TR(BR) - 3/98-99

# METHODOLOGY FOR THE WATER AVAILABILITY COMPUTATIONS UNDER DIFFERENT DATA AVAILABILITY SCENARIOS

आपो हि ष्टा मयोभुवः

NATIONAL INSTITUTE OF HYDROLOGY
JAL VIGYAN BHAWAN
ROORKEE - 247 667 (U.P.) INDIA

1998-99

# Preface

*Water availability generally refers to the volume of water available from a basin or stream at a specified point over a specified period of time. Determination of water availability or basin yield is required for solution of a number of water resource related problems including design of water resource projects, reservoir operation planning and determination of availability of water for irrigation, domestic, power supply or industrial use.*

*Determination of water availability requires sufficient length of runoff records for its correct estimation. However, at most of the sites in our country either shorter series of runoff is available or no record is available. To obtain sufficient length of recorded runoff series, the best suited alternative is therefore either to transfer the available information from a nearby site/s or to develop a suitable rainfall-runoff relationship for the basin and then to generate the long term runoff series based on available rainfall series. Broadly such situations can be grouped under three categories. 1) availability of sufficient length of rainfall-runoff record; 2) availability of short term runoff record and; 3) no record availability.*

*In this report, prepared by Sh. R. Mehrotra, Scientist, under the supervision of Sh. R. D. Singh, Divisional Head of Surface Water Analysis and Modelling Division, possible situations of data availability situations and their solutions are discussed. It also highlights the use of water balance modelling and stochastic and statistical modelling mainly for extension of streamflow records.*

*It is expected that this report may prove a useful tool to the practising hydrologists and field engineers involved in water resources project formulation.*

***( S.M. SETH )***
***DIRECTOR***

# CONTENTS

| Description | Page No. |
|---|---|
| List of Tables and Figures | i |
| Abstract | ii |
| INTRODUCTION | 1 |
| WATER AVAILABILITY COMPUTATIONS | 2 |
| DESCRIPTION OF METHODOLOGIES | 3 |
| DIRECT OBSERVATION METHOD | 3 |
| EXTENSION OF STREAMFLOW RECORD | 4 |
| EXTENSION WITH LONG TERM RUNOFF DATA AT NEIGHBOURING SITE | 4 |
| Double Mass Curve Method | 4 |
| Regression Analysis | 5 |
| Index-station Method | 5 |
| Langbein's Log-deviation Method | 6 |
| EXTENSION WITH SHORT TERM RUNOFF DATA AT THE PROPOSED SITE USING STOCHASTIC MODELLING | 7 |
| Classifications of Time Series Models | 8 |
| Data Generation for Daily Flows | 8 |
| *Markov Chain* | 9 |
| *Transition Probability Matrix Method* | 10 |
| *Alternating Renewal Process* | 12 |
| Data Generation for Monthly Flows | 14 |
| *Models Used for Monthly Streamflow Generation* | 14 |
| Monthly Markov Model | 14 |
| Thomas Fiering Model (TFM) | 15 |

| No | Description | Page No. |
|---|---|---|
| | Two Tier model (TTM) | 15 |
| | Method of Fragments | 16 |
| | Disaggregation models | 16 |
| | **Data Generation for Annual Flows** | 18 |
| | ***Generation of Synthetic Series*** | 23 |
| | **RAINFALL-RUNOFF SERIES METHOD** | 24 |
| | **Using statistical or correlation approach** | 24 |
| | ***Data availability scenarios*** | 24 |
| | Case I: Long term precipitation record along with a streamflow data for a few years at the site is available | 24 |
| | Case II: Long term precipitation record is available at the site along with precipitation and streamflow data for a few years at a neighbouring site is available | 24 |
| | Case III: Only precipitation record at the site is available | 25 |
| | Case IV: No record of any kind is available | 28 |
| | **Extension of records using Rainfall - Runoff modelling** | 28 |
| | ***Water balance approach*** | 28 |
| | ***Water balance or rainfall-runoff models*** | 29 |
| | ***Classification of hydrologic models*** | 30 |
| | Black box models | 31 |
| | Conceptual models | 31 |
| | Physically based models | 32 |

| No | Description | Page No. |
|---|---|---|
| | *Data availability scenarios* | 32 |
| | Case I: Long term precipitation record along with a streamflow data for a few years at the site is available | 33 |
| | Case II: Long term precipitation record is available at the site along with precipitation and streamflow data for a few years at a neighbouring site is available | 33 |
| | Case III: Only precipitation record at the site is available | 33 |
| | Case IV: No record of any kind is available | 33 |
| | **Monthly rainfall - runoff models for ungauged catchments** | 34 |
| | FLOW-DURATION CURVES FOR WATER AVAILABILITY COMPUTATIONS | 37 |
| | GENERATION OF LONG SEQUENCES OF FLOW VOLUMES | 40 |
| | INFLUENCE OF TIME INTERVAL FOR ORIGINAL RECORD | 40 |
| | DEVELOPMENT OF REGIONAL FLOW DURATION CURVE | 41 |
| | **Regional Flow Duration Model** | 42 |
| | CONCLUDING REMARKS | 45 |
| | *References* | 46 |
| | **Annexure - I** | |

# List of Tables and Figures

**No.** **Title** **Page No.**

**Table No.**

1. Identification Properties for AR, MA and ARMA Processes 20

2. Strange Rainfall-runoff Ratios 25

3. Monthly Losses as proposed by Khosla 27

4. Values of Runoff Coefficients 27

5 Average values of n(I) 36

**Figure No.**

1. Structure and schematic representation of a monthly rainfall-runoff model for ungauged catchment 35

# Abstract

Water availability generally refers to the volume of water available from a basin or stream at a specified point over a specified period of time. As it is computed for a specific time period it reflects the volumetric relationship between rainfall and runoff. Many factors viz. climatic and basin characteristics affect the water availability of a basin. Time and space distribution of rainfall, its intensity and duration, surface vegetation, soil moisture, soil characteristics, topography and drainage network are some of the important factors.

Determination of water availability or basin yield is required for solution of a number of water resource related problems: i) design of water resource projects, ii) determination of availability of water for irrigation, domestic, power supply or industrial use, iii) adjustment of long records of runoff for varying rainfall patters, and iv) reservoir operation planning.

This report elaborates the different data availability scenarios for computation of water availability of a basin and also provides the available methodologies under such situations. Broadly such situations can be grouped under three categories. 1) availability of sufficient length of rainfall-runoff record; 2) availability of short term runoff record and; 3) no record availability. Report also highlights the use of water balance modelling and stochastic and statistical modelling mainly for extension of streamflow records.

It is expected that this report may prove a useful tool to the practising hydrologists and field engineer involved in water resources project formulation.

# INTRODUCTION

Water is the most precious gift of the nature, the most crucial for sustaining life, is required for almost all types of activities. The development, conservation and use of water, therefore, forms one of the main elements in the country's development planning. Considering the fact that the available water resources of the country are now under stress with the increasing future demands, optimum and efficient utilization of available water resources based on scientific knowledge is becoming essential.

Water availability generally refers to the volume of water available from a basin or stream at a specified point over a specified period of time. As it is computed for a specific time period it reflects the volumetric relationship between rainfall and runoff. Many factors viz. climatic and basin characteristics affect the water availability of a basin. Time and space distribution of rainfall, its intensity and duration, surface vegetation, soil moisture, soil characteristics, topography and drainage network are some of the important factors.

Determination of water availability or basin yield is required for solution of a number of water resource related problems:

- Design of water resource projects,
- Determination of availability of water for irrigation, domestic, power supply or industrial use,
- Adjustment of long records of runoff for varying rainfall patters, and
- Reservoir operation planning

Generally, for most of the basins in the country, rainfall records of sufficient length are available. However, the available runoff data is either of short duration or has gaps due to missing data. The shorter length of data apart from other problems, always faces the problem of true representation of natural behaviour of the time series. In the absence of availability of sufficient length of recorded runoff series, the best suited alternative is either to use the information available at the neighbouring site/s or to develop a suitable rainfall-runoff relationship for the basin and then to generate the long term runoff series based on available rainfall series.

Based on limited data, hydrologists have proposed equations relating monthly runoff to factors such as monthly rainfall and monthly average temperature (Khosla, 1949, Murry, 1971). These equations and other similar equations developed elsewhere, such as that proposed by Mimikou & Rao (1983), have been used for planning purposes in cases where no other information is available. A few simple models have also been developed for estimating runoff volumes on ungauged catchments (Manley, 1978; Nathan & McMohan, 1990; Taffa Tulu, 1991). These models were meant for specific purposes to suit the specific needs.

In the present report an attempt has been made to list some of the possible alternatives of data availability. Also, possible/prevalent methodologies available or being used in the country for these situations keeping in view the planning and design aspect of a water resources project are discussed.

## WATER AVAILABILITY COMPUTATIONS

Determination of water availability is required for solution of number of water resources problems. There are several approaches to determine the water availability. Most of these approaches can be classified as either empirical approaches or continuous- time simulation approaches using the water balance equation. All these approaches are greatly influenced by the availability of data and the selection of period for which the water availability is to be determined. Generally with the increase in the time period, determination becomes easier. The time period of interest is generally equal to storm duration, a day, a month or a year. The approach followed is also influenced by the size of the basin and purpose for which it is to be applied.

For a systematic water availability computation for planning and design of a project, as a guide line, following rainfall and runoff data should be collected in order of preference as given below:

(a) Ten daily or monthly runoff data, i.e. the total of the daily runoff in 10 days/one month at the proposed site for at least 40 to 50 years.

or

(b)
1. Ten daily/ monthly rainfall data for at least 40 years for rain gauge stations located within or nearby the proposed site of the basin.
2. Ten daily/ monthly runoff observations at the proposed site for the last 5 to 10 years.

or

(c)
1. Ten daily/ monthly rainfall data of the catchment of the proposed site for the last 40 to 50 years.
2. Ten daily/ monthly runoff observation and concurrent rainfall data at the existing work upstream or downstream of proposed site for the last 5 to 10 years or more.

or

(d)
1. Ten daily/ monthly rainfall of the catchment for the last 40 to 50 years for the proposed site.
2. Ten daily/ monthly runoff observations and concurrent rainfall data on a nearby river for 5 to 10 years or more provided orographic conditions of the catchment at the works site are similar to that of the proposed site.

For preparation of a detailed project report CWC has specified some guidelines for preparation of hydrology aspect of the project report with regard to data requirement keeping in view the type of project. For the ready reference to the readers, a copy of the annexure of the report of CWC, relevant to hydrology is also given here as Annexure I.

## DESCRIPTION OF METHODOLOGIES

Depending upon the type of data available, the water availability can be computed from the following methods.

### DIRECT OBSERVATION METHOD

The method is applied when observed runoff-data at the point of interest is available for a sufficient period of time. The average of record over the available period for the particular period of time gives the average water availability for the basin. To compute the water availability for a particular dependability, the observed runoff is arranged in descending order and a flow duration curve is constructed. The synthetic period for a particular dependability is calculated from this arranged series.

## EXTENSION OF STREAMFLOW RECORD

For water resource assessment in the design and planning of water resources projects, the primary objective is frequently to obtain a long representative time series of streamflows from which a river basin project can be designed. However, in our country for most of the sites, the long-term data required for the design and planning of water resources projects is generally not available. The only option left before the hydrologists is either to extend the short-term data or to generate the data based on limited short-term historical record. We shall discuss here both the alternatives in detail.

The streamflow data can be extended with the help of two long-term records a) long-term runoff data at neighbouring site/s ; and b) long-term precipitation record at the concerned site.

### EXTENSION WITH LONG-TERM RUNOFF DATA AT NEIGHBOURING SITE

The following methods are used for the extension of runoff record using the long-term runoff record of the neighbouring site/s or site/s available either u/s or d/s of the site assuming that the site/s to be considered is/are located in the similar hydro-climatic zone.

1. Double-mass curve method.
2. Correlation with catchment areas.
3. Regression analysis.
4. Index-station method.
5. Langbein's log-deviation method.

### Double mass curve method

Double mass curve technique is frequently being used for checking the consistency of rainfall records and also the changes in rainfall record occurred as a result of changes in environment, shifting of the site etc. Similarly this technique is also used for determining the effect of changes in the runoff record due to land use changes or otherwise.

In this method, the cumulative streamflow of the station and the index station are plotted on a graph. The slope of the double mass-curve gives the relation of the streamflows at two stations.

The double mass curve is however of limited use for extending streamflow data because of its assumption of a constant slope of the correlation line, regardless of variation in size of yearly or monthly increments. As the two records of streamflow data may not necessarily correlate with each other as straight line, the method has limited applicability.

**Regression analysis**

Frequently, in hydrologic works, we are concerned with dependence of a given variable Y on one or more independent variables $X_1, X_2, X_3 \dots\dots\dots X_N$ so that we may predict the value of Y knowing the value of $X_1, X_2, X_3 \dots\dots\dots X_N$. This is done by regression analysis which determines the association between different variables.

Now, if $Q_s$ is the dependent variable while $Q_i$ is the independent variable then a regression line of the form $Q_s = a + bQ_i$ can be fitted. Here a and b are regression coefficients which are given by:

$$b = \frac{\Sigma Q_i Q_s - \frac{\Sigma Q_i \Sigma Q_s}{N}}{\Sigma Q_i^2 - \frac{(\Sigma Q_i)^2}{N}} \tag{1}$$

$$a = \bar{Q}_s - b\bar{Q}_i \tag{2}$$

Where

$$\bar{Q}_s = \frac{1}{N}\Sigma Q_s \tag{3}$$

$$\bar{Q}_i = \frac{1}{N}\Sigma Q_i \tag{4}$$

**Index station method**

The index station method makes use of duration curves. If the magnitude to be plotted is the discharge or flow of a stream, the duration curve is known as the flow duration curves. In a statistical sense, a flow duration curve is a cumulative frequency curve of a continuous time series. The slope of a flow duration curve depends upon the observation period used in the analysis, being much steeper with the mean daily data than with the annual data as the later tend to group and smoothen the variations of the short interval daily data.

In this method, two duration curves are prepared for the index station I. One for the long term available data ($I_l$) and another for the short term data, concurrent with the period of record at the short-term station S ($I_s$). Similarly, for the station S, one flow duration curve for the available record is constructed ($S_s$).

To develop the long term flow duration curve at station S, it is assumed that if some discharge Q has a duration of p% of the time on $I_s$ and p'% of the time on $I_l$, then the discharge Q' which has duration of p% of the time on $S_s$ will have duration p'% of the time on $S_l$. Thus the method assumes that for a given discharge Q at the index station, the change in percent of time from p to p' due to consideration of long term data (from $I_s$ to $I_l$) is the same at station S, had there been long term observed data at station S too. This works reasonably good if the two stations are situated in same hydro-climatic region.

**Langbein's Log-deviation method**

This is the most accepted and one of the procedures being used by the Central Water Commission for extension of short term data with the help of long term data at a nearby station. The method involves establishment of a correlation between discharges at long term and short term stations. Log of discharge values are used to remove the skewness and to avoid the effect of extremely high and low values. The usual practice is to correlate ten daily mean values or monthly mean values although flood peaks, daily means and annual means could also be used.

A relationship between the deviations of logrethim of values at the Index station with long term data (independent variable X) and logrethim of values at the station under consideration (dependent variable Y) is established. The regression equation takes the following form:

$$Y = a + b X \quad (5)$$

Here, X and Y values are deviations which are obtained by subtracting the mean of the series from the log values of original discharge values.

Regression coefficients are given by

$$a = 0 \text{ (As } \Sigma X = \Sigma Y = 0) \quad (6)$$

$$b = \Sigma XY / \Sigma X^2 \quad (7)$$

Correlation coefficient r is given by

$$r = \Sigma XY / [\Sigma X^2 \Sigma Y^2]^{1/2} \quad (8)$$

Now using the above relationship and concurrent records at both the stations, value of regression coefficient b is computed and then using the long term data of index station, corresponding values at the station under consideration are computed. Antilogrithms of the computed values gives the discharge values at the station under consideration.

## EXTENSION WITH SHORT-TERM RUNOFF DATA AT THE PROPOSED SITE USING STOCHASTIC MODELLING

Early studies by Hazen (1914) and Sudler (1927) showed the feasibility of using statistics and probability theory in analyzing river flow sequences. Barnes (1954) extended the early empiical studies of Hazen and Sudler and introduced the idea of synthetic generation off streamflow by using a table of normal random numbers. However, formal development of stochastic modelling started only with the introduction and application of autoregressive models for annual and seasonal streamflow generation (Thomas and Fiering, 1962; Yevjevich, 1963). Since than a vast body of literature has been published about the use of time series analysis for studying and solving many interesting problems.

The generation of long period synthetic streamflow data is potentially of great use in the design of water resource systems. Data generation procedures are used to provide equally likely flow sequences to the historic one to help the designer to prepare a qualitative picture of the probability of failure, storage relation etc. in simulation studies. Many models have been proposed by different investigators depending upon the time period of interest.

### Classification of time series models

First step in model construction is to sellect suitable classes or families of models from which the most appropriate model to fit to a given time data set can be eventually chosen by following the identification, estimation and diagnostic check stages of model development.

By defination a model is an approximation to reality. If one is aware of the important statistical characteristics of a data set, one can model the data as realistically as possible by considering one or more families of models which are capable of modelling these major characteristics. Two types of models which are commonly used in time series modelling are Markov chain model and the autoregressive model and its different forms. It has been reported in the literature Salas at.el (1980) that the judicious use of auto regressive (AR), auto regressive moving-average (ARMA), auto regressive integrated moving-average (ARIMA) and disaggregation models generally produce satisfactory results for most practical cases of operational hydrology. These models may be used for modelling time series of annual values as well as for intervals that are a fraction of the year.

### Data generation for daily flows

For daily flows Several stochastic models have been developed during the last two decades. Early studies assumed that each storm was made up of a random number of rain cells which occur in space and time according to a three dimensional point process and the resulting rainfall could be described by a Poisson process. Later studies, however, considered rainfall to occur in clusters and used the Neyman-Scott model in place of Poisson models. Cluster centres in the Neyman Scott model are not rain cells, but are just points around which the density of cells is larger than in other regions. Each cluster has associated with it a number of rain cells which is a random variable, independent and identically distributed for each cluster centre. In the cluster process the number of cells in a storm is randomly distributed and the cell arrival times are exponentially distributed.

Recently, the use of Poisson cluster process in stochastic modelling of rainfall has been investigated by several authors. The cell arrivals are modelled by a Poisson cluster process i.e. storm arrivals form a Poisson process and a cell arrival distribution is assigned to each storm; the depth and duration of the cell are modelled by exponential distributions.

The methodology used for generation of daily rainfall consists of two parts : the first determines the occurrence of dry and wet days and the second generates the rainfall depth on wet days. For reproducing the occurrence of the rain events, the techniques as discussed in the following sections have been used.

### *Markov Chain*

Many hydrologic time series exhibit significant serial correlation. That is, the value of the random variable under consideration at one time period is correlated with the values or the random variable at earlier time periods. The correlation of a random variable x at one time period with its value k time periods earlier is denoted by $\rho_x(k)$ and is called the kth order serial correlation. If $\rho_x(k)$ can be approximated by $\rho_x(k) = \rho^k_x(1)$ , then the time series of the random variable x might be modelled by a first order Markov process. A first order Markov process is defined by the equation:

$$X_{i+1} = \mu_x + \rho_x(1)\,(X_i - \mu_x) + \epsilon_{i+1} \tag{9}$$

where, $X_i$ is the value of the process at time i , $\mu_x$ is the mean of X, $\rho_x(1)$ is the first order serial correlation and $\epsilon_{i+1}$ is a random component with $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2$. This model states that the value of X in one time period is dependant only on the value of X in the preceding time period plus a random component. Further it is also assumed that $\epsilon_{i+1}$ is independent of $X_i$. If the distribution of X is $N(\mu_x, \sigma_x^2)$ then the distribution of $\epsilon$ is $N(0, \sigma_\epsilon^2)$. random value of $X_{i+1}$ can now be generated by selecting $\epsilon_{i+1}$ randomly from a $N(0, \sigma_\epsilon^2)$ distribution.

Thus a model for generating X s that are $N(\mu_x, \sigma_x^2)$ and follow the first order Markov model is:

$$X_{i+1} = \mu_x + \rho_x(1)\,(X_i - \mu_x) + t_{i+1}\, \sigma_x \sqrt{1-\rho_x^2\,(1)} \tag{10}$$

The procedure for generating a value for $X_{i+1}$ is to estimate $\mu_x$ , $\sigma_x$ and by X, $S_x$ and $r_x(1)$

respectively and then select $t_{i+1}$ at random from a N (0, 1) distribution and calculate $X_{i+1}$ based on X, $S_x$ and $r_x$ (1) and $X_i$ In this approach events are considered to belong to a certain number of states. No rain is one state and the other states which are wet may be only one or several. The probability of a day belonging to a certain state is dependent on the occurrence of belonging to a certain state is dependent on the occurrence of several previous states.

Above equation generates normally distributed X s with a mean of $\mu_x$ variance of $\sigma_x^2$ and first order serial correlation of $\rho_x(1)$. For a first order Markov process, the lag k serial correlation $\rho_x(k)$ is given by:

$$\rho_x(k) = \rho_x^{k}(1) \tag{11}$$

Thus, the correlogram exponentially decays from $\rho_x(0) = 1$ to $\rho_x(\infty) = 0$ according to above equation. If an observed correlogram has this property, the Markov model may be an appropriate generating model.

Haan et. al. (1976) developed a stochastic model based on a first order Markov Chain to simulate daily rainfall at a point. The model uses historical rainfall data to estimate the Markov transitional probabilities. The model was said to be capable of simulating daily rainfall records of any length based on the estimated transitional probabilities and the frequency distribution of rainfall amounts.

Roldan and Woolhiser (1982) and Woolhiser and Roldan (1982) used a first order Markov chain as the occurrence process and a mixed exponential distribution for the daily rainfall. The model was reported to have performed better than other alternatives.

***Transition Probability Matrix Method***

The essential features of this method are:

(i) the range over which rainfall is expected to vary is divided into set of discrete intervals.

(ii) a matrix corresponding to the intervals is built up from the observed rainfall sequences by tabulating the number of times the observed data went from state i to state j denoted by $n_{ij}$ as follows :

Final state

Starting state

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | $n_{00}$ | $n_{01}$ | $n_{02}$ | $n_{03}$ |
| 1 | $n_{10}$ | $n_{11}$ | $n_{12}$ | $n_{13}$ |
| 2 | $n_{20}$ | $n_{21}$ | $n_{22}$ | $n_{23}$ |
| 3 | $n_{30}$ | $n_{31}$ | $n_{32}$ | $n_{33}$ |

(iii) the m x m transitional probabilities matrix is $p = [p_{ij}]$ which is given by

Final state

Starting state

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | $p_{00}$ | $p_{01}$ | $p_{02}$ | $p_{03}$ |
| 1 | $p_{10}$ | $p_{11}$ | $p_{12}$ | $p_{13}$ |
| 2 | $p_{20}$ | $p_{21}$ | $p_{22}$ | $p_{23}$ |
| 3 | $p_{30}$ | $p_{31}$ | $p_{32}$ | $p_{33}$ |

where, $p_{00} = n_{00} / (n_{00} + n_{01} + n_{02} + \ldots\ldots)$ (12)

$p_{10} = n_{10} / (n_{10} + n_{11} + n_{12} + \ldots\ldots)$ (13)

and so on each $p_{ij}$ being obtained by dividing $n_{ij}$ by the corresponding row total of the $n_{ij}$.

(iv) After the transitional probabilities are estimated the next step is the simulation of rainfall using appropriate distribution. The synthetic sequences were generated by dividing the rainfall into a number of classes (intervals). The sequence of states are built up by selecting a pseudo random number between 0 and 1 and then assigning the state according as the value of u is less than or greater than $p_{01} + p_{00}$ and then moving to the next state and so on

After trying as many as 13 classes Haan et. al. (1976) found six classes to be a reasonable choice. Srikanthan and Mc Mohan (1982) used seven classes. The rainfall in the last class was generated using a shifted exponential distribution. Haan et. al. (1976) used a multi state 7 x 7 Markov chain model and employed a uniform distribution for each of the wet states except for the last for which an exponential distribution was assumed. The model was tested on data of seven rainfall stations in Kentucky. A separate transition probability matrix was used for each month. The class distribution for the states in the Markov chain were found by using geometric progression. The comparison of simulated rainfalls with observed rainfalls indicated that the model generated rainfalls exceeded the historical rainfalls on an average by about 2.5%.

### *Alternating Renewal Process*

This process consists of alternating wet and dry spells. Wet spells are assumed independent and belong to a particular distribution. Similarly dry spells are assumed independent and belong to another distribution. Further, the two random sequences independent. The following definitions apply :

(i) A wet spell is a sequence of wet days bounded on either side by a dry day.
(ii) A dry spell is defined likewise
(iii) Spells are assigned to the periods ( usually months or seasons) in which they begin.
(iv) A day is defined as wet if the rainfall exceeds a threshold value $\delta$ mm

The first two assumptions could be satisfied by analysing data on a monthly or seasonal basis. The third assumption is checked by computing the correlation between wet and dry spells in each month.

Several distributions can be fitted to the data to model the lengths of wet and dry spells. Commonly used distributions for wet and dry runs are; the truncated negative binomial distribution (TNBD) and the shifted negative binomial distribution (SNBD). The probability density function of the TNBD is given by:

$$P(x=K \mid x \geq 1) = \left[\frac{K+r-1}{K}\right] p^{r} \frac{(1-p)^{r}}{1-p^{r}} \tag{14}$$

in which x is the random variable, k is the length of the spell and p and r are parameters ($0 < p < 1$ ; $-1 < r$). If the rainfall is modelled on monthly basis, 24 parameters are required to be estimated for the above equation for the whole year. In general, the development of rainfall model based on the alternating renewal process requires long data series so that sufficient number of wet and dry spells could be included. Data of 25 to 30 years when used for only a season are known to have performed well.

The model for rainfall amounts must be one which describes the distribution of rainfall amounts on days when it rains. The distribution is highly skewed and since the wet day is defined as a day on which rainfall exceeds a threshold value $\delta$ mm, a shifted or truncated distribution would model the rainfall amounts well.

The shifted two parameter gamma distribution (SGD) has been frequently used to fit the rainfall amounts. The probability density function of this distribution is given by:

$$f(y) = \frac{\lambda^{\nu} y^{\nu} \exp(-y\lambda)}{\Gamma\nu} \tag{15}$$

in which y is the rainfall amount, $\Gamma$ is the gamma function, and $\nu$ and $\lambda$ are parameters.

The mean rainfall of a wet spell depends on the length of the spell (Buishand, 1978 ). To account for this, three types of rainfall are distinguished. These are :

(i) wet spells with a solitary wet day (type 0)
(ii) a wet day with one adjacent day also wet (type 1)
(iii) a wet day with both adjacent days wet (type 2)

The probability density function in the above equation is derived separately for each type of wet spell for each month. The estimation of parameters p and r in Eq. (14) and $\nu$ and $\lambda$ in Eq. (15) is done using the method of maximum likelihood.

The generation of daily rainfall sequences is carried out in two steps. First, the lengths of wet spells and dry spells are generated by coupling a uniform random number in the interval (0,1) to the cumulative distribution function of TNBD. For wet spells of type 0, type 1 or type 2 the appropriate random gamma variates are generated to obtain the values of daily rainfall.

## Data Generation for Monthly Flows

For monthly streamflow generation, many models such as autoregressive process based model (Thomas Fiering model) and models based on disaggregation processes have been proposed in the literature from time to time. Some of these commonly used models are discussed in the following section. In the autoregressive models two approaches are generally followed. In the first, seasonalities and periodicities in the monthly flows are first removed and then the resulting series is modelled. The second approach uses the Thomas-Fiering monthly model.

### *Models Used for Monthly Streamflow Generation*

The monthly streamflows models, developed in the past, can be broadly classified either as univariate model or multivariate model. Multivariate models are used when flows at more than one site need to be generated simultaneously and cross correlation properties of flows are required to be preserved. Univariate models considered here include Monthly Markov model, Thomas Fiering model and its modified forms and disaggregation process based models. Thomas Fiering model and its modified forms include (i) Thomas Fiering model (Thomas and Fiering, 1962) and (ii) Two-Tier model (Srikanthan and McMahon, 1982). Disaggregation process based models include (i) Method of fragments (Srikanthan and McMahon, 1982), (ii) Basic disaggregation model (Valencia and Schaake, 1973), (iii) Extended disaggregation model (Mejia and Rousselle, 1978), and (iv) Condensed disaggregation model (Lane, 1979).

(i) Monthly Markov model

After removing the peeriodicity in the time series either by parametric method or non-parametric method, lag one auto-correlation coefficient ($\rho$) of the resulting standardised series is computed. Then standard variates are computed using the equation :

$$X_{i+1} = \rho X_i + (1 - \rho^2)^{1/2} \epsilon_{i+1} \qquad (16)$$

Absolute monthly flows are obtained by using the monthly mean and standard deviation.

(ii) Thomas Fiering model (TFM)

The Thomas Fiering model (Thomas and Fiering, 1962) in its simplest form consists of twelve linear regression equations. Each month's flow is regressed upon previous month's flow. Using the Thomas Fiering notations, the model may be written as follows,

$$Q_{i+1} = \overline{Q_{j+1}} + r_j * \frac{\sigma_{j+1}}{\sigma_j} * (Q_i - \overline{Q_j}) + Z_i * \sigma_{j+1} * (1 - r_j^2)^{0.5} \qquad (17)$$

Where, $Q_i$, $Q_j$, $\sigma_j$ and $r_j$ are the flow of ith month, mean of the jth month flows, standard deviation of the jth month flows and correlation coefficient between j+1th and jth month respectively, and $Z_i$ is normally distributed random number N (0,1).

In equation (17), j varies from 1 to 12 and i varies from 1 to 12, 13 to 24 and so on.

(iii) Two Tier model (TTM)

When annual flows are obtained by summing 12 consecutive generated monthly flows by TFM, the resulting annual series in most cases, does not preserve annual parameters. To overcome this, in the TTM monthly flows generated by TFM are adjusted against the annual flows generated by an appropriate annual model (Srikanthan and McMahon, 1982) as:

$$Q'_{i,j} = \left[ \frac{Qa_i}{\sum_{j=1}^{12} Q_{i,j}} \right] * Q_{i,j} \qquad (18)$$

Where, $Qa_i$ is the generated annual flow for the ith year, $Q'_{i,j}$ and $Q_{i,j}$ are adjusted and original monthly flow for the jth month of ith year respectively.

The adjusted flows preserve the annual parameters at slight expense of the monthly parameters. The performance of the model can be further improved if the annual series obtained from an annual model and the series obtained by summing the monthly flows from the TFM are ranked separately. The monthly flows for each 12 month period are then adjusted using equation (18) against the annual flows obtained having the same rank as the summed monthly flows. The adjusted monthly flows are then rearranged according to the original sequence of the annual flows (Srikanthan and McMahon, 1982).

(iv) Method of Fragments (MF)

In method of fragments (Srikanthan and McMahon, 1982), observed monthly flows are first standardized year by year by dividing the monthly flows in a year by the corresponding annual flows. These standardized flows are known as fragments. Annual flows from the historic record (say n years) are ranked according to increasing magnitude and n classes are formed. Class 1 has a lower limit of zero and upper limit as average of first and second ranked annual flows. Other classes are also formed in a similar manner. The fragment obtained from the monthly flows corresponding to the smallest annual flow is assigned to class one and the fragment obtained from the monthly flows corresponding to the second smallest annual flow is assigned to class 2 and so on. The generated annual flows are then checked one by one for the class to which they belong and then disaggregated according to the fragment value assigned to that class.

(v) Disaggregation models

Disaggregation models (Valencia and Schaake, 1973) preserve relevant statistics at higher (annual) level as well as lower (seasonal) level and allow one to break down a already available series, into some shorter interval series. For instance an annual series may be disaggregated into monthly series. The key series is generated prior to disaggregation. Various forms of the disaggregation models are described in next section.

(a) Basic disaggregation model (BDM) as proposed by Valencia and Schaake, (1973), for a single site is given by,

$$Y = AX + B\epsilon \qquad (19)$$

Here X is the annual flow value and Y is a column matrix ($\omega$x1) containing the ($\omega$) seasonal flow values which sum to X. X and Y are transformed values with zero means. A is a ($\omega$x1) matrix of constant coefficients ; B is another ($\omega$x$\omega$) coefficient matrix and $\epsilon$ is a ($\omega$x1) matrix of independent identically distributed normal deviates. Matrices A and B are calculated as follows:

$$A = S_{YX} S^{-1}_{XX} \tag{20}$$

$$BB^T = S_{YY} - S_{YX} S^{-1}_{XX} S_{XY} \tag{21}$$

$S_{YY}$ is the matrix of co-variances among the seasonal values. $S_{YX}$ is the matrix of co-variances between the seasonal values and the annual values and $S_{XX}$ is the matrix of co-variances among the annual values. Lane (1979) gives solution technique to solve $BB^T$ for obtaining B.

(b) To preserve the co-variances of the first season of a year and any preceding season, Extended Disaggregation Model (EDM) was proposed by Mejia and Rousselle, (1976) as:

$$Y = AX + B\epsilon + CZ \tag{22}$$

Here Z is a column matrix ($\omega$'x1) containing as many seasonal values ($\omega$') from the previous years as are desired to be preserved and C is an additional parameter matrix ($\omega$x$\omega$'). A, B and C are the matrices of parameters and are estimated as follows (Salas, et. al.,1980).

$$A = (S_{YX} - S_{YZ} S^{-1}_{ZZ} S_{ZX})*(S_{XX} - S_{XZ} S^{-1}_{ZZ} S_{ZX})^{-1} \tag{23}$$

$$C = (S_{YZ} - AS_{XZ})S^{-1}_{ZZ} \tag{24}$$

$$BB^T = S_{YY} - AS_{XX} A^T - AS_{XZ} C^T - CS_{ZX} A^T - CS_{ZZ} C^T \tag{25}$$

or equivalently,

$$BB^T = S_{YY} - AS_{XY} - CS_{ZY} \tag{26}$$

(c) One drawback of basic and extended disaggregation models is their excessive number of parameters, because of the large number of cross correlations that they attempt to reproduce. To overcome this problem, Lane, (1979) proposed a model popularly known as Condensed Disaggregation Model (CDM), which essentially sets to zero several parameters of the extended model which are not important. Approach uses the extended lagged season of the previous year. The model is given as,

$$Y_\tau = A_\tau X + B_\tau \epsilon + C_\tau Y_{\tau-1} \tag{27}$$

Where $\tau$ denotes current season being generated. If there are $\omega$ seasons then $\omega$ equations are there and $\omega$ set of parameters $A_\tau$, $B_\tau$ and $C_\tau$ . For single site case all the parameter matrices are single element matrices.

Parameter of this model are estimated by

$$\begin{aligned} A_\tau = {} & [S_{YX}(\tau,\tau)\, S_{YY}(\tau,\tau-1) S^{-1}_{YY}(\tau-1,\tau-1) S_{YX}(\tau-1,\tau)] \\ & [S_{XX}(\tau,\tau) - S_{XY}(\tau,\tau-1) S^{-1}_{YY}(\tau-1,\tau-1) S_{YX}(\tau-1,\tau)]^{-1} \end{aligned} \tag{28}$$

$$C_\tau = [S_{YY}(\tau,\tau-1) - A_\tau\, S_{XY}(\tau,\tau-1)] S^{-1}_{YY}(\tau-1,\tau-1) \tag{29}$$

$$B_\tau B^T_\tau = S_{YY}(\tau,\tau) - A_\tau\, S_{XY}(\tau,\tau) - C_\tau\, S_{YY}(\tau-1,\tau) \tag{30}$$

Here $S_{YY}(\tau,\tau)$ is the matrix of co-variances among the current seasonal values. $S_{YX}(\tau,\tau)$ is the matrix of co-variances between the current seasonal values and the annual value associated with the current season and $S_{XY}(\tau,\tau-1)$ indicates the co-variance matrix between the annual value of the current season and the seasonal values of the previous season. $S_{XX}(\tau,\tau)$ is the matrix of co-variance between the annual values of the current season.

**Data generation for Annual Flows**

A variety of models are available for modelling of annual time series. Namely, one has to decide on one among the various alternative models, say AR, ARMA, ARIMA, Fractional Gaussian Noise (FGN), Broken Line (BL), Shifting Level (SL) or any other model that is available in stochastic hydrology.

However, ARMA models and their various forms i.e. AR(1) and AR(2) models are frequently used in hydrology for generation of annual flows. In the following section, ARMA model is descussed:

An annual hydrologic time series is represented by $x_t$. In general, this series may be non-normal distributed. If $x_t$ is not normal, an appropriate transformation may be used to transform $x_t$ into normal variable $y_t$. Considering the above, the mixed autoregressive-moving average model for fitting the annual hydrologic series $y_t$ is:

$$y_t = \mu + z_t \tag{31}$$

where $z_t$ is the ARMA(p,q) model.

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + \dots + \phi_p z_{t-p} + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \dots - \theta_q\epsilon_{t-q} \tag{32}$$

The autoregressive AR(1) model discussed above is a particular case of the ARMA model in which all $\theta$ coefficients are zero and $p=1$. The ARMA(1,1) model has been extensively used in hydrology. It has the form

$$z_t = \phi_1 z_{t-1} + \epsilon_t - \theta_1\epsilon_{t-1} \tag{33}$$

The parameters of the ARMA(p,q) model for $y_t$ ($\mu$, $\phi_1$, $\phi_2$ ..... $\phi_p$, $\theta_1$, $\theta_2$, $\theta_q$, $\sigma^2\epsilon$) are determined from observed annual hydrologc data. Here $\sigma^2\epsilon$ is the variance of the random variable $\epsilon_t$.

The fitting of ARMA(p,q) models to annual hydrologic series $z_t$ of length N comprises the following steps:

STEP (1) Calculate the sample mean $\bar{z}$ and variance $s^2$ of the series.

STEP (2) Calculate and plot the autocovariance function $c_k$, the autocorrelation coefficients, $r_k = \frac{c_k}{s^2}$, and the partial autocorrelation coefficients $\phi_k(k)$ for lags k going from 1 to at least N/4 but less than N. The partial autocorrelation function may be obtained recursively from

$$\phi_1(1) = r_1 \; ; \; \phi_2(2) = \frac{r_2 - r_1^2}{1 - r_1^2} \; ; \; \phi_2(1) = \frac{r_1(1 - r_2)}{1 - r_1^2} \tag{34}$$

$$\phi_{k+1}(k+1) = \frac{r_{k+1} - \sum_{j=1}^{k} \Phi_k(j)\, r_{k+1-j}}{1 - \sum_{j=1}^{k} \Phi_k(j)\, r_j} \tag{35}$$

$$\phi_{k+1}(j) = \phi_k(j) - \phi_{k+1}(k+1)\, \phi_k(k-j+1) \tag{36}$$

STEP (3) IDENTIFICATION: From the behaviour of the autocorrelation and partial autocorrelation functions and making use of the following table infer the order of the model, namely, the values of p and q which are likely to fit the series.

**Table 1 : Identification properties for AR, MA and ARMA Processes**

| Process | Autocorrelation | Partial Autocorrelation |
|---|---|---|
| AR(p) | Infinite in extent, consists of damped exponentials and/or damped waves. Attenuates as $\phi_k = \sum_{j=1}^{p} \phi_j \rho_{k-j}$ | Finite in extent, peaks at lags 1 through p then cuts off. |
| MA(q) | Finite in extent, peaks at lags 1 through q then cuts off. | Infinite in extent, consists of damped exponentials and/or damped waves. |
| ARMA(p,q) | Infinite in extent, first q-p lags; Irregular then damped exponentials and/or damped waves. Attenuates as $\rho_k = \sum_{i=1}^{p} \phi_i \rho_{k-i}$ $(k \geq q + 1)$ | Infinite in extent, first q-p lags; Irregular then damped exponentials and/or damped waves. |

STEP (4) Obtain the initial estimates of the p autoregressive parameters $\phi_1, \phi_2, \ldots\ldots, \phi_p$ solving the p Youle-Walker equations.

$$c_{q+1} = \phi_1 c_q + \phi_2 c_{q-1} + \ldots\ldots\ldots\ldots + \phi_p c_{q+1-p} \tag{37}$$

$$c_{q+2} = \phi_1 c_{q+1} + \phi_2 c_q + \ldots\ldots\ldots\ldots + \phi_p c_{q+2-p} \tag{38}$$

$$c_{q+p} = \phi_1 c_{q+p-1} + \phi_2 c_{q-p-2} + \ldots\ldots\ldots\ldots + \phi_p c_q \tag{39}$$

If the series $z_t$ does not have zero mean there is an overall constant $\theta_{00}$ on the right hand side of equation (32)

$$\theta_{00} = \bar{z}\ (1 - \sum_{i=1}^{p} \phi_i) \tag{40}$$

STEP (5) Obtain the initial estimates of the q moving average parameters. From the series

$$z_t' = z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \phi_3 z_{t-3} - \ldots\ldots - \phi_p z_{t-} \tag{41}$$

and calculate the autocovariance function $c_k'$ of the $z_t'$ series. It can be calculated as usual. Alternatively, Box and Jenkins (1976, p. 202) give the following formula for the $c_k'$ in terms of the $c_k$ of the $z_t$ series and the $\phi$ already available from steps (2) and (4), respectively:

$$c_j' = \sum_{i=0}^{p} \phi_i^2 c_j + \sum_{i=1}^{p} (\phi_0 \phi_1 + \phi_1 \phi_{i+1} + \ldots\ldots\ldots\ldots + \phi_{p-i} \phi_p)\ d_j \tag{42}$$

where

$$d_j = c_{j+i} + c_{j-i}\ ;\ j=0,\ 1,\ \ldots\ldots,\ q\ ;\ \phi_0 = -1 \tag{43}$$

With the $c_j'$, the $\theta$-parameters and the residual variance $\sigma^2\epsilon$ are obtained by solving following equations iteratively in which the initial values of the unknown parameters are assumed zero.

$$\sigma^2\epsilon = \frac{c_0'}{1+\theta_1^2+\theta_2^2+.....+\theta_q^2} \tag{44}$$

$$\theta_j = -\left(\frac{c_j'}{\sigma_\epsilon^2} - \theta_1\theta_{j+1} - \theta_2\theta_{j+2} - ..... - \theta_{q-j}\theta_q\right) \tag{45}$$

This completes the initial estimation of the parameters $\phi_1$, $\phi_2$ ..... $\phi_p$, $\theta_1$, $\theta_2$, $\theta_q$, $\sigma^2\epsilon$ and $\theta_{00}$. The first estimate of the model is thus

$$z_t = \theta_{00} + \sum_{i=1}^{p} \phi_i z_{t-i} + \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{46}$$

STEP (6) Obtain the maximum likelihood estimate of the parameters. Calculate the residuals

$\epsilon_j = 0$ ; $j=1$, ...... max(p,q)

if $p > q$,

$$\epsilon_{p+j} = \theta_{00} - \sum_{i=1}^{p} \phi_i z_{p+j-i} + \sum_{i=1}^{q} \theta_i \epsilon_{p+j-i} \tag{47}$$

$j=1, 2, ......, N-p$

Calculate the sum of squares

$$S = \sum_{t=1}^{N} \epsilon_t^2 \tag{48}$$

for several values of $\theta$ and $\phi$ around the initial estimates and obtain the values of the $\phi$'s and $\theta$'s for which S is minimum. This may be done graphically. For the ARMA(1,1) model one may plot the values of S on a $\phi$ - $\theta$ plane. Contours of equal values of S may be traced, and the minimum sum of squares is located, and the corresponding values of $\phi$ and $\theta$ are obtained. The variance of the residuals $\epsilon_t$ is $\sigma_\epsilon^2 = (1/N\ S)$. This graphical procedure has the advantage that it exhibits any pecularity that the sum of squares surface may have. After verification that the surface is free of anomolies this procedure is extended with the help of a computer programme to obtain the the minimum value of S.

The standard error of estimate of the parameters is obtained by taking the square root of the diagonal terms of the variance-covariance matrix.

STEP (7) Perform the Porte Manteau test to check that the $\epsilon_t$ is a normal independent variable. Calculate the auto-correlation function $r_k(\epsilon)$ of the residual series $\epsilon_t$ for the lags k going from 1 to L = N/10 + p + q. Calculate the statistic

$$Q = N \sum_{k=1}^{L} [r_k(\epsilon)]^2 \tag{49}$$

and determine if Q is less than the theoretical chi-square value with L-p-q degrees of freedom. If this test is not passed the model is rejected.

All the procedures outlined in steps (1) through (7) may be performed by means of a suitable computer programme developed for this purpose.

STEP (8) To select the final model, the previous procedure step (1) to step (7) may be performed for several models, for example AR(2), ARMA(1,1), ARMA(2,1) etc., for which the maximum likelihood estimation of the parameters is found to converge. The best model is found by calculating the Akaike information criterion

$$AIC(p,q) = N \ln(\sigma_\epsilon^2) + 2(p+q) \tag{50}$$

where $\sigma_\epsilon^2$ is maximum likelihood estimate of the residual variance obtained in step (6). The model having the minimum value of AIC(p,q) is selected.

***Generation of Synthetic Series***

Once the ARMA model has been selected, it may be used for generation of synthetic data. The series is generated by using the equation (49) as described above. It is necessary to give p initial values to start the algorithm. Generation of normal random numbers usually yields variates with zero mean and unit variance. These are multiplied by $\sigma_\epsilon$ to obtain random numbers with zero mean and variance $\sigma_\epsilon^2$. If the series has zero mean, the term $\epsilon_{oo}$ is zero. Finally inverse transform is performed to obtain the synthetically generated hydrologic variable. If the $z_t$ values generated in the previous step are centred so that $z_t = y_t - \bar{y}$, where, $y_t$ is the original hydrologic variable. After generating the $z_t$ series, $y_t$ is obtained from $y_t = z_t + \bar{y}$. If the $z_t$ is standardised such that $z_t = (y_t - \bar{y})/s_y$, where, $y_t$ is the original hydrologic variable, then after generating the $z_t$ series, $y_t$ is obtained from $y_t = z_t\, s_y + \bar{y}$.

### RAINFALL-RUNOFF SERIES METHOD

The method basically consists in extending the runoff data with the help of rainfall data by means of rainfall-runoff relationships developed either through their correlation analysis or using water balance approach.

### Using statistical or correlation approach

#### *Data availability scenarios*

Under statistical or correlation analysis, depending upon the availability of rainfall and runoff data, problem can be grouped under four cases. Water balance approach is being considered separately.

| | |
|---|---|
| Ist case: | Long term precipitation record along with a streamflow data for a few years at the site is available. |
| IInd case: | Long term precipitation record is available at the site along with precipitation and streamflow data for a few years at a neighbouring site is available. |
| IIIrd case: | Only precipitation record at the site is available. |
| IVth case: | No record of any kind is available. |

Case I

Here we have long term precipitation record besides a few years record of stremaflow at site. The procedure, therefore, is to establish statistical correlation between observed monthly rainfall and monthly runoff and plot it on a log-log graph for each month. If the relationship is not a straight line, it is then suitably extended to find out the runoff corresponding to weighted rainfall of each year.

Case II

Here we have the long term precipitation record at the proposed site besides a few years of runoff data at the neighbouring site. In this case, first the rainfall-runoff relationship at the neighbouring site is established and assuming that this relationship will hold good for the proposed site too, using the rainfall record of the proposed site, the long term runoff series is computed.

Case III

In this case, as only long term rainfall data is available, to compute the runoff we make use of following empirical formulae.

Strange tables: Strange evolved some ratios between rainfall and runoff based on data of Maharashtra state, India. He accounted for the geological conditions of the catchment as good, average and bad, while the surface conditions as dry, damp and wet prior to rain. The values recommended by him are given in the Table 2.

**Table 2: Strange Rainfall-runoff ratios**

| Daily rainfall (mm) | Runoff percentage yield when the original stage of ground is | | | | | |
|---|---|---|---|---|---|---|
| | Dry | | Damp | | Wet | |
| | Percentage | Yield (mm) | Percentage | Yield (mm) | Percentage | Yield (mm) |
| 5 | - | - | 4 | 0.2 | 7 | 0.35 |
| 10 | 1 | 00.10 | 5 | 0.5 | 10 | 1.00 |
| 20 | 2 | 00.40 | 9 | 1.8 | 15 | 3.00 |
| 25 | 3 | 00.75 | 11 | 2.75 | 18 | 4.50 |
| 30 | 4 | 01.20 | 13 | 3.90 | 20 | 6.00 |
| 40 | 7 | 02.80 | 18 | 7.20 | 28 | 11.20 |
| 50 | 10 | 05.00 | 22 | 11.00 | 34 | 17.00 |
| 60 | 14 | 08.46 | 28 | 16.80 | 41 | 24.60 |
| 70 | 18 | 12.61 | 33 | 25.10 | 48 | 33.60 |
| 80 | 22 | 17.60 | 39 | 31.20 | 55 | 44.00 |
| 90 | 25 | 22.50 | 44 | 39.60 | 62 | 55.80 |
| 100 | 30 | 30.00 | 50 | 50.00 | 70 | 70.00 |

Note: Depending upon the conditions of the catchment, up to 25% of the yield may be added or deducted.

(b) Inglis and De Souza's formulae: These formulae are applicable for the western ghats and plains of Maharashtra, India and are reproduced here

For ghat areas,

$$R = 0.85\,P - 30.5 \tag{51}$$

For plains,

$$R = \frac{(P - 17.8)\,P}{254} \tag{52}$$

Where R = Runoff (cm)
P = Precipitation (cm)

(c) Khosla's formula

$$R = P - \frac{T}{3.74} \tag{53}$$

Where R = Runoff (cm)
P = Precipitation (cm)
T = Mean temperature (°C)

To evaluate monthly runoff Khosla gave the following relationship

$$R_m = P_m - L_m \tag{54}$$

$$L_m = 5T_m \quad \text{when } T_m > 4.5\ ^\circ\mathrm{C} \tag{55}$$

Where $R_m$ = monthly runoff (mm)
$P_m$ = monthly precipitation (mm)
$L_m$ = Monthly losses (mm)

$T_m$ = Mean monthly temperature in °C

For $T_m < 4.5$ °C, losses are taken from the Table No. 3.

**Table 3 : Monthly losses as proposed by Khosla**

| | | | | | |
|---|---|---|---|---|---|
| $T_m$ (in °C) | 4.5 | -1 | -7 | -12 | -18 |
| $L_m$ (in mm) | 21 | 18 | 15 | 12.5 | 10 |

Similarly, for other regions other empirical relationships are available in the literature.

Runoff Coefficients

The runoff and rainfall can be correlated as below:

$$R = CP \tag{56}$$

Where R = Runoff (cm)

P = Precipitation (cm)

C = Runoff coefficient

The runoff coefficient depends on several factors affecting runoff i.e. soil type, land use, antecedent moisture conditions, intensity of rainfall and its duration etc.. The method is applicable for small watersheds. The usual values of C are given below in Table 4.

**Table 4 : Values of Runoff Coefficients**

| Type of area | Runoff coefficient, C |
|---|---|
| Urban | 0.3 - 0.5 |
| Forest | 0.05 - 0.20 |
| Commercial and Industrial Parks | 0.90 |
| Parks, Farms, Pastures | 0.05 - 0.30 |
| Asphalt or Concrete pavement | 0.85 |

Apart from above, the joint committee of the American Society of Civil Engineers and the Water Pollution Control Federation have recommended some values of the runoff coefficient C. These values are available in any of the standard text book on hydrology.

Although, the whole approach of runoff estimation using the runoff coefficient is extremely subjective and can not be advocated for estimating the runoff from specific storms, yet this practice is still being followed in the state departments to gather the first hand information about the runoff and for the design of storm drainage and small water-control structures.

Case IV

Under the circumstances when no record of any kind is available, the flood at the proposed site is estimated by making a regional flood-frequency analysis. Regional flood-frequency analysis consists of two steps. (a) plotting a curve between return period versus ratio of flood to mean flood; and (b) plotting a curve between the catchment area versus the mean flood.

The regional frequency curves have the most frequent applications in estimating the flood potential of ungauged catchments. Since such curves show the ratio of floods to the mean annual flood against the return period, it is necessary to make an estimate of the mean annual flood for the ungauged catchment. The mean annual flood depends on many variables, the most important and commonly available being the drainage area.

To compute the flood for the ungauged catchment, the mean flood corresponding to the catchment area is first computed from the curve of catchment area versus mean flood. Having known mean flood, using the relationship of ratio of flood to mean annual flood versus return period to the return period, for the desired return period ratio is find out. From the value of ratio and mean annual flood, flood of desired return period can be computed.

## Extension of Records Using Rainfall - Runoff Modelling

### *Water Balance Approach*

The methods discussed so far, provide runoff at the proposed site and then the dependable water availability was computed. A more sophisticate way to compute the reservoir inflow could be with the help of basin yield, which in general, refers to the quantity of water available from a stream at a given site over a specified duration of time. The summation over the specified time period of the continuous hydrograph of flow is carried out at a particular site on a stream. It is therefore a result of all hydrologic events responsible for flow such as storms of all durations and intensities, and the climatic, geologic and land-use practices. It includes streamflow from all the sources. Thus basin yield can be computed for the specific time interval and based on this desired dependability can be computed.

Basin yield of different watersheds can be compared by expressing it in terms of depth of water over the whole area of watershed. In this approach all other components of water balance i.e. precipitation evapotranspiration etc are also expressed in same units. The purpose of water balance study are to evaluate the net available water resources, both on the surface and sub-surface and to assess the existing water utilisation pattern and practices. Rainfall-runoff modelling or the water balance approach being most powerful and most common, is discussed in detail in the following section.

### *Water Balance or Rainfall-Runoff Models*

Use of rainfall-runoff modelling has been started since 10's or 20's whereas hydrologic simulations models of watershed based on physical and mathematical concepts have been developed since the beginning of the 1960's. Recent development in computers and analysis techniques have led to significant developments and application of mathematical and conceptual models in hydrology so as to help solving a variety of hydrological problems. The water balance models are very popular and are widely used for the assessment of surface runoff. These are also used for generating arbitrarily long runoff series which can be used for design of a water resource project.

Hydrologic models are symbolic or mathematical representation of known or assumed functions expressing the various components of a hydrologic cycle. However, the term hydrological model is often understood to be and is used more narrowly as a computer based mathematical model. With the current rapid developments within computer technology and hydrology the application of computer based hydrologic models can only continue to increase in the near future (Storm, 1989).

Various techniques are available in the literature for modelling hydrologic system. Simulation is one of them where a system is represented as a model and its behaviour is studied. Digital simulation is needed in a watershed research because it is a complex system to be analysed by exact mathematical techniques. In digital simulation, system model is developed by a number of mathematical expressions that represent the various processes of the system and simulation is done by using a computer. Models can be broadly grouped into two types:

i. Physical models : Those models which are built physically from concrete or other materials on certain scale come under this category. The disadvantages associated with physical models are :

- These are less flexible for representing changes to a watershed or river because it requires physical rebuilding, which is expensive.
- These models are not very useful for the full hydrologic cycle because of the vide range of physical and temporal scales of the processes to be modelled and because of calibration difficulties.
- Natural and man made changes in a river frequently initiate responses which can be propagated so for long distances both upstream and downstream. To create a physical model covering such a large area would be difficult and expensive.

ii. Mathematical Models : A mathematical model is a quantitative mathematical description of the processes or phenomena i.e. collection of mathematical equations (often partial differential equations), logical statements, boundary conditions and initial conditions, expressing relationships between input, variables and parameters. For example a model may be :

- a simple function to relate runoff to rainfall.
- a series of functions which attempt to reproduce all the steps in the runoff process.

### *Classification of Hydrologic Models*

Hydrological models can be classified in different ways. Broadly many of the models presented in the literature can be divided into deterministic and stochastic categories. A deterministic model is one in which the processes are modelled based on definite physical laws and no uncertainties in prediction are admitted. It has no component with stochastic behaviour i.e. the variables are free from random variation and have no distribution in probability. Deterministic models can be further classified according to whether the model gives a spatially lumped or distributed description of the catchment area, and whether the description of the hydrological processes is empirical, conceptual or fully physically based.

Another classification of stream flow models is : (i) Event based stream flow simulation models, and (ii) Continuous stream flow simulation Models (CSS). The event based stream flow simulation models are applied to simulate the flood events. On the other hand, continuous stream flow simulation (CSS) models are capable of providing the continuous output of stream flow generally at daily interval. An extensive listing of types of the CSS models and event based models is given by V. P Singh (1989).

The familiar classification of model classification is to classify them in three categories : a) Black box models, b) Lumped models and, c) Physically based models.

*Black box models*

Black box models are based on transfer functions which relate inputs with outputs. These models, as the name suggests, generally do not have any physical basis. Some commonly used black box models include the unit hydrograph based approaches, regression analysis, and time series models. These models have been in use for a long time. The parameters of these models are obtained either by analytical solution or through numerical optimization. The success of these models can be attributed mainly to simple mathematics, minimal computational requirements and acceptable results. However, it is not advisable to use these models for the input data which falls outside the calibration range. This is because in such cases the implicit understanding of the physical system and assumptions made, such as of linearity, may not remain valid and the modelling relies only on the mathematical equations, Anderson and Burt (1985). Due to this serious drawback, the use of these models for extreme events should always be made with utmost caution.

*Conceptual models*

Lumped conceptual models occupy an intermediate position between the fully physically- based approach and empirical black box analysis. Such models are formulated on the basis of a relatively small number of components, each of which is a simplified representation of one process element in the system being modelled. The catchment is a complex system where various physical, chemical, and biological processes take place continuously and govern the movement of water. In practice it is not feasible to model all these processes. Some simplifications have to be made either in the representation of the system or in the processes involved or both. The most common simplification made is spatial lumping and replacement of various components of the hydrological cycle by conceptual storage. It amounts to saying that the catchment system and its inputs and responses can be represented using the dimensions of depth and time. The within catchment variations of inputs and parameters are ignored. Due to this spatial averaging, the lumped model concept tends to be considered adequate only for small homogeneous catchment, Blackie and Eeles (1985). However, in practice they have been applied to sufficiently big and heterogeneous catchment. The computational requirements of these models are moderately small vis-a-vis the computational speeds available at a typical computer installation now-a-days.

*Physically based models*

The physically based models are based on our understanding of the physics of the hydrological processes which control the catchment response and use physically based equations to describe these processes. Also, these models are spatially distributed since the equations from which they are formed generally involve one or more space coordinates. A discretization of spatial and temporal coordinates is made and the solution is obtained at the node points of this discretized representation. This implies that these models can be used for forecasting the spatial as well temporal pattern of more than one hydrological variable. Such models require much of computational time and also require advance computers as well as a broad data base. Physically based distributed models do not consider the transfer of water in a catchment to take place in a few defined storage as in case of lumped conceptual models. From their physical basis such models can simulate the complete runoff regime, providing multiple outputs (e.g. river discharge, phreatic surface level and evaporation loss) while black box models can offer only one output. In these models transfer of mass momentum and energy are calculated directly from the governing partial differential equations which are solved using numerical methods, for example the St. Venant equations for surface flow, the Richards equation for unsaturated zone flow and the Boussinesq equation for ground water flow. As the input data and computational requirements are enormous, the use of these models for real-time forecasting has not reached the `production stage' so far, particularly for data availability situations prevalent in developing countries like India.

***Data Availability Scenarios***

Under this category, similar to statistical or correlation analysis, depending upon the availability of rainfall and runoff data, problem can be grouped under four following cases.

| | |
|---|---|
| Ist case: | Long term precipitation record along with a streamflow data for a few years at the site is available. |
| IInd case: | Long term precipitation record is available at the site along with precipitation and streamflow data for a few years at a neighbouring site is available. |
| IIIrd case: | Only precipitation record at the site is available. |
| IVth case: | No record of any kind is available. |

Case I

Under this case based on concurrent rainfall and runoff record and other details of the catchment a suitable water balance model is identified and calibrated. Once the model is calibrated then using the remaining rainfall series and the calibrated model runoff series is obtained.

Case II

In this case for the neighbouring catchment based on rainfall-runoff and other information a water balance model is identified and calibrated. Now, assuming that the neighbouring catchment and catchment under consideration are located in the same climatic zone and having similar catchment characteristics, the developed relationship for neighbouring catchment is used to obtain runoff series for the proposed site using the available rainfall series.

Case III

In this case, regional relationship for the region in which the catchment is located is developed. For it a simple model structure is first identified and then rainfall, runoff and other catchment characteristics required by the model are collected for the nearby catchment sites. For all these sites, calibrated values of parameters are find out and these values are then regionalised. These regionalised values of parameters and available precipitation records are used to get the desired runoff series. Alternatively, geomorphological characteristics of the catchment are collected and based on catchment area, travel time details and other information, geomorphological instantaneous unit hydrograph is developed which can be used further to get the flood hydrograph.

Case IV

If no record of any kind is available then only alternative is to make use of regional relationships developed for case III. Some relationships between catchment area and runoff, or other catchment characteristics and runoff can be established and using these characteristics, runoff can be computed. Here, as no information is available water balance approach can not be used as such.

For monthly time scale a few simple models have been developed for estimating runoff volumes on ungauged catchments or the catchments with scarce data (Khosla, 1949; Murry, 1971; Manley, 1978; Mimikou & Rao, 1983; Nathan and McMohan, 1990; Taffa Tulu, 1991; Jayasuria et al., 1991; Mehrotra, 1997; Agung et al., 1995 and Kothyari, 1995). Eric Servat and Alain Dezetter (1993) applied rainfall-runoff modelling in 20 catchments in northern Ivory Coast. Parameters of two conceptal models were characterised so that the models could be applied to ungauged catchments also. Variables were related to catchment land use and the rainfall distribution over the year Most of these models have their own limitations.

On annual basis also some models have been proposed to deal with conditions of no data availability. Mustonen (1967) included fall, winter, and summer precipitations, average annual temperature, potential evapotranspiration in summer, frost depth, percentage of drainage area with coarse soils, volume of forest growing stock, and change of soil moisture during water year in a multipule linear regression analysis to determine annual water yield. Along similar lines, Hann and Read (1970) correlated mean annual runoff with mean annual precipitation, watershed perimeter, and watershed relief ratio for small agriculture watersheds in Kentucky. Sucliffe and Carpenter (1967) used annual precipitation and elevation in their correlation study on a mountanious and semiarid area in western Iran. Hawley and McCuen (1982) related runoff to several physiographic characteristics from 605 watersheds in the western United States divided into five regions. Similar studies have been reported by Sharp et al. (1960); Harris et al. (1961) and Wang and Huber (1967) for ungauged watersheds.

### Monthly Rainfall-Runoff Models for ungauged catchments

(1) Based on extensive literature and studying the varius structures of monthly rainfall-runoff water balance type models, Mehrotra (1997) has proposed a simple structure of water balance model based on single parameter, which was applied and tested on some watersheds of Saurashtra region (arid and semi-arid watersheds) and some humid and semi-humid watersheds. The value of the parameter can be either found by making soil survey or by fitting the model on the

rainfall-runoff data of nearby gauged catchments. Mehrotra found the value of the parameter as 470 mm for Saurashtra region and average efficiency of the model was found as around 75 %. For humid and semi-humid watersheds, the value of parameter was found as 380 mm and mean efficiency as around 90%. The model structure is explained in brief in the following sections.

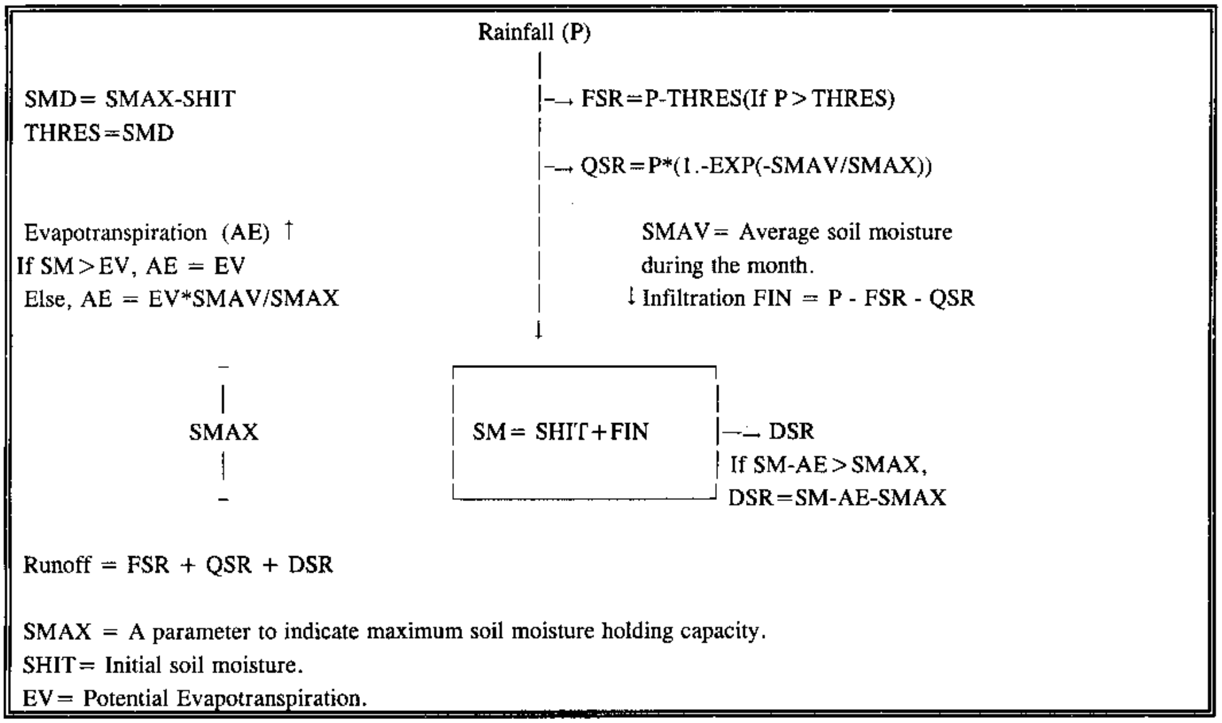

**Fig. 1 :** **Structure and schematic representation of a monthly rainfall-runoff model for ungauged catchment**

This model considers only soil storage. Single parameter SMAX is used to represent the soil moisture holding capacity of the soil storage. Fast surface runoff (FSR) is the portion of rainfall in excess of the soil moisture deficit of the soil storage. Quick surface runoff (QSR) depends on the average soil moisture condition of the soil storage. It follows an exponential function. Evaporation from the soil storage is governed by the average soil moisture available in the soil storage. The structure of the model is given in Fig. 1.

(2) Kothyari (1995) proposed a simple method for the estimation of monthly runoff for the monsoon season by analysing the data of 31 small catchments of India. The models proposed by him two parameters values of which can be estimated from the catchment area, the percentage of forest

cover, monthly average temperature. The goverinig equation takes the form:

$$R(I) = K(I)P(I) + K(I)^{n(I)} \{1 - K(I-1)\}P(I-1) \tag{57}$$

where R(I) = monthly runoff during the ith month; P(I) = monthly areal rainfall during the Ith month; P(I-1) is the monthly areal rainfall during the (I-1)th month; and K(I) and n(I) are parameters for the Ith month with $K(I) < 1.0$ and $n(I) > 1.0$. The coefficient K(I) and exponent n(I) are assumed to vary with time, I, and from one catchment to another. Value of K(I) is given by:

$$K=260.9T^{-2.02} F_A^{-0.05} A^{0.05} \tag{58}$$

where T is average monthly temperature of the catchment in °C, A is catchment area in km² and $F_A$ is the percentage of forest area in km². Values of n(I) for some basins may be obtained from the Table 5.

(3) A model to estimate runoff volumes from generally available geophysical data has been proposed by Agung and Cordery (1995). Input to the model consists of monthly rainfall, evaporation and one fitting parameter. Physically realistic relationships have been developed between other parameters and readily measurable catchment characteristics. The authors applied the model to 18 catchments in New South Wales, Australia.

**Table 5 : Average values of n(I)**

| Basin name | July | August | September | October |
|---|---|---|---|---|
| Damodar | 2.08 | 2.10 | 2.53 | 1.97 |
| Barakar | 2.35 | 2.26 | 2.47 | 2.22 |
| Chambal | 2.42 | 2.42 | 2.10 | 2.42 |
| Mayurakshi | 2.04 | 2.45 | 2.19 | 2.48 |
| Lower Bhawani | 1.70 | 2.05 | 2.60 | 2.15 |
| Ram Ganga | 2.15 | 2.02 | 2.08 | 2.25 |

## FLOW-DURATION CURVES FOR WATER AVAILABILITY COMPUTATIONS

The flow duration curve is a graphical tool which gives an idea about the variability of water flow in a very simplistic manner. The flows for various levels of dependability for these gauged sites may be estimated from this curve. Flow duration curves have been advocated for use in hydrologic studies such as hydropower, water supply, and irrigation planning and design. Additionally, flow duration curves have been applied to stream pollution and water quality management problems. With increasing attention focused on surface water-quality management, many agencies routinely require estimates of low-flow statistics to assure the maintenance of water-quality standards. The use of flow duration curves only provide the estimates of flow which may be available on an average at a particular location without any reference to the sequence in which these flows would be available. And therefore, the use of these flow duration curves must be restricted to the problems where one is interested in an estimate for an "on an average" availability of flow and not in the exact sequences of flows at the site of interest.

As the dimensions or size of a water resource project is based on the reliability of water flow at the site therefore proper estimation of water availability is therefore very essential. The water availability should be based for an irrigation project on 75%, for power projects on 90%, and for drinking water supply on 100%. Reliability can be estimated from the streamflow record, the characteristics of which can be depicted by flow duration curve. The flow duration curves have frequently been advocated for the use in hydrologic time series studies such as hydropower, water supply, and irrigation planning, design and management

For the preparation of a flow-duration curve, three situations may arise depending on the data availability : 1) observed data of sufficient length is available; 2) data of limited period is available and; 3) no records are available. The preparation of flow duration curve under these three circumstances of data availability is discussed below.

When observed data are available at the desired location, the simplest form of duration curve is constructed by listing all the items in the data under consideration in their relative order of magnitude (either increasing or decreasing). The diagram on which the curve is to be plotted is then divided into vertical segments of equal width, the number of segments being equal to the total number of items in the data. Each item is laid off to the proper vertical scale in order of magnitude, in its proper segment.

A horizontal line is drawn across each segment at the point thus noted, the result forming a "block diagram" or "histogram". If the centre points of the top of all blocks of such a diagram are joined by straight lines, a "duration curve" will be obtained. Each item of the duration series is then plotted at the mid point of its respective sub-division of the diagram. Thus, the "time" scale of the duration curve (axis of abscissas) has been subdivided into the same number of parts as there are separate items to be plotted. It is generally more convenient to plot these items on a percentage scale or %-of-time basis.

Analytically, this curve may be obtained by first arranging all the items in descending order of magnitude and then assigning each item a cumulative frequency equal to [i/(N+1)] where i = the position of a item in the ordered series and N is the total number of items in the record. A set of N points having y coordinate equal to magnitude of flow and x coordinate equal to cumulative frequency [i/(N+1)] may thus be obtained which if plotted results in a flow duration curve. Since it is only a sample of data which is available for the analysis, it is very difficult to infer about the characteristics of the population. In such situations advantage may be taken, if it may be assumed, on statistical basis, that a particular sample data belongs to a certain type of mathematical function. For series with statistically insignificant serial correlation, the flow duration characteristics may be modelled with a probability distribution model (gaussian, log-normal, etc.), as it is common practice when dealing with statistical data. In this, it is assumed that the frequency curve obtained from the flow data can be adequately represented by a suitable standard probability density function. Normal distribution function is one of the most commonly used distribution function. Where the data record length is very large this distribution may invariably be found to represent the observed frequency curve. An extension of using this distribution is to use the log-normal distribution in which the logarithms of the flow values tend to produce a normal distribution function.

When only short records are available, the flow-duration curve based on a short record is extended with the help of the flow duration curve of a nearby site with long-term data. Here, the procedure used could be any one of the following: 1) rigorous-data extension for a short-period site by correlation techniques; 2) regional analysis of flow duration characteristics and; 3) approximate adjustment of short-period flow duration statistics on the basis of short-term and long-term statistics of long-term sites.

When no records are available, the flow-duration curves of nearby sites can be transferred using specific discharge indices (flow per sq km). Losses and gains in different area/reaches involved in such a transfer may be considered subjectively as far as possible.

When it is required to find the flow duration curve for a location on the stream with either short term or no records then regional flow duration curves which are intended to be developed for a region as a whole are used. This region is a comparatively bigger area which is expected to be hydrometeorologically homogeneous in its behaviour. The regional model is evolved on the basis of data available for a few gauged sites in the same region or transposed from similar region in its vicinity. This regional model is then employed to know the flow duration curve for any ungauged location of interest in a region. Availability of such regional flow duration models would be of significant help in estimating the potential of hydro power in vast hilly regions of the country and also reducing the time of investigation and implementation of individual small hydro power projects besides building greater confidence while designing. The development of regional flow duration curve is further discussed in the separate section.

When the runoff record at a given point on the stream is to be compared with the record at another point on the same stream, or when the records on different streams are to be compared, it is desirable to divide the original streamflow values by the average runoff occurring at the point of record. This average being obtained from a record of as great a length as possible. A duration curve constructed from the reduced record will give the flow "in terms of mean runoff ($Q/Q_{mean}$)", and will reveal as much detail as is there in the one obtained from the original data, but in a more convenient form for a comparison with other duration curves

Expressing the flow in terms of mean flow has another advantage that in cases where the record length of flow data at individual sites is not adequate for any meaningful statistical analysis, the data of many sites may be pooled together. Since, the data at different sites is comparable as explained above there is no difficulty in pooling up the data. Now, having pooled up the data from many sites the number of data becomes sufficient for drawing some meaningful statistical inference.

These flow-duration curves are listed in a tabular form giving discharges at selected dependabilities (e.g. 5%, 10%, 15%, 30%, 50%, 75%, 90%, 95% and 97%) for all locations.

### GENERATION OF LONG SEQUENCES OF FLOW VOLUMES

Decide the overall approach whether historical or generated sequences of flow volumes are to be used. In case of limited available records, the stepwise procedure followed is:

1. Decide the method of extension of the short-term record. Compile the historical record of long-term related phenomenon (e.g. rainfall over the same area or flow records of a nearby site or sites) for an appropriate time unit.
2. Use statistical correlation (simple or multivariate) techniques to extend the record. Give details of transformation of data, correlation coefficients, explained variance and standard error of regression estimates. When the year has been divided in a number of time units, the correlation may be required for each time unit. However, for some lean periods, no correlation may be possible or necessary. Also, the runoff-to-runoff correlation can sometimes be considered stationary and a single correlation may hold for a number of time units.
3. Check the overall acceptability of correlations.
4. Consider the inclusion of a random component and generation of a single or alternative likely historical series including random components. This will be more important when correlations are not very strong, because non-inclusion of a random component will reduce the flow variance in the extended portion.

### INFLUENCE OF TIME INTERVAL FOR ORIGINAL RECORD

Flow duration curves for small projects, where very little storage is normally available, must atleast be prepared with the average daily flow data. However, many times the record of flow is available only at ten daily basis. Also, use of data on a daily basis involves an immense amount of work. It is often convenient to average these data for successive weeks, months, or years. Obviously, any lengthening of the time interval used for each item to be plotted in the duration curve will result in a loss of detail since in reality the flow will not generally be constant over the period of time for which they are averaged. It is important, therefore, to know what the relative effect will be on the accuracy of the duration curve used by such changes in the time interval.

As the time interval is increased in the length, the lower end of the duration curve will rise and the upper end will be lowered. This is to be expected, since a longer time interval used in averaging process must result in eliminating many of the smallest and the largest items in the record. The total area under each of the curves is the same, because this area represents the total volume of runoff during the entire period of record, and this volume is not affected by changing the time interval.

The relative effect of varying the time unit will not be the same with all the streams. Where the flow is not subject to sudden changes, it will be almost constant over considerable periods of time. With such a stream, the daily and weekly duration curve would be almost identical; and the monthly duration curve would not differ greatly from the daily curve. On the other hand, if the stream is "flashy", with floods lasting only a few hours, or days, there will be an appreciable difference between the daily and weekly curves; and the monthly curves will involve a considerable percentage of error as compared with the daily curve.

However, it is often necessary to use the longer time unit instead of the daily flow duration curve for two reasons: (1) With some streams, only the average flow at longer time unit are available over part or over all the period of record; and (2) in most preliminary investigations, the time and labour required to prepare the daily flow duration curves are so great that a longer period must be used. However, now a days with the advent of computers this difficulty is no faced felt and tendency of working with smaller time interval has thus increased. In any case, when the curve at longer time unit is used, its limitations should be recognized; and if the stream is at all "flashy", the results should be properly discounted when estimates of plant capacity and output are being made.

### DEVELOPMENT OF REGIONAL FLOW DURATION CURVE

The flow duration curve for a catchment, for which insignificant or no flow data is available, is established on the basis of regionalisation procedure. For this, a region is identified which is a comparatively bigger area than the individual ungauged catchments but adequately smaller so that homogeneous hydrometeorlogical conditions exist, on an average, across the region. All the gauged sites in the region are first identified. Then, on the basis of the flow characteristics at these sites, some relationship may be evolved which are representative of the conditions of flow regime throughout the region. This relationships usually incorporate a few parameters which may be evaluated for any ungauged site in the region on the basis of climatic and physiographic characteristics of the catchment. Keeping this in mind the Institute has proposed a methodology for development of regional flood duration curve which has been tested on some hilly regions of the country and produced satisfactory results (NIH,1998). This methodology is described in brief here.

## Regional Flow Duration Model

The methodology of developing the regional flow duration models for this study is based on the following considerations:

(a) The length of record of the flow data at most of the gauged sites is not very large. To draw valuable statistical inference the data from all the sites in the region is pooled up after expressing flow values at each site in terms of the mean flow. The mean flow is estimated from the limited flow data record of each individual gauged catchment and may not represent the long term mean at a site.

(b) The mean flow is considered to be a function of only the physiographic characteristics of the catchment and the model for this is evolved for each region individually. Though, ideally, the mean flow would be a function of climatic and geologic characteristics together with the physiographic characteristics but the climatic and geologic characteristics are not included in the model on account of (i) non-availability of geologic data at the scale of the gauged catchments, (ii) climatic and geologic data may not be available for individual ungauged catchments, which are very small in size. Thus, there would not be any advantage by including these factors when they are not known with adequate precession for ungauged catchments, (iii) variability in the long term average rainfall within the region may not be very high and so may not be very important.

On the basis of the above, a structure of the regional flow duration model is proposed for this study. The step-by-step procedure for developing this model, for any region of interest, is given hereunder:

(I) Identify all the gauged catchments in the region and collect observed flow and physiographic data for each of them.

(II) Process the flow data to get the average ten daily flow series for each gauged catchment. To avoid any bias this series must contain equal number of each of the 36 ten daily values of the year. Hence, the flow data series ($Q_i$) would always have multiple of 36 data points in it.

(III) Next, find the long term average ten daily flow ($Q_{mean}$) for each of the gauged catchments with the help of flow series available at step (II).

(IV) Non-dimensionalise the flow data series ($Q_i$) for each catchment by dividing each data value in the series by its respective long term average ($Q_{mean}$) obtained at step (III) to get the flow series in terms of mean flow ($Q_i/Q_{mean}$).

(V) Combine the non-dimensionalised flow data series of all the gauged catchments into one series.

(VI) Evaluate the basic characteristics of the combined series to see if the series is normally distributed. In most of the cases it would not be normal owing primarily to a shorter length of the data series. Transform the combined series obtained at step (V) using power transformation into another series which has the characteristics of normally distributed series. This is done as explained below:

Let Q and W implies the corresponding elements of original and the transformed series respectively. The power transformation is achieved using the transformation formula given by:

$$W = (Q^{\lambda} - 1)/\lambda \qquad \text{when } \lambda \neq 0 \qquad (59)$$

$$\text{and} \quad W = \ln(Q) \qquad \text{when } \lambda = 0 \qquad (60)$$

here $\lambda$ is an exponent which may either be obtained by trial and error procedure or any other suitable optimisation technique so as to give a normalised W series. The W series is considered to be a near normalized series for that value of λ which reduces the coefficient of skewness of W series to nearly zero and maintains the coefficient of kurtosis as 3.

Thus, for a region the value of parameter λ is obtained.

(VII) The statistics like mean ($\mu_W$) and standard deviation ($\sigma_W$) of the transformed series W are estimated using maximum likelihood method.

(VIII) Using the normal probability distribution, estimates of the flows in transformed domain ($W_D$) for various levels of dependabilities (D) like 25%, 50%, 60%, 75%, 80% and 90% etc. is made.

For this, first the standardised flow ($Z_D$) corresponding to any probability level (D) may be obtained by using widely available table for frequency factors for standard normal distribution ($\mu=0$, $\sigma=1$). This standard flow is converted for the case of a series having mean and standard deviation as $\mu_W$ and $\sigma_W$ respectively by using the transformation:

$$W_D = \mu_W + Z_D \sigma_W \tag{61}$$

(IX) Since, the estimates of flow (in terms of mean flow) at step (VIII) are in the transformed domain they are brought to the original domain by using the inverse transformation as:

$$\left(\frac{Q}{Q_{mean}}\right)_D = [W_D \lambda + 1]^{\frac{1}{\lambda}} \qquad \text{when } \lambda \neq 0 \tag{62}$$

and $$\left(\frac{Q}{Q_{mean}}\right)_D = \exp\,[W_D] \qquad \text{when } \lambda = 0 \tag{63}$$

(X) From the available physiographic data the catchment areas (A) for each gauged catchments may be obtained. Knowing the long term average flow ($Q_{mean}$) for each gauged catchment a relationship between catchment area (A) and the corresponding long term average flow ($Q_{mean}$) is established in the form:

$$Q_{mean} = C_1 A^2 + C_2 A \tag{64}$$

Here, $C_1$ and $C_2$ are the coefficients which are obtained by regression analysis.

(XI) Now, $Q_{mean}$ for any ungauged catchment in the region may be obtained by the regional relationship established at step (X). This value of $Q_{mean}$ when multiplied by the factor $(Q/Q_{mean})_D$ obtained at step (IX) gives the required D% dependable flow ($Q_D$) for that ungauged catchment.

(XII) The steps from (IX) to (XI) may be repeated to obtain the flow corresponding to any desired level of dependability for any ungauged catchment within the region.

Thus, a regional model may be evolved for any region of interest and then be subsequently employed to estimate the flows for desired levels of dependability for any ungauged catchment located within the region.

## CONCLUDING REMARKS

In this report water availability computations under different data availability scenarios are discussed. There is no standard method available which can be assigned for a particular case. Depending upon the availability of data, catchment information, availability of time, purpose for which this information is being collected and computing facilities available designer has to make a choice out of available methodologies. Rainfall- runoff relationship and runoff process being a complex process varies form catchment to catchment also from time to time. As data availability becomes poorer, results obtained using this data subjected to greater uncertainty. However, with the research and development in the area of hydrology, now-a-days more and more methods are being proposed to deal with scanty data availability situations.

## *References*

*Agung B. Ibrahim and Ian Cordery, 'estimation of recharge and runoff volumes from ungauged catchments in eastern Australia', Hydrological Sciences Journal, vol. 40(4), August, 1995, p.499-515.*

*Anderson, M. G. & T. P. Burt. eds. (1985), 'Hydrological forecasting', New York, John Wiley.*

*Barnes, F.B. 1954. Storage required for a city water supply. J. Inst. Eng., Australia, 26, pp. 198-203*

*Blackie, J.R. and Eles, C.W.O.: 1985, Lumped catchment models. In: Hydrological Forecasting, ed. M.G. Anderson & T.P. Burt, Wiley, New York, USA.*

*Eric Servat and Alain Dezetter. 1993, Rainfall-runoff modelling and water resources assessment in northwestern Ivory Coast. tentative extension to ungauged catchments. J. Hydrl, 148, pp. 231-248.*

*Haan, C.T. , D. M. Allen and J.O. Street, 1976. A Markov chain model of Daily Rainfall. Water Resources, Vol. 12, No. 3, pp. 443 - 449.*

*Hazen, 1. 1914. Storage to be provided in impounding reservoirs for municipal water supply. Trans. Amer. Soc. Civil Engg., 77, pp.1539-1669.*

*Jayasuria, L.N.N., McMahon, T.A. and O'Neill, I.C. 1991. Development of simplified two-parameter rainfall-runoff model with potential for ungauged catchment application. In:Proc. Hydrol. Water Resour. Symp., 498-503, Inst. Engrs. Aust. National Conf. Publ. no. 91/22.*

*Khosla, A. N., 'Analysis and utilization of data for the appraisal of water resources', J. Irrig. Power, India, Oct., 1949, p. 410-422.*

*Kothyari, U.C., 1995, 'Estimation of monthly runoff from small catchments in India', Hydrological Sciences Journal, vol. 40(4), August, 1995, p.533-541.*

*Lane W.L., (1979) Applied stochastic techniques (LAST computer package), user manual, Division of Planning Technical Services ,Bureau of Reclamation ,Denver , Colorado.*

*Manley, R.E., 1978. Simulation of flows in ungauged basins. Hydrol. Sci. Bull. 23, 85-100*

*Mehrotra, R. 1997, NIH Report No. CS(AR) 211, 'Comparison of monthly rainfall-runoff models', National Institute of Hydrology, Roorkee.*

*Mejia and Rousselle, J. (1976) Disaggregation models in hydrology revisited. Jour. Water Reso. Res.12, 2, 185-186.*

*Mimikou, M. & Rao, R. 'A regional monthly rainfall-runoff model', J. Wat. Resour. Plan. Manag., ASCE 109(1), 1983, p. 75-93*

*Murry, J. N. and Prsad, K. N., 'Relation between basin characteristics and discharges (peak and annual)' In: A study of Indian catchments, (7th Symp. Water Resources, Banglore, India) A17, 1971, p. 1-30.*

*Nathan, R.J. & McMohan, T.A. 1990. The estimation of low flow characteristics and yield from small ungauged rural catchments. AWRAC research project Report 85/105.*

*National Institute of Hydrology, 'Development of regional Flow Duration Models for Prepration of Zonal Plan for Small Hydro Power Projects', UNDP-GEF India, Hilly-Hydro Project, July, 1998*

*Roldan, J. and D.A. Woolhiser, 1982. Stochastic daily precipitation models (1). A comparison of occurrence processes. Water Resources Research, Vol. 18 No. 5, pp.1451-1461.*

*Salas, J.D., Delleur, J.W., Yevjevich, V., and Lane, W.L. (1980) Applied modelling of hydrologic time series, Water Resources Publications P O Box 2841, Littleton, Colorado 80161, USA.*

*Singh, V. P.: 1989, Hydrologic systems - Watershed Modelling, Vol.- II, Prentice Hall, New Jersey.*

*Srikanthan, R., and McMahon, T.A. (1982) Stochastic generation of monthly stream flow Jour. of the Hy. Div.,ASCE, Vol. 108, No. HY3, 419-441.*

*Sudler, C.E.. 1927. Storage required for the regulation of streamflow. Trans. Amer. Soc. Civil Engg., 91, pp. 622-660.*

*Taffa Tulu. 1991. Simulation of streamflows for ungauged catchments. J. Hydrol. 29, 3-17.*

*Thomas, Jr., H.A. and Fiering, M.B. (1962) Mathematical synthesis of streamflows sequences for the analysis of river basins by simulation. In : A.A. Maass M.M.Hufschmidt, R. Dorfman, H.A.*

*Valencia, D.R. and Schaake, J.C. (1973) Disaggregation processes in stochastic hydrology. Jour. Water Reso. Res.9, 3, 580-585.*

*Woolhiser, D.A. and J. Roldan, 1982. Stochastic daily precipitation models (2). A comparison of distribution.*

# ANNEXURE—I

## GUIDELINES FOR PREPARATION OF HYDROLOGY, VOLUME OF DETAILED PROJECT REPORT

CHAPTER

I. General Climate and Hydrology
II. Hydrological data requirements
III. Compilation and processing of Hydrological data
IV. Preparation of Hydrological inputs for Simulation
V. Preparation of Hydrological inputs for studies other than Simulation
VI. Simulation (working tables) for testing performance
VII. Effects of project development on Hydrologic regime

# CHAPTER 1

## 1.0 General climate and hydrology

This chapter shall provide sufficient information about climate and general hydrology of the hydrologic region and also provide specific information in respect of areas and reaches of interest appropriate to the plans of development.

NOTE :—Various areas and reaches of interest from the hydrologic point of view considering different developmental possibilities have been classified as E1 to E 13 as listed in Enclosure-A.

## 1.1 General information about region (refer E5—Enclosure-A)

1.1.1 Topography—types of climate seasons—type of monsoon causing rainfall—general hydrologic regime of the rivers—history of important historical storms and floods—geology—existing development of surface and ground water.

The above details shall be supported by the following maps and tables :

(a) One or more small index maps of size 25 × 20 cm (approximately) showing the boundaries of all the areas and reaches of interest for the various alternatives.

(b) Annual normal rainfall map of the region (scale 1 : 1,000,000).

(c) Tables or bar charts showing monthly normals and extremes of rainfall—number of rainy days—temperature, humidity—wind speed—pan evaporation—potential evapotranspiration (ETO) etc. for observatories (at least two in and around areas of interest); at least one of the observatory selected shall be in or nearest to the command area—(Refer climatological tables, other IMD publications etc.).

(d) Tables or bar charts giving average, maximum and minimum monthly seasonal and annual flow data for hydrological stations maintained by the State/CWC/GBWRO/other agencies (include stations with long period data).

## 1.2 Specific Information

Specific information required for different reaches and areas of interest relevant to the project shall be furnished.

### 1.2.1. Drainage basin (Refer E1 of Enclosure-A)

(a) Index map showing soil erodibility characteristics and infiltration.

The map shall show

(i) Land having gradient less than 1 per cent, 1-3 per cent and more than 3 per cent.

(ii) Present land use classifying land under forest cultivated area, fallow land, land under urban and other uses.

(iii) Soil types

NOTE :—In preparing the above maps information available in the revenue records District Gazetteers, Census report, Irrigation Commission's report shall be utilised.

Where sedimentation is of serious nature (Refer Chapter IV Item 4.3) land use data for two or more distinct periods shall be included as far as possible.

(b) Table giving the size of the drainage area at all control points indicating areas covered by lakes, swamps, permanent snow/glacier and points of diversion (natural or man made) for diverting flows from or into the drainage area.

The drainage area computations shall be made using sufficiently large scale maps (scale 1 : 200,000) and condensed maps showing all the features shall be furnished.

(c) Area altitude curves for orographic region having sizeable area above El 1000 m.

(d) Water quality data depending upon the nature of development.

### 1.2.2. Command area (Refer E2 of Enclosure-A)

(a) Table or charts presenting monthly normal rainfall and coefficient of variation for a few stations in or near command area.

(b) Monthly normal evaporation for few stations in or near the command.

(c) Infiltration characteristics of soils.

(d) Ground water behaviour supported with data of fluctua tion (maximum, minimum and average over a period).

### 1.2.3. Floods and drainage (Refer E3, E4 & E10 of Enclosure 3-A)

(a) River profile showing flood levels and river cross-section.

NOTE :—The Longitudinal section and cross-section through the reach of interest are required for making computations of flood profiles, working out gauge discharge and water rating curves, checking reservoir backwater studies, deciding channel capacities and for estimating historical flood discharges from flood marks by hydraulic calculations etc.

The river L-Section should be presented to such a vertical scale that the differences in normal low and high water will be presented by about 2 cm. It would show both banks, bed levels, normal, high and low flow levels, historical maximum flood levels with years of occurrence, and position of important towns, gauge discharge stations, bridges, existing and proposed structures as also the position of river cross-sections. River cross-sections covering the reach of interest shall also be presented. The spacing of cross-sections shall depend on river slope and uniformity of the channel. For smaller slopes or more uniform channels larger spacings can be adopted.

(b) Information of past floods and past events of drainage congestion giving levels, discharge flooded areas and depth and duration of submergences.

(c) Notes about flood protection and drainage works already sanctioned/executed and their performance supported by index plans.

(d) Notes about the problems of bank erosion, aggradation, degradation and meandering of rivers.

### 1.2.4. River Geometry (Refer E6 and E9 of Enclosure-A)

(a) River profile and cross-section and roughness coefficient for the reaches relevent to the project details (Refer note 1.2.3 above).

### 1.2.5. Ground water recharge (Refer E7 of Enclosure-A)

(a) Details about the ground water behaviour and infiltration characteristics of the soils in the recharge area (For type of information refer para 1.2.2 (d) above).

**1.2.6. Reservoir area (Refer E8 of Enclosure-A)**

(a) Monthly average pan evaporation data for a station in or near the area.

(b) Elevation-area-capacity curves and methodology used in the computation.

**1.2.7. Other water usage (Refer E11 & E13 of Enclosure-A)**

(a) Water quality data indicating both chemical and biologic quality temperature and other quality parameters. Indicate changes in these parameters from season to season.

**1.2.8 Navigation (Refer E12—Enclosure-A)**

(a) River profile and cross-section (Refer Notes 1.2.3 above).

(b) Low flow discharges and depth data.

(c) Historical changes in the levels and cross-section.

(d) Problems of bank erosion aggradation, degradation and meandering of river.

(e) River training works already sanctioned/executed and their performance.

**1.3.1. Data availability**

Description of the available meteorological and hydrological data supported by inventories in the form of bar diagrams indicating source—location and altitude of the station—drainage area where appropriate—period of availability in respect of stations within the areas of interest and surrounding regions shall be furnished in respect of the following :

(a) Rainfall and snowfall

(b) Pan evaporation

(b) Climatological parameters like temperature—humidity—wind—sunshine etc.

(d) River, gauge and discharge

(e) Sediment (suspended and bed load) inflow and grain size composition

(f) Water quality

NOTE :—A map to a scale of 1: 200,000 or 1: 1,000,000 depending upon the size of the area involved showing the location of relevant meteorological and hydrological stations shall be furnished.

# CHAPTER II

## 2.0 Hydrological Data Requirements

This chapter shall discuss the type and extent of Hydrological Inputs required for the proposed plan (s) of development.

### 2.1 Alternatives and classifications

The type and form of Hydrological Inputs for simulation (working tables) and other studies depend upon the type of structures (which can be classified based on the element of storage) and on the contemplated use of water and storage space.

The classification of alternative plans by storage have been indicated as A1 to A5 as listed in Enclosure-B and by use as B1 to B11 as indicated in Enclosure-C.

NOTE :—Alternative plans of development shall be discussed and their classification determined by storage or by use utilising the information classified in Enclosure-B and Enclosure-C

### 2.2 Inputs

#### 2.2.1. Type of inputs

The inputs required for simulation and other studies for the development in question shall be discussed.

The nature of inputs required has been grouped as C-1 for simulation studies and C-2 for studies other than simulation in Enclosure-D. The inputs required for the study of a particular type of development can be determined using Enclosure-E and Enclosure-F. Wherein various combinations are indicated

#### 2.2.2 Time unit for simulation studies

The time units applicable to the various type of projects are given in Enclosure-C. The information given in this Enclosure shall be utilised in deciding the time units of the hydrological inputs for particular type of development.

#### 2.2.3 Hydrological inputs

In fixing the length of hydrological inputs for simulation type of development and variability of inputs shall be kept in view. Brief Guidelines for fixing the minimum length of data required is given below :

| Type of project (Enclosure-B) | Minimum length of data for use in simulation |
|---|---|
| A1 and A2 | 10 years |
| A3 | 25 years |
| A4 | 40 years |
| A5 | Depending upon the predominant element (A1 to A4) |

### 2.3 Requirements of the inputs for the project

Taking into account the requirements of the project discussed under para 2.1 and 2.2 above the inputs (including required for simulation, flood studies and sedimentation) for the various components/at various control points shall be determined and discussed in details.

# CHAPTER III

## 3.0 Compilation and Processing of Basic Hydrological Data

3.1 Hydrological investigations specially carried out for the proposed project keeping in view the Guidelines given in Chapter 3.4.5 of the Detailed Project Report shall be discussed. The details of the specific data collected for the purpose shall be furnished.

### 3.2 Data from other sources

All the basic/processed hydrological data (flow, sedimentation, water quality etc.) available from the various sources as relevant to the project shall be collected, compiled and discussed. The source of such data collected shall be indicated at the appropriate place.

Where processed data is available need or otherwise of further processing of the data shall be indicated.

### 3.3 Processing of data

#### 3.3.1 Quality of data

(a) Methods of measurement observation of various types of hydrological and hydrometeorological data, standards followed, instruments used, frequency of observation etc. shall be discussed itemwise viz., flow, sedimentation, gauging, temperature, humidity, evaporation etc.

(b) Details of history of station, shifts in the location, shifts in the rate curves shall be identified. Sample calculations for discharge and sediment load shall be furnished. Mention shall be made as to whether discharge data is observed or estimated. Indicate methods of estimation.

(c) Discuss development of stage discharge curves at discharge site bringing out the extrapolations involved. The extrapolations shall be verified by other methods such as hydraulic calculations etc.

#### 3.3.2 Filling up of short data gaps

The method used shall be discussed. The following or some of the techniques which can be used for gap filling :

(a) Random choice from values observed for that period.

(b) Interpolation from adjoining values by plotting a smooth hydrograph (for runoff alone).

(c) Using the average proportion with normals for the adjoining stations.

(d) Double mass curve techiques.

(e) Correlations with adjoining stations either of the same hydrologic element, or of different hydrologic element.

(f) Auto correlation with earlier period at the same station.

(g) Any other.

### 3.4 Adjustment of records

3.4.1. The adjustments of flows (and sediment) to natural and virgin conditions for historical uses in the upper reaches and the manner in which this has been done shall be discussed duly supported by the withdrawal data, reservoir operation data and irrigation statistics. Where adjustments due to upstream storage(s) are made, such storage changes and evaporation losses are to be properly accounted for.

Apart from adding upstream withdrawals return flows have to be substracted.

NOTE:—(1) The adjustment of the observed flows/sediment data may not be nesessary if

(a) The utilisation by upstream projects has been same throughout the period of observation of flows and sediment

(b) If the pattern and quantum of usage has not changed appreciably or with a definite trend

(2) Adjustment with the flow and sediment records shall be required in other cases e.g. where apprreciable changes in land use have taken place

(3) Adjustment of floods and low flows to remove the effect of upstream regulation may be required where this is appreciable

### 3.5 Consistency of data

#### 3.5.1. Internal

The study of consistency of the observed data at specific control points and corrections, if any, made shall be checked and discussed

The check can be done by study of stage discharge relationship, and sedimentation rating curves for different periods Large veriations, if any, shall be investigated, corrected and explained suitably, if required.

#### 3.5.2 External

The consistency of the observed data shall be discussed with reference to the rainfall in the project catchment and observed data (yields and sediment loads) in adjacent locations/basins.

NOTE : The consistency can be checked by :

(a) Comparing monthly and annual rainfall with corresponding runoff.

(b) Comparing average annual specific flow experessed as ham/sq km with corresponding figures at other sites of the same river or adjacent basins.

(c) By comparing the hydrograph of daily discharge at the control point with adjacent sites etc.

(d) By use of double mass curve techniques.

Details of the study made for various hydrological observations at the control points and sites maintained by the CWC/GBWRO/States and other agencies shall be summarised and presented as follows :

(a) Average annual/seasonal/monthy flow volumes expressed as depth of water over drainage area.

(b) Average maximum and minimum discharge (cumec/sq km for concurrent period.

### 3.6 Presentation of data

#### 3.6.1 Data for simulation studies

The processed data shall be compiled and furnished keeping in view the hydrological inputs required for the studies for the development in question (Refer Enclosure-E and Enclosure-F).

The data shall be compiled for appropriate time unit (Refer Enclosure-G).

NOTE :—

(1) The avearge for each time unit and totals and averages for months/seasons/year (June to May) shall be furnished.

(2) Where gap filling has been carried out and basic data adjusted, suitable footnotes to the effect shall be given.

### 3.7 Data for studies other than simulation

Data on the annual maximum floods (peak discharges and levels) for all sites of interest shall be furnished for the entire period of record.

Flood hydrographs (plotted on the basis of hourly gauge observed for a few large events for all sites of interest shall be included. These should cover the entire rise and fall of the flood including three days period antecedent and following each flood. The concurrent daily and hourly rainfall data for all stations in and near the drainage area shall also be included.

# CHAPTER IV

## 4.0 Preparation of hydrologic inputs for simulation

This chapter shall discuss the details and results of the analysis made for preparation of the various hydrologic inouts required for simulation studies to supplement data presented in para 3.6.1. of Chapter-III.

## 4.1 Water inflows

### 4.1.1 Storage projects

The overall approach whether historical for generated sequences of flow volumes used shall be indicated.

### 4.1.1.1 Data Extension

The studies and methodology used for extending short-term runoff series to desired length of time (Chapter-II Para 2.2.3) shall be discussed covering details of type(s) of correlation transformation of data, correlation coefficient, standard error, etc. These studies can be done as follows:—

(a) Correlating runoff data with concurrent data on rainfall of long term stations in the same catchment or data of runoff of adjacent long-term station and applying these correlations developed to past data of long-term stations of rainfall/runoff.

(b) Such correlations shall be developed for each time unit selected.

Note :—

(1) Rainfall/runoff correlation may not be feasible or nessary for non-monsoon period.

(2) Overall Acceptability of correlation shall be checked.

(3) Random components may be considered where Correlations are not very strong.

### 4.1.1.2 Data generation

The approach used may be discussed giving the type of model and its suitability to the problem on hand, its parameters and their evaluation, validation of model and generation of flow data. Two approaches that can be used are :

(a) Stochastic modelling (Time series)

(b) Conceptual modelling

For the stochastic approach following details may also be included.

—Trends and cycles in the data, their physical justification and the necessity or otherwise of removing these.

—Auto-correlation in data, its physical explanation, need for modelling autocorrelation, possibilities of smoothening auto-correlation values from regional studies.

—Frequency distribution of random error component

—Generation of random numbers

—Where more than one site is involved, correlation between random error components of different series and method of flow generation at different sites

For data generation by conceptual Modelling, details of modelling input data (e.g. rainfall) may be includied together with compilation of output data in appropriate time units.

### 4.1.2 Diversion and small pondages

### 4.1.2.1 Extension of data

The studies and methodology used for extending short-term runoff series to the desired length of time (Chapter-II Para 2.2.3) shall be discussed covering details of type(s) of correlations, transformation of data, correlation co-efficient, standard error, etc.

Techniques as suggested under 4.1.1.1 are also generally applicable for extension of data but the time unit shall be of shorter duration (refer Enclosure-G).

## 4.2 Lake evaporation

Depth of lake evaporation shall be indicated with basis for selected time units (10-daily, fortnightly or monthly). These depths shall be worked out from the averages of long-term data of pan evaporation or climatiological data of a station close to the reservoir after adjusting these to the lake evaporation.

## 4.3 Sedimentation studies

### 4.3.1. Revised area capacity curves

The studies carried out to evaluate the effect of depletion of reservoirs' useful capacity on performance due to sedimentation shall be discussed giving details of methodology adopted, time, period considered, average annual rate of sedimentation and distribution of sediment volume (refer IS : 5477 (Part-II & III) 1969 & CBI&P Publication No. 89 and 19).

Note :—

(1) The studies may not be necessary for diversion structures/works and for storages where the ratio of annual sediment volume as compared to the gross storage is less than 0.1 percent

(2) Studies have to be carried out where the ratio of annual sediment volume to the gross storage is more than 0.1 percent and with percesion for rate exceeding 0.5 percent. In such cases, more than one (depending upon the seriousness of the sedimentation problem) revised adjusted area capacity curves may be required.

(3) Usually for irrigation projects while working out the adjusted area capacity curves a time period of 50 years is used. In case a different time period is used the same shall be justified.

### 4.3.2 Rate of sedimentation

Annual rate of sedimentation shall be estimated from the historical data by analysis of the sediment discharge observations adjusted to long-term conditions where necessary by means of sediment rating and flow duration curves and or hydrographic surveys of the nearby existing storage(s) in similar catchment(s).

Note :—

(1) Allowance shall be made for the anticipated changes in the rate of sedimentation due to the changes in the land management practices.

(2) Allowance shall also be made for the existing upstream projects or projects under construction. No allowance shall be made for future projects

## 4.4 Potential evapo-transpiration and rainfall

The number of stations considered, their locations, details of the data used and methodology adopted for working out the fortnightly weighted mean rainfall of the command shall be furnished and discussed. The methodology for computing ETO shall also be furnished.

This is an important factor for determining the releases at the canal head during different fortnights of the cropping season. The details regarding working out the crop water requirements have been discussed under command area development and modernisation (Volume-III) of this report.

### 4.5 Flood inputs

When planning is based on detailed simulation, flood inputs are required at all control points viz., reservoir site(s) and damage point(s). Further, simulation can be based on historical or generated data.

Where historical flood data are utilised. the methods of transferring the flood hydrographs of available gauging stations to required control points and the manner in which the sub-area flood hydrographs are obtained for controlled and uncontrolled parts of catchment may be discussed and details of studies included.

NOTE : Computation of historical flood hydrographs for sub-areas would involve channel routing. Method of routing and the co-efficients and assumptions may be indicated.

Where flood simulation is to be based on generated data additional details such as monthwise flood frequencies, relations between peak discharges and volumes and inter-relationships between different sub-area floods and lags, consistency of flood volumes and water inflows etc. may be discussed and incorporated in the models and random component considered.

### 4.6 Inputs for water quality

The water quality problems in the various reaches such as salinity control and other aspects for preservation of fish and wild life shall be indicated and details of the water quality characteristics and the water flows or discharges and the interrelationship of such characteristics at different locations shall be discussed including data extension in time and space.

### 4.7 Low flows inputs

The analysis of low flow data available for discharge sites of interest shall be discussed and the low flows determined for the required time units and locations. Method and details of extensioning in time and space shall also be given.

If any trends in low flows have been observed, it may be indicated whether the low flow data has been modified to allow for future changes due to these trends.

### 4.8 Surface to ground water recharge

The details of the analysis made for determining the infiltration characteristics of the recharge area and its variability with time and the estimated rainfall and evaporation during the recharge period may be discussed.

The methodology of preparing the inputs from the available data is also to be discussed.

# CHAPTER V

### 5.0 Preparation of hydrological inputs for studies other than simulation

This chapter shall discuss the studies and their results relating to design flood, design flood level and tail water rating curve etc.

### 5.1 Design Floods for safety of structures

(Recommended procedures given in CWC Mannual—"Estimation of Design Flood" shall be referred to).

5.1.1 The criteria for selection of design flood for each structure taking into account the importance of each structure shall be discussed. The selected floods may be.

(1) Probable maximum flood

(b) Standard project flood

(c) Flood of specified frequency (T-year flood).

5.1.2 Overall approach adopted

(a) Hydrometeorological (design storm and unit hydrograph) approach.

(b) Frequcency approach (including conversion of storm frequencies into flood frequencies).

5.1.2.1 Hydrometeorological approach

(a) Design storms

The details of the transposed storms/Depth Area Duration obtained from the storms considered transposable from those discussed in Chapter-I shall be included alongwith details of moisture adjustment and other types of maximisation, (if any), short period distribution of Storm rainfall and final depth—duration curve adopted for the design storm.

NOTE :

(1) For orographic area, where usually no transposition is done, depth duration analysis of historical storms over the problem drainage area to be made and given.

(2) For large basins the aerial pattern and time sequence for sub-area rainfall shall be discussed.

(3) For complex system (A5 of Enclosure-B) alternate positionings of storm centre will be required to work out sub-area-wise depth—duration curves in each case.

(b) Unit hydrographs

Details of analysis of flood hydrographs and their corresponding rainfall data including plotting of hydrographs from hourly river gauges and gauge discharge rating curves, separation of base flows, computation of mass curve and rainfall data of self-recording raingauges and details of derivation of unit hydrographs shall be given.

NOTE : Averaging and selection of unit hydrographs including maximization of unit hydrograph peak for increase of hydraulic efficiency shall be indicated.

Transfer of unit hydrographs to desired locations wherever made and details of synthetic unit hydrographs for ungauged area shall be included.

(c) Infiltration loss rates, runoff co-efficients:

Selection of infiltration loss rates, runoff co-efficient etc. shall be given based on information drived from observed flood hydrographs.

(d) Design flood hydrographs :

The synthesis of the flood hydrographs shall be indicated giving details of critical sequence of storm rainfall and antecedent storms, runoff and base flows adopted.

NOTE:—In very large catchments (say exceeding 5,000 sq km) the areas may be sub-divided and unit hydro-graphs/flood hydrographs prepared for each sub area by the procedure mentioned above and synthesised to the desired location by channel routing. Where upstream reservoirs exist, reservoir routing shall also be necessary.

5.1.2.2 Frequency Approach

The following shall be included:

(a) Details of analysis of observed series of annual maximum peak discharge—distribution and method of fitting adopted, plotting on probability paper—inclusion of known large historical floods to improve the estimates.

(b) Details of regional flood frequency study, if any.

(c) Reliability and consistency of frequency estimates, confidence intervals.

(d) Method adopted to draw the T-year flood hydrographs, where necessary.

Where long-term flood data are not available, storm frequencies are converted to flood frequency—details of compilation and frequency analysis of maximum rainfall series, infiltration rates and unit hydrographs, short period distribution of storm rain-fall and general flood synthesis criteria shall be furnished.

5.1.3. Comparison of design flood estimates

Comparison shall be made wherever possible with similar estimates for other projects in the region and by interpretation of the discharges from flood levels obtained by hydraulic calculations.

### 5.2 Design floods for determination of flood storage and flood control works.

5.2.1 Flood problem

The problem in various reaches downstream of the storage or at specific reaches affecting the command areas, shall be discussed. The channel capacity at each of the centres when damage begins to occur, shall be indicated.

5.2.2 Degree of protection

The degree of protection proposed shall be discussed.

The degree of protection will depend upon the magnitude of the average annual damage and the cost of the works to give the desired protection. The degree of protection is generally expressed in terms of protection against a flood of specific return period under natural conditions.

Normally, flood protection is provided for a known historical flood, a flood of specified return period, depending upon the assets protected. However, in case of protection of important cities, vital installations etc. are involved, Standard Project Flood may be considered.

5.2.3 Design flood for fixing flood storage and design of structures downstream.

The steps involved and details to be given shall be generally same as indicated under para 5.1 "Design of floods for safety of structures" except that in regions experiencing prolonged rainfall and for larger flood storages, series of floods may have to be considered.

Since design flood at the damage centre(s) may result from several combinations of floods from the controlled (c.i. above the flood storage dam) and uncontrolled parts of the total catchment, it may be necessary to study a few acceptable combinations of the controlled and uncontrolled sub-area floods.

The estimation of design flood at the damage centre(s) for post project conditions shall involve routing of the inflow flood hydrograph of controlled sub-area through the reservoir(s) and through the downstream river reach and combining with the uncontrolled sub-area flood. For reservoir routing, the assumption of initial levels, rules of operation (with or without the benefits of flood forecasting) shall have to be indicated. For river routing, assumption of routing coeffiicients will have to be indicated and justified based on observed rainfall and flood data.

Where channel hydraulics and channle storage characteristics are appreciably changed such as in the case of long embankments, effects of such changes may have to be considered on flood hydrographs, storage discharge relationships and water levels for post project conditions.

### 5.3 Studies for design of drainage in the command area

5.3.1 The problem

The problem of drainage in the command area and the need for surface drainage and sub-surface drainage, if any, shall be discussed.

5.3.2 Surface drainage

The design criteria proposed, shall be discussed. The frequency of rainfall adopted, season of rainfall considered, the nature of crops grown and the acceptable submergence depth and period, shall also be indicated.

NOTE :— The actual design of the drainage system, especially at the enfluence points and outfall points of large systems should take into account the different conditions of flow in the trunk and tributary drains and also outfalling rivers.

### 5.4 Design flood for diversion arrangements

The criteria for selection of construction design flood and studies made, shall be discussed in relation to the proposed plan of river diversion works, construction seasons and schedule etc. selection of peak and volume of flood hydrographs for design shall depend upon the nature of diversion arrangements. The design flood could either be of a specified return period or selected on the basis of economic considerations taking into account the relative risk involved in the occurrence of flood of varying frequency during the construction period and consequent damage and delay in the execution of work.

The design flood can be worked out according to the procedure indicated in para 5.1

### 5.5 Studies for determination of levels for locating structures on river banks and outlets in the dam.

5.5.1 Location of structures

The studies made for determining the levels for locating pumping plants, power houses, roads, bridges, etc. and elevation of the outlets as required, in the project, shall be discussed. The design criteria used and rating curves, if any, developed, shall also be indicated with details of studies.

The methodology of determination of flood magnitude and frequency studies is the same as in para 5.1.

5.5.2 Location of outlets

The details of the studies made shall be discussed.

NOTE (1) The studies for fixing outlet from sediment considerations, shall be as in para 4.3 of Chapter IV.

NOTE (2) The normal time period used for the purpose of studies is 100 years.

### 5.6 Tail water rating curves

The points at which rating curves are required and the approach adopted—hydrologic or hydraulic, shall be discussed.

NOTE:—For important structures, the upper and lower limits of the rating curves may also be computed from statistical methods or by computations using different rugosity co-efficients.

# CHAPTER VI

## SIMULATION STUDIES

This Chapter shall discuss the details of the simulation studies and the conclusions arrived therefrom.

### 6.1 Simulation studies (Working Tables)

The studies carried out for the alternative under consideration shall be discussed in detail explaining all the factors and assumptions that have been made.

Integrated tables shall be prepared in cases where the project under consideration will affect or be affected by other projects in the sub-basin or basin.

NOTE :— Such of the projects which will not have serious impact on the availability of flows can be ignored.

If necessary, allowances (approximate) can be made (as external constraints for meeting the requirements of upstream and downstream projects) while calculating the net inflows available for the projects under consideration without considering these as a part of the integrated system for purposes of simulation.

An indication whether such prohibitions will be applicable always or any allowance can be made during the period of scarcity shall be given.

While discussing the studies, the following shall be furnished:

(i) A schematic plan showing the various projects that have been considered while carrying out the studies, shall be furnished showing the control points, hydraulic structures points where inflows, outflows and return flows have been considered.

(ii) The time unit and the period of simulation with reasons for adopting them, shall be indicated. All the inputs prepared for the studies shall be presented.

Where latest technique of economic evaluation based on discounting procedures is being considered the period of simulation shall be in line with the rate of discounting adopted.

Where carry over storage is involved, it is desirable and necessary to consider a long time series containing cycles of dry years.

(iii) The series used in the simulation—single historical, many likey historical or synthetic—shall be indicated with reasons.

(iv) The various physical limits (constraint)—maximum and minimum limits of storages, diversion capacity of canal/water conductor systems installed capacity of power houses, discharging capacity of the spillways and outlets at different water levels etc. in the studies, shall be detailed.

(v) If control of quality of water etc. is involved in the studies, the manner in which this has been provided for has to be discussed.

(vi) The manner in which the losses/gains to the flows, have been accounted for and allowances made for changes in time distribution in cases involving travel of water over long distances or through storages, shall be discussed.

(vii) If return flows have been considered at any specific points, the basis on which this has been done, and the time span and pattern considered, shall be indicated.

(viii) The demand of all the projects considered in the system for simulation studies including that of the projects under consideration shall be listed along with their time pattern and basis (give suitable references to the documents/studies made). In addition it shall be indicated if the demands considered for existing and future projects are on the basis of any of the following:—

(a) Sanctioned or approved utilisations and legal right demands or

(b) historical actual use

(c) demands if any, based on reassessment of requirements of the existing projects.

(ix) The operation policies (priorities etc.) for different uses considered in the simulation studies, shall be indicated with reasons.

NOTE:—In case detailed study is based on economic evaluation where the entire period of simulation is taken into account for working out the average annual benefits, the firm and secondary demands, priority of uses, sharing of shortages etc. considered shall be discussed with basis.

(x) In the case of multipurpose projects involving flood control storages, rule curve(s) and flood release rules adopted in the studies, shall be indicated with basis thereof.

NOTE :—The flood release rules shall be framed so that if incoming floods turn out to be the spillway design flood, it can be negotiated safely without endangering the structure.

(xi) In case of multi-reservoir system rules of sharing of deficit and priorities of releases between reservoirs both for conservation and flood control purposes shall be indicated with basis.

### 6.2 Project performance

The results of the simulation studies shall be tabulated and discussed.

Performance can be expressed as:

1. The number of failure years compared with the total number of years considered in the studies to meet the demand of a particular use irrespective of the quantum and duration of failure.
2. Number of failure years compared with the total years considered in the studies by neglecting the failure of quantum below a particular quantity of failure for short periods or both.
3. Number of crop seasons in which failure takes place compared with total crop seasons, as in (a) and (b).
4. The number of succesive years of failures (exceeding two) in the entire period of simulation—usewise.

NOTE: —In case the project evaluation is carried out on the basis of economic analysis with discounted rates the following performance analysis shall be furnished:

(a) Average and annual quantum of shortages for use over the period of simulation.

(b) Present value of cost of shortages, indicating loss functions, discount rates etc. with justification.

The performance is to be discussed usewise and for each alternative plan. Where alternative flow series are considered performance shall also be indicated for each series and by giving average, maximum and minimum values of performance indices.

# CHAPTER VII

## 7.0 EFFECT OF PROJECT ON HYDROLOGIC REGIME

The following aspects shall be discussed under this chapter:

### 7.1 Effect on low flows

Likely changes (quantitatively) in low flows in different reaches of the river due to the project.

### 7.2 Effect on peak flood

The reaches where the flood peaks are reduced or become sharper due to the project and their quantitative effects.

The likely changes and their effects on existing facilities etc. as also likely changes in river hydrographs both on short and long term basis.

### 7.3 Effect on total runoff

The likely decrease in the total runoff yield of the basin due to increased evaporation from the altered water surface and evaporation in the command area.

### 7.4 Effect on sediment flows

Likely changes (quantitatively) on sediment flows downstream of the project and its effects on downstream structures, land fertility etc.

Enclosure A

## E—AREAS AND REACHES OF INTEREST

E-1 Drainage basins upto control points i.e. sites of hydraulic structures, hydrometric sites, flood damage points, confluence with large rivers etc.

E-2 Potential irrigation area

E-3 Potential flood damage area

E-4 Potential drainage congestion area

E-5 Hydrometeorologic region surrounding the project basin. The region E-5 system will thus include all other regions and reaches E-1 to E-4 and E-7 to E-13 described here and in addition will include surrounding areas of similar hydrometeorologic characteristics.

E-6 River system reach within and slightly upstream of a reservoir

E-7 Potential ground water recharge area

E-8 Reservoir submergence area

E-9 River system reach from a hydraulic structure to a downstream point which is a control point (another structure of a natural hydraulic control-causing critical flood or a point sufficiently downstream for friction controlled channels, or a confluence with major river or sea.

E-10 River reach through the area of potential flood damage or potential drainage damage.

E-11 River reach in which industrial or domestic water supply is contemplated and where the quantity and quality of water is to be monitored.

E-12 River reach in which navigation is to be sustained by monitoring low flows.

E-13 River reach in which water quality salinity of low flows area to be monitored for fish and wild life sustenance and for recreation.

Enclosure B

## A—CLASSIFICATION BY STORAGE BEHIND THE STRUCTURES

A-1 Diversion projects without pondage

A-2 Diversion projects with pondage

A-3 Within the year storage projects

A-4 'Over the year' storage projects

A-5 Complex systems involving combinations of 1 to 4 above mentioned.

Enclosure C

## B—CLASSIFICATION BY USE OF PROJECT

B-1 Irrigation

B-2 Hydropower

B-3 Water supply and industrial use

B-4 Navigation

B-5 Salinity Control

B-6 Water Quality Control

B-7 Recreation, fish and wild life

B-8 Flood control

B-9 Drainage

B-10 Surface to ground water recharge

B-11 Multipurpose.

Enclosure D

## C—TYPES OF HYDROLOGIC INPUTS REQUIRED

C-1 For simulation studies

C-1.1 Water inflows

C-1.2 Lake evaporation

C-1.3 Potential evapotranspiration and rainfall

C-1.4 Sediment inflows

C-1.5 Flood inputs

C-1.6 Water Quality inputs

C-1.7 Low flow inputs

C-1.8 Surface to ground water recharge

C-2 For studies other than simulation

C-2.1 Design floods for the safety of structures

C-2.2 Design floods and flood levels for flood control works

C-2.3 Design floods for design of drainage works

C-2.4 Design floods for planning construction and diversion arrangements

C-2.5 Studies for determination of levels for locating structures on river banks or for location of outlets

C-2.6 Tail water rating curves

Enclosure E

## TYPES OF HYDROLOGICAL INPUTS REQUIRED FOR SIMULATION (CLASSIFIED AS PER CONTEMPLATED USE)

| Use | Inputs |
|---|---|
| B 1* | C 1.1<br>C 1.2 (if storage is involved)<br>C 1.3 |
| B 2* and B 3* | C 1.1<br>C 1.2 (if storage is involved) |
| B 4 * to B 7* | C 1.1<br>C 1.2 (if storage is involved)<br>C 1.7 |
| B 8* | C 1.5 |
| B 9* | C 1.5 |
| B 10* | C 1.1<br>C 1.8 |
| B 11* | All depending on individual uses |

*If the project involves large pondage/storage input C 1.4 will also be required. Sediment inflows normally do not form direct input in the simulation from one time unit to another. Only the long term loss of storage in a 'time horizon' is considered and the revised area-capacity curve at the end of this time horizon is predicted. This revised area-capacity curve is used through out the period of simulation without any consideration for year to year changes.

Enclosure—I

## TYPE OF HYDROLOGICAL INPUTS REQUIRED FOR STUDIES OTHER THAN SIMULATION (CLASSIFIED AS PER STORAGE TYPES AND USE)

| Storage | Use | Hydrologic inputs required |
|---|---|---|
| A 1 | All | C 2.1<br>C 2.4<br>C 2.5<br>C 2.6 |
| A 3, A 4 & A 5 | B 1 to B 7,<br>B 10 and<br>B 11 | -do- |
| A 3, A 4 & A 5 | B 8 | Same as above<br>Also C 2.2 |
| Any | B 9 | C 2.3 |

Enclosure G

## TIME UNITS REQUIRED FOR SIMULATION (CLASSIFIED AS PER STORAGE TYPE AND USE)

| Type of Storage | Type of Use | Time unit required for simulation studies (except for studies of sediment inflow and deposition) |
|---|---|---|
| A 1 | B 2 to B 7 & B 10 | Instantaneous discharges every day, or at smaller units. |
| A 2 | B 2 to B 7 | 1 day to 10 days depending on the extent of pondage. |
| A 2 | B 1 | 3 days for upland crops, 10 days for paddies. If extra pondage at headworks in addition to natural storage on field is provided, larger units can be used. |
| A 3/A 2 | B 8 | 1 hour to 24 hours depending on the damping provided by the drainage basin to the storage. |
| A 2 | B 10 | 1 day to 10 days depending on the pondage. |
| A2 | B 11 | Minimum of individual time units required by each type of use. If flood control is involved much shorter interval (1 hr. to 24 hrs.) operation is required only for critical flood periods. |
| A 3 | B 1 to B 3 | Monthly. However, it may be sufficient to divide the year in 4 to 8 blocks by grouping together periods of definite storage accumulation and storage depletion type, and the periods which cannot be classified as such being kept as separate blocks. |
| A 3 | B 4 to B 7 | Same as above, but during critical low flows, shorter time unit of about 10 days to 1 month may be required to simulate droughts and extra releases for control of water quality, salinity etc. |
| A 3 | B 10 | Same as A 3 —B 1 to B 3 discussed above, in dry season, but in rainy season where extra recharge will be affected by rainfall, 1 day to 10 day working will be necessary. |
| A 3 | B 11 | Minimum of individual time units for each type of use. However, shorter time units required for use B 4 to B 7 or for B 8 will apply only during critical low flows or floods. |
| A 4 | B 1 to B 3 | Bi-seasonal (i.e. year divided in two blocks say monsoon and non-monsoon) or shorter blocks. |
| A 4 | B 4 to B 7 B 10 or B 11 | Same as above, but during critical low flows, short time operation as indicated for storage type A 3 and corresponding type of use may be adopted. |
| A 5 | All uses | Adopt the minimum of the time unit required by each of the component storages involved in the compled system, after considering the type of use through that storage. |
| | | However, structures or uses of minor importance in the overall system may not dictate the choice of time used to be adopted in total simulation of the system. |

***DIRECTOR : S M SETH***

***DIVISIONAL HEAD : R D SINGH***

***STUDY GROUP***
***R MEHROTRA, SCIENTIST***