267

Congress 1979

INTERNATIONAL ASSOCIATION FOR HYDRAULIC RESEARCH

Cagliari . 1979

Storage reservoir reliability and the Hurst phenomenon
(Subject A.b.)

V. Klemeš

Research Hydrologist, Inland Waters Directorate
Department of the Environment
Ottawa, Ontario, Canada

Synopsis

It is shown that the Hurst phenomenon, while having a significant influence on the magnitude of storage fluctuations in an infinite reservoir, has relatively little effect on performance reliability of finite reservoirs of the size typically encountered in practice.

Résumé

On démontre que le phénomène de Hurst a peu de répercussions sur la fiabilité du fonctionnement des réservoirs à capacité limitée qui sont de dimensions typiques, comme ceux que l'on rencontre en pratique, bien qu'il influence grandement l'importance des variations du volume d'eau dans un réservoir à capacité illimitée.

A.B.3

1. Introduction

The traditional way of estimating the storage capacity necessary for an on-stream reservoir in order to prevent the reservoir outflow from dropping below a certain flow Q (called safe draft, firm yield or target draft) has been to compute it from a "design inflow series" $\{X\}_n$ for the reservoir site using a simple water-accounting technique which has become known as the sequent peak procedure (see Klemeš[1] for a historic summary).

When the target draft Q is set equal to the mean inflow $\bar{X}_n$, the storage capacity obtained in this manner is equal to the range of a residual mass curve of the design inflow series. In such a case, the range is usually called the adjusted range since the mass curve to which it pertains has been computed with respect to the sample mean $\bar{X}_n$ and thereby 'adjusted' to reach zero at the end of the series of length n. A mass curve computed with respect to the true population mean $\bar{X}$ of a complete series $\{X\}$ generally does not reach zero at the end of a sample series of length n and, for this series, it yields a so-called 'crude' range. Thus the *crude range* for a series $\{X\}_n$ is given as

$$R_n = \max Y_t - \min Y_t \quad , \quad 0 \le t \le n \tag{1}$$

where Y_t is the residual mass curve $Y_t = \Sigma_{i=1}^{t}(X_i - \bar{X})$, $Y_0 = 0$, whereas the *adjusted range* for a series $\{X\}_n$ is given as

$$R_n^* = \max Y_t^* - \min Y_t^* \quad , \quad 0 \le t \le n \tag{2}$$

where $Y_t^* = \Sigma_{i=1}^{t}(X_i - \bar{X}_n)$, $Y_0^* = 0$. A pertinent graphical interpretation is shown in Fig. 1.

Since 1936, H.E. Hurst and his co-workers have conducted numerous studies on storage reservoirs on the Nile River, summarized in Hurst et al.[2]. These studies led to an investigation of the storage capacity that would have to be provided for a complete equalization of flow of the Nile in order to eliminate spilling as well as a reduction of the draft (the so-called full regulation). The analyses were based on the sequent peak procedure. Hence, for any given series of reservoir inflow $\{X\}_n$ (represented by annual flow totals), for which the equalized annual flow is equal to $\bar{X}_n$, the necessary storage capacity is equal to R_n^*. Sample series of a given length n, taken from the same basic series $\{X\}_N$, $N > n$, led to different values of R_n^*. One obvious reason for these differences was the different variability of flows in different samples. To eliminate this influence and to make the sample values of R_n^* mutually comparable, Hurst normalized R_n^* by the standard deviation S_n of the underlying sample series $\{X\}_n$, thereby introducing a new variable

$$R_n^{**} = R_n^*/S_n \tag{3}$$

now commonly known as the *rescaled adjusted range*. He then split the series $\{X\}_N$ into $k = N/n$ samples $\{X\}_n$ and plotted the logarithm of the average $\bar{R}_n^{**} = (1/k)\Sigma_{i=1}^{k} R_{n,i}^{**}$, against the logarithm of the sample length n.

---

[1] Klemeš, V. (in press). Applied storage theory in evolution. Advances in Hydroscience, vol. 12.

[2] Hurst, H.E., Black, R.P. and Y.M. Simaika (1965). Long-term storage. Constable, London.

The plot log $\bar{R}_n^{**}$ versus log n yielded two results important to storage reservoir design. The first was the fact that $\bar{R}_n^{**}$ increased with n and did not converge to any finite limit. This finding revealed a fundamental flaw in the traditional method of estimating reservoir storage capacity as the maximum depletion that a semi-infinite (bottomless) reservoir, fed with a single inflow series and subjected to a given draft Q, would experience during the period covered by the series. It showed that, on the average, the estimate of storage capacity obtained in this manner would increase with the length of the flow record employed - a feature not previously fully appreciated by engineers.

The second result was the finding that the rate of increase of $\bar{R}_n^{**}$ with n in observed series of geophysical phenomena was significantly higher than the stochastic process theory of that time predicted. According to the theory, the increase of $\bar{R}_n^{**}$ should be proportional to $n^{0.5}$ for large n, whereas Hurst showed that for natural phenomena such as river flows, precipitation, temperatures, tree rings and others it was proportional to n^h, where h was greater than 0.5 and on the average equal to about 0.72. The fact that typical natural stochastic series have $h > 0.5$ has become known as the *Hurst phenomenon*. The relevance to storage reservoir design resides in the fact that an estimate of reservoir storage capacity (obtained by the above described traditional method) is, on the average, lower when based on an inflow series which does not exhibit the Hurst phenomenon than it would be if based on a series which does exhibit it. Hence it has been suggested that reservoirs designed on the basis of simple mathematical models of streamflow series not exhibiting the Hurst phenomenon may have a significantly lower reliability then claimed because of the underestimation of their storage capacity. It is therefore deemed of paramount importance to base reservoir design on streamflow models exhibiting the Hurst phenomenon.

It is the purpose of this paper to demonstrate that the importance of the Hurst phenomenon for reservoir performance reliability has been exaggerated and that the main reason for this exaggeration can be traced to a fundamental error which Sudler[3] identified in Hazen's[4] storage reliability assessment. This was the application of results valid for a semi-infinite reservoir to a finite reservoir with a relatively small storage capacity.

2. Reliability of reservoir performance in dry periods

This reliability can be measured in many different ways, the three most common reliability characteristics being as follows:

1) *Annual reliability* (A), defined as the probability that reservoir outflow will not drop below the target draft during the year (any year); this concept was introduced by Hazen.[4]

2) *Time reliability* (T), defined as the percentage of time in which reservoir outflow is equal to higher than the target draft during some long period of time; this concept was also introduced by Hazen.[4]

---

[3] Sudler, C. (1927). Storage required for the regulation of streamflow. Trans. Amer. Soc. Civil Eng., 91, 622-660.

[4] Hazen, A. (1914). Storage to be provided in impounding reservoirs for municipal water supply. Trans. Amer. Soc. Civil Eng., 77, 1539-1640.

3) *Volume reliability* (V), defined as the amount of water actually supplied from the reservoir over some long period of time, expressed in percent of the total target water amount; this concept seems to have been first applied by Savarenskiy.[5]

For a constant draft, $A \le T \le V$, since usually not all failure periods (periods when reservoir outflow is less than the draft) extend over the whole year and reservoir outflow is not always reduced to zero during a failure period.

Reservoir design based on the sequent peak procedure implies a 100% reliability (of any of the three above types) with regard to the inflow series $\{X\}_n$ employed. On the other hand, the sampling variability of series $\{X\}_n$ formed from an underlying series $\{X\}_N$, $N \gg n$, implies that storage capacity K_n obtained in this manner (and equal to R_n^* for draft $Q = \bar{X}_n$) has a distribution. Hence no specific value of K_n guarantees a 100% reliability of the target draft during a period of length n with an absolute certainty; there always is some probability P_n that the given K_n may not be sufficiently large. This probability was introduced as another measure of reservoir performance reliability by Fiering.[6] He estimated it from an empirical distribution of K_n values obtained from a number of synthetic inflow series of length n via the sequent peak algorithm. Thus, for example, a value of K_n exceeded in 20% of cases is said to have an 80% reliability C_n or a 20% risk of failure P_n. Note that here the terms "failure" and "reliability" have a different meaning than above: a "failure" means here an inability of the reservoir to maintain the target draft without any interruption or reduction (i.e. without failure) throughout the whole period n, while "reliability" means its ability to do so.

Fiering himself pointed out that the main purpose of his analysis was to show, first of all, the inadequacy of using a single (historic) flow series for finding reservoir storage capacity, secondly, that the risk P_n may be a poor characteristic of reservoir performance as compared, for instance, with the volume reliability V. Unfortunately, his caveat went largely unnoticed and the distribution of K_n and hence that of R_n, R_n^* or R_n^{**} were later adopted by many authors as the basis for assessing reservoir performance reliability.

The inadequacy of this practice stems from the fact that P_n measures the risk that a reservoir will not meet an objective which is never aimed at in practice, namely a 100% reliable (in terms of A, T or V) performance during the whole of a design period lasting many decades. The important point is that, because of economic (and other) tradeoffs between benefits from water shortages averted and costs involved in averting them, the optimum streamflow regulation regime is always one in which some failures do take place. This makes the probability of a nonfailure operation C_n (or its complement $P_n = 1 - C_n$) a rather irrelevant performance characteristic. It is, for instance, quite likely that the probability of a given storage capacity meeting a specified draft without failure during an n-year period may decrease from, say, $C_n = 95\%$ for an inflow which does

---

[5] Savarenskiy, A.D. (1940). A method for streamflow control computations (in Russian). Gidrotekh. stroitelstvo, (2), 24-28.

[6] Fiering, M.B. (1967). Streamflow synthesis. Harvard University Press, Cambridge, Massachusetts.

not exhibit the Hurst phenomenon to C_n = 50% for an inflow which does exhibit it (Wallis and O'Connell[7]). Also, it may well happen that the storage capacity would have to be increased by, say, 50% if C_n in the latter case were to be brought up to the original value of 95%. However, it must be borne in mind that, while the two values of C_n may differ significantly, the reservoir performance reliability measured by A, $\bar{T}$ or V is likely to differ by an order of magnitude less for the two inflow processes. Moreover, the difference may often be within the margin of accuracy of the reliability estimate which is as a rule quite wide (Klemeš[8]).

The reason that, for the two inflow processes, the differences in the classical reliability measures are much lower than those in C_n is that the latter characteristic is derived on the assumption of an infinite or semi-infinite reservoir while the classical reliability measures assume a finite reservoir. The following example is intended to elucidate this problem.

3. Numerical Example

We shall consider two series of mean annual flows, series a which does not exhibit the Hurst phenomenon and thus has $h = 0.5$ and series b which, as far as can be inferred from the data, does and has $h = 0.75$. In order to be able to keep easily under control the sample statistics and thus to eliminate their influence on the result, we shall consider a simple case where the flow can take on only two different values, $X_1 = 1$ and $X_2 = 3$, both with equal probabilities $p_1 = p_2 = 0.5$, so that the mean is $\mu = 2$ and the standard deviation $\sigma = 1$ for both series a and b.

In the example, 36-year[9] samples from the two series will be used to examine the performance reliability of a reservoir the objective of which is to deliver a draft equal to the inflow mean, i.e. $Q = \mu = 2$. The two sample series are shown in the uppermost part of Fig. 2. They were constructed in such a way that their sample parameters are equal to the population parameters so that $\bar{X}(a) = \bar{X}(b) = \mu = 2$ and $S(a) = S(b) = \sigma = 1$. To eliminate the influence of short-term persistence (discussed in section 4) on the result, both sample series were designed to have the lag-one serial correlation coefficient (based on circular definition) $r_1 = 0$[10].

---

[7] Wallis, J.R. and P.E. O'Connell (1973). Firm reservoir yield - how reliable are historic hydrological records? Bull. Hydrol. Sci., 18 (2), 347-365.

[8] Klemeš, V. (1978). The unreliability of reliability estimates of storage reservoir performance based on short streamflow records. Proc. International Symposium on Risk and Reliability in Water Resources, Vol. 1, pp. 127-136, University of Waterloo, Waterloo, Ontario.

[9] This length has been chosen for two reasons: 36 years is close to the typical length of a historic flow record and the number 36 is easily divided into various subsamples of equal lengths which is a convenient feature in the evaluation of the Hurst phenomenon.

[10] The lag-one serial correlation in a two-variable series with $\sigma = 1$ is easy to control because $r_1 = (n_1 - n_2)/n$ (Klemeš[11]), where n is the sample size, n_1 is the number of pairs of equal successive flows and $n_2 = n - n_1$ is the number of pairs of unequal successive flows, i.e. the number of corner points of the residual mass curve of the series.

[11] Klemeš, V. (1969). Problems involved in taking the lag-one serial correlation coefficient into consideration in storage reservoir design (in Czech). Proc. National Conference on Large Dams, Vol. 1, pp. 114-130, ČSVTS, Prague.

A plot of log $\bar{R}_n^{**}$ versus log n for both series is shown in Fig. 3. Note that for $n = 36$, $R_n^{**} = R_n^{*} = R_n$ because $\bar{X}_n = \mu$ and $S_n = 1$ were chosen by design.

Now, using the sequent peak procedure, we obtain $K(a) = R(a) = 6$ for series a while series b leads to $K(b) = R(b) = 9$ (Fig. 2, second from top). Thus, in this case, we see that to obtain a failure-free performance for an inflow exhibiting the Hurst phenomenon, it would be necessary to raise the storage capacity by 50% as compared to the storage needed for an inflow which does not exhibit it. This result tends to make one believe that the neglect of the Hurst phenomenon in modelling the inflow series would severely impair the reservoir performance reliability.

To see whether it is so we turn to the lower half of Fig. 2. It shows storage fluctuations[12] for three finite reservoirs with different storage capacities operating on the two series and delivering the target draft Q = 2 mentioned above. The reservoir performance will be evaluated in terms of reliabilities T and V. Note that failures in performance occur in the periods during which the reservoir is empty since then the outflow must be reduced to the natural inflow which, in our dry periods, is $X_1 = 1$.

In the first case, the storage capacity is chosen to be $K = 6$. Naturally, it fully satisfies the need in series a so that both T_a and V_a are 100% On the other hand, in series b, two failure periods occur and the performance reliability drops to $T_b = 91.8\%$ and $V_b = 95.9\%$ (Table 1). Thus, while the original analysis of storage requirements indicated that a neglect of the Hurst phenomenon would lead to an underestimation of storage capacity by 50%, the actual drop in reliability would be only about 8% in terms of time and about 4% in terms of volume.

While such a difference could be of practical importance, it must be realized that it pertains to a storage capacity equalling three times the average annual runoff total, i.e. to a storage coefficient $s = 3$ - a value much too large to be encountered in most real-life situations. A more likely but still rather large value of s is about 1 (for example, $s = 1.4$ for the active storage of Lake Superior). This, in the present example corresponds to $K = 2$ for which case the difference in performance reliability reduces to 5.6% (for T) and 2.8% (for V) as seen in Table 1.

However, most reservoirs would have s substantially lower than 1 (the active storage of Lake Ontario is $s = 0.2$) and, with the decrease of s, the difference in reliability (for inflows of type a and type b) will decrease as well. In theory, the difference will tend to zero with $K \to 0$ since for a "no reservoir" situation the values of A, T and V depend only on the probability distribution of flow and not on its sequential structure. In practice, the difference will effectively disappear much sooner, usually for storage coefficients within the range of those for typical storage reservoirs. In the present example, the difference disappears for $K = 1$ corresponding to $s = 0.5$ (Fig. 2), bottom; Table 1).

---

[12] Steady-state conditions are assumed, i.e. an equivalence of the initial and the final storage states is postulated.

Table 1. Example of stationary flow-regulation reliability for inflow series of size $n = 36$ with two different values of the Hurst coefficient, a) $h = 0.5$, b) $h = 0.75$.

Time-based reliability, T (in %)

Reservoir storage capacity K	Inflow series		Reliability difference $T_a - T_b$
	a	b	
6	100.0	91.8	8.2
2	86.2	80.6	5.6
1	75.0	75.0	0
0	50.0	50.0	0

Volume-based reliability, V (in %)

Reservoir storage capacity K	Inflow series		Reliability difference $V_a - V_b$
	a	b	
6	100.0	95.9	4.1
2	93.1	90.3	2.8
1	87.5	87.5	0
0	75.0	75.0	0

4. A note on short-term persistence

It is virtually impossible to determine on the basis of a short sample series such as those encountered in historic flow records, whether or not a process exhibits the Hurst phenomenon. The reasons are 1) the extremely high sampling variability of R_n^{**}, 2) the practical difficulty, for small n, in approximating the relationship between R_n^{**} and n by a straight line and 3) the tendency for the apparent straight-line fit for small n to have a slope $h > 0.5$, even for processes not exhibiting the Hurst phenomenon.

Equally important is the fact that, in a short series, a pattern suggestive of the Hurst phenomenon (the clustering of extremes of the same sign) may well be the consequence of short-term persistence or may even be due to chance. The latter is in fact the case in series b which is a sample from a random (serially independent) series.

Fig. 4 (based on Klemeš[11]) shows all the possible combinations of the lag-one correlation coefficient r_1 and the range R_n (here equal to R_n^* and to R_n^{**} as well) for a series of length $n = 36$ composed of 18 values of $X_1 = 1$ and of 18 values of $X_2 = 3$. It can be seen that, despite a complete elimination of the sampling variability both in the mean (kept as $\bar{X}_n = 2$) and in the standard deviation (kept as $S_n = 1$), the spread of the possible R_{36} values is considerable, especially for positive lag-one correlation. Thus, for instance, for a random series, R_{36} can range from 2 to 10 and for a series with $r_1 \doteq 0.3$ from 3 to 13.

However, the main points which Fig. 4 is intended to demonstrate are 1) that a range $R_{36} = 6$, which is about average for a random series, can occur in an autocorrelated series with r_1 ranging from -0.45 to 0.65, and 2) that a range $R_{36} = 9$, which in combination with $r_1 = 0$ is rather unlikely and invokes the existence of the Hurst phenomenon, would be about average for an autocorrelated series with $r_1 \approx 0.5$ and quite likely for series with r_1 from 0.1 to 0.7. Thus it is obvious that the clustering effect within a short series can often be accounted for by lag-one serial correlation alone, without the necessity of invoking the Hurst phenomenon. To distinguish between these two causes, i.e. between short-term and long-term persistence, on the basis of a short record ($n < 100$) has been found difficult because of the weakness of the applicable statistical tests (Wallis and O'Connell[7]).

5. Conclusion

The Hurst phenomenon in a streamflow series, while having a significant influence on the fluctuation of storage in an infinite reservoir, has little effect on performance reliability of finite reservoirs with storage capacities typically encountered in practice. Even for reservoirs with an over-year storage, the neglect of the Hurst phenomenon in the inflow model will lead to a relatively small reduction in reservoir performance reliability in terms of the stationary probability of a failure year, of the percentage of the time with water shortage or of the percentage of the amount of the water desired. This relative insensitivity is due to the fact that constraints on storage in a finite reservoir frequently interrupt the long-term pattern of unconstrained storage fluctuations on the basis of which the Hurst phenomenon is defined, thereby reducing the difference between storage fluctuations caused by short-persistence and long-persistence inflows. Because of this reason, a low-order (usually the first-order) autoregressive process can be expected to provide a satisfactory model for an annual flow series to be used for the assessment of reservoir performance reliability expressed in terms of the three traditional indices A, T and V.

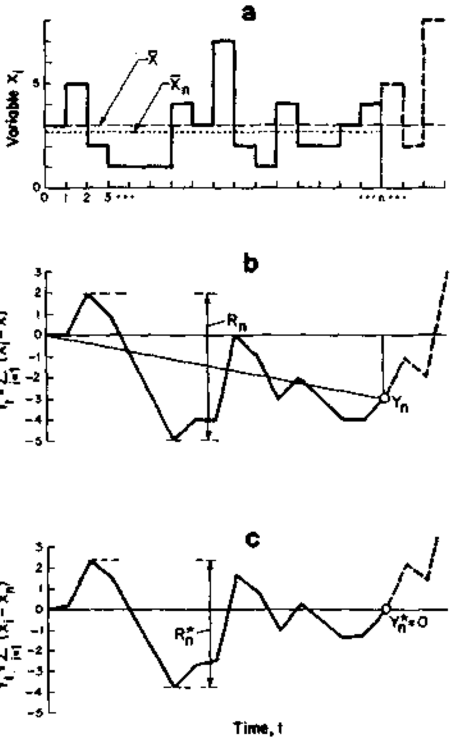

Fig. 1 Definition sketch for crude range R_n and adjusted range R_n^*.

Dessin de l'intervalle R_n à l'état brut et de l'intervalle R_n^* ajusté.

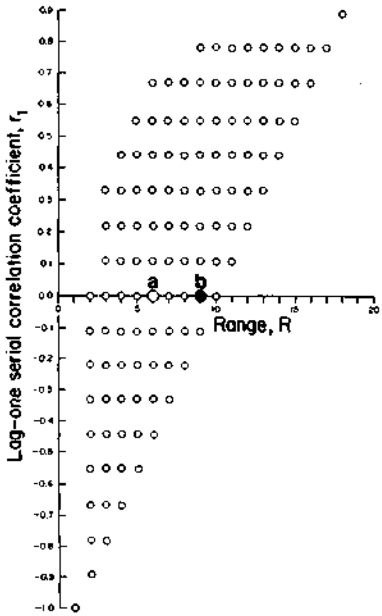

Fig. 4 Possible combinations of lag-one serial correlation coefficient r_1 and range R (equal to R^* and R^{**} in this case) for a series composed of 18 values $X_1 = 1$ and 18 values $X_2 = 3$.

Combinaisons possibles du premier coefficient de corrélation sérielle r_1 et de l'intervalle R (équivalent à R^* et R^{**} dans ce cas) pour une série composée de 18 valeurs $X_1 = 1$ et 18 valeurs $X_2 = 3$.

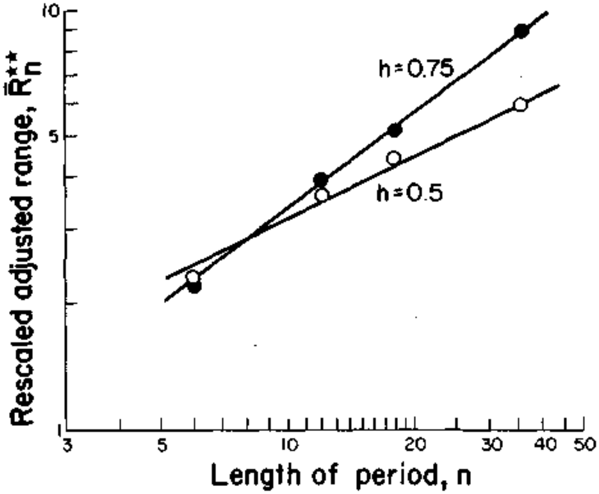

Fig. 3 Average value of rescaled adjusted range, $\bar{R}_n^{**}$, as a function of n for series a (empty circles) and series b (full circles).

Valeur moyenne de l'intervalle ajusté retracé à l'échelle, $\bar{R}_n^{**}$, en tant que fonction de n pour la série a (cercles vides) et la série b (cercles pleins).

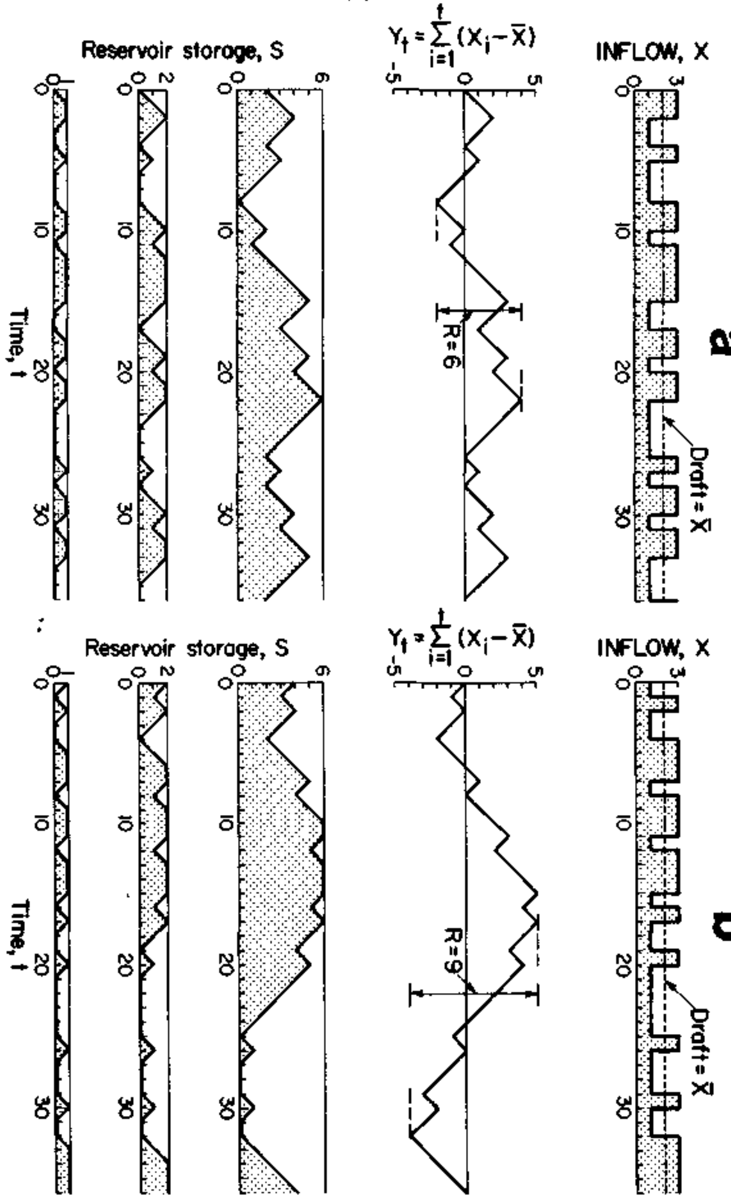

Fig. 2 Inflow series *a* and *b* (top); residual mass curves of series *a* and *b* (second from top); fluctuations of storage in reservoirs of storage capacity $K = 6$, 2 and 1 subject to inflow series *a* and *b* and to draft $Q = 2$ (lower half of the figure).

Séries *a* et *b* du débit entrant (première courbe); courbes des écarts cumulés résiduels des séries *a* et *b* (deuxième du haut); fluctuations du volume d'eau dans un réservoir d'une capacité de $K = 6$, 2 et 1 où arrivent les séries *a* et *b* du débit entrant et où le débit soutirable

Tom McMahon didn't believe my conclusions, so when I was on sabbatical at Monash, I suggested that we put them to a more realistic test. I suggested the simulation scheme and his grad. student Sri Srikanthan did the computations. The result was the WRR paper (1981, pp. 737-751)

over →

It would be an interesting exercise for a mathematically motivated grad. student (Master's degree?) to calculate the probabilities of all the points in Fig. 4.