Nordic Hydrology (1971), II, 109–129
Published by Munksgaard, Copenhagen, Denmark

STOCHASTIC MODELLING OF STORM PRECIPITATION

S. RAMASESHAN

Department of Civil Engineering, Indian Institute of Technology,
Kanpur, India

The distribution of precipitation in time during storms constitutes a complex stochastic process. Mathematical modelling of the process is quite difficult. The process is conceived as a finite-duration, quantitized-data, continuous-variable, stochastic process that can be represented as a finite time series.

Broad outlines of the procedure have been given in an earlier paper (Chow & Ramaseshan 1965). This paper deals with the details of formulating and fitting a suitable stochastic model for the process and the steps are illustrated with an example. The conclusions derived on the basis of the model are also discussed. The methods discussed in this paper may be useful for the study of transient stochastic processes such as floods, droughts, earthquakes, squalls, etc., and for the design of engineering systems to resist or control such processes.

Hydrologic processes such as rainfall and runoff are complex. They vary in space and in time, and also from event to event as, for example, from storm to storm. No two storms or floods in a basin are alike in duration or time distribution of precipitation or discharge. The precipitation in a storm and the discharge in a flood at any time t may be considered as a random variable whose values vary with t. Therefore they constitute stochastic processes, i. e., processes that change in time and are controlled by the laws of probability. The process of precipitation over a period of a month or a season is generally large enough in time that the process can be considered stationary, i. e., the parameters do not vary with time, and small enough in time that the seasonal variations from one season to another will not be important. Stochastic models for such cases have been successfully formulated (Le Cam 1961). However, the precipitation within a storm does not constitute a stationary process and so its modelling and analysis are more difficult.

When adequate streamflow data are available, they may be analysed statistically, and statistical or stochastic prediction of floods may be made. For large or medium basins, flood prediction is based on system analysis of hydrometeorological systems. They may require concurrent rainfall and runoff records. Very often it is seen that fairly extensive rainfall records and only limited runoff records are available. In such cases prediction is to be based on the available extensive rainfall and limited runoff data. Sequential generation and simulation techniques (Chow 1964) are to be used in such cases and they may require stochastic models for storm precipitation. Stochastic models are also necessary for directly making probabilistic predictions on the stochastic characteristics of storm precipitation and of floods.

GENERAL CONSIDERATIONS

Storm precipitation varies continuously in time and in the values it can assume. However, the measurement of the instantaneous rate of precipitation is very difficult; usually hourly totals or totals over other periods of time are measured either manually by observers in the field or in the offices from charts obtained from recording rain gauges. They are therefore available as quantitized data (Rameseshan 1964). Furthermore, the representation and analysis of a stochastic process which is continuous in its values and in time is very difficult at the present time (1971). Therefore a discrete time, continuous variable process is assumed.

Let the precipitation at time intervals 1,2, . . . t, during a storm be X_1, X_2, . . . X_t, as shown in Fig. 1. Annual storms are defined as those which produced maximum peak discharge in a water year. In this study, data of storm precipitation during annual storms from two basins were used for formulation and verification of the model. They are, respectively, the French Broad River basin at Bent Creek, North Carolina, and the Kaskaskia River basin at Shelbyville, Illinois. During some storms the records at some stations were incomplete in that only the total storm precipitation for a specific interval was known and not its distribution in time. A multiple linear regression relationship in terms of the concurrent precipitation at other nearby stations was used to estimate the

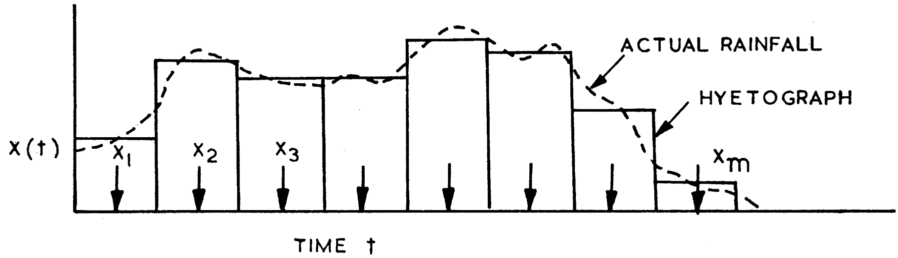

Fig. 1.
Stochastic model for time distribution of precipitation.

missing data. The Thiessen polygon method was used to calculate the average hourly precipitation over the basin.

During an intense storm there may be several storm bursts with lulls between them. Also there may be several storms in succession. In this study only the storm causing the maximum peak discharge is considered and all other precipitation is considered to be prior or subsequent precipitation. Their interaction with the annual storm in prolonging the floods is ignored for the purposes of this study. All the storms are considered to have the same duration of m hours. The assumed value of m is comparable to the largest duration of the storms on record, and for the French Broad River, $m = 36$. For many storms whose duration is less than m hours, there may be no rainfall for a part of m hours. Thus all storms have m hourly precipitation values and some of these values may be zero because the actual duration of the storm on record is less than m hours or because there were several bursts. The zero values due to the former cause may partly precede and partly succeed the observed data.

Storm Shifting (Chow & Ramaseshan (1965), Ramaseshan (1964))

The beginning of a storm, i. e., the time 1 of precipitation, may be considered as that time at which the precipitation starts or when the rate of precipitation exceeds a specified critical rate. In many cases there may be a drizzle before the main part of the storm or there may be a number of bursts some of which may have to be treated as prestorm precipitation. In such cases it will be very difficult to fix the time of beginning of a storm. Any criteria used for the above purpose should preferably be such that the resultant model is stable: there is a definite and consistent trend in the mean and the standard deviation of hourly precipitation; and they result in a simple, specific, and consistent model for the deterministic and random components of the hourly precipitation.

The assumption that time 1 corresponds to the beginning of the storm precipitation was found to be unsatisfactory, as the resultant model was erratic in its behaviour. A possible alternative may be to adjust the storms so that their median, mean, or mode occurs at the same time. Any such arrangement does not take into account the variation in duration and in the time distribution of precipitation in the different storms. The following criterion was found to be very satisfactory. In order to develop the best relative orientation of storms with one another for the time distribution, it is necessary to shift the beginning time of all storms so that the cross-correlation coefficients of their hourly rainfalls with one another are maximum. This is referred to as "storm shifting".

Shift Analysis

Consider N storms. Let X and Y be two of the annual storms in which the sequences of hourly precipitations are respectively x_t and y_t with $t = 1,2,\ldots m$. The cross-correlation coefficient of storm Y with storm X for a shift of v hours is given by

$$R_{xy}(v) = \frac{\sum_{t=1}^{m-v} x_t\, y_{t+v} - \frac{1}{m-v} \sum_{t=1}^{m-v} x_t \sum_{t=1}^{m-v} y_{t+v}}{(m-v-1)\, s_{x,t}\, s_{y,t+v}} \qquad (1)$$

in which $s_{x,t}$ is given by the equation

$$s^2_{x,t} = \frac{1}{m-v-1}\left[\sum_{i=1}^{m-v} x^2_t - \frac{1}{m-v}\left(\sum_{t=1}^{m-v} x_t\right)^2\right] \qquad (1a)$$

and $s^2_{y,t+v}$ is given by the same Eq. 1a except that x and t are replaced by y and $t+v$ respectively.

The time distribution of precipitation before and after a shift of 6 hours for storms of 1948 and 1949 in the French Broad River at Bent Creek, N.C., are shown in Fig. 2. The values of $R_{xy}(v)$ for different values of v are calculated. The larger the R_{yx}, the larger is the correlation between the sequences. It is generally seen that $R_{xy}(v)$ reaches a maximum for a specific value of v (Fig. 3), say v_o. When the second storm is shifted by v_o, the large cross-correlation coefficient indicates that, at this relative position, the time distributions of storm precipitation in the two storms are best related to each other. When all the storms are thus shifted to their best position with respect to every other storm, their relative orientation is considered to be the best. When v_o is negative, it is referred to as a "lag"; and when positive, a "lead". Thus Fig. 3 indicates a lead of 5 hours.

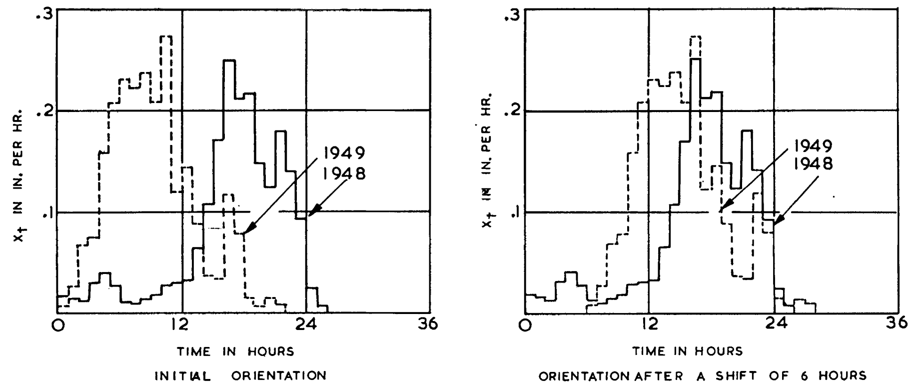

Fig. 2.
Alternative storm orientations.

There are two important factors which cause difficulties in arranging the storms in such relative position. The first is that there may be two or more bursts in a storm giving rise to two or more local maxima for the function $R_{xy}(v)$ or, when the function is flat, near v_0 and showing no significant variation over a wide range. This problem is overcome by considering only the value of v corresponding to the absolute maximum of $R_{xy}(v)$ or the centre of a wide range of v when the variation over the range is small. Also a value which is consistant with the other values is chosen instead of an inconsistent one. Secondly, not all the shifts indicated for a storm with reference to the other storms are consistent. In this case the relative shift indicated consistently by a large number of storms and estimated by the median or mode of the consistent values is chosen as the representative value. In order that the procedure be consistent and unique, the following steps may be followed (Chow & Ramaseshan 1966):

1. Initially the precipitations in all storms may start with time 1. In order to avoid many large shifts and to permit negative shifts without much loss of information and to reduce calculations, the storms may be initially shifted by arbitrary amounts. These shifts may be such that the storm precipitations are approximately similarly distributed and are centrally located with respect to the total duration of m hours. Data so oriented are referred to as preliminary data to distinguish them from the initial data. Preliminary data are used in shift analysis.

2. Let m be the duration of the N storms considered. Let $S_{i,k}$ be the "relative

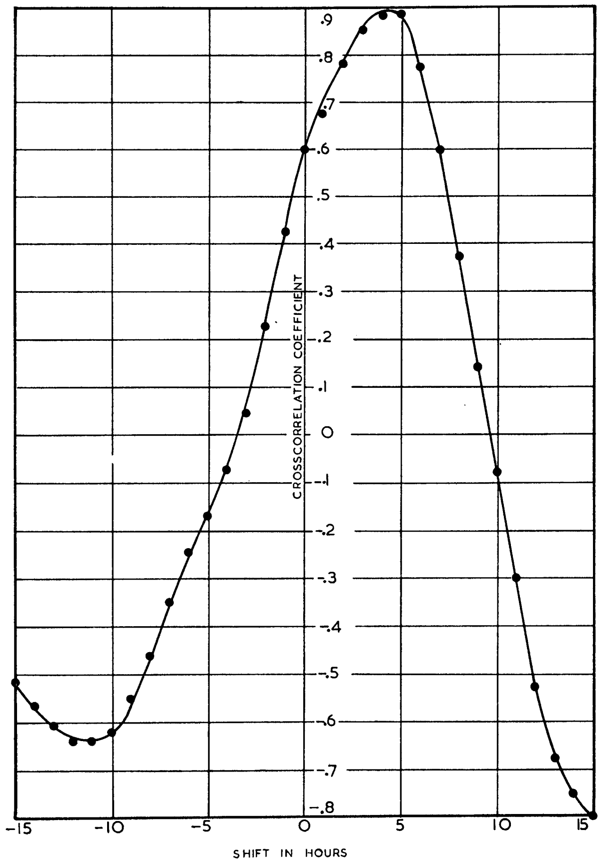

Fig. 3.
Variation of the cross-correlation coefficient with shift.

shift" of storm i with storm k for maximum cross-correlation. Calculate $S_{i,k}$ for all values of i with $i = 1,2, \ldots N$ and $i \leq k \leq N$. The results for the French Broad River Basin at Bent Creek, N.C., are shown in Table 1.

3. Storm 1 is assumed to have a shift equal to zero. The expected value of the relative shift of storm i with storm 1 is referred to as the "absolute shift". The absolute shift of storm i is denoted by $\bar{S}_{i,i}$ with $i \equiv 1,2, \ldots N$. $\bar{S}_{1,1}$ is assumed to be zero.

4. Denote the estimate of the absolute shift of storm k on the basis of storm i by $\bar{S}_{i,k}$. This is given by

$$\bar{S}_{i,k} = \bar{S}_{i,i} + S_{i,k} \tag{2}$$

As $\bar{S}_{1,1} \equiv 0$, $\bar{S}_{1,}k \equiv S_{1,k}$. The values of $\bar{S}_{1,k}$ with $k \equiv 2,3,\ldots N$ are determined and entered in Table 2.

5. The value of $\bar{S}_{1,2}$ is taken as the estimate of $\bar{S}_{2,2}$. Using this value, calculate the values of $\bar{S}_{2,k}$ with $k \equiv 3,4,\ldots N$ and enter in Table 2.

6. The values of $\bar{S}_{i,i}$ for $i \equiv 3,4,\ldots N$ can be calculated as follows: At the ith row there are (i–1) estimates of $\bar{S}_{i,i}$ given by $\bar{S}_{j,i}$ with $j = 1,2,\ldots N$. One or more of these values may differ significantly from the rest. They are considered to be inconsistent whereas the rest are considered to be "consistent" with one another. The variation between consistent values may be only a few hours. The inconsistent values are ignored. For $i \geqq 3$, $\bar{S}_{i,i}$ is generally considered to be equal to the median of the consistent values of $\bar{S}_{j,i}$ with $1 \leq j \leq i-1$. If $\bar{S}_{j,i}$ has a distinct skewed distribution, the mode may be chosen instead. The value of $\bar{S}_{i,i}$ is used to calculate $\bar{S}_{i,k}$ by Eq. 2. The above procedure is applied successively for all values of i from 3 to N.

For example, there are six estimates for $\bar{S}_{7,7}$. They are, respectively, 3,4,4,5,6, and 2. $\bar{S}_{7,7}$ is assumed to be 4.

7. The estimates of $\bar{S}_{i,i}$ obtained from the above procedure can be refined by iterations as follows. For any successive iteration, consider a value of i with $2 \leqq i \leq$ N. There are i–1 values of $\bar{S}_{j,i}$ with $j < i$, each of which is an estimate of $\bar{S}_{i,i}$.

Let $\bar{S}_{j,i} = \bar{S}_{j,j} - \bar{S}_{i,j}$ with $j = i + 1,\ldots N$. There are N–i estimates of $\bar{S}_{i,i}$ that can be obtained from $S_{i,k}$ with $i < k \leq N$ by using Eq. 2. Thus there are N–1 estimates of $\bar{S}_{i,i}$. Ignoring the inconsistent values, a better estimate of $\bar{S}_{i,i}$ can be calculated from the consistent values. This step is repeated for all values of i with $2 \leq i \leq N$ a number of times until all the values of $\bar{S}_{i,i}$ are consistent. The final result of shift analysis is shown in Table 2 for the French Broad River basin. Precipitation data for the two basins, mentioned above, indicated that the above steps constitute a consistent procedure. The iterative procedure of step 7 was developed subsequent to the later parts of the study and so was not incorporated in the earlier study (Chow & Ramaseshan 1965, Ramaseshan 1964).

The precipitation data are shifted as per the results of shift analysis. The mean, cumulative variance and the coefficient of variation of the hourly precipitation for the initial, preliminary, and shifted data are shown respectively in Figs. 4, 5,

Table 1.
Relative shifts for annual storm precipitation (French Broad River Basin).

Year	1935	36	37	38	39	40	41	42	43	44	45	46	47	48
1935		– 1	– 8	2	3	– 9	3	2	– 5	– 5	15	11	0	0
1936		0	–11	0	0	– 8	4	3	– 2	– 4	6	13	0	0
1937			0	8	9	13	– 9	–10	– 1	2	3	0	8	–13
1938				0	1	–11	5	7	– 7	– 6	> 15	10	– 2	– 1
1939					0	– 8	6	6	– 9	– 6	> 15	15	– 3	2
1940						0	12	> 15	4	13	10	0	10	10
1941							0	– 1	– 5	–11	12	8	– 6	– 3
1942								0	–11	– 7	12	9	– 5	– 5
1943									0	4	4	15	9	7
1944										0	– 2	> 15	4	5
1945											0	– 4	<–15	<–15
1946												0	–13	–12
1947													0	1
1948														0

Year	1949	50	51	52	53	54	55	56	57	58	59	60	61	62
1935	2	3	2	4	1	−1	1	1	2	7	2	6	5	7
1936	4	4	3	5	−1	0	−2	2	3	8	3	2	0	8
1937	−9	−9	−10	−10	−3	−11	8	−11	−9	−5	−10	−11	−13	9
1938	5	2	2	3	−2	4	−1	5	4	6	3	3	1	5
1939	4	2	3	3	−2	3	−4	4	5	5	3	3	2	5
1940	> 15	> 15	13	> 15	5	12	8	12	> 15	13	13	14	12	> 15
1941	0	−3	−1	−1	−6	−4	−8	−2	1	3	−1	−2	−4	1
1942	−1	−2	0	−3	−10	−3	−8	−1	0	1	−3	−3	−5	2
1943	10	13	10	10	−2	10	6	11	13	10	10	10	9	12
1944	4	8	7	9	5	7	3	10	11	6	3	10	8	11
1945	−13	−12	−14	−14	−6	< −15	−15	−13	−12	−11	< −15	< −15	< −15	−13
1946	−11	−11	−10	−12	−12	−12	< −15	−11	−10	−10	−12	−12	< −15	−12
1947	2	4	4	5	0	2	0	6	5	8	4	6	4	7
1948	5	3	3	4	−6	1	3	2	5	7	2	3	0	7
1949	0	4	−1	−1	−11	−3	−3	−2	0	3	−1	0	−4	2
1950		0	−1	1	−2	−2	−6	−2	2	4	−1	−3	−1	1
1951			0	0	−10	−1	−5	−1	2	4	−1	−1	−3	3
1952				0	−10	−1	−5	1	2	3	0	1	−3	3
1953					0	1	0	8	11	14	9	11	6	6
1954						0	−5	1	3	7	0	0	0	5
1955							0	6	9	10	5	6	6	8
1956								0	3	6	0	0	−2	3
1957									0	2	−3	−1	−5	1
1958										0	−5	−2	−10	−1
1959											0	0	−3	5
1960												0	−2	2
1961													0	4
1962														0

Table 2.

Absolute shift for annual storm precipitation (French Broad River Basin).

YEAR	1935	36	37	38	39	40	41	42	43	44	45	46	47	48
1935	0	-1	-8	2	3	-9	3	2	-5	-5	15	11	0	0
1936	1	0	-11	0	0	-8	4	3	-2	-4	6	13	0	0
1937	21	24	13	21	22	26	4	3	12	-15	-16	13	21	0
1938	-2	0	-8	0	1	-11	5	7	-7	-6	>15	10	-2	-1
1939	-3	0	-9	-1	0	-8	6	6	-9	-6	>15	15	-3	2
1940	-1	-2	-1	1	-2	-10	2	>5	-6	3	0	-10	0	0
1941	1	0	13	-1	-2	-8	4	6	-1	-7	16	12	-2	1
1942	3	2	5	-2	-1	<-10	3	5	-6	-2	17	14	0	0
1943	2	-5	-6	0	2	-11	-2	4	-7	-3	8	2	0	
1944	0	-1	-7	1	1	-8	-1	2	-9	-5	-7	>10	-1	0
1945	1	10	13	0	<1	0	4	4	12	14	16	12	∠1	∠1
1946	2	0	13	3	-2	0	5	4	-2	-2	17	13	0	1
1947	-1	-1	-9	3	2	-11	5	4	-10	-5	>14	14	-1	0
1948	0	0	13	1	-2	-10	3	5	-7	-5	>15	12	-1	0
1949	2	0	13	-1	0	∠-11	4	5	-6	0	17	15	2	-1
1950	0	-1	12	1	1	∠-12	6	5	-10	-5	15	14	-1	0
1951	1	0	13	1	0	-10	4	3	-7	-4	17	13	-1	0
1952	-1	-2	13	0	0	<-12	4	6	-7	-6	17	15	-2	-1
1953	-7	-5	-3	-4	-4	-11	5	4	-8	-11	0	6	-6	0
1954	2	1	12	-3	-2	-11	5	4	-9	-6	>16	13	-1	0
1955	-3	0	-10	-1	2	-10	6	6	-8	-5	>13	>13	-2	1
1956	1	0	13	-3	-2	-10	4	3	-9	-4	15	13	-4	0
1957	3	2	14	1	0	<-10	4	5	-8	-6	17	15	0	0
1958	0	-1	12	1	2	-6	4	6	-5	1	18	17	-1	0
1959	1	0	13	0	0	-10	4	6	-7	0	>18	15	-1	1
1960	-3	1	14	0	0	-11	5	6	-7	-7	>18	15	-3	0
1961	-5	0	13	-1	-2	-8	4	5	-9	-8	>15	>15	-4	0
1962	-1	-2	15	1	1	∠-9	5	4	-6	-5	19	18	-1	-1

YEAR	1949	50	51	52	53	54	55	56	57	58	59	60	61	62
1935	2	3	2	4	1	-1	1	1	2	7	2	6	5	7
1936	4	4	3	5	-1	0	-2	2	3	8	3	2	0	8
1937	4	4	3	3	10	2	21	2	4	2	3	2	0	4
1938	5	2	2	3	-2	4	-1	5	4	6	3	3	1	5
1939	4	2	3	3	-2	3	-4	4	5	5	3	3	2	5
1940	>5	>5	3	>5	-5	2	-2	2	>5	3	3	4	2	>5
1941	4	1	3	3	-7	0	-4	2	5	7	3	2	0	5
1942	4	3	5	2	-5	2	-3	4	5	6	2	2	0	7
1943	3	6	3	3	-9	3	-1	4	6	3	3	3	2	5
1944	-1	3	2	4	-13	2	-2	5	6	1	-2	5	3	6
1945	3	4	2	2	10	∠1	∠1	3	4	5	∠1	∠1	∠1	3
1946	2	2	3	1	1	1	∠-2	2	3	3	1	1	∠-2	1
1947	1	3	3	4	-1	1	-1	5	4	7	3	5	3	6
1948	5	3	3	4	-6	1	-3	2	5	7	2	3	0	7
1949	4	8	3	3	-7	1	-2	2	4	7	3	4	0	6
1950	-1	3	2	4	1	1	-3	1	5	7	2	0	2	4
1951	4	4	3	3	-7	2	-2	2	5	7	2	2	0	6
1952	4	2	3	3	-7	2	-2	2	5	7	2	2	0	6
1953	5	-4	4	4	-6	1	-6	2	5	8	3	5	0	5
1954	4	3	2	2	0	1	-4	2	4	8	1	1	1	6
1955	4	4	3	3	-2	3	-2	4	7	8	3	4	4	6
1956	4	4	3	3	-6	1	-4	2	5	8	2	2	0	5
1957	5	3	3	3	-6	2	-4	2	5	7	2	4	0	6
1958	4	3	3	4	-7	0	-3	1	5	7	2	5	-3	6
1959	4	4	4	3	-6	3	-2	3	6	8	3	3	0	8
1960	3	6	4	2	-4	3	-3	3	4	5	3	3	1	5
1961	4	1	3	3	-6	0	-6	2	5	10	3	2	0	4
1962	4	5	3	3	-6	1	-2	3	5	7	1	4	2	6

and 6. They indicate that shifting has adjusted the position of the annual storms so that the high hourly precipitations with high variances tend to cluster over the middle of the common storm duration. Apparently the shifted data have less, and more regular, variation than the unshifted data.

Storm shifting may not be an absolutely necessary step in the stochastic modelling of storm precipitation, but it generally improves the statistical models for the data; and because of the reduction in variance of the resulting model, it is to be preferred when sequential generation techniques are used in the analysis of the stochastic characteristics of storm precipitation.

STOCHASTIC MODELS FOR STORM PRECIPITATION

Let x_t be the t-th hourly rainfall in an annual storm with $t \equiv 1,2, \ldots m$ hours. Let ε_1, $\varepsilon_2 \ldots \varepsilon_m$ be mutually independent, random components of rainfall, i. e.,

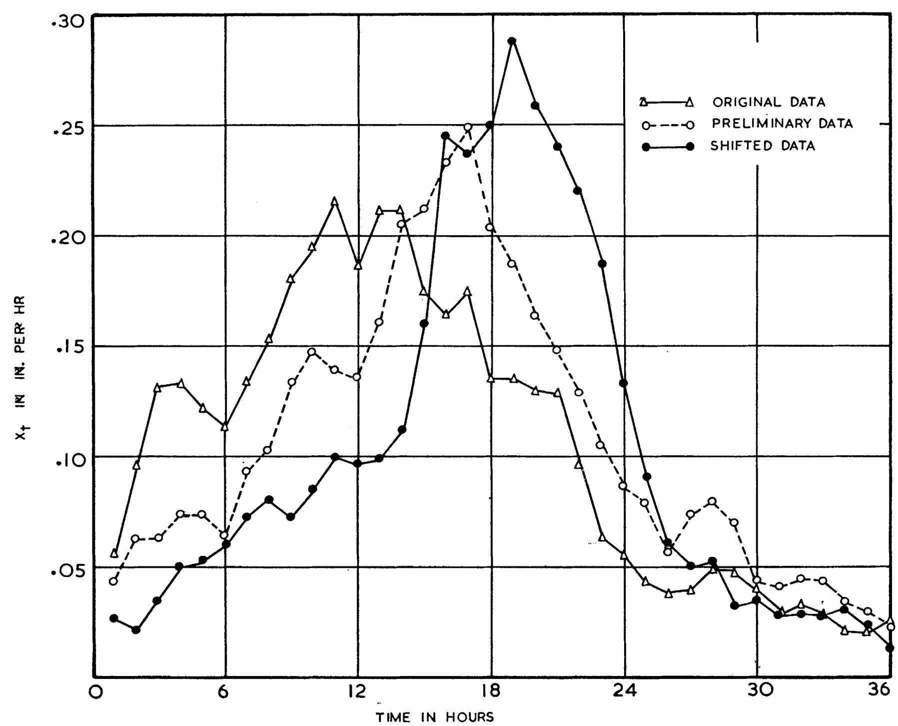

Fig. 4.
Variation of the mean hourly precipitation in an annual storm.

9*

ε_t is the random component of x_t ; x_t may be considered to be a function of the ε'_s, i. e.,

$$x_t = f(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_{t-1}, \varepsilon_t) \tag{3}$$

Generally this is assumed to be of a more specific form

$$x_t = f(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_{t-1}) + \varepsilon_t \tag{3a}$$

In this equation ε_t is the random component.

It follows that

$$x_1 \equiv \varepsilon_1 \tag{3b}$$

It is assumed that $f(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_{t-1})$ can be expressed in terms of x_t rather than ε_t. It is further assumed that the function f (.) of $(t-1)$ variables can be expressed as the sum of $(t-1)$ functions of one variable each and functions of more than one variable are not necessary. The assumptions lead to the relationship.

$$x_t \equiv f_{t,1}(x_{t-1}) + f_{t,2}(x_{t-2}) + \ldots + f_{t,t-1}(x_1) + \varepsilon_t \tag{4}$$

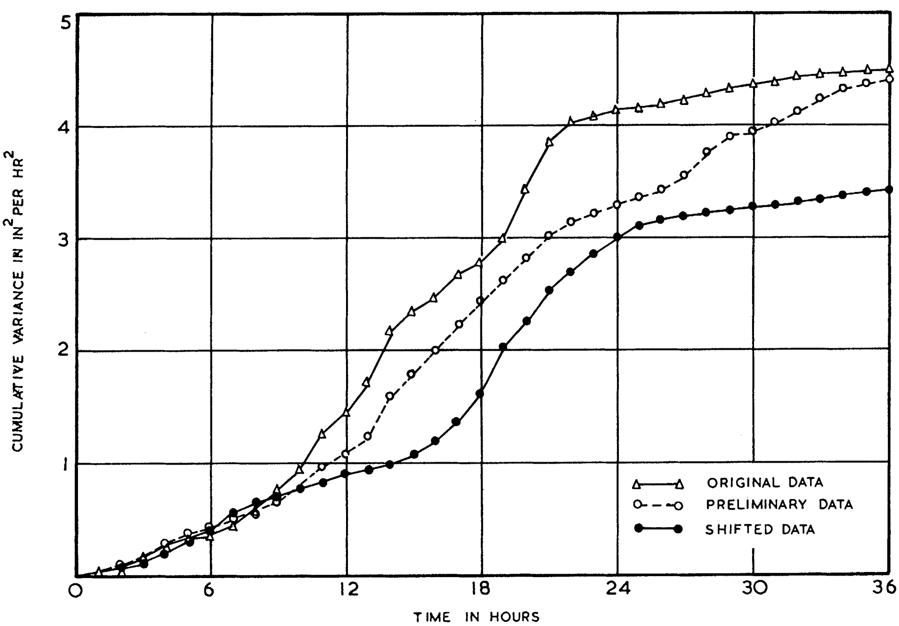

Fig. 5.

Cumulative variance of hourly precipitation in an annual storm.

In the absence of suitable expressions for the functions they may be replaced by polynomials of suitable degree and represented by the relationship

$$x_t = \sum_{j=1}^{t-1} \sum_{k=1}^{m_j} A_{tjk}\, x_j^k + \varepsilon_t \tag{5}$$

Several specific linear and nonlinear models of the form given above were tested.

Using the precipitation data for the French Broad River at Bent Creek, N.C., and the Kaskaskia River at Shelbyville, Illinois, the suitability of the different models was tested by regression analysis. The results indicate that precipitation at time t depends significantly (at the 5 % level) on the precipitation in the previous hour, but is independent of precipitation in the earlier hours considered

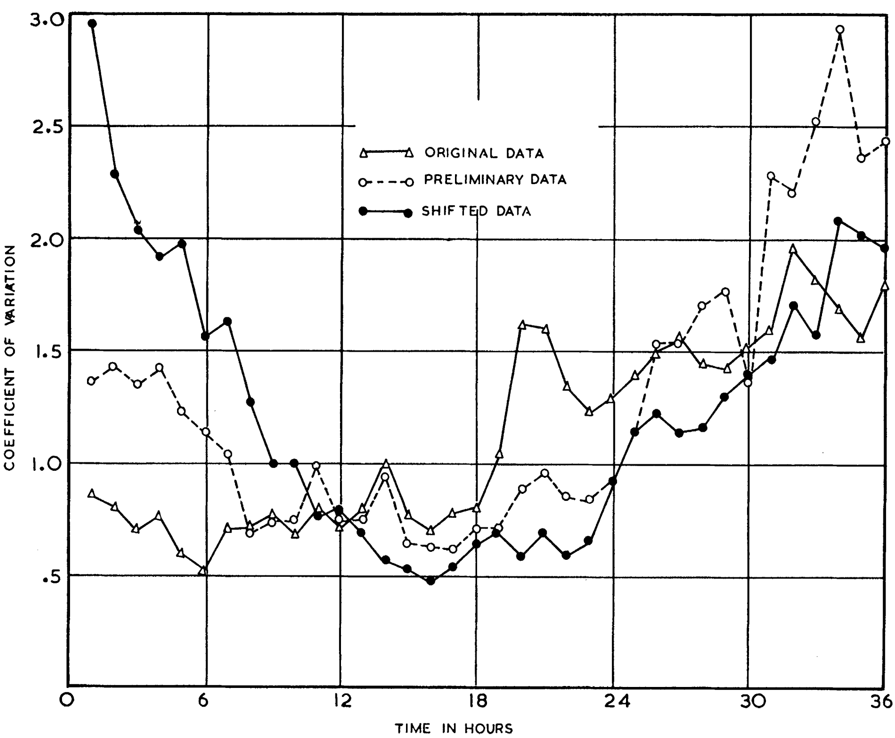

Fig. 6.
Coefficient of variation of hourly precipitation in an annual storm.

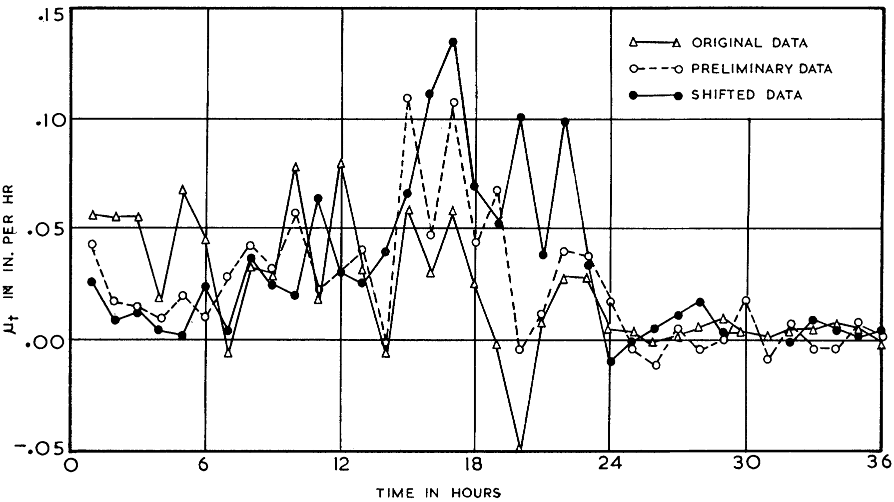

Fig. 7.
Variation of the mean value of the random component of hourly precipitation.

individually or collectively. The simplest model considered, viz., the Markov model,

$$x_t = r_t x_{t-1} + \varepsilon_t \tag{6}$$

was found to be consistently satisfactory for each t with $t = 2,3, \ldots m$.

The coefficient r_t is estimated for each value of t with $t \equiv 2,3 \ldots m$ by least squares linear regression of x_{t-1} on x_t. The trend component $r_t x_{t-1}$ can then be estimated and deducted from x_t to give the sample value of ε_t. For each value of t with $t = 2,3, \ldots m$, the sequence of values of ε_t are analysed for their first serial correlation coefficient. The result indicated that the component are random at the 5 % level. The random components ε_t are then analysed for their probability distribution. ε_t shows both positive and negative values and has a skewed distribution. An arbitrary correction k_t is added to it so that the corrected value ε' is always positive.

$$\varepsilon'_t = \varepsilon_t + k_t \tag{7}$$

x_t is then given by

$$x_t = r_t x_{t-1} - k_t + \varepsilon'_t \tag{7a}$$

Using the data for the two rivers, the mean and the standard deviation of the random components and the regression coefficients were studied for the initial, preliminary, and shifted data. They have shown that shifting has consistently and significantly improved the stability, regularity, and consistency of the proposed model (Figs. 7 and 8). Further the random components ε'_t consistently fit a log normal probability distribution (Fig. 9). The parameters of the distribution were estimated by the method of moments. After suitable correction the parameters for ε_t are shown in Table 3. Thus hourly storm precipitation can be considered as a finite, discrete-duration, quantitized-data, discrete-variable process. For the given basin it may be represented by a simple nonstationary Markov model with lognormally distributed, random components.

STOCHASTIC CHARACTERISTICS OF STORM PRECIPITATION

Several interesting stochastic characteristics of the storm precipitation process may be studied on the basis of the nonstationary Markov model which was found to be a satisfactory representation of the process of hourly storm precipitation in an annual storm.

Let ε_t, the random component of x_t, have a mean μ_t and a standard deviation σ_t. Total storm precipitation $\equiv \sum_{i=1}^{m} x_i$.

Expected total storm precipitation =

$$\sum_{i=1}^{m} \mu_i \left(1 + \sum_{n=i+1}^{m} \prod_{j=i+1}^{n} r_j\right) \tag{8}$$

$$= \sum_{i=1}^{m} \mu_i \,\mathrm{Sum}\,(i)$$

in which the cumulative coefficient for the mean

$\mathrm{Sum}\,(t) \equiv 1 + r_{t+1} + r_{t+1} r_{t+2} + \ldots$

$+ \ldots + r_{t+1} r_{t+2} \ldots r_m =$

$$1 + \sum_{n=t+1}^{m} \prod_{j=t+1}^{n} r_j \tag{8a}$$

Table 3.
Parameters and stochastic characteristics of annual storm precipitation in French Broad River Basin.

Time t	Random component ε_t		Markov regression coefficient	Cumulative coefficient		Variance due to ε_t	
	Mean μ_t (in/hr)	S. D. σ_t (in/hr)		Sum (T)	Variance Cum (T)	Cum (T) σ^2_t	Ratio of total variance
1	0.0260	0.0773	1.0000	5.4958	9.9915	0.0598	0.0296
2	0.0082	0.0295	0.5267	8.5362	16.0724	0.0406	0.0201
3	0.0127	0.0509	0.9944	7.5785	14.1569	0.0720	0.0357
4	0.0045	0.0244	1.3179	4.9917	8.9835	0.0847	0.0420
5	0.0018	0.0376	0.9930	4.0197	7.0394	0.0754	0.0374
6	0.0231	0.0597	0.7148	4.2243	7.4487	0.0673	0.0334
7	0.0036	0.0502	1.1449	2.8162	4.6324	0.0666	0.0330
8	0.0356	0.0717	0.6159	2.9488	4.8975	0.0519	0.0258
9	0.0247	0.0403	0.5969	3.2651	5.5302	0.0299	0.0148
10	0.0200	0.0572	0.8989	2.5198	4.0395	0.0309	0.0153
11	0.0636	0.0686	0.4191	3.6264	6.2529	0.0378	0.0187
12	0.0297	0.0576	0.6721	3.9076	6.8151	0.0412	0.0204
13	0.0259	0.0383	0.7513	3.8699	6.7399	0.0329	0.0163
14	0.0392	0.0408	0.8411	3.4123	5.8247	0.0298	0.0148
15	0.0657	0.0682	0.7771	3.1044	5.2089	0.0403	0.0200

16	0.1114	0.0926	0.8347	2.5213	4.0426	0.0564	0.0280
17	0.1363	0.1202	0.4096	3.7145	6.429	0.1079	0.0536
18	0.0688	0.1291	0.7629	3.5582	6.1165	0.1617	0.0802
19	0.0515	0.1339	0.9464	2.7031	4.4062	0.1833	0.0910
20	0.1014	0.1053	0.5436	3.1332	5.2664	0.1232	0.0611
21	0.0377	0.1217	0.7864	2.7126	4.4252	0.1295	0.0643
22	0.0979	0.0983	0.5124	3.3422	5.6845	0.0986	0.0489
23	0.0328	0.0866	0.6978	3.3566	5.7132	0.0911	0.0452
24	–0.01	0.0797	0.7628	3.0895	5.1790	0.0809	0.0402
25	0	0.0616	0.6777	3.0831	5.1663	0.0567	0.0281
26	0.0049	0.0389	0.6117	3.4056	5.8111	0.0326	0.0162
27	0.0108	0.0308	0.6609	3.6397	6.2794	0.0214	0.0106
28	0.017	0.0457	0.667	3.9574	6.9148	0.0249	0.0124
29	0.0032	0.0253	0.5730	5.1617	9.3235	0.017	0.0084
30	0.0035	0.0271	0.9869	4.2172	7.4343	0.0187	0.0093
31	0.0001	0.0136	0.8020	4.0113	7.0226	0.0126	0.0063
32	0.0012	0.0225	1.0482	2.8728	4.7457	0.0118	0.0058
33	0.0087	0.0296	0.6967	2.6882	4.3763	0.0091	0.0045
34	0.0061	0.0525	0.8711	1.9380	2.8760	0.0125	0.0025
35	0.0021	0.0164	0.6804	1.3787	1.7573	0.0040	0.0020
36	0.0044	0.0193	0.3787	1	1	0.0007	0.0003

The covariance of x_i with x_j is denoted by Covar (i,j). With $\bar{x}_i$ and $\bar{x}_j$ as the sample means of x_i and x_j respectively, they are estimated by Covar $(i, j) \equiv$

$$\frac{1}{N} \sum_{k=1}^{N} (x_{ik} - \bar{x}_i)\ (x_{jk} - \bar{x}_j)$$

in which N is the size of the sample. For the process of annual storm precipitation, the covariance matrix can be calculated (Ramaseshan 1964). Let cum (t) denote the cumulative coefficient for the variance of the total storm precipitation due to the random component at time t.

$$\text{Cum}\ (t) = 2\ \text{Sum}\ (t) - 1 \equiv \left[1 + 2 \sum_{n=t+1}^{m} \ \prod_{j=t+1}^{n} r_j \right] \tag{9}$$

The component of variance of the total storm precipitation due to ε_t, the random

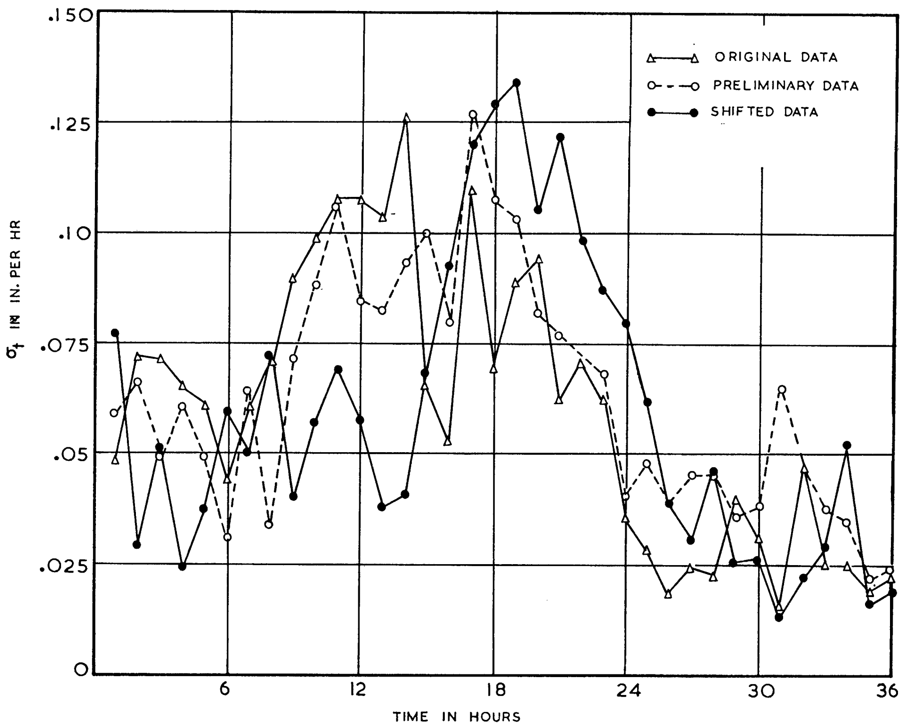

Fig. 8.

Variation of the standard deviation of the random component of hourly precipitation.

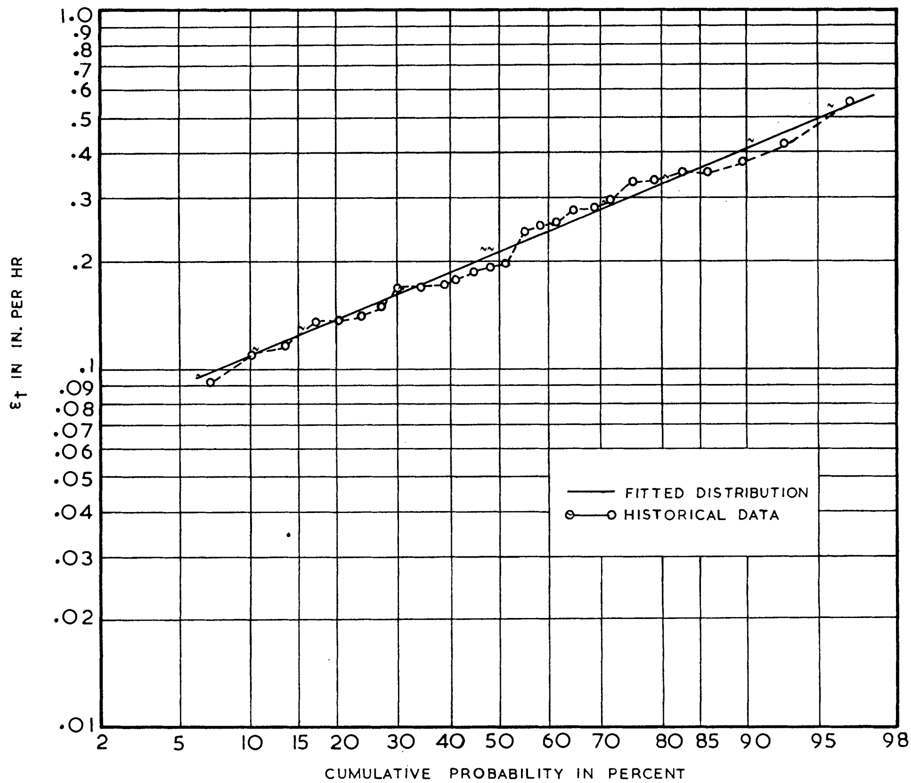

Fig. 9.
Cumulative probability distribution of the random component at time 17.

component at time $t \equiv \sigma_t^2$ cum (t). For the French Broad River Basin, the cumulative coefficients for the mean and the variance, sum (t), cum (t) and the component of the variance of total storm precipitation due to ε_t were calculated. The results are shown in Table 3.

The relative significance of the random components can be observed from the contribution of each random component to the mean and variance of the total storm precipitation. The time periods from 1 to 8 contributed about 26 % of the variance. Their contribution is high because of the large values of the cumulative coefficients for variance and the sum and moderate values of the standard deviation of the random components. The random components for the 10 hours from the sixteenth hour to the twenty-fifth hour contributed nearly

54 % of the variance. Their high contribution is due to the high variance of the random components and moderate values of the cumulative coefficients. The last ten hours contributed only 8 % of the variance, and the last 8 hours contributed only 4.4 % of the variance. For simplification they may be ignored and the process may be represented with $m = 28$. When complex stochastic characteristics of storms and floods are of interest, they may be determined by sequential generation and simulation procedures.

CONCLUSIONS

On the basis of the results from this study, the following conclusions may be derived.

1. Storm precipitation can be conceived of as a finite-duration, quantitized-data, continuous-variable, nonstationary, stochastic process. For the river basin considered, a simple Markov model with lognormally distributed, random components is found to be very satisfactory.

2. Storm precipitation, flood runoff, earthquakes, squalls, etc., are short-duration phenomena. For such processes, it may not be possible to identify the time when the phenomenon starts. Shift analysis is then found to be a useful tool in deriving a consistent, stable, and regular model for the process. Shift analysis is especially advisable when sequential generation and simulation techniques are used for analysing complex stochastic characteristics of the process.

3. Detailed steps in mathematical modelling of the processes can be standardized to yield consistent results.

4. From the parameters of the model fitted to the historical data, stochastic characteristics of the process may be analysed. The significance of precipitation at any time with respect to the total process can be evaluated. For any study, the significant duration of the process can then be determined. For the basin considered, a duration of 28 hours is found to be satisfactory.

5. Shift analysis is useful in generating sequential data of storm precipitation in a basin and subsequently verifying the data. Such data may be used in the analysis and design of water resources systems by simulation.

ACKNOWLEDGEMENT

The IBM 7094–1407 system, partially supported by grant GP-700 from the National Science Foundation, of the Digital Computer Laboratory and the facilities of the Statistical Services Unit, both of the University of Illinois, were used for most of the numerical analysis of this study. The author is very grateful to Dr. V. T. Chow, Professor of Civil Engineering, University of Illinois, for his valuable guidance and counsel when the study was conducted under his direction for the author's doctoral thesis. Some of the results were later revised, using the facilities of the Computer Centre, Indian Institute of Technology, Kanpur, India.

REFERENCES

Chow, V. T. (1964) *Handbook of applied hydrology.* McGraw Hill Book Co., Inc., New York.

Chow, V. T. & Ramaseshan, S. (1965) Sequential generation of rainfall and runoff data. *J. Hydraulics Div. Amer. Soc. Civil Eng. 91* (HY 4), 205–223.

Chow, V. T. & Ramaseshan, S. (1966) Closure to "Sequential generation of rainfall and runoff data". *J. Hydraulics Div. Amer. Soc. Civil Eng. 92* (HY 4), 162–165.

Le Cam, L. (1961) A stochastic description of precipitation. *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability,* ed. J. Von Neyman. University of California Press, Berkeley and Los Angeles, pp. 165–186.

Ramaseshan, S. (1964) A stochastic analysis of rainfall and runoff characteristics by sequential generation and simulation. Ph. D. thesis, University of Illinois, Urbana, Illinois.

Address:

Dr. S. Ramaseshan, Assistant Professor,

Department of Civil Engineering, Indian Institute of Civil Engineering,

IIT P.O.

Kanpur-16, India.

Received: January 1970.