REGRESSION BASED RAINFALL-RUNOFF MODELLING

INTRODUCTION

Many factors affect the runoff depending upon the period of its determination. Some of these factors are interdependent. These factors can be classified as (i) meteorologic factors, and (ii) watershed factors. Space-time distribution of precipitation amount, intensity and duration, and space time distribution of temperature are some of the important meteorological factors. Some important watershed factors include surface vegetation, soil moisture, soil characteristics, surface topography, and drainage density. In addition to these other factors, which include pondage of artificial reservoirs, diversion of water to the neighbouring basin or within the same basin to fulfil the water demands, cultivation and change in land use practices such as afforestation, deforestation or urbanization, etc., also influence the runoff considerably.

Determination of runoff is required for solution of a number of water resources problems. Prominent among them are:

(i) Design of storage facilities;
(ii) Determination of minimum amounts of water available for agricultural, industrial or municipal use;
(iii) Estimation of future dependable water supply for power generation under varying patterns of rainfall;
(vi) Planning irrigation operation;
(v) Design of irrigation projects, etc.

There are several approaches to determine the inflow. Most of these approaches can be broadly classified in two groups:

(a) Statistical and stochastic approaches;
(b) Deterministic approaches, which may be further classified in two sub groups:
    (i) Empirical approaches
    (ii) Watershed Modelling approach

In the present lecture, methodology for the development of regression based rainfall-runoff models has been discussed. Normally, larger the time period, simpler the determination. The time period of interest is generally equal to storm duration, a day, a month, a season, or a year.

VOLUMETRIC RAINFALL-RUNOFF RELATIONSHIP USING REGRESSION BASED APPROACHES

Volumetric rainfall-runoff relationships over the time periods of the day, month, season and year may be developed using the statistical and stochastic approaches. The development of daily rainfall-runoff relationships using this approach poses some difficulties. However, the problem of relating long term, say monthly, seasonal, or annual volumes of rainfall and runoff is relatively easier. Over a larger period of time, the averaging of a variety of rainfall storms tends to minimise the effect of rainfall intensity and antecedent moisture conditions on the volumetric relationship. Indeed in many cases a simple plot or linear relation may be adequate to define the relationship between annual volumes of rainfall and runoff if the water year is properly selected. In order to develop monthly, seasonal, and annual rainfall runoff relationships linear or non linear regression analysis may be carried out in different forms to relate the runoff with rainfall over the selected time periods and/or some other characteristics. Note that the records of rainfall-runoff used for developing such relationships should be homogeneous. In case some major man made changes occur in the catchment two different relationships must be accomplished:

(i) Relationship between the rainfall-runoff prior to the man made changes; and
(ii) Relationship between the rainfall-runoff after the man made changes.

In order to detect changes in hydrologic response of a watershed the hydrologists generally examine the mass curves for changes in slope. A mass curve is a plot of the accumulation over time of one variable versus the accumulation over time of a second variable. The time period usually selected for such computations is an year.

For developing the above relationships adequate record lengths are needed. In case the records are inadequate for any of the above two relationships, a single relationship may be developed relating the runoff with rainfall together with the factors representing the effects of the man made changes. Step wise regression may be performed to arrive at the suitable form of the rainfall runoff relationship.

Daily Rainfall-Runoff Relationship

Nash and Barsi (1983) developed a model which relates daily rainfall with daily runoff. The model, originally developed for daily flow forecasting on larger catchments exhibiting seasonality, may also be applied to estimate daily flow corresponding to given daily rainfall values. In the model it was assumed that in a year in which the rainfall on each day is the exactly the seasonal mean for that day, i_d , the corresponding discharges would also agree with their seasonal means, q_d. Hence:

$$i_d \rightarrow q_d \tag{1}$$

An attractive hypothesis was made considering the departures of the rainfall and the discharge from these seasonal means linearly related in any particular year:

$$i - i_d \rightarrow q - q_d \tag{2}$$

$$\text{or } x - y \tag{3}$$

where, $x = i - i_d$

$$y = q - q_d \tag{4}$$

For testing the hypothesis of linearity in the relationship of eq.(3), the values of i_d and q_d can be obtained by averaging the rainfall and the discharge records for each date d of over the years in the period of calibration, and smoothing by Fourier analysis. The seasonal values of i_d and q_d may be subtracted from the actual values of i and q on each day in order to obtain the departure series for x and y. Thus, the input and output series for x and y of length equal to the number of days in the calibration period are obtained.

Assuming that a general linear relationship with a memory length m exists between the x and y series, as obtained, it may be expressed as a linear multiple regression of y on the m previous x-values as independent variables.

$$y_i = h_i x_i + h_2 x_{i-1} + \ldots\ldots + h_m x_{i-m+1} \tag{5}$$

where, h = the vector of regression coefficients which represent the discrete series of pulse response.
u_i = the disturbance term.

Eq.(5) can also be expressed as:

$$y_i = \sum_{j=1}^{m} h_j x_{i-j+1} + u_i \tag{6}$$

Vector of h values, which are unknown, are estimated by method of least square after minimising the sum of error squares.

The standard errors for the estimates of h can be obtained using the following equation:

$$Se(\hat{h}) = \sqrt{V^{-1}S^2} \tag{7}$$

Vector of h values, which are unknown, are estimated by method of least square after minimising the sum of error squares.

where, $\hat{h}$ = vector of h values.
$Se(\hat{h})$ = standard error of vector, $\hat{h}$
$V = x^T x$ (8)

$$X = \begin{bmatrix} X_1 & 0 & \cdots\cdots & 0 \\ X_2 & X_1 & \cdots\cdots & 0 \\ X_3 & X_2 & \cdots\cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ X_m & X_{m-1} & \cdots & X_1 \\ X_n & X_{n-1} & \cdots & X_{n-m+1} \end{bmatrix} \tag{9}$$

and S is an unbiased estimate which is given by:

$$S = \sqrt{\sum_{i=1}^{n} u_i^2/(n-m)} \tag{10}$$

here n is the number of daily rainfall or runoff values during the calibration period.

The variance of $\hat{h}_i$ may be obtained by taking the i^{th} term of the principal diagonal of V^{-1} and multiplying by S^2. The standard error of h_i is the square root of its variance. It generally indicates the firmness in the estimation of H.

Having obtained the regression coefficients $\hat{h}$, the y values can be obtained using the following equation:

$$y_i = \sum_{j=1}^{m} h_j X_{i-j+1} \quad \textbf{(11)}$$

Finally the seasonal mean q_d is added with y values to give the estimates for $\hat{q}$ values. The difference between the observed and computed q values provides a series of residual errors for the calibration period. The series of residual errors may be analysed to identify the following persistence structure:

$$e_i = b_1 e_{i-1} + b_2 e_{i-1-1} + b_3 e_{i-1-2} + \ldots\ldots. \quad b_n e_{i-1-n+1} + E_i \quad (12)$$

where, l represent lag period to be identified from the analysis.

$b_1, b_2, \ldots, b_n$ are the regression coefficients to be obtained from least square analysis and E_i is the random component of mean zero and standard deviation 1. The estimated daily flow using eq.(11) are updated for the residual errors obtained from eq.(12).

Monthly Rainfall-Runoff Relationship For Gauged Catchments

In India more than 95% of the annual rainfall is received in monsoon season (normally from June to October). Thus, the rainfall-runoff relationships for monsoon months may be developed using linear rainfall-runoff model. However, during non-monsoon months (Nov-May) most of the runoff appears in the stream as a contribution of ground water reservoir towards stream, i.e. baseflow and contribution of the rainfall is almost negligible during this period. To the same extent few occasional thunder storms may contribute to the stream during non monsoon season. For partially fed basins melt runoff constitutes a part of the stream runoff. At the time of developing the monthly rainfall runoff relationships, it is necessary to identify the monsoon months for the study area as well as type of the basin i.e. fed or rain fed. If the basin is partially fed and partially rain fed, then monthly water equivalents are needed in addition to monthly rainfall data. The form of monthly rainfall-runoff relationships are given below for different conditions.

Monthly rainfall-runoff relationships

(a) Monsoon Months.

I. The simplest expression for runoff from a catchment, in terms of depth of water, is of the form:

$$RO_m = a\,(P_m - I_{am}) \quad (13a)$$

$$RO_m = a\,P_m - a\,I_{am}) \quad (13b)$$

$$RO_m = a\,P_m + b \quad (13c)$$

In the above equations, RO_m represent the runoff for a specific month, P_m is the rainfall for that month and I_{am} represent the initial abstraction of the specific month rainfall which does not become runoff. The coefficient a is the regression coefficient that scale the rainfall to the runoff. The coefficient b, which is also obtained from linear regression equals to $-aI_m$ knowing the values of a and b, interception loss for that specific month can be determined. The form of the relationship given by eq.(13c) is valid for small catchments wherein the contribution of the rainfall appears at the outlet of the catchment within the day.

II. The expression for runoff from large size catchment, in terms of depth of water, may be given in the following form:

$$RO_m = b_1 (P_m - I_{am}) + b_2 (P_{m-1} - I_{am-1}) \tag{14}$$

In eq.(14), RO_m is runoff for a specific month, P_m and P_{m-1} are precipitation in the specific month and a month prior to that month respectively, I_{am} and I_{am-1} represent the initial abstractions of the specific month rainfall and from the rainfall in the month prior to the specific month. The coefficients b_1 and b_2, of course, are the regression coefficients that scale the rainfall to the runoff. Eq.(14) may be expanded to:

$$RO_m = b_1 P_m + b_2 p_{m-1} - b_1 I_{am} - b_2 I_{am-1} \tag{15}$$

If a is substituted for the term $-(b_1 I_{am} + b_2 I_{am-1})$, eq (15) converts to:

$$RO_m = a + b_1 P_m + b_2 P_{m-1} \tag{16}$$

where $a = -(b_1 I_{am} + b_2 I_{am-1})$

The threshold values of I_{am} and I_{am-1} can not be determined exactly. They can only be determined if their relative values are known. For example, assuming $I_{am} = I_{am-1}$, the value of I_{am} may be estimated as:

$$\frac{a}{(b_1 + b_2)}$$

II. In addition to the above the forms of the monthly rainfall-runoff relationships, which could be tried, are given below:

$$RO_m = a + b_1 P_m + b_2 RO_{m-1} \tag{17}$$
$$RO_m = a P^b{}_m \tag{18}$$
$$RO_m = a (P_m - I_{am})^b \tag{19}$$
$$RO_m = a\, p_m^{b_1}\, RO_{m-1}^{b_2} \tag{20}$$

It is to be noted that a prior estimate for the initial abstraction, I_{am}, is necessary to develop the relationship of the form given by eq.(19) wherein only those records can be utilised that result the values of $(P_m - I_{am})$ greater than zero. Similarly, while developing the relationship of the form given by eq.(20), those records must be excluded which have P_m values equal to zero. Thus, the scope of developing the monthly rainfall-runoff relationships in the form given by eq.(19) and (20) are somewhat limited.

IV. In order to make accurate projections, it may be necessary to use a time-distributed model of the form:

$$RO_m = \bar{RO}_m + b\,(P_m - \bar{P}_m) \tag{21}$$

In the above equation, it is necessary to estimate the value of b for each month. It requires sufficient data for calibrating the coefficient b for each time period. Here $\bar{RO}_m$ and $\bar{P}_m$ represent the monthly mean runoff and precipitation respectively for the specific month. A time distributed model in the following form can also be used for making an accurate estimation of runoff particularly for large size catchments:

$$RO_m = \bar{RO}_m + b_1 (P_m - \bar{P}_m) + b_2 (P_{m-1} - \bar{P}_{m-1}) \tag{22}$$

(b) Non-monsoon months

During non-monsoon months the contribution of runoff resulting from the precipitation may not be that predominant. Therefore most of the relationships may be developed involving the runoff of the previous months. However, some relationships could be tried retaining the precipitation term in the equation and testing its significance in statistical sense before arriving at the definite conclusions about the form of the relationships for non-monsoon months. The possible forms of relationships which could be tried are:

$$RO_m = a + b\, RO_{m-1} \tag{23}$$

$$RO_m = \bar{RO}_m + b\, (RO_{m-1} - \bar{RO}_{m-1}) \tag{24}$$

$$RO_m = \bar{RO}_m + b_1 (RO_{m-1} - \bar{RO}_{m-1}) + b_2 (RO_{m-2} - \bar{RO}_{m-2}) \tag{25}$$

$$RO_m = a + b_1\, RO_{m-1} + b_2\, RO_{m-2} \tag{26}$$

$$RO_m = a + b_1\, P_m + b_2\, RO_{m-1} \tag{27}$$

$$RO_m = a\, (RO_{m-1})^b \tag{28}$$

The relationships for non-monsoon months can also be developed based on non-monsoon flows and annual flows, computed using available data of monthly flows. Non-monsoon flows (RO_{NON}) is usually taken as total of runoff values for seven non-monsoon months within a year. Total runoff for twelve months of a year represents annual flow (RO_{AN}). Two relationships may be obtained in the following steps:

(i) Develop the following relationship between RO_{NON} and RO_{AN}:

$$RO_{NON} = K\, (RO_{AN}) \tag{29}$$

The value of constants K may be obtained as a ratio of average non-monsoon flow to average annual flow for a site.

(ii) Distribute non-monsoon flows, RO_{NON}, in each of seven months using the following form of relationships:

$$RO_m = K_i\, (RO_{NON}) \tag{30}$$

The value of K_i for each of seven months may be evaluated as a ratio of average monthly flow for the concerned month to average non-monsoon flow for particular site.

Monthly rainfall runoff relationships for ungauged catchments

The runoff records are not available for an ungauged catchment. For such catchments it is not possible to develop the monthly rainfall-runoff relationships using the methodology discussed above. It involves the regionalization of the regression coefficients estimated for different gauged catchments of a hydro-meteorologically homogeneous region. intercept component of the regression equation may be related with the physiographic characteristics of the catchments. However, the regional values of the slope components may be determined taking their median values from different gauged catchments. The step by step procedures to develop the regional monthly rainfall-runoff relationships are explained taking the following form of relationship:

$$RO_m = \bar{RO}_m + b(P_m - \bar{P}_m) \qquad (31)$$

Step (i) : Identify the hydrometeorologically homogeneous region wherein the ungauged catchment is located.

Step (ii): Analyse the monthly rainfall-runoff records of all the gauged catchments in the region and develop monthly rainfall-runoff relationships for them in the form given by eq.(31) for different months.

Step (iii): For a specific month relate the intercept term with the physiographic characteristics of the catchments such as area, length, and drainage density, etc.

Step (iv): For the same month, find out the median value of b from its estimates obtained for different gauged catchments or take the value of b for a catchment having almost similar hydrologic characteristics as that of the ungauged catchments.

Step (v): Repeat step (iii) and (iv) for different months in order to derive the regional monthly rainfall runoff relationships for each month.

Step (vi): For an ungauged catchment, estimate RO_m using the relationship developed at step (iii) for the specific month.

Step (vii): Derive the monthly rainfall-runoff relationship in the form of eq.(31) for the ungauged catchment putting the value of $\bar{RO}_m$ obtained from step (vi) and median value of b obtained from step (iv) for the specific month.

Step (viii): Repeat step (vi) and (vii) for different months.

Seasonal Rainfall Runoff Relationships

A water year consists of twelve months starting from Ist June of the current calendar year upto 31st May of the next calender year. The sum of the runoff for five months i.e. June, July, August, September and October represent monsoon runoff (RO_{MON}). Total runoff value for seven non-monsoon months, i.e. November, December, January, February, March, April and May is considered to be non-monsoon runoff (RO_{NON}), as discussed earlier. Thus the sum of monsoon as well as non-monsoon runoff represents the annual runoff (RO_{AN}) which can also be computed as the total runoff of twelve months in a water year. Different form of relationships may be developed for monsoon and non-monsoon runoff.

(a) Monsoon season rainfall-runoff relationship

The following form of the rainfall runoff relationships may be considered for monsoon season:

$$RO_{MON} = a + b\,P_{MON} \qquad (32)$$
$$RO_{MON} = a\,.\,RO^{b}_{AN} \qquad (33)$$
$$RO_{MON} = a\,(RO_{AN})^{b} \qquad (34)$$

$$RO_{MON} = a\,P_{AN} \tag{35}$$
$$RO_{MON} = a\,RO_{AN} + b \tag{36}$$
$$RO_{MON} = a\,P_{AN} + b \tag{37}$$
$$RO_{MON} = a\,(P_{AN} - P_o)^b \tag{38}$$

Here, P_{MON} = Rainfall for monsoon season,
P_{AN} = Annual rainfall, and
P_o = Threshold value of the rainfall, below which no runoff occurs. It is considered to be lost as initial loss without contributing the runoff.

(b) Non-monsoon season rainfall-runoff relationship

For non-monsoon season, the following form of the relationships may be considered.

$$RO_{NON} = a.\ RO_{AN} \tag{39}$$
$$RO_{NON} = a.\ RO_{MON} \tag{40}$$
$$RO_{NON} = a(RO_{MON})^b \tag{41}$$
$$RO_{NON} = a(RO_{AN})^b \tag{42}$$
$$RO_{NON} = a(P_{AN})^b \tag{43}$$
$$RO_{NON} = a(P_{AN}) \tag{44}$$
$$RO_{NON} = a\,P_{MON} \tag{45}$$
$$RO_{NON} = a.P_{MON} + b.P_{NON}. \tag{46}$$
$$RO_{NON} = a.(P_{MON})^b \tag{47}$$
$$RO_{NON} = a.P_{MON} + b \tag{48}$$
$$RO_{NON} = a.(P_{MON} - P_O)^b \tag{49}$$
$$RO_{NON} = a.P_{MON} + b \tag{50}$$
$$RO_{NON} = a.(P_{MON} - P_o)^b \tag{51}$$
$$RO_{NON} = A.\ P_{MON} + B.\ P_{NON}\ . + C \tag{52}$$

Here, P_{MON} represents the rainfall during non-monsoon period.

Annual Rainfall-Runoff Relationships For The Gauged Catchments

For annual (water year) rainfall-runoff relationship the following types of equations may be tried:

$$RO_{NON} = a\,P_{AN} \tag{53}$$
$$RO_{NON} = a\,P_{AN} + b \tag{54}$$
$$RO_{NON} = a\,P_{AN}{}^b \tag{55}$$
$$RO_{AN} = .a\ .(P_{AN} - P_o)^b \tag{56}$$

Seasonal and annual rainfall-runoff relationships for ungauged catchments

Seasonal and annual rainfall-runoff relationships for ungauged catchments can be developed by regionalizing the regression coefficients, involved in their respective relationships using the methodology discussed for monthly rainfall-runoff relationships for ungauged catchments.

REMARKS

In this lecture different forms of the relationships have been presented for monthly, seasonal, and annual rainfall-runoff under the first approach. Daily rainfall-runoff relationship has also been presented utilizing the regression based approach. The suitable relationships can also be developed for other periods such as weekly, ten daily etc. not discussed in the lecture after trying the various form of the relationships using the first approach.

REFERENCES

1. Chow, V.T., 1964. *Handbook of Applied Hydrology.* McGraw-Hill, New York.

2. Haan, C.T., 1972). *A water yield model for small watersheds.* Water Resources Research, 8(1): 58-69.

3. Kohler, M.A., 1963a. *Rainfall-Runoff model.* International Association of Scientific Hydrology Publication, No.63, pp.473-491.

4. Linsley, R.K., M.A.Kohler and Paulhus, J.L.H., 194. *Applied Hydrology.* New York: McGraw-Hill.

5. Linsley, R.K., M.A.Kohler and Paulhus, J.L.H., 1975. *Hydrology for engineers.* 2nd ed., 258-60 and 265-76, New York, McGraw-Hill.

6. Majtenyi, S.I., 1972. *A model to predict mean annual watershed discharge.* Jour. of the Hydraulics Division, Proc. of the American Society of Civil Engineer, 98, (HY7): 1171-86).

7. Pike, J.G., 1964. *The estimation of annual runoff from meteorological data in tropical climate.* Jour. of Hydrology, 2: 116-23.

** *** **