WATER RESOURCES RESEARCH, VOL. 29, NO. 8, PAGES 2545–2559, AUGUST 1993

Conceptual-Stochastic Modeling of Seasonal Runoff Using Autoregressive Moving Average Models and Different Scales of Aggregation

PIERLUIGI CLAPS

Dipartimento di Ingegneria e Fisica dell'Ambiente, Universita' della Basilicata, Potenza, Italy

FABIO ROSSI

Istituto di Ingegneria Civile, Universita' di Salerno, Fisciano, Italy

COSIMO VITALE

Istituto di Statistica, Universita' di Salerno, Fisciano, Italy

The statistical and phenomenological aspects of the runoff process observed on different scales of aggregation are taken as a priori information for the conceptually based stochastic modeling of seasonal runoff. Runoff is considered as the sum of two groundwater components, with over-year and subannual response lag, and of a purely random component representing the direct runoff. This scheme is equivalent to a linear system, with two parallel linear reservoirs plus a zero lag linear channel. The system output is the runoff, and the input is the effective rainfall, considered proportional to the direct runoff. Assuming the effective rainfall as a non-Gaussian periodic independent process and considering nonseasonal groundwater parameters, this conceptualization leads to an autoregressive and moving average (2, 2) stochastic process with periodic independent residual. Stochastic model parameters are directly related to the linear system coefficients, and the effective rainfall structure can be determined from the estimated model residual. In order to obtain parameter estimates consistent with the conceptual constraints, two estimation stages, on an annual and a seasonal basis, and an iterative procedure are needed. The model was applied to a number of time series of monthly streamflows in the Apennine regions of Italy with promising results.

1. INTRODUCTION

Most of the literature on stochastic modeling of seasonal streamflow consists of purely empirical models, which do not take into account physical information on the phenomenon. These models generally follow the approach proposed by *Box and Jenkins* [1970], who introduced autoregressive moving average (ARMA) models [e.g., *Salas et al.*, 1980].

Time series of seasonal runoff are characterized by periodic variability in several statistical characteristics. In other words, river flows for a given season of the year are statistically similar from year to year but may vary considerably across seasons. A common approach to seasonal data modeling is first to deseasonalize the series and then fit an appropriate nonseasonal stochastic model to the deseasonalized data [e.g., *Delleur et al.*, 1976; *Salas et al.*, 1980, section 5.3]. Moreover, it is usually desirable to deal with approximately normally distributed and homoscedastic model residuals, so that the data are transformed generally by taking logarithms. Attempts to refer to nontransformed variables were indeed made by *Fernandez and Salas* [1986, 1990] and *Sim* [1987], who proposed linear stochastic models with gamma marginal distribution, namely GAR and GARMA models.

With the aim of reproducing the autocorrelation structure of the seasonal hydrological series, particular attention has been directed to linear models with periodic parameters, since the approach proposed by *Thomas and Fiering* [1962]. The Thomas-Fiering model corresponds to a PAR(1) (periodic autoregressive model of order 1). *Tao and Delleur* [1976] proposed later the use of PARMA models, to which successive significant contributions were made by, among others, *Hirsch* [1979], *Salas et al.* [1982], and *Jimenez et al.* [1989], with particular reference to the PARMA(1, 1) model. Parameter estimation techniques of the periodic AR and ARMA models are quite cumbersome, particularly in the case of PARMA models, and some diagnostic checking procedures [e.g., *Salas et al.*, 1980] are only approximations of those developed for stationary, constant parameter models. Moreover, given the great number of parameters to be estimated, models with periodic parameters fail when applied to short series [*Kottegoda*, 1980].

Statistical identification techniques of time series models are generally based on the shape of the total and partial autocorrelation function. Due to the periodicity displayed by the autocorrelation function these techniques are difficult to apply to seasonal time series, unless the PARMA class of models is chosen a priori [*Jimenez et al.*, 1989].

As an alternative to the empirical model identification, a priori information on the runoff phenomenon can be used to build linear stochastic models founded on a conceptual interpretation of the process. The advantages of using conceptually based stochastic models can be summarized as follows: (1) The use of a priori information provides objective criteria for identification of model type and order, implicitly determining parsimony in the number of parameters; (2) links can be established between stochastic and

Paper number 93WR00550.
0043-1397/93/93WR-00550$05.00

conceptual parameters, and thus separate estimates of the latter allow validation of stochastic parameters, which is particularly important in situations of limited data; (3) it is viable, in principle, to evaluate model parameters at ungauged stations.

This paper aims to contribute in defining a systematic approach for the construction of conceptually based univariate stochastic models of seasonal runoff. Guidelines for model identification are discussed in section 2 of the paper, which refers to the appendices for the mathematical derivations. In section 3, the probabilistic model of the residual is analyzed. Sections 4 and 5 relate respectively to estimation and verification of the model. Application of the proposed model to seven time series of monthly runoff of basins located in the Apennine regions of Italy is the subject matter of section 6.

2. Conceptually Based ARMA Models of Seasonal Runoff

Links between conceptual and stochastic representation of the runoff process have often been reported in literature. The first stochastic models proposed for the runoff process were actually conceptually based, as summarized by *Klemeš* [1978]. On this subject, additional notable contributions concerning linear conceptual models include papers by *Spolia and Chander* [1974], *Salas et al.* [1981], *Koch* [1985], *Vandewiele and Dom* [1989], and *Salas and Obeysekera* [1992].

An important contribution to this topic, on a seasonal basis, has come from *Moss and Bryson* [1974], who derived a bivariate ARMA(1, 1) model with periodic parameters from the mass balance equations of a conceptual model of the runoff process. Univariate models with periodic parameters, namely PAR and PARMA models, were shown by *Salas and Obeysekera* [1992] to derive from a conceptual representation of the seasonal runoff. In particular, a PARMA(p + 1, q + 1) process was shown to result from the Thomas-Fiering model of a watershed [see *Fiering*, 1967] given the hypothesis of seasonal precipitation as following an ARMA(p, q) process. The PARMA(1, 1) process arises from the more reasonable hypothesis of rainfall following a periodic independent stochastic process. In the above context, a deeper insight into the basis for a conceptual representation of the runoff process is needed for a rational choice of PARMA model type and order.

2.1. *A Conceptual Rationale for Model Identification*

In this paper a conceptually based univariate stochastic model of seasonal runoff is proposed, based on the statistical and phenomenological aspects of the process displayed on different scales of aggregation. According to this rationale, which somewhat resembles an approach suggested by *Klemeš* [1983], seasonal runoff can be considered as the sum of components characterized by different response lags to precipitation. In general, it can be assumed that the streamflow is made up of one purely random component, one component with subannual lag and one with over-year lag.

The random component is the direct runoff and is assumed proportional to the effective rainfall $I_t = P_t - E_t$, where P_t is the precipitation and E_t is the evapotranspiration in season t. The effective rainfall is considered as a periodic independent stochastic process, in analogy to the assumption made by *Yevjevich and Karplus* [1973] with regard to the monthly precipitation. Proportionality between I_t and direct runoff corresponds to the hypothesis of very short response time of the processes involved in the direct runoff in comparison to the length of the seasonal interval.

The subannual lag component is considered as arising from groundwater storages with an average lag time of a few months and is responsible for the short-term persistence displayed by the seasonal time series. The over-year lag component is considered as coming from groundwater storages with an average lag time of a few years and is responsible for the long-term persistence displayed by the time series, which is more readily observable when the series are aggregated on an annual scale. Both of these two storage elements are considered as linear reservoirs, so that the system which transforms the effective rainfall in runoff is considered linear. This assumption limits the analysis to linear stochastic models.

With reference to the watersheds considered in this application, which are part of the Apennine region of Italy, the presence of two groundwater components in the runoff series is first of all recognizable by observing the runoff in the dry season and noting the rate of decrease in the discharge. In fact, observation of the series of daily runoff on a semilogarithmic scale highlights that during spring and summer, corresponding to the dry season, discharges decrease over time following approximately straight lines of different slopes. A year of records of a series with an over-year groundwater component is shown in Figure 1a, in comparison with one relative to an ephemeral stream (Figure 1b) which practically lacks discharges in the dry season. Another way of recognizing the existence of the over-year component is the presence of significant autocorrelation in annual runoff series, particularly when two distinct seasons, wet and dry, exist. In this case, the year can be started at the end of the dry season (hydrologic year) and the correlation can be ascribed only to the presence of an over-year lag storage [*Rossi and Silvagni*, 1980].

Summarizing, the above considerations address a conceptual model in which runoff is the sum of the outlets of two parallel linear reservoirs and of a periodic independent stochastic component. This scheme is depicted in Figure 2 and will be analyzed in detail in the ensuing discussions.

2.2. *Proposed Model*

In this analysis, runoff is considered as the result of a linear transformation of the effective rainfall, which is unknown since the approach here is univariate. However, the rainfall–net rainfall transformation is not part of the model. This avoids considering the evapotranspiration as a fixed proportion of the total precipitation, as assumed in the Thomas-Fiering watershed model.

With reference to Figure 2, discussed more fully in Appendix B, net rainfall I_t is subdivided into the following parts: aI_t, recharge to the over-year groundwater; bI_t, recharge to the subannual groundwater; and $(1 - a - b)I_t$, direct runoff. Coefficients a and b are called the recharge coefficient of the over-year and subannual groundwater, respectively. In general, they are variable, depending on the degree of soil saturation. The question whether or not to consider constant a and b will be discussed later.

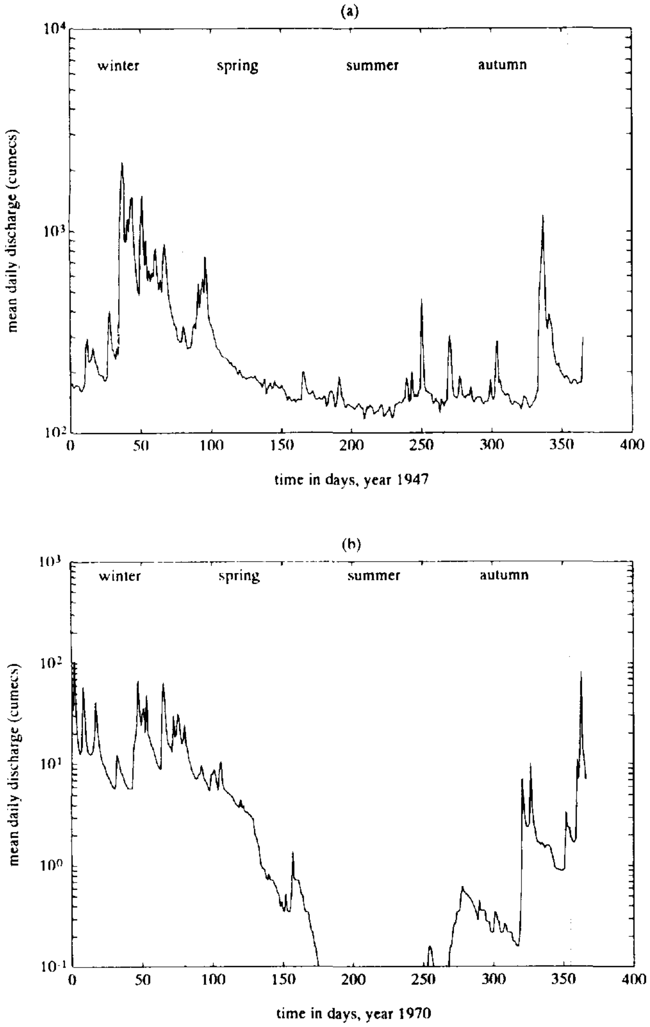

Fig. 1. Daily streamflows in semilogarithmic scale: (*a*) River Tevere; (*b*) River Tammaro. Streamflows are in units of cubic meters per second.

Each of the recharge volumes aI_t and bI_t contribute both to the storage volumes V_t and W_t, respectively, and to the outflow, with proportions defined by the recharge recession coefficients r_k and r_q defined in Appendix A. These coefficients are incorporated in the conceptual scheme essentially to take advantage of possible a priori knowledge of the form of the recharge over time, which can be important particularly when considering runoff on the annual scale. Coefficients r_k and r_q depend on the storage coefficients k and q, respectively, and on the a priori within-season distribution of precipitation. They are considered as predefined constants, while the conceptual model parameters are a, b, k, and q.

Considering nonseasonal parameters and periodic independent stochastic input, the outlet of the linear system of Figure 2 corresponds to an ARMA(2, 2) stochastic process with periodic independent residual, PIR-ARMA(2, 2), as shown in Appendix B. This stochastic process is expressed as

$$d_t - \Phi_1 d_{t-1} - \Phi_2 d_{t-2} = \varepsilon_t - \Theta_1 \varepsilon_{t-1} - \Theta_2 \varepsilon_{t-2} \qquad (1)$$

where $d_t = D_t - E[D_t]$ is the zero-mean runoff, and ε_t is the zero-mean residual. Also given in Appendix B are explicit relationships between the four conceptual parameters and the AR and MA parameters, which are Φ_1, Φ_2 and Θ_1, Θ_2 respectively.

For ephemeral streams, which have a negligible over-year groundwater component, the conceptual model is made up of only one linear reservoir plus a linear channel with zero lag (Figure 3). As noted in Appendix B, based on the hypothesis of periodic independent input, runoff corresponds in this case to a PIR-ARMA(1, 1) stochastic process

$$d_t - \Phi d_{t-1} = \varepsilon_t - \Theta \varepsilon_{t-1} \qquad (2)$$

Again, parameters Φ and Θ are directly linked to the recharge and storage conceptual parameters.

As underlined by *Salas et al.* [1981], processes in (1) and (2) are to be considered as restricted ARMA processes. In fact, due to the underlying conceptual hypotheses, their parameter spaces, given in Appendix B for the PIR-ARMA(2, 2) process and by the above authors for process (2), are subspaces of those corresponding to the general ARMA processes.

2.3. *Possible Improvements in the Model Structure*

As shown before, a constant parameter PIR-ARMA(2, 2) model is identified for seasonal streamflows. The hypothesis of nonseasonality of the conceptual parameters requires some discussion with reference to the alternative of considering all parameters as seasonally variable. In this regard, the increase in model accuracy arising when considering seasonal variability of the storage coefficients k and q does not seem remarkable. In fact, even in the context of models with periodic parameters some authors [e.g., *Moss and Bryson*, 1974; *Hirsch*, 1979; *Salas and Obeysekera*, 1992], who consider the presence of only one reservoir, tend to assume the storage coefficient constant over seasons.

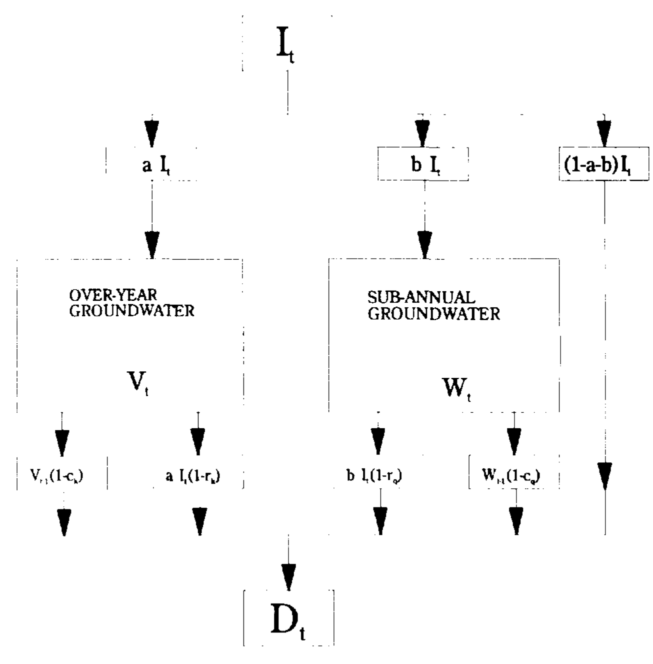

Fig. 2. Representation of the linear system underlying the formation of monthly streamflows.

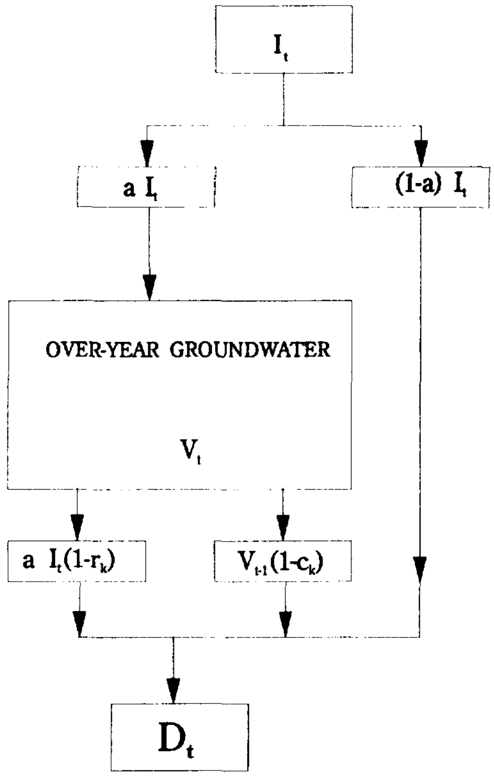

Fig. 3. Linear system representing the response of one linear reservoir plus a zero lag linear channel (e.g., conceptual model of annual runoff).

Regarding possible refinements of the model that can arise by relaxing the hypothesis of constancy of parameters a and b, it is first worth remarking that in the proposed model coefficients a and b do not represent proportions in the rainfall–net rainfall transformation but indicate percentages of the net rainfall transformed in each of the three components of runoff. As discussed in Appendix B, by considering seasonal variability only for recharge coefficients the PIR-ARMA(2, 2) becomes a PARMA(2, 2) process with nonseasonal AR parameters. Similarly, when considering ephemeral streams, the PIR-ARMA(1, 1) becomes a PARMA(1, 1) process with nonseasonal AR parameter.

Thus PARMA models with nonseasonal AR parameters can be a suitable, more refined, alternative to PIR-ARMA models. However, the increase in complexity when adopting models with periodic parameters is remarkable, given that, excluding mean, variance, and residual variance, in the case of monthly data the PIR-ARMA(2, 2) has four parameters while the PARMA(2, 2) with nonseasonal AR coefficients has 26 parameters.

2.4. *Models of the Aggregated Runoff Process*

An obvious requirement for a model built for a given time scale should be its compatibility with the models developed for the aggregated scales. As *Obeysekera and Salas* [1986] have commented, this requisite has seldom been addressed. Based on this point, some comments on the runoff process on the annual scale are in order.

It has long been recognized that the long-term persistence effects displayed by annual runoff series, and particularly the Hurst effect [e.g., *Lloyd*, 1967], can be adequately reproduced by the ARMA(1, 1) model [*O'Connell*, 1971]. This empirical finding was later substantiated by *Rossi and Silvagni* [1980], *Salas and Smith* [1981], and *Salas et al.* [1981] based on a conceptual interpretation of the phenomenon on that scale.

In particular, *Rossi and Silvagni* [1980] gave more credit to the ARMA(1, 1) model with respect to higher-order ARMA models. They showed that analyzing data in the hydrologic year, in the presence of definite wet and dry seasons, avoids the consideration of spurious correlation effects. Moreover, the restricted parameter space of the conceptually based ARMA(1, 1) model given by *Salas et al.* [1981] exactly matches the conditions in which, according to *O'Connell* [1971], the Hurst effect can be reproduced by the ARMA(1, 1) model. Incidentally, the conceptual model used by *Salas et al.* [1981] is modified here (Appendix B) so that both evaluation of conceptual parameters from the estimated stochastic coefficients and evaluation of the net rainfall from the estimated residual process are feasible.

It has been also shown [*Rossi and Silvagni*, 1980; *O'Connell et al.*, 1991] that the ARMA(1, 1) model is not always needed to describe the structure of annual runoff series. In fact, *Rossi and Silvagni* [1980], analyzing series from rivers in southern Italy, showed that in the absence of a significant over-year groundwater component, runoff in the hydrologic year can be considered as an independent random process, well reproduced by a Box-Cox transformation of the normal distribution. It may be concluded that annual runoff can be considered as generally following an ARMA(1, 1) process and that in some cases, for instance when considering ephemeral streams, runoff in the hydrologic year can be regarded as a white noise process.

The issue of determining the structure of aggregated processes has been addressed by a number of researchers [e.g., *Kavvas et al.*, 1977; *Vecchia et al.*, 1983; *Obeysekera and Salas*, 1986]. In particular, *Vecchia et al.* [1983] derived the ARMA(1, 1) process from the aggregation of the PARMA(1, 1). By extension of this finding, the aggregation of a PARMA(2, 2) results in an ARMA(2, 2) process.

Considering an annual aggregation scale, within the conceptual representation suggested here (see Figure 2), it is not possible to distinguish the subannual groundwater response from that of the other subannual lag components, so that they are all included in annual direct runoff. Consequently, a conceptual system made up of only one linear reservoir plus a zero lag linear channel (Figure 3) is assumed. Considering the annual net rainfall as a stationary independent stochastic process, this conceptual representation corresponds to an ARMA(1, 1) process.

3. PROBABILISTIC MODEL OF THE RESIDUAL

The marginal probability distribution of the residual ε_t of a stochastic model is usually analyzed empirically, either by transforming the runoff series in order to reduce ε_t to normality or by building models which incorporate a gamma distribution for residuals. A different approach to the analysis of the residual is attempted when a conceptual meaning is imposed.

The residual ε_t of the PIR-ARMA model is a periodic independent random process, i.e., its probability distribution parameters are considered as varying with the seasons. As shown in Appendix B, ε_t represents an estimate of the effective rainfall I_t. Because the conceptual model and its stochastic form (1) are only an approximation of reality, it must be assumed that in the residual of an actual PIR-ARMA model there is an error term, which can induce negative values in the effective rainfall. Consequently, (1) can be reconsidered using the actual stochastic model residual ε'_t:

$$d_t - \Phi_1 d_{t-1} - \Phi_2 d_{t-2} = \varepsilon'_t - \Theta_1 \varepsilon'_{t-1} - \Theta_2 \varepsilon'_{t-2} \qquad (3)$$

with $\varepsilon'_t = \varepsilon_t + \xi_t$, ε_t being the component having conceptual meaning and ξ_t representing a Gaussian error term, with zero mean and variance σ^2_ξ. These components may be considered uncorrelated and the following hold:

$$E[\varepsilon'_t] = E[\varepsilon_t] \qquad \sigma^2_{\varepsilon'} = \sigma^2_\varepsilon + \sigma^2_\xi \qquad (4)$$

Therefore reference can be made to the "estimate with error" of the net rainfall:

$$I'_t - \mu_{I'} = \frac{\varepsilon'_t}{c} \qquad (5)$$

where $\mu_{I'}$ equals $E[D_t]$ and $c = (1 - ar_k - br_q)$, and to the "conceptual estimate" of the net rainfall:

$$I_t - \mu_I = \frac{\varepsilon_t}{c} \qquad (6)$$

which, given (4), are related by

$$\mu_{I'} = \mu_I$$
$$\sigma^2_{I'} = \frac{\sigma^2_\varepsilon}{c^2} + \frac{\sigma^2_\xi}{c^2} = \sigma^2_I + \frac{\sigma^2_\xi}{c^2} \qquad (7)$$

Because of its meaning, the variable I_t should assume only positive values and present finite probability at zero. A reasonable probabilistic representation of I_t results by considering the variable as the sum of a Poissonian number of events with exponentially distributed intensity [e.g., *Benjamin and Cornell*, 1970]. The corresponding probability density function has the expression

$$P[I = 0] = e^{-\nu} \qquad I = 0$$
$$f_I(I) = e^{-\lambda I - \nu}(\nu\lambda/I)^{1/2}\mathcal{I}_1[2(\lambda I \nu)^{1/2}] \qquad I > 0 \qquad (8)$$

where λ is the exponential parameter, ν is the Poisson parameter and $\mathcal{I}_1(x)$ is the modified Bessel function of order 1. The above probability function is called the Bessel distribution. This representation is more realistic than the gamma distribution, which results from the sum of a fixed number of exponentially distributed events. Using $\beta = 1/\lambda$ as the exponential parameter, which can be expressed in millimeters, relations between the sample moments and the distribution parameters are

$$\mu_I = \nu\beta \qquad \sigma^2_I = 2\nu\beta^2 \qquad \gamma_I = \frac{3}{(2\nu)^{1/2}} \qquad (9)$$

with γ_I as the skewness coefficient. Relations in (9) refer only to the conceptual estimate I_t of the net rainfall.

The probabilistic model of the net rainfall estimated with error, I'_t, is the sum of a Bessel and a normal random variable, with parameters ν, β and $\sigma^2_0 = \sigma^2_\xi/c^2$. Moments of this distribution are [*Claps*, 1992]

$$\mu_{I'} = \nu\beta \qquad \sigma^2_{I'} = \sigma^2_0 + 2\nu\beta^2 \qquad \gamma_{I'} = \frac{6\nu}{\left(\frac{\sigma^2_0}{\beta^2} + 2\nu\right)^{3/2}} \qquad (10)$$

The above relations can be used to estimate parameters for the probability distribution of the residual of each season.

4. Parameter Estimation of the PIR-ARMA Model

The analysis of the phenomenological aspects of the seasonal runoff clearly shows that the variability of mean, variance, and autocorrelation is directly or indirectly due to the seasonal variability of rainfall. For this reason, the series is not deseasonalized, since this operation not only does not completely eliminate periodicity in the autocorrelation function [*Tao and Delleur*, 1976] but also causes the removal of characters in the series which have a subannual evolution, such as the effects of the subannual lag groundwater. Moreover, in order to preserve the formal correspondence between the conceptual and stochastic representations of the process, data are not transformed.

4.1. *Standard Parameter Estimation*

The estimates of stochastic parameters of the PIR-ARMA model are obtained by means of the least squares method, using the TSP statistical package (Quantitative Micro Software, Irvine, California, 1982). As shown by *Pierce* [1971], least squares parameter estimates of a PIR-ARMA model have finite variance and are asymptotically normal but are not the most efficient in a statistical sense, i.e., mean square error is only asymptotically minimum among the linear estimates, the asymptotic condition corresponding to white noise residual. A lowering in the statistical efficiency is the price one must pay in order to safeguard the formal structure of the conceptualization in the stochastic model.

In the estimation of stochastic parameters it must be recognized that the model considered is a restricted ARMA model, because relations (B22)–(B25) between conceptual and stochastic parameters constitute constraints on the admissible region of stochastic parameters. This leads to more serious estimation problems. In fact, under these conceptual constraints, parameters Φ_1 and Φ_2 make sense only near the nonstationarity condition (see Appendix B) where their estimates can show high variability [*Box and Jenkins*, 1970, paragraph 6.3.5.].

At the application stage, the simultaneous estimation of PIR-ARMA(2, 2) parameters was found to be unreliable, in the sense that estimates of the AR parameters resulted outside of their admissible space. As underlined above, the reasons for this unreliability are ascribed to the nature of conceptual constraints. This is also in agreement with a comment by *Acton* [1970] on the ill conditioning of the problem of estimating all four parameters of a system whose response is the sum of two exponentials.

4.2. *Parameter Estimation Over Different Aggregation Scales*

An approach for parameter estimation based on the use of information from different aggregation scales is proposed, as

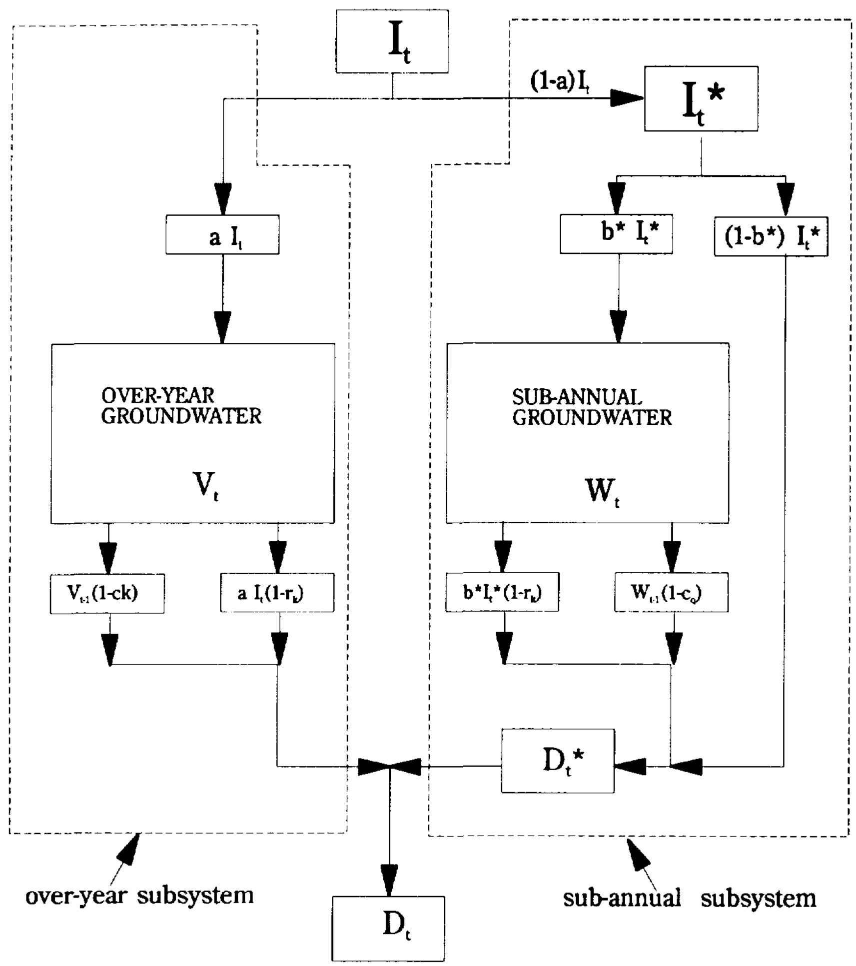

Fig. 4. Linear system of monthly streamflows reshaped as a combination of two subsystems.

an alternative to the usual method discussed above. Parameters related to the over-year groundwater are estimated on the annual scale while those related to subannual groundwater are estimated on the seasonal scale. Information is transferred from the annual to the seasonal aggregation scale by means of the conceptual hypotheses, so as to increase the reliability of estimates and the compatibility with the aggregated scale.

To better distinguish the conceptual schemes considered on the different scales, the original conceptual model was reshaped (Figure 4) highlighting two subsystems: one with over-year response, made up of only the over-year groundwater, and one with subannual response. These subsystems are considered separately in the estimation stage.

Estimation of the over-year groundwater parameters on the annual time scale. On the annual scale, the ARMA(1, 1) model is selected according to what was discussed in section 2. To preserve the conceptual meaning of parameters, data records, referring to the hydrologic year, are not transformed. Estimation of the AR and MA stochastic parameters allows the determination of coefficients a and k and the estimation of the annual net rainfall series I_t by means of relations (B11). Coefficients k, expressed in years, and a are assumed valid for all of the lower scales.

The hypothesis on the form of the recharge function leading to the value of r_k must be specified to accomplish the evaluation of conceptual parameters. On the annual time scale, the form of the recharge function $r(\tau)$ depends on the climatological regime of the site being investigated, which can show minima and maxima of rainfall in different seasons. As a first approximation, an impulse recharge was considered with occurrence time T set in the centroid of the runoff regime curve, used as an approximation of the net rainfall regime curve. These curves designate the diagrams of the nondimentional monthly averages of runoff and net rainfall. In this case, (A6) is used to compute r_k. In addition, to better assess the variability of r_k, the previous estimate is compared to the value of r_k obtained by means of (A8), valid for uniform net rainfall distribution.

If only two distinct seasons exist, one dry and one wet, a more refined evaluation of this coefficient could be achieved by deriving the analytical expression of r_k for a gamma recharge curve, assuming that the shape of a gamma function is well suited for the reproduction of the runoff regime. However, given that the centroid of the recharge distribution must be preserved by the gamma curve, the consequent value of r_k will assume intermediate values between the two values corresponding to the previous hypotheses.

Estimation of the subannual groundwater parameters on the seasonal time scale. The conceptual subsystem with subannual response, as shown in Figure 4, has the same structure as the conceptual model of Figure 3, formerly considered for the annual runoff. In this case, the input to the system is the fraction $I_t^* = (1 - a)I_t$ of the effective rainfall, while runoff is the difference between the total runoff D_t and the over-year groundwater runoff G_t. Given I_t^* as a periodic independent process, D_t^* is a PIR-ARMA(1, 1) process. Relations (B10) and (B11) between conceptual and stochastic parameters obtained for the scheme in Figure 3 hold with substitution of symbols k, a, and I_t with q, $b^* = b/(1 - a)$, and I_t^*, respectively.

The possibility of estimating parameters of the PIR-ARMA(1, 1) model is conditioned by the determination of the runoff series D_t^*, which is unknown. In fact, the over-year groundwater runoff G_t can be obtained only when the seasonal net rainfall series is determined. Since the evaluation of the net rainfall series is a result of the parameter estimation itself, an iterative estimation procedure is needed.

The iterative procedure is characterized by (1) preliminary estimation of trial net rainfall; (2) calculation of the over-year groundwater runoff and of the D_t^* series due to the trial input; (3) estimation of the PIR-ARMA(1, 1) model and of the updated net input; and (4) comparison of the updated and the trial input: if too different, return to step 2 with the updated series as trial input. These steps are commented on in detail below.

Step 1: A preliminary seasonal net rainfall series can be obtained by disaggregating the net rainfall series estimated on the annual basis, with the constraint of preserving the shape of the dimensionless seasonal runoff mean. Practically, if I_j is the net rainfall estimated for the year j, the values corresponding to the season τ of year j are $I_{j,\tau} = I_j g(\tau)$, where $g(\tau) = \mu_\tau/\mu$, μ_τ being the mean of runoff in season τ and μ being the general runoff mean.

Step 2: Equation (B4), obtained with reference to a linear reservoir, expressed in transfer function form, gives the over-year groundwater runoff G_t as a response to the trial input aI_t:

$$G_t = c_k G_{t-1} + aI_t(1 - r_k) + aI_{t-1}(r_k - c_k) \qquad (11)$$

where $c_k = e^{-1/k}$ and a is estimated on the annual scale. To use this expression on the seasonal scale, k is expressed in seasons and r_k is computed with (A8), as will be discussed hereafter.

To obtain the G_t series from (11), initial values of G_t and I_t are needed. If there are two distinct climatic seasons, it is a fairly straightforward task to define the initial values starting the series at the end of the dry season. However, for long series records, accuracy in the determination of these data is not essential. For short series, estimation of the initial values could be made by means of backcasting techniques.

The seasonal subsystem runoff series is obtained as $D_t^* = D_t - G_t$. Possible negative values of D_t^* are not adjusted in this step, because this can alter the convergence of the procedure. When the convergence is achieved, negative values of D_t^* can be due to an incorrect estimate of the initial value of G_t or to an unsatisfactory estimate of the coefficient a.

Step 3: The PIR-ARMA(1, 1) model is fitted to the zero mean series d_t^* of the subannual subsystem runoff obtained in the previous step. Estimates of the stochastic parameters and of the residual $\varepsilon_t'^*$ allow the calculation of b^* and q as well as the evaluation of the net rainfall series:

$$I_t'^* = \frac{\varepsilon_t'^*}{(1 - b^* r_q)} + \mu^* \qquad (12)$$

where $\mu^* = E[I_t^*] = E[D_t^*]$. Note that reference is always made to the net rainfall estimated with error.

Actually, (12) is only the simplest relation which allows correspondence of the first two moments between residual and net input as reported in (B12). The same correspondence could be reached in different ways and could have an impact in reducing errors in the estimated series, which are responsible for negative values found in $I_t'^*$.

The a priori evaluation of the recharge recession coefficients on the seasonal scale requires some comments. Coefficients r_k and r_q are calculated, once k and q are known, through (A8), which corresponds to the hypothesis of uniform net rainfall during the season. This hypothesis is indirectly discussed by *Moss and Bryson* [1974] with reference to the effects of the rainfall distribution on mean and variance of a recharge recession coefficient, say r_q. It is shown that both when the number of storms per interval is relatively high and when the number of events approaches zero, the mean of r_q is practically equal to that corresponding to uniform recharge. Given this, a uniform within-season recharge seems more reasonable than any other possible predetermined recharge function.

The updated net input to the whole system is obtained by simply restoring mean and variance with the position

$$I_t' = \frac{I_t'^* - E[D_t^*]}{(1 - a)} + E[D_t] \qquad (13)$$

Step 4: The updated net input series I_t' is compared with the trial series used in step 2. Iterations are stopped when the standard deviation of the two series differ by less than 5%. Otherwise, the updated series is used as trial input in step 2.

When procedure convergence is reached the following are determined: (1) the conceptual model coefficients, i.e. the characteristics of the linear system; (2) the coefficients Φ_1, Φ_2, Θ_1, Θ_2 of the PIR-ARMA(2, 2) stochastic model, through (B18)–(B21); and (3) the residual variance $\sigma_{\varepsilon'}^2$ of the PIR-ARMA(2, 2) model, from the residual variance $\sigma_{\varepsilon'*}^2$ of the PIR-ARMA(1, 1) model of the subannual subsystem:

$$\sigma_{\varepsilon'}^2 = \sigma_{\varepsilon'*}^2 \left(\frac{1 - ar_k - br_q}{1 - a - br_q} \right)^2 \qquad (14)$$

5. VERIFICATION OF THE CONCEPTUALLY BASED PIR-ARMA MODEL

The verification of conceptually based stochastic models assumes different aspects, with respect both to objectives and to methods, from the diagnostic checking operated on empirically identified models. In fact, the aim of the Box-Jenkins approach to time series modeling, in which the series must "speak for themselves," is to achieve optimal statistical representation of the data analyzed. A model which is identified for a given river flow series based on this approach is not valid in general for different rivers. Conversely, a conceptually based model is built in order to achieve the best representation of the data consistent with an interpretation of the runoff phenomenon, so that its validity can be general.

Model verification in the Box-Jenkins approach is made with tests on parameter significance and on independence of

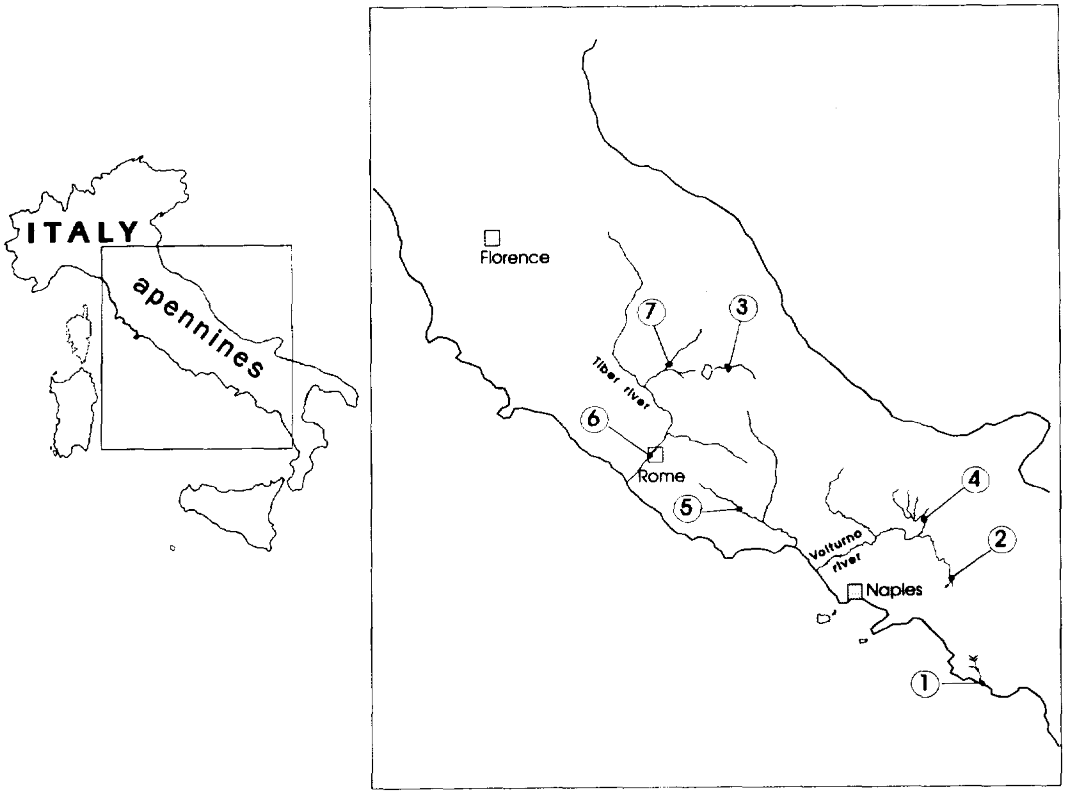

Fig. 5. Locations of the gauging stations considered in Italy. The numbers relate to the codes assigned in Table 1.

residuals, whose results may or may not lead to the decision of rejecting the model. If the model is rejected, an alternative one is to be selected, which will be less complex in general if there is lack of parameter significance while a more complex model is generally needed if the hypothesis of independence of residual is not verified. Thus, in some way, diagnostic checking can be considered as part of the model identification stage. This is not exactly the case when conceptual hypotheses are invoked for model identification, because only limited alternative models are compatible with the conceptualization of the runoff phenomenon.

In the previous sections, a limited number of conceptually based models have been identified for seasonal streamflows, namely ARMA models with periodic independent residual (PIR-ARMA) or models also with periodic stochastic parameters (PARMA). In this paper, model construction is considered only with reference to PIR-ARMA models, because of their simpler structure.

TABLE 1. Characteristics of the Stations and Time Series Considered

Code	Station	Area, km^2	Record Length, years	Mean Annual Rainfall, mm
1	Alento at Casalvelino	284	13	521
2	Calore Irpino at Montella	123	26	609
3	Giovenco at Pescina	139	11	277
4	Tammaro at Pago Veiano	555	13	350
5	Sacco at Ceccano	922	12	483
6	Tiber at Rome	16,545	50	448
7	Nera at Torre Orsina	1,445	25	606

On the annual scale, the ARMA(1, 1) model is first tested for comparison with the possible alternative of a white noise model. According to the scheme presented in the previous sections, when passing to the seasonal scale these two alternatives lead respectively to a PIR-ARMA(2, 2) and to a PIR-ARMA(1, 1) model.

The validity of the ARMA(1, 1) model on the annual scale is essentially verified by means of the autocorrelation function and of the significance of the autoregressive parameter estimate. The observance of the conceptual constraints must be also guaranteed. Once it is decided that the ARMA(1, 1) formulation, rather than the white noise formulation, is the model for the annual data at hand, testing procedures can help the assessment of the reliability of estimates.

At the application stage it is often observed that the significance of the moving average parameter Θ is lower than that of the AR parameter Φ. However, a high standard error of Θ cannot lead to the rejection of the ARMA(1, 1) model and the consequent selection of an AR(1) model, because the latter does not correspond to the conceptual scheme.

Low significance of the Θ estimate implies low reliability of the estimated recharge coefficient a. To support the estimate of a, an index similar to the base flow index [*Institute of Hydrology*, 1980] is introduced. It is called the deep flow index (DFI) and represents the ratio between the average over-year groundwater runoff and the average total runoff. The former term is calculated as the mean of the annual minima of the mean monthly discharge (smoothed minima technique). This corresponds to the assumption that

TABLE 2. Group 1 Series: Parameter Estimates of the ARMA(1, 1) Stochastic Model Applied to Annual Runoff Data

Station	Φ	Standard Error Φ	Θ	Standard Error Θ	R^2
3	0.712	0.687	0.42	0.777	0.216
6	0.744	0.438	0.536	0.464	0.09
7	0.788	0.366	0.421	0.424	0.261

TABLE 4. Group 1 Series: Stochastic Parameter Estimates of the PIR-ARMA(2, 2) Model Applied to Monthly Runoff Data

Station	Φ_1	Φ_2	Θ_1	Θ_2
3	1.61	−0.621	1.35	−0.387
6	1.52	−0.535	1.16	−0.202
7	1.74	−0.741	1.19	−0.228

groundwater runoff is linearly variable with time and that the annual minimum of the mean monthly discharge can be considered as being produced only by deep groundwater runoff. The validity of the DFI as a measure of the over-year groundwater runoff can be recognized if we consider that this runoff component is underestimated in the wet season and overestimated in the dry season, thus giving rise to opposite sign errors which minimize the global error.

On the seasonal scale, whatever the model of annual data, two parameters are to be estimated, corresponding to the storage and recharge parameters of the subannual groundwater. Here again, MA estimates are generally less significant than the AR ones. So, the estimate of the recharge parameter b can result in low reliability. A rough verification of the estimate of b can be made by seeing if the reconstructed series of the global groundwater runoff component correctly reproduces the recession periods and the streamflow minima of the observed series. This check is only preliminary to a simulation study, which is a more general tool for the global testing of model hypotheses.

Goodness-of-fit tests on stochastic models are generally based on the analysis of residuals. Based on the results of these tests one can decide to use more complex models. On the seasonal scale, more complex alternative models to the PIR-ARMA arise by considering seasonal variability of parameters: PARMA(1, 1) model in case of independence of annual data and PARMA(2, 2) model in the other case. These alternatives can be considered with regard to the analysis of different aspects of the residual.

In the classical time series analysis the residual is tested with respect to independence, normality, and stationarity. While standard tests of independence can be applied to the PIR-ARMA model residual, a different point of view is to be considered when dealing with its probabilistic structure, not to mention its periodic characteristics. The residual is in fact considered here as the sum of a conceptual and an error components, so that its analysis is completed by checks on both components.

Validation of the conceptual part, considered as effective rainfall, could be made either by a comparison with the general characteristics of the total precipitation in the basin upstream or by using information from other stations located in climatically and geologically similar regions. This issue is currently being investigated.

The characteristics of the error term ξ_t are related to the model performance, which can be considered good if $\mu_\xi \approx 0$ and if σ_ξ^2 is constant among seasons. In addition, the smaller σ_ξ^2, the smaller the global unexplained variance. High seasonal variability of σ_ξ^2 can be considered an index of the low efficacy of the hypothesis of nonseasonal parameters.

Assessment of alternative models with periodic parameters is outside the objectives of this paper. However, the use of PARMA models appears hampered by awkward parameter estimation. In particular, estimation of the PARMA(2, 2) model with seasonal or nonseasonal AR parameters is practically unmanageable, while the PARMA(1, 1) model is suitable, in principle, only for streamflow series without the over-year groundwater component. Moreover, the problem of dealing with conceptually constrained estimates is yet to be faced for this class of models. If significant inadequacies are found in the model verification, before deciding to use a more complex model which may be theoretically capable of more detailed explanation of the statistical characteristics of the series, the available amount of data must be taken into account with respect to the requirement of reliable estimation of parameters.

6. PIR-ARMA MODEL APPLICATION AND TESTING

The proposed procedure was applied to seven time series of monthly runoff of rivers in central southern Italy (see Figure 5 and Table 1). For the selected rivers no considerable regulations or diversions during the observation period were reported and the snowmelt runoff contribution can be neglected.

The watersheds under study are all characterized by the geology and climate of the Apennine mountains. One geological feature of the Apennines relevant to the runoff process on a monthly and annual scale is the presence of several great fractured carbonate massifs, containing large aquifers at their base. The climate of the Apennines is characterized by two distinct seasons: a rainy season, during autumn and winter, and a dry season, during spring and summer. The hydrologic year thus begins on October 1.

TABLE 3. Group 1 Series: Summary of the Conceptual Model Parameters of Annual Runoff

Station	k	r_k^*	$r_k^\dagger$	a^*	$a^\dagger$	DFI
3	2.94	0.83	0.85	0.61	0.59	0.62
6	3.38	0.85	0.87	0.53	0.52	0.53
7	4.2	0.89	0.89	0.71	0.71	0.77

*Hypothesis of impulsive input in the centrum of the runoff regime.
†Uniform within-year input.

TABLE 5. Group 2 Series: Stochastic Parameter Estimates of the PIR-ARMA(1, 1) Model Applied to Monthly Runoff Data

Station	Φ	Standard Error Φ	Θ	Standard Error Θ
1	0.576	0.098647	−0.098	0.127273
2	0.55	0.075862	−0.073	0.095425
4	0.579	0.101224	−0.068	0.128302
5	0.526	0.126747	−0.032	0.151659

Watersheds upstream from the stations considered can be subdivided into two categories, according to their hydrogeologic characteristics and to the relevance of runoff in the dry season. The first category, which will be referred to as group 1 and includes the stations 3, 6, and 7, refers to watersheds partly made up of highly permeable carbonate formations and presenting remarkable runoff in the dry season; group 2 includes stations 1, 2, 4, and 5, whose watersheds can be classified as quite impermeable over their whole area, in which runoff is very low in the dry season and whose data do not display significant autocorrelation on the annual scale. Basins of group 2 are mainly made up of clay-type geologic formations, and present hill morphology with moderately steep slopes. All the basins considered present high climatic homogeneity, and within each group notable similarity in the hydrogeological structure can be assumed.

For all the basins under study, monthly series were analyzed, because the monthly scale was considered to be consistent with the hypothesis of randomness of the direct runoff component.

According to model identification criteria on the annual scale discussed previously, the ARMA(1, 1) model was fitted to the three series of annual runoff from rivers belonging to group 1, which present remarkable over-year groundwater contribution. Estimates of the stochastic parameters together with their standard errors and values of the R^2 are shown in Table 2. Estimates of the AR parameter Φ resulted in values that are much more statistically significant than those of the MA parameter Θ, as can be recognized from the standard error values. The low R^2 values, particularly for station 6, point out the high overall variability of the process on the annual scale.

Table 3 shows the values of conceptual parameters a and k obtained by using (B11). As can be observed, the order of magnitude of k was of 3–4 years, which is reasonable as a measure of over-year persistence. Two estimates of coefficient a are obtained for different hypotheses of distribution of the recharge, as anticipated in Section 4. These estimates differ only slightly, so that a more careful definition of the input distribution does not appear to be a point of major importance.

TABLE 6. Group 1 and 2 Stations: Estimates of the Subannual Subsystem Conceptual Parameters

Station	Group	q	b	r_q
1	2	1.84	0.76	0.77
2	2	1.67	0.77	0.75
3	1	2.23	0.277	0.807
4	2	1.84	0.79	0.77
5	2	1.53	0.73	0.73
6	1	1.66	0.363	0.75
7	1	3.56	0.256	0.869

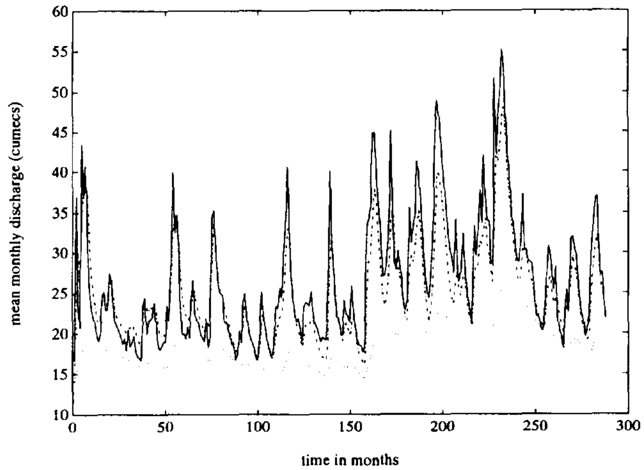

Fig. 6. River Nera. Period: October 1946 to September 1969. Observed monthly streamflow series (solid), reconstructed over-year groundwater runoff (dotted), and reconstructed total groundwater runoff (dash-dot). Streamflows are in units of cubic meters per second.

The estimate of the recharge parameter a was checked through the DFI index. The good agreement shown in Table 3 between coefficient a and DFI indicates that high standard errors of the estimates of Θ do not exclude a reasonable result in terms of the recharge coefficient estimate.

The PIR-ARMA(2, 2) model was then fitted to the series of monthly runoff of the stations in group 1. Table 4 shows the estimates of the stochastic parameters, obtained through the iterative procedure shown in section 4. The PIR-ARMA(1, 1) model was fitted to the group 2 series, and Table 5 reports the estimates of parameters Φ and Θ as well as their standard errors. As can be observed, the estimates of Θ still have high standard errors in comparison with those of Φ, indicating a relatively higher uncertainty on the corresponding values of b.

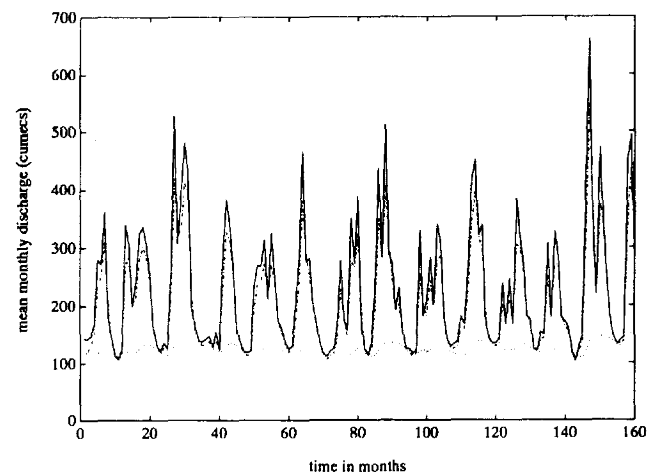

Fig. 7. River Tevere. Period: October 1921 to January 1935. Observed monthly streamflow series (solid), reconstructed over-year groundwater runoff (dotted), and reconstructed total groundwater runoff (dash-dot). Streamflows are in units of cubic meters per second.

TABLE 7. Box-Pierce Portemanteau Test Applied to Groups 1 and 2 Series

Station	Lag	f = lag-p-q	X_f^2	Q_{stat}
1	24	22	33.9	20.8
2	50	48	65.2	52.3
3	20	16	26.3	21.0
4	24	22	33.9	37.0
5	22	20	31.4	25.4
6	98	94	117.4	103.2
7	48	44	60.5	55.9

For all series the autocorrelation is calculated up to lag $N/6$, with N equal to the number of recorded data. The condition $Q_{stat} < X_f^2$ is satisfied by all stations except station 4.

Table 6 shows the values of the estimated subannual subsystem parameters for all the series considered. The values of the storage constant q are, as expected, of the order of magnitude of a few months. Tables 5 and 6 show that the parameter estimates relative to the group 2 series are characterized by remarkable similarities. Values for the storage constant q and the recharge coefficient b fall within a narrow range, which is consistent with the substantial hydrogeological uniformity of the basins in this group.

By using the net input estimated by the means of (5), the reconstructed subannual and over-year groundwater runoff series were examined, partly so as to check the validity of the estimate of the recharge parameter b. An example of the reconstruction of groundwater runoff is shown in Figures 6 and 7, relative to the Nera and Tiber rivers, respectively. The figures show that the reconstructed series are consistent with the hypotheses made, particularly in correspondence to the minima of the observed series. Some inconsistencies emerged in the first months of the record, mainly due to the uncertainty of the initial value to assign to the over-year groundwater component.

Commenting on the estimates of parameter b and the reconstruction of the global groundwater component, it can be said that the percentage of this component with respect to the total runoff seems quite high. This outcome could be either a typical effect of the aggregation of the data or the result of interference between the subsurface runoff, which has a response lag of much less than 1 month, and the subannual groundwater component, which is considered as deriving from groundwater storages with seasonal recession. An interpretation of the global meaning of the subannual component could be obtained by analyzing data on a lower scale of aggregation. The use of an appropriate conceptual-stochastic model on the daily scale, which is under study as an extension of the approach proposed here, might be helpful for this purpose.

A test on the uncorrelation of residuals and a comparative rating of the explained variance were made as standard procedures of verification of the global efficiency of the stochastic component of the model. The Box-Pierce portemanteau test [e.g., *Salas et al.*, 1980] was first applied to the standardized residuals. As shown in Table 7, the hypothesis of overall uncorrelation of residuals is met by all series but one.

Under the proposed conceptually based framework, stochastic models selected on an empirical basis cannot be considered as possible alternatives to the PIR-ARMA models. However, a comparison can be useful in order to provide a relative evaluation of the PIR-ARMA model performances. This was done by using as a competing model the popular PAR(1), which performs particularly well from a statistical point of view [e.g., *Noakes et al.*, 1985; *Jimenez et al.*, 1989]. The PAR(1) model was fitted to the logarithms of the runoff series (as is generally suggested for applications) after deseasonalizing by monthly mean subtraction. The comparison was based on the explained variance R^2 of the stochastic component, equal to $(1 - \sigma_{\varepsilon'}^2/\sigma^2)$ with $\sigma_{\varepsilon'}^2$ as the residual variance of the stochastic model and σ^2 as the variance of the runoff series. The probabilistic analysis of the residual was not taken into account. Excluding monthly means, variances, and residual variances, the PIR-ARMA(2, 2) model requires the estimation of four stochastic parameters, reduced to two for the PIR-ARMA(1, 1) applied to ephemeral streams, while the PAR(1) model requires the estimation of 12 autoregressive parameters. Results reported in Table 8 show that the proposed model performs significantly better than the PAR(1) model, with a 30% average improvement in R^2. No correction is made on the R^2 to account for the different number of parameters to estimate in the two models.

From Table 8 it can be noted that the stochastic component and the periodic variability of the mean of the proposed model explain 49–74% of the total variance of the series. The residual variance is partly due to the natural variability of the phenomenon and partly to an error term. Consequently, a measure of model performance in which all the parts of the model with conceptual meaning are taken into account is a total R^2, equal to $(1 - \sigma_\xi^2/\sigma^2)$ where σ_ξ^2 is the variance of the error component of the residual. In the cases examined the total R^2 was found to vary between 0.671 and 0.841.

Referring to these results, the proposed model seems to perform reasonably well both from the point of view of verification of the underlying conceptual hypotheses and from that of the statistical descriptive ability. However, the error variance σ_ξ^2 was found to be rather variable between the months, so that a future comparison with PARMA models is in order. Moreover, a thorough evaluation of the model's statistical efficiency, in terms of reproduction of the statistical characteristics of the observed series, can be achieved only by means of an extensive simulation study.

TABLE 8. Comparison of Explained Variance R^2 for PAR(1) and PIR-ARMA Models

	Station						
	1	2	3	4	5	6	7
R^2 PAR(1)	0.384	0.390	0.417	0.322	0.236	0.363	0.682
R^2 PIR-ARMA(2, 2)			0.551			0.496	0.745
R^2 PIR-ARMA(1, 1)	0.550	0.540		0.597	0.488		

7. Concluding Remarks

Based on a phenomenological interpretation of the runoff process, considered on different scales of aggregation, a conceptual framework is proposed for univariate linear stochastic model building of seasonal runoff. Streamflow is regarded as the sum of two groundwater components with different lag time and of a purely random term. Model residual is proportional to the conceptual system input, signifying effective rainfall, and is considered as a periodic independent stochastic process. In this conceptual scheme, the hypothesis of linearity is invoked only with regard to the properties of aquifers. Under these hypothesis the model of seasonal runoff is a constant parameter ARMA(2, 2) model with periodic independent residual (PIR-ARMA). For ephemeral streams, lacking the over-year groundwater component responsible for the long-term persistence effects, the proposed conceptual scheme leads to a PIR-ARMA(1, 1) model. In this framework, compatibility with the stochastic model of the aggregated runoff, typically considered on an annual basis, is guaranteed by the identification method itself.

More refined but more complex PARMA(2, 2) and PARMA(1, 1) models arise from the PIR-ARMA(2, 2) and PIR-ARMA(1, 1), respectively, when considering the seasonal variability of parameters.

The recognition of the lack of an over-year groundwater component may be difficult in environments where the climate shows no distinct wet and dry seasons, so the notion of the hydrologic year cannot be used. In these environments, however, one can take advantage of the a priori information on the net rainfall distribution used in the annual runoff model.

Explicit relationships are established between conceptual and stochastic parameters, as a first essential step toward the possibility of validating the conceptual hypotheses on which the model is built. Parameter estimation is performed on different aggregation scales by means of an iterative procedure.

The residual is considered to consist of a noise term, considered as a zero-mean Gaussian variable, and of a conceptual component, denoting effective rainfall and distributed according to a physically consistent compound Bessel probability function.

Model application to monthly runoff series in central southern Italy showed that the PIR-ARMA models are suitable for the interpretation of the process, at least to a first level of testing.

A comprehensive simulation study and a comparison with the corresponding PARMA models, useful to fully assess the statistical efficiency of the PIR-ARMA models, are left to future developments. Likewise, analysis of runoff data on a lower aggregation scale, e.g., daily, is forthcoming, with the added intention of a better understanding of the meaning of the recharge and the recession parameters of the subannual groundwater.

From the basis on which the conceptual framework is built, it is possible to set up a stochastic model for the runoff on a time scale in which the effects of the subsurface runoff component are also explicitly considered. This scale must be considerably larger than the surface runoff response time, so that it can be considered as an independent process. The subsequent conceptual model will be made up of three linear reservoirs in parallel plus a zero lag linear channel. For a time scale of the order of a few days, a stochastic model selected in the class of the shot noise models [see *Murrone et al.*, 1992] will better reflect the structure of the runoff process.

Future research made possible by the proposed framework may also be in the direction of the construction of a bivariate model in which total rainfall is a known input process and of runoff modeling in ungauged stations by means of regional analysis.

Appendix A: Recharge Recession Coefficient

The continuous time response of a linear reservoir to an input $r(\tau)$ is [e.g., *Chow et al.*, 1988]

$$Q(t) = Q_0 e^{-t/k} + \int_0^t \frac{1}{k} e^{-(t-\tau)/k} r(\tau)\, d\tau \qquad \text{(A1)}$$

where Q is the discharge, k is the reservoir storage coefficient and $Q_0 = Q(t = 0)$. Integration of (A1) over a unit time interval gives

$$D_1 = \int_0^1 Q(t) = \int_0^1 Q_0 e^{-t/k}\, dt + \int_0^1 \frac{1}{k} e^{-t/k} \int_0^t e^{\tau/k} r(\tau)\, d\tau\, dt \qquad \text{(A2)}$$

where D_1 is the outflow volume. Equation (A2) becomes, given the initial volume $V_0 = kQ_0$, by definition of linear reservoir,

$$D_1 = (1 - e^{-1/k})V_0 + \int_0^1 \frac{1}{k} e^{-t/k} \int_0^t e^{\tau/k}\, r(\tau)\, d\tau\, dt \qquad \text{(A3)}$$

Introducing the recharge recession coefficient,

$$r_k = 1 - \int_0^1 \frac{1}{k} e^{-t/k} \int_0^t \frac{1}{R} e^{\tau/k}\, r(\tau)\, d\tau\, dt \qquad \text{(A4)}$$

where $R = \int_0^1 r(\tau)\, d\tau$ is the recharge volume, (A3) can be expressed as

$$D_t = (1 - c_k)V_{t-1} + (1 - r_k)R_t \qquad \text{(A5)}$$

assuming, for the generic $(t - 1, t)$ unit time interval, D_t as the runoff, V_{t-1} as the storage at the beginning of the interval, R_t as the recharge in the interval and $c_k = e^{-1/k}$.

Assuming $r(\tau)$ as an impulse recharge in time T, with $0 \le T \le 1$, (A4) becomes [*Moss and Bryson*, 1974]

$$r_k = e^{-(1-T)/k} \qquad \text{(A6)}$$

while, in case of uniform recharge $r(\tau) = R/\Delta$ over a subinterval of duration $\Delta \le 1$ and initial time T_i, the expression is

$$r_k = \frac{k}{\Delta}(1 - e^{-\Delta/k})[1 - (1 - e^{(1-T_i-\Delta)/k})] \qquad \text{(A7)}$$

For Δ approaching zero, (A7) reduces to (A6) while for $T_i = 0$ and $\Delta = 1$ (A7) becomes

$$r_k = k(1 - e^{-1/k}) \tag{A8}$$

For $\Delta = 0$ and $T = 0$, which corresponds to impulse recharge at the beginning of the interval, (A7) gives

$$r_k = e^{-1/k} = c_k \tag{A9}$$

Appendix B: Stochastic Representation of Conceptual Model Forms

B1. *Linear Reservoir*

The simplest conceptual model, made up of a single linear reservoir, can be assumed as the model of the runoff of a spring. The outlet D_t of the reservoir in a unit time interval $(t - 1, t)$ is expressed in (A5) as a function of the initial storage V_{t-1} and of the recharge R_t in the interval. The mass balance equation for the groundwater storage is

$$V_t = c_k V_{t-1} + r_k R_t \tag{B1}$$

where, again, $c_k = e^{-1/k}$. Substituting into (B1) the expression of V_{t-1} resulting from (A5), one obtains

$$V_t = \frac{c_k D_t}{1 - c_k} - \frac{c_k(1 - r_k)R_t}{1 - c_k} + r_k R_t \tag{B2}$$

Considering in (B2) V_{t-1}, D_{t-1}, and R_{t-1} instead of V_t, D_t, and R_t, which does not affect generality, and resubstituting V_{t-1} as obtained from (A5), gives

$$\frac{D_t}{1 - c_k} - \frac{(1 - r_k)R_t}{1 - c_k} = \frac{c_k D_{t-1}}{1 - c_k} - \frac{c_k(1 - r_k)R_{t-1}}{1 - c_k} + r_k R_{t-1} \tag{B3}$$

and, rearranging,

$$D_t - c_k D_{t-1} = (1 - r_k)R_t - (c_k - r_k)R_{t-1} \tag{B4}$$

If R_t is considered as an independent stochastic process, (B4) represents an ARMA(1, 1) process. Introducing the variables $d_t = D_t - E[D_t]$ and $\varepsilon_t = (1 - r_k)(R_t - E[R_t])$, with mean zero, this model is written, in Box-Jenkins notation,

$$d_t - \Phi d_{t-1} = \varepsilon_t - \Theta \varepsilon_{t-1} \tag{B5}$$

where $\Phi = c_k$ and $\Theta = (c_k - r_k)/(1 - r_k)$ are the autoregressive and moving average parameters, respectively. Under these latter conditions it can be shown that $E[D_t] = E[R_t]$.

If the within-period distribution of recharge is predetermined, as assumed in this paper, coefficients Φ and Θ of (B5) are not independent. In the limit case of impulse recharge at the beginning of the interval, r_k equals c_k (see Appendix A) and (B4) simplifies to

$$D_t - c_k D_{t-1} = (1 - c_k)R_t \tag{B6}$$

which is equivalent to an AR(1) process.

On a seasonal basis, if the storage constant is considered as varying with the seasons, (B6) becomes a PAR(1) model, as shown by *Salas and Obeysekera* [1992] with reference to a similar conceptual scheme, that can be written as

$$d_{n,m} - \Phi_m d_{n,m-1} = \varepsilon_{n,m} \tag{B7}$$

in which n denotes the year and m denotes the season. Coefficients Φ_m relate to the seasonal $c_k(m)$ as [*Salas and Obeysekera*, 1992]

$$\Phi_m = c_k(m - 1)[1 - c_k(m)]/[1 - c_k(m - 1)] \tag{B8}$$

where $c_k(m) = e^{-1/k(m)}$.

B2. *Linear Reservoir Plus Zero Lag Linear Channel*

This conceptual system was analyzed by *Salas et al.* [1981] with reference to the annual streamflows. It is revisited here by considering the input as the effective rainfall, instead of total precipitation, and by taking into account the effects due to the distribution of the within-period recharge.

With reference to Figure 3, the groundwater recharge R_t in the interval t is considered as the fraction aI_t of the effective rainfall I_t, while $(1 - a)I_t$ is the direct runoff. Including the direct runoff into the mass balance equations (A5) and (B1) and rearranging, an ARMA(1, 1) process

$$d_t - \Phi d_{t-1} = \varepsilon_t - \Theta \varepsilon_{t-1} \tag{B9}$$

is obtained for the runoff, as shown by *Salas et al.*, [1981]. In the hypotheses made here, (B9) holds with positions

$$\Phi = c_k = e^{-1/k} \qquad \Theta = \frac{c_k - ar_k}{1 - ar_k} \qquad \varepsilon_t = (1 - ar_k)i_t \tag{B10}$$

and, correspondingly,

$$a = \frac{\Phi - \Theta}{r_k(1 - \Theta)} \qquad k = \frac{1}{\ln \Phi} \qquad i_t = \frac{\varepsilon_t}{1 - ar_k} \tag{B11}$$

where $i_t = I_t - E[I_t]$, and ε_t is the zero-mean residual. If σ_ε^2 is the residual variance, the variance of the net rainfall process is

$$\sigma_I^2 = \frac{\sigma_\varepsilon^2}{(1 - ar_k)^2} \tag{B12}$$

Due to the conditions $0 \le a \le 1$ and $k \ge 0$ and to (B10) the parameter space of the process (B9), given by *Salas et al.* [1981], is restricted with respect to that of a general ARMA(1, 1) model. Process (B9) thus belongs to the class of restricted ARMA processes.

With reference to seasonal runoff, *Salas and Obeysekera* [1992] showed that, considering parameters a and k as varying with the seasons, the scheme depicted in Figure 3 leads to a PARMA(1, 1) process. If only one parameter is assumed nonseasonal, a PARMA(1, 1) process with nonseasonal AR or MA parameter is obtained. In all cases, retaining the results from the authors cited, explicit relationships similar to those given in (B10) and (B11) can be obtained between a_m, k_m and Φ_m, Θ_m.

B3. *Two Parallel Linear Reservoirs Plus Zero Lag Linear Channel*

This case is represented in Figure 2 where, as an immediate extension of the previous scheme, b, q, W, and r_q have the same meaning as a, k, V, and r_k, and, accordingly, $c_q = e^{-1/q}$. For this conceptual system, runoff D_t is given by

$$D_t = (1 - c_k)V_{t-1} + (1 - c_q)W_{t-1} + a(1 - r_k)I_t + b(1 - r_q)I_t + (1 - a - b)I_t \quad \text{(B13)}$$

The volume balance equations for the groundwater storage are

$$V_t = c_k V_{t-1} + ar_k I_t \quad \text{(B14)}$$

$$W_t = c_q W_{t-1} + br_q I_t \quad \text{(B15)}$$

Putting in (B13) the expressions of W_{t-1} and V_{t-1} obtained from (B14) and (B15) and rearranging results in one equation in D_t, D_{t-1}, D_{t-2}, I_t, I_{t-1}, I_{t-2}:

$$D_t - (c_k + c_q)D_{t-1} + (c_k c_q)D_{t-2} = (1 - ar_k - br_q)I_t - [c_k + c_q - ar_k(1 + c_q) - br_q(1 + c_k)]I_{t-1} - (ar_k c_q + br_q c_k - c_k c_q)I_{t-2} \quad \text{(B16)}$$

With reference to the zero-mean variables d_t and i_t, if $\varepsilon_t = (1 - ar_k - br_q)i_t$ is an independent stochastic process, the above representation is equivalent to an ARMA(2, 2) process,

$$d_t - \Phi_1 d_{t-1} - \Phi_2 d_{t-2} = \varepsilon_t - \Theta_1 \varepsilon_{t-1} - \Theta_2 \varepsilon_{t-2} \quad \text{(B17)}$$

Following the reasoning by *Salas and Obeysekera* [1992] and letting a, b, q, and k vary with the seasons, a PARMA(2, 2) process arises.

Relations between conceptual and stochastic parameters for the ARMA(2, 2) are

$$\Phi_1 = c_k + c_q \quad \text{(B18)}$$

$$\Phi_2 = -c_k c_q \quad \text{(B19)}$$

$$\Theta_1 = \frac{c_k + c_q - ar_k(1 + c_q) - br_q(1 + c_k)}{1 - ar_k - br_q} \quad \text{(B20)}$$

$$\Theta_2 = \frac{ar_k c_q + br_q c_k - c_k c_q}{1 - ar_k - br_q} \quad \text{(B21)}$$

$$c_k = \frac{\Phi_1 + (\Phi_1^2 + 4\Phi_2)^{1/2}}{2} \quad \text{(B22)}$$

$$c_q = \frac{\Phi_1 - (\Phi_1^2 + 4\Phi_2)^{1/2}}{2} \quad \text{(B23)}$$

$$a = \frac{M - N - Mbr_q}{Mr_k} \quad \text{(B24)}$$

$$b = [-(\Theta_1 - \Theta_2)N + (\Phi_1 - \Phi_2)M + (1 + 2c_q)(N - M)] \cdot [2M(c_k - c_q)r_q]^{-1} \quad \text{(B25)}$$

where $N = (1 - \Phi_1 - \Phi_2)$ and $M = (1 - \Theta_1 - \Theta_2)$. Relations (B18)–(B21) or (B22)–(B25) ensure equivalence of the means of D_t and I_t [*Claps*, 1990]. The net rainfall variance can be obtained from the residual variance σ_ε^2 through the relation $\sigma_I^2 = \sigma_\varepsilon^2/(1 - ar_k - br_q)^2$.

In the general ARMA(2, 2) model, stationarity and invertibility conditions provide admissible regions for the AR and MA parameters. Limiting the analysis to the autoregressive

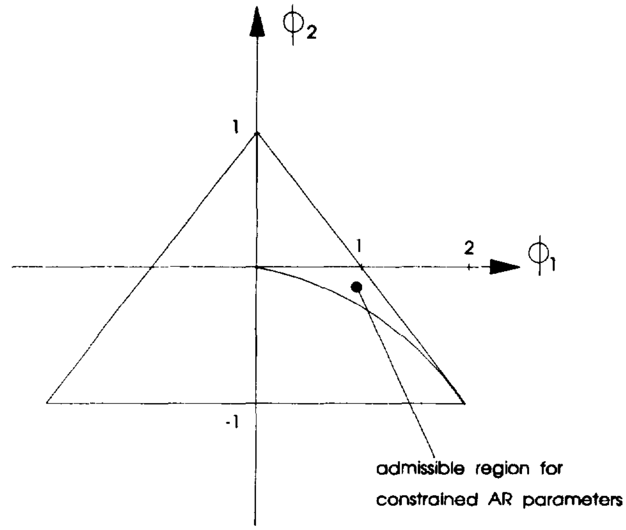

Fig. 8. Admissible region for autoregressive parameters of the restricted PIR-ARMA(2, 2) model.

parameters Φ_1 and Φ_2, a triangular admissible region results [*Box and Jenkins*, 1970, paragraph 3.2.4.] from the stationarity conditions $-1 < \Phi_2 < 1$; $\Phi_2 - \Phi_1 < 1$; and $\Phi_2 + \Phi_1 < 1$. Given that $0 < c_q < 1$ and $0 < c_k < 1$, (B22)–(B23) provide conceptual constraints for Φ_1 and Φ_2:

$$\Phi_1^2 + 4\Phi_2 > 0 \qquad \Phi_2 < 0 \qquad \Phi_1 > 0 \quad \text{(B26)}$$

which restrict the admissible space of the two parameters to the region indicated in Figure 8. Again, the conceptual representation determines a restricted ARMA process. It is interesting to remark that if one of the storage coefficients, for instance k, is large, c_k approaches unity and this leads to $\Phi_1 + \Phi_2 \approx 1$, which is close to the nonstationarity condition $\Phi_1 + \Phi_2 = 1$. In other words, the outlet of a system with a very slow response (such as a reservoir with a recession coefficient $c_k \approx 1$) may be confused with a trend.

Acknowledgments. This work was supported in part by funds granted by the Italian National Research Council–Group for Prevention of Hydrogeological Disasters, grant 90.01431.42. Comments and suggestions by three anonymous reviewers are acknowledged and appreciated.

REFERENCES

Acton, F. S., *Numerical Methods That Work*, p. 252, Harper Collins, New York, 1970.

Benjamin, J. R., and C. A. Cornell, *Probability, Statistics and Decision for Civil Engineers*, p. 310, McGraw-Hill, New York, 1970.

Box, G. E., and G. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, Calif., 1970.

Chow, V. T., D. R. Maidment, and L. W. Mays, *Handbook of Applied Hydrology*, p. 208, McGraw-Hill, New York, 1988.

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

Claps, P., Sulla validazione di un modello dei deflussi a base concettuale, paper presented at XXIII Convegno di Idraulica e Costruzioni Idrauliche, Dept. of Civ. Eng., Univ. of Florence, Florence, Italy, 1992.

Delleur, J. W., P. C. Tao, and M. L. Kavvas, An evaluation of the practicality and the complexity of some rainfall and runoff time series models, *Water Resour. Res.*, *12*(5), 953–970, 1976.

Fernandez, B., and J. D. Salas, Periodic gamma-autoregressive processes for operational hydrology, *Water Resour. Res.*, *22*(10), 1385–1396, 1986.

Fernandez, B., and J. D. Salas, Gamma-autoregressive models for stream-flow simulation, *J. Hydraul. Eng.*, *116*(HY11), 1403–1414, 1990.

Fiering, M. B., *Streamflow Synthesis*, pp. 69–70, Harvard University Press, Cambridge, Mass., 1967.

Hirsch, R. M., Synthetic hydrology and water supply reliability, *Water Resour. Res.*, *15*(6), 1603–1615, 1979.

Institute of Hydrology, Low flow studies, research report, Wallingford, Oxon, England, 1980.

Jimenez, C., A. I. McLeod, and K. W. Hipel, Kalman filter estimation for periodic autoregressive-moving average models, *Stochastic Hydrol. Hydraul.*, *3*, 227–240, 1989.

Kavvas, M. L., L. J. Cote, and J. W. Delleur, Time resolution of the hydrologic time series, *J. Hydrol.*, *32*, 347–361, 1977.

Klemeš, V., Physically based stochastic hydrologic analysis, *Advances in Hydroscience*, vol. 11, edited by V. T. Chow, pp. 285–352, Academic, San Diego, Calif., 1978.

Klemeš, V., Conceptualization and scale in hydrology, *J. Hydrol.*, *65*, 1–23, 1983.

Koch, R. W., A stochastic streamflow model based on physical principles, *Water Resour. Res.*, *21*(4), 545–553, 1985.

Kottegoda, N. T., *Stochastic Water Resources Technology*, p. 141, MacMillan, London, 1980.

Lloyd, E. H., Stochastic reservoir theory, in *Advances in Hydroscience*, vol. 4, edited by V. T. Chow, pp. 281–339, Academic, San Diego, Calif., 1967.

Moss, M. E., and M. C. Bryson, Autocorrelation structure of monthly streamflows, *Water Resour. Res.*, *10*(4), 737–744, 1974.

Murrone, F., F. Rossi, and P. Claps, 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, 1992.

Noakes, D. J., A. I. McLeod, and K. W. Hipel, Forecasting monthly riverflow time series, *Int. J. Forecast.*, *1*, 179–190, 1985.

Obeysekera, J. T. B., and J. D. Salas, Modeling of aggregated hydrologic time series, *J. Hydrol.*, *86*, 197–219, 1986.

O'Connell, P. E., A simple stochastic modeling of Hurst's law, paper presented at International Symposium on Mathematical Models in Hydrology, Int. Assoc. of Hydrol. Sci., Warsaw, 1971.

O'Connell, P. E., D. G. Thorn, and R. S. MacKenzie, Multisite synthetic streamflow modeling in a semi-arid region (abstract), *Ann. Geophys.*, *Spec. Issue*, *9*, C512, 1991.

Pierce, D. A., Least squares estimation of the regression model with autoregressive–moving average errors, *Biometrika*, *58*, 299–312, 1971.

Rossi, F., and G. Silvagni, Analysis of annual runoff series, paper presented at Third Symposium on Stochastic Hydraulics, Int. Assoc. for Hydraul. Res., Tokyo, 1980.

Salas, J. D., and J. T. B. Obeysekera, Conceptual basis of seasonal streamflows time series models, *J. Hydraul. Eng.*, *118*(HY8), 1186–1194, 1992.

Salas, J. D., and R. A. Smith, Physical basis of stochastic models of annual flows, *Water Resour. Res.*, *17*(2), 428–430, 1981.

Salas, J. D., J. W. Delleur, V. Yevjevich, and W. L. Lane, *Applied Modeling of Hydrologic Time Series*, Water Resources Publications, Littleton, Colo., 1980.

Salas, J. D., J. T. B. Obeysekera, and R. A. Smith, Identification of streamflow stochastic models, *J. Hydraul. Eng.*, *107*(HY7), 853–866, 1981.

Salas, J. D., D. C. Boes, and R. A. Smith, Estimation of ARMA models with seasonal parameters, *Water Resour. Res.*, *18*(4), 1006–1010, 1982.

Sim, C. H., A mixed gamma ARMA(1, 1) model for river flow time series, *Water Resour. Res.*, *23*(1), 32–36, 1987.

Spolia, S., and S. Chander, Modeling of surface runoff systems by an ARMA model, *J. Hydrol.*, *22*, 317–322, 1974.

Tao, P. C., and J. W. Delleur, Seasonal and nonseasonal ARMA models in hydrology, *J. Hydraul. Eng.*, *102*(HY10), 1541–1559, 1976.

Thomas, H., and M. B. Fiering, Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, in *Design of Water Resource Systems*, edited by A. Mass et al., pp. 459–463, Harvard University Press, Cambridge, Mass., 1962.

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

Vecchia, A. V., J. T. B. Obeysekera, J. D. Salas, and D. C. Boes, Aggregation and estimation of low-order periodic ARMA models, *Water Resour. Res.*, *19*(5), 1297–1306, 1983.

Yevjevich, V., and A. K. Karplus, Area-time structure of the monthly precipitation process, *Hydrol. Pap. no. 64*, Colo. State Univ., Fort Collins, 1973.

P. Claps, Dipartimento di Ingegneria e Fisica dell'Ambiente, Universita' della Basilicata, Via della Tecnica, 3, Potenza, I-85100 Italy.

F. Rossi, Istituto di Ingegneria Civile, Universita' di Salerno, Via Ponte Don Melillo, Fisciano, I-84084 Italy.

C. Vitale, Istituto di Statistica, Universita' di Salerno, Via Ponte Don Melillo, Fisciano, I-84084 Italy.

(Received June 1, 1992;
revised February 15, 1993;
accepted March 3, 1993.)