International Symposium on
Hydrology of Mountainous Areas
Shimla (INDIA) May 28-30, 1992

Linear Perturbation Rainfall-Runoff Model for Western Ghats Basins

E.J. James*, K.E. Sreedharan*, S.G. Mayya and Govinde Gowde****
*GWRDM, Kozhikode 673 571
** KREC, Srinivasangar 574 157

SYNOPSIS

The Linear Perturbation Model (LPM), developed in University College, Galway, is applied to five river basins of the Western Ghats region in Kerala. The runoff values estimated using the model for larger basins compared well with the observed values in the calibration and verification modes. The LPM is used to reduce the dependence of linearity and increase the dependence on observed seasonal behaviour. The model has been applied in non-parameteric form; parameter optimization is carried out by the method of least squares. This model is acceptable for estimating the runoff values for the larger river basins of the Western Ghats region.

1.0 INTRODUCTION

1.1 For rivers with a relatively predictable seasonal variation, a considerable improvement on conventional rainfall-runoff models may be obtained by concentrating on the relationship between the series of departures from the seasonal behaviour in rainfall and discharge, as input and output respectively. Conventionally, the total rainfall series is treated as input and the total discharge series as output. The LPM may be viewed as an attempt to overcome the basic weakness inherent in classical unit hydrograph approach, wherein the departures (storm runoff) from the base flow are related by a linear model to the corresponding rainfall loss series (UCG,1985).

1.2 For LPM, the input to the model may be a single lumped input (either rainfall or upstream hydrograph) or multiple lumped inputs. In the present case, the LPM is used in the context of rainfall-runoff modelling with rainfall as the total daily input series (I_T) and the corresponding downstream runoff as the total daily output series (Q_T). The objective of this study is to apply the LPM for different selected river basins of Kerala and to establish the level of acceptability for these river basins.

2.0 MODEL DESCRIPTION

2.1 Denoting the total daily outflow series by QT, the elements of this series may also be specified by day and year as $q_{d,r}$ where d is the day and r the year. An estimate of the seasonal mean daily outflow series is then given by

$$q_d = \frac{1}{n} (q_{d,1} + \cdots\cdots q_{d,n}) = \frac{1}{n} \sum_{r=1}^{n} q_{d,r} \qquad \cdots\cdots (1)$$

for d = 1,2,...365
where n is the number of years in the calibration period. In practice, the estimated q_d is smoothed by discrete Fourier Series analysis. Since the true seasonal function is smooth by nature, the resulting smoothed seasonal mean daily outflow series (repeating every 365 days), constitutes the 'seasonal model forecast', Q_s. The output departure series, Q_D is then defined as

$$Q_D = (Q_T - Q_S) \qquad \ldots(2)$$

The outflow departure series Q_D has positive and negative values with zero mean value. Similarly, denoting the total daily input series by I_T and its corresponding smoothed seasonal, daily mean series by I_S, the outflow departure series is given by

$$I_D = (I_T - I_S) \qquad \ldots(3)$$

The seasonal smoothed components, I_S and Q_S, having been uniquely determined for the calibration period are then used directly in the verification period, with $(I_T - I_S) = I_D$ used as input to the linear element to obtain Q_D and thus the model forecast for the verification period $Q_T = (Q_D + Q_S)$. Implicit in the LPM structure outline above is the assumption that if, in a particular year, the total daily input series, I_T, coincides with the 'true' seasonal mean daily input series, I_S, then the total outflow series forecast, Q_T, would also coincide with the 'true' seasonal mean daily outflow forecast Q_S, .. that is with the seasonal model forecast.

2.2 The model (Nash and Bari, 1983) is based on the following assumptions:
(i) If in a particular year, each input function is equal, for each day of the year to its expected value for that date, the output will also equal its expectation for that date, ie, if the expected values of rainfall and discharge on each date d are obtained by id and qd respectively then, the input id produces the output qd:

$$(i_d \longrightarrow q_d)$$

(ii) Perturbations or departures from the date expected input values (i_d) are linearly related to the corresponding perturbations or departures from the date expected output values (q_d):

$$\{(i-i_d) \dashrightarrow (q-q_d)\}$$

For a single input series, the LPM may be described by

$$Y_i = \sum_{j=1}^{m} X_{i-j+1} h_j + e_i \qquad \ldots(4)$$

$i=1,2,\ldots\ldots\ldots,n;\ d=1,2,\ldots\ldots\ldots\ldots,365$

where $y_i = q_i - q_d$; $x_i = I_i - I_d$.

In the context of linear input - output model, the equation is

$$Y_n = \sum_{j=1}^{m} X_{n-j+1} h_j + e_n \qquad \ldots\ldots(5)$$

where Y_n is the output (runoff), X_n the input (rainfall), e_n the error, h_j the j_{th} ordinate of the pulse response function, m the memory length which implies that the effect of any input X will be lost for m days (365 element of q_d and i_d series being repeated for each year of the record).

2.3 The linear element of the LPM in the form of non-parametric unconstrained linear model, defined by the convolution summation relation, is

$$Y_t = h_1 X_t + h\, x_{t-1} + \ldots\ldots + h_m x_{t-m+1} + e_t \qquad \ldots(6)$$

In the matric form,

$$Y = Xh + e \qquad \ldots \quad (7)$$

where Y is the vector of Q_D element, the matric x the band matrix of the discrete pulse response series, e the vector of residual model errors and m the memory length. The ordinary least squares estimate of the h vector is

$$\hat{h} = [X^T X]\, X^T Y \qquad \ldots(8)$$

The variance of the vector, h, as given by

$$(h) = [X^T X]^{-1} \sigma^2 \qquad \ldots(9)$$

where an unbaised estimate of σ^2, the variance of the residuals, e_t, is given by

$$\sigma^2 = \frac{1}{N-2m+1} \sum_{i=m}^{n} e_i^2 \qquad \ldots(10)$$

The variance of the elements h_i, namely var (hi), is the product of σ^2 and the corresponding ith element of the principal diagonal of the $[X^T X]^{-1}$ matrix (N is the total number of the elements in the set) For example, if $[X^T X]^{-1}$ is denoted by M having elements m_{ij}, then Var (hi)= $m_{ij} \sigma^2$. Thus the standard error of elements of the h_i element is

$$SE\ (h_i) = \sqrt{Var\ (h_i)} = \sigma\ m_{ij} \quad \ldots(11)$$

2.4 In many perennial streams, discharge continues even for months after the cessation of rainfall. The time interval between the occurrence of the rainfall and the time when its effect on the stream flow finally ceases, is known as the memory length, when the operation is represented by the equation:

$$Y_i = \sum_{j=1}^{m} X_{i-j+1} h_j + e_i \quad \ldots.(12)$$

where m is memory length of the system and this must be estimated prior to the formulation of equation. The need to develop a completely objective procedure for the estimation of the memory length is not essential. The relationship of each of the last few ordinates to their respective standard errors may be used as a guide to the memory length, which may be taken as ceasing when the estimated ordinates are not significantly different from zero.

2.5 Due to wide fluctuations in the observed rainfall and runoff data, it is desirable to smooth both rainfall and runoff series by Harmonic Analysis for the calibration period.The choice of the number of harmonics to use is admittedly somewhat subjective, usually based on the variance accounted for by these harmonics, or simply by adopting the first four harmonics.

2.6 Criteria which express model accuracy are generally linked with the objective used for optimising its parameters. A commonly used objective function even for nonlinear or conceptual models is to minimise the sum of squares of differences between the observed (y) and the estimated (y) discharges, when the summation is taken over the whole of the calibration period:

$$F = \sum (y-\hat{y})^2 \quad \ldots(13)$$

where F is an index of residual error which reflects the extent to which a model is successful in reproducing the observed phenomenon. Nash and Sutcliffe (1970) defined R^2 analogous to the coefficient of determination as one minus the proportion of the initial variance represented by F. The initial variance, Fo, may be defined as

$$F_o = \sum (y-\bar{y})^2 \quad \ldots(14)$$

where $\bar{y}$ is the average estimated discharge and

$$\bar{y} = \frac{1}{n} \sum_{i=i}^{n} y_i \qquad ...(15)$$

where n is the number of data points and

$$R^2 = \frac{F_o - F}{F_o} \text{, lies between 0 and 1} \qquad ...(16)$$

R^2 in this case is identical to the coefficient of determination in the liner regression analysis and varies between zero and one. When applied in the verification period, the initial variance Fo is calculated as the sum of squares of deviation from the mean of the calibration period, as the criteria attempts to compare the sum of squares of model errors which would occur when, in the absence of any model (the no model situation) the only forecast which could be made for the verification period would be the mean value of the variable in the calibration period. As a result, R^2 may take negative values in the verification period, when the model test produces forecasts which are worse estimators than is the mean of the calibration period. The error term (e_i) is calculated as

$$e = \sum_{i=1}^{N} (Q_o - Q_c)^2 \qquad ..(17)$$

where Q_o is the observed discharge, Q_c the computed discharge, N the number of hydrograph ordinates. To minimise the error, adjustment can be made in unit hydrograph ordinates to give the best solution. Since the square of the error in the observed discharge Q_o and the computed discharges Q_c is minimised, it is called the least square method.

3.0 APPLICATION, CALIBRATION AND VERIFICATION

3.1 Daily rainfall and corresponding runoff data were available from the Chaliyar and the Kallada river basins for 10 years, from the Manimala and the Meenachil for 8 years and from the Chandragiri for 9 years. The areas of the catchments and the data depth used for calibration and verification are given in Table 1.

Table 1. Rainfall and runoff data used

Catchment	Area (km^2)	Calibration period	Verification period
Chaliyar	1580	1973-80	1981-82
Kallada	1196	1974-81	1982-83
Manimala	816	1972-77	1978-79
Chandragiri	570	1975-81	1982-83
Meenachil	430	1972-77	1978-79

3.2 This model given by equation (4) was applied as follows:
(i) Seasonal mean rainfall and seasonal mean discharge were calculated for the periods of calibration and smoothing was done by the method of unconstrained Fourier analysis using the first four harmonics;
(ii) The smoothed seasonal mean values were subtracted from the observed rainfall and discharge series for the period of calibration, to yield the time series of the perturbations x and y;
(iii) The pulse response function was estimated by the method of ordinary least squares;
(iv) The pulse responses were convoluted with the rainfall perturbations to obtain the estimated outflow perturbation series;
(v) The estimated discharge series was calculated by adding the seasonal mean discharge to the estimated outflow perturbation series; and
(vi) The differences between observed and computed discharges were squared and summed; the efficiency R^2 is calculated.

A computer program developed at University College, Galway, was made use for the computations.

3.3 The optimum value of memory length (m) is chosen by trial and error method. The efficiency of the model for verification period is verified by the efficiency criteria.

4.0 RESULTS AND DISCUSSIONS

4.1 The results of fitting the pulse response function for unconstrained LPM are presented in Tables 2-6. The pulse response function for the Chaliyar basin is shown in Fig 1; this is estimated by the method of least squares.

Table 2. Pulse response function - Chaliyar basin

Memory length (days)	1	2	3	4	5	6	7	8	9
Pulse response 1/day	0.256	0.362	0.133	0.070	0.019	0.046	0.031	0.018	0.037
Estimated Error	0.013	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016

Table 3. Pulse response function - Kallada basin

Memory Length (days)	1	2	3	4	5	6	7	8
Pulse response 1/day	0.182	0.223	0.185	0.163	0.126	0.073	0.024	0.024
Estimated Error	0.015	0.016	0.016	0.016	0.016	0.016	0.016	0.015

Table 4. Pulse response function - Manimala basin

Memory length (days)	1	2	3	4	5	6	7	8
Pulse response 1/Day	0.156	0.171	0.176	0.169	0.133	0.076	0.059	0.059
Estimated Error	0.016	0.017	0.017	0.017	0.017	0.017	0.017	0.016

Table 5. Pulse response function - Chandragiri basin

Memory length (Days0	1	2	3	4	5	6
Pulse response 1/Day	0.427	0.278	0.105	0.108	0.054	0.007
Estimated Error	0.031	0.033	0.033	0.033	0.033	0.033

Table 6. Pulse response function - Meenachil basin

Memory length (days)	1	2	3	4	5	6	7	8	9
Pulse response 1/Day	0.256	0.208	0.146	0.104	0.084	0.076	0.047	0.053	0.025
Estimated Error	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013

4.2 For the Chaliyar basin, the optimum memory length is 10 days, ie, effect of rainfall on the stream flow finally ceases within 10 days; the effect is more on the second day and gradually decreases as indicated in Fig 1. For the Meenachil basin, the effect of rainfall on the stream flow is more on the first day and then it gradually decreases.

4.3 The model gives good results during calibration and verification periods for basins with larger areas (Table 7). For example, for the Chaliyar basin with 1580.00 sq km area, the efficiency during calibration and verification was of the order of 80.50% and 73.58% respectively. In the case of the basin with smallest area, namely Meenachil (430.00 sq km), the efficiency in calibration and verification period was of 58.07% and 20.03% respectively.

There has been a uniform reduction in efficiency with decrease in the size of the basin. A comparison of observed and estimated discharges for the Chaliyar is given in Fig 2. The estimated and observed runoff values compare well for the larger

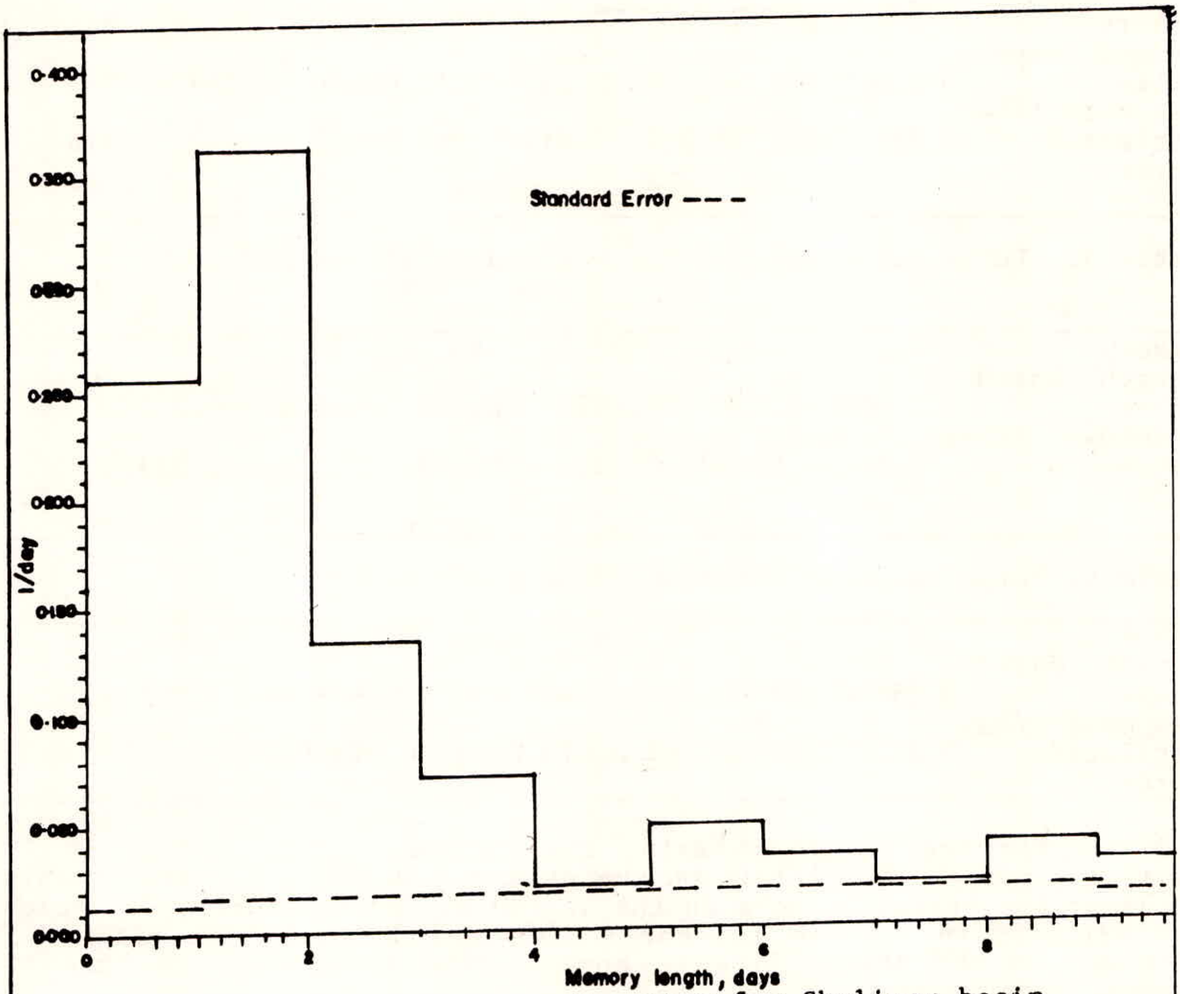

Figure 1. Pulse response function for Chaliyar basin

basins studied namely, the Chaliyar, the Kallada and the Manimala.

Table 7. Comparison of model efficiency (R^2%)

Name of catchment	Area (km^2)	Memory length (m)	Calibration period	Verification period
Chaliyar	1580	10	85.50	73.58
Kallada	1196	8	60.07	55.95
Manimala	816	11	58.20	42.34
Chandragiri	570	6	58.09	39.64
Meenachil	430	9	60.93	20.03

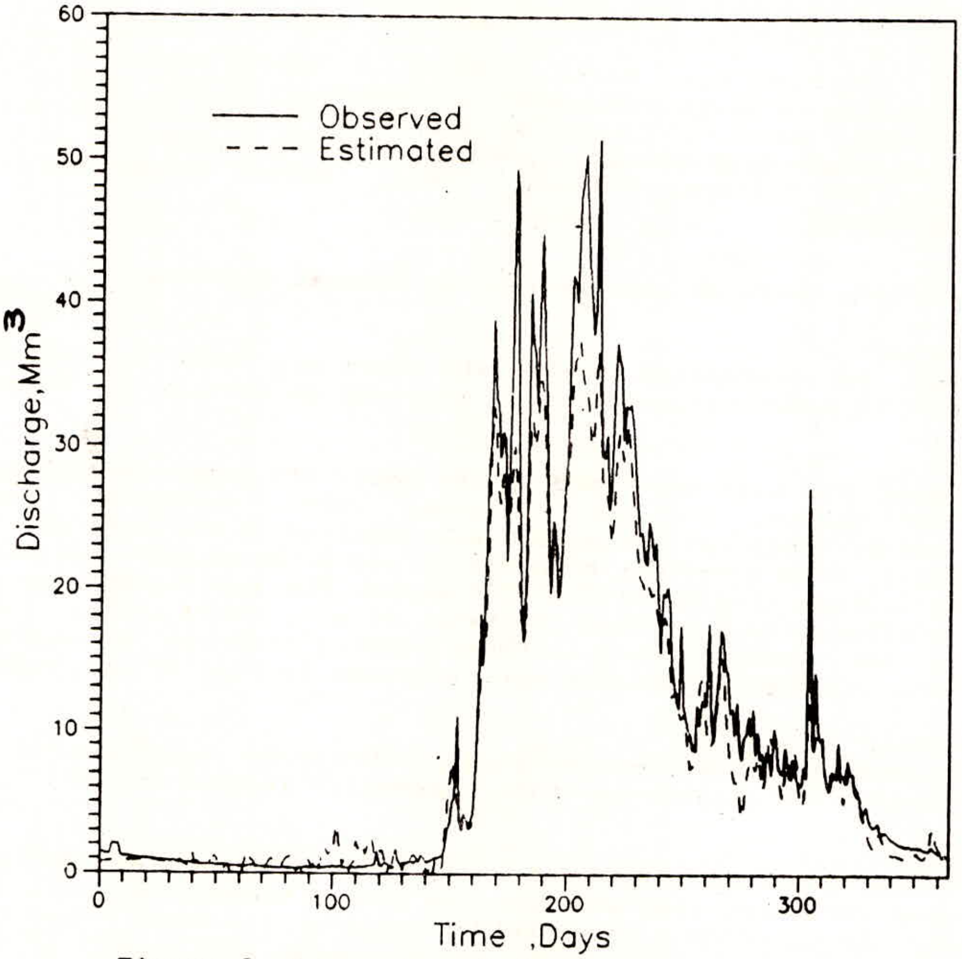

Figure 2. Comparison between estimated and observed runoff - Chaliyar basin

4.4 Comaprison of daily mean runoff values for calibration and verification periods are given in Table 8. The daily mean observed and estimated values for calibration and verification periods are equal for all the basins, except for the Meenachil.

Table 8. Comparison of daily mean discharge - calibration and verification mode

Name of catchment	Daily Mean Discharge (Mm ^{3})			
	Calibration		Verification	
	Observed Series	Estimated Series	Observed Series	Estimated Series
Chaliyar	0.6295	0.6298	0.6420	0.5455
Kallada	0.3246	0.3248	0.2069	0.2206
Manimala	0.3749	0.3747	0.2337	0.3637
Chandragiri	0.2107	0.2008	0.1626	0.2024
Meenachil	0.6693	0.6696	0.4525	0.7176

REFERENCES

1. Nash, J.E. and B.I. Bari, (1983), 'A hybrid model for flow forecasting on large catchment', J. hydrol., Vol.65, pp 125-137.
2. Nash, J.E. and J.V. Sutcliffe, (1970), 'River flow forecasting through conceptual models, J. Hydol., Vol.10, P.283
3. University College, Galway, (1985), Lecture notes on flood forecasting.

Comments by authors on observations by General Reporter

Observation 1
It would be interesting to know why there has been a uniform reduction in model efficiency with decrease in size of the basin.

Authors
This model has been developed in the Dept. of Engg. Hydrology, Univ. of Galway, Ireland, (Kachroo, 1988). Workshops were organised in which various models were applied to the data of 14 catchments, from a variety of climatic and geographical regions, ranging in size from 19 to 1,31,500 sq km. The results of some of these analyses are published (Kachroo, et al., 1988, 1992a, 1992 b). It has been observed in these studies that is a uniform reduction in model efficiency with decrease in size of the basin.

Observation 2
Authors have used different sets of calibration ......Even Fig. 1 furnished by the authors is quite suggestive in this regard.

Authors
The General Reporters' views are well taken. We could not do this since data were not available for the period mentioned. It is not possible to develop a completely objective procedure for the estimation of the memory length. Little damage will result if the chosen length is somewhat greater than the actual. Small or zero values will be obtained for the values beyond the actual memory length. However, an underestimation of m is likely to produce h values which fail to decrease monotonically towards zero, (Kachroo, 1992 b).

References
Kachroo, R K (1988). Rainfall runoff and flood routing models for use as forecasting tools with special reference to South Asia. Journal of Applied Hydrology, Vol. 1. No. 1, pp. 63-91.
Kachroo,R K and Liang, C G (1992 a). River flow forecasting part 2. Algebraic development of linear modelling techniqaues. Journal of Hydrology, Vol.133, No.1-2, pp 17-40.
Kachroo,RK, Sea,CH, Warsi,MS, Jemeneez,H and Saxena,RP, (1992 b). River flow forecasting. Part 3 Application of linear techniques in modelling rainfall-runoff transformations. Journal of Hydrology, Vol.133, no.1-2, pp. 41-47.
Kachroo, R K, Liang, C G and O'Connor, K M (1988). Application of the linear perturbation model (LPM) to flood routing on the Mekong river. Hydrological Sciences Journal, Vol.32, No. 2, pp 193-214.