Journal of Hydrology, 47 (1980) 235—249
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

[3]

A STOCHASTIC MODEL FOR DAILY STREAMFLOW*

JERSON KELMAN

Centro de Pesquisas de Energia Elétrica (CEPEL), Grupo ELETROBRÁS, 20.000 Rio de Janeiro R.J. (Brazil)

(Received November 7, 1979; accepted for publication November 14, 1979)

ABSTRACT

Kelman, J., 1980. A stochastic model for daily streamflow. J. Hydrol., 47: 235—249.

A model for the description and generation of samples for daily streamflow is developed. The basic assumption is that the rising and falling limbs of hydrographs are to be modeled separately due to the fact that they translate different physical processes. The rising limb is mostly due to factors external to the watershed. It can be modeled similarly for precipitation. On the other hand, the falling limb is mostly governed by the emptying-water from the watershed. The model assumes the conceptual representation of the watershed as two linear reservoirs. Any sequence of recession discharge is then a stochastic output from these two reservoirs. The model was tested for a case study and results are satisfactory.

INTRODUCTION

In this paper a new approach for the stochastic modeling of daily streamflow is introduced. It should be pointed out at the outset that no universality is claimed for the model to be described. In fact, the attempt to develop a general model may have been the reason for the failures of previous efforts to model daily flows. It is hardly conceivable that a simple scheme could model equally well the streams fed by snow melt and streams draining a tropical catchment, to give only an example. The model to be described here refers to catchments for which the direct runoff plays an important role in the composition of the total flow. Nevertheless, each catchment that qualifies for such a description must be studied on a case-by-case basis.

A dual approach is used, in the sense that the positive and the negative first derivatives of the streamflow process can be modeled by two alternating intermittent stochastic processes.

In this paper the conceptual framework of the model is set up first, and then a technique developed with the help of the case study of the Powell

* Presented at the International Seminar held at the Simon Bolivar University, Caracas, Venezuela, March 20—23, 1978.

0022-1694/80/0000—0000/$02.25

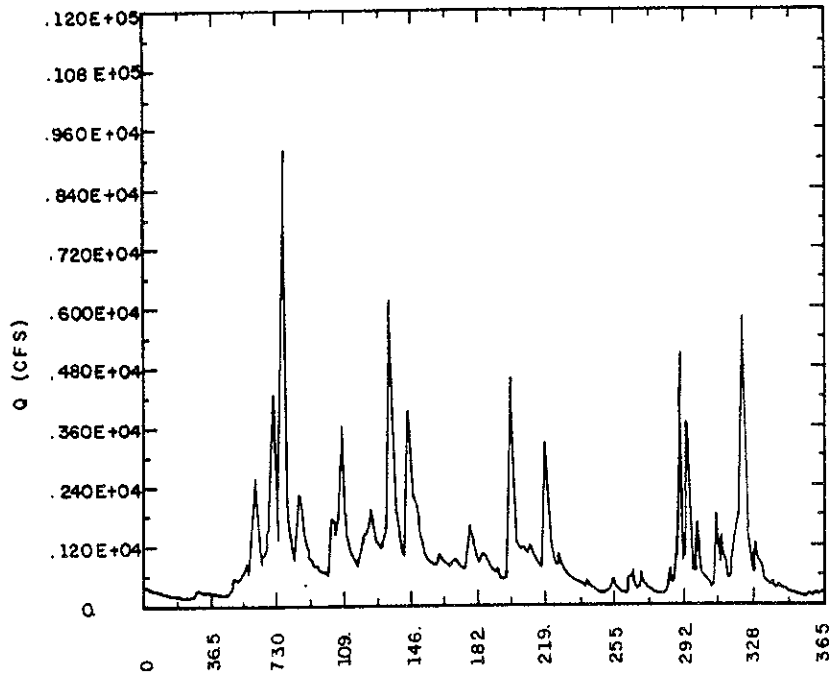

Fig. 1. Daily flow hydrograph of the Powell River for the year of 1921.

River, near Arthur, Tennessee. This river is described by Quimpo (1967) as having an accurate record from 1921 to 1960. The outlet drains an area of 1767 km^2 and is located at 36°32′N latitude and 83°38′W longitude. The mean daily flow is 32.0 $m^3\ s^{-1}$ (1116 cfs). For a better insight into the type of streamflow studied, Fig. 1 shows the hydrograph for the year of 1921, which is a fairly typical hydrograph.

THE CONCEPTUAL FRAMEWORK

The runoff at the outlet of a watershed is considered to be the sum of three components, namely:

$$q(t) = q_1(t) + q_2(t) + q_3(t) \tag{1}$$

Conceptually, these components have different physical characteristics, such as in the case of underground flow and surface flow. Therefore, it is expected that these components will exhibit also different stochastic characteristics. Fig. 2 gives an illustration of how the runoff formation is conceived in this study. $q_1(t)$ is the outflow from reservoir *1*, which simulates the groundwater storage; $q_2(t)$ is the outflow from reservoir *2*, which simulates the lumped storages of: (1) surface detention storage; (2) bank storage; and (3) channel storage; and $q_3(t)$ is the direct runoff, which is composed mainly

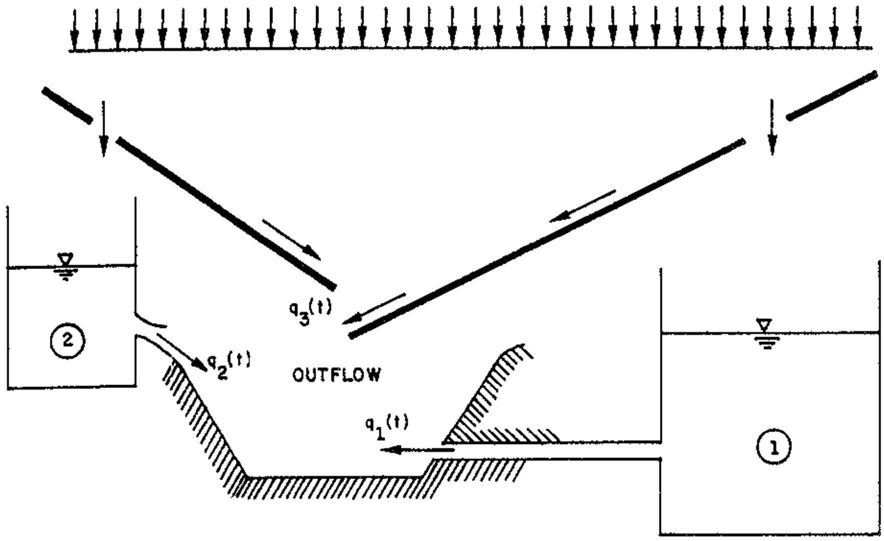

Fig. 2. Schematic representation of components in streamflow.

of the surface runoff and the precipitation over the stream surfaces. Like daily precipitation, daily direct runoff is an intermittent process.

There is no doubt that representing the retention capacity of a watershed by only two reservoirs is an oversimplification of the real situation. However, it is better than assuming the homogeneity of the whole process, as is usually done.

Ideally $q_3(t)$ depends mostly on factors external to a watershed. It can be thought of as the immediate response of a catchment to the precipitation events. It is shown by Kelman (1977) that precipitation can be successfully modeled by the stochastic process as represented here in Fig. 3. $\{\xi(t)\}$ are independent random variables with standard normal distribution, $\{Z(t)\}$ being:

$$Z(t) = \mu(\tau) + \frac{\sigma(\tau)}{\sigma(\tau-1)} \rho(\tau)[Z(t-1) - \mu(\tau-1)] + \sigma(\tau)[1-\rho^2(\tau)]^{1/2}\xi(t)$$

$\{Y(t)\}$ being:

$$Y(t) = Z(t)I_{(0,\infty)}[Z(t)]$$

and $\{X(t)\}$ being:

$$X(t) = Y(t)^{1/\alpha(\tau)}$$

$I_{(0,\infty)}(\cdot)$ is the indicator function; and $\mu(\tau)$, $\sigma(\tau)$, $\rho(\tau)$ and $\alpha(\tau)$ are periodic functions to be estimated from data, $\tau = 1, 2, \ldots, w$; w is the number of seasons in which the year is divided.

It seems reasonable to model $\{q_3(t)\}$ in the same way as the precipitation process, $\{X(t)\}$. The estimation procedure is described in the reference mentioned. However, no realization of the $q_3(t)$ process is available. The fact is

Fig. 3. Representation of the intermittent model.

that only the series of the total discharge, $q(t)$, is measured. There is no way of splitting $q(t)$ into exactly its three components, $q_1(t)$, $q_2(t)$ and $q_3(t)$. A somewhat arbitrary assumption is then necessary. It is possible that some modification would lead to a more realistic representation of the phenomena. The assumption is:

$$q_3(t) = \max[0, q(t) - q(t-1)] \tag{2}$$

Eq. 2 says that the direct runoff is either zero or it is equal to the positive increment of total discharge. In fact, if $q_3(t) > 0$ one could expect that the reservoirs are partially replenished on the day t, and therefore it is likely that:

$$q_1(t+1) + q_2(t+1) > q_1(t) + q_2(t)$$

or

$$[q_1(t+1) - q_1(t)] + [q_2(t+1) - q_2(t)] > 0 \tag{3}$$

Eq. 2 simply says that the above positive quantity is equal to $q_3(t)$, or that:

$$q_3(t) = [q_1(t+1) - q_1(t)] + [q_2(t+1) - q_2(t)] \quad \text{for } q_3(t) > 0 \tag{4}$$

From eqs. 1 and 4 one can see that, whenever $q_3(t) > 0$:

$$q_1(t) + q_2(t) = q(t-1) \tag{5}$$

Hence, any rising limb of the hydrograph, say from day t_0 to day t_f can be obtained if the value of $q(t_0)$ as well as of the succession $q_3(t_0), \ldots, q_3(t_f)$ are known. In order to have a rising limb all the values in the succession $q_3(t_0), \ldots, q_3(t_f)$ should be positive. How to resolve the problem of the falling limbs of the hydrographs will be shown later. Next the process $q_3(t)$ for the Powell River is studied in more detail.

DAILY STREAMFLOW MODEL OF THE POWELL RIVER

Direct runoff

The observed streamflow data of the Powell River were processed following eq. 2 to produce the time series $q_3(t)$. (The same symbol is used for convenience, either for the stochastic process or for the corresponding time series.) The data were divided into 26 seasons each 14 days long, adding up to 364 days. For each season the parameters μ_τ, σ_τ, ρ_τ and α_τ were esti-

TABLE I

Results for goodness-of-fit statistics for the 26 seasons of daily flows of the Powell River

Period (τ)	From To	$-\mu_\tau$	σ_τ	ρ_τ	α_τ	X^2 (d.f.)
1	1 Oct.—14 Oct.	3.6296	6.9689	0.2413	0.4312	8.01(1)†
2	15 Oct.—28 Oct,	3.2556	6.8980	0.7113	0.4098	10.57(4)*
3	29 Oct.—11 Nov.	3.3940	7.1673	0.6098	0.3883	8.76(3)*
4	12 Nov.—25 Nov.	4.3443	9.8158	0.5737	0.3848	5.33(6)
5	26 Nov.—9 Dec.	4.2393	13.5182	0.6620	0.4123	14.74(9)
6	10 Dec.—23 Dec.	11.2918	22.7644	0.6352	0.4500	6.19(11)
7	24 Dec.—6 Jan.	13.5328	37.9861	0.6001	0.5197	20.54(13)
8	7 Jan.—20 Jan.	15.2014	37.4007	0.6905	0.5067	18.57(14)
9	21 Jan.—3 Feb.	19.3112	43.0868	0.5325	0.5038	12.78(14)
10	4 Feb.—17 Feb.	28.4365	67.6299	0.5148	0.5626	17.57(16)
11	18 Feb.—3 Mar.	36.3807	68.4997	0.6526	0.5623	13.55(14)
12	4 Mar.—17 Mar.	43.4370	96.6109	0.6247	0.6158	14.82(15)
13	18 Mar.—31 Mar.	29.4476	57.0287	0.5679	0.5421	10.52(14)
14	1 Apr.—14 Apr.	39.4548	60.2830	0.5625	0.5800	12.72(9)
15	15 Apr.—28 Apr.	35.1393	52.9360	0.5672	0.5483	14.30(10)
16	29 Apr.—12 May	21.0269	35.0319	0.6524	0.5284	12.63(9)
17	13 May—26 May	19.7310	27.8554	0.5145	0.4901	7.00(7)
18	27 May—9 Jun.	14.4111	20.8414	0.5221	0.4903	14.14(8)
19	10 Jun.—23 Jun.	12.2671	26.4863	0.4879	0.5550	14.32(9)
20	24 Jun.—7 Jul.	7.0326	14.4152	0.3834	0.4371	18.79(10)*
21	8 Jul.—21 Jul.	8.0993	18.3048	0.4296	0.4800	14.91(10)
22	22 Jul.—4 Aug.	8.9538	17.0723	0.2473	0.5132	9.86(6)
23	5 Aug.—18 Aug.	6.7849	15.0933	0.4047	0.4694	24.88(9)†
24	19 Aug.—1 Sep.	10.3664	15.5428	0.5306	0.4953	8.70(5)
25	2 Sep.—15 Sep.	7.2767	11.7339	0.3901	0.4904	9.17(3)*
26	16 Sep.—29 Sep.	7.6774	10.2326	0.3397	0.4761	7.30(2)*

* The test is rejected at the 5% significance level.
† The test is rejected at the 1% significance level.

mated. The χ^2 goodness-of-fit statistic was computed for each of the 26 marginal distributions (one for each season). The results are shown in Table I. The seasons marked with an asterisk are those which have the goodness-of-fit of marginal distribution rejected at the 5% significance level. Those marked with a dagger are the cases with rejection also at the 1% significance level. The number of rejections was high: seven and two cases at the 5% and 1% levels, respectively. August—November, roughly the autumn, seems to be the time of the year for which the positive increments were invalidly fitted by the model. Later it will be shown that this problem is serious enough to impede reliable working of the model for this specific season. However, it is likely that one will be more concerned with studying spring and summer, rather than the autumn, due to the timing of the floods. Data from Table I were used to produce the periodic functions that represent the time variation

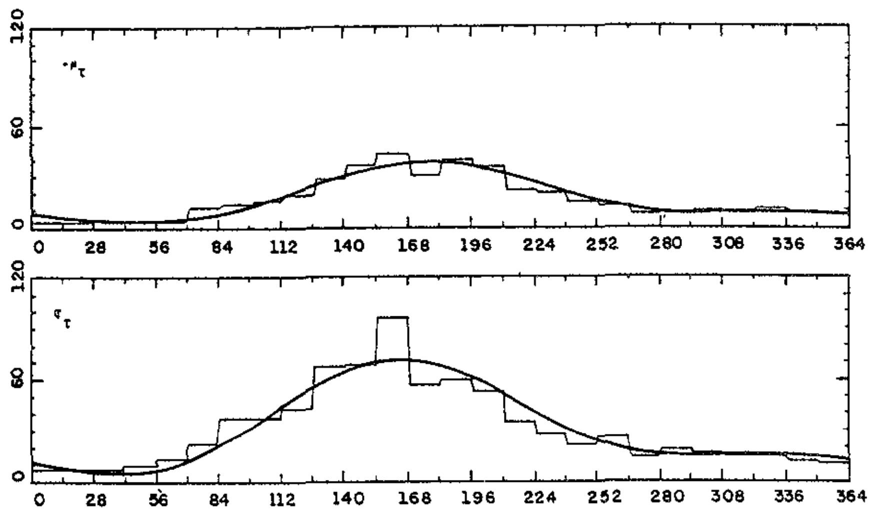

Fig. 4. The periodic μ_τ and σ_τ for daily values of the Powell River.

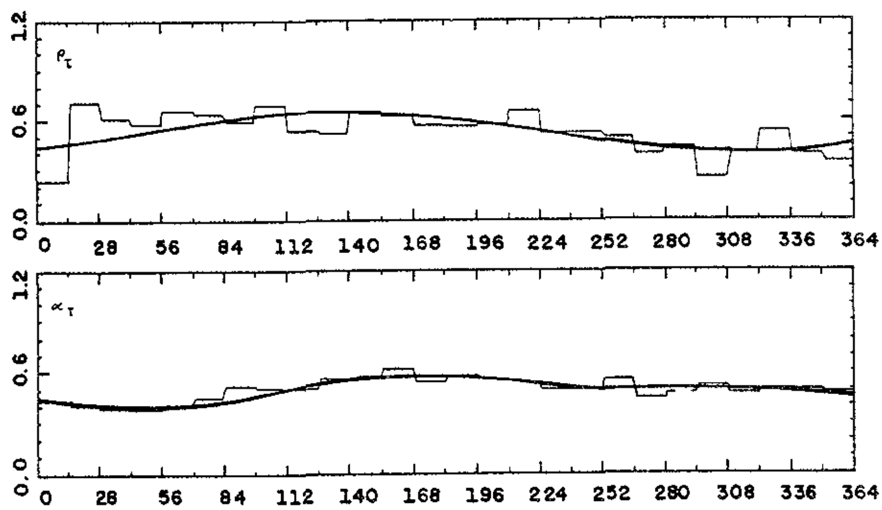

Fig. 5. The periodic ρ_τ and α_τ for daily values of the Powell River.

of each one of the parameters. These are seen in Figs. 4 and 5 for the Powell River (the step function and its smoothed version).

Outflow from the watershed storage

It was seen previously that, according to the model proposed, any falling limb of the hydrograph is the result of emptying the two reservoirs. The hydrograph values decrease only when $q_3(t) = 0$. Hence, the hydrograph recession curve is nearly independent of the characteristics of storm which causes the hydrograph rise. Only the states of the reservoirs, as well as their operating rules are relevant for this analysis. Description of reservoirs is then needed. It is assumed that both reservoirs are linear, meaning that the output $q_i(t)$, $i = 1$ and 2, is proportional to the storage $S_i(t)$. Or:

$$q_i(t) = K_i S_i(t), \qquad (i = 1 \text{ and } 2) \tag{6}$$

During the recession part of the hydrograph the input to reservoirs is zero with the continuity equation expressed in a simple form as:

$$q_i(t) = -\mathrm{d}S_i(t)/\mathrm{d}t, \qquad (i = 1 \text{ and } 2) \tag{7}$$

If eq. 6 is differentiated with respect to time t and then eq. 6 used, we obtain:

$$\mathrm{d}q_i(t)/\mathrm{d}t = -K_i q_i(t), \qquad (i = 1 \text{ and } 2)$$

or

$$\mathrm{d}q_i(t)/q_i(t) = -K_i \mathrm{d}t, \qquad (i = 1 \text{ and } 2) \tag{8}$$

Integrating of eq. 8 between 0 and t yields:

$$\ln[q_i(t)/q_i(0)] = -K_i t, \qquad (i = 1 \text{ and } 2)$$

or

$$q_i(t) = q_i(0) \exp(-K_i t) \tag{9}$$

Eq. 9 is the well-known exponential recession curve. It is obvious that the outflow discharge from the ith linear reservoir, during a recession period, depends only on the initial discharge $q_i(0)$ and on the reservoir characteristic K_i. Therefore any recession curve can be expressed by:

$$q(t) = q_1(0) \exp(-K_1 t) + q_2(0) \exp(-K_2 t), \qquad t \leqslant \Upsilon \tag{10}$$

where for convenience $t = 0$ indicates the beginning of the recession curve, and Υ is the length of the recession considered. For:

$$W = q_1(0)/q(0) \tag{11}$$

eq. 10 may be rewritten as:

$$q(t) = q(0)[W \exp(-K_1 t) + (1 - W) \exp(-K_2 t)] \tag{12}$$

or for $\gamma_1 = \exp(-K_1)$ and $\gamma_2 = \exp(-K_2)$:

$$q(t) = q(0)[W\gamma_1^t + (1 - W)\gamma_2^t] \tag{13}$$

K_1 and K_2 are constants that must be estimated. On the other hand, W indicates how the maintenance of the hydrograph is split between the two reservoirs, after a storm has occurred. Since the initial states of reservoirs are expected to vary from one recession curve to another, W cannot be conceived as a constant; rather its visualization as a random variable seems feasible. Therefore, in order to use eq. 12 in the generation of new samples, not only the values of K_1 and K_2 must be known but also the probability distribution of W, with $q(0)$ always known.

It is reasonable to estimate K_1 and K_2 in such a way that the theoretical recession curves will resemble the observed recession curves. In the more specific terms, the estimation of K_1 and K_2 should be taken in the framework of the following optimization problem:

$$\min_{K_1, K_2} \sum_{r=1}^{n} \sum_{t=1}^{\Upsilon(r)} (q'(t,r) - q'(0,r)[w(r) \exp(-K_1 t) + \{1 - w(r)\} \exp(-K_2 t)])^2 \tag{14}$$

where $\Upsilon(r)$ is the length of the rth recession curve; n is the number of recession curves in the historic data; $q'(t,r)$ is the observed discharge on the tth day of the rth recession curve; and $w(r)$ is the outcome of the random variable W, associated with the rth recession curve.

For any pair (K_1, K_2) the objective function of eq. 14 can only be evaluated if the outcomes $w(r)$, $r = 1, 2, \ldots, n$ are known. Again, it is reasonable to assume that each $w(r)$ is such that the differences between the rth theoretical and the observed recession curve values are minimized. By this reasoning, each $w(r)$ can be found by solving the equation:

$$\frac{\delta}{\delta w(r)} \left[\sum_{t=1}^{\Upsilon(r)} (q'(t,r) - q'(0,r)[w(r) \exp(-K_1 t) + \{1 - w(r)\} \exp(-K_2 t)]) \right] = 0 \tag{15}$$

or

$$w(r) = \frac{\Sigma_{t=1}^{\Upsilon(r)} q'(t,r) - q(0,r) \Sigma_{t=1}^{\Upsilon(r)} \exp(-K_2 t)}{q(0,r) \Sigma_{t=1}^{\Upsilon(r)} [\exp(-K_1 t) - \exp(-K_2 t)]} \tag{16}$$

Several numerical algorithms are available for solving the optimization problem defined by eq. 14. Among them is the Rosen algorithm as a quite convenient one. It is a mountain climbing type of technique, based on the gradient projection method. A detailed description of the algorithm is given by Kuester and Mize (1973). Here it is sufficient to say that the only requirements for the algorithm are: (1) objective function, which is given by eq. 14; (2) first derivatives of the objective function, which can be obtained by a proper use of eq. 14; and linear constraints, given as:

$$\left. \begin{array}{lll} 0 < \gamma_1 < 1 & \text{or} & 0 < K_1 < \infty \\ \text{and} & & \\ 0 < \gamma_2 < 1 & \text{or} & 0 < K_2 < \infty \end{array} \right| \tag{17}$$

Attention is called to the fact that each time the value of the pair (K_1, K_2) is changed, the observations $w(r)$ are re-assessed by using eq. 16. Also, one should expect from the way the conceptual model was set that $\gamma_1 > \gamma_2$ (or $K_1 < K_2$), although this does not constitute a constraint. For the Powell River daily flow data, the application of the algorithm yields:

$$\left.\begin{aligned} \gamma_1 &= 0.8971 \rightarrow K_1 = 0.1086/\text{day} \rightarrow 1/K_1 = 9.2091 \text{ days} \\ \gamma_2 &= 0.5029 \rightarrow K_2 = 0.6874/\text{day} \rightarrow 1/K_2 = 1.4548 \text{ days} \end{aligned}\right| \qquad (18)$$

It is of interest to check how the theoretical recession functions obtained by the above procedure, fit their observed counterparts. Fig. 6 gives this visual comparison for the recession curves of the daily flow series of the Powell River during the year 1921 for recessions which were longer than four days. This choice is an arbitrary selection, imposed by the practical difficulty of plotting all the recessions registered in 40 years. Attention is called to the fact that in general the curves would not be well fitted by straight lines. This means that the representation of the watershed storage by a single linear reservoir would not be appropriate.

Once the values K_1 and K_2 are estimated the next problem is how to describe statistically the random variable W. The set of outcomes of this variable is simultaneously obtained with K_1 and K_2. In principle, one could expect any outcome w to lie between 0 and 1. A value of $w > 1$ would indicate a reversion of the direction of flow related to the second reservoir. Analogously, $w < 0$ would indicate a reversion of the direction of flow coming from the first reservoir. These flow reversions are anticipated to be rare, but when one of them does occur, it is necessary to assert the rules which govern the inflow hydrographs, rather than the outflow hydrographs. This leads to the assumption that the characteristics of flow either from the reservoir to outlet or from the outlet to the reservoir are identical.

Qualitatively, one could expect $E[W \mid q(0)]$ to be small whenever the initial discharge $q(0)$ is large. Indeed high flows are associated with high retention in the storages that the second reservoir is supposed to represent. Consequently, its share of the flow supply should be initially higher than the flow supply which corresponds to the first reservoir. The first reservoir is characterized by a high storage capacity, which makes its contribution, $q_1(t)$, reasonably stable. Whenever the initial discharge is small, it is likely that the total flow will be sustained entirely by the outflow from the first reservoir, i.e.:

$$\lim_{q(0) \rightarrow 0} E[W \mid q(0)] = 1$$

A mathematical representation that fits the above qualitative descriptions is given by:

$$E[W \mid q(0)] = \exp[-\Psi q(0)], \qquad \Psi > 0 \qquad (19)$$

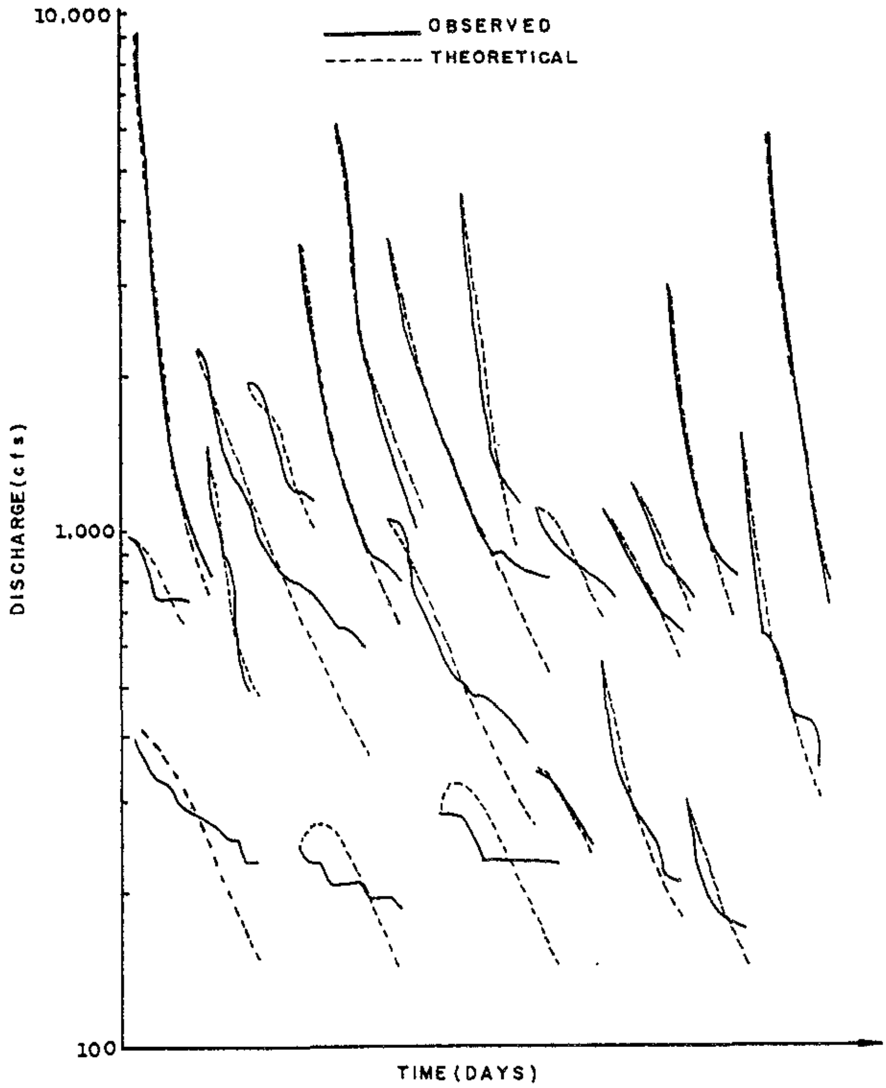

Fig. 6. Comparison between theoretical and observed recession curves, of daily flow series of the Powell River, for the year 1921.

For each historic recession curve one pair of values $[q(0,r), w(r)]$ is available, where r stands for the rth recession. These pairs can then be used to estimate the value of Ψ by the least-squares method. For the daily flow sequences of the Powell River, the value of Ψ is 0.000160. The coefficient of correlation between $q(0)$ and $\log w$ is -0.6737.

In general, the random variable W may be expressed by:

$$W = \exp[-\Psi q(0)] + Z \tag{20}$$

where Z is another random variable.

For each recession the corresponding outcomes of Z can be obtained by solving eq. 20 for Z. In the case of the daily flow series of the Powell River, we obtain:

$$z(r) = w(r) - \exp[-0.000160\, q(0,r)], \qquad r = 1, 2, \ldots, n \tag{21}$$

The next thing to do is to test whether the sample of Z may be considered as drawn from a normal probability distribution. This was tested for daily flows of the Powell River. The χ^2 goodness-of-fit test statistic is 42.70, with 36 degrees of freedom. Therefore, the hypothesis of normality could not be rejected at the 5% significance level. The sample mean and standard deviation of Z are 0.07334 and 0.25604, respectively. With these last estimates and test one can then generate the new series.

TESTING THE MODEL

The utility of the model depends on its capacity to generate new series, which are to be considered the outcomes of the same stochastic process from which the historic series is observed. The generation procedure is performed in the following steps:

Step 1. Generate the intermittent process $q_3(t)$. This is accomplished by following the procedure outlined in Fig. 2. The parameters used are shown in Figs. 4 and 5.

Step 2. Select a value of the discharge for the commencement of new samples. The mean discharge is a good choice for this value.

Step 3. Generate for each day, according to the following rules:

(a) If $q_3(t) > 0$; take step 3(b); otherwise, go to step 3(c).

(b) Make $q(t) = q(t-1) + q_3(t)$, and go back to step 3(a).

(c) If $q_3(t-1) > 0$, go to step 3(d); otherwise, go to step 3(e).

(d) Find $E[W|q(0)]$, given by eq. 19. For the daily flow series of the Powell River compute it by $E[W|q(t-1)] = \exp[-0.00016q(t-1)]$. Sample from the normal distribution a value for z. For the Powell River Z comes out as $N(0.07334, 0.06556)$. Then find w by eq. 20, and define $\delta_1 = wq(t-1)$ and $\delta_2 = (1-w)q(t-1)$.

(e) Make $\eta_1 = \gamma_1\delta_1$ and $\eta_2 = \gamma_2\delta_2$, so that $q(t) = \eta_1 + \eta_2$. Go to step 3(f).

(f) Make $\delta_1 = \eta_1$ and $\delta_2 = \eta_2$, and go back to step 3(a).

The above step-by-step procedure was used to generate 40 years of data for daily flows of the Powell River. The hydrograph for the first year of the generated series is plotted in Fig. 7. This particular hydrograph year is given because the first year of the historic record had previously been used. When the model is valid, Figs. 1 and 7 show two different realizations of the

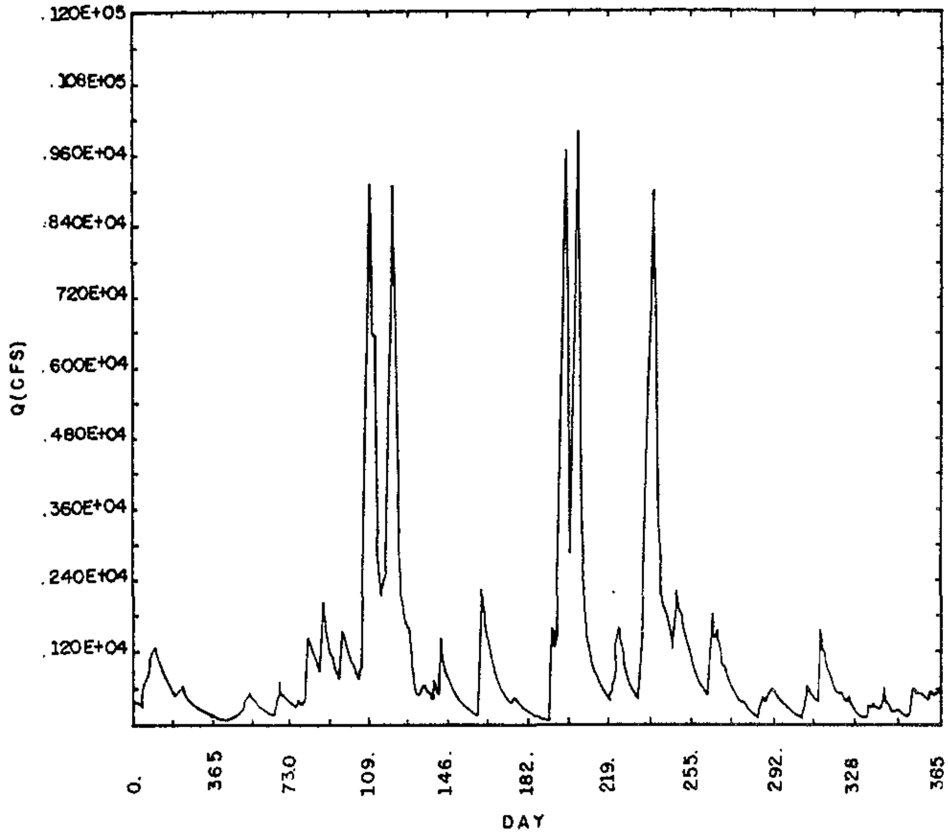

Fig. 7. A generated daily flow hydrograph of the Powell River.

same hydrologic process. These samples are different but the pattern of the series is expected to be similar. This approach is based on a subjective inference with the individual assessment, whether the hydrograph of Fig. 7 looks like the historic sample of Fig. 1, in a general hydrologic sense.

On a month-to-month basis, the random variables which are likely to be relevant for the evaluation of the validness of the model are: (1) the maximum daily discharge for each particular month; and (2) the mean daily discharge for each particular month. For each of these two random variables a matrix of observations with 40 rows (years) and 12 columns (months) was constructed out of the historic and generated samples. Let us designate these matrices by $\{F_{ij}\}$; $i = 1, 2, \ldots, 40; j = 1, 2 \ldots, 12$. For month j the sample marginal distributions are available for the historic and generated series. The Smirnov two-sample test can then be applied. It says that:

$$D = \max_x |S_1(x) - S_2(x)| \qquad (22)$$

has some distribution which 95% quantile is given approximately by:

TABLE II

Maximum daily flows for each month

Month	Mean		Standard deviation		D
	hist.	gen.	hist.	gen.	
Oct.	843.2	758.0	1,649.6	568.2	0.350*
Nov.	2,574.3	833.6	3,522.9	990.8	0.275
Dec.	5,328.4	4,615.7	4,706.5	5,235.4	0.200
Jan.	7,890.0	12,347.2	6,190.5	11,286.5	0.275
Feb.	8,615.6	11,248.0	5,377.4	9,995.9	0.125
Mar.	7,204.2	7,888.8	4,565.6	7,156.4	0.150
Apr.	5,000.1	6,350.9	3,270.7	5,755.6	0.100
May	3,772.6	4,638.5	3,944.9	3,164.6	0.250
Jun.	2,320.4	2,487.5	2,910.2	2,153.4	0.225
Jul.	2,289.5	1,408.7	2,324.7	1,002.4	0.225
Aug.	1,501.4	1,680.9	1,745.8	970.1	0.300
Sep.	747.8	1,346.8	904.2	838.0	0.550*

* The test is rejected at the 5% significance level.

TABLE III

Mean daily flows for each month

Month	Mean		Standard deviation		D
	hist.	gen.	hist.	gen.	
Oct.	235.2	293.8	192.2	233.7	0.200
Nov.	584.0	196.4	586.1	160.4	0.425*
Dec.	1,288.4	1,274.9	1,053.9	1,200.0	0.125
Jan.	1,981.2	3,394.2	1,224.6	2,390.9	0.300
Feb.	2,396.9	3,186.6	1,217.7	2,648.6	0.175
Mar.	2,310.3	2,450.3	1,136.8	1,943.8	0.225
Apr.	1,612.4	1,698.6	753.6	1,290.9	0.150
May	1,122.7	1,563.0	786.5	1,065.7	0.300
Jun.	652.9	880.0	492.1	736.0	0.200
Jul.	610.3	550.3	440.2	374.3	0.125
Aug.	437.0	707.1	384.7	400.7	0.400*
Sep.	241.7	578.8	173.2	352.2	0.550*

* The test is rejected at the 5% significance level.

$$d_{95} = 1.358[(n_1 + n_2)/n_1 n_2]^{1/2} \tag{23}$$

where $S_1(x)$ is the sample c.d.f. of the historic sequence and $S_2(x)$ is its counterpart for the generated one. The sample sizes are n_1 and n_2. For $n_1 = n_2 = 40$, $d_{95} = 0.304$.

The results of the test are displayed in the last columns of Tables II and III. In these tables the values of:

$$\bar{F}_j = \frac{1}{40} \sum_{i=1}^{40} F_{ij} \qquad \text{and} \qquad \text{std}(F_j) = \frac{1}{39} \left[\sum_{i=1}^{40} (F_{ij} - \bar{F}_j)^2 \right]^{1/2} \qquad (24), (25)$$

are also shown for the historic and generated series, respectively.

The deviations marked by an asterisk are those of the rejection of the hypothesis of statistical equality of samples, at the 5% significance level. Using jointly the results given in the two tables, one can see that the period of time between August and November is characterized by a rejection of the model. The remainder of the year shows the model to be accepted. Previously, while studying the process $q_3(t)$, it was found that a reasonable fit could not be obtained for the autumn data. This is probably also the reason for a poor performance of the overall model during this specific season.

CONCLUSIONS AND RECOMMENDATIONS

The proposed streamflow model is able to generate new samples with the complex characteristics of daily streamflow. The intermittent model fairly fits the positive first differences of daily streamflow. Representation of the recession parts of hydrographs as a stochastic output from the two linear reservoirs is acceptable.

Several further research possibilities of the dual streamflow model look promising, such as:

(1) If the direct-runoff $q_3(t)$ is supposed to represent the portion of the input to the watershed which is not retained by any river basin storage, it is likely that the parameters of $q_3(t)$ are strongly related to those that define the precipitation for the area. A joint study of the two processes could yield results valid for regional applications.

(2) The constants K_1 and K_2, associated with the linear reservoirs, are estimated by an iterative algorithm. They define the operation rules of the two reservoirs. As these two reservoirs conceptually represent the watershed retention capacity, K_1 and K_2 must be related to physiographic characteristics of the catchment. Therefore an estimation procedure could be developed to use this additional information.

ACKNOWLEDGMENTS

This study is part of a Ph.D. dissertation submitted by the author to Colorado State University, Fort Collins, Colorado. Thanks go to Dr. Vujica Yevjevich, Professor of Civil Engineering, and Dr. D. C. Boes, Associate Professor of Statistics, for their suggestions and guidance. The financial support was provided by the National Research Council of Brazil (CNPq).

REFERENCES

Kelman, J., 1977. Stochastic modeling of hydrologic intermittent daily processes. Colo. State Univ., Fort Collins, Colo., Hydrol. Pap. No. 89.
Kuester, J.L. and Mize, J.H., 1973. Optimization Techniques with Fortran. McGraw-Hill, New York, N.Y.
Quimpo, R.G., 1967. Stochastic model of daily river flow sequences. Colo. State Univ., Fort Collins, Colo., Hydrol. Pap. No. 18.