Letters

Applications of the Theory of Runs to Hydrology

IGNACIO RODRÍGUEZ-ITURBE*

Universidad del Zulia, Maracaibo, Venezuela

INTRODUCTION

The problem of runs is one of long history among hydrologists. Classically a run is defined as a sequence of observations of the same kind preceded and succeeded by one or more observations of a different kind.

Considering a sequence of random variables x_n, $n = 0, 1, 2, \ldots$ and choosing a constant x_0, one may arbitrarily classify the nth year as a surplus year if $x_n > x_0$ and call $x_n - x_0$ the surplus. Similarly if $x_n \leq x_0$, the nth year is a deficit year with the deficit $x_0 - x_n$ [*Downer et al.*, 1967].

A consecutive sequence of k surplus years preceded and succeeded by a deficit year is called a run length of length k, and the sum of the surpluses $x_n - x_0$ over such a run length is called a surplus run sum. A deficit run length and a deficit run sum are similarly defined.

Four specific problems will be studied here:

1. The mean number of times a river may reach an arbitrary discharge during an arbitrary time.
2. The mean time that a river spends between successive upcrossings or downcrossings of an arbitrary level.
3. The mean duration of an upward excursion of a river over an arbitrary level and the mean duration of a downward excursion of a river below an arbitrary level. This is equivalent to studying the mean run length and the mean deficit run length with respect to an arbitrary level.
4. The mean volume of water that a river may carry during an arbitrary time over an arbitrary stage.

* Now at the Instituto Venezolano de Investigaciones Cientificas, Caracas, Venezuela.

The operating policy of a water resources system is influenced by the sequence of hydrologic events. Consider two series of hydrologic data having the same long-term average and variance. One series differs from the other by having longer runs of very low flow counterbalanced by runs of very high inflow. Although in both the storage requirements for a given degree of regulation may be similar, the operating policies would be quite different and as a result the benefits derived from the system under different operating policies would be markedly dissimilar [*Buras*, 1966]. Part of the theory used in this study has been previously applied with good results to a series of monthly levels of the Orinoco River [*Rodríguez-Iturbe*, 1968]. In this brief report the hydrologic applications of the theory have been expanded and are illustrated here, making use of annual discharge data of the Rhine River at Basle.

RHINE RIVER STUDY

If the number C_u of crossings is finite, then $C_u = U_u + D_u$, U_u and D_u being the number of upcrossings and downcrossings, respectively, of an arbitrary level u.

Let $\xi(t)$ denote a real-valued, normal, continuous parameter, stationary process having for convenience zero mean. The $2i$th moment of the spectral function $F(\omega)$ is denoted by λ_{2i}

$$\lambda_{2i} = \int_0^\infty \omega^{2i}\, dF(\omega), \quad i = 0, 1, 2, \cdots \tag{1}$$

where ω is the angular frequency.

The mean number of crossings of the level u by $\xi(t)$ during the interval $(0,T)$ is given by *Rice* [1954] as

$$E[C_u(0, T)] = \frac{T}{\pi}\left(\frac{\lambda_2}{\lambda_0}\right)^{1/2} e^{-u^2/2\lambda_0} \tag{2}$$

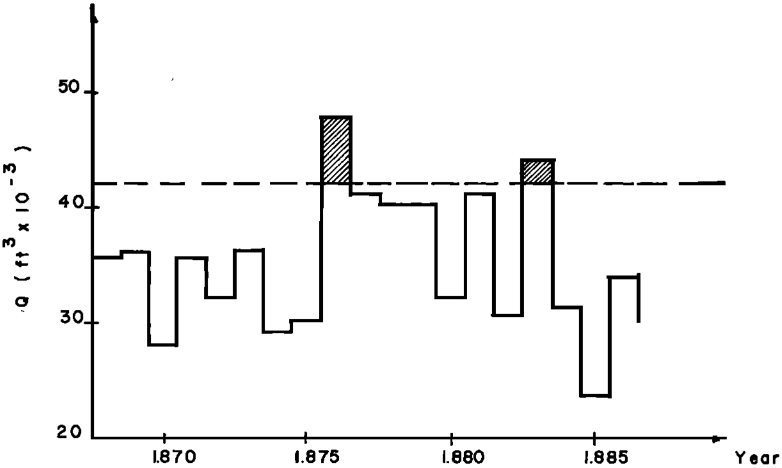

Fig. 1. Representative portion of the series of annual flows of the Rhine River at Basle.

where λ_2 may be estimated by the method given by *Rodríguez-Iturbe* [1968] and λ_0 is the variance of $\xi(t)$.

Equation 2 was applied to the series of annual flows of the Rhine River at Basle constructed with the 150 years of data given by *Yevjevich* [1963]. Figure 1 shows part of this series, which has a mean of 36,253 ft^3/sec and a variance of 33.10 × 16^6 (ft^3/sec)^2. A χ^2 test performed for this series shows there is a prob-

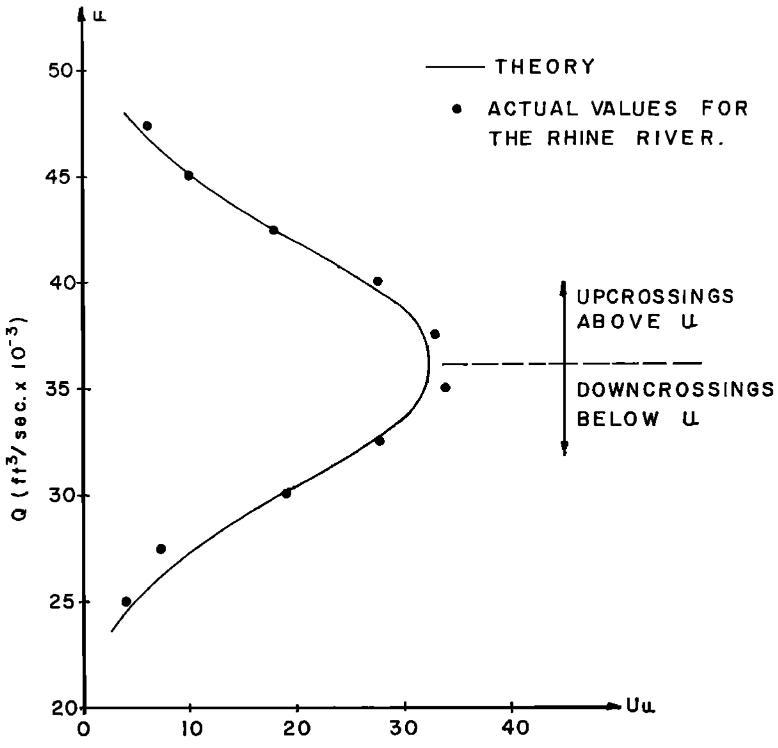

Fig. 2. Mean number of upcrossings at different levels. Series of annual flows of the Rhine River at Basle.

ability of 92% that such a sample comes from a normal population. Figure 2 shows the theoretical and actual number of upcrossings at different levels for the series of annual flows of the Rhine River.

Let $F_1(t)$ denote the distribution function of the duration of an upward excursion over a certain level or in other words the distribution of the mean run length, and let $F_2(t)$ be the distribution function of the duration of the interval betwen an arbitrarily chosen upcrossing and the next upcrossing. The functions $F_1(t)$ and $F_2(t)$ have not yet been obtained in an explicit manner, but *Cramer and Leadbetter* [1967] have calculated exact expressions for their moments, which for the mean values have been simplified by *Rodríguez-Iturbe* [1968] in order to apply them to hydrologic problems. For hydrologic time series the mean values of $F_1(t)$ and $F_2(t)$ may be written as

$$\int_0^\infty t\, dF_1(t) = \mu^{-1}[P\{\xi(0) > u\}] \qquad (3)$$

$$\int_0^\infty t\, dF_2(t) = \mu^{-1} \qquad (4)$$

where

$$\mu = E[U_u(1)] = \frac{1}{2\pi}\left(\frac{\lambda_2}{\lambda_0}\right)^{1/2} e^{-u^2/2\lambda_0} \qquad (5)$$

In the case that $\xi(t)$ has a nonzero mean m it is only necessary to replace u by $u - m$ in the previous formulas, since $\xi(t)$ crosses the level u when $\xi(t) - m$ crosses $u - m$.

Figures 3 and 4 shows the theoretical curves obtained from equations 3 and 4 and the actual

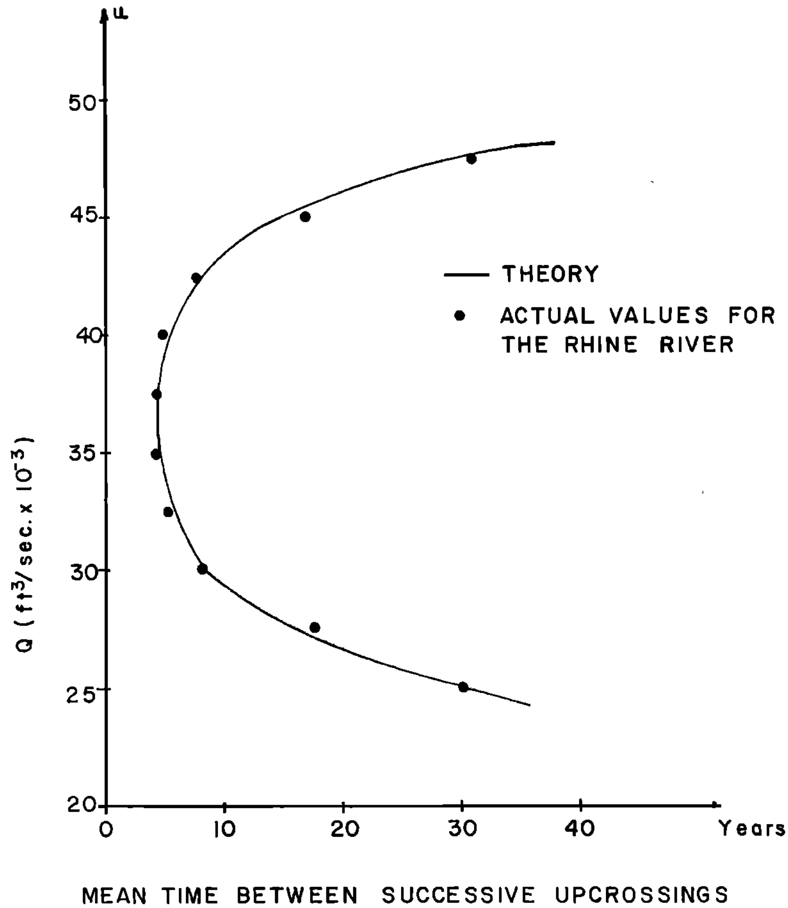

Fig. 3. Mean time between successive upcrossings of a given level. Series of annual flows of the Rhine River at Basle.

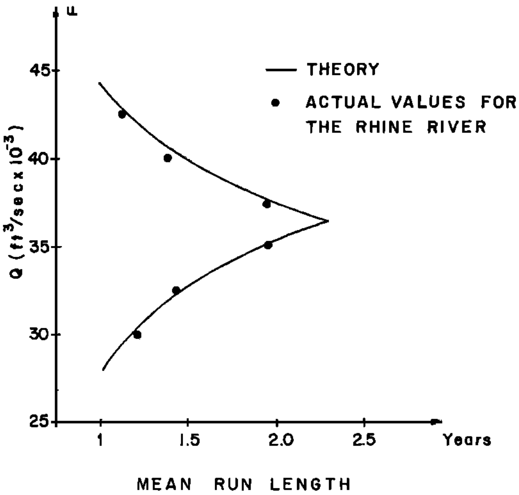

Fig. 4. Mean duration of the length of an upward excursion over a certain level. Series of annual flows of the Rhine River at Basle.

values for the series of annual flows of the Rhine River.

The area cut off above an arbitrary level by the process has been investigated by *Cramer and Leadbetter* [1967] under the name of Z_n — exceedance measures. Defining for any nonnegative integer n

$$\left.\begin{aligned} \eta_n(t) &= (\xi(t) - u)^n \quad \text{if} \quad \xi(t) > 0 \\ &= 0 \qquad\qquad\quad \text{otherwise} \end{aligned}\right\} \tag{6}$$

and

$$Z_n(t) = \frac{1}{T}\int_0^T \eta_n(t)\,dt \quad n = 0, 1, 2, \cdots \tag{7}$$

Z_0 will be the proportion of time of $(0,T)$ that $\xi(t)$ spends above the level u. $TZ_1(T)$ is just the area cut off by the process above the level u in $0 \leq t \leq T$ as indicated by the shaded area in Figure 1.

The mean of $Z_n(t)$ is given by

$$\left.\begin{aligned} E[Z_n(t)] &= \lambda_0^{n/2}\int_K^\infty (x - K)^n\Phi(x)\,dx \\ &\text{where} \\ \Phi(x) &= (2\pi)^{-1/2}e^{-x^2/2} \\ K &= (u - m)\lambda_0^{-1/2} \end{aligned}\right\} \tag{8}$$

and m represents the mean of the process. For the area cut off above a level one gets

$$\begin{aligned} E[TZ_1(T)] &= T\lambda_0^{1/2}\left[\int_K^\infty x\Phi(x)\,dx - K\int_K^\infty \Phi(x)\,dx\right] \\ &= T\lambda_0^{1/2}\left[\int_K^\infty -\Phi'(x)\,dx - K\int_K^\infty \Phi(x)\,dx\right] \\ &= T\lambda_0^{1/2}[\Phi(K) - K + KF(K)] \end{aligned} \tag{9}$$

where $F(\cdot)$ denotes the normal cumulative distribution function. The variance of $Z_n(t)$ has been evaluated by *Cramer and Leadbetter* [1967], but it will not be used in this paper.

The volume of water that the Rhine carried during 150 years over several arbitrary levels has been computed in two ways: by using equation 9 and by measuring the shaded area shown in Figure 4 by means of a planimeter.

Figure 5 shows a comparison of the results obtained where there is an average difference of 3% between the theoretical curve and the computed points.

COMMENTS

Stationarity and normality are the main assumptions for the practical use of the tech-

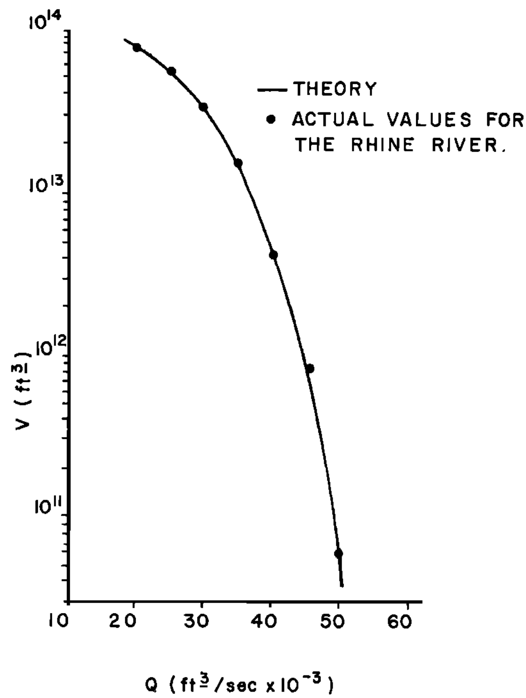

Fig. 5. Mean volume of water discharged by the Rhine River over different levels during a period of 150 years.

niques applied in this paper. The subject of stationarity poses questions such as the length of the time scale to which one has to test stationarity when dealing with hydrologic time series, and it also raises questions such as the statistical character of cycles in hydrology. The existence of the annual cycle in many hydrologic data does not necessarily make these series nonstationary. A time series with a cyclic component may or may not be stationary [*Rodríguez-Iturbe,* 1968]. Some considerations about stationarity when dealing with hydrologic time series will be the subject of a future paper by the author, and the problems previously noted will be studied in detail. *Cramer and Leadbetter* [1967] have extended the theory used in this paper for the case that the nonstationary character of a time series is due to a cyclic component.

Although many hydrologic data are not Gaussian, the approximation appears to provide reasonable estimates of the crossing and run characteristics of some time series [*Nordin,* 1968]. One should note that although the probability distributions of the variables studied in this paper are not known, the moments are useful not only in their own right but also in providing bounds on, and approximations to, probabilities of interest.

Acknowledgments. The writer thanks Dr. Vujica Yevjevich of Colorado State University and Dr. Carl F. Nordin of the U. S. Geological Survey for their helpful comments on a preliminary draft of this paper.

REFERENCES

Buras, N., Dynamic programming in water resources development, in *Advances in Hydroscience,* vol. 3, edited by Ven Te Chow, Academic Press, New York, 1966.

Cramer, H., and M. R. Leadbetter, *Stationary and Related Stochastic Processes,* 348 pp., John Wiley, New York, 1967.

Downer, R. N., M. M. Siddiqui, and V. Yevjevich, Application of runs to hydrologic droughts, *Proc. Int. Hydrol. Symp.,* Fort Collins, Colorado, *1,* 496–505, 1967.

Nordin, Carl F., Computer analysis of bed configurations in sand channels, paper presented at the American Society of Civil Engineers Hydraulics Division 16th Annual Speciality Conference, 1968.

Rice, S. O., Mathematical analysis of random noise, in *Selected Papers on Noise and Stochastic Processes,* edited by N. Wax, Dover, New York, 1954.

Rodríguez-Iturbe, I., A modern statistical study of monthly levels of the Orinoco river, *Bull. Int. Ass. Sci. Hydrol., 4,* 25–41, 1968.

Yevjevich, V. M., Fluctuations of wet and dry years, 1., 55 pp., *Colorado State Univ. Hydrol. Pap. 1,* 1963.

(Manuscript received October 2, 1968;
revised March 24, 1969.)