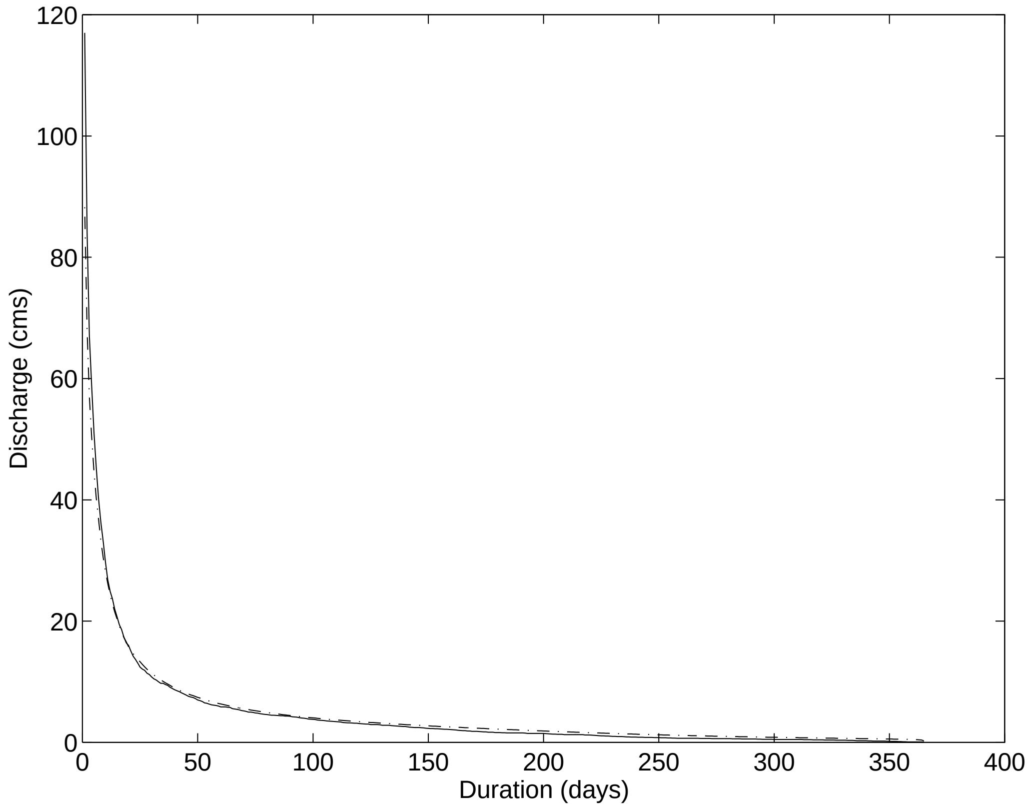

Figure 4. Series # 1: Comparison of flow duration curves:

(—) historical flows, (-·-·) generated flows

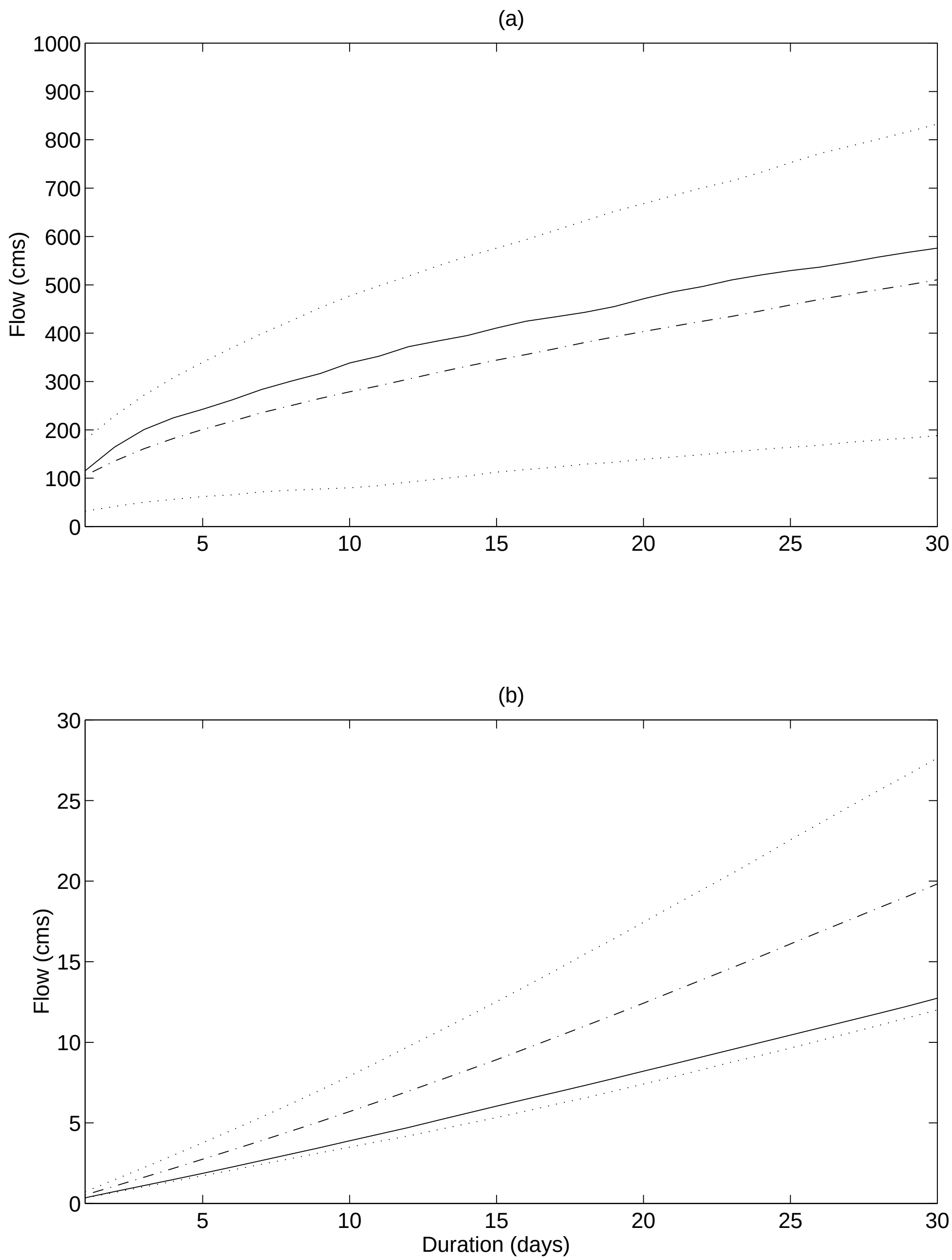

Figure 5. Series # 1: comparison of mean annual maximum (a) and minimum (b) flows over fixed durations; (—) historical flows, (-·-·) generated flows, (·····) 2-std. bands

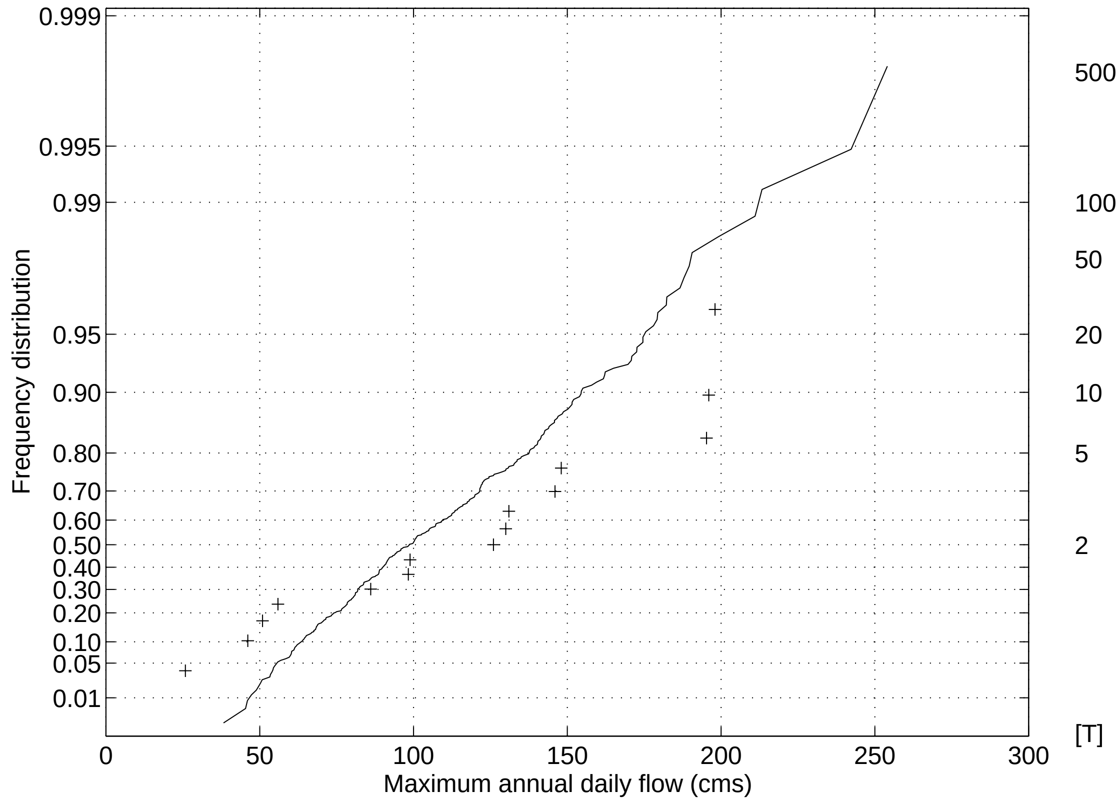

Figure 6. Series # 1: Gumbel probability plot of annual maximum daily flows; (o) historical flows, (---) generated flows