CHAPTER 1
INTRODUCTION

1.0 GENERAL

By definition, *an ungauged basin* (or catchment) is one with inadequate records, in terms of both data quantity and quality, of hydrological observations to enable computation of hydrological variables of interest, at the appropriate spatial and temporal scales, and to the accuracy acceptable for practical applications. Hydrological variables here may refer to evaporation, infiltration, rainfall, runoff, and sub-surface flow. However, many processes that are of interest from hydrologic point of view are difficult to observe routinely and unambiguously. Streamflow measurement is one such variable that can be measured at a gauging site of a basin with some confidence. So from practical point of view, the definition for an ungauged basin is usually limited to the basins with inadequate streamflow measuring facilities at the project site and also with scanty or no streamflow records. In true sense, almost every basin is an ungauged basin as per this definition. However, from a broader perspective, major river basins or catchments in India have been gauged for determination of hydrological variables but medium and small sized catchments are mostly ungauged. Still, there are several major catchments in different parts of the country that have not yet been gauged for measurement of runoff, sediment etc. and in some catchments the existing gauging networks are being discontinued (Kothyari, 2004). This is due to economic constraints that do not justify detailed hydrological and meteorological investigations at every new site on large scale and on long term basis for a large country like India. Thus, there is a need for methods that can be utilized for realistic estimation of such hydrological variables for ungauged catchments. Chapter on IAHS decade on 'Predictions on Ungauged Basins (PUB)' addresses scientific problems that requires major breakthrough in its theoretical foundations and critically examine the existing approaches to hydrological predictions (Sivapalan et al. 2004).

However, a major dilemma facing hydrologists is how best to predict hydrologic variables in ungauged basins where little or no information is available. The primary cause of difficulty with such predictions is the high degree of heterogeneity of the land surface condition, soils, vegetation, land use, flood producing mechanisms e.g. snow fed or rain fed, and space-time variability of the climate inputs. Though a lot of advances have been made in developing methods for delineating homogeneous and heterogeneous regions on basis of specific hydrologic variables, but there remain numerous basic problems that still need to be resolved before these methods can be used universally. One of the drawbacks is that the existing method for estimating the degree of heterogeneity requires sufficient field data at regional scales, which is generally not available. Also, in cases where sufficient data is available, representation of hydrologic processes resulting due from the heterogeneities of landscape properties and climatic change is difficult (Nagesh and Maity, 2004). Given the above discussed limitations, the present practice of predictions in ungauged catchments in regard to streamflow variable, which is the interest of this study, is discussed in the section that follows.

MAHANADI RIVER

The river traverses Cuttack district in an east-west direction. Just before entering Cuttack city, it gives off a large distributary called the Kathjori. The city of Cuttack stands on the spit separating the two channels. The Kathjori then throws off many streams like the Kuakhai, Devi and Surua which fall into the Bay of Bengal after entering Puri district. The Kathjori itself falls into the sea as the Jotdar. Other distributaries of Mahanadi include the Paika, Birupa, Chitroptala iver, Genguti and Nun. The Birupa then goes on to join the Brahmani River at Krushnanagar and enters the Bay of Bengal at Dhamra. The Mahanadi proper enters the sea via several channels near Paradeep at False Point, Jagatsinghpur. The combined Delta of the Mahanadi's numerous distributaries and the Brahmani is one of the largest in India.

An average annual surface water potential of 66.9 km³ has been assessed in this basin. Out of this, 50.0 km³ is utilisable water. Culturable area in the basin is about 80,000 km², which is 4% of the total culturable area of the country .Present use of surface water in the basin is 17.0 km³. Live storage capacity in the basin has increased significantly since independence. From just about 0.8 km³ in the pre-plan period, the total live storage capacity of the completed projects has increased to 8.5 km³. In addition, a substantial storage quantity of over 5.4 km³ would be created on completion of projects under construction. Additional storage to the tune of over 11.0 km³ would become available on execution of projects under consideration. The hydropower potential of the basin has been assessed as 627 MW at 60% load factor.

At its peak during the monsoon, the Mahanadi has a discharge rate of 2 million cubic ft. per second, almost as much as the much larger *Ganges*. However owing to its seasonal nature the river is mostly a narrow channel flanked by wide sand banks for most of the year. The Mahanadi was notorious for its devastating floods for much of recorded history. However the construction of the *Hirakud Dam* has greatly altered the situation. Today a network of canals, barrages and check dams keep the river well in control. However heavy rain can still cause large scale flooding as evidenced in September, 2008 when 16 people died as the river breached its banks.

In 2011, September heavy down pour of rain caused flash flood and many mud dwellings in more than 25 villages above Hirakud Dam which were never got affected before; both in Chhattisgarh and Orissa have melted down due to back water, which could not pass through the river.

1.1 STATUS

For many small catchments, the stream flow data are limited, and for ungauged basins it is not available. Under such circumstances, regional unit hydrograph and regional food frequency methods along with regional empirical formula developed using the data of gauged catchments in the region are used to arrive at design flood for the ungauged catchment. Regional unit hydrographs for such regions are derived using their physical, hydrological and storm characteristics. Research Designs and Standard Organization (RDSO) and Central Water Commission (CWC) documented a report in 1980 recommending regional hydrograph parameters and SUH relationships. This covers a part of the study area (proposed for this project) i.e. Mahanadi sub-zone (3d). In addition to this, the CWC has recommended criteria to be adopted for estimation of design flood for waterways of rail and road bridges in North Bengal, which doesn't cover the present study area.

In India regional flood frequency studies have been carried out using conventional methods such as USGS method, regression based methods. For some typical regions attempts have been made to study application of the new approaches in the studies conducted at some of the Indian research institutions and academic organizations. National Institute of Hydrology, Roorkee and IIT, Roorkee has brought out a state of art report on 'Regionalization of Hydrological parameters' that covers the status of the rational and empirical methods practices in India for regional analysis of flood. From the report it is inferred that no systematic study of this nature has been conducted in the proposed study area. Kumar et al., (1997) carried out a detailed regional flood frequency studies for zone-3, where a part of the present study area (3d) was included. Similarly, no systematic study has so far been attempted for development regional flow duration curves in the proposed study area.

Under the present practices, synthetic unit hydrographs, regional unit hydrograph, regional food frequency methods along with regional empirical formula developed using the data of gauged catchments in the region are used to arrive at design flood for the ungauged catchment. Synthetic unit hydrographs (SUH) are derived using the well established empirical formula which relates the hydrograph parameters with the physiographic characteristics of a catchment. These parameters are than used to derive the complete shape of the unit hydrograph through some trial procedures; so in a way this involves some subjectivity. Regional unit hydrographs for a region are derived using their physical, hydrological and storm characteristics and in the process both SUH approach and physical based models are used. Similarly, in regional food frequency methods the parameters of a probability distribution function that is valid for a hydrological homogeneous are regionalized to be used for an ungauged catchment in the region. In addition to the above mentioned methods, sometime simple empirical formulae are derived for a region where the flow characteristics are related to the physical properties of a catchment. These formulae are used to compute certain flow characteristics (like peak discharge, time to peak of a UH) of a ungauged catchment. Research Designs and Standard Organization (RDSO) and Central Water Commission (CWC) documented a report in 1980 (Goel and Subhash Chander, 2002; CWC, 1980, 1982) recommending regional hydrograph parameters and SUH relationships. This covers a part of the study area (proposed for this project) i.e. Mahanadi sub-zone (3d). In addition to this, the CWC has recommended criteria to be adopted for estimation of design flood for waterways of rail and road bridges in North Bengal, which doesn't cover the present study area. Some of the popular methods adopted in recent times for design flood estimation in ungauged catchments are describes in the following section.

1.1.1 FLOW DURATION CURVE (FDC)

The flow-duration curve provides a convenient means for studying the flow characteristics of streams and for comparing or°, basin with another. Various uses of the flow-duration curve A flow-duration curve (FDC) represents the relationship between the magnitude and frequency of daily, weekly, monthly (or some other time interval of) streamflow for a particular river basin, providing an estimate of the percentage of time a given streamflow was equaled or exceeded over a historical period. An FDC provides a simple, yet comprehensive, graphicalview of the overall historical variability associated with streamflow in a river basin.

An FDC is the complement of the cumulative distribution function (cdf) of daily streamflow. Each value of discharge Q has a corresponding exceedance probability p, and an FDC is simply a plot of Qp, the pth quantile or percentile of daily stream flow versus exceedance probability p, where p is defined by

$$p = 1 - P\{Q :s; q\} \quad \text{(la)}$$

$$p = 1 - FQ(q) \quad \text{(lb)}$$

The quantile Qp is a function of the observed streamflows, and since this function depends upon empirical observations, it is often termed the empirical quantile function. Statisticians term the complement of the cdf the "survival" distribution function. The 'term survival results from the fact that most applications involve survival data that arise in various fields.

1.1.2 Regional Flow Duration Curve (RFDC)

Regional estimation of flow duration characteristics is important for water resources development at the small catchment scale. Regional analyses often fail to adequately represent

the variability of the flow regime in small catchments (<50 km^2), especially in remote mountainous regions where the calibration data are sparse and refer to much larger catchment scales. This study suggests an approach in which regional data are combined with actual instantaneous flow data to construct a representative daily flow duration curve for small catchments. A regional dimensionless flow duration curve (FDC) is developed for a hydrologically homogeneous region in Western–Northwestern Greece and used to estimate the FDC in two small mountainous catchments within the region. A number of instantaneous flow measurements available at the two sites are used in a statistical representation of the flow regime from which an estimate of the mean annual flow at the sites is made, allowing the construction of the FDC from the regional curve. Results obtained are in good agreement with observed data and show significant estimation improvement over other methods commonly employed in the study region. A sensitivity analysis using Monte Carlo simulation is performed to establish reasonable sampling requirements for small ungauged catchments in the study area and similar mountainous areas in the Mediterranean region.

1.1.3 Hydrograph

A hydrograph is a graph showing the rate of flow (discharge) versus time past a specific point in a river, or other channel or conduit carrying flow. The rate of flow is typically expressed in cubic meters or cubic feet per second (cumecs or cfs). It can also refer to a graph showing the volume of water reaching a particular outfall, or location in a sewerage network, graphs are commonly used in the design of sewerage, more specifically, the design of surface water sewerage systems and combined sewers.

In surface water hydrology, a hydrograph is a time record of the discharge of a stream, river or watershed outlet. Rainfall and/or snowmelt is typically the main driver of watershed

discharge; a hydrograph records how a watershed responds to these and other drivers of watershed hydrology.

A watershed's response to rainfall depends on a variety of factors which affect the shape of a hydrograph:

a) Watershed topography and dimensions

b) Soil permeability and thickness

c) Geology (e.g., bedrock permeability and transmissivity)

d) The area of a basin receiving rainfall

e) Land use and land cover (e.g., vegetation types and density, impervious areas)

f) Drainage density and other river geometrics

g) Duration and intensity of precipitation, and form of precipitation (rain vs. snow)

h) Temperature, wind speed, time of year, and other effects on precipitation type and rates of evapo-transpiration

i) Initial conditions (e.g., the degree of saturation of the soil and aquifers)

j) Storage opportunities in the drainage network (e.g., lakes, reservoirs, wetlands, channel and bank storage capacity). For example Fig 1a

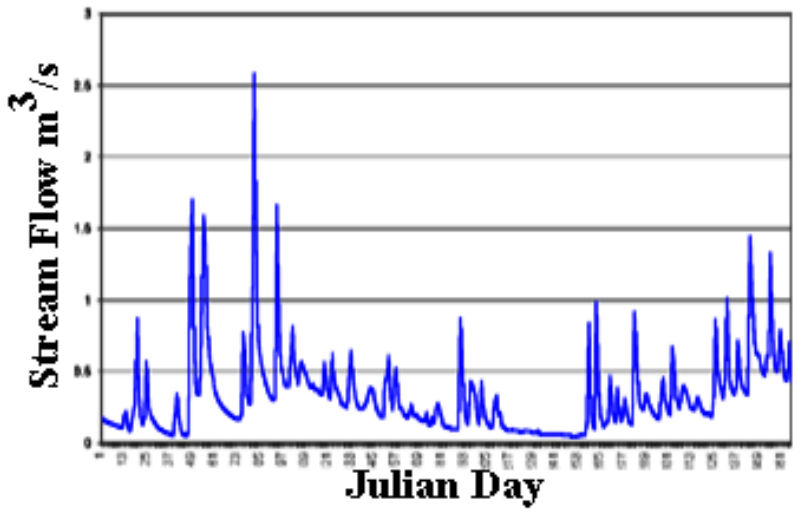

Figure 1a Stream hydrograph. Increases in stream flow follow rainfall or snowmelt.

1.1.4 Unit Hydrograph

A *unit hydrograph* (UH) is the hypothetical unit response of a watershed (in terms of runoff volume and timing) to a unit input of rainfall. It can be defined as the *direct runoff hydrograph* (DRH) resulting from one unit (e.g., one cm or one inch) of *effective rainfall* occurring uniformly over that watershed at a uniform rate over a unit period of time.

A unit hydrograph (UH) is a hydrograph of direct surface runoff resulting at a given location on a stream from a unit rainfall excess amount occurring in unit time uniformly over the catchment area up to that location. The excess rainfall (ER) excludes losses i.e. hydrological abstractions, infiltration losses from total rainfall and unit rainfall excess volume equals one mm may be considered as a standard value since rainfall is measured in mm. However, in certain cases UH is derived for one cm ER. The selection of unit time depends on the duration of storm and size of the catchment area. For example, for small catchments, periods of 1 or 2 hours can be assumed and for larger catchments, 3, 4, 6, or even 12 hours can be adopted. CWC has recommended the following for practical uses:

A unit hydrograph can be interpreted as a multiplier that converts rainfall excess to direct surface runoff. The direct surface runoff (DSRO) is the streamflow hydrograph excluding base-flow contribution. Since, a unit hydrograph depicts the time distribution of flows, its multiplying effect varies with time. In real-world application, the unit hydrograph is applied to each block of rainfall excess and the resulting hydrographs from each block are added for computing direct surface runoff hydrographs, to which base-flows are further added to obtain total flood hydrographs.

1.1.5 Synthetic Unit Hydrographs

Synthetic unit hydrographs (SUH) are of great significance in determination of flood peak and runoff volume, especially from ungauged watersheds. Singh (1988) provided a good review of several methods dealing with the SUH derivation. The qualifier 'synthetic' denotes that the unit hydrographs (UH) are obtained without using watershed's rainfall-runoff data. These synthetic or artificial unit hydrographs can be characterized by their simplicity and ease in construction. These require fewer amounts of data and yield a smooth and single valued shape corresponding to unit runoff volume, which is essential for UH derivation. In practice, an SUH is derived from few salient points of the unit hydrograph by fitting a smooth curve manually. The methods of Snyder (1938) and Espey (1974) are a few examples among others. These utilize empirical equations to estimate salient points of the hydrograph, such as peak flow (Q_p), lag time (t_L), time base (t_B), UH widths at $0.5Q_p$ and $0.75Q_p$.

1.1.6 Regional Unit Hydrograph

Regional unit hydrographs are used for a catchment where sufficient flood data is not available to derive a UH. In such cases the physical and storm characteristics are used to derive the UH. The procedure used for this purpose involve the derivation of the parameters that describes the unit hydrograph for the gauged catchment and then the development of regional relationship between the unit hydrograph parameters with the pertinent physiographic and storm characteristics of the catchment. The catchments considered for such regional study has to be within a hydrologically homogeneous region. Regional parameters for the UH are developed using the information of catchments response characteristics like soil, geology and land-use based on the data available for the gauged catchments in nearby hydro- meteorologically similar regions for modeling their hydrological response. Regionalization of the parameters is, however, a very tedious task to accomplish since the hydrological behavior of many nearby catchments

have to be ascertained before being confident about the values of the parameters. Generally the sequential steps that is followed to develop a regional unit hydrograph are : selection of a catchment, split sample test for the region, derivation of UH for the gauged catchments in the region, derivation of representative UH for the region, split sample test for the regional UH, development of regional UH relationships. Both physical processes based approach such as O'Kelly's approach (1955), Nash's approach (1960) , and Clark's approach (1945) and empirical or regression approach such as Snyder's approach (1938), Gray's approach (1961), SCS method (1958), and Espey's approach (1974) are used to develop regional UH relationships.

1.1. 7 Flood Frequency Analysis

Flood frequency analyses are used to predict design floods for sites along a river. The technique involves using observed annual peak flow discharge data to calculate statistical information such as mean values, standard deviations, skewness, and recurrence intervals. These statistical data are then used to construct frequency distributions, which are graphs and tables that tell the likelihood of various discharges as a function of recurrence interval or exceedence probability. In many problems in hydrology, the data consists of measurements on a single random variable; hence we must deal with univariate analysis and estimation. The objective of univariate analysis is to analyze measurements on the random variable, which is called sample information, and identify the statistical population from which we can reasonably expect the sample measurements to have come. After the underlying population has been identified, one can make probabilistic statements about the future occurrences of the random variable, this represents univariate estimation. It is important to remember that univariate estimation is based on the assumed population and not the sample, the sample is used only to identify the population.

Hydrologic processes such as rainfall, snowfall, floods, droughts etc. are usually investigated by analyzing their records of observations. Many characteristics of these processes may not represent definite relationship. For example, if you plot instantaneous peak discharges from each year for a river, a rather erratic graph is obtained. The variation of peak discharge from one year to another can not be explained by fitting a definite relationship, which we call as deterministic relationship. For the purpose of hydrologic analysis, the annual peak discharge is then considered to be a random variable. Methods of probability and statistics are employed for analysis of random variables. In this chapter, some elementary probability distributions are presented, which are used for frequency analysis in hydrology.

The purpose of the frequency analysis is to estimate the design flood for desired recurrence interval assuming the sample data follow a theoretical frequency distribution. It is assumed that the sample data is a true representative of the population. It is generally seen that minimum 30 to 40 years of records are needed in order to carry out flood frequency analysis to the at site data for estimating the floods in extrapolation range, somewhat, within the desired accuracy. In case the length of records is too short, it represents inadequate data situation and at site flood frequency analysis fails to provide the reliable and consistent flood estimates. The regional flood frequency curves together with at site mean is generally able to provide more reliable and consistent estimates of floods under the inadequate data situation. For ungauged catchment, the regional flood frequency analysis approach is the only way to estimate the flood, for desired recurrence interval for which a regional relationship between mean annual peak flood and catchment characteristics is developed along with the regional frequency curves.

The estimation of flood frequency curves in ungauged basins is an important field of application of advanced research concerning the knowledge on physical processes and the

development of statistical tools. Regional statistical analysis and physically-consistent derivation of probability density functions (pdf) are the key fields to rely on for providing robustness to estimation of flood quantiles, and for transferring hydrological information between basins. In this context, physically consistent reasoning applied to the regional statistical analysis can be closely connected with some basis of a geomorphoclimatic approach for derivation of the flood pdf

1.1. 8 Regional Flood Frequency Approach

The methods that have been outlined in section 1.2.1 and 1.2.2 uses short-term rainfall-runoff data of gauged catchments to derive a UH, and subsequently these are used to develop a regional UH. Finally, the design flood is estimated using the computed UH with the corresponding probable maximum precipitation or storm characteristics of that catchment. An alternate approach for the estimation of design flood is by statistical flood frequency analysis using the annual maximum flow data series or annual maximum series (AMS) of a catchment. In this approach, at site data together with regional data is utilized to provide most consistent and reliable flood estimates for the gauged sites with limited data records. And for ungauged sites with no record, only regional data is used for flood frequency analysis. As such there are essentially two types of models generally used in flood frequency analysis literature: (i) annual flood series (AFS) or AMS models and (ii) partial duration series (PDS) models. Regional flood frequency method explicitly incorporates a homogeneity test in the process of selecting the collection of stations that comprise the region for an ungauged site. In India regional flood frequency studies have been carried out using conventional methods such as USGS method (or index flood method), and regression based methods. Very few attempts have been made to study application of the new approaches in the studies.

1.1.9 Partial Duration Series (PDS) or Peak Over Threshold (POT) Method

The AMS model considers the annual maximum flood of each year in a series that has as many elements as there are years in the data record, notwithstanding that the secondary events in one year may exceed the annual maxima of other years. In some regions, annual maximum floods observed in dry years may be very small, and inclusion of these events can significantly alter the outcome of the extreme value analysis. In contrast, the PDS model avoids all these drawbacks by considering all flood peaks above a certain threshold level; hence it is also referred to as peak over threshold (POT) method. Clearly in the PDS model, more than one flood per year may be included, and exceptionally low annual maximum values that may cause problems in parameter estimation are not included. In AMS model the observed flood frequency of the annual number of exceedances are fitted with a single probability distribution function (*pdf*) which is eventually used to estimate recurrence flood at higher return periods. Whereas in PDS models, two distribution functions are used; one to count the number of times the flood peaks are above a pre defined threshold, and the other *pdf* is used to fit the flood exceedances. In the limit, when the threshold is increased, recurrence interval of large events computed using AMS and PDS models tend to converge. The PDS models are often used for flood frequency analysis (FFA) when there is a paucity of data (Madsen et al*.,* 1994; 1995; 1997; Birikunadavyi and Rousselle, 1997; Lang, 1999; Onoz and Bayazit, 2001).

1.2 Use of Remote Sensing Techniques for Information

Remote sensing can be defined as any process whereby information is gathered about an object, area or phenomenon without being in contact with it. Thus, a remote sensing device records response which is based on many characteristics of the land surface, including natural and artificial cover. An interpreter uses the element of tone, texture, pattern, shape, size, shadow, site and association to derive information about land cover. The generation of remotely sensed

data/images by various types of sensor flown aboard different platforms at varying heights above the terrain and at different times of the day and the year does not lead to a simple classification system. It is often believed that no single classification could be used with all types of imagery and all scales. To date, the most successful attempt in developing a general purpose classification. The basis of using remote sensing data for change detection is that changes in land cover result in changes in radiance values which can be remotely sensed. Techniques to perform change detection with satellite imagery have become numerous as a result of increasing versatility in manipulating digital data and increasing computer power. The sun synchronous satellites have a synoptic coverage. The ground area covered by the satellite's passes can be obtained by referring to its path and row. *Sensors* are devices used for making observations to operate and produce outputs, which are either representative of the observed area

It has been noted over time through series of studies that Landsat Thematic Mapper is adequate for general extensive synoptic coverage of large areas. As a result, this reduces the need for expensive and time consuming ground surveys conducted for validation of data. Generally, satellite imagery is able to provide more frequent data collection on a regular basis unlike aerial photographs which although may provide more geometrically accurate maps, is limited in respect to its extent of coverage and expensive; which means, it is not often used. The procedure adopted in this project work forms the basis for deriving statistics of land use dynamics and subsequently in the overall, the findings. Remote sensing technology in recent years has proved to be of great importance in acquiring data for effective resources management and hence could also be applied to coastal environment monitoring and management (Ramachandran, 1993, Ramachandran et.al., 1997, 1998). Further, the application of GIS (Geographical Information System) in analyzing the trends and estimating the changes that have occurred in different themes helps in management decision making process.

For the study, satellite images of Mahanadi sub-zones in Orissa and Madhya Pradesh states were acquired for three Epochs; 2001 and 2008. All these were obtained from National Remote Sensing Centre, Department of Space under Government of India. The images has the following format Satellite IRS-1C and Sensor LISS-3, Satellite: IRS-P6 and Sensor: LISS-3. These were used with the corresponding area toposheets some acquired from Survey of India, Dehradun and few i.e. restricted ones were used from the RS Lab, NIH Roorkee. All the details are given in the following chapters.

1.2.1 Application of Remote Sensing Technique in Hydrology

I. Soil moisture estimation: Soil moisture is an integral quantity in hydrology that represents the average conditions in a finite volume of soil. In this paper, a novel regression technique called Support Vector Machine (SVM) is presented and applied to soil moisture estimation using remote sensing data.

II. Drainage basin mapping and watershed modeling: A drainage basin is an extent or an area of land where surface water from rain and melting snow or ice converges to a single point, usually the exit of the basin, where the waters join another water body, such as a river, lake, reservoir, estuary, wetland, sea, or ocean. Other terms that are used to describe a drainage basin are catchment, catchment area, catchment basin, drainage area, river basin and water basin

III. Measuring snow thickness: Snow is a form of precipitation, but, in hydrology it is treated somewhat differently because of the lag between when it falls and when it produces runoff, groundwater recharge, and is involved in other hydrologic processes. The hydrologic interest in snow is mostly in mid- to higher latitudes and in mountainous areas where a seasonal accumulation of a snowpack is followed by an often lengthy melt period that sometimes lasts

months. During the accumulation period there is usually little or no snowmelt. Precipitation falling as snow (and sometimes rain) is temporarily stored in the snowpack until the melt season begins. The hydrologist generally wants to know how much water is stored in a basin in the form of snow. The hydrologist will also be concerned with the areal distribution of the snow, its condition and the presence of liquid water in it

IV. River /delta change detection: Information about change is necessary for updating land cover maps and the management of natural resources. The information may be obtained by visiting sites on the ground and/ or extracting it from remotely sensed data. For many of the physical and cultural features on the landscape there are optimal time periods during which these features may best be observed. Remotely sensed data acquired at the fixed time interval becomes an important factor. Many researches have been undertaken to develop methods of obtaining change information. Change detected from different temporal images usually reflect natural and human activity impact each other and then can be used to study how to form the regional geographic feature.

V. River and lake ice monitoring: Presence and thickness of lake ice has a significant impact on fish ecology, particularly in terms of over-wintering habitats for certain species. The Lake Ice Service team of Hatfield Consultants Ltd. will be focusing its efforts on studying the dynamics of the ice in lakes and rivers of two study sites in Ungava Bay, Northern Quebec. The Payne River on the west and George River on the eastern side of the bay are two of the main habitats for Arctic Char (Salvelinus alpinus) in the region, providing nutrition and livelihoods to local Inuit communities.

VI. Flood mapping and monitoring: A constant watch was kept on the flood situation in the country through hydrological and meteorological information from various sources. Major flood/cyclones are mapped and monitored with the satellite data from optical and microwave satellites. Flood maps showing Flood Inundation information at state, district and detailed levels are generated. Flood damage statistics like district-wise flood inundated area, submerged roads & railways and submerged crop area are also estimated.

VII. Irrigation canal leakage detection: Remote sensing has shown promise as a tool for quick and cost-effective detection of leaks in aqueducts (Nells,1982; Pickerill and Malthus, 1998). One limitation is the resolution which is required (within a few meters at most) in order to detect canal leakage. Earth-observing satellite systems are limited in the spatial resolution.

VIII. Determining snow-water equivalent: Snow Water Equivalent (SWE) is a common snowpack measurement. It is the amount of water contained within the snowpack. It can be thought of as the depth of water that would theoretically result if you melted the entire snowpack instaneously. To determine the depth of snow using snow water equivalent and density, use the following formula: [SWE] ÷ [Density] = Snow Depth (Density must be in decimal form. For example: 25% = 0.25)

IX. Snow pack monitoring / delineation of extent: Areas of the Canadian high plains, the Montana and North Dakota high plains, and the steppes of central Russia have been studied in an effort to determine the utility of space borne microwave radiometers for monitoring snow depths in different geographic areas. Significant regression relationships between snow depth and microwave brightness temperatures were developed for each of these homogeneous areas. In each of the study areas investigated in this paper, Nimbus-6 (0.81 cm) ESMR data produced

higher correlations than Nimbus-5 (1.55 cm) ESMR data in relating microwave brightness temperature to snow depth. It is difficult to extrapolate relationships between microwave brightness temperature and snow depth from one area to another because different geographic areas are likely to have different snowpack conditions.

1.3 PARAMETER ESTIMATION USING OPTIMIZATION METHODS

The procedure proposed here for estimating the parameters of models used for this study is a combination of regression and an iterative procedure. During the last two decades, the tools that engineers and scientists work with have improved significantly. The rapid growth of computing power has enabled the researchers to develop effective modeling tools. An improved parameter estimation procedure has been developed by using optimization techniques and applied to estimate the parameters of the probability distribution function (pdf) and the cumulative distribution function (cdf). As a result, an improved estimation method was found. These approaches are generally used for calibrating fold frequency analysis for calculating return period flood and sometimes in ungauged and partially gauged basins when we use peak over threshold methods. Though distributions with more parameters can better fit data sets of different patterns, the associated parameter estimates often have large standard errors which result in wide confidence intervals for the quantile estimates. Furthermore, such distribution function may be difficult to express in a close inverse form of the cumulative distribution function, giving rise to problem in parameter estimation by maximum likelihood (ML) and probability weighted moment (PWM) methods. The analytical expressions of variance of the *T*-year event estimates using method of moments (MOM), method of maximum likelihood, and probability weighted moment estimation methods are checked using Monte Carlo experiments and optimization techniques by minimizing both the relative root average square error.

In case of unit hydrograph derivation for flood computing analysis, most of the available methods for synthetic unit hydrograph (SUH) derivation involve manual, subjective fitting of a hydrograph through few data points. Because of cumbersome trials involved, the generated unit hydrograph is often left unadjusted for unit runoff volume. Two approaches are generally followed for deriving a UH from recorded flood hydrograph and simultaneous rainfall records. The first one is a non-parametric approach based on a discretization technique, i.e., determination of a model at a finite number of discrete points. This includes the least square method (Snyder, 1955), matrix inversion (Eagleson et al., 1966), linear programming (Collins, 1939), non-linear programming (Mays and Taur, 1982), and recently Yang and Han (2006) used a transfer function method for this purpose. The second one is a parametric approach that fits some prescriptive functional curves with limited number of parameters, and these parameters are estimated by means of optimization using an objective function. For instance, Nash (1959) derived an IUH based on the concept of *n*-linear reservoirs of equal storage coefficient, and showed that the shape of IUH can be represented by a two-parameter Gamma distribution (2PGD). This study uses three methods viz, non-linear optimization, Neuro-Fuzzy and Neural networks for estimating the parameters.

1.4 GAPS

Several methods for synthetic unit hydrographs are available in the literature. Some of them like Snyder (1938), Gray (1961), and SCS (1957) involve manual, subjective fitting of a hydrograph through few data points. Because it is difficult, the generated unit hydrograph is often left unadjusted for unit runoff volume. Some probability distribution functions (*pdf*) has the flexibility in shape (skew on both sides) similar to an UH, and area under its curve is unity which are pre requisite for any UH. In spite of these properties, very few studies are found in literature that have focused and explored the potentiality of probability distribution functions

(*pdf*) to develop a synthetic unit hydrograph. In some cases theses might yield more SUH more conveniently and accurately than the popular methods.

The procedure of developing a regional unit hydrograph is similar to that adopted for a SUH, the only difference being the former is a used universally and the later is for a region; hence a regional unit hydrograph can be used for an ungauged catchment in a hydrologically homogeneous region for which it is developed. Different approaches are followed for regional UH development such as Snyder's approach (1938), Gray's approach (1961), SCS method (1958), Espey's approach (1974), O'Kelly's approach (1955), Nash's approach (1960) , and Clark's approach (1945). The first four approaches are purely synthetic methods where the flow data of all the gauging sites in the region are analyzed in order to evolve suitable relationship between UH parameters and catchment characteristics. Though the relationships for these synthetic methods were initially developed for a specific region (Viessman et al., 1989), they are universally used in spite of the inbuilt errors they carry in predicting the overall shape of the UH. In methods given by O'Kelly, Nash, and Clark, the parameters of the corresponding conceptual models are determined for gauged basins in the region and subsequently relationships established between parameters and catchment characteristics. In applying these methods, it is always necessary to validate these relationships for a new time period and for additional gauged catchments which are not considered for their development. Keeping these constraints in view, there is need for improving the accuracy of the results by including latest conceptual models for SUH derivation and also to simplify the parameter estimation technique of these models for a region, which will permit their unique definition from basin characteristics and eliminates the need for model calibration with streamflow records.

Regional flood frequency is advantageous when the sites forming a region have similar frequency distribution. In practice, however, available sample size is usually not large enough to unequivocally identify the frequency distribution. Therefore, most regional flood frequency analysis procedures fit to the data a distribution whose form is specified apart from a finite number of undetermined parameters. Over the years, researchers world wide have suggested the use of regional regression to improve the estimation of normalized flood quantiles used in the IF method. Zrinji and Burn (1994) and Burn et al. (1997) presented a framework for regional flood frequency analysis using region of influence (ROI) that is applicable for estimating extreme flow quantiles at ungauged catchments. However, in heterogeneous regions, when a particular site has substantially different characteristics from the rest of the region or in circumstances where the record length of data is short at gauge of individual sites, ROI is difficult to apply. Further, in areas where there are few gauged sites, it may be difficult to identify enough homogeneous sites to apply regional flood frequency methods.

Partial duration series (PDS) or Peak over threshold (POT) models generally employ a Poisson distribution (PD) to count for the number of annual flood exceedances above threshold, and a generalized pareto distribution (GPD) which has the exponential distribution (ED) as a special case is used to fit the flood exceedances (e.g., Davison and Smith, 1990; *R*osbjerg et al., 1992; Madsen et al*.,* 1994; 1995; 1997). However, a criticism against use of PD for estimating the number of floods exceeding a threshold is that the number of peaks occurring each year was not a Poisson variate since its variance was significantly greater than its mean which violates the basic assumption of PD (NERC, 1975; Cunnane, 1979). Onoz and Bayazit (2001) observed that in certain cases the variance of observed number of exceedances is significantly different, usually larger than the mean. In most cases the annual number of occurrences of peaks above a threshold has a variance larger than its mean. Similarly the advantage of ED over other

distributions in fitting the flood exceedances is that it has a single parameter that is to be estimated, and in some cases, it gives a more precise prediction of the *T*-year return period event (Rosbjerg et al., 1992) ,however the demerit is its inability to accommodate its shape to provide a good fit across many data sets due to the single parameter and its lack of memory property that makes the distribution insensitive in the upper tail (Montgomery and Runger, 1994). Though distributions with more parameters can better fit to the data set of any shape, the associated parameter estimates often have large standard errors which result in wide confidence intervals for the quantile estimates. Furthermore, the distribution function is difficult to estimate in a close (inverse) form giving rise to problem in parameter estimation.

1.5 OBJECTIVES

1. To calibrate and validate an event based model based on unit hydrograph approach analyzing the available data of flood events for the gauged catchments in the region.
2. To identify few robust flood frequency distributions that may be used for the computation of return period flood for the gauged catchments in the region.
3. To develop regional flood formulae using statistical correlation of the observed peak characteristics with important catchment and storm characteristics, for the estimation of the peak, and time to peak for the ungauged catchments in the region.
4. To develop regional unit hydrograph, and regional flood frequency analysis procedures utilizing the available data and methodologies.
5. To develop methodology for the regionalization of the hydrological parameters for the computation of the water availability for the ungauged catchments in the region.

CHAPTER 2
STUDY AREA AND STREAM MORPHOLOGY

2.0 MAHANADI RIVER SYSTEM

Mahanadi is a major sedimentary basin located along the east coast of India. Mahanadi sub zone 3(d) (CWC, 1997) is located between22 80^0 25'N - 87^0 E longitude 19^0 15'- 230 35'N latitude. It extends over an area of 141,589 km which is nearly 4.3% of total geographical area of the country. Mahanadi River rises from the district of Madhya Pradesh and flows about 851 km before it joins the Bay of Bengal (Fig 2.1a). The Mahanadi River flows slowly for 560 miles (900 km) and has an estimated drainage area of 51,000 square miles (132,100 square km). The area of the drainage basin is 141,464 km². The interior coastal plain has a relatively low elevation. The average elevation of the drainage basin is 426 m, with a maximum of 877 m and a minimum of 193 m. The average elevation of the drainage basin is 426 m, with a maximum of 877 m and a minimum of 193 m. The bulk of the precipitation is in the July to September period [800 to over 1200 mm] with January to February precipitation of less than 50 mm. The average discharge 2,119 m^3/s (74,832 cu ft/s) - max 56,700 m^3/s (2,002,342 cu ft/s). It deposits more silt than almost any other river in the Indian subcontinent.

The Mahanadi river basin lies north-east of Deccan plateau. Mahanadi basin is bounded on the North by the Central India hills, on the South and east by the Eastern Ghats and on the west by the Maikala range. The upper basin is a saucer shaped depression known as *Chhattisgarh*. The basin is circular in shape with a diameter of about 400 km and an exit passage of about 160 km long and 60 km wide. There are four well defined physical regions in the basin viz.: (i) The Northern plateau, (ii) The Eastern Ghats, (iii) The Coastal plains and (iv) The Erosional plains of the Central table Land. The Northern plateau and Eastern Ghats are well forested hilly regions. The coastal plains stretching over the districts of Cuttack and Puri cover

the large delta formed by the river Mahanadi and is a fertile area well suited for intensive cultivation. The erosional plains of the Central Table Land are traversed by the river Mahanadi and its tributaries.

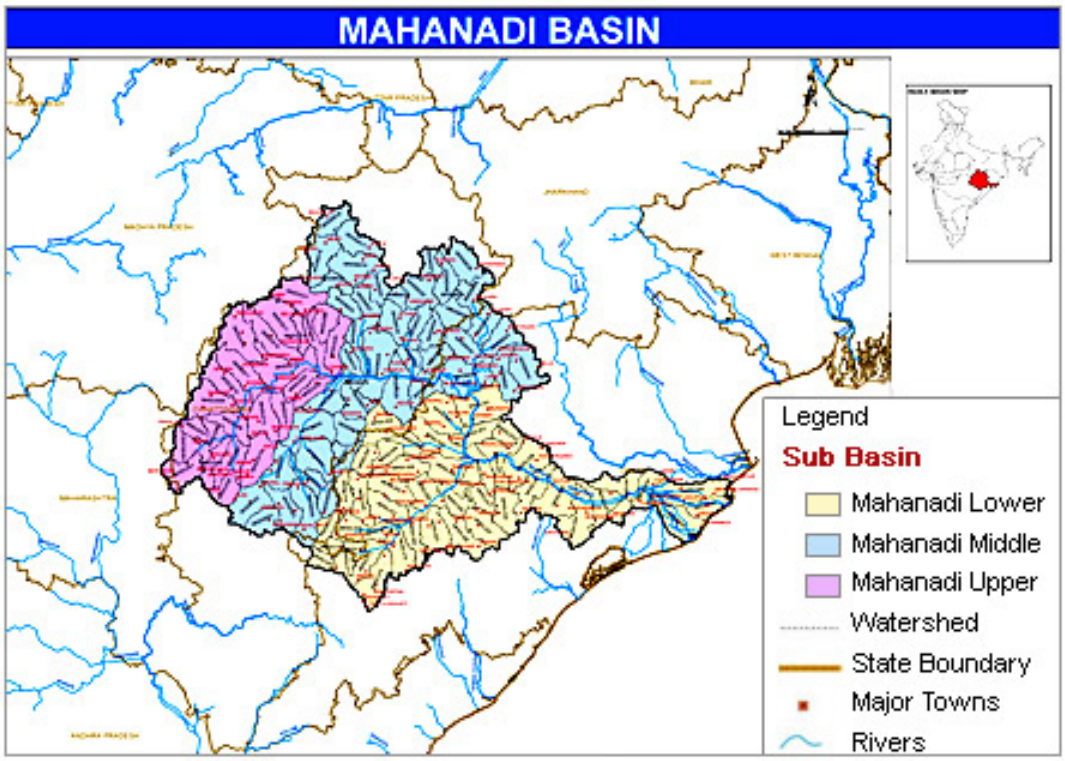

Figure 2.1a Showing the area covered by Mahanadi Basin

The Mahanadi is one of the major peninsular rivers of India draining into Bay of Bengal. It rises in a pool, 6 km from *Pharsiya* village near *Nagri* town in *Raipur* district of *Chhattisgarh* State at an elevation of 457 m. The Mahanadi splits into several streams just before falling into the Bay of Bengal. The total length of the river is about 851 km out of which 357 km is in *Chhattisgarh* and 494 km is in Orissa. The important tributaries of river Mahanadi are: *(i) Sheonath, (ii) Jonk, (iii) Hasdeo, (iv) Mand, (v) Ib, (vi) Tel, and (vii) Ong*. The drainage area of the Mahanadi is *141589*

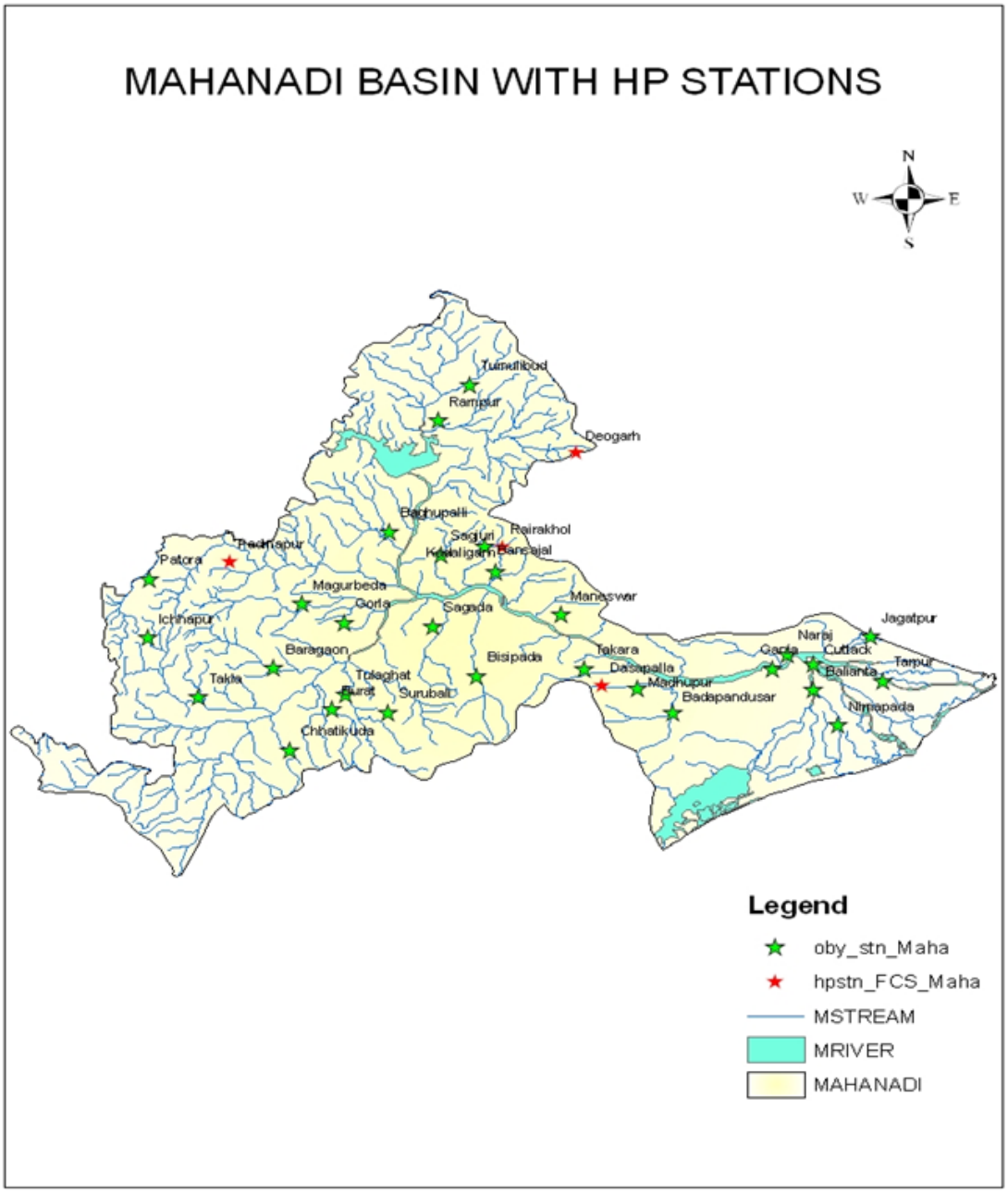

Figure 2.1b River system and Observatory sites under Hydrology project in Mahanadi basin.

km^2 extending in *Madhya Pradesh* (*107* km^2), *Chhattisgarh* (*75229* km^2), *Orissa* (65889 km^2), *Maharashtra* (*238* km^2) and *Jharkhand* (126 km^2) States. The proposed Mahanadi region study

pertains to the basin that is inside the state of Orissa. The catchment area of the Mahanadi up to *Nimapara* observatory site is *65628* km^2.

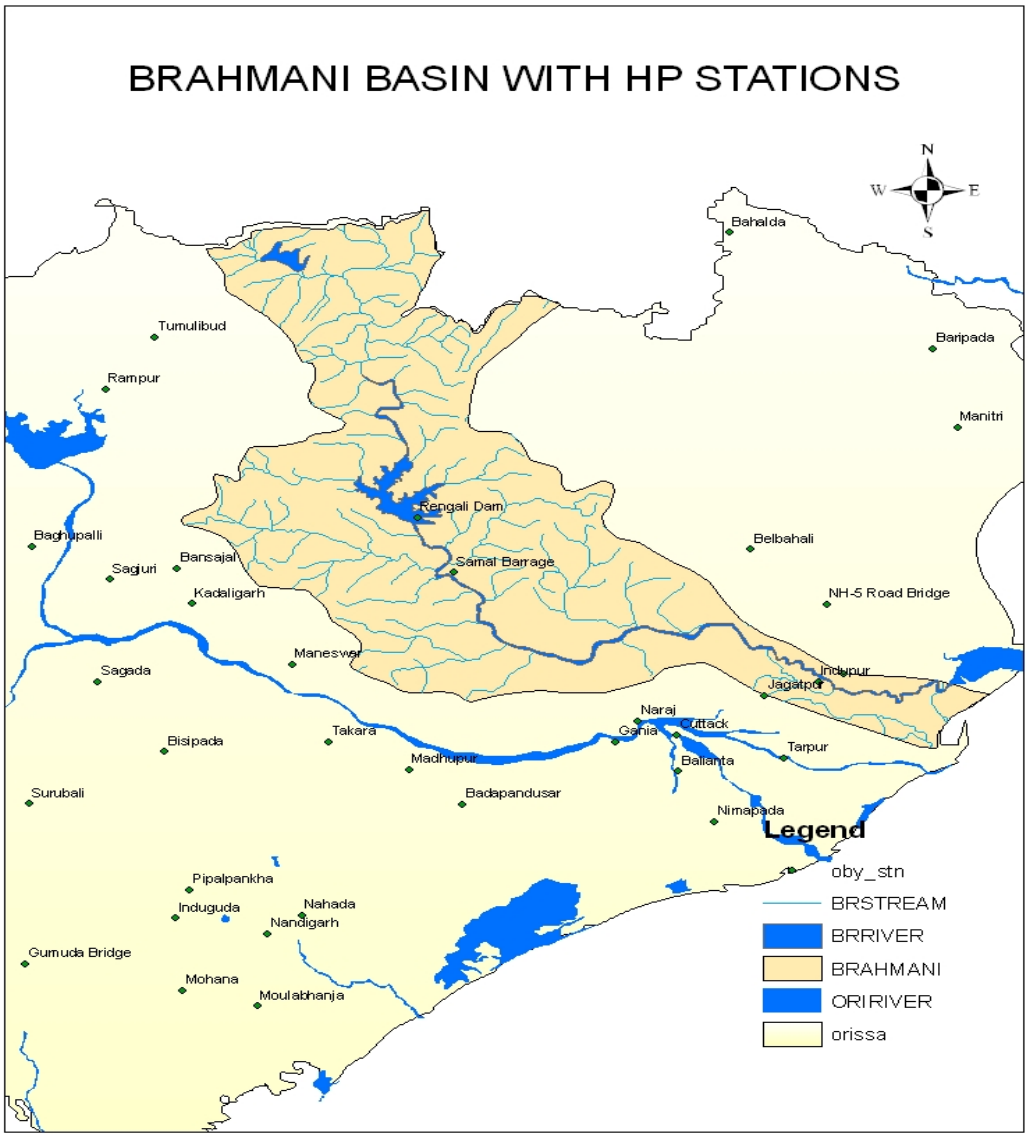

Figure 2.1c River system and Observatory sites under Hydrology project in Brahmani basin.

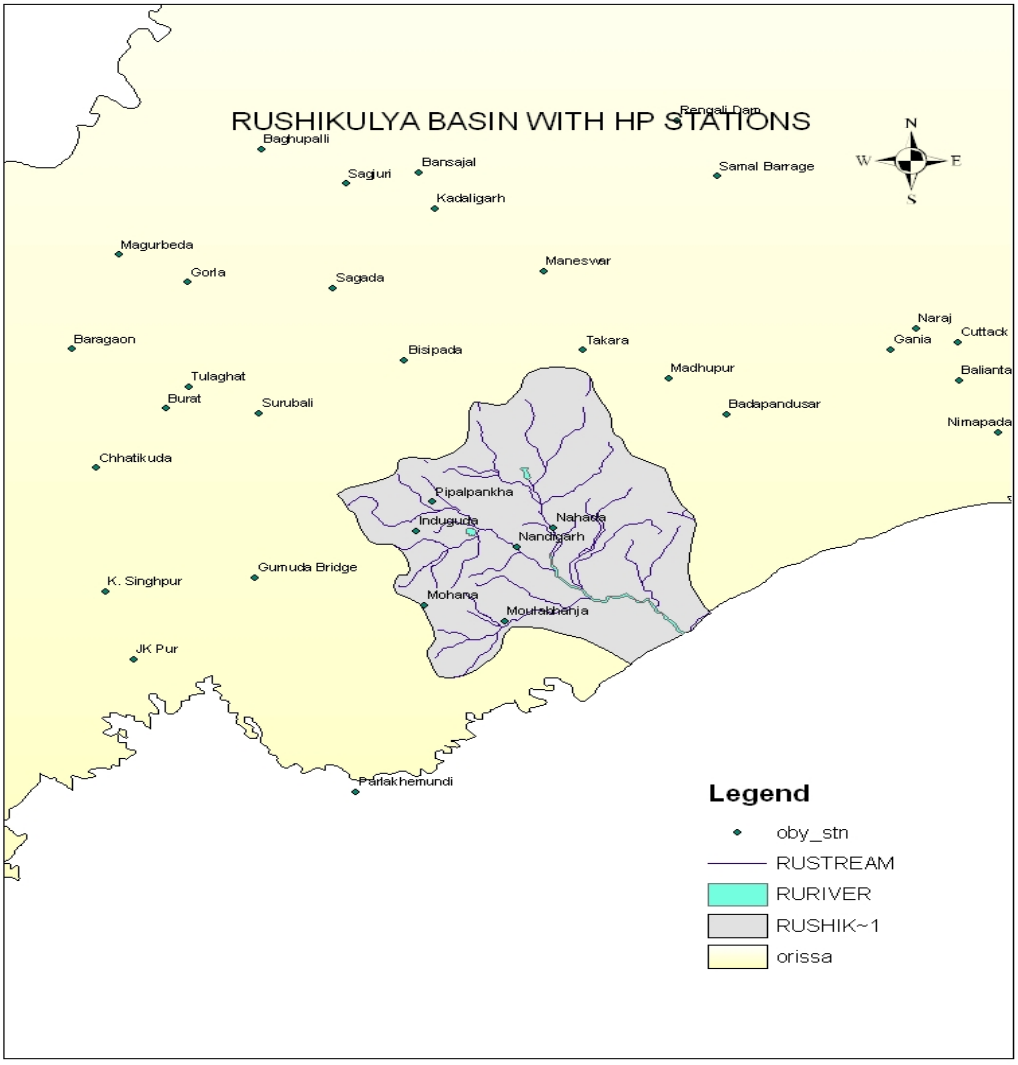

Figure 2.1d River system and Observatory sites under Hydrology project in Rushikulya basin.

2.1 Mahandi-Brahmani-Rushukulya River Basin

The *Brahmani* is formed by the confluence of the rivers *South Koel* and *Sankh* near the major industrial town of Raurkela at 22 15'N and 84 47' E.. It is the second largest river in Orissa. The Brahmani Basin lies between latitude 20° 28' to 23° 35' N and longitude 83° 52' to 87° 30' E in the in the states of *Orissa, Chhattisgarh*, and *Jharkhand* .Both of these sources are in the Chota Nagpur Plateau. Brahmani basin is situated between Mahanadi Basin (on the right) and Baitarani Basin (on the left). It then passes the town of *Bonaigarh* in *Angul* district before

being dammed at Rengali. A large reservoir of the same name is created as a consequence. It then flows through the cities of *Talcher* and *Dhenkanal* before splitting up into two streams. The branch stream called Kimiria receives the waters of the *Birupa* (a distributary of the *Mahanadi*, *Kelo* and *Genguti* before re joining the main stream at *Indupur*. A distributary called *Maipara* branches off here to join the Bay of Bengal. At about 480 km long, the Brahmani is the second longest river in Orissa after the *Mahanadi.* However if its constituent rivers are included its length extends to about 799 km, 541 of which is in Orissa. It has a catchment area of about *39,033* km^2 in Orissa. The river system and observatory sites are given in Figure 2.1 b.

The *Rushikulya* rises from the *Rushyamala* hills of the Eastern Ghats in Kandhamal district and flows in the south east direction and falls into the Bay of Bengal near Chatrapur. It has no delta in its mouth and the total c*atchment area:* 8,963 Km^2. There are several streams between the river Mahanadi and Rushikulya flowing east and draining into *Chilika* Lake i.e. lake under Bay of Bengal. The basin is roughly fan shaped and is undulating and sloping hillocks and eroded mounds. The entire area can be grouped under flat plains with isolated hills and long ridges. There are several streams flowing between Mahanadi and Rushikulya river systems and draining into the *Chilika* Lake. The important streams are: (i) *Todinala, (ii) Dultal dhara nadi, (iii) Kharia or Khaljhor nadi, (iv) Kusumi nadi, (v) Salia nadi and (vi) Khalhor nadi*. The river system and observatory sites are given in Figure 2.1 c.

2.1.1 Topography

Mahanadi is a major sedimentary basin located along the east coast of India (Figure 2.1a), and rises from the district of Madhya Pradesh and flows about 851 km before it joins the Bay of Bengal. The area that is covered in the state of Orissa is 65628 km^2. The length of the stream along Orissa is 494 km and the Mahanadi River system i.e. *Mahanadi, Brahmani,*

Baitarani and Dhamara rivers with a sediment load to the basin of the order of 7.10 x 109 kg/yr (Subramanian, 1978). *Mahanadi* River Delta encompasses an area of 0.1 km, which counts nearly 4.3% of the total geographical area of the country. The geography of Mahanadi River is credited for irrigating a fertile valley where crops of rice, oilseed, and sugarcane are grown.

2.1.2 Geology

Geographically, the delta of River Mahanadi can be divided into four zones, namely, the Northern Plateau, the Eastern Ghats, the Coastal Plain and the Erosional Plains of Central Table Land. The first two zones are the hilly regions, while the coastal plain is the central interior region of the delta, which is traversed by the river and its tributaries. The main soil types found in the basin are red and yellow soils, mixed red and black soils, laterite soils and deltaic soils. *Mahanadi* basin predominantly consists of Archaean rocks represented by folded Khondalites, Granite gneisses and Charnockite. They are inter banded and the first two appear to grade into one another. Field relationship of these rocks is complex and it is difficult to assign any age relationship. It is generally agreed that the rocks were deposited during Archaean era (2000 - 2500 Million years) and were folded and metamorphosed by at least two tectonic activities. The rocks have experienced metamorphic conditions of amphibolites to granulate fancies. Magmatization appears to have played a major role in resulting present State of rocks. The area has been reported to have experienced block faulting during *Gondwana* times and area upstream of *Badmul* appears to be a *Graben*. The area of the basin covering the streams between Mahanadi and Rushikulya mostly comprises of consolidated gneiss rock formation and semi-consolidated tertiary sedimentary. Alluvium is also found in the catchment. Alluvial track extends almost to entire length of coastal area. Groundwater is available in the region of alluvium in the confined and unconfined aquifers below 300 m.

2.1.3 Hydrogeology

Annual, seasonal, monthly and daily variation in the water and sediment discharge in the Mahanadi River is as follows: the water discharge varies from 9.61 to 1809.71 m^3s^{-1} and the total suspended matter varies from 130 to 806 mg l^{-1}. The smaller tributaries upstream carry higher sediment concentration. No yearly cyclic pattern in the water flow in the basin is noticed. More than 95% of the sediment discharge takes place in the monsoon, of which the discharge during July, August and September is 90% of the annual total. On certain days of the year, the sediment discharge accounts for 10–15% of the total annual load. Water discharge, rainfall, geology of the basin and the smaller tributaries upstream seem to control the bulk of the sediment discharge. The Mahanadi annually delivers 15.74 million tonnes of sediment to the Bay of Bengal and more than 80% of the sediment load is carried in the coarse silt fraction. Bulk of suspended sediment is carried in fine fraction. Its mineralogy is characterized by clay minerals *illite, kaolinite, smectite* and *chlorite* in addition to quartz and felspar. The river annually transports a cumulative load of 33.59 million tons into the Bay of Bengal. The dissolved load constitutes less than 25% of the total. Present study reveals that Mechanical load is strongly correlated with discharge, elevation and area while a weak correlation exists between bed slope and mechanical load. The ratio of mechanical to chemical erosion rate an index of erodibility of the basin seems to be related to rock type in the catchment area. Higher values of erodibility index correspond to higher elevations, lesser area and hard silicate rocks of sub basins. On the other hand, sub-basins with low to moderate elevation and chemically active sedimentary rocks are characterized by low values. The river has registered 14% decrease in sediment load transport over a decade due to anthropogenic activities.

2.1.4 Climate

The Mahanadi River Basin is located in the monsoon region of India, and its mean annual flow is 66,640 million m^3. An earlier study revealed that the surface air temperature over the basin has increased at a rate of 1.1 C° per century, which is highly significant, however, there has been no considerable change in the precipitation regime, though a slight decreasing trend is observed. In this study, an analysis of the trends in the runoff of the upper catchment and the whole catchment gauged sites is presented. The results show a steady decrease in the river flows at some locations during the 30 year period of the study; this is significant statistically at the 1 per cent level. In order to increase confidence in this result, the time series of these indices also show a clear declining trend during the period 1901–80. The main result obtained here is that climate warming that occurred over the basin, without being offset by an increase in precipitation (in fact a slight decrease in precipitation was observed), has resulted in a gradual decrease of river flows of the upper catchment as well as of the entire basin during the period 1970–1998. Nevertheless, it is hoped that recent efforts toward development of nested regional climate models may soon make it possible to have acceptable climate simulations for regions and sub regions of this scale. The results of this article may be useful at that stage in hydrological forecasting studies for the basin.

About 90% of the annual rainfall is received during the monsoon period, i.e. from July to October. The average maximum and minimum temperatures are 29 °C and 21°C respectively. The temperature at Raipur varies from 12-42 °C and Cuttack is 14 – 40 °C. The monthly mean relative humidity data of the basin (for five IMD stations) show that the maximum and minimum values of humidity are 95% and 9% during monsoon and summer seasons respectively. The maximum and minimum wind velocities are 16.1 km/hr and 1.0 km/hr respectively at Cuttack and Puri districts. Usually the maximum cloud cover is observed during the months of July or

August whereas minimum cloud cover is observed during December. The monthly average coefficients of sunshine vary from 0.469 and 0.736. The monthly average evaporation at Cuttack varies from 3.0 cm (in December) to 7.5 cm (in May).

2.1.5 Rainfall

Besides the raingauge stations maintained by IMD, there are about 200 observatory stations maintained by state and Hydrology Project within whole Mahandi basin. Salient raingauges available recently for the upper and lower Mahanadi region is given in Table 2.1. The normal annual rainfall of the *Mahanadi* catchment is 141.7 cm. The maximum, minimum and the average annual weighted rainfall values are computed to be 3669.8 mm (*Bulandarpara* in 1944), 660.0 mm and 1458 mm respectively for the period from 1901 to 2008. Figure 3.1(a-c) in the next chapter discuss in details about the number of rain gauge stations that exist in and around Mahanadi maintained by IMD, CWC and state observatory.

2.1.6 Gauging and Discharge (GD) Sites in Mahanadi, Brahmani and Rushukulya basins

Given the large size of the Mahanadi basin and plenty of available water, it is natural that a number of projects have been constructed for utilization of these resources. For this study three types of runoff data are used: (i) monthly runoff series, (ii) Annual maximum series (AMS), and (iii) short –term (hourly) data. Monthly data is used for water balance, consistency check, and rainfall-runoff modeling and the annual maximum data are used for return period flood estimations using AMS modeling. All the runoff (daily, monthly average) were collected from CWC and some of the dat was also collected during tours from state hydrometry (Director of Hydrometry, Water Resources Department, Govt. of Orissa). The salient floods in thirty nine gauging sites attempted in this study is briefly given in Table 2.2 (a-b).

Table 2.1 Rain gauge stations in Mahanadi Zone (R/F is daily mean rainfall in MM)

STATION	R/F	STATION	R/F	STATION	R/F	STATION	R/F
Upper Mahanadi Basin							
Ambabehna		Dharmjaygarh	8.2	Kurubhatta	10.6	Sankara	1.4
Ambikapur	0.0	Ghatora	12.2	Laikera	46.0	Saradihi	43.8
Andhiyokore		Ghorari	25.8	Manedragarh	3.8	Seorinarayan	2.6
Baikanthapur	0.0	Hemgir	14.0	Mahupalli	19.2	Simga	17.8
Bamandini	0.8	Jagadalpur	6.2	Nandaghat	3.0	Sundergarh	31.6
Bangodam	4.2	Jamadarpalli	24.8	Nuapara		Surajgarh	10.6
Bargaon	16.	Jamankira	22.6	Paramanpur	6.8	Tarapur	0.0
Basantpur	8.0	Jharsuguda	4.6	Pendraraoad	0.6	Thettang	3.8
Boranda	8.2	Kelo	4.0	Raipur	23.6		
Champua		Korba	0.2	Rajim	6.2		
Deogan	6.0	Kuchinda	23.6	Rampur	2.6		
Lower Mahanadi Basin							
Alipingal	63.6	Daspalla	38.4	Mundali	40.0	Salebhatta	16.2
Armpur	37.2	Dunguripali		Naraj	43.0	Sambalpur	18.2
Athagarh	68.0	Gop	31.0	Naringhpur	33.0	Sohela	
Athamalik	30.4	Harabhanga	19.0	Nayagarh	51.0	Sonepur	15.0
Baliguda	18.0	Hirakud	14.6	Nimapara	46.2	Tangi	38.0
Banapur	37.0	Jaipatna	35.0	Padmapur		Tikabali	43.0
Khandapada	71.0	Junagarh	56.0	Padmavati	86.4	Tikarpara	53.0
Barmul	29.6	Kakatpur	68.0	Paikamal		Titilagarh	3.4
Belgaon	16.4	Kantamal	22.0	Paradeep	162.0	Pipiii	4.2
Bhawanipatna	48.0	Kendrapara	75.0	Patnagarh			
Bhubaneshwar	50.6	Kesinga	12.6	Jatamundai	94.0		
Bijepur		Kairmal	15.1	Phulbani	15.2		
Bolangir	34.0	Khriar	43.0	Puri	48.8		
Boudhraj	20.0	Komna	98.6	Rairakhol	19.0		
Burla	14.6	Lanjgarh	105.0	R.K.Nagar	22.6		
Cuttack	37.8	M.Rampur	112.0	Ranpur	34.4		

In spite of water management through various dams, diversion works like weirs and channels, flooding is caused primarily due to the Mahanadi River so far as Orissa is concerned. The entire deltaic region of Mahanadi River beyond *Munduli* Barrage intercepting a catchment of 48700 sq km (Mishra and Behera, 2009) gets affected by medium to severe flood almost every year causing immense loss to life and property. In addition, a study on historic flood events in Mahanadi Delta (1969 to 2011) reveals that 69% major floods are caused due to contribution from downstream uncontrolled catchment beyond Hirakud reservoir.

Table 2.2a Gauging Discharge sites maintained by CWC on Mahanadi basin and nearby regions.

Sr. N.		District	Cat. Area (sq Km)	River Name / Tributary Type	Period	Latitude/ Longitude	Length of Stream (Km.)
1	Anandpur	Keonjhar	8570	Baitarani	1972-2009	21 12 40-86 07 21	76
2	Altuma	Dhenkanal	830	Brahmani / Ramiyal	1990-2009	20 55 01-85 31 08	265
3	Champua	Keonjhar	1710	Baitarani	1990-2009	22 03 57-85 40 56	76
4	Gomlai	Sundargarh	21950	Brahmani	1979-2009	21 48 00-84 57 30	52
5	Jaraikela	Sundargarh	9160	Brahmani/ Kelo	1972-2009	22 19 08-85 06 19	49
6	Jenapur	Jaipur	33955	Brahmani	1980-2009	22 55 00-86 00 00	276
7	Pamposh	Sundargarh	19448	Brahmani	2000-2009	22 10 00-84 50 00	89
8	Talcher	Angul	29750	Brahmani	1986-1996	20 57 00-85 15 00	174
9	Kantamal	Boudh	19600	Mahanadi / Tel	1990-2009	20 38 49-83 43 45	115
10	Kesinga	Kalahandi	11960	Mahanadi / Tel	1979-2009	20 11 15-83 13 30	94
11	Salebhata	Bolangir	4650	Mahanadi / Ong	1973-2009	20 59 00-83 32 30	49
12	Sundargarh	Sundargarh	5870	Mahanadi / IB	1972-2009	22 15 45-84 03 00	142
13	Sukuma	Bolangir	1365	Mahanadi / Tel / Suktel	1989-2003	20 48 30-83 30 00	
14	Tikerpada	Bolangir		Mahanadi / Tel	2000-2009		89

Table 2.2b Details of the 25 small catchments taken from CWC (1997) report

Sl. No.	Bridge No.	Name of Stream	Name of Section	Period	Location	Catchment Area (Sq. Km.)
					Latitude/ Longitude	
1	7	Bhadon	Jharsuguda – Titlagarh	1957-1986	21 47 06-84 02 00	3108
2	121	Kelo	Jharsuguda – Bilaspur	1966-1986	21 57 45-83 18 18	1150
3	489	Karo	Bondamunda – Ranchi	1967-1986	22 58 24-84 56 36	0823
4	12	Lilagar	Kharagpur – Nagpur	1958-1989	22 02 00-82 20 00	0666
5	195	Jira	Sambalpur - Titlagarh	1977-1986	21 16 00-83 38 18	0615
6	235	Ranjhor	Sambalpur - Titlagarh	1976-1989	21 07 48-83 35 12	0312
7	332 (ii)	Parri	Raipur – Nagpur	1965-1986	21 06 36-81 03 48	0225
8	385	Sandur	Raipur – Viziram	1966-1989	20 04 36-83 21 00	0194
9	69	Borai	Jharsuguda – Bilaspur	1977-1989	22 00 36-82 58 42	0173
10	59 (B)	Karwar	Kharagpur – Nagpur	1967-1989	22 01 00-82 53 42	0136
11	698	Bisra	Bondamunda – Nagpur	1965-1989	22 15 18-84 57 12	0113
12	48	Barajhor	Nergundi – Talchir	1957-1989	20 41 00-85 35 12	0109
13	79	Kisindajh	Nergundi – Talchir	1957-1986	20 49 06-85 17 24	0067
14	37	Barajhor	Nergundi – Talchir	1967-1989	20 34 36-85 43 48	0064
15	154	Aherajhor	Kharagpur – Nagpur	1968-1986	21 51 38-83 41 35	0058
16	59 (S)	Dungajhor	Kharagpur – Nagpur	1957-1989	22 08 13-84 25 37	0047
17	139	Karo	Rajkharaswar – Gua	1957-1989	22 15 10-85 25 00	0374
18	90	Phaljhor	Bhilai – Dhallirajhare	1975-1986	20 50 12-8117 00	0193
19	332 (i)	Pitakalia	Kharagpur – Walter	1966-1989	21 14 48-86 38 24	0175
20	66 (K)	Hailania	Bilaspur - Katni	1958-1986	22 40 00-81 56 00	0154
21	478	Gokana	Bilaspur – Raipur	1957-1989	22 03 12-82 08 24	0144
22	25	Nun	Kharagpur – Puri	1957-1986	20 03 30-85 45 00	0132
23	40 (K)	Sargood	Bilaspur - Katni	1966-1989	22 32 54-81 56 42	0115
24	176	Sildana	Raipur - Vizanagram	1957-1986	20 50 48-85 32 24	0066
25	42	Barjhor	Nergundi - Talchir	1970-1989	20 37 10-85 39 00	0049

2.1.7 Major Projects and Previous Studies

In the context of flood scenario, the Mahanadi system can be broadly divided into two distinct reaches: (i) Upper Mahanadi (area upstream of *Mundili* barrage, intercepting a catchment of 132100 sq km) (Mishra and Behera, 2009, which does not have any significant flood problem (ii) Lower Mahanadi (area downstream of *Mundili* barrage, intercepting a catchment of 9304 sq km). The key area downstream of Hirakud up to Munduli intercepting a catchment of 48700 sq km is mainly responsible for flood havoc in the deltaic area of Mahanadi. Figure 1 shows the details of catchments of Mahanadi Basin inside and outside of Orisssa and Figure 2 shows the schematic diagram of Mahanadi lower reach beyond Munduli up to Bay of Bengal respectively.

Catchment degradation in its various forms continues without effective control measures due partly to uncertainty regarding the adverse effects on water resources. This uncertainty again arises from lack of adequate hydrological data that should enable quantification of effects of specific land use practices on quality and quantity of water resources. In addition, floods and droughts occur with frequencies and magnitudes that are poorly defined in Mahanadi basin because of lack of relevant hydrological data. Lack of adequate hydrological data introduces uncertainty in both the design and management of water resources systems. Much of *Mahanadi* sub-basins i.e. has as a result of low conversion of rainfall to runoff a precarious balance between available water resources and water

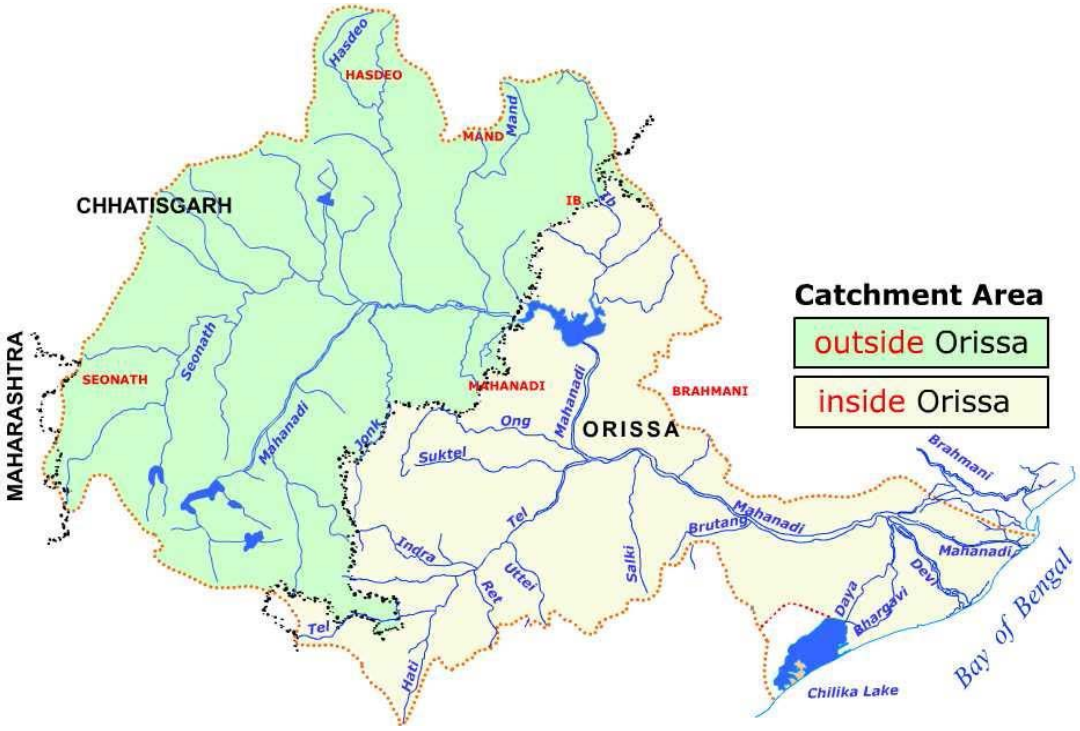

Figure 2.1e Details of Catchments of Mahanadi System inside and outside of Orissa (Source: Parhi et al. 2009)

demand. Given below (Table 2.2.) briefly gives the small catchments and gauging sites that are being maintained and some data are available for further studies. These have been considered for this study with additional data collected from Director of Hydrometry, Government of Orissa.

The major existing/ ongoing projects in the Mahanadi basin while in the upper Mahanadi-basin the existing projects are mostly of medium and small in size except two major projects. The major project in the upper Mahandi basin is *Hirakud* and *Ravishankar Sagar* dams. The *Saroda* and *Indravati* storage reservoir forms a part of the lower sub-basin also. The details rather in brief are summarized here below.

Hirakud Dam

Hirakud is one of the earliest and prestigious major multi-purpose river valley projects in India after independence. Commissioned in 1957, the reservoir is situated a little downstream of the confluence of Mahanadi with its tributary Ib, 15 km upstream of Sambalpur town. Situated

within the geographical ordinates of 21° 30' and 21° 50' N, and 83° 30' and 84° 05' E, the reservoir has a water spread area of 719.63 km^2 at FRL. The 1,248 m long masonry dam is 61 m high and this, along with the earthen dams, has a combined length of 25.8 km. At the dam site, the maximum annual runoff was 91,900 Mm^3 while the minimum annual runoff was 12,400 Mm^3. The total catchment area up to the dam is 83,400 km^2. With gross storage capacity of 5,818 Mm^3, this is one of the biggest reservoirs. The spillway capacity at FRL is 41,428 cumec.

The average annual rainfall in the region is 152 cm. More than 65% of the vast catchment area stretching over the central Indian plateau is fertile land area. The project has been designed to serve three purposes, namely flood control, irrigation and power. In addition it was planned to supplement supplies to the old irrigation system in the Mahanadi delta. Now, the reservoir serves the irrigation needs of 2,640.38 km^2 of land. The water released through power house irrigates further 4,360 km^2 of CCA in Mahanadi delta. A hydropower plant at the dam has 307.5 MW of installed capacity. Besides, the reservoir produces a fish crop of 350 ton every year. The reservoir space is also used to provide flood protection to 9,500 km^2 of delta area in district of Cuttack and Puri.

Table 2.3 Salient features of Hirakud reservoir

Elevation at FRL	192.024 m
Elevation at DSL	179.830 m
Gross storage capacity	8,136 Mm^3
Live storage capacity	5,818 Mm^3
Dead storage capacity	2,318 Mm^3
Water spread area at FRL	719.63 km^2
Length of the masonry dam	1,248 m
Number of sluices	64
Number of crest gates	34
Crest level of spillway dam	185.928 m
Maximum spillway capacity	41,428 cumec

Ravishankar Sagar

The Ravishankar Sagar project (RSP) dam was constructed in 1978 on the Mahanadi River, at 20° 38' N, 81° 34' E. RSP is about 92 km south of the city of Raipur. It is a multi-purpose reservoir which serves irrigation, hydro-electric power-generation and the industrial requirements of the Bhilai Steel Plant. At the FRL of 348.7 m, the reservoir surface area is 95.40 km^2. The total catchment area is estimated at 3,620 km^2, of which 625 km^2 is intercepted by the upstream dam, Dudhawa and 486 km^2 by Murumsilli reservoir. At full level, the reservoir storage capacity is 909 Mm^3. The maximum depth of the reservoir is about 32 m. The off-taking channels carry water in the order of 11,000 to 30,000 cumec in the peak season. However, the outlets are rarely completely closed. Water level fluctuates by 3 to 5 m in a year. The sediments of Ravishankar Sagar are poor in nutrients and organic matter.

Dudhawa Reservoir

The Dudhawa reservoir is situated at 81° 45' 21" E longitude and 20° 18' 1" N latitude across Mahanadi River near Dudhawa village about 21 km west of Sihawa near the origin of Mahanadi river and 29 km east of Kanker. The reservoir is in Dhamtari district of Chhattisgarh state. The construction of the project started in 1953-54 and it was commissioned in 1963-64. This reservoir is designed to supply water to Ravishankar Sagar Project complex thereby increasing its irrigation potential. Water will also be provided to additional culturable areas under the command of the existing Mahanadi Tandula canal system. The maximum height of this earthen dam is 24.53 m and length is 2,906.43 m. Two subsidiary bunds of the dam have heights of 6.61 m and 2.83 m and lengths 568.42 and 426.70 m, respectively. The catchment area of the reservoir is 625.27 km^2 and gross command area is 566.80 km^2. At the Full Reservoir Level (FRL), the submergence area of the reservoir is 44.80 km^2.

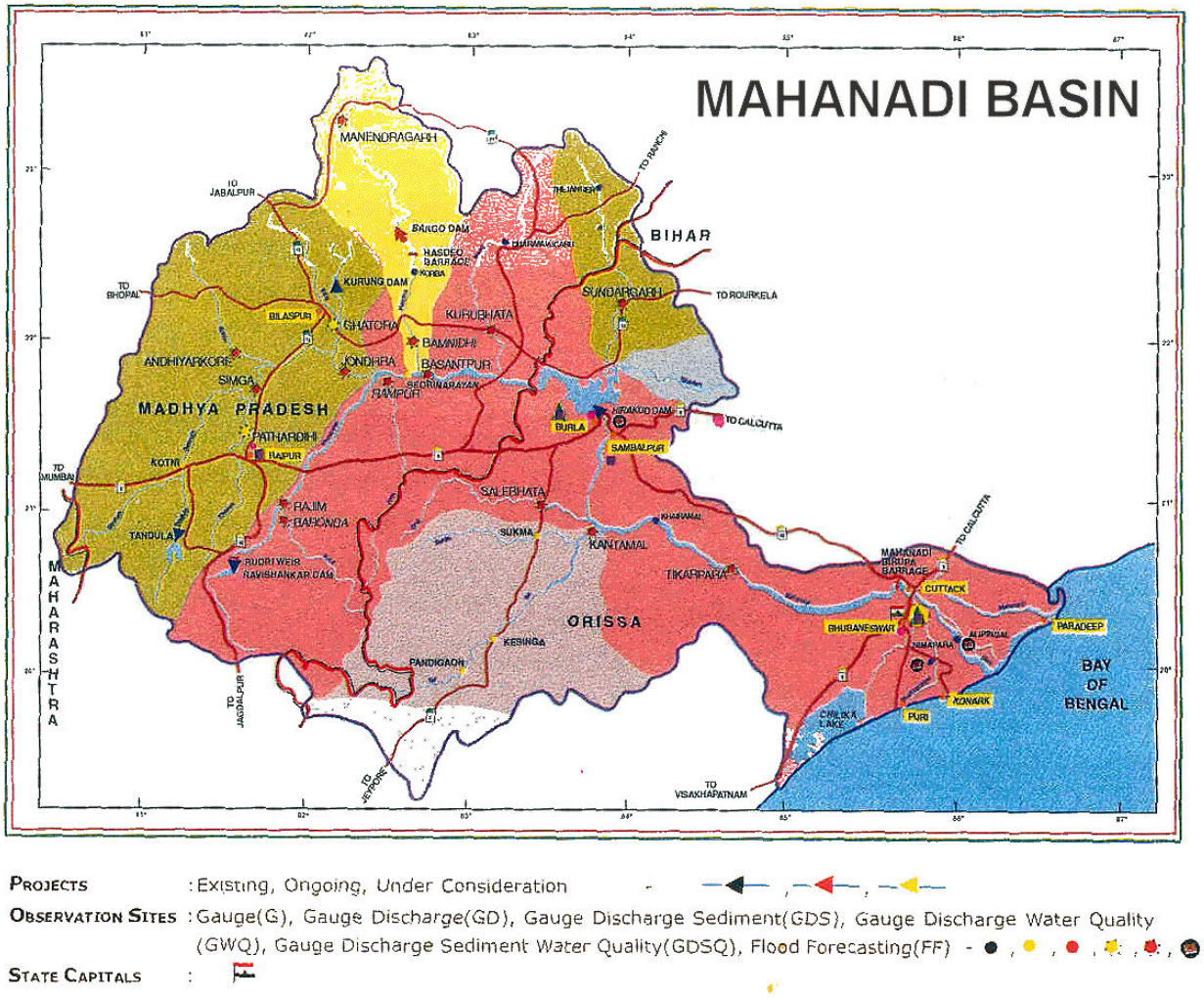

Figure 2.2 Projects status on Mahanadi Basin

Sondur Reservoir

The Sondur reservoir is constructed at 82° 6' E longitude and 20° 14' N latitude across Sondur River a tributary of Mahanadi. Located near Gram Machka, Nagri block, Dhamtari district of Chhattisgarh state, the dam was constructed in the year 1988. The catchment area of Sondur River up to the dam site is 518 km^2. Major portion of the catchment lies in Dhamtari district of Chhattisgarh and Koraput district of Orissa state. Sondur project comprises of a 3.33 km long composite dam. This consists of a 191.25 m long masonry dam at the center which

includes overflow and non overflow portions; the rest is earthen dam at both the flanks. The spillway has 5 radial gates of size 15m × 10m each. Irrigation sluice is provided at the left flank. The project is also designed to supply water to RSP complex through Dudhawa reservoir thereby augmenting the irrigation potential of RSP complex for irrigation. It would also provide irrigation to about 122.60 km^2 of Kharif and Rabi crops in Sihawa Nagri block. The designed rate of sedimentation in gross storage is 0.357 mm /year. Anicuts built on Mahanadi and Baitarani Systems are functional as part of Orissa Canal system, which currently irrigates 3,500 km^2 in Mahanadi and Baitarani basins mostly and also in Brahmani basin by exporting water from Mahanadi. Two of the canals of "Orissa Canal System" starting from Birupa weir (on Birupa River, a branch of Mahanadi River) namely Kendrapara and Pattamundai canals start from the right of the river and proceed to irrigate to the right of Birupa. The third, High Level Canal, Range I originating from the Left of the weir crosses Birupa-Brahmani watershed to enter the Brahmani basin.

Two tributaries, namely Ong and Tel, join Mahanadi downstream of Hirakud dam. Both these carry large volumes of flow during monsoons. The average monsoon runoff at Tikarpara site is 65,636.00 MCM. In absence of any large storage dam lot of flow of Mahanadi goes to the sea and a small portion is utilized in Mahanadi delta. There is an imperative need for construction of a large terminal storage to conserve the precious resource and utilize it to meet the reasonable needs of the basin and transfer the surplus to water short areas. A dam site at Manibhadra dam has been identified and investigated but no progress seems to have been made. Opposition to the proposed dam is on account of large population displacement and other reasons. Live storage capacity of the proposed dam is 6,000 Mm^3. Studies have shown that Mahanadi has surplus water (Govt. of Orissa has not agreed with this conclusion) that is proposed to be transferred from Manibhadra reservoir to the Dowlaiswaram barrage on the

Godavari. In fact, the proposed Manibhadra dam is the starting point of the peninsular component of NWDA's ILR proposal.

Hasdeo Bango

Minimata Hasdeo Bango is a multipurpose storage reservoir on Hasdeo River, tributary of Mahanadi River, 70 km from Korba, in Korba District, Chhatisgarh. The catchment area at the dam is 6,730 km^2. The masonry gravity dam is 87 m high. The FRL and the MDDL of the reservoir are 359.66 m and 329.79 m and it has a live storage capacity of 3,040 MCM. Mean annual inflow to the reservoir is 3,540 MCM. The power house has 3 units of 40 MW each and a firm power of 20 MW.

Tandula

This is an important project of Chattisgarh state. The dam is located in Balod tehsil of Durg district at about 5 km from the Balod city. A dam was completed on the confluence of Sukha Nala and Tandula River in 1921, with a catchment area of 827.2 sq. km. The gross, live, and dead storage capacities of the reservoir are 312.25 MCM, 302.28 MCM and 99.67 MCM respectively. For the reservoir, the highest flood level, the FRL, and MDDL are 333.415 m, 332.19 m, and 320.445 m. A canal takes of from the dam to provide irrigation to 68,219 ha of Kharif crop. Main canal and distributaries run for about 110 km and the length of minors is 880 km. The monsoon rainfall in the command is about 1,293 mm.

In the 1950s, it was realized that the Tandula reservoir is unable to meet the demands of the command and hence a reservoir, named Gondali reservoir, was created on Jujhara Nala in 1957 and a supplementary canal of 9 km length was constructed to supply water from Gondali

reservoir to Tandula reservoir. After construction of Bhilai Steel Plant in 1956, water is being supplied to this plant from the Gondali reservoir and supply for irrigation has been stopped.

CHAPTER 3
PROCESSING AND ANALYSIS OF HYDRO-METEREOLOGICAL DATA

3.1 PROCESSING OF RAINFALL DATA

The precipitation on the basin is in the form of rainfall and the intensity of rainfall varies fairly in time and space as observed from the database. The inventory of existing database for the basin and the daily mean rainfall for the year 2008 is given in Table 3.1.

Table 3.1 Rain gauge stations in Mahanadi Zone (R/F is daily mean rainfall in MM)

STATION	R/F	STATION	R/F	STATION	R/F	STATION	R/F
Upper Mahanadi Basin							
Ambabehna		Dharmjaygarh	8.2	Kurubhatta	10.6	Sankara	1.4
Ambikapur	0.0	Ghatora	12.2	Laikera	46.0	Saradihi	43.8
Andhiyokore		Ghorari	25.8	Manedragarh	3.8	Seorinarayan	2.6
Baikanthapur	0.0	Hemgir	14.0	Mahupalli	19.2	Simga	17.8
Bamandini	0.8	Jagadalpur	6.2	Nandaghat	3.0	Sundergarh	31.6
Bangodam	4.2	Jamadarpalli	24.8	Nuapara		Surajgarh	10.6
Bargaon	16.	Jamankira	22.6	Paramanpur	6.8	Tarapur	0.0
Basantpur	8.0	Jharsuguda	4.6	Pendraraoad	0.6	Thettang	3.8
Boranda	8.2	Kelo	4.0	Raipur	23.6		
Champua		Korba	0.2	Rajim	6.2		
Deogan	6.0	Kuchinda	23.6	Rampur	2.6		
Lower Mahanadi Basin							
Alipingal	63.6	Daspalla	38.4	Mundali	40.0	Salebhatta	16.2
Armpur	37.2	Dunguripali		Naraj	43.0	Sambalpur	18.2
Athagarh	68.0	Gop	31.0	Naringhpur	33.0	Sohela	
Athamalik	30.4	Harabhanga	19.0	Nayagarh	51.0	Sonepur	15.0
Baliguda	18.0	Hirakud	14.6	Nimapara	46.2	Tangi	38.0
Banapur	37.0	Jaipatna	35.0	Padmapur		Tikabali	43.0
Khandapada	71.0	Junagarh	56.0	Padmavati	86.4	Tikarpara	53.0
Barmul	29.6	Kakatpur	68.0	Paikamal		Titilagarh	3.4
Belgaon	16.4	Kantamal	22.0	Paradeep	162.0	Pipiii	4.2
Bhawanipatna	48.0	Kendrapara	75.0	Patnagarh			
Bhubaneshwar	50.6	Kesinga	12.6	Jatamundai	94.0		
Bijepur		Kairmal	15.1	Phulbani	15.2		
Bolangir	34.0	Khriar	43.0	Puri	48.8		
Boudhraj	20.0	Komna	98.6	Rairakhol	19.0		
Burla	14.6	Lanjgarh	105.0	R.K.Nagar	22.6		
Cuttack	37.8	M.Rampur	112.0	Ranpur	34.4		

Figure 3.1(a-c) shows two more nearby basins that is also considered for certain analysis in this project, and they are Brahmani and Rushikulya basins. There are ninety eight rain gauge stations

that exist in and around Mahanadi are being maintained by IMD, CWC and state observatory. These data have been used for computing mean rainfall over four sub-basins: (i) Up to *Patheria*/ *Garakota*; (ii) Up to *Gasiabad*, (iii) Up to *Daudhan,* and (iv) Up to *Banda*. The temporal variations of annual rainfall for four stations within Mahanadi basin are shown in Figure 3.2(a-b). The distribution of the existing rain gauges in the basin taking into account four specific stations is shown in this Figure

The mean rainfall has been calculated using two methods: (i) Arithmetic mean (AM) and (ii) Thiessen Polygon (TP); these are discussed briefly in the subsequent sections in this chapter. Before that, the consistency of the monthly rainfall series of the available rain gauge stations was checked and the results are described in the following section.

3.1.1 Tests for Consistency of Rainfall Series

Consistency of a rain gauge station is used to cross check if the conditions relevant to the recording of a rain gauge station have undergone a significant change during the period of record. Then inconsistency would arise in the rainfall data of that station. This inconsistency would be noticed from the time the significant changes have taken place. Some of the common causes for inconsistency of record are:

i) Shifting of a rain gauge station to a new location,

ii) The neighborhood of the station undergoing a marked change,

iii) Occurrence of observational error from a certain date.

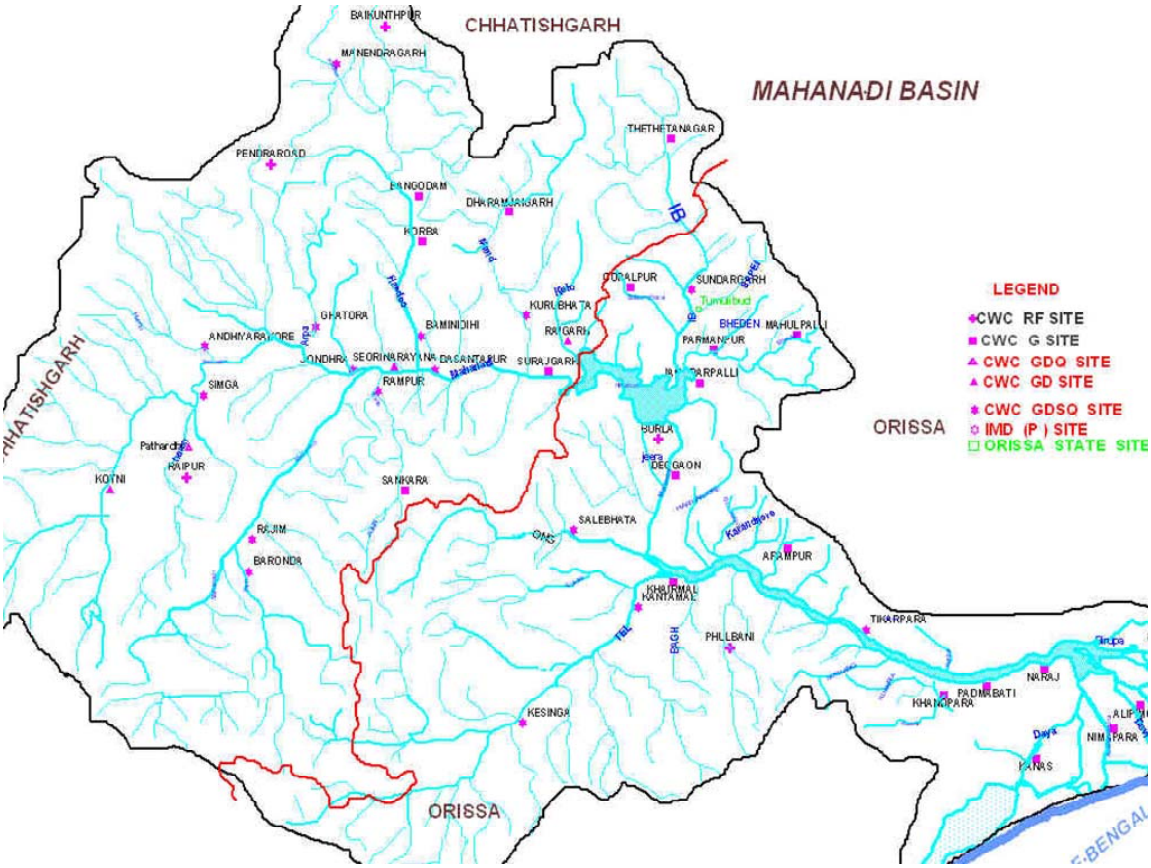

Figure 3.1a Distribution of the existing rain gauges and observatory sites in Mahandi basin.

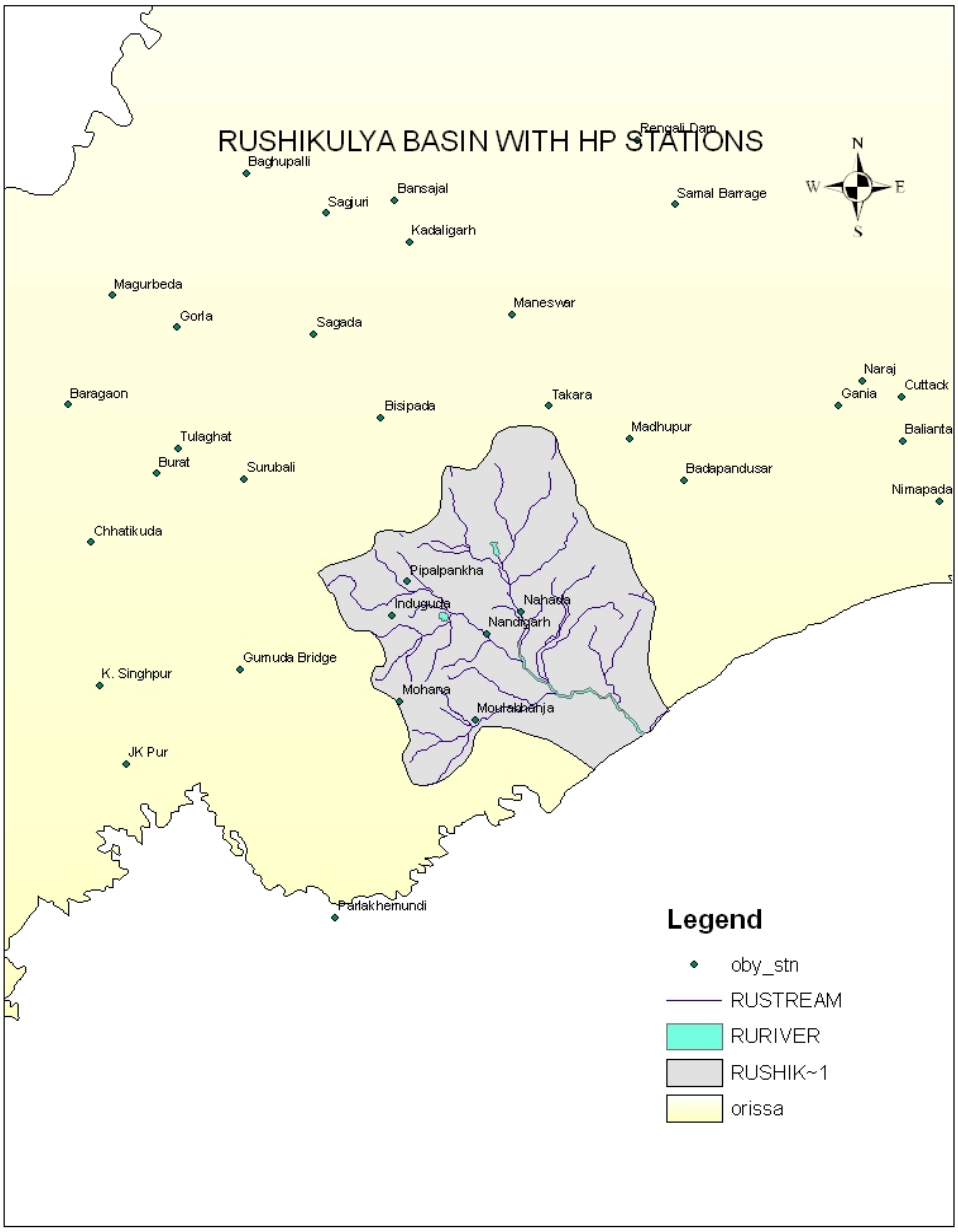

Figure 3.1b Distribution of the existing rain gauges and observatory sites in Rushikulya basin.

Figure 3.1c Distribution of the existing rain gauges and observatory sites in Brahmani basin.

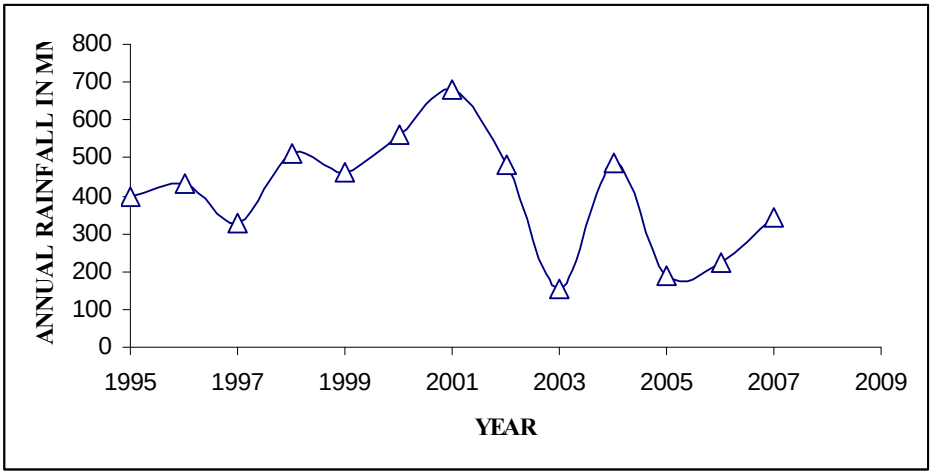

Figure 3.2a Temporal variation of rainfall for mean of *kakatpur* and *Kendrapada* (1995-2007)

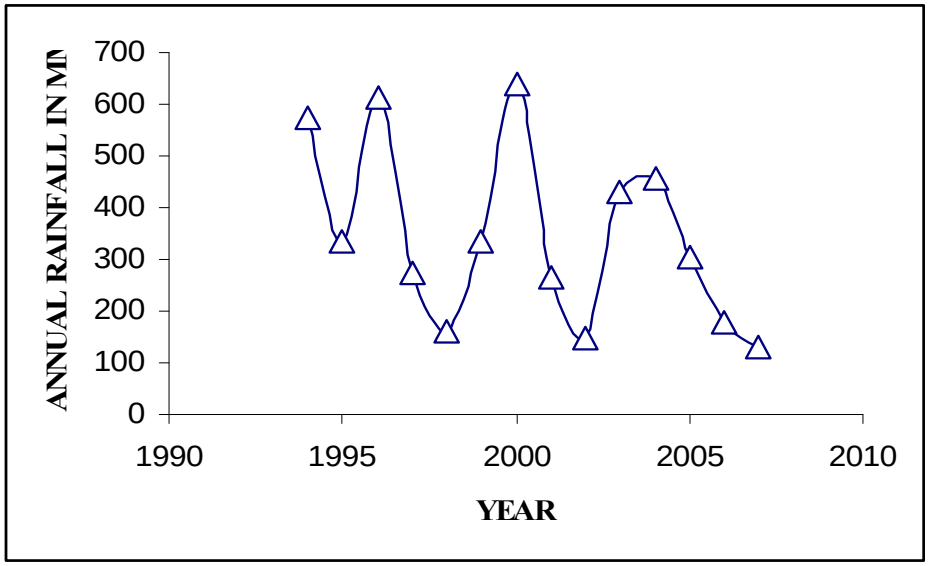

Figure 3.2 b Temporal variation of rainfall for mean of *Hemgir* and *Jagadalpur* (1994-2007)

For this study, the *double mass curve technique* is used to check the consistency of a rainfall series record. This technique is based on the principle that when each recorded data comes from the same parent population, they are consistent.

A group of stations in neighborhood of the index station (for which the consistency test is required) is selected. The data of the annual (or monthly mean) rainfall of the index station X and also the average rainfall of the group of base stations covering a long period is arranged in the reverse chronological order, i.e. the latest record as the first entry and the oldest record as the last entry in the list. The accumulated precipitation of the station X and the accumulated values of the average of the group of base stations are calculated starting from the latest record. Individual values are plotted against cumulative ones for various consecutive time periods. A break in the slope of the resulting plot indicates a change in the precipitation regime of station X. The precipitation values at station X beyond the period of change of regime is corrected by using the following relation,

$$\text{Correction factor} = \frac{X_A}{X_B} \qquad (3.1)$$

where X_A and X_B are the interception lengths in y and x-axis respectively. In this way the older records are brought up the new regime of the station. It is apparent that the more homogeneous the base station records are, the more accurate will be the corrected values at station X. A change in slope is normally taken as significant only where it persists for more than five years. Otherwise the records are presumed to be consistent. The double mass curve is also helpful in checking arithmetic errors in transferring rainfall data from one record to another.

For the study area, the consistency in monthly rainfall series is checked for all the raingauge stations (Table 3.1). From the results, it is observed that there is significant change in slope for the following eight stations:

i) Alipingal

ii) Ghatora

iii) Sundergarh

iv) Laikera

v) Jamadarpalli

vi) Kuchinda

vii) Daspalla

viii) Naringhpur

The rainfall series for the above eight stations showed inconsistency mainly during the month of September. The corrections for the specific period where these rainfall series show a bend in the double mass curve is done using Eq. (3.1). The monthly data of Sundergarh is used as an example to show the procedure to rectify the inconsistency in the data. The double mass curve using the monthly data for September for this station is shown in Figure 3.3 (a).

As observed from Figure3.3 (a), the double mass curve has a bend starting at coordinate X_1 = 6680.22, Y_1 = 5737.1up to X_2 =13643.2, Y_2 = 11517.6. The slope i.e., correction factor for this part as per Eq. (3.1) is $(Y_1-Y_2) / (X_1-X_2)$ = 1.23. The corrections are applied after (X_1, X_2) as follows:

Corrected cumulative rainfall at index station $(Y_i) = Y_{i-1} + (X_i – X_{i-1})$ (correction factor), e.g., the corrected values of monthly rainfall value is:

station (Y_i) = 5661.19 + (6680.22 – 6473.12) x (0.83) = 6070.37

and the procedure is repeated for the next succeeding values. The corrected double mass curve for Sundergarh for the September month is shown in Figure 3.3. (b). The inconsistency found in two more stations i.e. *Kuchinda* and *Daspalla* are reported using double mass curve in Figure

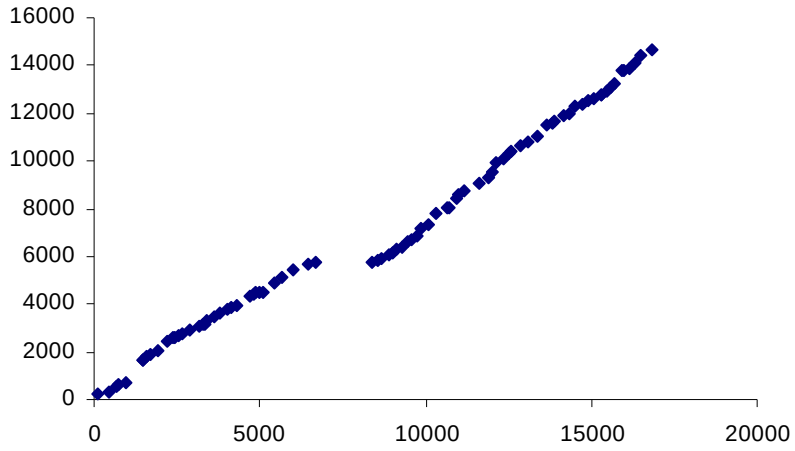

Figure 3.3a Double mass curve for S*undergarh* raingage in Mahandi basin for September

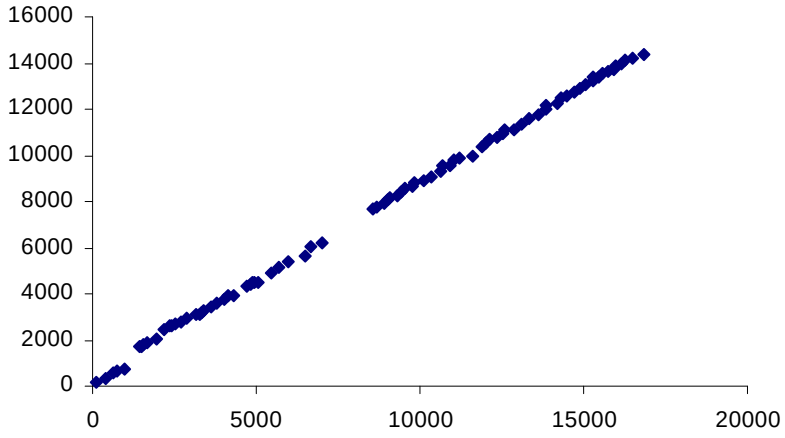

Figure 3.3b Corrected Double mass curve for S*undergarh* raingage in Mahandi basin for September

3.4 (a) and (b), respectively. After completing the consistency check and correcting the records, average aerial rainfall was computed.

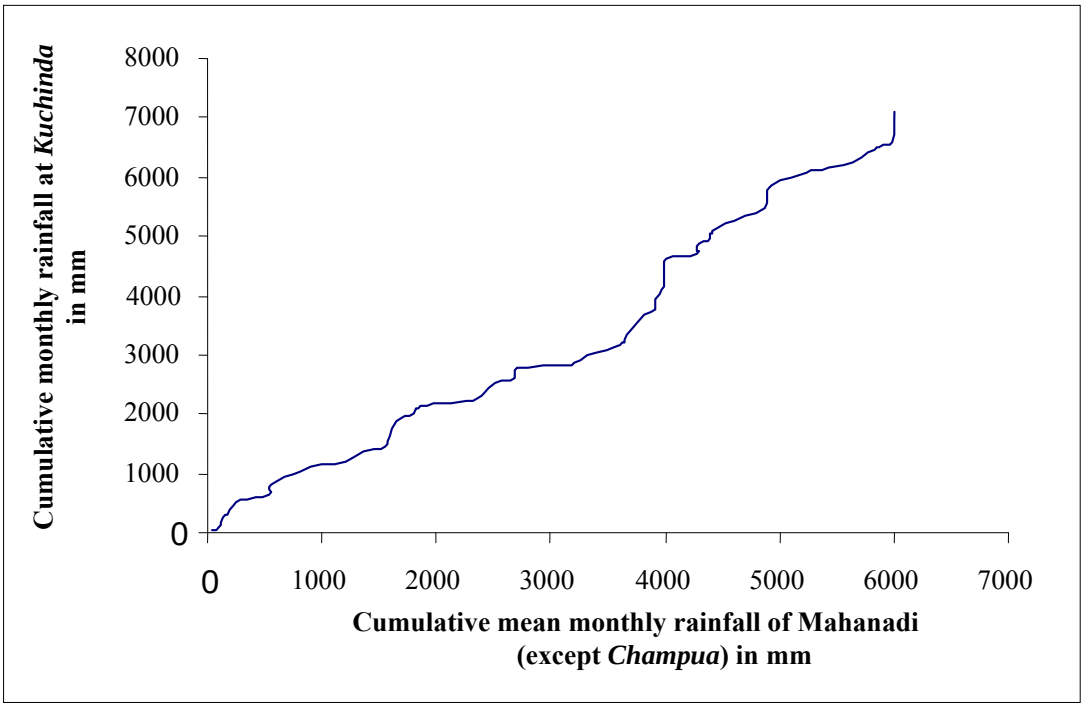

Figure 3.4 a Double mass curve for *Kuchinda* raingage in Mahandi basin for September

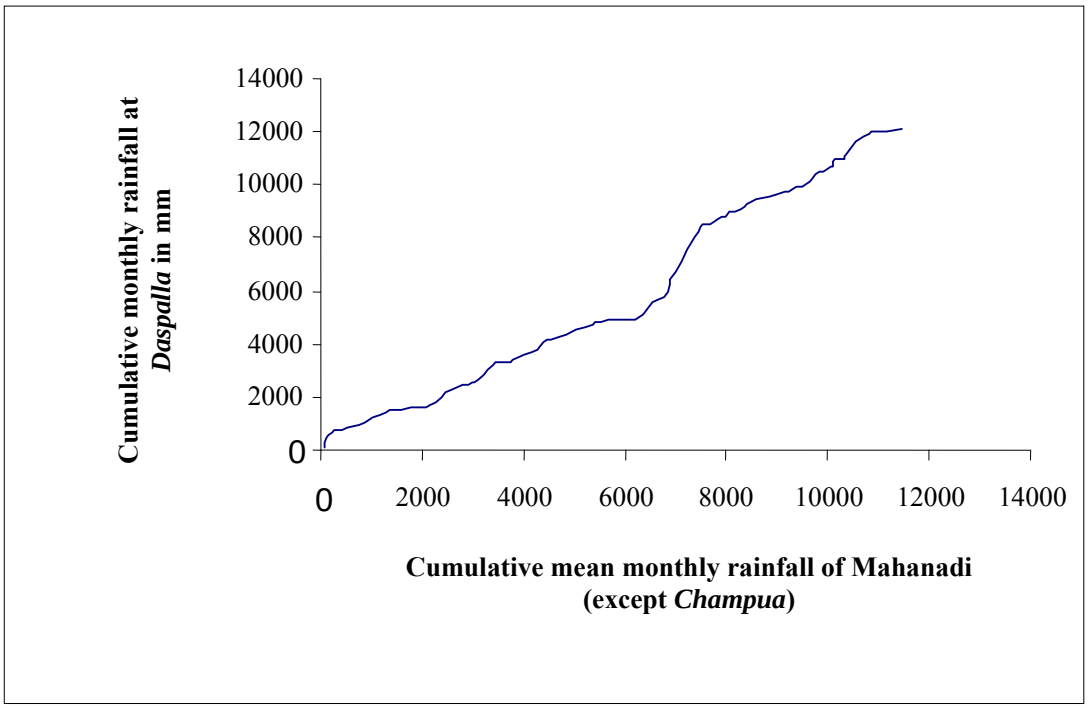

Figure 3.4 b Double mass curve for *Daspalla* raingage in Mahandi basin for September

3.1.2 Arithmetic Mean (AM) Method

In any hydrological analysis the knowledge of the total rainfall over an area is required. However, a rain gauge represents only point sampling of the areal distribution of a storm. Generally three methods are practiced to convert the point rainfall values at various stations into an average value over an area. These are: (i) arithmetic average or mean (AM) method, (ii) isohyetal method, and (iii) Thiessen polygon (TP) method. In the present study, the AM and TP methods have been used to calculate the total rainfall over the study area. In case of arithmetic mean (AM) method, the arithmetic average of the available raingage records is taken to be the rainfall over the basin. Though a detailed analysis using AM method is done for the study area, Figure 3.5 presents the result of AM method for the Mahandi basin up to *Champua* for the month of July.

Table 3.2: Corrections for the inconsistency rainfall records for Champua (Mahanadi basin) for the month of September in 1983

Cumulative Rainfall for the Month of Sept, 1996 at Champua			
PREVIOUS		**CORRECTED**	
Y-axis (index station)	X-axis(Mean rainfall of rest of the stations)	Y-axis (index station)	X-axis(Mean rainfall of rest of the stations)
4868.62	5566.89	4868.62	5566.89
4933.94	5745.00	5220.52	5745.00
4940.34	7216.47	5368.35	6020.00
4986.33	7341.34	6589.67	7341.34
5099.39	7443.61	6693.31	7443.61
5216.92	7642.04	6778.20	7642.04
5306.44	7747.89	6942.89	7747.89
5412.31	7827.01	7030.75	7827.01

3.1.3 Thiessen Polygon (TP) Method

In this method, the rainfall recorded at each station is given a weightage on the basis of an area closest to the station. In this method, Thiessen polygons are drawn around each of the stations by drawing perpendicular bisectors to the lines joining raingage stations. The value of the rainfall at the station is assumed to be acting over that polygon area. Calculation of the average rainfall over the area is done using the following equation:

$$\overline{P} = \frac{P_1A_1 + P_2A_2 + P_3A_3 + ...P_nA_n}{(A_1 + A_2 + A_3 + ... + A_n)} \quad (3.2)$$

where, A_1, A_2,...A_n are respective areas of Thiessen polygons.

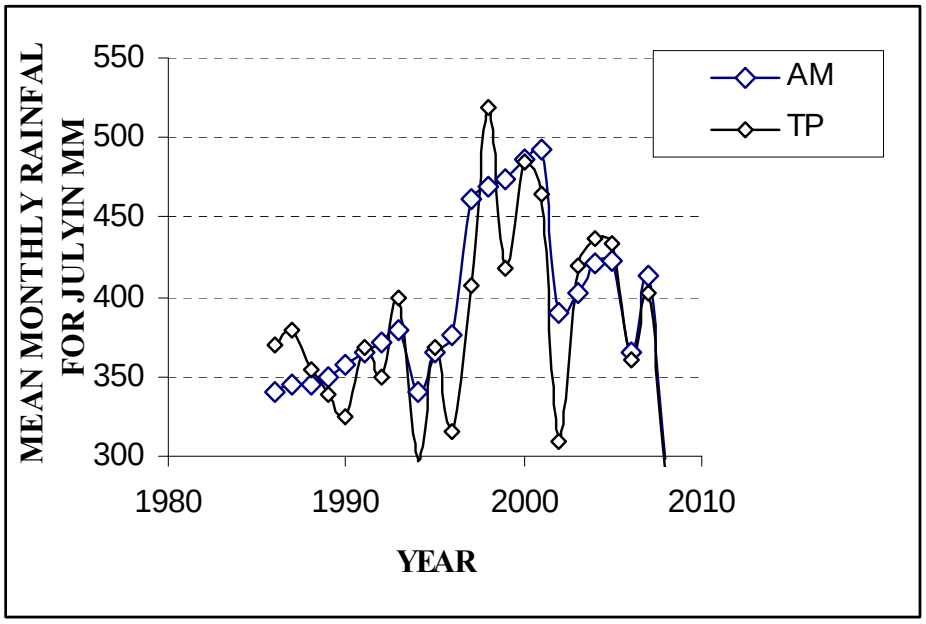

Figure 3.5 Mean monthly rainfall up to *Champua* for the month of July using Thiessen polygon (TP) and arithmetic mean (AM) method

Since the water availability analysis is to be carried out on monthly basis, the daily data were collected from IMD web site and officially from Department of Water Resources (Govt. of

Orissa) and from the earlier reports, were aggregated to monthly data. However, for the purpose of design flood calculations, short term data of four available stations are used separately for the historic periods. This is discussed separately in chapter five. From the data supplied so far, it was observed that the period of availability of data varies for the 21 stations (the length of record is not same for all the stations). Hence to compute the total rainfall for the Mahanadi basin up to fourteen gauging sites, and its sub catchments (given in the table in the nest section in Processing of runoff data) , it was necessary to work on different combinations of rain gauges having concurrent period of data availability.

A computer program was developed for estimation of average areal rainfall of the catchment and sub-catchments using the Thiessen polygon method. For each month, the program forms the combination of rain gauges as per data availability and estimates the Thiessen weights for the stations in the combination. Table 3.3 shows an example of the comparison of AM and TP calculation of mean rainfall (July) for Champua gauging sites. These weights are then used along with the rainfall data of the network to estimate the average areal rainfall. The sample results of the program were checked using manual calculations with the help of a spreadsheet. After the results were found satisfactory, the program was used to estimate monthly average areal rainfall for the period 1986-2007 for the Mahandi basin up to its sub catchments Few samples of these combinations and their respective weightages are reported in Table 3.4 for total monthly rainfall up to *Anandpur* GD site. The results of Thiessen polygon (TP) were compared with the results of the arithmetic mean method and the results show a difference of about ± 10 percent.

Further for this analysis the monthly mean rainfall values obtained using the TP method were compared with corresponding values of mean rainfall derived using isohyetal map for

some randomly selected periods. The comparison gave fairly good approximation between the two. Thus on basis of the above results and following the recommendations of Tabios-III and Salas (1985), Croley and Hartmann (1985), the TP method was adopted to compute mean rainfall values in this study.

Table 3.3 Estimation of average rainfall for Mahanadi basin up to *Champua* GD site

AM	TP	AM	TP	AM	TP	AM	TP
340.8	369.2	371.4	349.0	468.5	519.1	421.4	437.2
345.3	379.4	378.8	399.3	474.4	417.5	423.5	434.2
345.8	355.1	340.8	296.9	486	484.3	365.2	361.7
349.1	338.6	365.4	369.0	492.1	465.2	413.1	403.1
356.9	325.1	375.7	315.2	389.6	308.7	298.7	291.2
365.1	368.4	461.2	406.4	402.1	419.7		

Table 3.4 Estimation of average rainfall for Mahandi basin up to *Anandpur* GD site

Gauge Name	With all data available (e.g., Jul., 1994)			With the data of 14 gauges (e.g., Jul., 1995)			With the data of 10 gauges (e.g., Sep., 1995)		
	Rainfall	Weight	Effective	Rainfall	Weight	Effective	Rainfall	Weight	Effective
Ambabehna	209.7	0.0	0.0	168.6	0.0	0.0	120.3	0.0	0.0
Ambikapur	361.0	0.0	0.0	0.0	2.6	0.0	143.1	0.0	0.0
Bamandini	343.7	5.9	21.6	97.8	5.1	5.7	43.9	5.1	2.6
Bargaon	332.7	7.6	32.4	118.6	7.6	10.5	0.0	9.0	0.0
Basantpur	323.5	4.1	16.2	0.0	0.0	0.0	0.0	9.0	0.0
Boranda	295.1	1.8	0.0	144.0	5.4	0.0	83.5	2.0	0.0
Champua	369.6	6.0	28.6	116.2	7.1	9.6	83.3	0.0	9.6
Ghorari	588.3	9.0	68.2	160.3	9.0	16.8	49.6	1.1	5.2
Hemgir	433.1	4.3	16.5	139.4	3.0	4.8	0.0	9.0	0.0
Jharsuguda	466.1	11.5	69.0	90.4	13.3	14.0	36.9	11.4	7.3
Kelo	555.0	3.7	26.3	163.3	2.6	7.0	109.2	7.9	10.0
Sundergarh	225.4	0.0	0.0	180.7	0.0	0.0	0.0	2.5	0.0
Surajgarh	324.3	3.8	15.6	181.5	3.8	7.9	0.0	4.5	0.0
Tarapur	479.4	3.5	10.2	151.7	2.7	2.9	0.0	9.0	0.0
Kantamal	240.9	5.6	17.3	96.6	5.6	6.3	86.1	4.6	10.6
Kendrapara	251.2	1.1	0.0	103.8	0.0	0.0	0.0	0.0	0.0
Kesinga	257.7	0.0	0.0	164.4	2.8	0.0	72.5	0.0	0.0
Burla	206.2	9.9	26.1	165.7	9.9	19.0	100.7	10.1	11.8
Cuttack	589.1	8.1	16.6	200.4	4.2	5.1	84.4	2.2	2.1
Salebhatta	645.5	7.8	64.6	161.9	7.8	14.7	202.3	7.8	18.4
Sambalpur	681.4	5.5	48.0	178.6	6.4	13.3	138.5	6.4	6.4
Avg. rainfall		100.00	477.3		100.00	137.7		100.00	84.1

3.1.4 Filling of Missing Rainfall Data

Since the Thiessen polygon method is used here for different combinations, i.e., for the available station data (among the existing 21 stations) to formulate a combination, this procedure obviates calculations for filling the missing data. However, the normal annual precipitation method which is popularly employed for filling up of missing data is used here for double mass curve computation as discussed in the following section. The normal rainfall of a station, as defined by National Weather Service, is the mean monthly/annual rainfall over a 30 years period, and the notation used is N_i where i is the station index. The results of normal monthly rainfall values for one station i.e. *Kesinga* in Mahanadi basin are given in Table 3.5.

Table 3.5 Normal monthly rainfall in MM for *Kesinga* (Mahanadi).

YEAR	JAN	FEB	MAR	APR	MAY	JUN.	JUL.	AUG	SEPT	OCT	NOV	DEC.
1974-2002	9.2	13.1	9.9	2.5	3.1	72.5	194.7	266.2	133.3	23.5	2.9	4.9
1966-1996	14.7	8.1	4.5	1.8	2.0	59.3	234.9	216.3	120.3	16.1	9.7	4.4
1980-2007	7.1	6.7	4.1	2.9	5.4	69.6	274.7	234.6	116.2	24.5	10.8	5.5

3.2 PROCESSING OF RUNOFF DATA

For this study three types of runoff data are used: (i) monthly runoff series, (ii) Annual maximum series (AMS), and (iii) short –term (hourly) data. Monthly data is used for water balance, consistency check, and rainfall-runoff modeling and the annual maximum data are used for return period flood estimations using AMS modeling. The short term data, i.e., the hourly runoff data are used for design flood calculation using the unit hydrograph technique. The following section discusses in detail the consistency of the given observed data. The salient floods in thirty nine gauging sites attempted in this study is briefly given in Table 3.6. The first part of the table shows the mean runoff of twenty five bridge catchments taken from CWC

(1982) and the details of duration of the data and other characteristics of the small catchments are given in Appendix-1. The last part of the table that shows fourteen CWC gauging sites names are collected from CWC and details of duration of the data and other characteristics of the catchments engulfing the GD sites are given in Appendix-2. The annual observed and the virgin flow series at *Altuma* gauging site are shown in Figure 3.6.

3.2.1 Comparison Plots Using Residual Series

The simplest and often the most helpful means of identifying anomalies between stations is the plotting of comparative time series. There will of course be differences in the plots depending on the contributing catchment area, differing rainfall over the basins and differing response to rainfall. However, gross differences between plots can be identified. The most helpful comparisons are between sequential stations on the same river. The series may be shifted relative to each other with respect to lag times from rainfall to runoff or the wave travel time in a channel. Comparison of series may permit the acceptance of values flagged as suspect because they fell outside the warning ranges, when viewed as stage or when validated as a single station. When two or more stations display the same behavior, there is a strong evidence to suggest that the values are correct.

Residual series is the plot of differences between data of two stations. For instance, plots of Altuma and Anandpur data are used as a comparison of two sets of time series. Both the original time series (virgin flow series) and their residuals are plotted in the same Figure this is done for (i) monsoon (July-Sept) and (ii) non-monsoon (Oct-June) periods separately.

Table 3.6 General hydrological data for Mahanadi.

Flood data of 25 small catchments from CWC (1997) report							
Sl.	**Br-No.**	**Q_{av} (m^3/s)**	**Name of Stream**	**Sl**	**Br-No.**	**Q_{av} (m^3/s)**	**Name of Stream**
1	7	559	Bhedon	14	478	197	Gokena
2	121	332	Kelo	15	59B	41	Karwar
3	489	271	Karo	16	25	247	Nun
4	12	112	Lilagar	17	40K	131	Sargood
5	195	123	Jira	18	698	52	Bisra
6	139	176	Karo	19	48	21	Barajhor
7	235	53	Ranjhor	20	79	27	Kisindajhor
8	332NGP	85	Parri	21	176	82	Sildha
9	385	58	Sandul	22	37	24	Barajhor
10	90	130	Phaljhar	23	154	23	Aherajhor
11	332KGP	72	Pitakalia	24	42	51	Barajhor
12	69	52	Borai	25	59(S)	42	Dungajhor
13	66k	260	Mailania				

Fourteen Gauging sites maintained by CWC							
Sl	**GD site**	**River/Tributary**	**Q_{av} (m^3/s)**	**Sl**	**GD site**	**River/Tributary**	**Q_{av} (m^3/s)**
1	Altuma	Ramiyal	17.1	8	Kesinga	Mahanadi	320.7
2	Anandpur	Brahmani	181.2	9	Kantamal	Mahanadi	265.8
3	Gomlai	Brahmani	109.3	10	Salebhatta	Mahanadi/Ong	113.2
4	Jaraikela	kelo	189.5	11	Jenapur	Brahmani	505.6
5	Pamposh	Brahmani	485.1	12	Sukma	Mahanadi / Tel	23.5
6	Champua	Brahmani	38.3	13	Talcher	Brahmani	440.2
7	Sundergarh	Mahanadi/Ib	66.7	14	Tikerpada	Mahanadi	1530.4

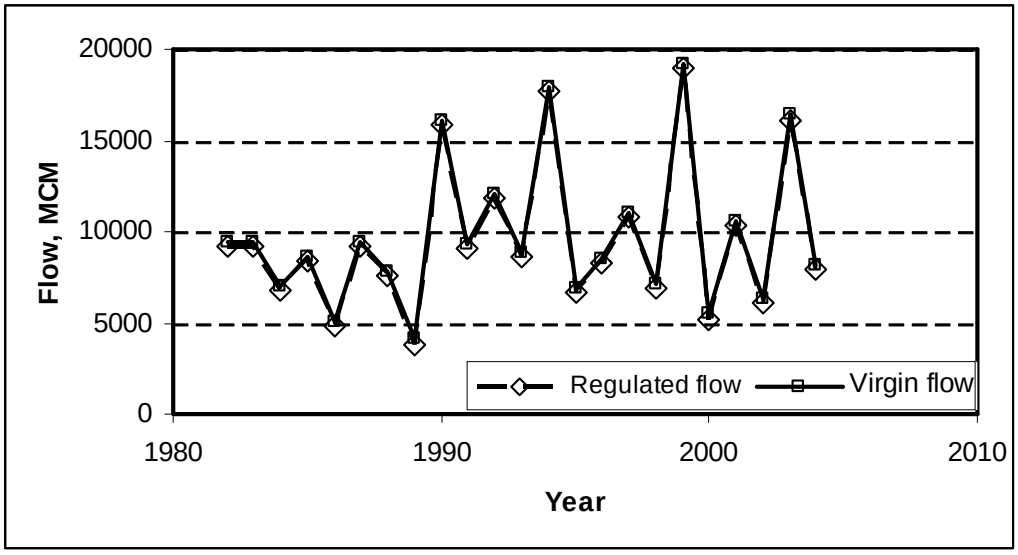

Figure 3.6 Annual flow series at Altuma (1982-2004)

Any anomalous behavior should be further investigated. Sharp negative peaks may be eliminated from the plot by applying the appropriate time shift between the stations or to carry

out the analysis at a higher aggregation level. The results are shown in Figure 3.7 (a)-(f) for the months of July, Aug, Sept, Oct, Nov-May, respectively.

The residual series, i.e. the difference in flow between Altuma and Anandpur should always have positive values since Anandpur is downstream of Mahanadi and encompasses more catchment area than Altuma. However, the results as seen from the residual series plots show that there are many negative values for all the months; negative values are less frequent for August and September months. The reason for this might be due to one or more of the following: (i) wrong observations at one or both stations, (ii) incorrect estimation of monthly upstream utilization. The gauge-discharge recording at Altuma might be affected by backwater effect as it lies just up stream of pick up weir (PUW). It was also informed that at Altuma, during lean season observation are taken at the site while during the monsoon season, gauging is carried out near the bridge site. Incorrect reporting/estimation of upstream utilization might be another reason for getting negative residual series since the data of upstream utilization is important in computing the virgin flow series. During monsoon there is small or no diversion of flow from the upstream diversions and the magnitude of river flow is quite large. Therefore, the utilization estimates do not materially affect the calculation of the virgin flow series. This might be one of the reasons why the monsoon flows, especially for the August and September months, shows fewer negative values in the residual series. The authenticity of runoff series at Altuma and Anandpur were further checked using the double mass curve, t and F-tests. Both the tests (*t* and *F* tests) are applicable to hydrological time series for testing shift in mean and standard deviation over the years. First the series is divided into two sub-series with approximately half of the data in each of the sub-series. Then the sub-series are treated as samples and the tests are carried out as above. In many cases where there is anomaly between two sets of data, the t-test

may not be able to recognize this but this can be pointed by the test of their variance using F-test.

3.2.2 Double Mass Curves

Double mass curve analysis has been used to check trends or non-homogeneity between flow records at neighboring stations and is normally used with aggregated series. The plots of monthly cumulative runoff between data of two stations, i.e. Kantamal and Kesinga are used as a consistency check of the two sets of time series and the double mass curves are plotted in Figure 3.8 (a-f). This is done for (i) individual monsoon months (July-Sept) and (ii) non-monsoon (Oct-June) periods separately. For consistent data records the curve should plot as a straight line; a bend or crank in the curve means anomalous behavior and should be further investigated. The results of July, September and non-monsoon months are observed to be having certain anomaly. The observations were similar in case of residual plots.

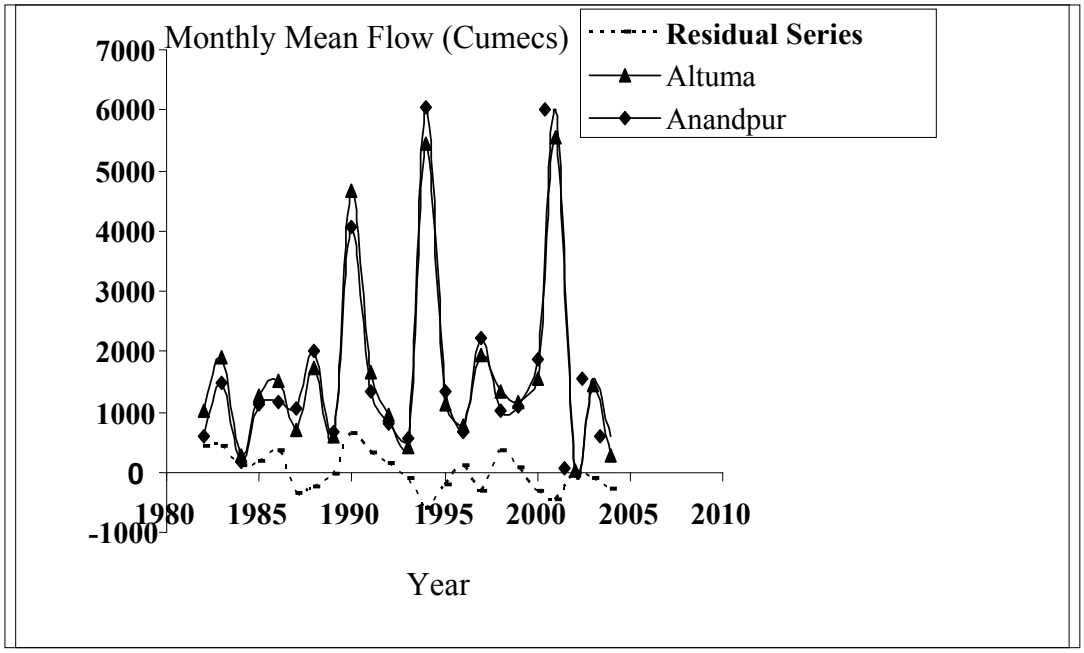

Figure 3.7(a) Residual flow series for the month of July-September (Monsoon) between Altuma and Anandpur (1981-2004).

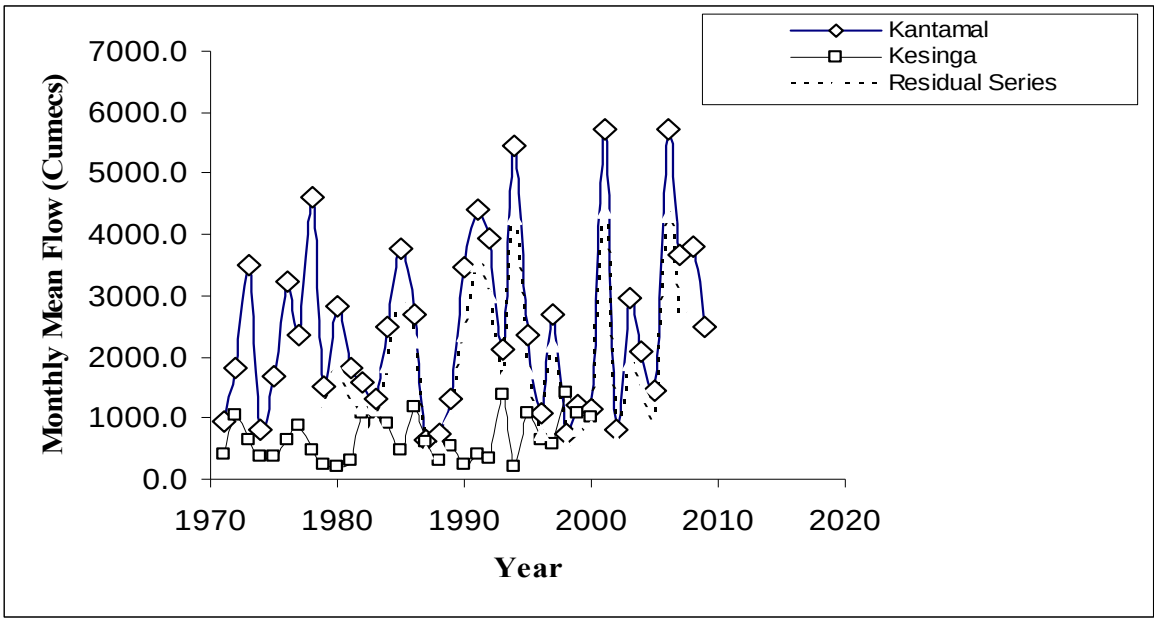

Figure 3.7(b) Residual flow series for the month of July-September (Monsoon) between Kantamal and Jaraikela (1971-2009).

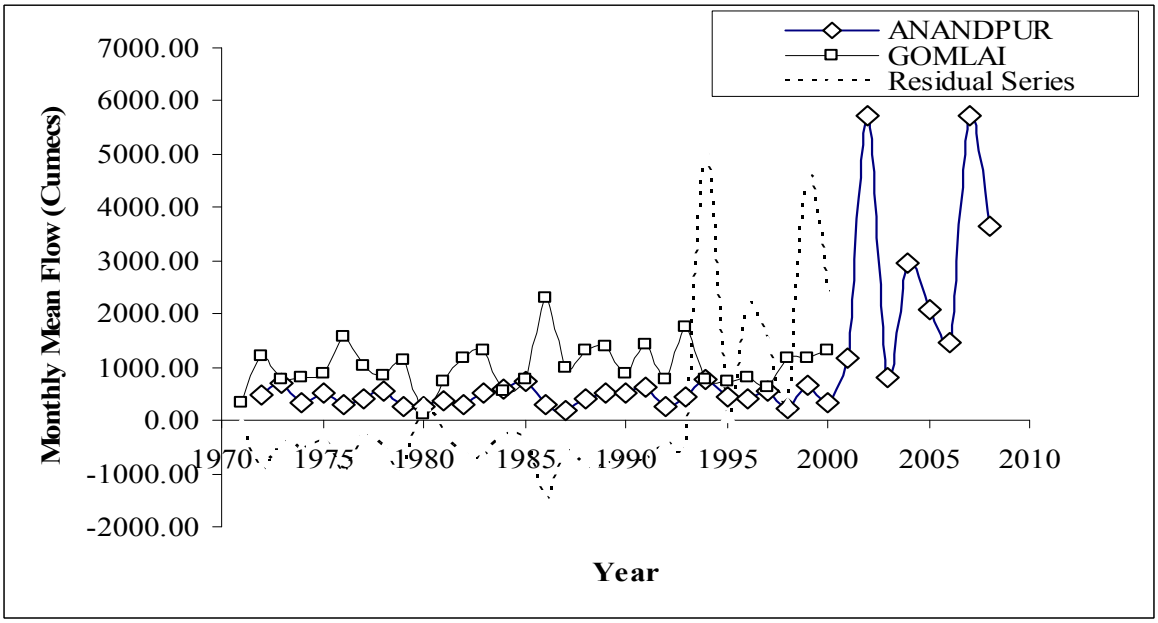

Figure 3.7 (c) Residual flow series for the month of July-September (Monsoon) between for Anandpur and Gomlai

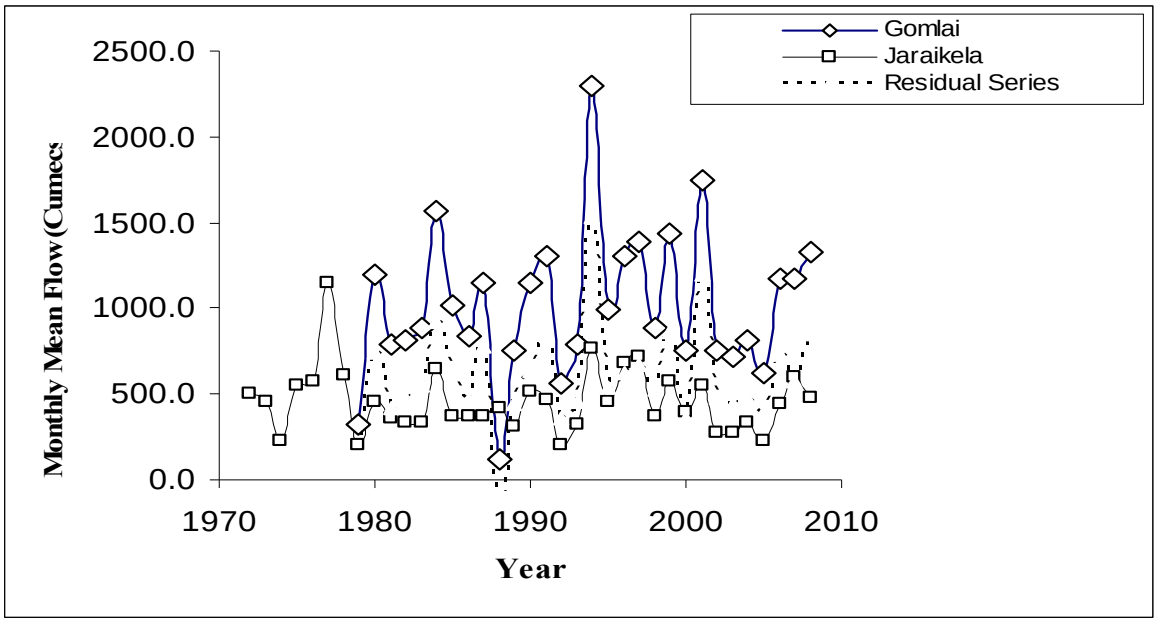

Figure 3.7 (d) Residual flow series for the month of July-September (Monsoon) between for Gomlai **a**nd Jaraikela

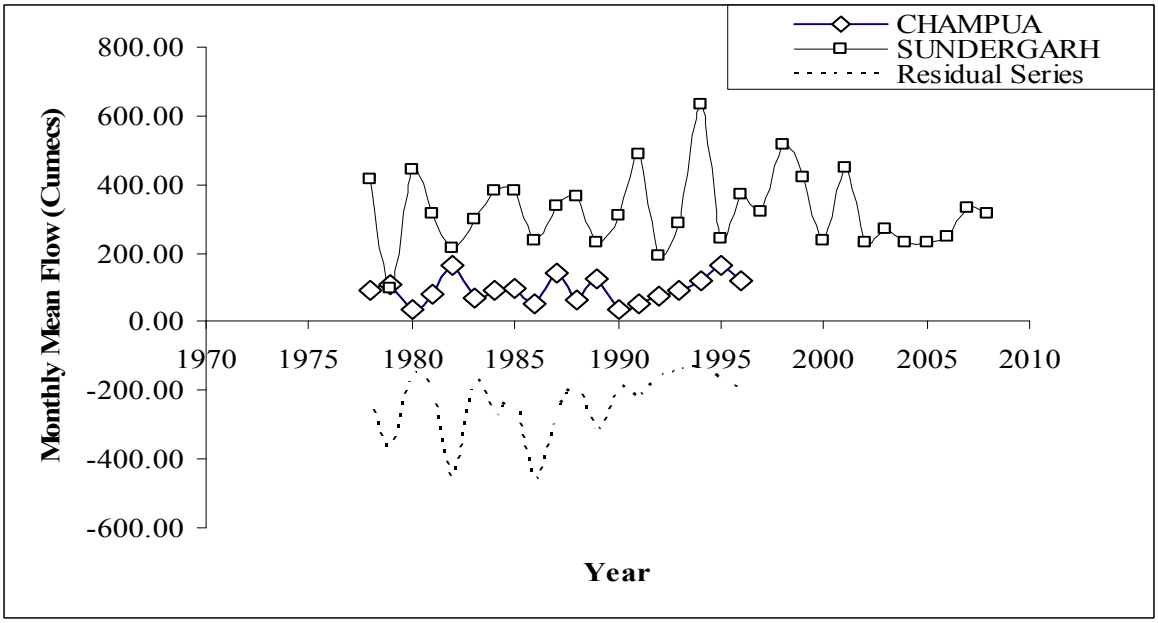

Figure 3.7 (e) Residual flow series for the month of July-September (Monsoon) between Champua and Sundergarh

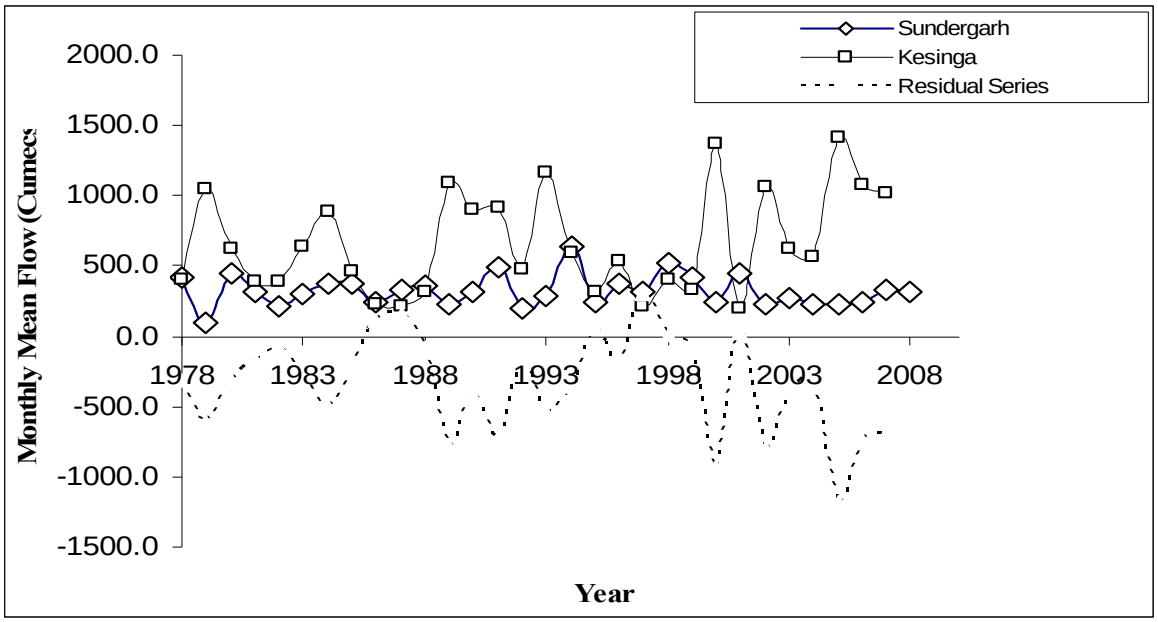

Figure 3.7 (f) Residual flow series for the month of July-September (Monsoon) between Sundergarh and Kesinga

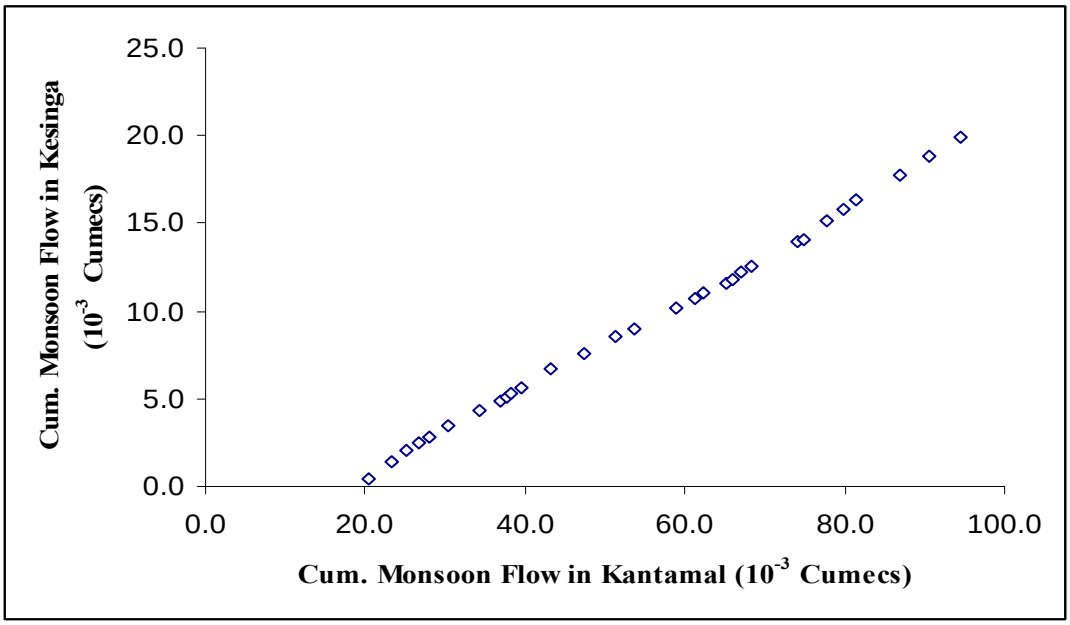

Figure 3.8a Double mass curve for the month of Nov to May using monthly runoff series at Kesinga and Kantamal.

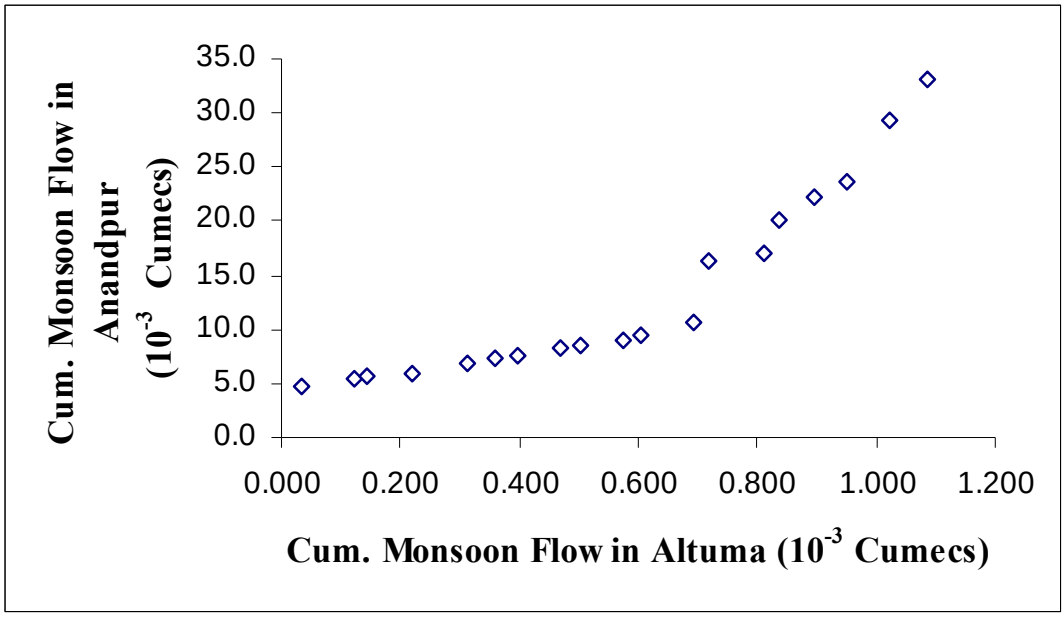

Figure 3.8b Double mass curve for the month of Nov to May using monthly runoff series at Altuma and Anandpur.

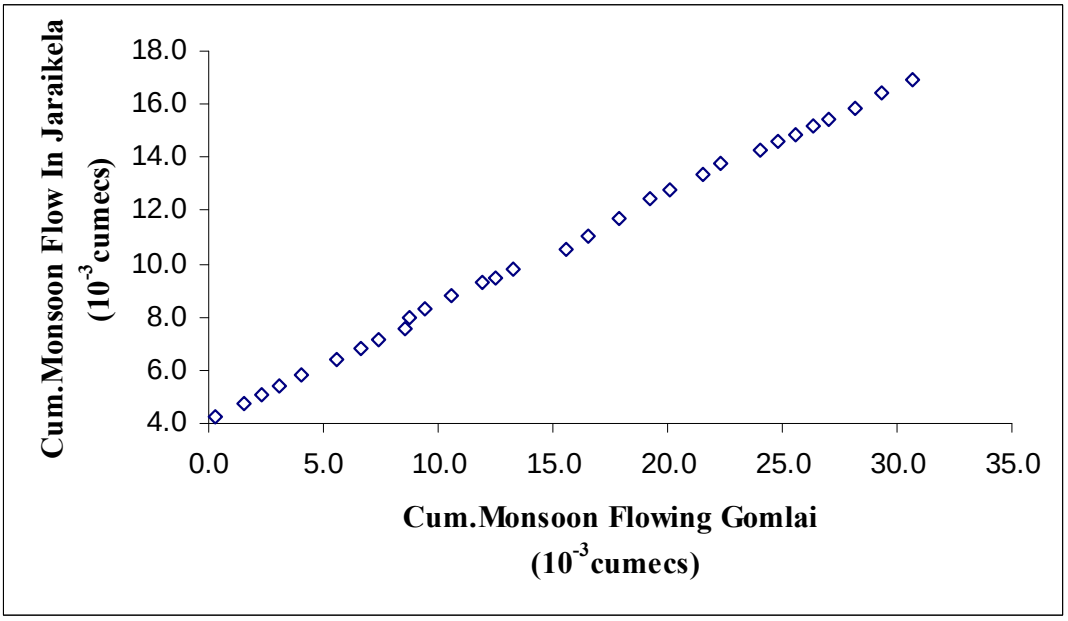

Figure 3.8c Double mass curve for the month of Nov to May using monthly runoff series at Anandpur and Gomlai

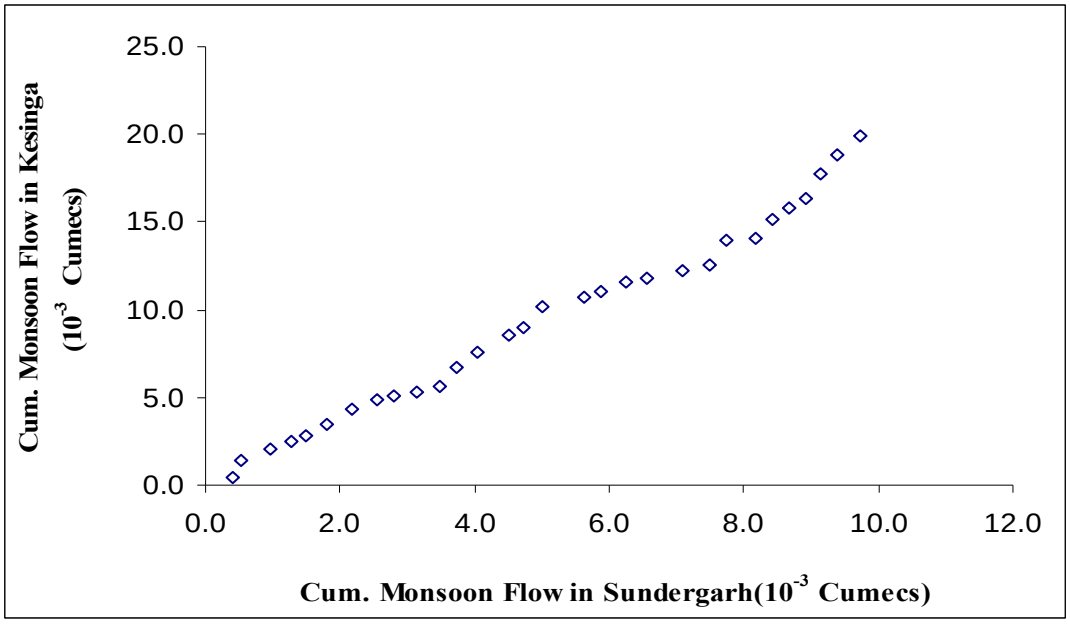

Figure 3.8d Double mass curve for the month of Nov to May using monthly runoff series at Sundergarh and Kesinga

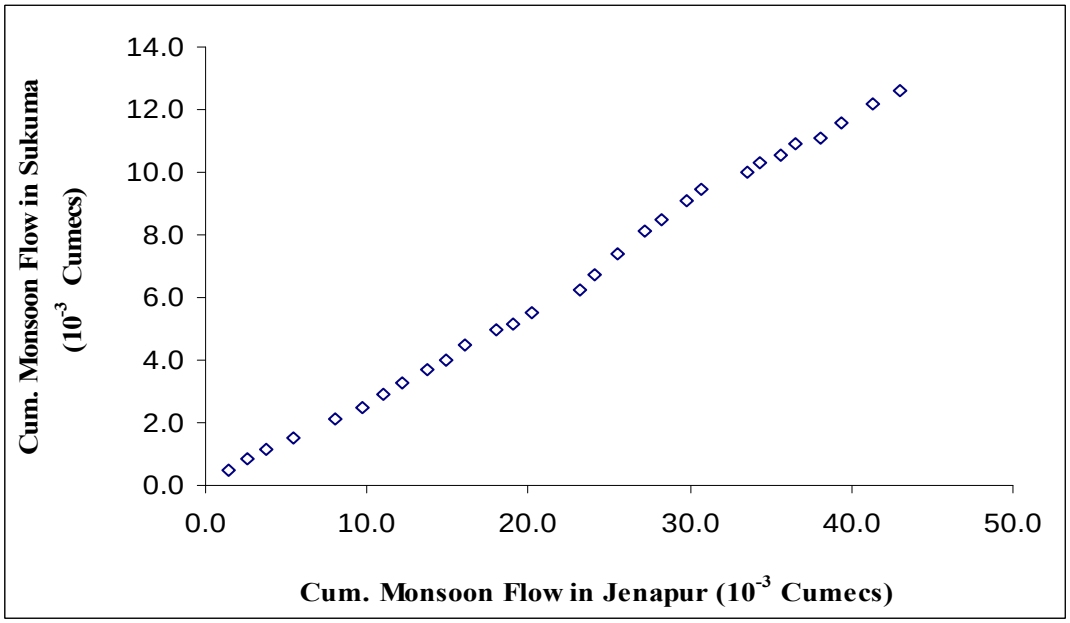

Figure 3.8e Double mass curve for the month of Nov to May using monthly runoff series at Jenapur and Sukma

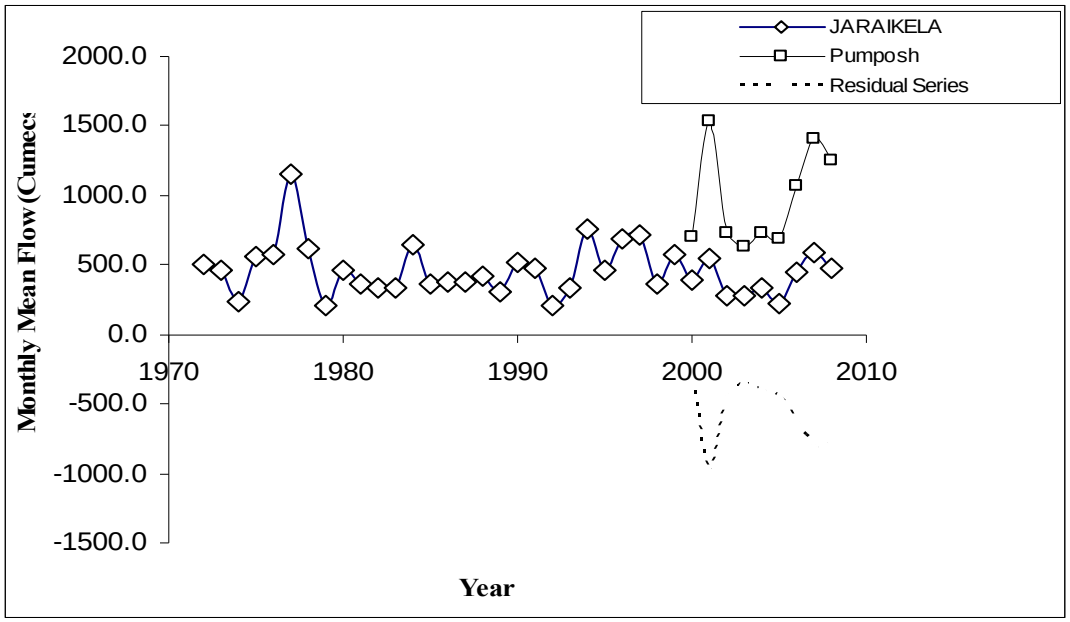

Figure 3.8f Double mass curve for the month of Nov to May using monthly runoff series at Jaraikela and Pamposh

3.2.3 t-Test

The most common parametric test used to check whether or not two samples are from the same population is the *t*- test. The main assumptions of this test are: (i) the observations are independent, (ii) the observations are drawn from normally distributed populations, and (iii) these populations have the same variance. Hence this test is useful to determine whether the mean of the samples are significantly different from each other. Thus, the t-test indicates whether both the series belong to the same population or not. According to this test, the't' statistic of the samples is determined by:

$$t = \frac{\left|\overline{X_1} - \overline{X_2}\right|}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (3.3)$$

where, $\overline{X_1}$ and $\overline{X_2}$ are the arithmetic mean of the two samples of size n_1 and n_2 respectively; S is the unknown population standard deviation estimated from the samples variances s_1 and s_2 as:

$$S = \frac{(n_1 - 1)s_1 + (n_2 - 1)s_2}{n_1 + n_2 - 2} \quad (3.4)$$

If the statistic *t* is less than the tabulated value of Student's distribution at some chosen significance level α and n_1+n_2-1 degrees of freedom then the hypothesis that 'the means of both the samples are not significantly different' may be accepted at the chosen significance level.

The results of t-test for Anandpur and Gomlai and for Altuma and Anandpur are reported in Table 3.7 (a) and 3.7 (b), respectively. For Anandpur, the available series for 1971-2007 was split into two parts: the first for 1971-1990 and the second for 1991-2007. Similarly, the available series for Gomlai is for the period 1972-2008 and in this case, the t-test was carried out by splitting the series in two parts: the first for 1972-1982 and the second for 1982-2008. The results of the given test taking other random pairs are given in Appendix- A. The pairs are as per the following order: . (i) *Altuma* and *Anandpur*, (ii) *Anandpur* and *Gomlai*, (iii) *Gomlai* and *Jaraikela*, (iv) *Jaraikela* and *Pamposh*, (v) *Pamposh* and *Champua*, (vi) *Champua* and *Sundergarh*, (vii) *Sundergarh* and *Kesinga*, (viii) *Kesinga* and *Kantamal*, (ix) *Kantamal* and *Tikerpada*, (x) *Kesinga* and *Kantamal*, (xi) *Kantamal* and *Salebhata*, (xii) *Salebhata* and *Jenapur*, (xiii) *Jenapur* and *Sukma*, (xiv) *Sukma* and *Talcher* , and (xv) *Talcher* and *Tikerpada*.

From the results it is observed that the hypothesis that the mean of both the series belong to the same population is accepted for all the cases. Next, the t-test was conducted by taking the flow series of both the stations to check whether both belong to the same population having the same mean or not. The results are reported in Table 3.6, which shows acceptance of the hypothesis for most cases except for the month of April which consistently fails at all the significant levels of test.

Table 3.7 (a) Consistency Test for Anandpur and Gomlai using t-test

Test of Consistency by t-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	-1.035	1.721	Accepted	2.080	Accepted	2.831	Accepted
Jan	-1.330	1.721	Accepted	2.080	Accepted	2.831	Accepted
Feb	0.847	1.721	Accepted	2.080	Accepted	2.831	Accepted
Mar	0.617	1.721	Accepted	2.080	Accepted	2.831	Accepted
April	0.674	1.721	Accepted	2.080	Accepted	2.831	Accepted
May	0.064	1.721	Accepted	2.080	Accepted	2.831	Accepted
June	0.079	1.721	Accepted	2.080	Accepted	2.831	Accepted
July	-1.180	1.721	Accepted	2.080	Accepted	2.831	Accepted
Aug	-0.345	1.721	Accepted	2.080	Accepted	2.831	Accepted
Sep	-0.477	1.721	Accepted	2.080	Accepted	2.831	Accepted
Oct	-0.681	1.721	Accepted	2.080	Accepted	2.831	Accepted
Nov	-1.127	1.721	Accepted	2.080	Accepted	2.831	Accepted
Dec	-0.960	1.721	Accepted	2.080	Accepted	2.831	Accepted

Table 3.7 (b) Consistency test for Altuma and Anandpur using t-test

Test of Consistency by t-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	0.734	1.681	Accepted	2.017	Accepted	2.695	Accepted
Jan	0.103	1.681	Accepted	2.017	Accepted	2.695	Accepted
Feb	-0.291	1.681	Accepted	2.017	Accepted	2.695	Accepted
Mar	-0.414	1.681	Accepted	2.017	Accepted	2.695	Accepted
April	0.949	1.681	Accepted	2.017	Accepted	2.695	Accepted
May	0.751	1.681	Accepted	2.017	Accepted	2.695	Accepted
June	-0.164	1.681	Accepted	2.017	Accepted	2.695	Accepted
July	0.704	1.681	Accepted	2.017	Accepted	2.695	Accepted
Aug	0.948	1.681	Accepted	2.017	Accepted	2.695	Accepted
Sept	-0.194	1.681	Accepted	2.017	Accepted	2.695	Accepted
Oct	0.269	1.681	Accepted	2.017	Accepted	2.695	Accepted
Nov	0.083	1.681	Accepted	2.017	Accepted	2.695	Accepted
Dec	-0.917	1.681	Accepted	2.017	Accepted	2.695	Accepted

This means that for the month of April, the monthly runoff series at these two sites do not have the same mean as the population mean. These results show that there is some sort of anomaly in the data of Kantamal and Salebhata, Champua and Sundergarh, Sundergarh and Kesinga, Kesinga and Kantamal, or both. This was further checked using the F-test.

Table 3.7 c Consistency test for Altuma and Anandpur using t-test

Test of consistency by t-test							
	Level -->	0.100		0.050		0.010	
	t value	Table t	Result	Table t	Result	Table t	Result
Annual	0.602	1.680	Accepted	2.015	Accepted	2.692	Accepted
Jan	1.217	1.680	Accepted	2.015	Accepted	2.692	Accepted
Feb	1.631	1.680	Rejected	2.015	Accepted	2.692	Accepted
Mar	1.639	1.680	Accepted	2.015	Accepted	2.692	Accepted
April	3.727	1.680	Rejected	2.015	Rejected	2.692	Rejected
May	0.273	1.680	Accepted	2.015	Accepted	2.692	Accepted
June	-0.357	1.680	Accepted	2.015	Accepted	2.692	Accepted
July	-0.054	1.680	Accepted	2.015	Accepted	2.692	Accepted
Aug	0.041	1.680	Accepted	2.015	Accepted	2.692	Accepted
Sept	0.868	1.680	Accepted	2.015	Accepted	2.692	Accepted
Oct	0.628	1.680	Accepted	2.015	Accepted	2.692	Accepted
Nov	0.477	1.680	Accepted	2.015	Accepted	2.692	Accepted
Dec	0.336	1.680	Accepted	2.015	Accepted	2.692	Accepted

Table 3.7 d Consistency test for Anandpur and Gomlai using t-test

Test of Consistency by t-test							
	Level -->	0.100		0.050		0.010	
	t value	Table t	Result	Table t	Result	Table t	Result
Annual	0.654	1.681	Accepted	2.017	Accepted	2.695	Accepted
Jan	0.163	1.681	Accepted	2.017	Accepted	2.695	Accepted
Feb	-0.233	1.681	Accepted	2.017	Accepted	2.695	Accepted
Mar	-0.413	1.681	Accepted	2.017	Accepted	2.695	Accepted
April	0.849	1.681	Accepted	2.017	Accepted	2.695	Accepted
May	0.751	1.681	Accepted	2.017	Accepted	2.695	Accepted
June	-0.164	1.681	Accepted	2.017	Accepted	2.695	Accepted
July	0.604	1.681	Accepted	2.017	Accepted	2.695	Accepted
Aug	0.948	1.681	Accepted	2.017	Accepted	2.695	Accepted
Sept	-0.174	1.681	Accepted	2.017	Accepted	2.695	Accepted
Oct	0.269	1.681	Accepted	2.017	Accepted	2.695	Accepted
Nov	0.073	1.681	Accepted	2.017	Accepted	2.695	Accepted
Dec	-0.817	1.681	Accepted	2.017	Accepted	2.695	Accepted

3.2.4 F-test

F-test is commonly used for testing whether or not the variances of two samples are significantly different. According to this test, the *F* statistic of the samples is determined as:

$$F = \frac{s_1^2}{s_2^2} \tag{3.5}$$

If the computed F is less than the tabulated value of F distribution at some chosen significance level α, and n_1-1 and n_2-1 degrees of freedom then the hypothesis that 'the variances of both the samples are not significantly different' may be accepted at the chosen significance level. Table 3.8(a) shows the result of the F-test for the runoff series at Anandpur and Table 3.8 (b) for Gomlai. From the results, it is observed that the hypothesis whether the variances of both the split series belong to the same population is rejected for most of the cases. This means that there is inconsistency in the monthly flow series at the two gauging sites with respect to the population variance.

Table 3.8 Consistency test for Kesinga and Kantamal using F-test

Test of Consistency by F-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	1.931	1.681	Rejected	2.017	Accepted	2.695	Accepted
Jan	34.683	1.681	Rejected	2.017	Rejected	2.695	Rejected
Feb	50.682	1.681	Rejected	2.017	Rejected	2.695	Rejected
Mar	9.502	1.681	Rejected	2.017	Rejected	2.695	Rejected
April	8.472	1.681	Rejected	2.017	Rejected	2.695	Rejected
May	13.448	1.681	Rejected	2.017	Rejected	2.695	Rejected
June	4.133	1.681	Rejected	2.017	Rejected	2.695	Rejected
July	2.375	1.681	Rejected	2.017	Rejected	2.695	Accepted
Aug	9.830	1.681	Rejected	2.017	Rejected	2.695	Rejected
Sept	6.553	1.681	Rejected	2.017	Rejected	2.695	Rejected
Oct	7.212	1.681	Rejected	2.017	Rejected	2.695	Rejected
Nov	2.982	1.681	Rejected	2.017	Rejected	2.695	Rejected
Dec	27.734	1.681	Rejected	2.017	Rejected	2.695	Rejected

Table 3.8 reports the results of F-test taking the flow series of both the stations to check whether both belong to the same population having the same variance or not. The results show the hypothesis accepted in most cases; all the months shows acceptance of the hypothesis at 1 % significance level. This infers that the flow series at both sites are affected by the same type of error either the systematic errors in observation or errors in estimation of water utilization.

Table 3.8a Consistency Test for Anandpur by Using F-test

Test of Consistency by F-test							
	Level -->	**0.10**		**0.050**		**0.010**	
	F value	**Table F**	**Result**	**Table F**	**Result**	**Table F**	**Result**
Annual	2.517	2.248	Rejected	2.854	Accepted	4.539	Accepted
Jan	1.592	2.248	Accepted	2.854	Accepted	4.539	Accepted
Feb	1.442	2.248	Accepted	2.854	Accepted	4.539	Accepted
Mar	2.550	2.302	Rejected	2.943	Accepted	4.772	Accepted
April	6.741	2.302	Rejected	2.943	Rejected	4.772	Rejected
May	1.386	2.248	Accepted	2.854	Accepted	4.539	Accepted
June	3.038	2.302	Rejected	2.943	Rejected	4.772	Accepted
July	4.173	2.248	Rejected	2.854	Rejected	4.539	Accepted
Aug	1.634	2.248	Accepted	2.854	Accepted	4.539	Accepted
Sept	2.832	2.248	Rejected	2.854	Accepted	4.539	Accepted
Oct	3.487	2.248	Rejected	2.854	Rejected	4.539	Accepted
Nov	7.830	2.248	Rejected	2.854	Rejected	4.539	Rejected
Dec	89.070	2.248	Rejected	2.854	Rejected	4.539	Rejected

Table 3.8 b Consistency test for Gomlai using F-test

Test of Consistency by F-Test							
	Level -->	**0.10**		**0.050**		**0.010**	
	F value	**Table F**	**Result**	**Table F**	**Result**	**Table F**	**Result**
Annual	1.867	1.761	Rejected	2.073	Accepted	2.837	Accepted
Jan	1.189	1.751	Accepted	2.059	Accepted	2.805	Accepted
Feb	2.201	1.751	Rejected	2.059	Rejected	2.805	Accepted
Mar	1.402	1.751	Accepted	2.059	Accepted	2.805	Accepted
April	2.695	1.761	Rejected	2.073	Rejected	2.837	Accepted
May	1.821	1.751	Rejected	2.059	Accepted	2.805	Accepted
June	1.832	1.751	Rejected	2.059	Accepted	2.805	Accepted
July	2.364	1.761	Rejected	2.073	Rejected	2.837	Accepted
Aug	1.825	1.761	Rejected	2.073	Accepted	2.837	Accepted
Sept	1.576	1.751	Accepted	2.059	Accepted	2.805	Accepted
Oct	1.253	1.751	Accepted	2.059	Accepted	2.805	Accepted
Nov	3.407	1.751	Rejected	2.059	Rejected	2.805	Rejected
Dec	22.746	1.751	Rejected	2.059	Rejected	2.805	Rejected

Table 3.9 Consistency test for Anandpur and Gomlai using F-test

Test of Consistency by F-Test							
	Level -->	0.10		0.050		0.010	
	F value	Table F	Result	Table F	Result	Table F	Result
Annual	1.08	1.744	Accepted	2.048	Accepted	2.785	Accepted
Jan	2.062	1.744	Rejected	2.048	Rejected	2.785	Accepted
Feb	1.035	1.744	Accepted	2.048	Accepted	2.785	Accepted
Mar	1.026	1.744	Accepted	2.048	Accepted	2.785	Accepted
April	1.419	1.744	Accepted	2.048	Accepted	2.785	Accepted
May	1.308	1.744	Accepted	2.048	Accepted	2.785	Accepted
June	1.425	1.744	Accepted	2.048	Accepted	2.785	Accepted
July	1.183	1.744	Accepted	2.048	Accepted	2.785	Accepted
Aug	1.753	1.744	Rejected	2.048	Accepted	2.785	Accepted
Sept	1.245	1.744	Accepted	2.048	Accepted	2.785	Accepted
Oct	1.521	1.744	Accepted	2.048	Accepted	2.785	Accepted
Nov	1.332	1.744	Accepted	2.048	Accepted	2.785	Accepted
Dec	1.674	1.744	Accepted	2.048	Accepted	2.785	Accepted

3.2.5 Time Series Analysis

The time series analysis is done here so as to estimate the time series parameters of the flow series. These might be useful if the series is to be extended with stochastic model. In case a stochastic model is planned to be used to simulate monthly flows using the periodicity observed in the available monthly discharge data, certain pre-requisite analysis is mandatory. These are discussed briefly in the following section.

The annual yield series at a gauging site can be treated as a stochastic time series and can be analyzed using time models such as autoregressive (AR), moving average (MA) and autoregressive moving average (ARseries mMA) models. For time series modeling of the annual yield series the basic statistics of the series like mean, variance, autocorrelation and partial autocorrelation are analyzed for identification of a suitable time series model for the series. These statistics are estimated using the following formulae.

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$s^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2$$

$$c_k = \frac{1}{N}\sum_{i=i}^{N-k}\left(x_i - \bar{x}\right)\left(x_{i+k} - \bar{x}\right), \qquad 0 \le k \le N$$

$$r_k = \frac{c_k}{c_0} \tag{3.6}$$

where, $\bar{x}$ is the mean, s^2 is the variance, c_k is the auto-covariance at lag 'k', and r_k is the autocorrelation at lag 'k'. The partial autocorrelations are estimated using Yule-Walker equations. Using the above equations the characteristics of the annual yield series at Kesinga and Kantamal were analyzed. The partial autocorrelation at different lags for annual yield of Kesinga and Kantamal are plotted in Figs. 3.9, respectively.

For the Kesinga data, the partial autocorrelation was significant even at higher lags which indicates that an AR model of higher order may be required to model the series. However, the same trend was not observed for the yield series at Kantamal for which the partial auto-correlations of lags higher than lag 1 are quite small. This indicates that an AR (1) model may be applicable for the Kantamal series. Four AR models with 1, 2, 3, and 4 parameters were applied for the both series and the model parameters were estimated from the series and diagnostic tests were carried out (Tables 3.10 (a) and 3.10(b). The autocorrelation of historical data is compared with the model autocorrelation (Figs. 3.12 (a) and (b). It is evident from the Tables and the figures that an AR (2) model fits well for the yield series at Kesinga whereas the AR (1) model approximately reproduces the historical autocorrelation for the Kantamal series

Table 3.10 (a) Performance of AR models for Kesinga

Model	AR(1)	AR(2)	AR(3)	AR(4)
Parameters	-0.318	-0.219	-0.25198	-0.23548
		0.29218	0.31622	0.35425
			0.10837	0.07974
				-0.10007
Noise variance	0.89785	0.83899	0.84835	0.86014
PM lack of fit test for residuals	Accepted	Accepted	Accepted	Accepted
AIC	-2.7411	-3.5488	-0.9074	1.8232

Table 3.10 (b) Performance of AR models for Kantamal

Model	AR(1)	AR(2)	AR(3)	AR(4)
Parameters	-0.411	-0.36009	-0.35598	-0.35271
		0.12059	0.10923	0.09573
			-0.03204	0.00835
				0.10543
Noise variance	0.81649	0.88329	0.88192	0.91674
PM lack of fit test for residuals	Accepted	Accepted	Accepted	Rejected
AIC	-2.4602	0.3938	3.5179	6.3482

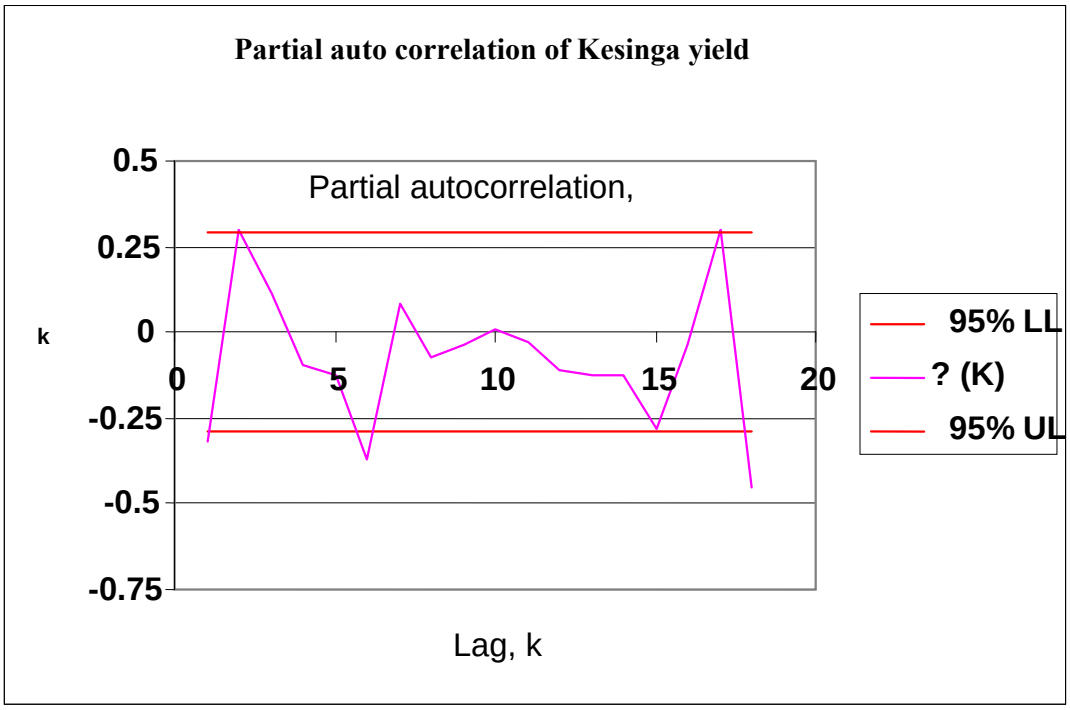

Figure 3.9 Partial auto correlation for annual yield series of Kesinga

Table 3.11 Consistency test for Altuma and Anandpur using T-test

Test of Consistency by T-test							
	Level -->	0.100		0.050		0.010	
	t value	Table t	Result	Table t	Result	Table t	Result
Annual	-9.721	1.681	Accepted	2.017	Accepted	2.695	Accepted
Jan	-4.114	1.681	Accepted	2.017	Accepted	2.695	Accepted
Feb	-7.676	1.681	Accepted	2.017	Accepted	2.695	Accepted
Mar	-14.494	1.681	Accepted	2.017	Accepted	2.695	Accepted
April	-18.869	1.681	Accepted	2.017	Accepted	2.695	Accepted
May	-13.670	1.681	Accepted	2.017	Accepted	2.695	Accepted
June	-1.792	1.681	Accepted	2.017	Accepted	2.695	Accepted
July	-9.086	1.681	Accepted	2.017	Accepted	2.695	Accepted
Aug	-4.718	1.681	Accepted	2.017	Accepted	2.695	Accepted
Sept	-15.275	1.681	Accepted	2.017	Accepted	2.695	Accepted
Oct	-17.294	1.681	Accepted	2.017	Accepted	2.695	Accepted
Nov	-12.723	1.681	Accepted	2.017	Accepted	2.695	Accepted
Dec	-10.398	1.681	Accepted	2.017	Accepted	2.695	Accepted

Table 3.12 Consistency test for Kesinga and Kantamal using T-test

Test of Consistency by t-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	-3.872	1.681	Accepted	2.017	Accepted	2.695	Accepted
Jan	-2.527	1.681	Accepted	2.017	Accepted	2.695	Accepted
Feb	1.757	1.681	Rejected	2.017	Accepted	2.695	Accepted
Mar	0.688	1.681	Accepted	2.017	Accepted	2.695	Accepted
April	3.790	1.681	Rejected	2.017	Rejected	2.695	Rejected
May	5.360	1.681	Rejected	2.017	Rejected	2.695	Rejected
June	-2.599	1.681	Accepted	2.017	Accepted	2.695	Accepted
July	-3.300	1.681	Accepted	2.017	Accepted	2.695	Accepted
Aug	-2.500	1.681	Accepted	2.017	Accepted	2.695	Accepted
Sept	-5.394	1.681	Accepted	2.017	Accepted	2.695	Accepted
Oct	-4.755	1.681	Accepted	2.017	Accepted	2.695	Accepted
Nov	-15.832	1.681	Accepted	2.017	Accepted	2.695	Accepted
Dec	-7.406	1.681	Accepted	2.017	Accepted	2.695	Accepted

Table 3.13 Consistency test for Pamposh and Champua using t-test

Test of Consistency by t-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	**3.664**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Jan	**3.385**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Feb	**8.653**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Mar	**14.872**	1.681	Rejected	2.017	Rejected	2.695	Rejected
April	**14.321**	1.681	Rejected	2.017	Rejected	2.695	Rejected
May	**16.065**	1.681	Rejected	2.017	Rejected	2.695	Rejected
June	**2.431**	1.681	Rejected	2.017	Rejected	2.695	Accepted
July	**8.002**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Aug	**6.289**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Sept	**11.512**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Oct	**18.485**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Nov	**8.101**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Dec	**7.592**	1.681	Rejected	2.017	Rejected	2.695	Rejected

Table 3.14 Consistency test for Gomlai and Jaraikela using t-test

Test of Consistency by t-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	1.931	1.681	Rejected	2.017	Accepted	2.695	Accepted
Jan	34.683	1.681	Rejected	2.017	Rejected	2.695	Rejected
Feb	50.682	1.681	Rejected	2.017	Rejected	2.695	Rejected
Mar	9.502	1.681	Rejected	2.017	Rejected	2.695	Rejected
April	8.472	1.681	Rejected	2.017	Rejected	2.695	Rejected
May	13.448	1.681	Rejected	2.017	Rejected	2.695	Rejected
June	4.133	1.681	Rejected	2.017	Rejected	2.695	Rejected
July	2.375	1.681	Rejected	2.017	Rejected	2.695	Accepted
Aug	9.830	1.681	Rejected	2.017	Rejected	2.695	Rejected
Sept	6.553	1.681	Rejected	2.017	Rejected	2.695	Rejected
Oct	7.212	1.681	Rejected	2.017	Rejected	2.695	Rejected
Nov	2.982	1.681	Rejected	2.017	Rejected	2.695	Rejected
Dec	27.734	1.681	Rejected	2.017	Rejected	2.695	Rejected

Table 3.15 Consistency test for Kantamal and Salebhatta using t-test

Test of Consistency by t-test							
	Level -->	**0.100**		**0.050**		**0.010**	
	t value	**Table t**	**Result**	**Table t**	**Result**	**Table t**	**Result**
Annual	**7.070**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Jan	**10.059**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Feb	**13.902**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Mar	**17.118**	1.681	Rejected	2.017	Rejected	2.695	Rejected
April	**16.803**	1.681	Rejected	2.017	Rejected	2.695	Rejected
May	**7.448**	1.681	Rejected	2.017	Rejected	2.695	Rejected
June	**2.983**	1.681	Rejected	2.017	Rejected	2.695	Rejected
July	**3.270**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Aug	**4.289**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Sept	**10.086**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Oct	**11.532**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Nov	**28.900**	1.681	Rejected	2.017	Rejected	2.695	Rejected
Dec	**47.432**	1.681	Rejected	2.017	Rejected	2.695	Rejected

T-distribution

It is similar to the normal distribution but its shape depends upon the size of sample. As the number of observations in the sample increases, the t-distribution approaches the normal distribution. An important concept in statistical tests pertains to degree of freedom. Actually, in statistical tests, the same sample is used to estimate the parameters of the distribution as well as perform the test. Thus, there are multiple uses of observations. This limitation is tackled by using the concept of degrees of freedom. *Degree of freedom* is the number of observations in a sample less the number of parameters being estimated. For example, if we have 20 data which are used to estimate two parameters, the degree of freedom will be 18. The tables of t-distribution list the value of t statistics corresponding to various levels of significance and the degrees of freedom

Most F-tests arise by considering a decomposition of the variability in a collection of data in terms of sums of squares. The test statistic in an F-test is the ratio of two scaled sums of squares reflecting different sources of variability. These sums of squares are constructed so that

Table 3.16 Performance of AR models for Altuma and Anandpur

Model	**AR(1)**	**AR(2)**	**AR(3)**	**AR(4)**
Parameters	0.894	0.652	0.639	0.646
		0.278	0.246	0.285
			0.050	0.155
				-0.169
Noise variance	0.89785	0.83899	0.84835	0.86014
PM lack of fit test for residuals	Accepted	Accepted	Accepted	Accepted
AIC	-2.7411	-3.5488	-0.9074	1.8232

Table 3.17 Performance of AR models for Kantamal and Salebhatta

Model	AR(1)	AR(2)	AR(3)	AR(4)
Parameters	-0.318	-0.219	-0.25198	-0.23548
		0.29218	0.31622	0.35425
			0.10837	0.07974
				-0.10007
Noise variance	0.89785	0.83899	0.84835	0.86014
PM lack of fit test for residuals	Accepted	Accepted	Accepted	Accepted
AIC	-2.7411	-3.5488	-0.9074	1.8232

Table 3.18 Performance of AR models for Pamposh and Champua

Model	AR(1)	AR(2)	AR(3)	AR(4)
Parameters	-0.318	-0.219	-0.25198	-0.23548
		0.29218	0.31622	0.35425
			0.10837	0.07974
				-0.10007
Noise variance	0.89785	0.83899	0.84835	0.86014
PM lack of fit test for residuals	Accepted	Accepted	Accepted	Accepted
AIC	-2.7411	-3.5488	-0.9074	1.8232

the statistic tends to be greater when the null hypothesis is not true. In order for the statistic to follow the F-distribution under the null hypothesis, the sums of squares should be statistically independent, and each should follow a scaled chi-squared distribution. The latter condition is guaranteed if the data values are independent and normally distributed with a common variance

Table 3.19 Performance of AR models for Gomlai and Jaraikela

Model	**AR(1)**	**AR(2)**	**AR(3)**	**AR(4)**
Parameters	0.847	1.082	1.126	1.105
		-0.283	-0.449	-0.390
			0.159	0.002
				0.157
Noise variance	0.89785	0.83899	0.84835	0.86014
PM lack of fit test for residuals	Accepted	Accepted	Accepted	Accepted
AIC	-2.7411	-3.5488	-0.9074	1.8232

The F Test

Equality of variances are determined by employing a distribution known as F distribution (here F stands for Sir Ronald Fisher, an eminent statistician who developed the analysis of variance). This crucial ratio of the between-group to the within-group variance estimate is called F ratio:

$$F = s_1^2 / s_2^2$$

where s is standard deviation.

Obviously if the number of observations is small, the sample variances will vary more from trial to trial. Therefore, the shape of the *F* distribution would be expected to change with changes in sample size. This brings us back to the idea of degrees of freedom *v.* But in this situation, the *F* distribution is dependent upon two values of *v,* one associated with each variance in the ratio. Also, the distribution cannot be negative because it is the ratio of two positive numbers. If the sample size is large, the average of the ratios should be close to 1.0.

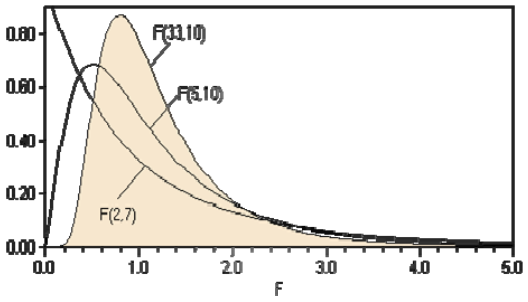

Figure 3.10 Plot of F-distribution for selected values of parameters.

The shape of F-distribution would change with sample size. Recall the idea of degrees of freedom. The distribution can't be negative – it is the ratio of two positive numbers. For any pair of variances, two ratios can be computed *(s_1/s_2* and *s_2/s_1)*. The larger variance is to be placed in numerator so that the ratio will always be greater than 1.0. Only one-tailed tests need be utilized and the alternative hypothesis actually is a statement that the absolute difference between the two sample variances is greater than expected if the population variances are equal. If sample size is large, average of ratios is close to 1.0. We may hypothesize that two samples are drawn from populations having equal variances. After computing F ratio, we can then can ascertain probability of obtaining, by chance, that specific value from two samples from one normal population. If it is unlikely that such a ratio could be obtained, this is an indication that samples came from different populations having different variances.

Because the *F* distribution describes the probabilities of obtaining specified ratio of sample variances drawn from the same population, it can be used to test the equality of variances which we obtain in statistical sampling. We may hypothesize that two samples are drawn from populations that have equal variances. After computing the *F* ratio, we can then can ascertain the

probability of obtaining, by chance, that specific value from two samples from one normal population. If it is unlikely that such a ratio could be obtained, this is an indication that the samples came from different populations having different variances.

For any pair of variances, two ratios can be computed *(s_1/s_2* and *s_2/s_1)*. We may decide that the larger variance will be placed in the numerator, so that the ratio will always be greater than 1.0 and the statistical tests can be simplified. Only one-tailed tests need be utilized and the alternative hypothesis actually is a statement that the absolute difference between the two sample variances is greater than expected if the population variances are equal

Example: Consider a comparison between the two sample sets of discharge measurements (21 observations at each site) on two gauging sites on a stream. We are interested in determining if the variation in discharge is the same at the two sites. Let us adopt a level of significance of 5%. Thus, we are willing to face the risk of concluding that the discharges are different when actually they are the same one time out of every twenty trials.

The variances of the two samples may be computed using the observed data. Then, the *F* ratio between the two may be calculated by eq. (given above). Note that s_1 is the larger variance and s_2 is the smaller. Our null hypothesis is

$$H_0 : \sigma_1^2 = \sigma_2^2$$

against the alternative hypothesis

$$H_0 : \sigma_1^2 \neq \sigma_2^2$$

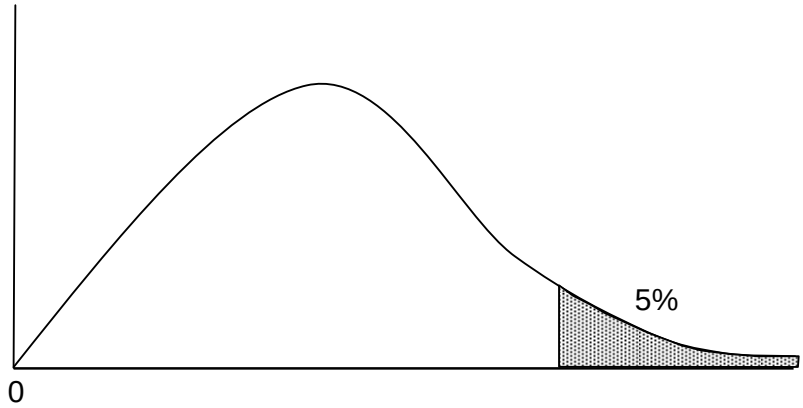

Figure 3.11 A typical *F* distribution with v_1 = 15 and v_2 = 25 degrees of freedom, with critical region (shown by shading) that contains 5% of the area under the curve. Critical value of *F* = 2.09.

The null hypothesis states that the parent populations of the two samples have equal variances; the alternative hypothesis states that they are not equal. Degrees of freedom associated with this test are (n_1 - 1) for ν_1 and (n_2 - 1) for v_2. The critical value of F with v_1 = 20 and v_2 = 20 degrees of freedom and a level of significance of 5% (α = 0.05) can be found from Table 2.14a. This value is 2.12.The value of *F* calculated from (2.36) will fall into one of two areas. If the calculated value of *F* exceeds 2.12, the null hypothesis is rejected and we conclude that the variation in discharge is not the same in the two groups. If the calculated value is less than 2.12, we would have no evidence to conclude that the variances are different.

In most practical situations, we do not have knowledge of the true parameters of the population and these are estimated from the data. While comparing two samples, it is necessary to verify whether their variances are equivalent statistically. If they are and the samples have been selected without any bias, one can proceed to conduct other statistical tests.

As an example, consider the following problem: snow and ice collected from permanently frozen parts of the Earth contain small quantities of micron-sized bits of dust called micro particles. Individual grains range in size from about 0.5 to 3.0 μ; these have been injected into the atmosphere by many agents, including volcanic eruptions, dust storms, and micrometeorite rainfall. The particles are so small they would remain suspended indefinitely, but they are scrubbed out by snow because they serve as nuclei for tiny ice crystals. The ice in turn is incorporated into the permanent snowfields of Polar Regions. It has been postulated that the concentration of micro particles in snow should be uniform over the Earth, because of mixing of the atmosphere and the manner by which micro particles are removed from the air. The theory, if true, has significance for the advisability of atmospheric testing of nuclear weapons, so two suites of snow samples have been collected carefully from the Greenland ice cap and from Antarctica. Under controlled conditions, the snow has been melted and the quantity of contained micro particles determined by an electron particle classifier.

Assuming the samples have been collected without bias and the distribution of microparticles is normal throughout the snowfields, the first step is to test the equality of variances in the two sample sets. The hypothesis and alternative are

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_0 : \sigma_1^2 \neq \sigma_2^2$$

If the variances are not significantly different, the next step in the procedure is to test equality of means. The appropriate test is the t-test. For obvious reasons, the level of significance attached to this test cannot be higher than the significance attached to the test of equality of variances. The appropriate hypothesis and alternative are

$$H_0 : \mu_1 = \mu_2$$

$H_0 : \mu_1 \neq \mu_2$

because there is no reason to suppose that one region should have a larger mean than the other. If the variances and means cannot be distinguished (that is, the null hypotheses cannot be rejected), there is no statistical evidence to suggest that

F distribution is not symmetrical but is positively skewed. This is true of F distributions in general. The reason for the positive skew is that an F distribution is a distribution of ratios of variances. Variances are always positive numbers and a ratio of positive numbers can never be less than 0. Of course, there is a rare chance that the ratio is a very high number. Thus the F ratio cannot be lower than 0 and can be quite high, but most F ratios tend to fall just on the positive side of 0.

Also notice that we use the F distribution as a comparison distribution for determining whether the between-group estimate of the population variance is sufficiently bigger (never smaller) than the within-group estimate of the population variance. That is, we are always interested in the right-hand tail of the distribution, and all F tests in the analysis of variance are technically one-tailed tests. However, the meaning of a one-tailed test is not really the same as with a t test because in the analysis of variance, it does not make sense to talk about a direction of a difference, only of whether the variation among the means is greater than would be expected if the null hypothesis were true.

CHAPTER 4
USE OF DEFINED RATING CURVE METHODS

4.0 INTRODUCTION

A *rating table or curve* is a relationship between stage (S) and discharge (Q) at a cross section of a river. In most cases, data from stream gages are collected as stage data. In order to model the streams and rivers, the data needs to be expressed as stream flow using rating tables. Conversely, the output from a hydrologic model is a flow, which can then be expressed as stage for dissemination to the public. The development of rating curve involves two steps. In the first step the relationship between stage and discharge is established by measuring the stage and corresponding discharge in the river. And in the second part, stage of river is measured and discharge is calculated by using the relationship established in the first part. For such cases, a series of streamflow measurements using a current meter are plotted versus the accompanying stage and a smooth curve is drawn through the points. The measured value of discharges when plotted against the corresponding stages show a definite relationship between the two and it represents the integrated effect of a wide range of channel and flow parameters. The combined effect of these parameters is termed control. If the S-Q relationship for a gauging section is constant and does not change with time, the control is said to be permanent. If it changes with time, it is known as shifting control.

There can be significant scatter around this curve. Because of this, when using a rating curve, it is good to keep in mind that the discharge read from the curve is the most likely value, but it could be a little different from the measured value. Also, since rating curves are developed with few stage/discharge measurements, and measurements of high flows are rare, there can be significant errors in rating curves at high levels, especially around record level flows. Rating curves usually have a break point, which is around the stage at which the river spreads out of its

banks or it could be at a lower stage if the river bed cross section changes dramatically. Above that stage, the river does not rise as fast, given that other conditions remain constant.

4.1 STAGE-DISCHARGE CONTROL

Continuous records of discharge at gauging stations are computed by applying the discharge rating for the stream to records of stage. Discharge ratings may be simple or complex, depending on the number of variables needed to define the stage-discharge relation. This chapter is concerned with ratings in which the discharge can be related to stage alone. The terms *rating, rating curve, stage rating, and stage-discharge relation* are synonymous and are used here interchangeably. Discharge ratings for gauging stations are usually determined empirically by means of periodic measurements of discharge and stage. The discharge measurements are usually made by current meter. Measured discharge is then plotted against concurrent stage on graph paper to define the rating curve. At a new station many discharge measurements are needed to define the stage-discharge relation throughout the entire range of stage.

The relation of stage to discharge is usually controlled by a section or reach of channel downstream from the gauge that is known as the station control. A section control may be natural or man-made; it may be a ledge of rock across the channel, a boulder-covered riffle, an overflow dam, or any other physical feature capable of maintaining a fairly stable relation between stage and discharge. Section controls are often effective only at low discharges and are completely submerged by channel control at medium and high discharges. Channel control consists of all the physical features of the channel that determine the stage of the river at a given point for a given rate of flow. These features include the size, slope, roughness, alignment, constrictions and expansions, and shape of the channel. The reach of channel that acts as the

control may lengthen as the discharge increases, introducing new features that affect the stage-discharge relation. Knowledge of the channel features that control the stage-discharge.

4.2 TYPES OF STATION CONTROLS

The development of rating curve involves two steps. In the first step the relationship between stage and discharge is established by measuring the stage and corresponding discharge in the river. And in the second part, stage of river is measured and discharge is calculated by using the relationship established in the first part. Stage is measured by reading gauge installed in the river. If the stage discharge relationship doesn't change with time then it is called *permanent control* and if this relationship changes it is called *shifting control.* In a nutshell, the character of rating curve depends on type of control, governed by:

a) Geometry of the cross-section

b) Physical features of the river d/s

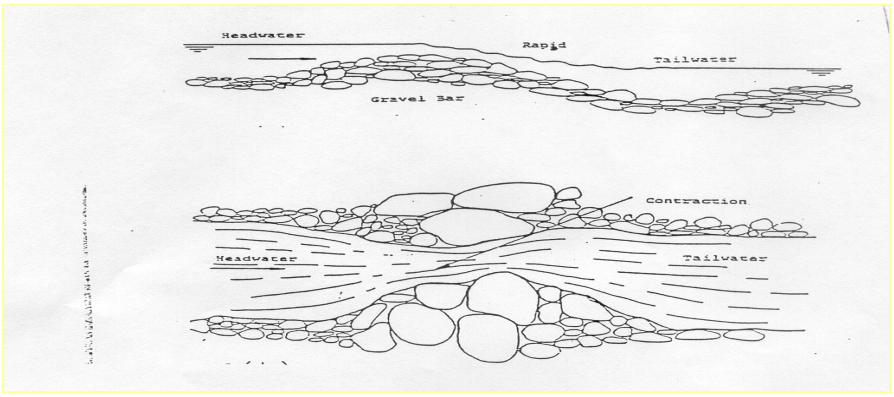

Figure 4.1a Control configuration in natural channel

and station controls are classified in the following ways:

a) Section and channel controls

b) Natural and artificial controls

c) Complete, compound and partial controls

d) Permanent and shifting controls

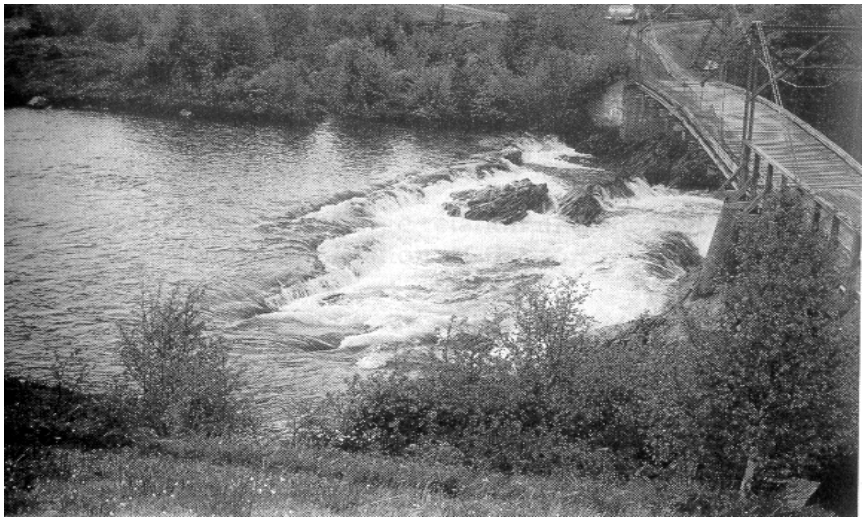

Figure 4.1 b Section control

Figure 4.1c Partial channel control

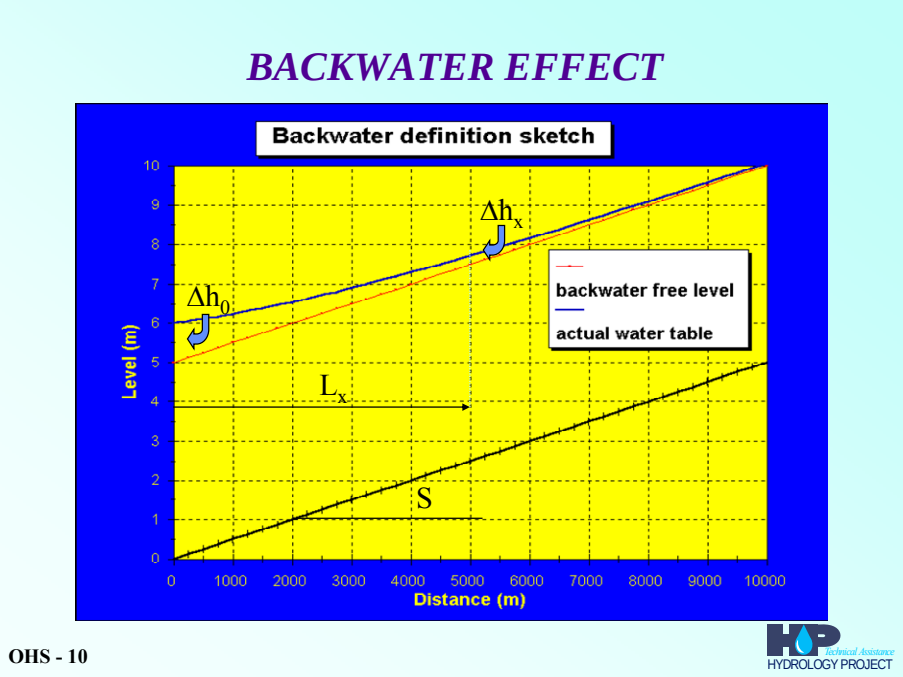

Figure 4.2 Backwater effect

4.3 EXTENT OF CHANNEL CONTROL

Variable backwater given in Figure 4.2, causes variable energy slope for the same stage. Hence, discharge is a function of both stage and of slope. The slope-stage-discharge relation is generally is the energy slope approximated by water level slope. The first order approximation of backwater effect in a rectangular channel can be given as follows:

at x = 0: $h_0 = he + \Delta h_0$

at x = Lx: $h_x = h_e + \Delta h_x$

Backwater: $\Delta h_x = \Delta h_0 . \exp [(-3.S.L_x)/(h_e(1-Fr^2)]$

Where;

Froude: $Fr^2 = u^2/(gh)$ often << 1

Manning: $Q = K_m B h_e^{5/3} S^{1/2}$

So with q = Q/B: $he = \{q/(K_m S^{1/2})\}^{3/5}$ and $\ln(\Delta h_x/\Delta h_0) = -3.S.L_x/h_e$

At: $\Delta h_x/\Delta h_0 = 0.05$: $L_x = h_e/S$ (4.1)

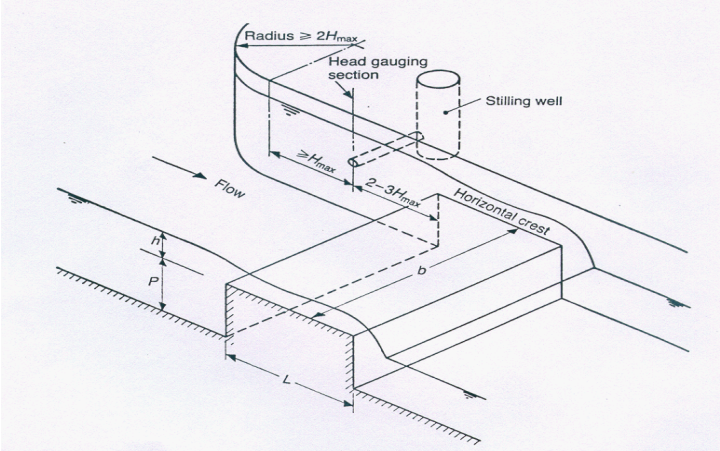

Figure 4.3 Artificial Control to accommodate this, generally are used artificial controls given in the above Figure

4.4 METHODS OF DEVELOPING RATING CURVES

4.4.1 General

A majority of streams and rivers, especially non-alluvial rivers exhibit permanent control. For such a case, the relationship between the stage and the discharge is a single-valued relation which is expressed as:

$$Q = a\ (H - H0)_b \qquad (4.2)$$

in which; Q = stream discharge, H = gauge height (stage), H_o = a constant which represent the gauge reading corresponding to zero discharge, a and b are rating curve constants. This relationship can be expressed graphically by plotting the observed stage against the corresponding discharge values in arithmetic or logarithmic plot. A typical rating curve is shown in Figure 4.4 (a) and (b). Logarithmic plotting is advantageous as gauge-discharge relationship forms a straight line in logarithmic coordinates. The advantage of using the double logarithmic plot is two fold: (i) firstly, the plot would produce a straight line since the general form of rating

curve is parabolic, and (ii) secondly, different straight lines allow to further grouping of data. This is shown in an example in Figure 4.4. A part of the entire range of stage may form a straight line. It gives an indication about the stage at which the slope of the straight line changes if more than one lines are used to fit the data points.

While plotting the data on double log plot a prior knowledge about the value of H_o is necessary. As a first approximation the value of H_o is assumed to be the level of the bottom of the channel as determined from the cross section of the gauging station. Marginal adjustment in the values of H_o may be required in order to produce a straight line giving better fit to the plotted points. There is a possibility that more than one straight lines are fitted if so required to represent the changing conditions at different stages.

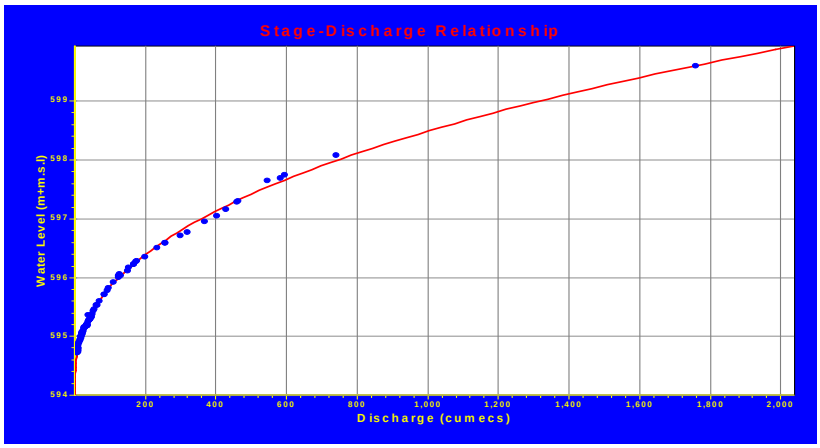

Figure 4.4a A typical rating curve

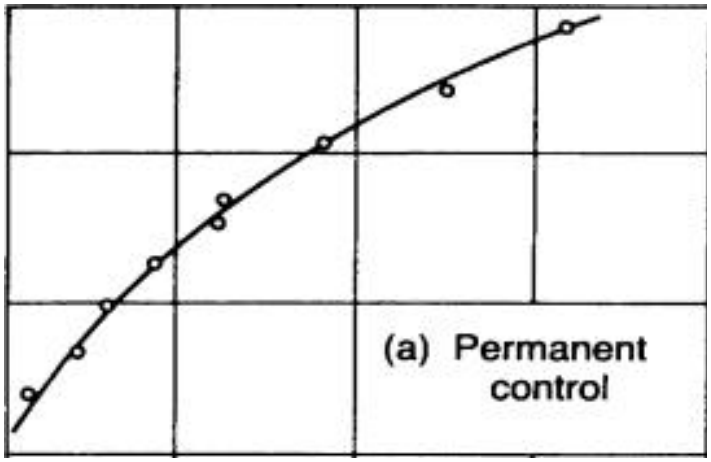

Figure 4.4b A typical rating curve in a permanent control

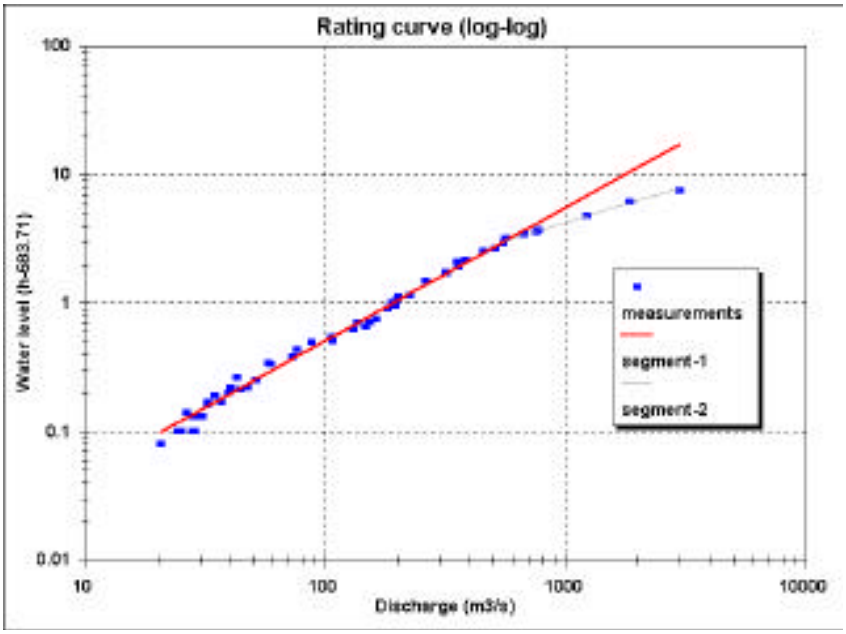

Figure 4.5 Double logarithmic plot of rating curve showing a distinct break

4.4.2 Establishing the Stage Discharge Rating Curve

If Q and h are discharge and water level, then the relationship can be analytically expressed as:

$$Q = f(h) \quad (4.3)$$

Where; *f(h)* is an algebraic function of water level. A graphical stage discharge curve helps in visualizing the relationship and to transform stages manually to discharges whereas an algebraic relationship can be advantageously used for analytical transformation.

A simple stage discharge relation is one where discharge depends upon stage only. A complex rating curve occurs where additional variables such as the slope of the energy line or the rate of change of stage with respect to time are required to define the relationship. The need for a particular type of rating curve can be ascertained by first plotting the observed stage and discharge data on a simple orthogonal plot. The scatter in the plot gives a fairly good assessment of the type of stage-discharge relationship required for the cross section. Examples of the scatter plots obtained for various conditions are illustrated below. If there is negligible scatter in the plotted points and it is possible to draw a smooth single valued curve through the plotted points then a simple rating curve is required. This is shown in Figure 4.4 (a, b).

However, if scatter is not negligible then it requires further probing to determine the cause of such higher scatter. There are four distinct possibilities:

a) *Backwater Effects:* If the station is affected by the variable backwater conditions arising due for example to tidal influences or to high flows in a tributary joining downstream, then , in such cases, if the plotted points are annotated with the corresponding slope of energy line (surface slope for uniform flows) then a definite pattern can be observed. A smooth curve passing through those points having normal slopes at various depths is drawn first. It can then be seen that the points with greater variation in slopes from the corresponding normal slopes are located farther from the curve. This is as shown in Figure 4.6 (a) and (b).

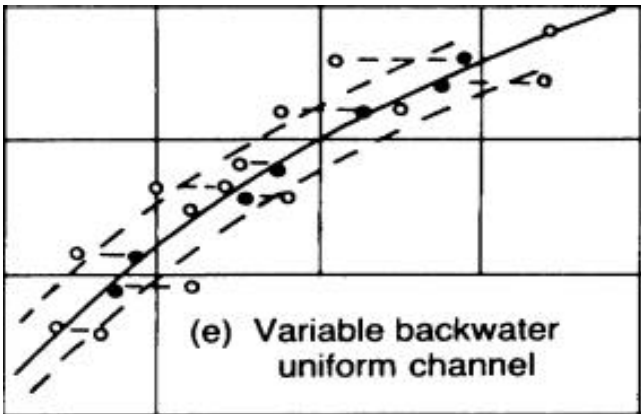

Figure 4.6a Rating curve affected by variable backwater (uniform channel)

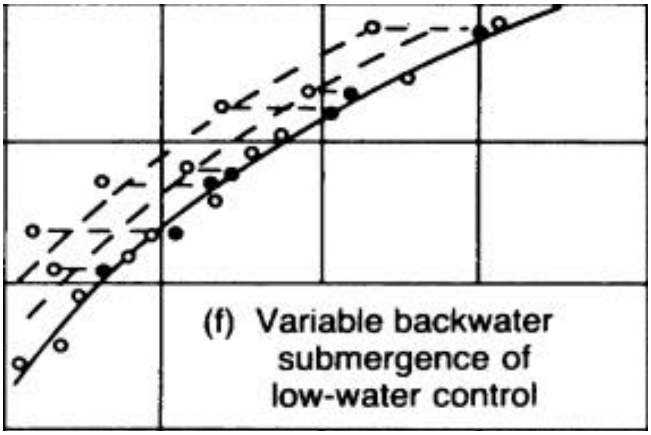

Figure 4.6b Rating curve affected by variable backwater (submergence of low water control)

b) *Local acceleration effect:* If the stage discharge rating is affected by the variation in the local acceleration due to unsteady flow, in such case, the plotted points can be annotated with the corresponding rate of change of slope with respect to time. A smooth curve (steady state curve) passing through those points having the least values of rate of change of stage is drawn first. It can then be seen that all those points having positive values of rate of change of stage are towards the right side of the curve and those with negative values are towards the left of it. Also, the distance from the steady curve

increases with the increase in the magnitude of the rate of change of stage. This is as shown in Figure 4.6 (c).

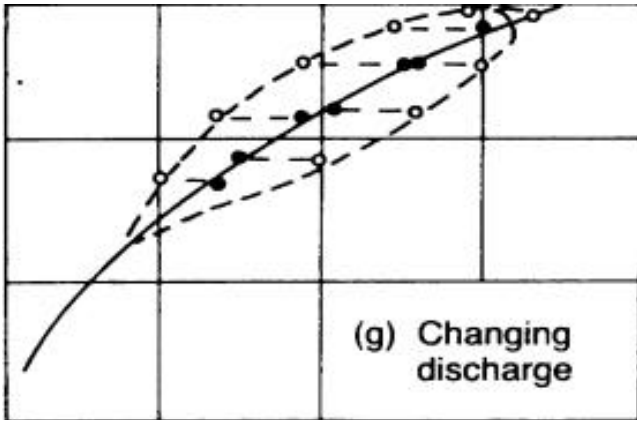

Figure 4.6c Rating curve affected by unsteady flow

c) *Scouring effect:* The stage discharge rating is affected by scouring of the bed or changes in vegetation characteristics. A shifting bed results in a wide scatter of points on the graph. The changes are erratic and may be progressive or may fluctuate from scour in one event and deposition in another. Examples are shown in Figure 4.6 (d). If no suitable explanation can be given for the amount of scatter present in the plot, then it can perhaps be attributed to the observational errors. Such errors can occur due to non-standard procedures for stage discharge observations.

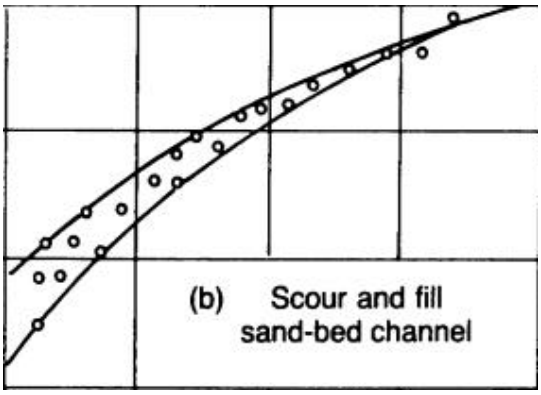

Figure 4.6d Stage-discharge relations affected by scour and fill

4.5 TYPE OF RATING CURVE

Based on the interpretation of scatter of the stage discharge data, the appropriate type of rating curve is fitted, and there are four main cases as described as follows:

a) *Simple rating curve*: If simple stage discharge rating is warranted then either single channel or compound channel rating curve is fitted according to whether the flow occur essentially in the main channel or also extends to the flood plains.

b) *Rating curve with backwater corrections*: If the stage discharge data is affected by the backwater effect then the rating curve incorporating the backwater effects is to be established. This requires additional information on the fall of stage with respect to an auxiliary stage gauging station.

c) *Rating curve with unsteady flow correction:* If the flows are affected by the unsteadiness in the flow then the rating curve incorporating the unsteady flow effects is established. This requires information on the rate of change of stage with respect to time corresponding to each stage discharge data.

d) *Rating curve with shift adjustment:* A rating curve with shift adjustment is warranted in case the flows are affected by scouring and variable vegetation effects.

4.5.1 Fitting of Single Channel Simple Rating Curve

Single channel simple rating curve is fitted in those circumstances when the flow is contained the main channel section and can be assumed to be fairly steady. There is no indication of any backwater affecting the relationship. The bed of the river also does not significantly change so as create any shifts in the stage discharge relationship. The scatter plot of the stage and discharge data shows a very little scatter if the observational errors are not

significant. The scatter plot of stage discharge data in such situations typically is as shown in Figure 4.4(b). The fitting of simple rating curves can conveniently be considered under the following headings:

a) Equations used and their physical basis,

b) Determination of datum correction(s)

c) Number and range of rating curve segments

d) Determination of rating curve coefficients

e) Estimation of uncertainty in the stage discharge relationship

The procedures and step by step approach is discussed in the following section.

4.5.2 Least Square Method

The best values of a and b for a given range of stage can be obtained by the least-square-error method. Thus by taking logarithms, Eq. (4.2) may be represented as

$$\log Q = \log a + b \log (H - H_0) \tag{4.4}$$

or $Y = m X + c$

in which the dependent variable Y = log Q, independent variable X = log ($H-H_o$) and c = log a. The values of the coefficients for the best-fit straight line using data of N observations of X and Y is:

$$m = \frac{N\sum_{i=1}^{N}(X_i Y_i) - (\sum_{i=1}^{N} X_i)(\sum_{i=1}^{N} Y_i)}{N\sum_{i=1}^{N}(X_i)^2 - (\sum_{i=1}^{N} X_i)^2}$$

$$c = \frac{\sum_{i=1}^{N} Y_i - m\sum_{i=1}^{N} X_i}{N} \tag{4.5}$$

The Eqn. (4.4) is known as the rating equation of the stream and can be used for estimating the discharge Q of the stream for a given gauge reading H within the range of data used in its derivation. The constant H_0 representing the stage (gauge height) for zero discharge in the stream is a hypothetical parameter and can not be measured in the field. As such, its determination poses some difficulties. Different alternative methods are available for its determination. However generally it is found by extrapolating the rating curve by eye judgement to find H_0 as the value of H corresponding to Q=0. Using the value of H_o, plot log Q vs. log (H - H_o) and verify whether the data plots as a straight line. If not, select another value in the neighbourhood of previously assumed value and by trial and error find an acceptable value of H_o which gives a straight line plot of log Q vs. log (H-H_o). Rating curve is established by concurrent measurements of stage (H) and discharge (Q) covering expected range of river stages at section over a period of time. If Q-H rating curve not unique, then additional information required on (i) Slope of water level (backwater), (ii) Hydrograph in unsteady flow condition.

Example 4.1

Develop a rating curve for the river using the available stage-discharge data during the period January 1997 to December 1997. The rating curve using the power type of equation given in Eq. (4.2) may be used for this case.

Table 4.1 Data given in the example

No.	Water level (M)	Monthly mean Q (m^3/s)		Difference	Relative difference error (%)
		Observed	Computed		
1	594.800	9.530	8.541	0.989	10.38
2	595.370	36.480	46.661	-10.181	-27.91
3	596.060	127.820	136.679	-8.859	-6.93
4	596.510	231.400	226.659	4.741	2.05
5	598.080	738.850	783.019	-44.169	-5.98
6	597.700	583.340	610.35	-27.019	-4.63

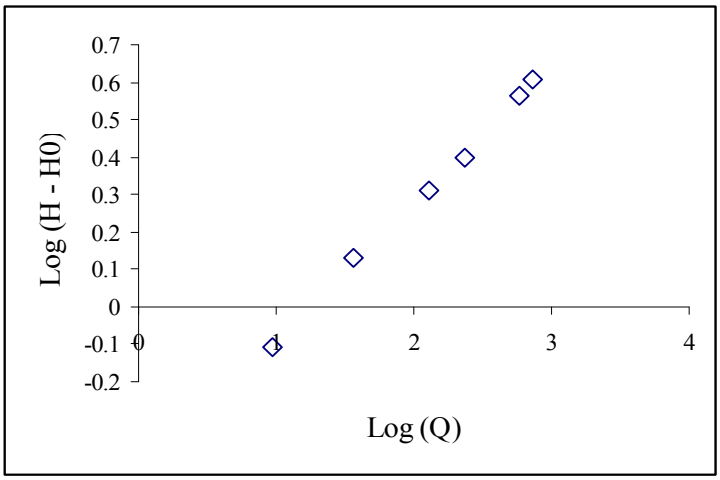

Figure 4.7 A typical rating curve

a = -594.025, b = 2.531, c = 22.63

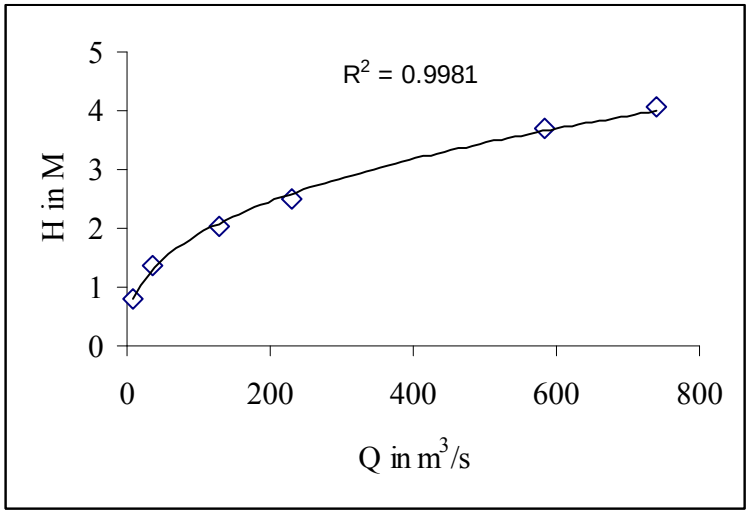

Figure 4.8 A typical rating curve

A simple way of judging the results with criteria of goodness-of-fit using the following

Relative difference error (%) = [Observed - Computed]/ Observed x 100

This is given in the last column of table. A minus sign indicates the over-estimated value and positive the under-estimated value i.e. the estimated or computed monthly discharge value.

4.5.3 Fitting of Single Channel Simple Rating Curve

Single channel simple rating curve is fitted in those circumstances when the flow is contained the main channel section and can be assumed to be fairly steady. There is no indication of any backwater affecting the relationship. The bed of the river also does not significantly change so as create any shifts in the stage discharge relationship. The scatter plot of the stage and discharge data shows a very little scatter if the observational errors are not significant. The scatter plot of stage discharge data in such situations typically is as shown in Figure 4.4(b) and 4.5. The fitting of simple rating curves can conveniently be considered under the following headings:

a) equations used and their physical basis,

b) determination of datum correction(s)

c) number and range of rating curve segments

d) determination of rating curve coefficients

e) estimation of uncertainty in the stage discharge relationship

4.5.4 Equations Used and Their Physical Basis

Two types of algebraic equations are commonly fitted to stage discharge data are:

(1) P*ower type equation* which is most commonly used,

$$Q = c\,(h - a)^b \qquad (4.6)$$

(2) *Parabolic type of equation*

$$Q = c_2\,(h_w + a)^2 + c_1\,(h_w + a) + c_0 \qquad (4.7\ a)$$

Where: Q = discharge (m^3/sec)

h = measured water level (m)

a = water level (m) corresponding to Q = 0

c_i = coefficients derived for the relationship corresponding to the station characteristics

It is anticipated that the power type equation is most frequently used in India and is recommended. Taking logarithms of the power type equation results in a straight line relationship of the form:

$$\log (Q) = \log (c) + b \log (h + a) \qquad (4.7\ b)$$

or $Y = A + B X$

That is, if sets of discharge (Q) and the effective stage (h + a) are plotted on the double log scale, they will represent a straight line. Coefficients A and B of the straight line fit are functions of a and b. Since values of a and b can vary at different depths owing to changes in physical characteristics (effective roughness and geometry) at different depths, one or more straight lines will fit the data on double log plot. This is illustrated in Figure 4.8, which shows a distinct break in the nature of fit in two water level ranges. A plot of the cross section at the gauging section is also often helpful to interpret the changes in the characteristics at different levels.

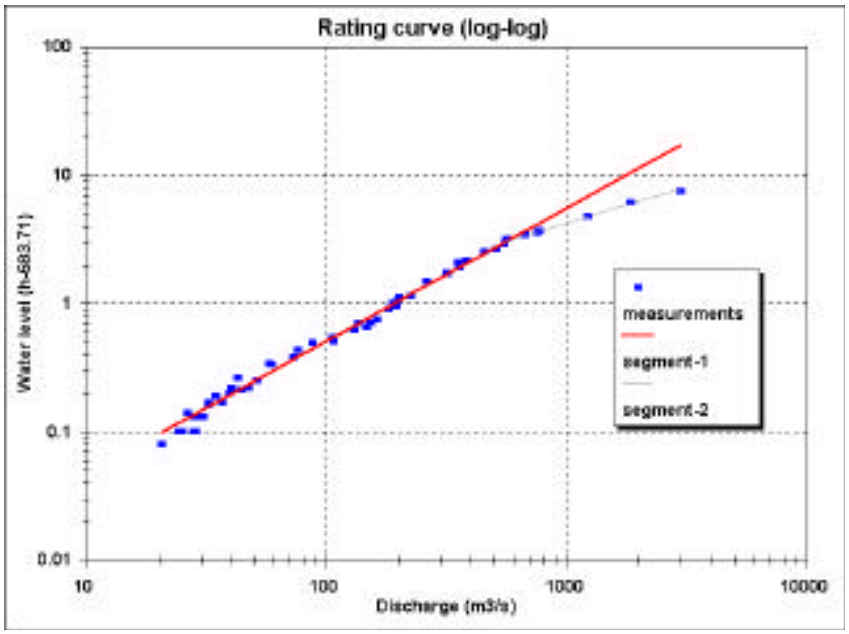

Figure 4.9 Double logarithmic plot of rating curve showing a distinct break

The relationship between rating curve parameters and physical conditions is also evident if the power and parabolic equations are compared with Manning's equation for determining discharges in steady flow situations. The Manning's equation can be given as:

$$Q = (1/n)A(R)^{2/3}\sqrt{S} \qquad (4.8)$$

Q = function of (roughness and slope) (depth and geometry)

Hence, the coefficients a, c and d are some measures of roughness and geometry of the control and b is a measure of the geometry of the section at various depths. Changes in the channel resistance and slope with stage, however, will affect the exponent b. The net result of these factors is that the exponent for relatively wide rivers with channel control will vary from about 1.3 to 1.8. For relatively deep narrow rivers with section control, the exponent will commonly be greater than 2 and sometimes exceed a value of 3. Note that for compound channels with flow over the floodplain or braided channels over a limited range of level, very high values of the exponent are sometimes found (>5).

4.5.5 Determination of Datum Correction (a)

The datum correction (a) corresponds to that value of water level for which the flow is zero. From Eq. (4.6) it can be seen that for Q = 0, (h + a) = 0 which means: a = -h.

Physically, this level corresponds to the zero flow condition at the control effective at the measuring section. The exact location of the effective control is easily determined for artificial controls or where the control is well defined by a rock ledge forming a section control. For the

channel controlled gauging station, the level of deepest point opposite the gauge may give a reasonable indication of datum correction. In some cases identification of the datum correction may be impractical especially where the control is compound and channel control shifts progressively downstream at higher flows. Note that the datum correction may change between different controls and different segments of the rating curve. For upper segments the datum correction is effectively the level of zero flow had that control applied down to zero flow; it is thus a nominal value and not physically ascertainable.

Alternative analytical methods of assessing "a" are therefore commonly used and methods for estimating the datum correction are as follows:

a) trial and error procedure
b) arithmetic procedure
c) computer-based optimization

However, where possible, the estimates should be verified during field visits and inspection of longitudinal and cross sectional profiles at the measuring section.

Trial and error procedure

This was the method most commonly used before the advent of computer-based methods. The stage discharge observations are plotted on double log plot and a median line fitted through them. This fitted line usually is a curved line. However, as explained above, if the stages are adjusted for zero flow condition, i.e. datum correction a, then this line should be a straight line. This is achieved by taking a trial value of "a" and plotting (h + a), the adjusted stage, and discharge data

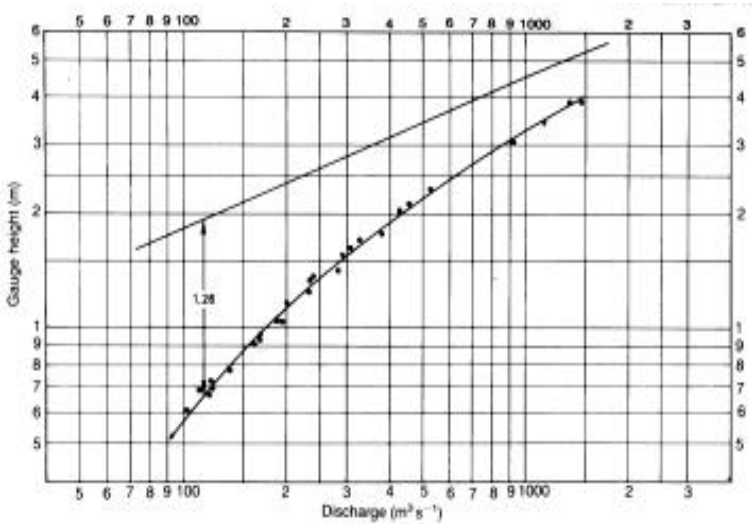

Figure 4.10 Determination of datum correction (a) by trial and error

on the same double log plot. It can be seen that if the unadjusted stage discharge plot is concave downwards then a positive trial value of "a" is needed to make it a straight line. And conversely, a negative trial value is needed to make the line straight if the curve is concave upwards. A few values of "a" can be tried to attain a straight line fit for the plotted points of adjusted stage discharge data. The procedure is illustrated in Figure 4.10. This procedure was slow but quite effective when done earlier manually. However, making use of general spreadsheet software (having graphical provision) for such trial and error procedure can be very convenient and faster now.

Arithmetic procedure

This procedure is based on expressing the datum correction "a" in terms of observed water levels. This is possible by way elimination of coefficients b and c from the power type equation between gauge and discharge using simple mathematical manipulation. From the median curve fitting the stage discharge observations, two points are selected in the lower and

upper range (Q_1 and Q_3) whereas the third point Q_2 is computed from $Q_2^{\ 2} = (Q_1) \times (Q_3)$, such that:

$$\frac{Q_1}{Q_2} = \frac{Q_2}{Q_3} \tag{4.9}$$

If the corresponding gauge heights for these discharges read from the plot are h_1, h_2 and h_3 then using the power type, we obtain:

$$\frac{c(h_1 + a)}{c(h_2 + a)} = \frac{c(h_2 + a)}{c(h_3 + a)} \tag{4.10}$$

Which yields:

$$a = \frac{h_2^{\ 2} - h_1 h_3}{h_1 + h_3 - 2h_2} \tag{4.11}$$

From this equation an estimated value of "a" can be obtained directly. This procedure is known as Johnson method which is described in the WMO Operational Hydrology manual on stream gauging **(Report No. 13, 1980).**

Optimization procedure

This procedure is suitable for automatic data processing using computer and "a" is obtained by optimization. The first trial value of the datum correction "a" is either input by the user based on the field survey or from the computerized Johnson method described above. Next, this first estimate of "a" is varied within 2 m so as to obtain a minimum mean square error in the fit. This is a purely mathematical procedure and probably gives the best results on the basis of observed stage discharge data but it is important to make sure that the result is confirmed where

possible by physical explanation of the control at the gauging location. The procedure is repeated for each segment of the rating curve.

4.5.6 Number and Ranges of Rating Curve Segments

After the datum correction "a" has been established, the next step is to determine if the rating curve is composed of one or more segments. This is normally selected by the user rather than done automatically by computer. It is done by plotting the adjusted stage, (h-a) or simply "h" where there are multiple segments, and discharge data on the double log scale. This scatter plot can be drawn manually or by computer and the plot is inspected for breaking points. Since for (h-a), on double log scale the plotted points will align as straight lines, breaks are readily identified. The value of "h" at the breaking points give the first estimate of the water levels at which changes in the nature of the rating curve are expected. The number and water level ranges for which different rating curves are to be established is thus noted.

4.5.7 Determination of Rating Curve Coefficients

A least square method is normally employed for estimating the rating curve coefficients. For example, for the power type equation, taking a and b as the estimates of the constants of the straight line fitted to the scatter of points in double log scale, the estimated value of the logarithm of the discharge can be obtained as:

$$Y = \alpha + \beta X \tag{4.12}$$

The least square method minimizes the sum of square of deviations between the logarithms of measured discharges and the estimated discharges obtained from the fitted rating curve. Considering the sum of square the error as E, we can write:

$$E = \sum_{i=1}^{N} (Y_i - Y_i)^2 = \sum_{i=1}^{N} (Y_i - \alpha - \beta X_i)^2 \quad (4.13)$$

Here i denote the individual observed point and N is the total number of observed stage discharge data. Since this error is to be minimum, the slope of partial derivatives of this error with respect to

the constants must be zero. In other words:

$$\frac{\partial E}{\partial \alpha} = \frac{\partial \{\sum_{i=1}^{N} (Y_i - \alpha - \beta X_i)^2}{\partial \alpha} = 0 \quad (4.14)$$

and

$$\frac{\partial E}{\partial \beta} = \frac{\partial \{\sum_{i=1}^{N} (Y_i - \alpha - \beta X_i)^2}{\partial \beta} = 0 \quad (4.15)$$

This results in two algebraic equations of the form:

$$\sum_{i=1}^{N} Y_i - \alpha N - \beta \sum_{i=1}^{N} X_i = 0 \quad (4.16)$$

and

$$\sum_{i=1}^{N} (X_i Y_i) - \alpha \sum_{i=1}^{N} X_i - \beta \sum_{i=1}^{N} (X_i)^2 = 0 \quad (4.17)$$

All the quantities in the above equations are known except a and b. Solving the two equations yield:

$$\beta = \frac{N \sum_{i=1}^{N} (X_i Y_i) - (\sum_{i=1}^{N} X_i)(\sum_{i=1}^{N} (Y_i)}{N \sum_{i=1}^{N} (X_i)^2 - (\sum_{i=1}^{N} X_i)^2} = 0 \quad (4.18)$$

and

$$\alpha = \frac{\sum_{i=1}^{N} Y_i - \beta \sum_{i=1}^{N} X_i}{N} = 0 \qquad (4.19)$$

$$b = \beta \qquad \& \qquad c = 10^{\alpha} \qquad (4.20)$$

4.6 COMPOUND CHANNEL RATING CURVE

If the flood plains carry flow over the full cross section, the discharge (for very wide channels) consists of two parts:

$$Q_{river} = (hB_r)(K_{mr} h^{2/3} S^{1/2}) \qquad (4.21a)$$

and

$$Q_{floodplain} = (h - h_1)(B - B_r)[K_{mf} (h - h_i)^{2/3} S^{1/2}) \qquad (4.21b)$$

Assuming that the floodplain has the same slope as the river bed, the total discharge becomes:

$$Q_{total} = hB_r (Kmrh^{2/3} S^{1/2}) + (h - h_1)(B - B_r)[K_{mf} (h - h_i)^{2/3} S^{1/2}) \qquad (4.21c)$$

This is illustrated in Figure 2.9. The rating curve changes significantly as soon as the flood plain at level h-h1 is flooded, especially if the ratio of the storage width B to the width of the river bed Br is large. The rating curve for this situation of a compound channel is determined by considering the flow through the floodplain portion separately. This is done to avoid large values of the exponent b and extremely low values for the parameter c in the power equation for the rating curve in the main channel portion.

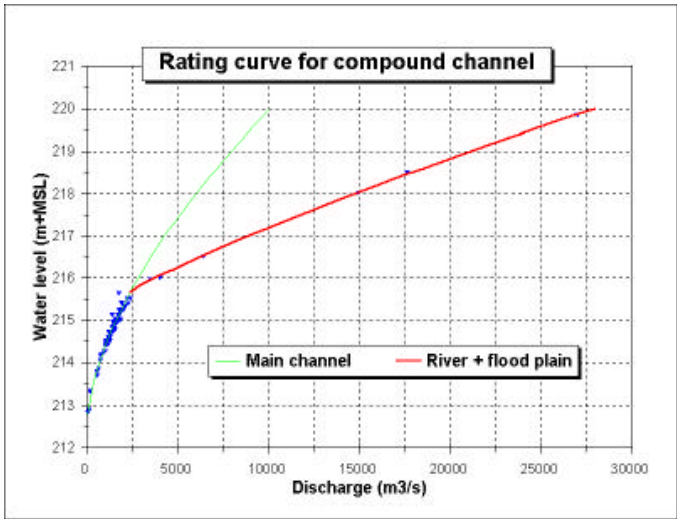

Figure 4.11 Example of rating curve for compound cross-section

The last water level range considered for fitting rating curve is treated for the flood plain water levels. First, the river discharge Q*r* will be computed for this last interval by using the parameters computed for the one but last interval. Then a temporary flood plain discharge Q*f* is computed by subtracting *Qr* from the observed discharge (*Qobs*) for the last water level interval, i.e.

$$Q_f = Q_{obs} - Q_r \tag{4.22}$$

This discharge *Qf* will then be separately used to fit a rating curve for the water levels corresponding to the flood plains. The total discharge in the flood plain is then calculated as the sum of discharges given by the rating curve of the one but last segment applied for water levels in the flood plains and the rating curve established separately for the flood plains.

The rating curve presented in Fig 2.9 for

For h < 215.67 m + MSL: Q = 315.2(h-212.38)1.706

For h > 215.67 m + MSL: Q = 315.2(h-212.38)1.706 + 3337.4(h-215.67)1.145

Hence the last part in the second equation is the contribution of the flood plain to the total river flow.

4.7 RATING CURVE WITH BACKWATER CORRECTION

When the control at the gauging station is influenced by other controls downstream, then the unique relationship between stage and discharge at the gauging station is not maintained. Backwater is an important consideration in streamflow site selection and sites having backwater effects should be avoided if possible. However, many existing stations in India are subject to variable backwater effects and require special methods of discharge determination. Typical examples of backwater effects on gauging stations and the rating curve are as follows:

a) by regulation of water course downstream,
b) level of water in the main river at the confluence downstream,
c) level of water in a reservoir downstream,
d) variable tidal effect occurring downstream of a gauging station,
e) Downstream constriction with a variable capacity at any level due to weed growth etc.
f) rivers with return of over bank flow

Backwater from variable controls downstream from the station influences the water surface slope at the station for given stage. When the backwater from the downstream control results in lowering the water surface slope, a smaller discharge passes through the gauging station for the same stage. On the other hand, if the surface slope increases, as in the case of sudden drawdown through a regulator downstream, a greater discharge passes for the same stage.

The presence of backwater does not allow the use of a simple unique rating curve. Variable backwater causes a variable energy slope for the same stage. Discharge is thus a function of both stage and slope and the relation is termed as slope-stage-discharge relation. The stage is measured continuously at the main gauging station. The slope is estimated by continuously observing the stage at an additional gauge station, called the auxiliary gauge station. The auxiliary gauge station is established some distance downstream of the main station. Time synchronization in the observations at the gauges is necessary for precise estimation of slope. The distance between these gauges is kept such that it gives an adequate representation of the slope at the main station and at the same time the uncertainty in the estimation is also smaller. When both main and auxiliary gauges are set to the same datum, the difference between the two stages directly gives the fall in the water surface. Thus, the fall between the main and the auxiliary stations is taken as the measure of surface slope. This fall is taken as the third parameter in the relationship and the rating is therefore also called *stage-fall-discharge relation*.

Discharge using Manning's equation can be expressed as:

$$Q = K_m R^{2/3} S^{1/2} A \tag{4.23}$$

Energy slope represented by the surface water slope can be represented by the fall in level between the main gauge and the auxiliary gauge. The slope-stage-discharge or stage fall-discharge method is represented by:

$$\frac{Q_m}{Q_r} = \left(\frac{S_m}{S_r}\right)^p = \left(\frac{F_m}{F_r}\right)^p \tag{4.24}$$

Where

Qm is the measured (backwater affected) discharge

Qr is a reference discharge

Fm is the measured fall

Fr is a reference fall

p is a power parameter between 0.4 and 0.6

From the Manning's equation given above, the exponent "p" would be expected to be ½. The fall (F) or the slope (S = F/L) is obtained by the observing the water levels at the main and auxiliary gauge. Since, there is no assurance that the water surface profile between these gauges is a straight line, the effective value of the exponent can be different from ½ and must be determined empirically. An initial plot of the stage discharge relationship (either manually or by computer) with values of fall against each observation, will show whether the relationship is affected by variable slope, and whether this occurs at all stages or is affected only when the fall reduces below a particular value. In the absence of any channel control, the discharge would be affected by variable fall at all times and the correction is applied by the constant fall method. When the discharge is affected only when the fall reduces below a given value the normal (or limiting) fall method is used.

4.8 CONSTANT FALL METHOD

The constant fall method is applied when the stage-discharge relation is affected by variable fall at all times and for all stages. The fall applicable to each discharge measurement is determined and plotted with each stage discharge observation on the plot. If the observed falls do not vary greatly, an average value (reference fall or constant fall) Fr is selected. Manual computation For manual computation an iterative graphical procedure is used. Two curves are used (Figs. 2.10 and 2.11):

1. All measurements with fall of about *Fr* are fitted with a curve as a simple stage discharge relation (Figure2.10). This gives a relation between the measured stage *h* and the reference discharge *Qr*.
2. A second relation, called the adjustment curve, either between the measured fall, *Fm*, or the ratio of the measured fall for each gauging and the constant *fall (Fm / Fr*), and the discharge ratio (*Qm / Qr*) (Figure 2.11)

This second curve is then used to refine the stage discharge relationship by calculating *Qr* from known values of *Qm* and *Fm/Fr* and then re plotting h against *Qr*. A few iterations may be done to refine the two curves.

The discharge at any time can be then be computed as follows:

1. For the observed fall (*Fm*) calculate the ratio (*Fm/Fr*)
2. Read the ratio (*Qm / Qr*) from the adjustment curve against the calculated value of (*Fm/Fr*)
3. Multiply the ratio (*Qm / Qr*) with the reference discharge *Qr* obtained for the measured stage *h* from the curve between stage *h* and reference discharge *Qr*.

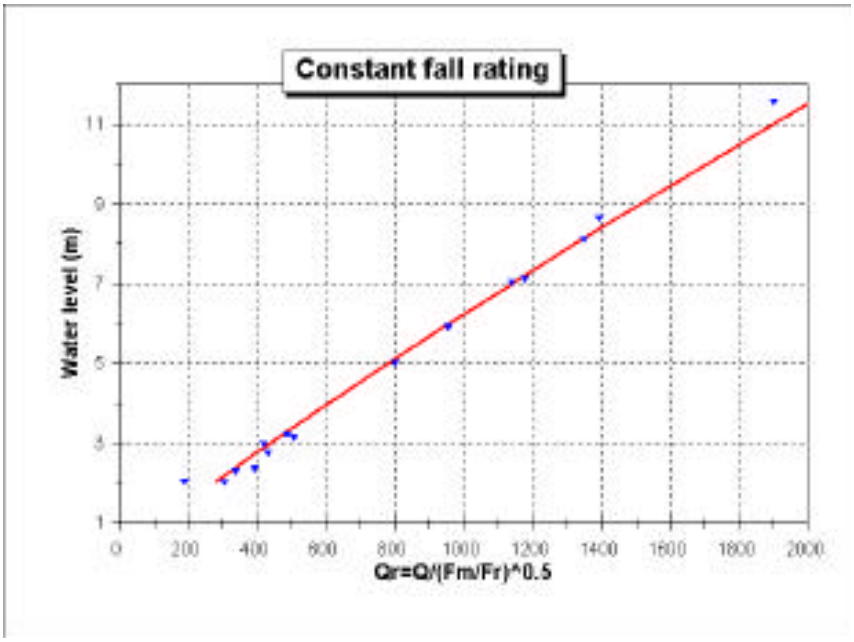

Figure 4.12 Qr = f (h) in constant fall rating

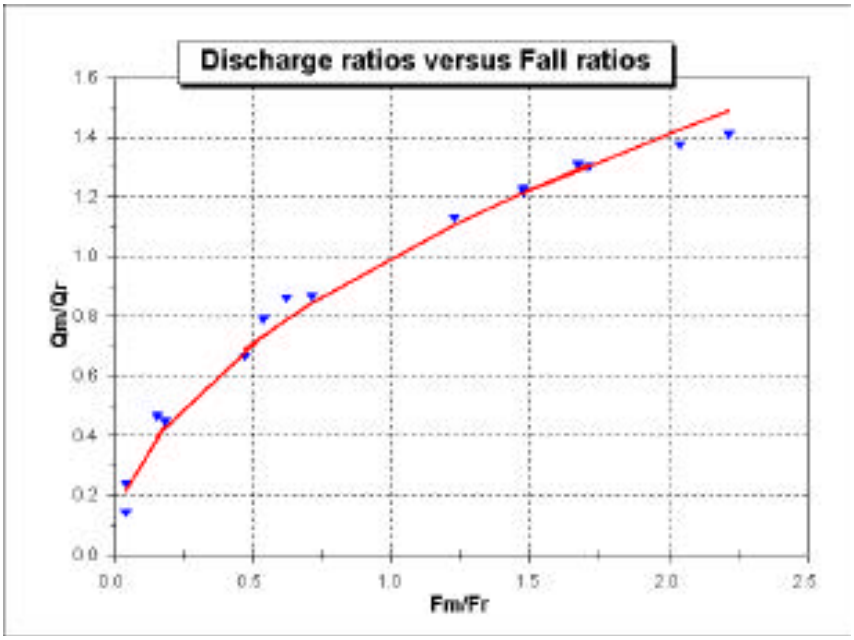

Figure 4.13 $Q_m / Q_r = f(F_m/F_r)$

4.9 NORMAL FALL METHOD

The normal or limiting fall method is used when there are times when backwater is not present at the station. Examples are when a downstream reservoir is drawn down or where there is low water in a downstream tributary or main river. The simplified procedure for such cases is given below step by step.

a) Compute the backwater-free rating curve using selected current meter gauging (the Qr -h relationship)

b) Using values of *Qr* derived from (1) and *Fr* derived from following relation

$$F_r = F_m (\frac{Q_r}{Q_m})^{1/p} \quad (4.25)$$

A parabola is fitted to the reference fall in relation to stage *(h)* as:

$$F_r = a + bh + ch^2 \quad (4.26)$$

The parameter *p* is optimized between 0.4 and 0.6.

The discharge at any time, corresponding to the measured stage *h* and fall *Fm,* is then calculated by:

a) Obtaining *Fr* for the observed *h* from the parabolic relation between *h* and *Fr*

b) Obtaining *Qr* from the backwater free relationship established between *h* and *Qr*

c) Then calculating discharge corresponding to measured stage *h* as:

$$Q = Q_r (\frac{F_m}{F_r})^p \quad (4.27)$$

4.10 RATING CURVE WITH UNSTEADY FLOW CORRECTION

Gauging stations not subjected to variable slope because of backwater may still be affected by variations in the water surface slope due to high rates of change in stage. This occurs when the flow is highly unsteady and the water level is changing rapidly. At stream gauging

stations located in a reach where the slope is very flat, the stage-discharge relation is frequently affected by the superimposed slope of the rising and falling limb of the passing flood wave. During the rising stage, the velocity and discharge are normally greater than they would be for the same stage under steady flow conditions. Similarly, during the falling stage the discharge is normally less for any given gauge height than it is when the stage is constant. This is due to the fact that the approaching velocities in the advancing portion of the wave are larger than in a steady uniform flow at the corresponding stages. In the receding phase of the flood wave the converse situation occurs with reduced approach Velocities giving lower discharges than in equivalent steady state case.Thus, the stage discharge relationship for an unsteady flow will not be a single-valued relationship as in steady flow but it will be a looped curve as shown in the example below. The looping in the stage discharge curve is also called hysteresis in the stage-discharge relationship. From the curve it can be easily seen that at the same stage, more discharge passes through the river during rising stages than in the falling ones.

4.11 APPLICATION IN THE STUDY AREA

4.11.1 General

The Mahanadi river basin and the small catchments and gauging sites that are being taken for this study is open channel involving variety of stages and hydraulic characteristics. The empirical, or also theoretical, relationship existing between the water-surface stage (i.e. the water level) and the simultaneous flow discharge in an open channel is known as *stage-discharge relation* or *rating curve,* or also just *rating*. These expressions are synonymous and they can be used interchangeably. The rating curve is a very important tool in surface hydrology because the reliability of discharge data values is highly dependent on a satisfactory stage-discharge relationship at the gauging station. Although the preparation of rating curves seems to

be an essentially empiric task, a wide theoretical background is needed to create a reliable tool to switch from measured water height to discharge.

Several methods have been discussed in the earlier section in this chapter that is to improve data fitting, but generally they have not adequately assessed the fundamentals of stage-discharge ratings based on fluid mechanics. As a consequence, several difficulties with stage-discharge ratings have been recognized. For example, in some cases, the relation between stage and discharge is not unique. The water surface slope, in fact, produces different discharges for the same stage. The present results for the study area does not touch upon all the methods discussed earlier in this section, however, it refers to the existing technical side and obviously dealing with a hydraulics subject that are widely use hydraulics concepts as, for example, uniform or normal flow, steady and unsteady flow, sediment transport and so on. Therefore it is assumed that the users are aware of this matter.

4.11.2 Fitting of single channel simple rating curve

For practical purposes the discharge rating must be developed by the application of adjustment factors that relate unsteady flow to steady flow. Omitting the acceleration terms in the dynamic flow equation the relation between the unsteady and steady discharge is expressed in the form:

$$Q_m = Q_r \sqrt{\left(1 + \frac{h}{cS_0} dh / dt\right)} \qquad (4.28)$$

Where Qm is measured discharge

Q_r is estimated steady state discharge from the rating curve

c is wave velocity (celerity)

S_0 is energy slope for steady state flow

dh/dt is rate of change of stage derived from the difference in gauge height at the beginning and end of a gauging (+ for rising ; - for falling)

Q_r is the steady state discharge and is obtained by establishing a rating curve as a median Curve through the uncorrected stage discharge observations or using those observations for which the rate of change of stage had been negligible. Care is taken to see that there is sufficient number of gauging on rising and falling limbs if the unsteady state observations are considered while establishing the steady state rating curve.

Rearranging the above equation gives:

$$\frac{1}{cS_0} = \frac{(Q_m / Q_r)^2 - 1}{dh / dt} \tag{4.29}$$

The quantity (*dh/dt*) is obtained by knowing the stage at the beginning and end of the stage discharge observation or from the continuous stage record. Thus the value of factor ($1/cS_0$) can be obtained by the above relationship for every observed stage. The factor ($1/cS_0$) varies with stage and a parabola is fitted to its estimated values and stage as:

$$\frac{1}{cS_0} = a + bh + ch^2 \tag{4.30}$$

A minimum stage h min is specified beyond which the above relation is valid. A maximum value of factor ($1/cS_0$) is also specified so that unacceptably high value can be avoided from taking part in the fitting of the parabola.

Thus unsteady flow corrections can be estimated by the following steps:

a) Measured discharge is plotted against stage and beside each plotted point is noted the value of *dh/dt* for the measurement (+ or -)

b) A trial *Qs* rating curve representing the steady flow condition where *dh/dt* equals zero is

fitted to the plotted discharge measurements.

c) A steady state discharge Q_r is then estimated from the curve for each discharge measurement and Q_m , Q_r and *dh/dt* are together used in the Equation 35 to compute corresponding values of the adjustment factor ($1 / cS_0$)

d) Computed values of *1 / cS0* are then plotted against stage and a smooth (parabolic) curve is fitted to the plotted points. For obtaining unsteady flow discharge from the steady rating curve the following steps are followed:

a) obtain the steady state flow *Qr* for the measured stage *h*

b) obtain factor (*1/cS0*) by substituting stage *h* in the parabolic relation between the two

c) obtain (*dh/dt*) from stage discharge observation timings or continuous stage records

d) substitute the above three quantities in the Equation 35 to obtain the true unsteady flow discharge

It is apparent from the above discussions and relationships that the effects of unsteady flow on the rating are mainly observed in larger rivers with very flat bed slopes (with channel control extending far downstream) together with significant rate change in the flow rates. For rivers with steep slopes, the looping effect is rarely of practical consequence. Although there will be variations depending on the catchment climate and topography, the potential effects of rapidly changing discharge on the rating should be investigated in rivers with a slope of 1 metre/ km or less. Possibility of a significant unsteady effect (say more than 8–10%) can be judged easily by making a rough estimate of ratio of unsteady flow value with that of the steady flow value.

4.12 RESULTS AND DISCUSSION

Data sets for the stations on the river Mahanadi are given in Table (4.1). In a set of points from the rating curve are plotted and to obtain the linear fit, however, a least-squares procedure has been used in log-log space that means that as this opens out the points very much at the low-flow end, that they contribute more than their real importance. This is illustrated by plotting on linear axes, as shown in Figure 4.14, using both the same data and line of best fit, now curved. A higher degree polynomial fit in terms of the logarithms of the variables could be implemented in this study, and better fits could be obtained. However, standards and other sources recommend a simple linear plot such as have been shown in Figures in Appendix (4A). Rating curved developed for fourteen sites is explained below.

4.12.1 Rating Curve for *Anandpur* GD site

For Anandpur GD site, the range of gauging record during monsoon varies from 32.62-39.27, with datum at RL= 32.33m. The rating curve has a goodness of fit (R^2) equal to 0.9453 and for non-monsoon it varies from 32.3- 34.2, with datum at RL= 31.12m. The rating curve has a goodness of fit (R^2) equal to 0.1243. The rating curve has been developed using Parabolic type of equation (Eq. 4.7a) with three parameter. The parameters a, b and c are 48.954, 32.33, and 1.1822 respectively with, R^2 equal to 0.9453. The fitting line is approximately straight meaning that there is no back water corrections and any steep or mild slope in the curve. Maximum runoff in Monsoon range from a maximum of 572.8 and a minimum of 73.82.

Table 4.2 WL-Q relationships for Monsoon period at Anandpur GD site (July-Sept) during 1972-2009

Days	WL	Q	WLav	Qest	Days	WL	Q	WLav	Qest
1	7.51	388.11	6.64	459.03	17	7.55	498.53	7.17	502.92
2	7.20	408.03	6.65	459.92	18	10.41	528.46	7.25	508.87
3	6.83	492.33	6.73	466.29	19	8.86	544.32	7.17	502.78
4	6.59	526.21	6.90	479.87	20	7.91	572.80	7.10	497.12
5	6.32	538.39	7.08	495.19	21	10.09	540.12	7.11	497.75
6	6.22	528.12	7.14	499.75	22	9.21	499.02	7.09	495.60
7	6.51	564.80	7.16	501.62	23	9.07	498.85	7.06	493.70
8	6.63	507.66	7.06	493.24	24	8.60	463.21	6.83	474.82
9	6.80	449.45	6.81	472.92	25	7.81	394.01	5.54	370.38
10	6.73	397.39	6.76	468.53	26	7.78	380.84	6.48	446.01
11	6.79	423.60	6.98	486.97	27	7.39	455.55	6.64	459.03
12	6.87	450.17	7.04	491.66	28	6.86	483.19	6.84	475.07
13	6.77	501.49	7.15	501.17	29	6.52	455.90	6.80	471.88
14	6.37	562.53	7.51	530.70	30	6.48	451.32	6.83	474.30
15	6.60	546.64	7.49	529.24	31	3.95	73.82	1.39	72.11
16	6.99	527.58	7.29	512.17					

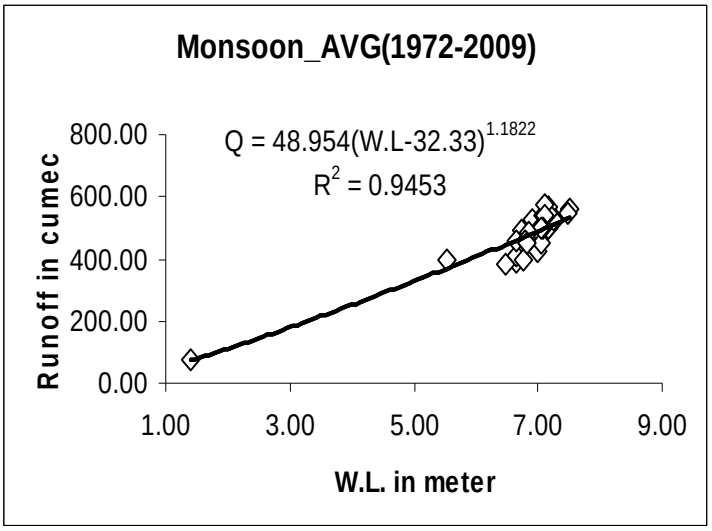

Figure 4.14a Rating curve for Anandpur site (July-Sept); R2 for Q & Q_{est} = 0.9453

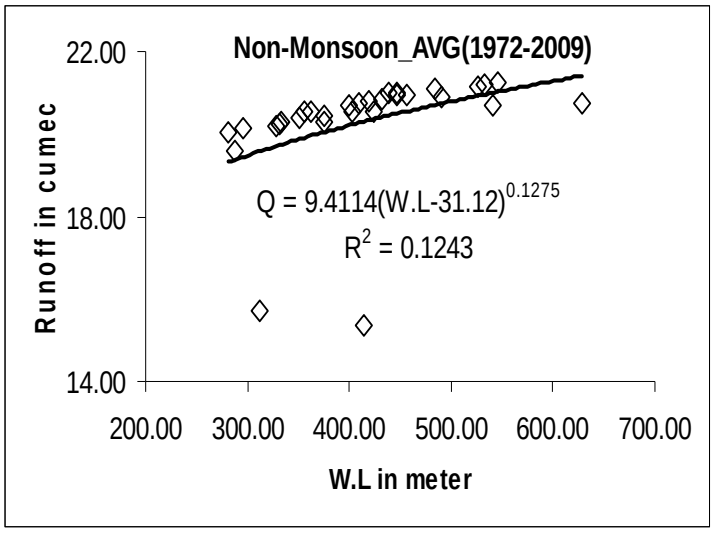

Figure 4.14b Rating curve for Anandpur (Oct-Nov); R2 for Q & Q_{est} = 0.1243

4.12.2 Rating Curve for *Altuma* GD site

For Altuma GD site, the range of gauging record during monsoon varies from 46.05-50.14, with datum at RL= 45.40m. The rating curve has a goodness of fit (R^2) equal to 0.7359 and for non-monsoon it varies from 46.03- 49.16, with datum at RL= 36.30. The rating curve has a goodness of fit (R^2) equal to 0.4707 (Refer Figure4.1 (a) and Fig 4.1 (b) Appendix 4a). The rating curve has been developed using Parabolic type of equation (Eq. 4.7a) with three parameter. The parameters a, b and c are 26.412, 45.40 and 2.1218 respectively with, R^2 0.7338.The fitting line is approximately straight meaning that there is back water corrections and any steep or mild slope in the curve. Maximum Runoff for Altuma GD Site in monsoon range from a maximum of 127.36 and a minimum of 34.77.

Table 4.3 Gauge for monsoon and non-monsoon periods using CWC Guidelines for fourteen GD sited in Mahanadi.

Sl	Gauging site	Gauge during Monsoon		Gauge during Non-Monsoon	
		Minimum	Maximum	Minimum	Maximum
1	Altuma	46.05	50.14	46.03	49.16
2	Anandpur	32.62	39.27	32.3	34.2
3	Champua	367.48	375.92	367.03	373.39
4	Gomlai	138.65	147	138.57	142.28
5	Jaraikela	186.59	193.01	185.30	190.09
6	Jenapur	17.03	23.36	16.26	21.24
7	Kantamal	101	131.73	102.63	124.58
8	Kesinga	168.12	178.5	168.09	171.69
9	Pamposh	171.32	180.2	171.37	174.5
11	Salebhata	130.4	137.62	130.22	132.64
12	Sukma	150	156.55	150	154.15
13	Sundergarh	216.17	222.14	216.86	218.67
14	Talcher	53.94	59.93	54.52	56.24
15	Tikerpada	54.47	73.2	54	58.39

4.12.3 Rating Curve for *Champua* GD site

Similarly for the Champua GD site, the gauging record during monsoon range from a minimum of 367.48 to a maximum of 375.92, with datum at RL value at 370.32m. Corresponding rating curve has a value of goodness of fit (R^2) that is equal to 0.845 and for non-monsoon it varies from a minimum of 367.03 to a maximum of 373.39, with datum at RL= 370.12m. The rating curve has a goodness of fit (R^2) equal to 0.3135.The rating curve is developed using the parabolic equation given above. The parameters calculated using the rating curve are a, b and c are 30.185, 370 and 2.045 respectively with, R^2 =0.845(Refer Figure4.3 (a) and Fig 4.3(b) Appendix 4a). The ratin curve come out to be a straight meaning that there is no back water corrections and any steep or mild slope in the curve. Maximum Runoff for Champua GD Site in monsoon range from a maximum of 113 and a minimum of 78.08.

4.12.4 Rating Curve for *Gomlai* GD site

For Gomlai GD site, the range of gauging record during monsoon varies from 138.65-147, with datum at RL equals to 138.3m with a goodness of fit (R^2) equal to 0.8958 and for non-monsoon it varies from 138.57 to 142.28, with datum at RL= 138.7. The rating curve formed using the above equations has a goodness of fit (R^2) equal to 0.0469 (Refer Figure4.4 (a) and Fig 4.4(b) Appendix 4a). The parameters a, b and c for the rating curve are 239.79, 138.3 and 1.6532 respectively with, R^2 is 0.8958. The rating curve comes out to be a straight line meaning that there are no back water corrections. The value of maximum Runoff for Gomlai GD Site in monsoon range from a maximum of 1311.96 and a minimum of 943.83.

Table 4.4 Rating Curves for monsoon and non-monsoon periods using CWC Guidelines for fourteen GD sited in Mahanadi.

GD sites	**Coefficients in Q = A (WL-B) ^{C}**							
	Monsoon (July-Sept)				**Non-monsoon (Oct-June)**			
	A	B	C	R^2	A	B	C	R^2
Altuma	26.412	45.40	2.1218	0.7338	15.424	363	-0.185	0.4707
Anandpur	48.954	32.33	1.1822	0.9453	9.4114	31.12	0.1275	0.1243
Champua	30.185	370	2.045	0.845	2.6615	365	5.5823	0.9331
Gomlai	239.79	138.3	1.6532	0.8958	151.15	137.2	-0.161	0.0469
Jaraikela	9.0421	185.11	3.2072	0.8265	0.8969	185.11	6.6653	0.9202
Jenapur	579.1	1753	1.382	0.885	168.9	16.31	1.209	0.300
Kantamal	6.0168	152.11	7.0053	0.8276	17.02	118.02	3.1334	0.5256
Kesinga	71.046	168.12	2.9604	0.7036	0.688	167.3	10.952	0.4599
Panposh	47.483	170.2	2.761	0.9394	16.592	170.2	6.266	0.4602
Salebhata	473.8	17.67	1.45	0.822	0.819	129.37	20.53	0.573
Sukma	6.0168	152.11	7.0053	0.8276	6.006	152.11	6.0947	0.8318
Sundergarh	89.37	216.6	2.407	0.582	19.7	215.7	1.756	0.446
Talcher	441.5	54.37	1.471	0.715	193.3	53.96	1.038	0.935
Tikarpara	889.3	55.53	1.186	0.946	321.6	53.67	1.719	0.849

4.12.5 Rating Curve for *Jaraikela* GD site

The range of gauging record during monsoon varies from 186.59-193.01 for Jaraikela GD site, with datum at RL= 185.11m. The rating curve has a goodness of fit (R^2) equal to 0.8265 and for non-monsoon it varies from minimum of 185.30 and a maximum of 190.09, with datum at RL= 185.11. The rating curve for non-monsoon has a goodness of fit (R^2) equal to 0.9202. The parameters a, b and c comes out to be 9.0421, 185.11, and 3.2072 respectively with, R^2 0.8265 (Refer Figure4.5 (a) and Fig 4.5(b) Appendix 4a) using the parabolic equation. Hera also similar to above result the fitting line is approximately straight meaning that there are also no back water corrections or any steep or mild slope in the curve. Here the maximum Runoff for Gomlai GD Site in monsoon ranges from a maximum of 596.58 and a minimum of 356.57.

4.12.6 Rating Curve for *Jenapur* GD site

Now for Jenapur GD site, the value of gauging record during monsoon range from a minimum of 17.03 to a maximum of 23.36, having datum at RL value of 16.31m. The rating curve for monsoon has a goodness of fit (R^2) equal to 0.8265 and for non-monsoon the gauging record varies from 16.26 to 21.24, with datum at RL value of 16.31 for this goodness of fit (R^2) equal to 0.9202 (Refer Figure4.6 (a) and Fig 4.6(b) Appendix 4a). For developing the Rating Curve the parabolic equation given above is used. The parameters a, b and c are 579.1, 1753 and 1.382 respectively with, R^2 0.8265. There is no steep or mild slope in the curve that a means it's a straight line which means there is no back water correction.

4.12.7 Rating Curve for *Kantamal* GD site

For the rating curve for Kantamal GD site, the gauging record during monsoon varies from 101 to 131.73, with datum value at RL= 152.11m. The rating curve for Kantamal has a goodness of fit (R^2) equal to 0.8265. Similarly for Non-Monsoon the gauging record varies

from 102.63 to124.58, with datum at RL=118.02. Here the rating curve has a goodness of fit (R^2) equal to 0.5256 (Refer Figure4.7 (a) and Fig 4.7(b) Appendix 4a).Similar set of equation have been use to develop the rating curve as above. The parameters a, b and c are 6.0168, 152.11 and 7.0053 respectively with a R^2 value of 0.8276.. The line formed after forming rating curve is almost a straight meaning that there are no back water corrections similar to above.

4.12.8 Rating Curve for *Kesinga* GD site

For Kesinga GD site, the range of gauging record during monsoon varies from 168.12-178.5, with datum at RL equals to 168.12m. Here the rating curve has a goodness of fit (R^2) equal to 0.7036 and for non-monsoon the gauging record varies from a minimum of 168.09 to a maximum of 171.69, with corresponding datum at RL=167.3 and the goodness of fit (R^2) is equal to 0.4599 (Refer Figure4.8 (a) and Fig 4.8(b) Appendix 4a). The rating curve has been developed using Parabolic type of equation with three parameter a, b and c for which the corresponding values are 71.046, 168.12 and 2.9604 respectively with R^2 0.7036. Here the fitting line is approximately straight meaning that there are no back water corrections.

4.12.9 Rating Curve for *Pamposh* GD site

Similarly for Pamposh GD site the gauging record during monsoon varies from 171.32 to 180.2, with datum at RL equals to 170.2m. With goodness of fit (R^2) equal to 0.9394 and similarly for non-monsoon it varies from 171.37 to174.5, with datum at RL=170.2. Here the rating curve has a goodness of fit (R^2) equals to 0.4602 (Refer Figure4.9 (a) and Fig 4.9(b) Appendix 4a). The rating curve has been developed using Parabolic type of equation with three parameter a, b and c having the corresponding values 47.483, 170.2 and 2.761 respectively with the value of R^2 0.9394. Here also similar to above graph the fitting line is almost a straight line meaning that there is no back water correction too.

4.12.10 Rating Curve for *Salebhata* GD site

Now for Salebhata GD site, the gauging record during monsoon ranges from 130.4 to 137.62, with datum at RL= 129.37m. The rating curve has a goodness of fit (R^2) equal to 0.822 and the gauging record for non-monsoon varies from 130.22 to 132.64, with datum at RL=129.37. For Non-Monsoon the rating curve has a goodness of fit (R^2) equal to 0.573 (Refer Figure4.10 (a) and Fig 4.10(b) Appendix 4a). The parameters for the rating curve using the parabolic equation has parameters a, b and c with values 473.8, 17.67 and 1.45 respectively with, R^2 0.822. The fitting line is approximately straight meaning that there are no back water corrections.

4.12.11 Rating Curve for *Sukma* GD site

Here the range of gauging record during monsoon varies from 150.4 to 156.55, with datum at RL equals to 152.11m. with the goodness of fit (R^2) equal to 0.8276 and for non-monsoon the gauging record varies from 150 to 154.15, with datum at RL equals to 152.11. Here rating curve has a goodness of fit (R^2) equals to 0.8318 (Refer Figure4.11 (a) and Fig 4.11(b) Appendix 4a). The rating curve has been developed using parabolic type of equation with three parameter. The parameters a, b and c are having the corresponding values are 6.0168, 152.11 and 7.0053 respectively with, R^2 0.8276. There is no back water correction in this case also as the rating curve come out to be almost a straight line.

4.12.12 Rating Curve for *Sundergarh* GD site

For Sundergarh GD site, the range of gauging record during monsoon varies from a maximum of 216.17 to 222.14, with datum at RL equals to 152.11m. The rating curve has a goodness of fit (R^2) equal to 0.582. For non-monsoon it varies from 216.86 to 218.67, with datum at RL=152.11. The rating curve has a goodness of fit (R^2) equal to 0.446 (Refer

Figure4.12 (a) and Fig 4.12(b) Appendix 4a). The rating curve has been developed using Parabolic type of equation with three parameter a, b and c having values 689.37, 216.6 and 2.40 respectively with, R^2 0.582. The fitting line is approximately straight meaning that there is no back water corrections and any steep or mild slope in the curve.

4.12.13 Rating Curve for *Talcher* GD site

For Talcher GD site, the range of gauging record during monsoon varies from 53.94-59.93, with datum at RL= 53.96m. The rating curve has a goodness of fit (R^2) equal to 0.582 and for non-monsoon it varies from 54.52- 56.24, with datum at RL=54.17. The rating curve has a goodness of fit (R^2) equal to 0.436 (Refer Figure4.13 (a) and Fig 4.13(b) Appendix 4a). The rating curve has been developed using Parabolic type of equation with three parameter. The parameters a, b and c are 441.5, 54.37 and 1.471 respectively with, R^2 0.715. The fitting line is approximately straight meaning that there is no back water corrections and any steep or mild slope in the curve. The fitting line is approximately straight meaning that there is no back water corrections and any steep or mild slope in the curve.

4.12.14 Rating Curve for *Tikerpada* GD site

For Talcher GD site, the range of gauging record during monsoon varies from 54.47-73.2, with datum at RL= 55.53m. The rating curve has a goodness of fit (R^2) equal to 0.946 and for non-monsoon it varies from 54- 58.39, with datum at RL=53.67. The rating curve has a goodness of fit (R^2) equal to 0.849 (Refer Figure4.14 (a) and Fig 4.14(b) Appendix 4a). The rating curve has been developed using parabolic type of equation with three parameter. The parameters a, b and c are 889.3, 55.53 and 1.186 respectively with, R^2 321.6. The fitting line is approximately straight meaning that there is no back water corrections and any steep or mild

slope in the curve. The fitting line is approximately straight meaning that there is no back water corrections and any steep or mild slope in the curve.

4.13 REMARKS AND CONCLUSION

When developing rating curves based on existing data the traditional power curve model may not be the best choice. Depending on the situation, a polynomial or quadratic power curve model may provide a fit that statistically better than the power curve model. The power curve model tends to perform better when data points are closely spaced and no data points lay far from the main cluster. The traditional power curve method tends to overestimate high flows. Also using the r2 value alone could be misleading in judging which model is better. F-tests, runs tests and Ryan-Joiner tests are important also for evaluating the relative performance of each model. The choice of which model to use depends on the situation of the data.

Rating-curves count a number of practical applications in hydrology, hydraulics and water resources management. For instance, hydrological rainfall-runoff models are usually parameterized on the basis of concurrent observations of rainfall and discharge; discharge observations in turn are generally derived from water-level observation by means of a rating-curve. Roughness coefficients of mathematical hydrodynamic models are calibrated by simulating historical events that are usually described in terms of boundary conditions, which include discharge hydrographs.

The results show that there was no back water correction for any sites and any steep or mild slope in the curve. The Rating curve for both Monsoon and Non-Monsoon were almost a straight line. The Value of R^2 for monsoon range from a maximum of 0.946 that is for

Tikerpada and a minimum of 0.582 that was Talcher. Similarly for Non-Monsoon the value of R^2 range from 0.902 to 0.342.

As a concluding remark, it is worth highlighting here that the sign of the bias associated with the approaches to the construction of rating-curves considered in this study (i.e. the Traditional and Constrained approaches) cannot be determined a priori. Regardless of the sign of the expected bias associated with Traditional approaches (underestimation or overestimation) our study clearly points out that bias and overall uncertainty associated with rating-curves can be dramatically reduced by constraining the identification of rating-curve with information resulting form simplified hydraulic modeling, with significant advantages for practical applications.

CHAPTER 5
DEVELOPMENT OF FLOW DURATION CURVES FOR GAUGED BASINS

5.0 INTRODUCTION

The stream flow component of the hydrologic cycle is very important for any water resources project. However, in most circumstances, water resources projects will have to be constructed at locations on a river where no measured stream flow records will be available. Since direct measurement of stream flow cannot be made at all locations and at all times, hydrological studies have concentrated on modeling techniques to estimate stream flow. Experimental and analytical studies have been taken up to characterize this component and develop techniques to estimate them.

Water Availability generally refers to the volume of water available from a basin or stream at a specified point over a specific 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 f 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 number of water resources related problems such as; (i) Design of water resources project, (ii) Determination of availability of water for irrigation, domestic, power supply or industrial use, (iii) Adjustment of long records of runoff for varying rainfall patterns and (iv) Reservoir operation planning.

Design of water resources projects across rivers; require information of several stream flow characteristics such as; catchment yield (daily, month, annual) peak discharge or design flood, low flows etc. For safe and cost effective engineering design of projects/hydraulics

structures the temporal variabilities of these characteristics have to be evaluated so that frequency of occurrences or return periods may be calculated. The flow duration curve has been identified as one of the most important descriptor of stream flow variability and find number of applications in engineering hydrology. A Flow Duration Curve (FDC) is a graphical representation of the relationship between a giver stream discharge and the percentage of time this discharge is equaled or exceeded. It is an important signature of the hydrologic response of the catchment and can also be used in variety of hydrologic analyses related to design and operation of water use and control structures.

5.1 FLOW DURATION CURVE

Flow Duration Curve (FDC) is perhaps the most basic form of data representation, which has been used in flow calculation. It shows graphically the relationship between any given discharge and percentage time that the given discharge is being equaled or exceeded. The curve can be drawn for daily or monthly flow data or for any consecutive 'D' days or month period. Discharge Q corresponding to a particular exceedance frequency are symbolized as Q_d where d is the percent of the time the specified mean daily discharge id equaled or exceeded 10% of time; 50% of time the flows are less than this values. A FDC is plot of discharge against the percent of time the flow is equaled or exceeded. This curve is also known as Discharge-Frequency Curve.

A very simple method of data transfer is the FDC or cumulative frequency curves of flow rates. The FDC technique is very useful for regional studies and for extrapolation of short records. The extrapolation of short records rests on the concept of regional hydrological similarity. The cumulative frequency curve combines number flow characteristic into one curve. The FDC can be prepared for any period but most important flow duration is one complete year.

Application of FDC's are of interest for many hydrological problems related to hydropower generation, river and reservoir sedimentation, water quality assessment, water allocation and habitat suitability.

5.1.1 Significance of flow duration curves

FDC's have long been used as means of summarizing the catchment hydrologic response. More recently, they have been used to validate the outputs of hydrologic models and/or compare observed and modeled hydrologic response. The end user of river flow data may have no knowledge of the quality control measures applied in their generation. Therefore an independent quality indicator is needed to give confidence in their use. The use of long-term flow duration curves as an indicator of data quality is proposed. This method visually highlights irregularities in river flow data and enables the type and location of the error to be readily located. It has been suggested that FDC's may provide a hydrologic 'signature' of a catchment. The hydrologic signature has been used in top-down modeling approaches to determine the appropriate level of complexity required in a rainfall-runoff model. It is hypothesized that this signature may reflect properties of the catchment, and may therefore be of potential use in regionalization studies.

Significant land use change such as dam construction, rapid urbanization etc, can have high level impact on FDC, implying that FDC may be used as an indicator of land use change in a catchment. Finally mirror-image nature of the FDC may imply that it possible to predict high flows in a particular catchment based on analysis of low flows of the catchment. The flow regime of a river can be describe using the flow duration curve (FDC), which represents the frequency distribution of flow and can be derived from gauged data. The shape of the flow-duration curve is determined by the hydrologic and geologic characteristics of the drainage area,

and the curve may be used to study the hydrologic response of a drainage basin to various types and distribution of inputs. A curve with steep slope results from stream flow that varies markedly and is largely fed by direct runoff, whereas a curve with flat slope results from stream flow that is well sustained by surface releases or ground water discharges. The slope f the lower end of the duration curve, i.e, low flow characteristics, shows the behavior of the perennial storage in the drainage basin; a float slope of the lower end indicates a large amount of storage and a steep slope indicates a negligible amount. In unregulated streams, the distribution of low flows is controlled chiefly by the geology of the basin. Thus, the lower end of the flow-duration curve is often used to study the effect of geology on the groundwater runoff to the stream.

5.1.2 Data required for development of flow duration curves

The length of data required for the development of flow duration curves depend upon the type of scheme, type of development and variability of inputs. General guidelines regarding the minimum length of data required for some of the projects are given below:

Table 5.1 Guidelines for the minimum length of data required for some of the projects

Sl. no	Type of Project	Minimum Length of Data
1	Diversion project	10 yrs
2	Within the year storage	25 yrs
3	Over the year storage project	40 yrs
4	Complex system involving combination of above	Depending upon the predominant element

The above guidelines are only illustrative and not exhaustive. Sometimes the rainfall records are available for long period; however, the runoff records are available for short period. In such a situation the rainfall-runoff relationships may be developed based on the available rainfall-runoff records for the concurrent period. Then these relationships are used to

generate the long term Runoff records corresponding to the available long term rainfall records. Many a times, all such data are not generally available and it becomes necessary to use data nearby sites.

Considering these aspects, the data requirements for major water resources projects may be summarized as follows:

1. Stream flow data of the desired specific duration at the proposed site for atleast 40 to 50 years, or
2. Rainfall data of specific duration for atleast 40 to 50 years for rain gauge stations influencing the catchment of the proposed site as well as stream flow data of specific duration at the proposed site for the last 5 to 10 years, or
3. Rainfall data of specific duration for the catchment of the proposed site for the last 40 to 50 years and flow data of the specific duration and concurrent rainfall data of existing work located upstream or downstream of the proposed site for the last 5 to 10 year or more or;
4. Rainfall data of specific duration for the catchment for the last 40 to 50 years for the proposed site and flow data and concurrent rainfall data of specific duration at the existing works on a nearby river for 5 to 10 year or more, provided orographic conditions of the catchment at the works are similar to that of the proposed site.

In case of the flow data are not virgin because of the construction of water resources projects upstream of the gauging site, the operational data such as reservoir regulation which include the outflows from the spillway and releases for various uses, etc. is required. If the flow data time series consists of the records for the period prior as well as after the construction of the structure, the flow series is considered to be non-homogeneous. Necessary modifications have to be made

to the records in order to make them homogeneous. For the development of flow duration curve and computation of dependable flow for ungagued catchments some important catchment and climatic characteristics are required. The catchment characteristics are derived from the toposheets covering the drainage area of the catchment.

5.1.3 Construction of flow duration curves

A flow duration curve is usually constructed from a time series of a river flow data for a catchment that is thought to be representative of the underlying natural variability in river flows within that catchment. The reliability of the information extracted from a FDC is influenced by;

1. The reliability of the Hydrometry at the point of measurement
2. Sampling error. The error associated with the assumption that he sample of river flow data used to construct the FDC is representative of the underlying natural variability

FDC may be constructed using different time resolutions of streamflow data: annual, monthly or daily. FDCs constructed on the basis of daily flow time series provide the most detailed way of examining duration characteristics of a river. Curves may also be constructed using some other time intervals, e.g. from *m*-day or *m*-month average flow time series. In this case, prior to FDC construction, a moving average approach is used to construct a new time series of *m*-day or *m*-month averaged flows from initially available daily or monthly data. Details on FDC construction and interpretation are provided in many sources (e.g. Searcy, 1959; Institute of Hydrology, 1980; McMahon and Mein, 1986).

The flows may be expressed in actual flow units, as percentages/ratios of MAR, MDF or some other 'index flow', or divided by the catchment area. Such normalization facilitates the comparison between different catchments, since it reduces the differences in FDCs caused by

differences in catchment area or MAR. Consequently, the effects of other factors on the shape of FDCs (aridity, geology, and anthropogenic factors) may be inspected. FDCs may be calculated: (i) on the basis of the whole available record period ('period of record FDC'), or 'long-term average annual FDCs'; (ii) on the basis of all similar calendar months from the whole record period (e.g. all Januaries—'long-term average monthly FDC' or FDC of a monthly 'window'. FDCs may also be constructed using all similar seasons from the whole record period (long-term average seasonal FDCs), for a particular season or a particular month (e.g. January 1990).

However, construction of FDC using the stream flows are done generally by two methods, they are;

1. Flow duration curve by plotting position method
2. Flow duration curve by Class-Interval method

5.1.4 FDC's by Plotting position formulae

The construction of a FDC using the stream flow observation can be performed through non-parametric plotting position formulae method. The procedure consists of following steps:

a). The observed stream flows are ranked to produce a set of ordered stream flows from the largest to the smallest observation, respectively;

b). Each ordered stream flow is then plotted against its corresponding duration

c). The percentage probability of the flow magnitude being equaled or exceeded is calculated using the formulae

$$P_e = \frac{m}{N+1} * 100 \qquad (5.1)$$

Where, m = rank given to the ordered stream flows

N = Sample length or total number of stream flows

Pe = Percentage probability of the flow magnitude being equaled or exceeded

d). Plot the discharge (Q) vs Pe, the resulting graph is the flow duration curve.

5.1.5 FDC's by class-interval method

The procedure for the class-interval method is explained below

a). Determine the size classes for tabulating the daily discharges :- Examine the range of flow values present in the data. Determine the number of log cycle (powers of 10) the data spans, and then for tallying data use the class intervals suggested in the table below.

b). Construct a table of size classes for tallying daily discharge; arrange the categories from largest to smallest.

c). Tally the discharge into the size categories. From appropriate sources of mean daily discharge data tally the number of flows in each discharge size category into the boxes on the flow duration compilation table. After you have finished tallying all the discharges, add them across to get the total number of days in each size category. Enter these totals as the Total days

d). Determine the percent of time and cumulative percent of time. Determine the percent of time each flow category represents by dividing total days in each class by the total number of days encompassed by our data.

Add the individual percents sequentially downward to find the cumulative percent of time that discharges equal or exceed the lower limit of each size category. The very lowest category will have a value of 100% since all the discharges will be greater than it.

e). Plot the curve on log-probability paper. On log-probability paper plot the water discharge corresponding to the lower boundary of each class versus the cumulative % of time that discharges equal or exceed that category. The discharges are plotted on the log scale and the cumulative percent on the probability scale. It is customary to have 0.01% at the left had edge of the graph and 99.99% at the right hand edge. Note that it is impossible to actually plot 0% or 100% on the paper, but you an safely plot your lowest flow at 99.99%.

f). Draw a smooth follow the dots curve through all the points. This is the flow duration curve.

5.2 DEVELOPMENT OF FLOW DURATION CURVES

(a) Case 1:

For gauged catchments if the data available corresponds to the situation (a), as discussed in section above, the flow duration curves from daily flow data may be developed in the following steps:

1) Choose a constant width class interval (ci) such that about 25 to 30 classes are formed;
2) Assign each day's discharge to its appropriate class interval;
3) Count the total number of days in each class interval.
4) Cumulate the number of days in each class interval to get the number of days above the lower limit of each class interval.
5) Compute the probabilities of exceedance dividing the quantities obtained from step (iv) by the total number of days in the record e.g. 365 if one year record is considered for the construction of flow duration curve.
6) Multiply the probabilities of exceedance obtained from step (5) by 100 to get percentage exceedance.

7) Plot the probabilities of exceedance in percentage against the corresponding lower bound of class interval on linear graph paper. Sometimes the flow duration curve better approximates to a straight line if log normal probability paper is used in place of linear graph paper.

(b) Case 2:

In case the data items are not sufficient enough to define the class intervals, the flow duration curves using monthly flow data or any other duration larger than daily may be developed in the following steps:

1) Arrange the flow data in descending order;
2) Assign the probability of exceedances to each data item obtained from step (1) using the Weibull plotting position formula:

$$P = \frac{m}{N+1} x\ 100 \qquad (5.2)$$

where m = 1 for the highest flow values and N is the number of data items (or variate).

Note: If the flow duration curve is required to be linearized on normal probability paper or log normal probability paper, the probability of exceedences may be assigned using the Blom's Plotting position formula:

$$P = \frac{m - 0.375}{N + 0.250} x\ 100 \qquad (5.3)$$

3) Plot the ranked flow values against the probabilities of exceedences computed using eq. 5.1) on linear graph paper to get the flow duration curve.

Note: Use normal probability paper, if the required dependable flow (or probability of exceedance) is to be extrapolated. Try to fit either normal distribution or log normal

distribution in order to linearized the flow duration curve. Here the probabilities of exceedance may be computed using Eq. 5.2 for the purpose of plotting. Fitting of other theoretical frequency distribution may also be tried.

(c) Case 3:

If the data situation corresponds to (b), as discussed in section 6.1.1 under data requirement, the steps involved in development of flow duration curves are as follows:

1) Develop the rainfall runoff relationship for the specific duration utilising the available data for the concurrent period;
2) Compute the long term flow data of the specific duration using the developed relationship at step (1) and long term available rainfall data.
3) Develop the flow duration curve either using the procedure stated for Case 1 or for Case 2.

(d) Case 4:

If the data availability situation corresponds to either (c) or (d), as discussed in section 6.1.1 under data requirement, the following steps may be followed for the development of the flow duration curve.

1) Develop the rainfall-runoff relationship for the existing site for the specific duration analysing the available rainfall-runoff records for the concurrent period;
2) Develop the flow duration curve using the procedure described either for Case 1 or Case 2.
3) Divide the flow values of flow duration curve by the catchment area of the existing project site.

Multiply the flow values obtained from step (3) by the catchment area of the proposed site for which the flow duration curve is required to be developed.

5.3 USE OF FLOW DURATION CURVES

The first use of FDCs can be traced back to Clemens Herschel (Foster, 1934) in the 1880s. The use of FDCs is widespread in hydrology, water resources research and hydraulic engineering. Vogel & Fennessey (1994, 1995) provide a brief history of the applications with extensive references. Smakhtin's recent review (2001) documents the use of FDCs in low-flow hydrology. Recent applications include: studies of the impact of land-use/cover change on streamflow from paired catchment studies (Best *et al*., 2003; Croker *et al*., 2003; Lane *et al*., 2003), regionalization studies (Fennessey & Vogel, 1990; Meunier, 2001); and in prediction for ungauged basins (Yu & Yang, 2000; Holmes *et al*., 2002; Yu *et al*., 2002).

5.4 RESULTS AND DISCUSSION

5.4.1 Application to Gauging Sub catchments in Mahanadi Region

This study has been carried out for the small and sub-catchments of Mahanadi River, almost 885 kms long. Figure 2.2 shows the details of catchments of Mahanadi system inside and outside of Orissa (Source: Parhi et al. 2009). The total catchment extends over an area of 141,600 sq kms. Figure 5.1 shows which region each of the catchments is in the Mahanadi basin. As in the figure, the whole region is divided into three physio-hydrological regions i.e. upper, lower and central as described in the earlier studies and reports. River system and observatory sites under Hydrology project (HP2) in Mahandi basin are in detail shown in Figure 5.1 a-b. Figure 5.1 (c) and (d) shows the nearby important basins of Brahmani and Rushikulya basins that significantly contribute to the hydrological impacts of the Mahandi basin, especially at the outlet at Bay-of-Bengal. Table 5.1 shows the distribution of rain gauge stations in

Mahanadi Zone, and gives the daily mean rainfall in MM at each station mean rainfall values. The mean is in the basin is about 715mm (period 1921–2004); precipitations are concentrated in the monsoon period) July-September) and October-June months are usually rainless.

The total-period method that each river would need to be considered in the context of (i) the natural flow regimes of nearby rivers, and (ii) the main purpose for which it is regulated. However, the method described here needs further development to determine whether fewer flow statistics can achieve acceptable levels of flow characterization. To determine a suitable flow regime we must first be able to characterize the manageable aspects of flow regime in terms of five elements: (i) magnitude of discharge; (ii) frequency of occurrence, such as the frequency of floods of certain magnitude; (iii) duration, or the period of time a specified flow condition is sustained; (iv) timing, or predictability of flows of a defined magnitude; (v) rate of change, or the speed at which flow changes from one magnitude to another. There are many ways of incorporating these five elements into variables that characterize the flow regime, as can be seen from the large number of hydrological variables used in recent studies, (e.g. Hughes & James 1989; Arthington 1994; Richter *et al*. 1996; Clausen & Biggs 1997; Puckridge *et al*. 1998; Toner & Keddy 1998).

5.4.2 Flow Duration Curves for Monsoon and Non-Monsoon Period Data

The results presented in this section used two seasonal i.e. monsoon (July-Sept) and Non-Monsoon (Oct-June) periods to construct flow-duration curves. Monthly streamflows for fourteen gauging sites as given in Table. 5.2 (a) have been collected for from Central Water Commission, and annual maximum series for twenty five bridge catchments (Table 5.2 b) and other geomorphological with salient statistical properties from CWC reports (CWC, 1982a, 1997). The GD sites with the corresponding data type and period is shown in Table 5.2.

Table 5.2. Hydrological Flow data for Mahanadi.

Fourteen Gauging sites maintained by CWC (Date type: Daily gauge with corresponding mean flow)							
Sl	**GD site**	**River/Tributary**	**Period**	**Sl**	**GD site**	**River/Tributary**	**Period**
1	Altuma	Ramiyal	1990-2009	8	Kesinga	Mahanadi	1979-2009
2	Anandpur	Brahmani	1972-2009	9	Kantamal	Mahanadi	1990-2009
3	Gomlai	Brahmani	1979-2009	10	Salebhatta	Mahanadi/Ong	1973-2009
4	Jaraikela	kelo	1972-2009	11	Jenapur	Brahmani	1980-2009
5	Pamposh	Brahmani	2000-2009	12	Sukma	Mahanadi / Tel	1989-2003
6	Champua	Brahmani	1990-2009	13	Talcher	Brahmani	2000-2009
7	Sundergarh	Mahanadi/Ib	1972-2009	14	Tikerpada	Mahanadi	2000-2009

As observed, the working period of most gauging sites starts in the middle '70 since only few stations give runoff data previous to this year. This fact suggested the opportunity to limit the analysis to the 30-years period ranging from 1975 to 2007 in case of daily/monthly flows of fourteen gauging sites maintained by CWC . In case of twenty five bridge catchments, three unregulated basins with at least less i.e. five or ten years of data have been excluded in this section reducing the number of stations used in this study to 23. For these basins the mean daily streamflows is not available, however, the values recorded are annual maximum series (AMS) and the record length is 33 years while the mean sample size is about 20 years. These data are used in flood frequency analysis for return period flood estimation along with fourteen basins as mentioned earlier.

5.4.2.1 Monsoon Period

The Flow duration curve for monsoon period is from July to September because of the data survey that is available in Mahanadi region, the maximum flow (more than average) is frequently occurs during this period. The discussion in regards to FDC for all the fourteen GD sites ranged in size from 3.4 to 1844 km^2 is discussed in the following section.

FDC in Gomlai GD site

This table shows the mean flow of *Gomlai* GD site of Mahanadi basin .For this the mean flow of the monsoon period is considered i.e. mean monthly data during July-Sept. The flow duration curve for various probability of exceedences are computed using (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formulae.

Table 5.3 Flow duration curve for monsoon period (July-Sept) in Gomlai GD site

Rank (n)	Mean flow (m^3/s)	p (>Q)			
		WEIBULL	BLOM	GRINGORTON	CUNNANE
1	3591.31	3.1	2	1.8	1.92
2	2763.16	6.3	5.2	5.0	5.13
3	2294.10	9.4	8.4	8.2	8.33
4	1961.71	12.5	11.6	11.4	11.54
5	1724.71	15.6	14.8	14.7	14.74
6	1581.29	18.8	18	17.9	17.95
7	1473.75	21.9	21.2	21.1	21.15
8	1370.47	25.0	24.4	24.3	24.36
9	1285.22	28.1	27.6	27.5	27.56
10	1208.34	31.3	30.8	30.7	30.77
11	1145.18	34.4	34	33.9	33.97
12	1078.81	37.5	37.2	37.1	37.18
13	1008.52	40.6	40.4	40.4	40.38
14	954.29	43.8	43.6	43.6	43.59
15	899.32	46.9	46.8	46.8	46.79
16	836.12	50.0	50	50.0	50.00
17	794.42	53.1	53.2	53.2	53.21
18	757.79	56.3	56.4	56.4	56.41
19	717.94	59.4	59.6	59.6	59.62
20	679.23	62.5	62.8	62.9	62.82
21	640.81	65.6	66	66.1	66.03
22	604.65	68.8	69.2	69.3	69.23
23	565.80	71.9	72.4	72.5	72.44
24	527.29	75.0	75.6	75.7	75.64
25	497.74	78.1	78.8	78.9	78.85
26	467.87	81.3	82	82.1	82.05
27	429.79	84.4	85.2	85.3	85.26
28	395.20	87.5	88.4	88.6	88.46
29	353.44	90.6	91.6	91.8	91.67
30	307.58	93.8	94.8	95.0	94.87
31	298.85	96.9	98	98.2	98.08

Inspecting FDC developed using all the four methods (Fig 5.1) there is a probability of more than ninety seven percent is that a flow of around 300 cumecs exceeds the GD site during the monsoon period. Similarly a flow exceeding 1000 cumecs has a probability of fifty percent and 3000 cumecs flow exceedance comes to around ten percent probability.

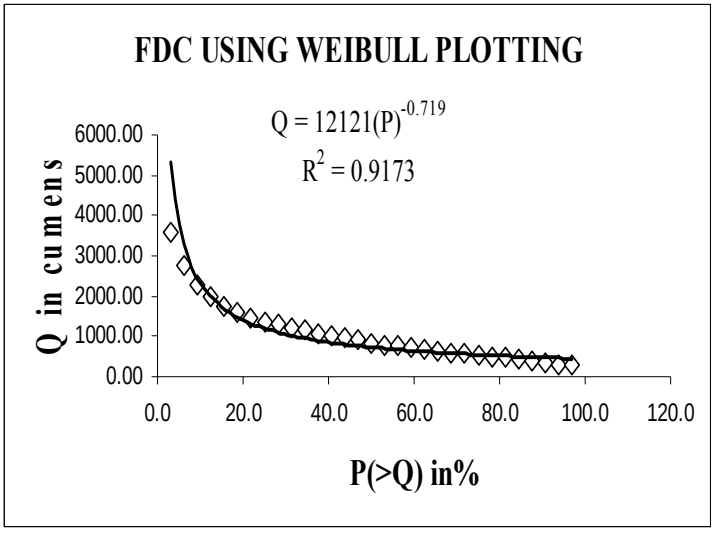

Figure 5.1 a Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for *Gomlai* GD site

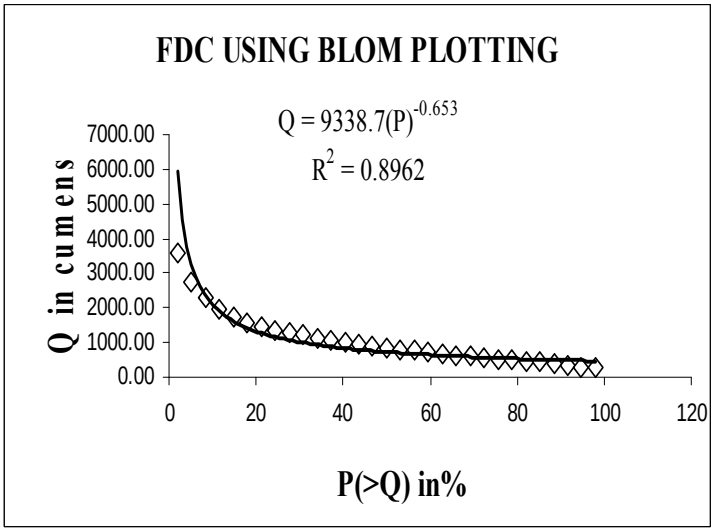

Figure 5.1 b Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for *Gomlai* GD site

One could argue that both curves in Figure 5.1 (a) and (d) are almost equivalent, but that the estimator Q(>P) has the advantage of being easily implemented on a computer and leads to significantly smoother quantile functions than the traditional small samples of ten or forty years data record. When the equation is fitted to FDC as it can be seen that the R^2 values exceeding 0.89 (Weibull), 0.89 (Blom), with most of them having poorer fits than the more perennial catchments. The slope of the lower end of the duration curve (80-95 %) shows a flat slope which indicates flow greater than 300 cumecs up to 500 cumecs has consistent probability of exceedance and might be a large amount of storage; and at the upper end has a steep slope (3-20 %) that indicates a negligible amount in Figure5.1 d.

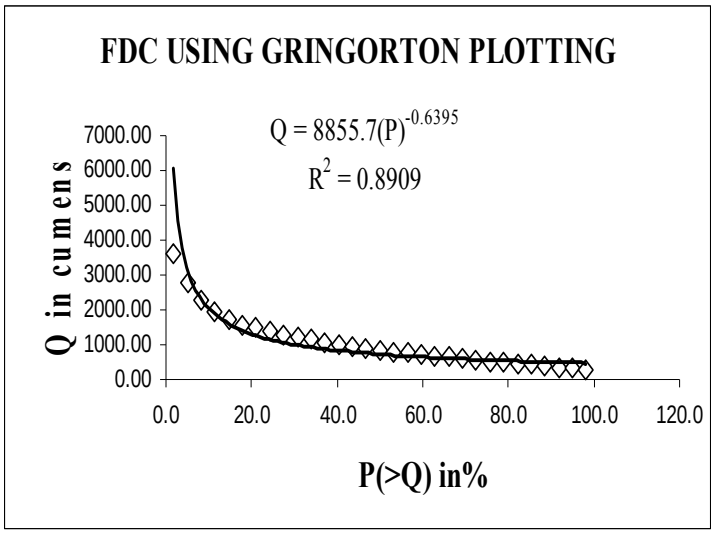

Figure 5.1 c Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for Gomlai GD site

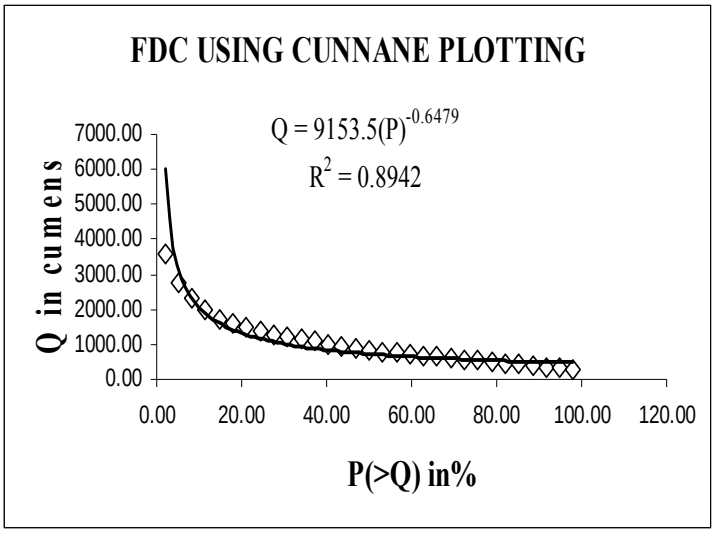

Figure 5.1 d Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for *Gomlai* GD site

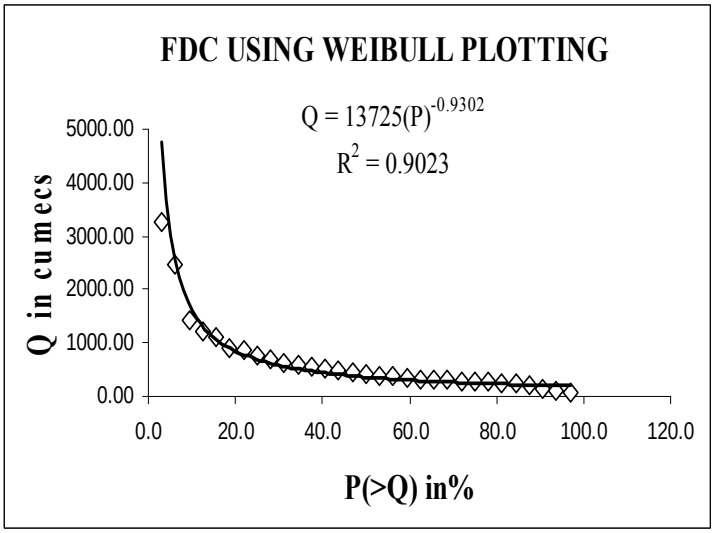

Figure 5.2 a Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for Kesinga GD Site

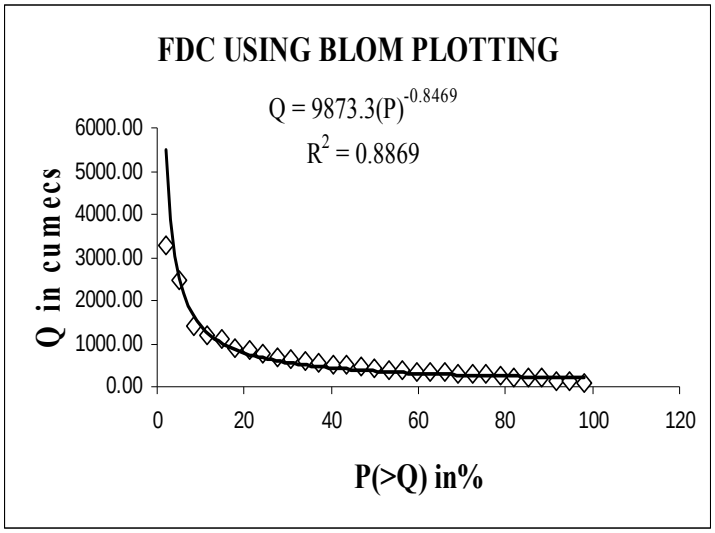

Figure 5.2 b Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for Kesinga GD Sit

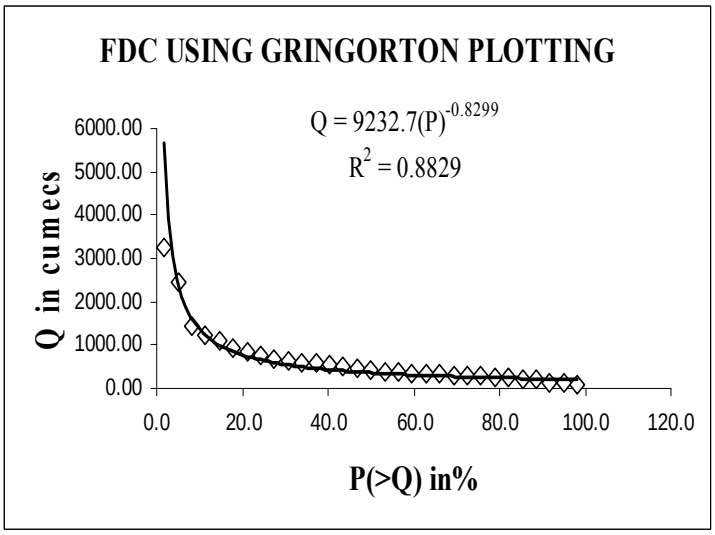

Figure 5.2 c Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for Kesinga GD Site

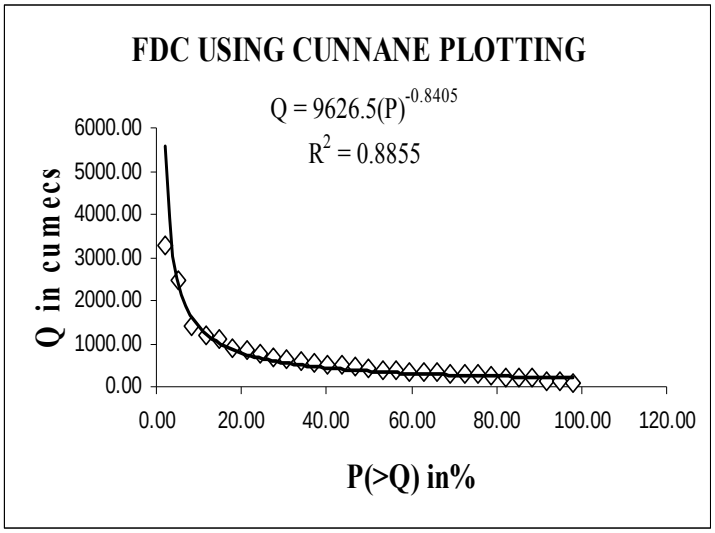

Figure 5.2 d Flow Duration Curve using average for Monsoon (July-Sept) season (1979 - 2009) for Kesinga GD Site

FDC in Kesinga GD site

The FDC during monsoon period at Kesinga GD side is developed using mean flow for the available period 1979 -2009. Fig 5.2 (a)-(d) (referred in Appendix 5 b) shows FDC using Weibull, Blom Gringorton and Cunnane using normalized mean flow, respectively. Figures show a clear picture of the variation of the river flow at this site during the monsoon period and also the probability of flow exceedance in X-axis and corresponding flow magnitude at the site. Flow duration curve using normalized mean monsoon flow at *Kesinga* GD site (1979-2009) has fit indexed with R^2 near 0.8855 for monsoon. The slope of the lower end of the duration curve shows a flat slope (80-96%) which indicates a large amount of storage; and at the upper end is a steep slope (20-40%) that indicates a negligible amount in Figure5.2 d.

FDC in Altuma GD site

Figure 5.3 (Appendix 5 b) for *Altuma* GD site with fifteen years period of- record using two different periods of record. Figure 5.2 illustrates how sensitive the lower tail of an FDC can

be to the chosen period of record. The period of record 1989-2001 contains the drought that was more drought experienced, hence the FDC for these two periods are significantly different. The R^2 value is greater than 0.9 then the flow duration curve is suitably straight for this range of percentiles. The figure shows flat slope (40-97 %) steep slope (30-50) that indicates a negligible amount.

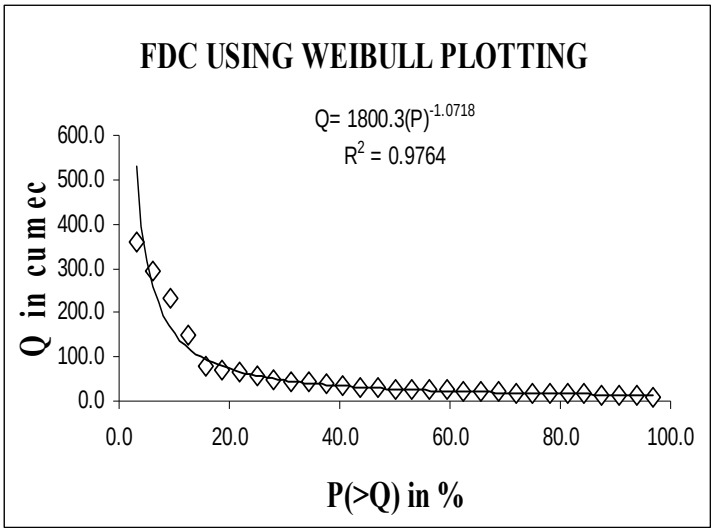

Figure 5.3 a Flow Duration Curve using average for Monsoon (July-Sept) season (1990 - 2009) for Altuma GD Site

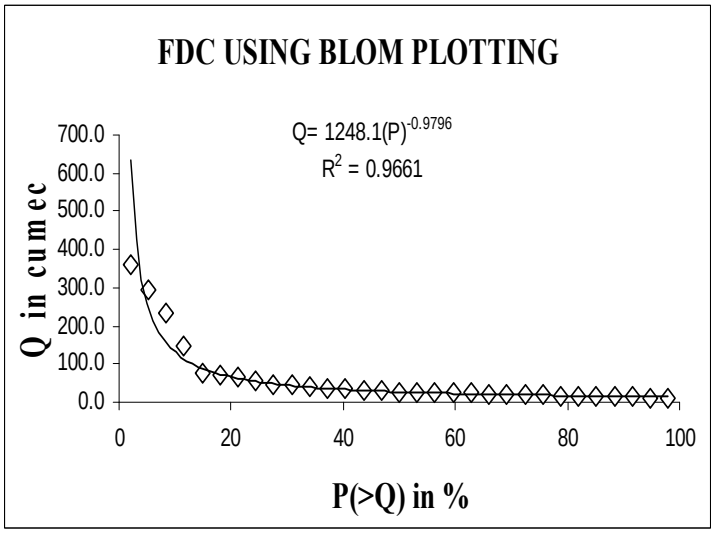

Figure 5.3 b Flow Duration Curve using average for Monsoon (July-Sept) season (1990 - 2009) for Altuma GD Site

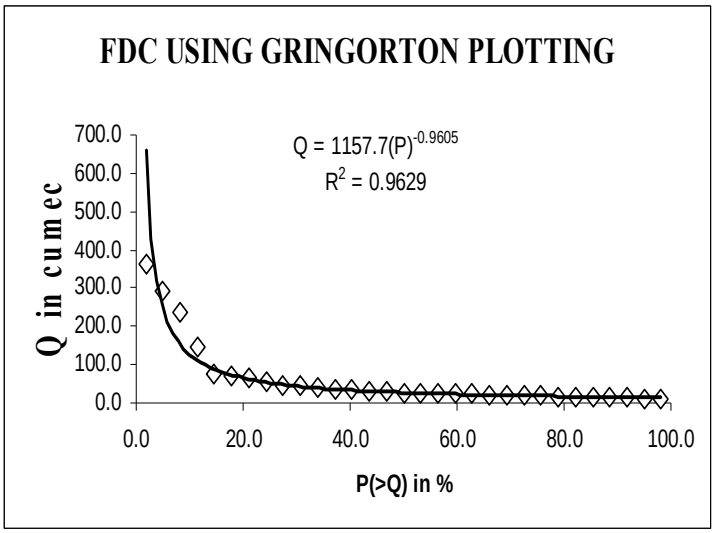

Figure 5.3 c Flow Duration Curve using average for Monsoon (July-Sept) season (1990 - 2009) for Altuma GD Site

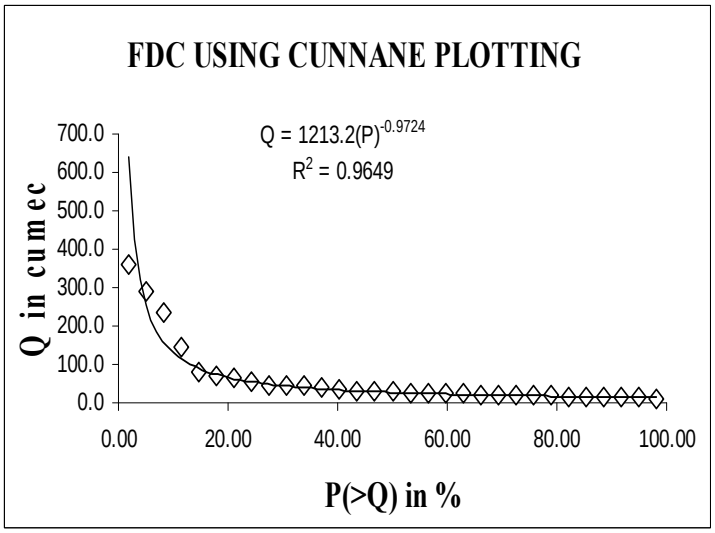

Figure 5.3 d Flow Duration Curve using average for Monsoon (July-Sept) season (1990 - 2009) for Altuma GD Site

FDC in Kantamai GD site

Figures in Appendix 5 b plot the FDC using Weibull, Blom Gringorton and Cunnane for the monsoon period, respectively. From the Table 5.6 (Appendix 5a) and the corresponding

figures shows a apparent image of the difference of the river flow at this site during the monsoon period and also the possibility of a approximate flow exceedance during the monsoon period. On inspecting the FDC developed using all the four methods. a flow exceeding 600 cumecs has a probability of fifty percent and 120 cumecs flow exceedance comes to around ten percent probability. The Figure 5.7 (a)-(d) in Appendix 5 b shows flat slope (80-96 %) steep slope (50-75%) that indicates a negligible amount.

FDC in Jenapur GD site

The FDC during monsoon period at *Jenapur* GD side is developed using mean flow for the available period 1979 -2009. Figure 5.6 (a)-(d) (Appendix 5 b) shows FDC using Weibull, Blom Gringorton and Cunnane using mean flow, respectively. Figures show a flat slope (80-92 %) steep slope (30-40%) that indicates a insignificant amount. Similarly, Figure 5.10 (a)-(d), 5.11 (a)-(d) in appendix 5 b of Salebhatta shows flat slope (60-90 %) steep slope (50-65%) that indicates a slight amount and Sukma shows flat slope (60-92 %) steep slope (20-30%) that indicates a minor amount.

5.4.2.2 Non- Monsoon Period

In case the adequate records of stream flow are available at the site, the flow duration curves may be easily developed analysing those records. However, for the ungauged sites or sites having inadequate stream flow data the regional flow duration curves are developed and utilized for the estimation of available water resources. This is discussed in the next chapter dealing on regional basis. Figure 5.2b (Appendix 5b) shows the FDC using average for monsoon period at Anandpur GD taking all the data during 1972 -2009. As a long period data length i.e. 37 years, it becomes easier for to analyze and discuss the methods and the results in brief for others to understand. Figure 5.1 (a)-(d) in Appendix 5(b) shows FDC using Weibull,

Blom Gringorton and Cunnane for the monsoon period, respectively. From the Table 5.1 (appendix 5a) and the corresponding figures shows a clear picture of the variation of the river flow at this site during the monsoon period and also the probability of a particulate flow exceedance during the monsoon period. Inspecting FDC developed using all the four methods, there is a probability of more than ninety seven percent is that a flow of around 110 cumecs exceeds the GD site during the monsoon period. Similarly a flow exceeding 300 cumecs has a probability of fifty percent and 750 cumecs flow exceedance comes to around ten percent probability. Looking at the highest flow given in Table 5.1 in Appendix 5(a) a runoff of 2070.26 cumecs and all the four methods gives a probability of exceedance of about 2 percent with Weibull having the highest probability of 3.1 percent and Gringorton plotting method with 1.9.

FDC in Altuma GD site

The FDC during non monsoon period at Altuma GD side is developed using mean flow for the available period 1990 -2009. Figure 5.4 (a)-(d) shows FDC using Weibull, Blom Gringorton and Cunnane using normalized mean flow, respectively. Figures show a clear picture of the variation of the river flow at this site during the monsoon period and also the probability of flow exceedance in X-axis and corresponding flow magnitude at the site. Flow duration curve using normalized mean monsoon flow at *Altuma* GD site (1990-2009) has fit indexed with R^2 near 0.8855 for monsoon. In respective figures in FDC, the slope of the lower end of the duration curve shows a flat slope (80-90%) which indicates a large amount of storage; and at the upper end is a steep slope i.e. 20-25% indicates a negligible amount. Figures shows a clear picture of the difference of the stream flow at this site during the monsoon period and also the probability of flow exceedance in X-axis P (>3,>50, and>97) % exceedance and the equivalent mean flow.

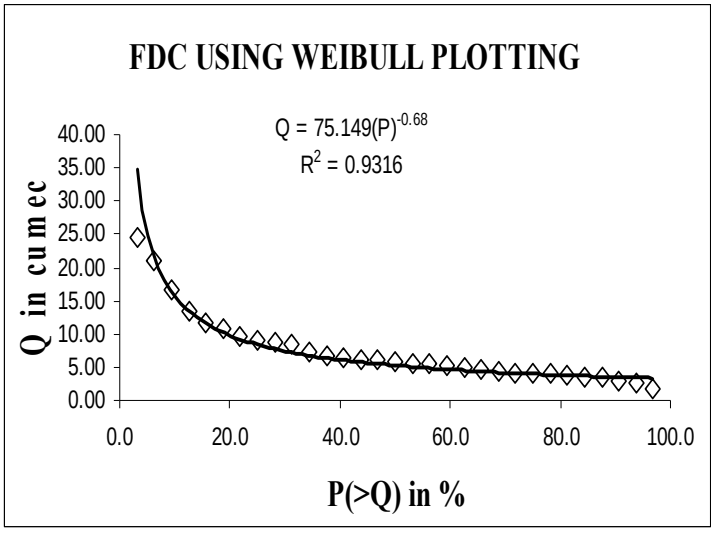

Figure 5.4 a Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09) for *Altuma* GD Site

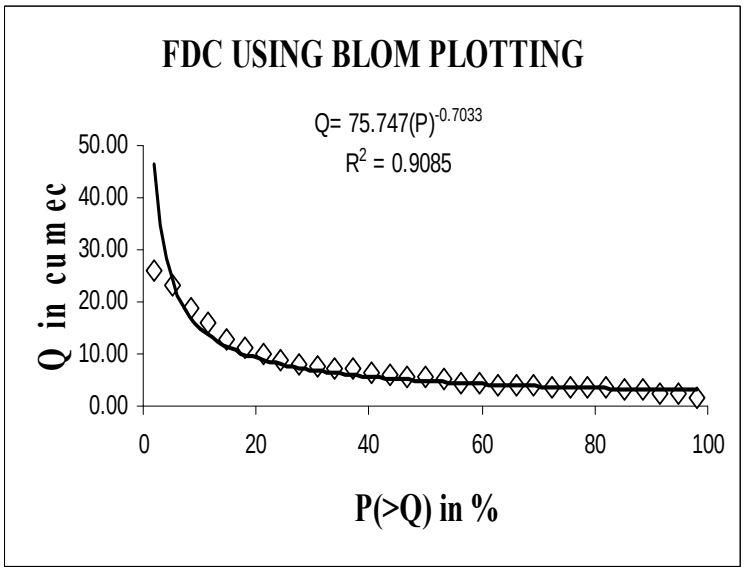

Figure 5.4 b Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09) for *Altuma* GD Site

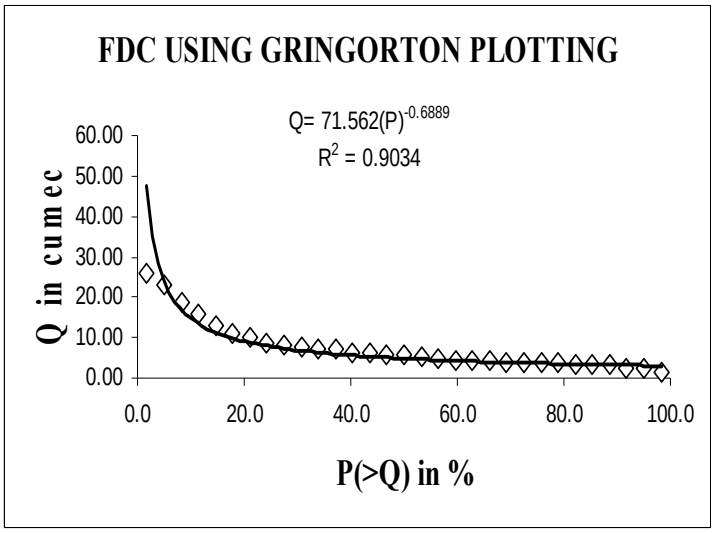

Figure 5.4 c Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09) for *Altuma* GD Site

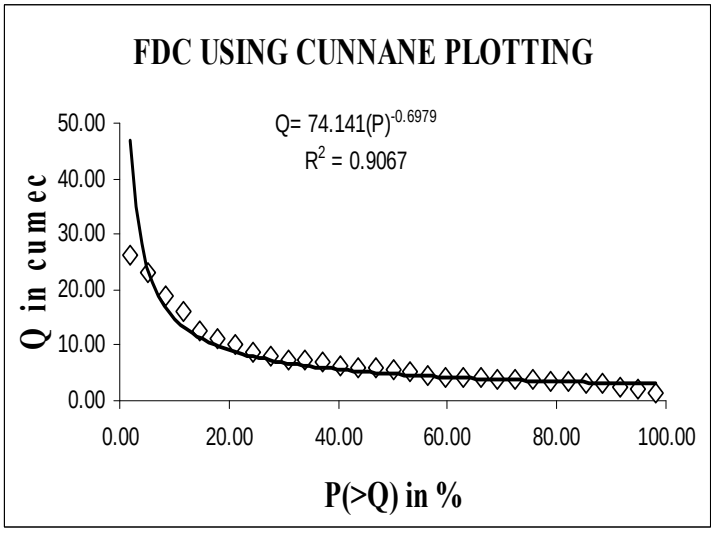

Figure 5.4 d Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09) for *Altuma* GD Site

FDC in Jaraikela GD site

The flow duration curve for various return periods calculated by four methods i.e. (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formula at monthly time intervals.

Flow duration curve of each water year was constructed by plotting and arranging the daily discharge values in descending order. The characteristics of probability distribution for daily discharge values extracted at 5% of the time in a given N year data are graphically presented by plotting magnitude versus probability.

Table 5.4 Flow duration curve for monsoon period (July-Sept) in *Jaraikela* GD site

Rank (n)	Mean flow (m^3/s)	Flow duration curve using following methods to get p(>Q)			
		WEIBULL	BLOM	GRINGORTON	CUNNANE
1	1981.686	3.1	2	1.8	1.92
2	1194.788	6.3	5.2	5.0	5.13
3	985.9832	9.4	8.4	8.2	8.33
4	838.6894	12.5	11.6	11.4	11.54
5	750.7525	15.6	14.8	14.7	14.74
6	669.2251	18.8	18	17.9	17.95
7	611.1414	21.9	21.2	21.1	21.15
8	565.8185	25.0	24.4	24.3	24.36
9	512.8982	28.1	27.6	27.5	27.56
10	485.0769	31.3	30.8	30.7	30.77
11	453.7554	34.4	34	33.9	33.97
12	429.6097	37.5	37.2	37.1	37.18
13	403.3569	40.6	40.4	40.4	40.38
14	380.2367	43.8	43.6	43.6	43.59
15	360.2362	46.9	46.8	46.8	46.79
16	340.6172	50.0	50	50.0	50.00
17	322.3813	53.1	53.2	53.2	53.21
18	302.53	56.3	56.4	56.4	56.41
19	286.1154	59.4	59.6	59.6	59.62
20	269.9351	62.5	62.8	62.9	62.82
21	253.7372	65.6	66	66.1	66.03
22	237.8145	68.8	69.2	69.3	69.23
23	225.0796	71.9	72.4	72.5	72.44
24	211.7854	75.0	75.6	75.7	75.64
25	200.5019	78.1	78.8	78.9	78.85
26	185.7582	81.3	82	82.1	82.05
27	172.5709	84.4	85.2	85.3	85.26
28	158.7221	87.5	88.4	88.6	88.46
29	144.7095	90.6	91.6	91.8	91.67
30	130.5284	93.8	94.8	95.0	94.87
31	113.1675	96.9	98	98.2	98.08

This table shows the relationship between the catchment area and the mean flows of different thirteen gauging site of Mahanadi basin .For this we take the common of the mean flow of the monsoon period. The flow duration curve for various return periods calculated by four methods i.e. (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formula at monthly time intervals .The Flow duration curve were made by using the Weibull formula P = m / (N+1)*100 at various probability of flow i.e. P (>3, >50, and >97) % exceedance and the corresponding mean flow. In Blom method we use the formula P = ((m - 0.375) / (N + 0.25))*100 at various probability of flow i.e. P (>2, >50, and >98) % exceedance and the corresponding mean flow. In Gringorton and Cunnane method we use the formula P = ((I - 0.44) / (N + 0.12))*100 and P = ((I - 0.4) / (N + 0.2))*100 at various probability i.e. P (>1.8, >50, and >98.2) % and P (>1.92, >50, and >98.08) % exceedance respectively

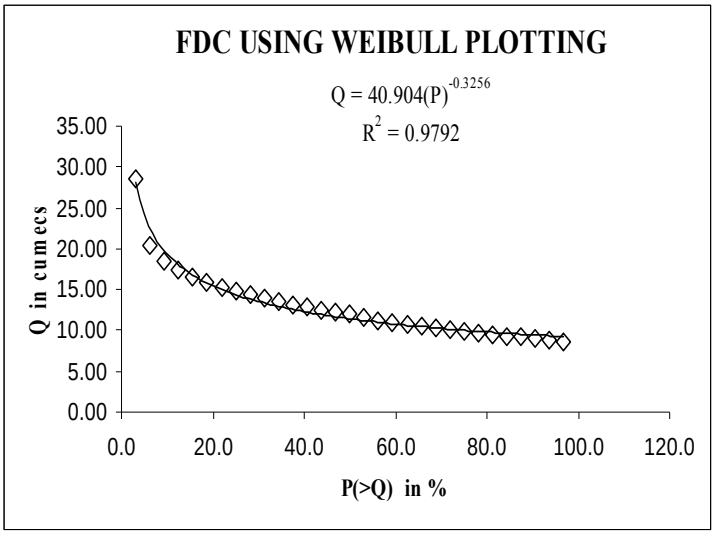

Figure 5.5 a Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1979-09) for *Champua* GD Site

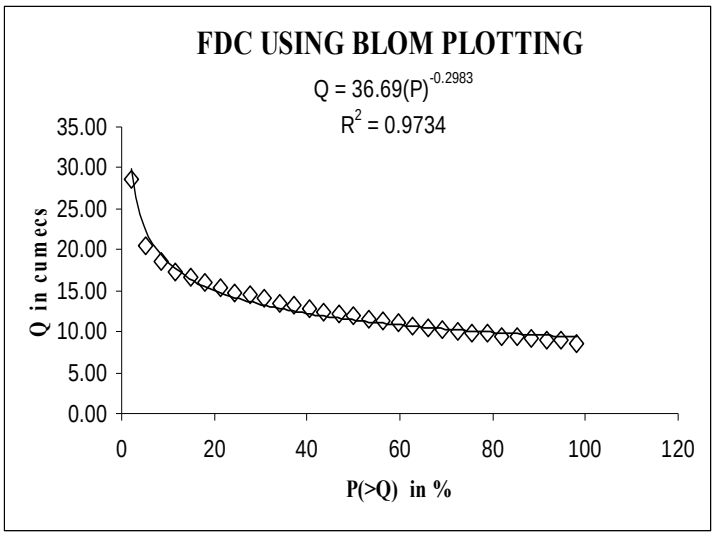

Figure 5.5 b Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1979-09) for *Champua* GD Site

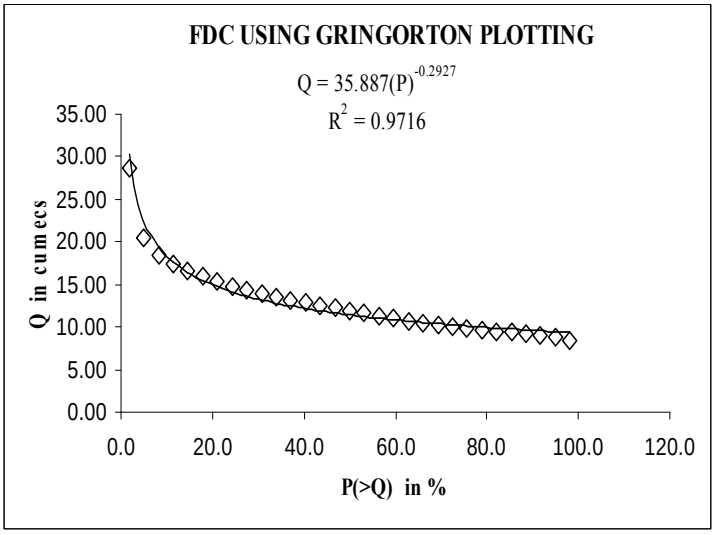

Figure 5.5 c Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1979- 09) for *Champua* GD Site

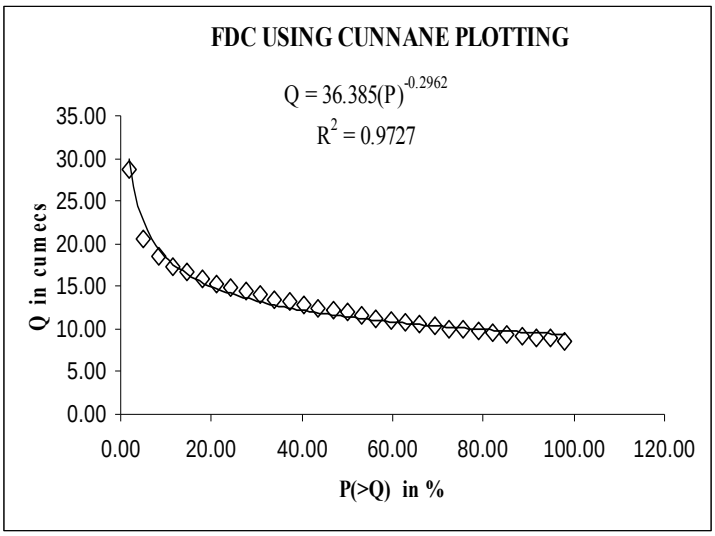

Figure 5.5 d Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1979-09) for *Champua* GD Site

FDC in Champua GD site

The FDC during non monsoon period is developed using mean flow for the available period 1979 -2009. The table in Appendix 5a shows FDC using Weibull, Blom, Gringorton and Cunnane mean flow. Figure 5.5 shows flat slope (60-80%) steep slope (20-30%) that indicates a negligible amount.

FDC in Kesinga GD site

The FDC during monsoon period at Kesinga GD side is developed using mean flow for the available period 1979 -2009. Figure 5.8 (e)-(h) (Appendix 5 b) shows FDC using Weibull, Blom Gringorton and Cunnane using normalized mean flow, respectively. Flow duration curve using non monsoon mean flow at *Kesinga* GD site (1979-2009) has fit indexed with R^2 near 0.9868 for nonmonsoon. The slope of the lower end of the duration curve shows a flat slope (40-80%) which indicates a large amount of storage; and at the upper end is a steep slope (30-40%) that indicates a small amount.

From the results, it can be note that only 34.2% of the daily flows exceeded the mean over the period of record. This is not unusual, and this result emphasizes how misleading it can be to use the mean as a measure of central tendency for highly skewed data such as daily stream flow. Daily streamflows are so highly skewed that ordinary product moment ratios such as the coefficient of variation and skewness are remarkably biased and should be avoided, even for samples with tens of thousands of flow observations. Although FDCs are appealing for depicting the hydrologic response of a river basin, they can be misleading because the autocorrelation structure of streamflow series is sometimes effectively removed from the plot. In this case this has been avoided.

It should always be understood when viewing an FDC, that streamflow behaves the way it is illustrated in the hydrograph that is given in Chapter 9 and 10. Although FDCs have a long and rich history in hydrology, they are sometimes criticized because, traditionally, their interpretation depends on the particular period of record on which they are based. If one considers n individual FDCs, each corresponding to one of the individual n years of record, then one may treat those n annual FDCs in much the same way one treats a sequence of annual maximum streamflow series (AMS). Viewed in that context, the FDC becomes a generalization of the distribution of daily streamflow where the distribution of annual maximum flood flows and annual minimum low flows are simply special cases drawn from either end of the complete annual-based FDC. This is in clarity discussed in Chapter 14 where flood frequency analysis for the above fourteen stations AMS flows are considered. In the present estimator (shown in Figure 5.1 and Appendix 5b) yields a slightly smoother and more representative FDC than the piecewise linear empirical cumulative histogram advocated by Searcy (1959) and others, even for a large sample such as this one.

Table 5.5: Flow duration curves using four plotting position formulae for Monsoon (July-Sept) using thirteen GD sited in Mahanadi.

GD sites	Coefficients in $Q = A (P)^B$											
	Weibull			Blom's			Gringorton			Cunnane		
	A	B	R^2	A	B	R^2	A	B	R^2	A	B	R^2
Gomlai	12121	-0.71	0.917	9338.7	-0.653	0.8962	8855.7	-0.6395	.8909	9153.5	-0.64	.894
Kesinga	13725	-0.93	0.902	9873.3	-0.846	0.8869	9232.7	-0.8229	.8299	9626.5	-0.84	.885
Altuma	1800	-1.07	0.976	1248.1	-0.979	0.966	1157.7	-0.9605	.9629	1213.2	-0.97	.964
Kantamal	54933	-0.94	0.946	39211	-0.860	0.928	36603	-0.8427	.9237	38206	-0.86	.926
Jaraikela	6142	-0.77	0.943	4673	-0.708	0.9283	4420	-0.693	.9243	4576	-0.70	.926
Sukma	1055	-0.85	0.974	786.09	-0.78	0.9622	740.22	-0.7648	.9588	768.46	-0.77	.960
Jenapur	11763	-0.60	0.921	9424	-0.54	0.898	9008	-0.53	0.892	9265	-0.54	.896
Sundergrh	4437	-0.78	0.943	3368	-0.71	0.928	3184	-0.70	0.924	3298	-0.70	.926
Talcher	5204	-0.14	0.906	4473	-0.37	0.882	4338	-0.36	0.876	4422	-0.37	.880
Salebhata	7865	-1.12	0.935	5260	-1.02	0.917	4846	-1.00	0.912	5100	-1.01	.915
Tikarpara	41033	-0.65	0.910	32239	-0.59	0.885	30692	-0.58	0.879	31646	-0.58	.883
Champua	774.8	-0.62	0.950	624.5	-0.565	0.9369	597	-0.5546	0.933	614.27	-0.56	0.93
Pamposh	774.8	-0.62	0.950	10306	-0.72	0.903	97317	-0.706	0.898	10085	-0.71	0.90
Anandpur												

Table 5.6: Flow duration curves using four plotting position formulae for Non-monsoon (Oct-June) using thirteen GD sited in Mahanadi.

GD sites	Coefficients in $Q = A (P)^B$											
	Weibull			Blom's			Gringorton			Cunnane		
	A	B	R^2	A	B	R^2	A	B	R^2	A	B	R^2
Gomlai	1472	-0.84	0.979	1101.8	-0.753	0.9645	1038.5	-0.7403	.9604	1077.5	-0.74	.963
Kesinga	1414	-1.05	0.986	981.9	-0.960	0.9735	911.31	-0.9419	.9698	954.66	-0.95	.972
Altuma	75.14	-0.68	0.931	75.747	-0.703	0.9085	71.562	-0.6889	.9034	74.141	-0.69	.906
Kantamal	2700	-0.65	0.935	2110	-0.596	0.910	2119.8	-0.575	0.927	2872.0	-0.89	.968
Jaraikela	2777	-0.65	0.955	2210	-0.596	0.940	2109.8	-0.584	0.937	2172	-0.59	.939
Sukma	162.5	-0.84	0.954	133.23	-0.845	0.9669	124.81	-0.828	0.963	129.99	-0.83	.965
Jenapur	926.5	-0.42	0.970	798.1	-0.38	0.953	774.2	-0.37	0.948	789.0	-0.38	.951
Sundergar	120.5	-0.61	0.994	98.1	-0.56	0.989	94.06	-0.55	0.987	96.56	-0.56	.988
Talcher	793.0	-0.30	0.951	711.0	-0.28	0.935	695.4	-0.27	0.930	705.1	-0.27	.933
Salebhata	222.3	-1.00	0.965	336.2	-1.10	0.980	229.0	-1.0	0.966	211.7	-0.98	.962
Tikarpara	1979	-0.34	0.990	1779	-0.31	0.979	1737	-0.30	0.976	1763	-0.30	.978
Champua	40.90	-0.32	0.979	36.69	-0.298	0.9734	35.887	-0.2927	0.971	36.385	-0.29	0.97
Pamposh	673.5	-0.59	0.983	551.68	-0.544	0.9767	529.62	-0.534	0.974	543.26	-0.50	0.97
Anandpur												

One could argue that both curves in Figure 5.1 (a) and (d) are almost equivalent, but that the estimator Q(>P) has the advantage of being easily implemented on a computer and leads to significantly smoother quantile functions than the traditional small samples of ten or fifteen years data record. The significant differences between the period-of-record FDC and either the mean or median annual FDC occurs because the period-of record FDC is highly sensitive to the

hydrologic extremes associated with the particular period of record chosen, whereas the mean and median annual FDCs are not nearly as sensitive. This effect is explored in Figure 5.3 (Appendix 5b) for *Altuma* GD site with fifteen years period of- record using two different periods of record. Figure 5.2 illustrates how sensitive the lower tail of an FDC can be to the chosen period of record. The period of record 1989-2001 contains the drought that was more drought experienced, hence the FDC for these two periods are significantly different.

5.4.3 Application to Fourteen Catchment Data in Mahanadi River

Table 5.5 and 5.6 shows the relationship between the catchment area and the mean flow of different thirteen gauging sites of Mahanadi basin. For this work, the average of the mean flow of the monsoon period is considered. The flow duration curve for various return periods is calculated by four methods i.e. (i) Weibull (ii) Blom (iii) Gringorton and (iv) Cunnane plotting formulae at monthly time intervals. Flow duration curve for each water year was constructed by plotting and arranging the daily discharge values in descending order. The characteristics of probability distribution for daily discharge values extracted at 5% of the time in a given N year data are graphically presented by plotting magnitude versus probability.

The moving-day FDC values will move towards the mean flow value, resulting in a curve with a flatter slope than the 1-day FDC. This is analogous to increasing the permeability of the catchment when considering the synthetic curves shown in Figure 5.1. It was postulated that, if the FDCs for an individual gauging station with long records are plotted using a range of moving-day averages, their envelope should take the form of Figure 5.1. If the envelope shape does not conform, t then further investigation is needed to check whether there are data or calibration errors.

Errors in stage data are most noticeable at the two ends of the FDCs, especially in the curves formed from the smaller moving day averages. These FDCs tend to deviate considerably from the general shapes of Figures. 5.1-5.5. Loss of record can result in the logger recording a value of zero and hence periods can be assigned zero flow rather than being omitted. This will result in a discontinuity at the lower end of the curve. Overestimation of the stage will result in an extremely large flow and the affected moving averages will deviate from the general curve at its low probability end. Incorrect calibration affecting complete ranges of flows will cause distortion to the envelope and steep gradients in each of FDCs. The location of the distortion indicates the range over which calibration is in doubt. Upward distortion of the envelope is caused by overestimation of the flow and downward by underestimation. The change from good calibration to bad usually occurs slowly and the effect increases the further into the poor range the percentage exceeded goes. The change from poor to good calibration is usually evidenced by a steep change in the gradient of the FDC. The envelopes from flow duration curves for less than six years tend to be less smooth and are the result of variations in annual flows and the lack of gauging from which to calibrate the section.

Thus, visual inspection of the envelope of FDCs formed by using moving-day averaged flow data derived from the records for a station can indicate errors in the records of stage or problems with the calibration of the stage–discharge relationship for the station. The rating equations are average ones, as the effects of hysteresis have been ignored due to the practicalities of gauging in the field and of developing a computer program which can differentiate between rising and falling stage.

5.5 APPLICATION OF FDCS IN FLOW ASSESSMENT

In this study fourteen stations (given in Table 5.2) were considered for developing flow duration curve. The catchments investigated ranged in size from 3.4 to 1844 km^2 and were

predominantly rural, 40 with less than 2.5% urbanization, seven with between 2.5 and 10% and four with more than 10%. The four heavily urbanized catchments are all small (> 20 km2) with one exception, basin encompassing the Champua GD site (62 km^2). The reason is that the referred stations recorded flows on daily basis or sometimes monthly scale for reasonable period i.e. 25-30 years (Table 5.2). This becomes easy to develop FDC using four methods as described above and simple to analyze the flow pattern.

The flow duration curve for various return periods were developed using the characteristics of distribution of probability plots of stream calculated by four methods i.e. (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formula at daily time intervals from 0 to 100 percent given in the Appendix 5 (a). For stations of the fourteen gauged basins having several years of record, a flow duration curve can be constructed using mean values over periods longer than one day i.e. for ten day and monthly mean values. For this case mean daily flows data is used in developing a standard FDC that are independent. Here the daily flows for this referred period (July-Sept) is first averaged for a single years say 1972, and a mean is taken for this mean daily flow over the available period (1972-2009). The case initially considered here is for the data of Anandpur GD station given in Table 5.2. The table shows the ranking of the mean flows during monsoon period.

5.5.1 FDC Using Normalized Stream flow

The moving-day FDC values frequently move towards the mean flow value, resulting in a curve with a flatter slope than the 1-day FDC. The log of normalized streamflow is taken here for Anandpur GD site record (as give in Table 5.2 in Appendix 5a) versus percent exceedence is shown in Figure 5.2 (a)-(d) (Appendix 5b) by the fitted trend line. Based on the shape of this

curve, a logarithmic function appears to be most appropriate on to choose to represent it. The chosen function is given below.

$y = a (P) + b$, where $y = \log (Q)$ (5.4)

P = percent exceedence

and a is constant parameter for P, and b is a constant controlling the slope of the FDC. The fit to the observed data is shown by the solid line in Figure 5.6 a-c using data ar *Anandpur* GD site.

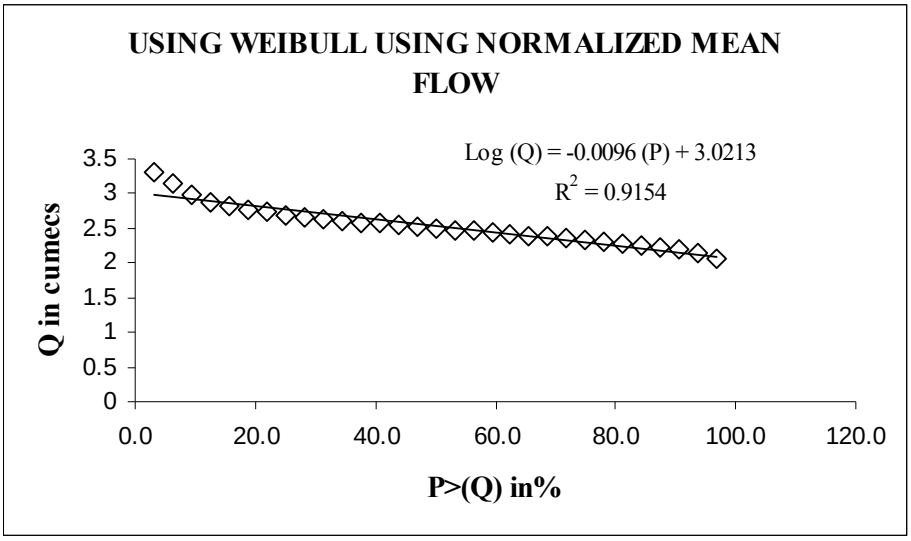

Figure 5.6 a Flow Duration Curve using normalized mean monsoon flow at *Anandpur* GD site (1972-2009)

It can be seen that in general, the two- parameter logarithmic FDC model produces a very good fit to the observed flow duration curve, with all r-squared values exceeding 0.9. However, an examination of the goodness of fit of the models is illuminating. Figures show the relationship between the cease to flow point and the goodness of fit (R^2). It will be seen that more perennial catchments (flowing more than 80% of the time) all have R^2 values greater than 0.91. However, more ephemeral catchments (flowing less than 80% of the time) have R^2 values ranging from 0.944 to 0.994, with most of them having poorer fits than the more perennial catchments. This

indicates that the procedure described here for using the FDC to define the hydrologic response of a catchment are more likely to work in humid, perennial catchments than in arid ephemeral catchments. As many of the sub-catchments of the Mahandi basin (region) examined in this study are in the sub-humid and rainy (during monsoon) tropics, it suggests that this technique will be more applicable in the more humid parts and possibly the wet tropics.

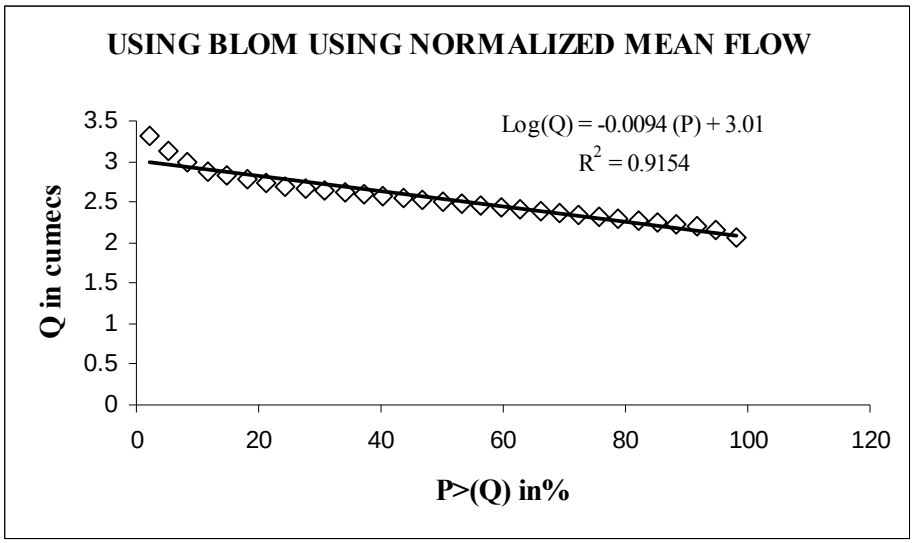

Figure 5.6 b Flow Duration Curve using normalized mean monsoon flow at *Anandpur* GD site (1972-2009)

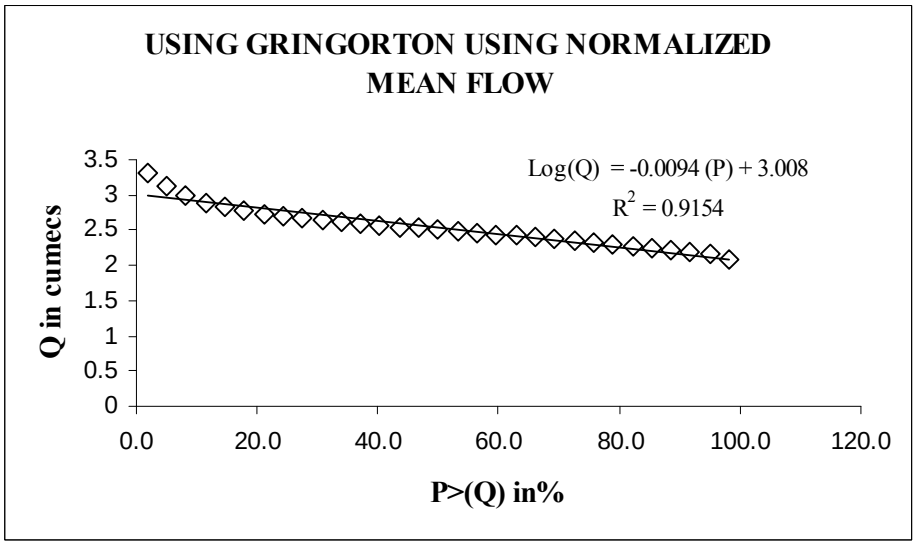

Figure 5.6 b Flow Duration Curve using normalized mean monsoon flow at *Anandpur* GD site (1972-2009)

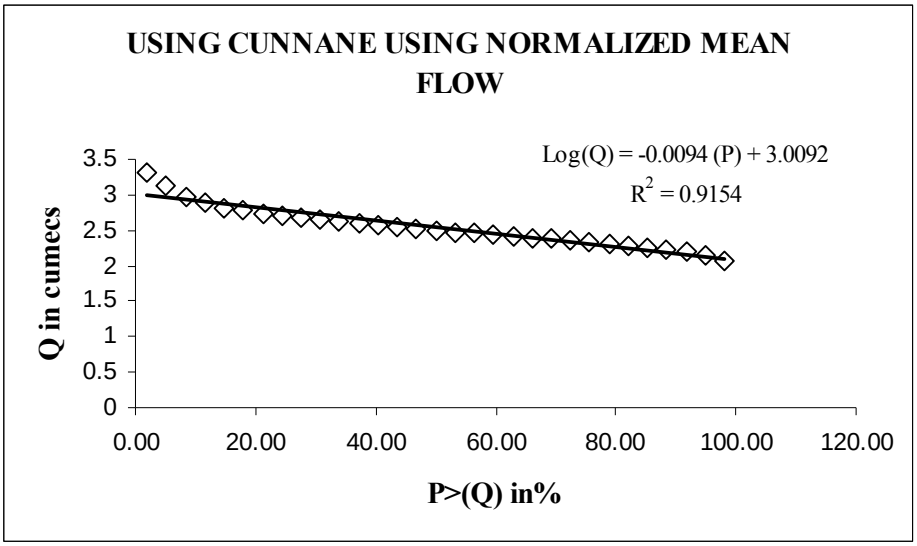

Figure 5.6 c Flow Duration Curve using normalized mean monsoon flow at *Anandpur* GD site (1972-2009)

The FDC during monsoon period at *Champua* GD side is developed using mean flow for the available period 1990 -2009. Figure 5.3.3 (a)-(d) (referred in Appendix 5 c) shows FDC using Weibull, Blom Gringorton and Cunnane using normalized mean flow, respectively. Figures show a clear picture of the variation of the river flow at this site during the monsoon period and also the probability of flow exceedance in X-axis and corresponding flow magnitude at the site. Flow duration curve using normalized mean monsoon flow at *Champua* GD site (1990-2009) has fit indexed with R^2 near 0.8815 for monsoon and 0.9608 for non-monsoon.

A flow-duration curve (FDC) represents the relationship between the magnitude and frequency of monthly streamflow for a particular river basin, providing an estimate of the percentage of time a given streamflow was equaled or exceeded over a historical period..

Figure 5.3.3 shows a variation of the river flow at this site during the monsoon period and also the probability of a particulate flow exceedance during the monsoon period. Inspecting FDC developed using all the four methods, there is a ninety seven percent probability where the flow exceeds 150 cumecs at the GD site. Similarly, flow exceeding 300 cumecs has a probability of fifty percent and around ten percent probability when flow exceeded 750 cumecs. The highest flow value given in Figure 5.3.3 is 350 cumecs. Figure 5.3.6 illustrates the disparity of the river flow at this site during the monsoon period and the probability of a particulate flow exceedance during the monsoon period. Inspecting the FDC developed using all the above mentioned four methods, there is a ninety seven percent probability where the flow exceeded 20 cumecs at the GD site. Similarly, flow values exceeding 50 cumecs & 150 cumecs has the probability of fifty percent and ten percent. In Jaraikela, the values of R^2 coefficient during Monsoon and Non-monsoon periods are recorded as 0.8293 & 0.663 (Figure 5.3.7).

Figure 5.3.9 & 5.3.10 demonstrates the FDC & probability of a particulate flow exceedance for *Janepur* gauging site during the monsoon & non-monsoon periods. The FDC was calculated d using above mentioned four methods. The flow value exceeding 800 cumecs & 4500 cumecs depicted the probability of fifty percent & ten percent. The highest flow value given in Figure 5.3.10 is 4500 cumecs. Similarly, Figure 5.3.11 to 5.3.18 explains the FDC & probability of a particulate flow exceedance values for Kantamal, Kesinga, Salebhata, Sudergarh, Talcher and Tikerpada sites during the monsoon and non-monsoon periods.

5.5.2 Data errors

Flow duration curve analysis identifies intervals, which can be used as a general indicator of hydrologic conditions e.g. wet (monsoon period) versus dry (non-monsoon) and severity like floods during monsoon and draught during summer. Flow duration curve intervals

can be grouped into several broad categories or zones that provide additional insight about conditions and patterns associated with the impairment. A common way to look at the duration curve is by dividing it into five zones. The high zone is centered at the 5th percentile, while the low zone is centered at the 95th percentile, and to be easier ranges can be adjusted, depending on local hydrology issues being addressed. Although five zones are commonly used to derive additional information from FDCs, the number of zones and range of frequency values are decided based on local hydrologic conditions.

Frequencies of the order of Q(95), the range of flow values in any one of the different moving day averages is small. For this reason, normally only the smaller duration moving-day averages are affected by errors such as short periods of loss of record (zero flow). For example, errors in 30 separate days spread out throughout the record. For example, in this study the GD site named *Kantamal* during the period 1976-1978 (May-June) is (range : maximum-9.7 and minimum 0.9 m^3/s) not connected as one or two blocks which have reduced the flow from say 0.2 m^3 /s to zero will affect 30 -day values by 0.2 m^3/ s-1, thirty 7-day values by 0.029 m^3/ s and thirty 30-day values by 0.003 m^3 /s. Thus effects are becoming increasingly small, so that the larger moving-day averages can absorb the errors. There are various ways of checking the data, but in this instance the minimum flow values were suspect i.e. many values were zero (no flow). The quickest way to check was by using the annual flow summaries. If any of the three different moving-day averages were below 0.35m^3/s, the flow values and the days on which they occur were noted and the computer stage plots were compared against the autographic chart. (If there are no charts, then the computer plot must be inspected in greater detail to see if there are any anomalies.) When the errors were found and corrected, the FDCs were re-calculated to see if the error had been eliminated. Figure 5.3.11 (Appendix 5c) for *Kantamal* and FDC is seen in Appendix 5b. shows the corrected FDC envelope.

5.5.3 Calibration errors

Hydrologic simulation combines the physical characteristics of the watershed and the observed meteorologic data series to produce the simulated hydrologic response. All watersheds have similar hydrologic components, but they are generally present in different combinations; thus different hydrologic responses occur on individual watersheds. It simulates runoff from four components: surface runoff from impervious areas directly connected to the channel network, surface runoff from pervious areas, and interflow from pervious areas, and groundwater flow. Since the historic streamflow is not divided into these four units, the relative relationship among these components must be inferred from the examination of many events over several years of continuous simulation. A complete hydrologic calibration involves a successive examination of the following four characteristics of the watershed hydrology, in the following order: (1) annual water balance, (2) seasonal and monthly flow volumes, (3) base flow, and (4) storm events. Simulated and observed values for each characteristic are examined and critical parameters are adjusted to improve or attain acceptable levels of agreement (discussed further below).

The station calibration at the high-flow end was checked and found to be poor, as the rating equation was overestimating the high flows. This part of the flow-duration was re-calibrated and the FDCs re-calculated. The resulting FDC envelope was found to be good and can be seen in Figure 5.3.15 for *Sukma* GD site. It should be noted that not all discontinuities in the FDC envelope are necessarily due to calibration errors. Irregularities occur in the range of 30–60% exceedence levels. This is not due to calibration problems but is caused by the flow regime in the river.

5.6 DEGREE OF CONFIDENCE

The flow data stored on a computer are unlikely ever to be totally correct, no matter how much quality control has been carried out. It may not be possible to calibrate the station throughout the whole range of flow, or the station may have only an interim rating equation. The FDC envelope will give an indication of how good the record quality is. If the envelope is not good at a particular point or range then the confidence in that part of the flow data is diminished. Irregularities occur below the 5% exceedance range. The data above this range appears good and can be used with confidence, should the end user have no access to the database to make the necessary corrections. Although there are minor fluctuations in the lower end of the FDC, the flow quality in this region of the envelope seems fairly good. Therefore, if the user is interested in the lower flow range in this study then flows below 3 m^3 /s (or 50% of the range), then the data could be used with reasonable confidence. However, it would not be advisable to use the data in the high range. Most data users do not have the facilities to correct the data even if they wanted to. However by using the FDCs in this way, they can check that the data quality is satisfactory for their use.

Consider the construction of an FDC or empirical quantile function from n observations of streamflow qi' where i = 1,..., n. If the streamflows are ranked, then the set of order statistics q (i)' where i = 1,..., n, results where q(l) is the largest and q(n) is the smallest observation. Even before the era of computers, quantiles and associated FDCs could be estimated from one or two order statistics. For example, the simplest empirical quantile function, or quantile estimator, is obtained from a single order statistic using

$$Qp,l = q(i) \text{ if } i = [(n + l)p] \quad (5.5a)$$

$$Qp.l = q(i+l) \text{ if } i < [(n + l)p] \text{ (2b)} \quad (5.5b)$$

where the quantity in brackets [(n + l)p] denotes the integer component of (n + l)p that is always less than or equal to (n + l)p. We recommend setting the smallest possible observation q(n+l) equal to zero, the natural minimum for streamflow. If the observations are not bounded above, the estimator Qp.l is undefined for values ofp that lead to i = [(n + l) p] = 0, since q(O) = 00. Essentially, Qp.l is equivalent to plotting the ordered observations q(i) versus an estimate of their plotting positions Pi' where Pi = i/(n + 1) is an estimate of the exceedance probability p in (1) known as the Weibull plotting position. The Weibull plotting position provides an unbiased estimate of 1 -FQ (q), regardless of the underlying probability distribution from which streamflows arise.

The estimator Qp.l yields a slightly smoother and more representative FDC than the Piecewise linear empirical cumulative histogram advocated by Searcy (1959) and others, even for a large sample such as this one. One could argue that both curves in Figure 5.2 a and 5.3 a (for *Gomlai* and *Kesinga* data using Weibull plotting) are almost equivalent, but that the estimator Qp.l has the advantage of being easily implemented on a computer and leads to significantly smoother quantile functions than the traditional ogive for small samples.The length of data required for the development of flow duration curves depend upon the type of scheme, type of development and variability of inputs. All these above discussion are implemented in Chapter 12 titled *estimation of confidence intervals and uncertainty of predictions*.

5.7. REMARKS

Flow duration curve reflects the nature of streams and gives an idea about the dependable flows at the gauging site over a particular period for planning the various water resources projects. In this section, daily data of fourteen gauging sites have been analyzed for a

range of about 35 years (1972-2009) period. In this case the fitting of the mean monthly runoff (Q) at gauging site at a time step and the probability of exceedences (p %) plotted using four empirical relations viz. Weibull, Blom, Gringorton and Cunnane that are useful for estimating and interpreting FDCs. The FDCs are constructed using non-linear power fitting to provide a generalized description of hydrologic frequency analysis using average recurrence intervals. Though the logarithmic transform on both sides of the fitted curve may be made linearized as commonly used in the past.

The fitting of the curve (FDC) to the plotted position points may be adjudged by COD(R^2) –measure values. These are discussed in details for all the study area catchments. On examining in at micro scale, the GD site on Altuma exceeding 0.9764 (Weibull), 0.9661 (Blom), 0.9629 (Gringorton) and 0.9649 (Cunnane) respectively. This is a good fit as $R^2 > 0.9$. There is a probability of more than ninety seven percent is that a flow of around 300 cumecs exceeds the GD site during the monsoon period. Similarly a flow exceeding 1000 cumecs has a probability of fifty percent and 3000 cumecs flow exceedance comes to around ten percent probability. Similarly for non-monsoon of Altuma there is a probability of more than ninety seven percent is that a flow of around 39 cumecs ,84 cumecs has a probability of fifty percent and 326 cumecs flow exceedance comes to around ten percent probability for non-monsoon period. The curve showed a steep slope throughout that denotes a highly variable stream with flow largely from direct runoff. The slope of the lower end of the duration curve shows a flat slope which indicates a large amount of storage; and at the upper end is a steep slope that indicates a negligible amount. Examination from Table 5.5 for monsoon periods, it indicate that the lower tail of FDCs during monsoon period for Anandpur GD site are highly sensitive to the 1987-1992 period of record used Non-Monsoon (Oct-June) season (1990-09) for *Altuma* GD Site (table 5.6). Further, the FDC is more sensitive to parameter B on both the tails (upper and lower) in

case of *Kesinga, Kantamal, Sukma* and *Salebhata*. In the referred sites B > 0.8 and in case of *Kesinga and Kantamal* both parameters A and B are high making it more sensitive to these values. For the non-monsoon period FDC parameters is more sensitive in two sites i.e. *Kesinga,* and *Salebhata*.

In this case or rather in many cases in earlier studies on this topic, an unbiased estimate of the expected probability of exceedance associated with the smallest observed average daily streamflow q (n) is p = n/(n + 1) = 365/366 = 0.99726. Though, the FDC envelope will give an indication of how good the record quantity is, however, if the envelope is not good at a particular point or range then the confidence in that part of the flow data is diminished. An example of this is the Flow exceeding p (%) using daily data of thirteen catchments in Mahanadi basin for Monsoon (July-Sept) is discussed in Table 12.4 and Figure 12.5 in Chapter 12 where the flow exceeding p (%) at both upper and lower tail of the FDC is examined for thirteen catchments and are related with catchment area considering data of monsoon and non-monsoon period.

In this case, irregularities occur below the 5% exceedance range. The data above this range appears good and can be used with confidence, should the end user have no access to the database to make the necessary corrections. Although there are minor fluctuations in the lower end of the FDC, the flow quality in this region of the envelope seems fairly good. Therefore, if the user is interested in the lower flow range, i.e. flows below 3 $m^3 s^{-1}$ (or 50% of the range), then the data could be used with reasonable confidence. However, it would not be advisable to use the data in the high range. Most data users do not have the facilities to correct the data even if they wanted to. However by using the FDCs in this way, they can check that the data quality is satisfactory for their use.

CHAPTER 6
DEVELOPMENT OF REGIONAL FLOW DURATION CURVES FOR UNGAUGED BASINS

6.0 INTRODUCTION

Assessment of the quantity of water in time and space is one of the most important aspects in planning, design and operation of the project. In order to ensure the success of a project, it is necessary to plan it such that desired quantity of water is available on most of the time. Necessary analysis is required to be carried out for identifying the characteristics of the flows which are essentially required in decision making process. The flow duration curve developed by analyzing the stream flow time series provides very useful information in this regard. The flow duration curve shows graphically the relationship between any given discharge and percentage of time that discharges is exceeded. The curve can be drawn from daily or monthly flow data or for any consecutive N day or month period. Thus the flow duration curve is simply the cumulative frequency distribution function of average stream flows occurring during a specified interval of time.

Methodology for the development of flow duration curve depends upon the availability of stream flow data at the site under consideration. In case the adequate records of stream flow are available at the site, the flow duration curves may be easily developed analyzing those records. However, for the Ungauged sites or sites having inadequate stream flow data the regional flow duration curves are developed and utilized for the estimation of available water resources. In this lecture, data requirements for the development of flow duration curve are discussed. Methodologies for the development of flow duration curves for gauged as well as Ungauged catchments are described. Some important characteristics of the flow duration curve are also described together with its use.

6.1 DEVELOPMENT OF REGIONAL FLOW DURATION CURVES

Most of the small catchments are ungauged. The methodology discussed for the development of flow duration curve for such catchments will not be the same as discussed under Section 3.1 for the gauged catchments. For ungauged catchments regional flow duration curves are developed based on the available rainfall-runoff records for the gauged catchments of the region which is considered to be hydro-meteorologically homogeneous. The regional flow duration curve is used to estimate the dependable flows for the ungauged catchment. Different methods for developing regional flow duration curve have been evolved by various investigators. Out of those three methods are outlines here under.

6.1.1 Method 1: Regionalization of the parameters of chosen probability distribution for individual gauged sites

As discussed under Section 5.1, the flow duration curve for the gauged catchments may be linearized fitting either normal distribution or log normal distribution. Theoretically any distribution may be fitted to linearize the flow duration curve and best fit distribution may be chosen for further application. Let the parameters of the fitted distribution are p_1, p_2.... p_n. Then these parameters may be estimated for the gauged catchments of the region and the regional relationships may be developed relating to the parameters with the physiographic and climatological characteristics. Step by step procedure involved in this method is given as follows:

(i) Identify the hydro-meteorological homogeneous region wherein the ungauged catchment is located.

(ii) Analyse the flow records of specific duration for each gauged catchment of the region fitting a chosen frequency distribution (say Normal or log normal distribution) and estimate the parameters of the distribution for each gauge catchments.

(iii) Evaluate the physiographic characteristics of the gauged as well as ungauged catchment from the toposheet.

(iv) Compute the basic statistics of the rainfall data such as mean, standard deviation, etc. for the different catchments of the region (gauged as well as ungauged) if the rainfall data are available.

(v) Develop the different relationships for the region using linear regression approach wherein the parameters of the distribution are considered as dependable variable (one at a time) and physiographic and/or climatological factors as independent variable(s).

(vi) Estimate the parameters of the selected distribution for the ungauged catchment utilising the relationships developed at step (v).

(vii) Develop the flow duration curve for the ungauged catchment substituting the values of the parameters in the cumulative density function of the chosen distribution.

For example, if the Normal distribution is fitted to develop the flow duration curve for gauged catchments of the region, the parameters $\overline{Q}$ and σ_Q may be related with their catchment characteristics. Typical forms of such equations for $\overline{Q}$ and σ_Q are given as:

$$\sigma_Q = C_2 + \frac{C_3}{H} \tag{6.1}$$

$$\overline{Q} = C_1 \log (A) \tag{6.2}$$

Where $\overline{Q}$ = Mean of the flow series (parameter of normal distribution; σ_Q =Standard deviation of the flow series (parameter of normal distribution); A = Catchment area, H = Basin Relief, and C_1, C_2, and C_3 = the constants.

These constants are evaluated using simple linear regression approach to the data of gauged catchments.

6.1.2 Method 2: Regionalization of the parameters of a chosen probability distribution derived for the region as a whole

Sometime, either adequate flow records for gauged catchments are not available or numbers of gauged catchments are limited. It makes the development of regional flow duration curve using method 1 erroneous. Under these circumstances it is considered appropriate to make all the flow data values of individual sites non-dimensional by dividing it by the mean of the flow occurring at that site. Now the non-dimensionalised flow data series of all the gauged sites are clubbed together to provide a simple series representing the region. While clubbing them together, it is presumed that these non-dimensional flow series for each gauging site in the region is a sample drawn from the same population. A single series, thus obtained, for the region may be analyzed and a chosen probability distribution may be fitted. It results in the parameters of the distribution which may be considered as the regional parameters. Now these parameters may be used to get the non-dimensionalised flow for any level of dependability. In order to develop the flow duration curve for an ungauged catchment, the estimate for mean flow is required. It may be multiplied by the non-dimensional flow values for having the required flow duration curve. It necessitates the development of a regional relationship for the mean flow relating the values of the mean flows of the gauged catchments with their catchment characteristics. The step by step procedure for this method is given as follows:

(i) Identify the hydro-meteorologically homogeneous region.

(ii) Compute the mean of the flow values ($\overline{Q}$) of the specific duration for each gauged catchment.

(iii) Divide the flow values (Q) for each gauging site by the mean ($\overline{Q}$) of the respective gauging sites.

(iv) Consider a single sample of non-dimensional ($Q/\overline{Q}$) combining these together for all the gauging sites.

(v) Fit a chosen distribution (say Normal or log normal distribution) to the sample obtained from step (iv) and estimate the parameters of the distribution.

(vi) Develop the relationship for mean flow relating the mean flow of the gauged catchments with their catchment characteristics using linear regression approach. A typical relationship for mean flow considering only catchment area as the catchment characteristic may be given as:

$$\overline{Q} = a\, A^{b} \qquad (6.3)$$

Where, $\overline{Q}$ = Mean flow for different sites; A = Catchment area of the respective sites;

a & b = Constants for the region to be evaluated using linear regression approach.

(vii) Develop the regional flow duration curve for $Q/\overline{Q}$ values substituting the parameters of the distribution obtained at step (v) in the cumulative density function of the fitted distribution.

(viii) Estimate the mean flow for the ungauged sites using the relationship developed at step (vi).

(ix) Develop the flow duration curves for the ungauged sites multiplying the ordinates of the regional flow duration curve of $Q/\overline{Q}$ values by the mean flow for the respective sites.

6.1.3 Method 3: Regionalization of the dependable flows

Some times one may be interested in the evaluation of dependable flows corresponding to the limited number of probability of exceedances (say only for 50% and 90% dependable flows). In such a situation the dependable flows itself may be regionalised rather than regionalising the flow duration curve. For the accuracy of this method adequate numbers of gauging sites having the flow series of the specific duration for a sufficiently long period are required. However, under the inadequate data situation either method 1 or method 2 may be used to regionalise the flow duration curve.

The step by step procedure for the regionalisation of the dependable flows by this method is given below:

(i) Identify the hydro-meteorological homogeneous region;

(ii) Develop the flow duration curve for each gauged catchment using the methodology discussed under Section 5.1.

(iii) Compute the desired dependable flows for each gauged catchments from the flow duration curves of the respective gauging site developed at step (ii).

(iv) Develop the regional relationship for each dependable flow relating the dependable flows of the gauged catchments with the pertinent catchment characteristics using linear regression approach.

The forms of the typical regional relationships for 50% and 90% dependable flows are given below:

$$Q_{50} = a_1 A^{a}_2 H^{a}_3 \quad (6.4)$$

$$Q_{90} = b_1 A^{b}_2 H^{b}_3 \quad (6.5)$$

Where, Q_{50} and Q_{90} are the 50% and 90% dependable flows respectively. A and H represent the catchment area and relief respectively. a_1, a_2, a_3, b_1, b_2, and b_3 are the constants to be evaluated by using linear regression approach.

6.2 RESULTS AND DISCUSSION

The results presented in this section used two seasonal i.e. monsoon (July-Sept) and Non-Monsoon (Oct-June) periods to construct regional flow-duration curves. Monthly stream flows for fourteen gauging sites with period of data as collected for from Central Water Commission is given in Table 6.1. The other geomorphological characteristics with salient statistical properties can be viewed in CWC (1982a, 1997) reports and Chapter 2 and 3.

Table 6.1 Hydrological Flow data for Mahanadi.

site Name	GD Site	District	Cat. Area (sq Km)	Mean flow (m^3/s)	Period
1	Anandpur	Keonjhar	8570	2070.26	1972-2009
2	Altuma	Dhenkanal	830	317.44	1990-2009
3	Champua	Keonjhar	1710	118.07	1990-2009
4	Gomlai	Sundargarh	21950	325.82	1979-2009
5	Jaraikela	Sundargarh	9160	72.85	1972-2009
6	Jenapur	Jaipur	33955	28.75	1980-2009
7	Pamposh	Sundargarh	19448	3591.31	2000-2009
8	Talcher	Angul	29750	836.12	1986-1996
9	Kantamal	Boudh	19600	298.85	1990-2009
10	Kesinga	Kalahandi	11960	1981.686	1979-2009
11	Salebhata	Bolangir	4650	340.6172	1973-2009
12	Sundargarh	Sundargarh	5870	113.1675	1972-2009
13	Sukuma	Bolangir	1365	4040.29	1989-2003
14	Tikerpada	Bolangir		1231.21	2000-2009

As observed, the working period of most gauging sites starts in the middle ‘70 since only few stations give runoff data previous to this year. This fact suggested the opportunity to limit the analysis to the 30-years period ranging from 1975 to 2007 in case of daily/monthly flows of thirteen gauging sites maintained by CWC . For these basins the mean daily stream flows is not

available, however, the values recorded are annual maximum series (AMS) and the record length is 33 years while the mean sample size is about 20 years. These data are used in regional flood frequency analysis for return period flood estimation along with fourteen basins as mentioned earlier. The table 6.1 represents the total geographical characteristic of Mahanadi basin.

6.2.1 Regional Flow Duration Curves for Monsoon Period

The regional flow duration curve (RFDC) shows graphically the relationship between any given discharge or flow and percentage of time that discharge is exceeded. The curve is drawn from daily flow data or for any consecutive N month period. Regional flow duration curve of each year may be constructed by plotting and arranging the monthly daily discharge values in descending order as per their ranking i.e. maximum value of runoff has a rank 1 and the minimum in the last position. Thus the RFDC is simply the cumulative frequency distribution function of average stream flows occurring during a specified interval of time. Mean annual daily river discharge data can be summarized in the form of a regional flow duration curve which relates flow to the percentage of the time that it is exceeded in the record.

In this study fourteen stations (given in Table 6.1) were considered for developing regional flow duration curve. The catchments investigated ranged in size from 3.4 to 1844 km^2 and were predominantly rural, 40 with less than 2.5% urbanization, seven with between 2.5 and 10% and four with more than 10%. The four heavily urbanized catchments are all small (> 20 km^2) with one exception, the reason is that the referred stations recorded flows on daily basis or sometimes monthly scale for reasonable period i.e. 25-30 years (Table 6.1). This becomes easy to develop RFDC using methods as described above and simple to analyze the flow pattern.

The regional flow duration curve is plotted using flow on a logarithmic scale as the ordinate and percentage of time discharge exceeded on a probability scale as the abscissa.

Discharge or flow referred many times in this section i.e. Q corresponding to a particular exceedance frequency are symbolized as p(>Q) that gives the percent of the time the specified mean monsoon flow equaled or exceeded. In this chapter, the RFDC is done using mean monthly data for a monsoon period taken from July to September. The reason is that the maximum flow (more than average) in Mahanadi basin frequently occurs during this period i.e. July-September. In rest of the year, the flow frequently is less than the mean or sometimes too low. This shows the relationship between the cease to flow point and the goodness of fit measured by COD (R^2). The probability of flow excedance p(>Q) calculated using (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formulae.

The Excel function Percentile is used on the array of the total mean flow record to get the 1-98 percentile flows. The results are given in Table 6.2. The area under the flow-duration curve is a measure of the discharge available 100 percent of the time. Dividing the area by 100 (base of the curve 100 percent of the time) gives the average ordinate which, multiplied by the scale factor is the mean discharge. Similarly, the area under a portion of the curve, divided by the percent of time of that portion, represents the mean flow during the particular percent of time. This property of the curve has important applications in such studies, but finding the scale factor of a flow-duration curve plotted on logarithmic- probability paper is somewhat complicated.

The slope of the flow duration curve is calculated using the Excel function SLOPE ( ), and the R^2 value is calculated using the Excel function RSQ ( ). If the R^2 value is greater than 0.9 then the flow duration curve is suitably straight for this range of percentiles and the value of the slope and R^2 are written to the output file.

Table 6.2 Probability of flow excedance p (>Q) for the Mahanadi region. (Using mean monsoon flow for months of July-Sept taking 14 GD sites)

Mean Q in m^3/s	Rank	Probability of flow exceedence p(>Q) using			
		WEIBULL	BLOM	GRINGORTON	CUNNANE
4484.65	1	3.1	2	1.8	1.92
3768.85	2	6.3	5.2	5.0	5.13
3346.08	3	9.4	8.4	8.2	8.33
3027.82	4	12.5	11.6	11.4	11.54
2777.26	5	15.6	14.8	14.7	14.74
2205.35	6	18.8	18	17.9	17.95
1710.65	7	21.9	21.2	21.1	21.15
1627.57	8	25.0	24.4	24.3	24.36
1373.35	9	28.1	27.6	27.5	27.56
1311.47	10	31.3	30.8	30.7	30.77
1258.91	11	34.4	34	33.9	33.97
1207.54	12	37.5	37.2	37.1	37.18
1130.38	13	40.6	40.4	40.4	40.38
1088.30	14	43.8	43.6	43.6	43.59
1037.53	15	46.9	46.8	46.8	46.79
986.14	16	50.0	50	50.0	50.00
949.72	17	53.1	53.2	53.2	53.21
916.42	18	56.3	56.4	56.4	56.41
864.72	19	59.4	59.6	59.6	59.62
831.47	20	62.5	62.8	62.9	62.82
800.89	21	65.6	66	66.1	66.03
748.11	22	68.8	69.2	69.3	69.23
721.04	23	71.9	72.4	72.5	72.44
701.91	24	75.0	75.6	75.7	75.64
656.86	25	78.1	78.8	78.9	78.85
633.62	26	81.3	82	82.1	82.05
597.67	27	84.4	85.2	85.3	85.26
565.51	28	87.5	88.4	88.6	88.46
534.19	29	90.6	91.6	91.8	91.67
507.04	30	93.8	94.8	95.0	94.87
316.16	31	96.9	98	98.2	98.08

A flow duration curve is usually straight for this range of percentile values, but for higher and lower flows it usually becomes non-linear due to the effect of extreme flows. It is for this reason that only suitably linear central sections of the flow duration curve are considered. Inspecting RFDC developed using all the four methods, (Figure 6.2 a-d); the curve is slightly represented by a straight line through this central portion of the curve. The curve has a steep slope throughout that denotes a highly variable stream with flow largely from direct runoff. The slope of the lower end of the duration curve shows a flat slope which indicates a large amount of

storage; and at the upper end is a steep slope that indicates a negligible amount. When the equation is fitted to RFDC as per described in Method 2, it can be seen that the R^2 values exceeding 0.9382 (Weibull), 0.9182 (Blom), 0.913 (Gringorton) and 0.9163 (Cunnane) respectively. This is a good fit as $R^2 > 0.9$. However, an examination of the goodness of fit of the models is illuminating. It will be seen that more perennial catchments (flowing more than 80% of the time) all have R^2 values greater than 0.91. However, more ephemeral catchments (flowing less than 80% of the time) have R^2 values ranging from 0.9382 to 0.9163, with most of them having poorer fits than the more perennial catchments. More perennial catchments flows more than 80% of the time and has R^2 values greater than 0.974.

For examining the RFDC further, $R^2 = 0.9$ was set as a minimum for the points on this range of the curve. For $R^2 < 0.9$, the flow duration curve was not considered to be suitably linear and its slope was not recorded. From the figure, there is a probability of more than ninety seven percent is that a flow of around 300 cumecs exceeds the GD site during the monsoon period. Similarly a flow exceeding 1000 cumecs has a probability of fifty percent and 3000 cumecs flow exceedance comes to around ten percent probability (Vogel and Fennessey 1993). One could argue that both curves in Figure 6.1 (a) and (d) are almost equivalent, but that the estimator Q(>P) has the advantage of being easily implemented on a computer and leads to significantly smoother quantile functions than the traditional small samples of ten or fifteen years data record.

6.2.2 Regional Flow Duration Curves for Non-Monsoon Period

The regional flow duration curve were developed using four plotting formula methods i.e. (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton at monthly time intervals from 0 to 100 percent. For stations of the fourteen gauged basins having several years of record, a regional flow duration curve can be constructed using mean values of monthly mean values.

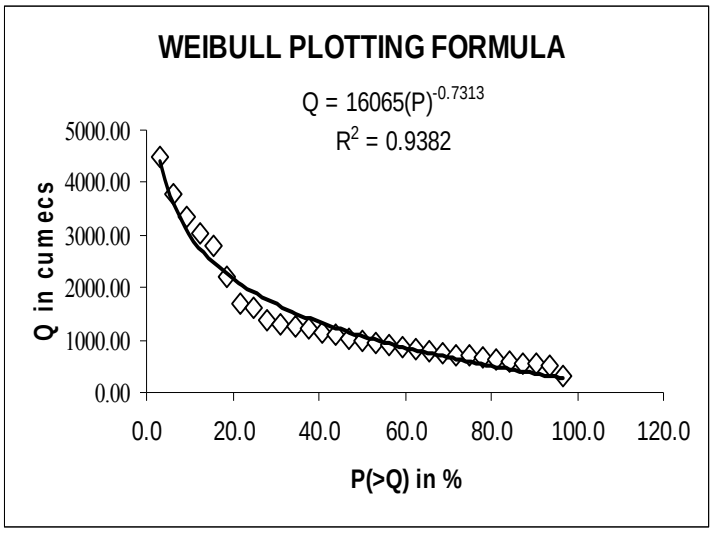

Figure 6.1 a. Regional Flow Duration Curve using average for Monsoon (July-Sept) season (1990 -2009)

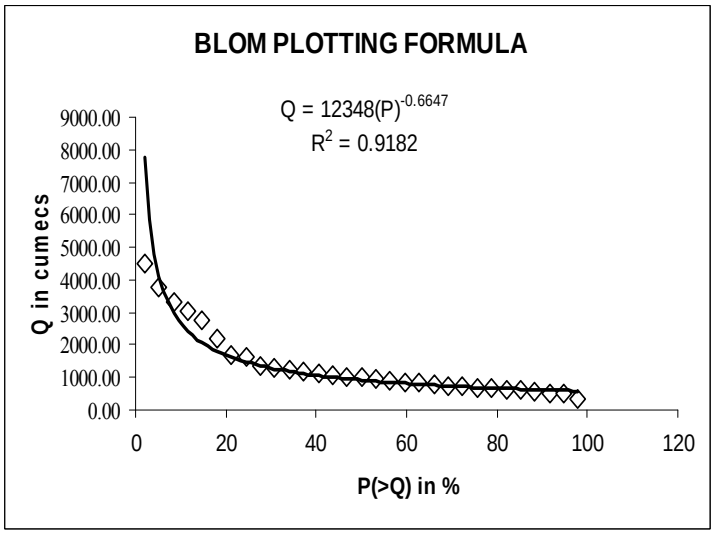

Figure 6.1 b. Regional Flow Duration Curve using average for Monsoon (July-Sept) season (1990 -2009)

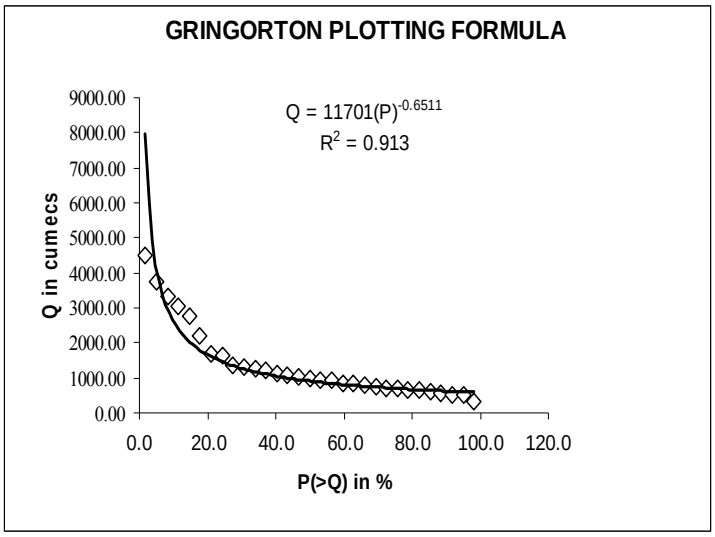

Figure 6.1 c. Regional Flow Duration Curve using average for Monsoon (July-Sept) season (1990 -2009)

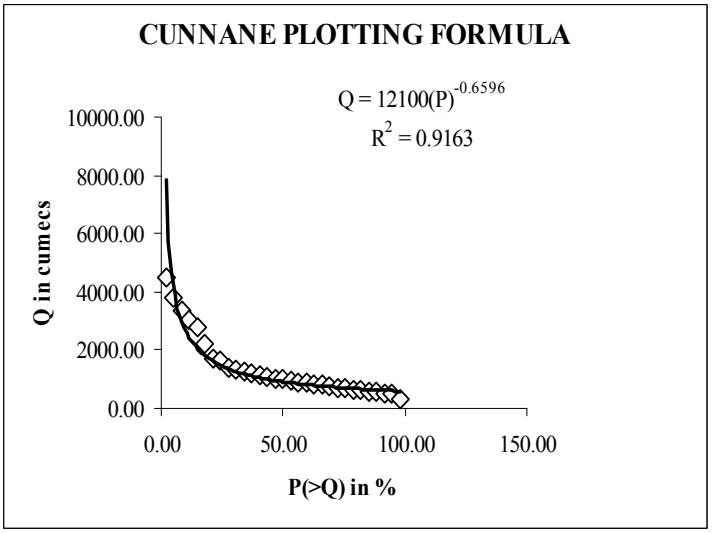

Figure 6.1 d. Regional Flow Duration Curve using average for Monsoon (July-Sept) season (1990 -2009)

For this case mean monthly data is used in developing a standard RFDC that are independent. Here the monthly mean flows for this referred period (Oct-June) is first averaged for a years, by taking the mean of all the fourteen gauging site in decreasing order.

RFDC developed using all the four methods, there is a probability of more than ninety seven percent is that a flow of around 69 cumecs exceeds the GD site during the non-monsoon period. Similarly a flow exceeding 154 cumecs has a probability of fifty percent and 326 cumecs flow exceedance comes to around ten percent probability. One could argue that both curves in Figure 6.1 (e) and (h) are almost equivalent, but that the estimator Q(>P) has the advantage of being easily implemented on a computer and leads to significantly smoother quantile functions than the traditional small samples of ten or fifteen years data record.

When the equation is fitted to RFDC as per described in Method 2, it can be seen that the R^2 values exceeding 0.9382 (Weibull), 0.9289 (Blom) 0.9201 (Gringorton) and 0.9225 (Cunnane) respectively for the non-monsoon period. The curve has a steep slope throughout that denotes a highly variable stream with flow largely from direct runoff. The slope of the lower end of the duration curve shows a flat slope which indicates a large amount of storage; and at the upper end is a steep slope that indicates a negligible amount.

6.3 RFDC Using Normalized Stream flow

The moving-day RFDC values frequently move towards the mean flow value, resulting in a curve with a flatter slope than the 1-day RFDC. The log of normalized streamflow is taken here for all GD site record (as give in Table 6.1 in Appendix 6) versus percent exceedance is shown in Figure 6.2 (a)-(d) by the fitted trend line. Based on the shape of this curve, a logarithmic function appears to be most appropriate on to choose to represent it. The chosen function is given below.

$y = a (P) + b$

Where:

$y = \log (Q)$

P = percent exceedence

And a is constant parameter for P, and b is a constant controlling the slope of the FDC

The fit to the observed data is shown by the solid line in Figure 6.2(a).

Table 6.3 Probability of flow excedance p (>Q) for the Mahanadi region. (Using mean non-monsoon flow for months of Oct-Nov taking 14 GD sites)

Mean Q in m^3/s	Rank	Probability of flow exceedence p(>Q) using			
		WEIBULL	BLOM	GRINGORTON	CUNNANE
510.72	1	3.1	2	1.8	1.92
370.40	2	6.3	5.2	5.0	5.13
325.42	3	9.4	8.4	8.2	8.33
298.73	4	12.5	11.6	11.4	11.54
275.32	5	15.6	14.8	14.7	14.74
260.17	6	18.8	18	17.9	17.95
248.04	7	21.9	21.2	21.1	21.15
238.10	8	25.0	24.4	24.3	24.36
228.16	9	28.1	27.6	27.5	27.56
203.10	10	31.3	30.8	30.7	30.77
193.23	11	34.4	34	33.9	33.97
187.27	12	37.5	37.2	37.1	37.18
182.38	13	40.6	40.4	40.4	40.38
177.36	14	43.8	43.6	43.6	43.59
172.82	15	46.9	46.8	46.8	46.79
153.59	16	50.0	50	50.0	50.00
149.72	17	53.1	53.2	53.2	53.21
143.81	18	56.3	56.4	56.4	56.41
138.93	19	59.4	59.6	59.6	59.62
136.83	20	62.5	62.8	62.9	62.82
133.91	21	65.6	66	66.1	66.03
131.72	22	68.8	69.2	69.3	69.23
127.97	23	71.9	72.4	72.5	72.44
125.89	24	75.0	75.6	75.7	75.64
122.63	25	78.1	78.8	78.9	78.85
120.58	26	81.3	82	82.1	82.05
116.85	27	84.4	85.2	85.3	85.26
114.53	28	87.5	88.4	88.6	88.46
110.19	29	90.6	91.6	91.8	91.67
106.98	30	93.8	94.8	95.0	94.87
68.38	31	96.9	98	98.2	98.08

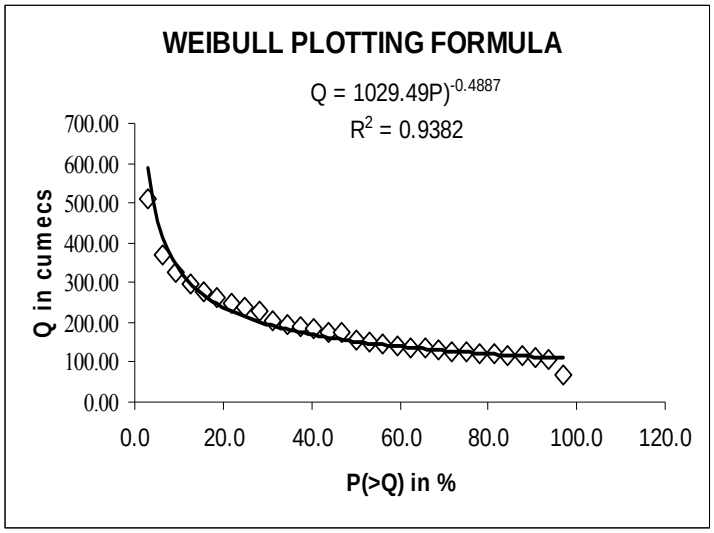

Figure 6.1 e. Regional Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09)

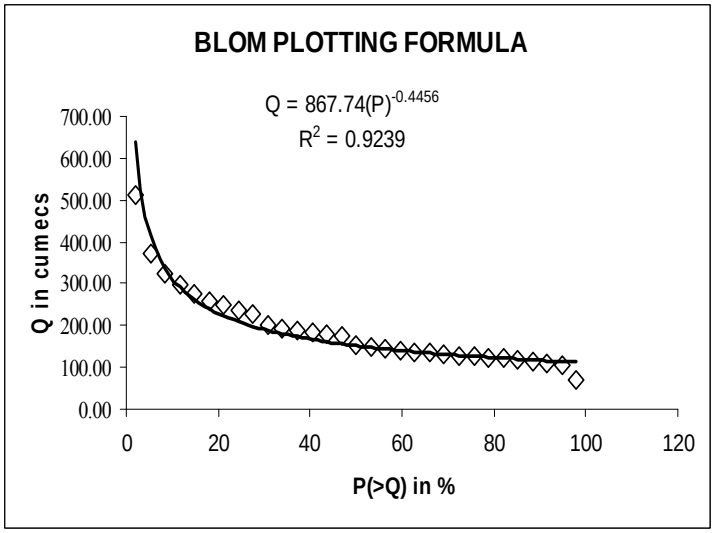

Figure 6.1 f. Regional Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09)

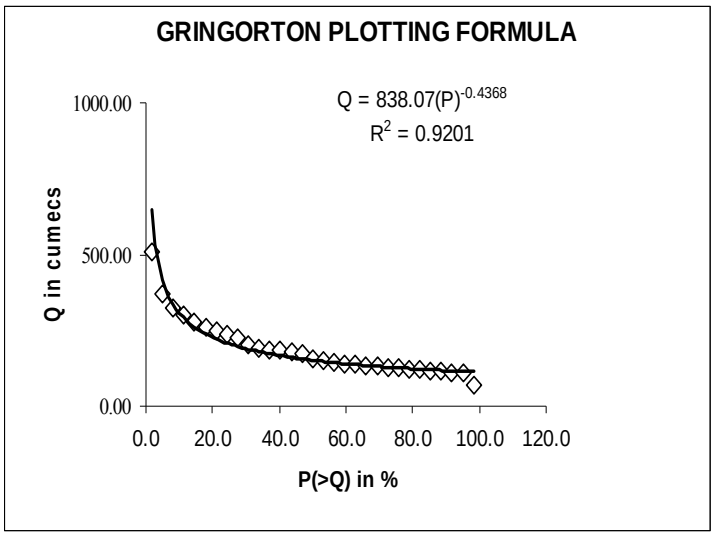

Figure 6.1 g. Regional Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09)

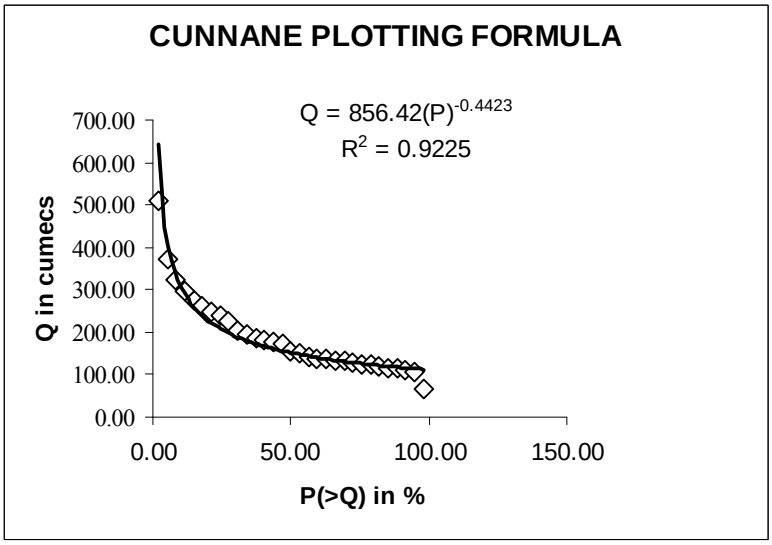

Figure 6.1 h Regional Flow Duration Curve using average for Non-Monsoon (Oct-June) season (1990-09)

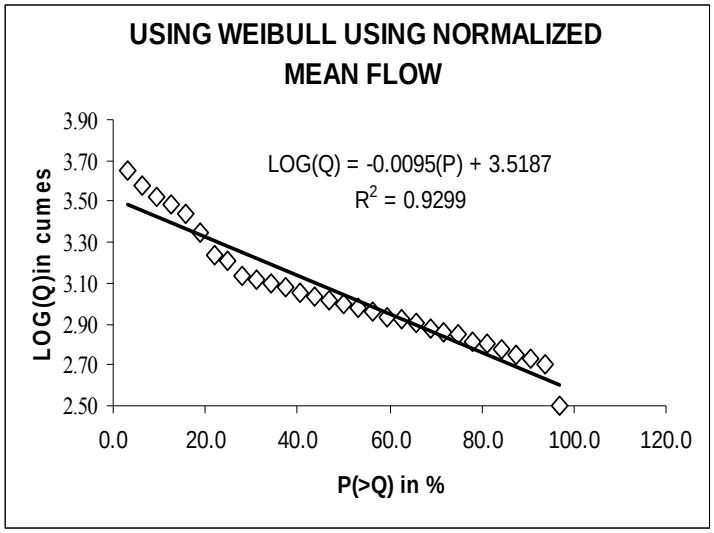

Figure 6.2 a Regional Flow Duration Curve using normalized mean for monsoon season (1972-2009)

It can be seen that in general, the 2- parameter logarithmic RFDC model produces a very good fit to the observed flow duration curve, with all r-squared values exceeding 0.9. However, an examination of the goodness of fit of the models is illuminating. The relationship between the cease to flow point and the goodness of fit (R^2). It will be seen that more perennial catchments (flowing more than 80% of the time) all have R^2 values greater than 0.91. However, more ephemeral catchments (flowing less than 80% of the time) have R^2 values ranging from 0.944 to 0.994, with most of them having poorer fits than the more perennial catchments. This indicates that the procedure described here for using the RFDC to define the hydrologic response of a catchment are more likely to work in humid, perennial catchments than in arid ephemeral catchments. As many of the sub-catchments of the Mahanadi basin (region) examined in this study are in the sub-humid and rainy (during monsoon) tropics, it suggests that this technique will be more applicable in the more humid parts and possibly the wet tropics.

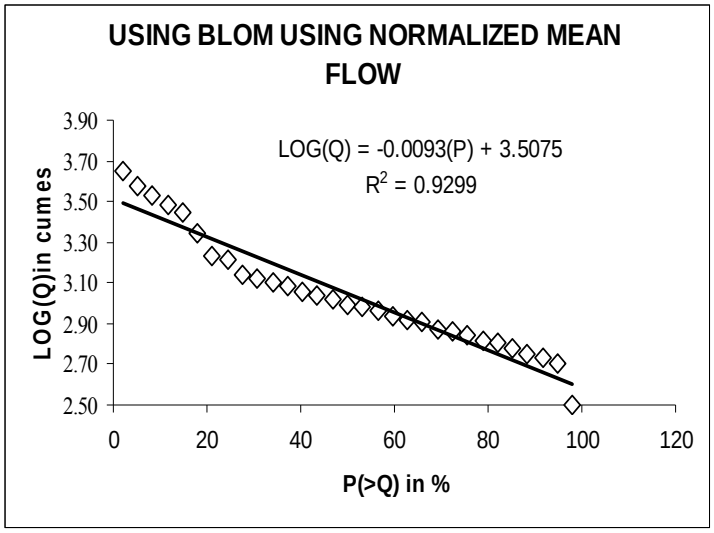

Figure 6.2 b Regional Flow Duration Curve using normalized mean for monsoon season (1972-2009)

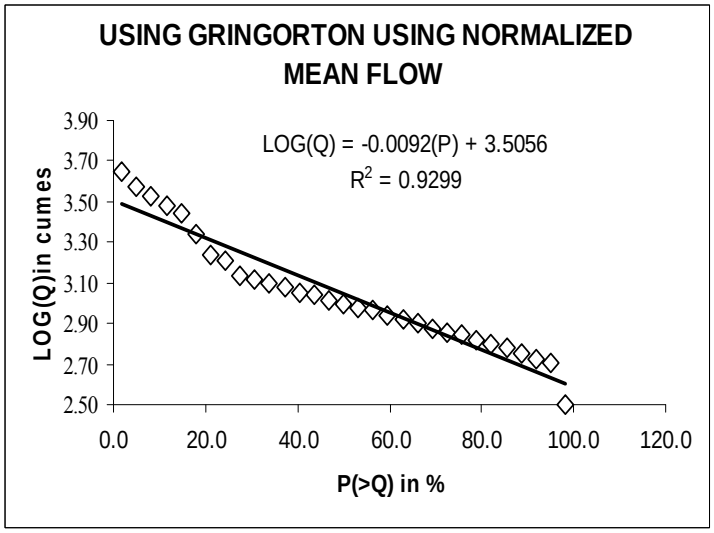

Figure 6.2 c Regional Flow Duration Curve using normalized mean for monsoon season (1972-2009)

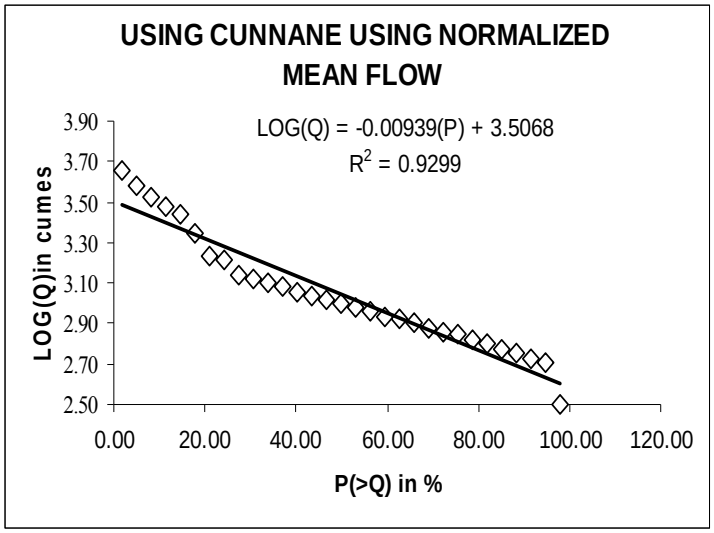

Figure 6.2 d Regional Flow Duration Curve using normalized mean for monsoon season (1972-2009)

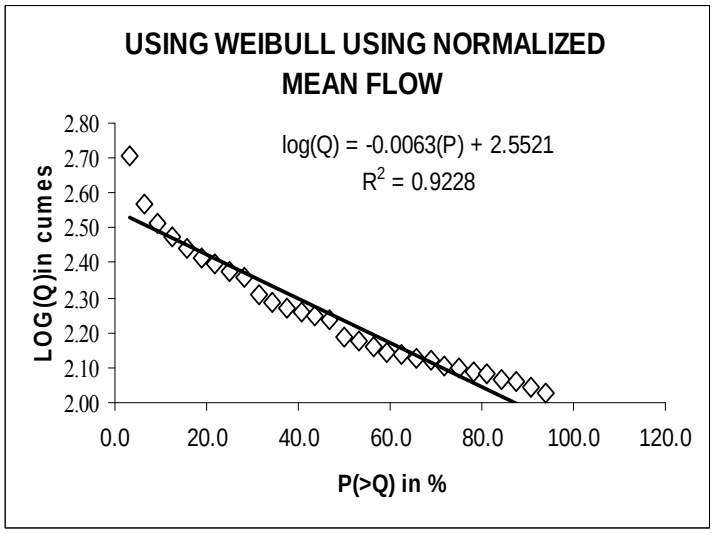

Figure 6.2 e. Regional Flow Duration Curve using normalized mean for non-monsoon (Oct-June) season (1972-2009)

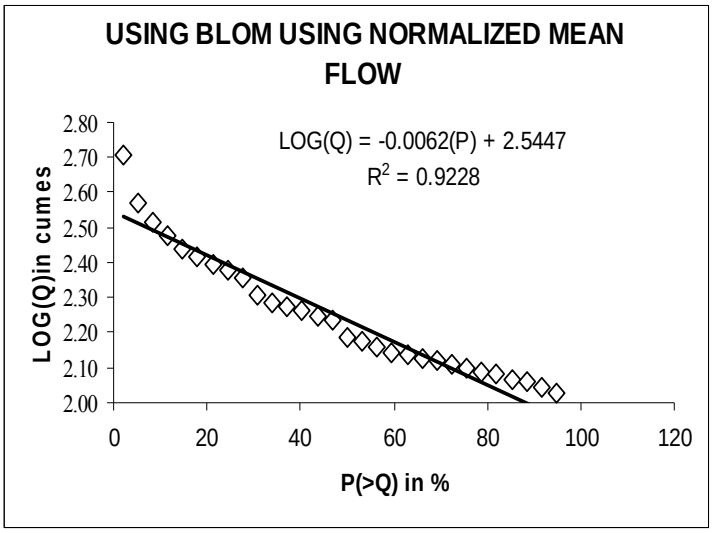

Figure 6.2 f. Regional Flow Duration Curve using normalized mean for non-monsoon (Oct-June) season (1972-2009)

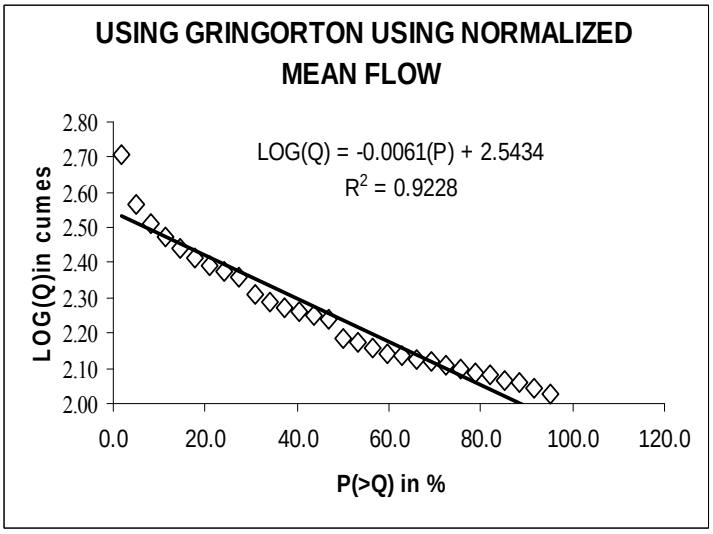

Figure 6.2 g. Regional Flow Duration Curve using normalized mean for non-monsoon (Oct-June) season (1972-2009)

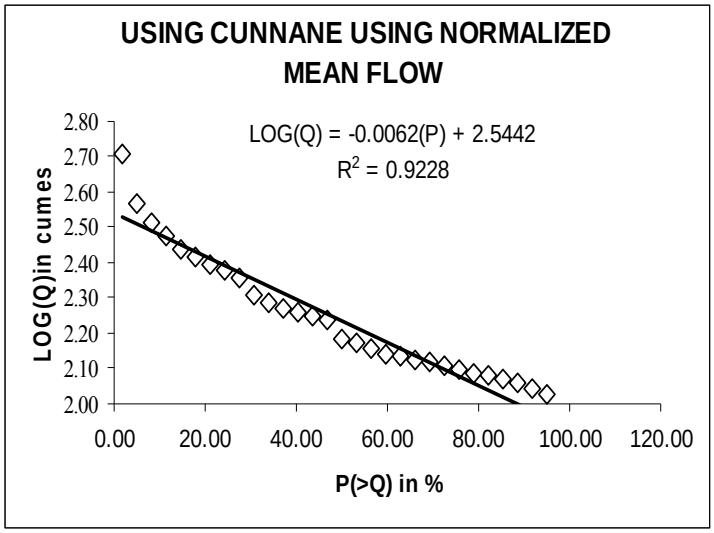

Figure 6.2 h. Regional Flow Duration Curve using normalized mean for non-monsoon (Oct-June) season (1972-2009)

The RFDC during monsoon period is developed using mean flow for the available period 1972 -2009. Figure 6.2(a)-(d) (referred in Appendix 6) shows RFDC using Weibull, Blom Gringorton and Cunnane uses normalized mean flow, respectively. Figures6.2 (a) show a clear picture of the variation of the river flow at this site during the monsoon period and also the probability of flow exceedance in X-axis and corresponding flow magnitude at the site. Regional Flow duration curve using normalized mean monsoon period (1972-2009) has fit indexed with R^2 near 0.93 for monsoon and 0.92 for non-monsoon period.

A Regional flow-duration curve (RFDC) represents the relationship between the magnitude and frequency of monthly streamflow for a particular river basin, providing an estimate of the percentage of time a given streamflow was equaled or exceeded over a historical period. Figure 6.2(a) shows a variation of the river flow at this site during the monsoon period and also the probability of a particulate flow exceedance during the monsoon period. Inspecting RFDC developed using all the four methods, there is a ninety seven percent

probability where the flow exceeds 150 cumecs at the GD site. Similarly, flow exceeding 300 cumecs has a probability of fifty percent and around ten percent probability when flow exceeded 750 cumecs.

Incorrect calibration affecting complete ranges of flows will cause distortion to the envelope and steep gradients in each of RFDCs. The location of the distortion indicates the range over which calibration is in doubt. Upward distortion of the envelope is caused by overestimation of the flow and downward by underestimation. The change from good calibration to bad usually occurs slowly and the effect increases the further into the poor range the percentage exceeded goes.

The change from poor to good calibration is usually evidenced by a steep change in the gradient of the RFDC. The envelopes from Regional flow duration curves for less than six years tend to be less smooth and are the result of variations in annual flows and the lack of gauging from which to calibrate the section. Instantaneous flow gauging is frequently implemented to gain insight into the flow characteristics of small ungauged catchments. This study proposed a combination approach in which regionally derived information on flow duration characteristics is complemented by on-site instantaneous flow data to improve the estimated FDC. The instantaneous flow samples are treated as a case of a censored data problem with an (unknown) upper threshold.

The lognormal distribution, valid for most of the observed daily flow records in the study region, was fitted to the data points and the mean annual flow was computed using the properties of the distribution under the assumption that the sampled median flow is a good approximation of the actual median flow provided that flow sampling is not restricted to base

flow conditions. The FDC for a typical small catchment in the study region having an adequate flow sampling record was estimated, yielding better results than other simpler drainage area methods, while in another examined case, where the number of samples was limited, results were not much different than those achieved by simpler methods. The procedure presented is not difficult to employ and the computations are simple. The estimation procedure has been found accurate only if a sample of sufficient size is available (n>25) which in addition should be approximately log normally distributed. The latter prerequisite is, however, expected to be usually fulfilled in the study region. Therefore a recommendation for flow sampling of small catchments in the study region would be the collection of a sample comprised of not less than 25 and up to 48 flow values, which, given a semi-monthly time-step, amounts to site monitoring twice per month for a period of 1–2 hydrological years. If small hydropower development is considered for example, a 2-year monitoring period (at most) is quite feasible since a similar period of time is usually needed for preliminary studies and investigations, licensing approvals and environmental permitting.

Further research is required, using instantaneous flow samples of different sampling frequency and size, possibly originating from catchments with different base flow characteristics within the study region. The incorporation of more sites and longer records in the regional FDC will allow the construction of FDCs applicable to different base flow regimes and will generally improve significantly the regional flow regime representation. The findings of this study are primarily of interest for the region under study; however, the results may be of use for other mountainous Mediterranean regions exhibiting the same climatic and flow regime characteristics.

6. 4 REMARKS AND CONCLUSION

Regional Flow duration curve reflects the nature of streams and very useful means for the estimation of available water for different dependabilities, primarily developed for catchments where the data measurements are short or ungauged for inaccessible or other reasons. Whenever adequate lengths of records are available, the regional flow duration may be developed analyzing the available flow records at the site. RFDC in this study focused on catchments in Mahandi basin and the relationships were developed using flow data for fourteen catchments. From the results, the R^2 values were closely to 0.9382 (Weibull), 0.9182 (Blom), 0.913 (Gringorton) and 0.9163 (Cunnane) respectively. These are believed to be a good fit as $R^2 > 0.9$, and it is observed that the perennial catchments (flowing more than 80% of the time) have R^2 values greater than 0.91, and more ephemeral catchments (flowing less than 80% of the time) have R^2 values ranging from 0.9382 to 0.9163, with most of them having poorer fits than the more perennial catchments.. There is a probability of more than ninety seven percent in case of respective flows of around 300 cumecs during the monsoon period. Similarly a flow exceeding 1000 cumecs has a probability of fifty percent and 3000 cumecs flow exceedance comes to around ten percent probability. Similarly for non-monsoon period as taken in this study, there is a probability of more than ninety seven percent is that a flow of around 69 cumecs,154 cumecs has a probability of fifty percent and 326 cumecs flow exceedance comes to around ten percent probability. The curve showed a steep slope throughout that denotes a highly variable stream with flow largely from direct runoff. The slope of the lower end of the duration curve shows a flat slope which indicates a large amount of storage; and at the upper end is a steep slope that indicates a negligible amount. Stepwise multiple linear regression software may be utilized for developing the regional flow duration curves considering the different catchment characteristics as independent variables. Nevertheless, this points also has been pointed out in Chapter 12 titled uncertainty and confidence bands in FDC. In this referred section, the FDC for individual

catchments are averaged for two bands i.e. lower (nearly 5 % exceedance) and upper (nearly 97 % exceedance) and the corresponding runoff values are related with catchment area. Though the regional flow duration curves are generally used for planning the various water resources projects, but the optimum utilization of the project, therefore, depends on how much accurate the regional flow duration curve is developed.

.

CHAPTER 7
DESIGN STORM ANALYSIS FOR THE STUDY AREA

7.0 GENERAL

While planning the construction of medium and large dams for irrigation, hydropower generation, and/or flood control, it becomes necessary to design the structure to pass a certain magnitude of critical flood generally known as design flood. The design flood to be adopted is generally decided by the type of the hydrologic structure. To this end, the structures are classified as:

i) Large or medium structures such as dams,

ii) Medium structures like barrages and minor roads and railway bridges, and

iii) Small or minor structures like cross drainage works and minor irrigation tanks and minor road bridges.

According to the criteria laid down by Central Water Commission, New Delhi, the following types of spillway design flood are recommended for major and medium dams and other small structures.

Table 7.1 Recommendations for major and medium dams

Type of Dam	Capacity (Million m^3)	Hydraulic head (m)	Inflow design flood
Small	0.5 – 10	7.5 - 12	100 year
Intermediate	10 – 60	12 - 30	Standard Project Flood
Large	> 60	> 30	Probable Maximum Flood

For minor structures, a flood of 50 or 100 year frequency is adopted depending on the importance of the structure.

Since long term runoff data are generally not available, rainfall data of longer period are used to estimate the 'design storm' for the estimation of design flood using a suitable rainfall-runoff model. Thus, rainstorm analysis is a first step in the design storm estimation procedure.

7.1 DESIGN STORM

The design storm or rainfall is the critical rainfall hyetograph. It is developed for the design of specific type of structures described above. It has following three components:

i) Rainfall amount,

ii) Arial distribution of rainfall, and

iii) Time distribution of rainfall.

These are generally described with the aid of the following components:

(a) Scale of design storm

The scale of design storm is linked with the scale of the design flood which, in turn, depends on the size of the structure and the risk associated with it.

(b) Types of design storm

The basic types of design storm include those derived from statistical and physical approaches. Statistical approach involves frequency analysis and the physical approach involves the use of a physical model. Generally, two types of storms are considered for design purposes. They are the 'Probable Maximum Precipitation (PMP)' or 'Maximum Probable Storm' and 'Standard Project Storm (SPS)'.

Probable Maximum Precipitation (PMP): For a given duration, PMP is theoretically the greatest depth of precipitation that is physically possible over a given area at a particular geographical location at a certain time of the year.

Standard Project Storm: It is the rainstorm that is reasonably capable of occurring in the region of problem basin. It is generally the most severe rainstorm that has occurred in the region during the period of available records.

7.1.1 Duration of design storm

The duration of design rainfall is determined considering the size of the drainage basin, duration of the flood, and the type of design structures. It is necessary that the selected duration be at least as long as the supply duration. Based on studies conducted on small watersheds in USA, it was concluded that the ratio of the volume of the rainstorm to peak discharge might be the most appropriate duration of design rainfall.

7.1.2 Storm Selection and Analysis

The first step in rainstorm analysis is a thorough understanding of the meteorology of large storms in the region of analysis. A judicious selection of a few storms which are representative of the whole catchment should be made. Before deciding the appropriate storm depth for design purposes, storm analysis based on either basin centered or storm centered is carried out based on the guidelines discussed below. To arrive at dependable estimates of design storm depths, it is necessary to make a judicious selection of pertinent storms over the river catchment or sub-catchment. To this end, it is essential to be aware of the size and location of the catchment under study, nature of the project envisaged, etc. The preliminary selection of the storm is based on the review of relevant records such as:

i) Daily rainfall data,

ii) Depression/storm tracks,

iii) Recent and historical storm details, and

iv) Flood and discharge data.

The first step is to list all the rain periods of maximum rainstorms, such as 1 day, 2 day, 3 day, 4 day, and 5 day major rainstorms, for long term data of all raingauge stations. The next step is to fix up an appropriate threshold value for storm selection keeping in mind the following:

a) If the catchment lies in a semi-arid region, a lower depth of rainfall is adopted as threshold while a higher depth is taken for humid regions. In India, the threshold value of 25 cm is taken as appropriate for humid regions, and 5 cm for semi-arid regions.
b) For smaller catchment areas, a higher threshold value and for larger catchment areas, a smaller threshold value is used. An average catchment depth of 10 cm, for catchment areas up to 5000 sq. km, 5 to 8 cm for catchment areas between 5000 and 10000 sq. km, and 5 cm for catchment areas greater than 10000 sq. km may be appropriate.

For the storm periods listed above, the arithmetic average data of raingauges located within and around the basin are worked out and compared with the threshold value. All such storms whose daily depths equal or exceed the threshold value are considered for further analysis. The storms selected on the basis of the threshold criterion are then subjected to further analysis by preparing isohyetal maps. The map scale generally used for preparation of isohyetal maps is 1 cm = 10 km. It is customary to prepare the total storm depth maps of 1 day, 2 days, and 3 days depending on the storm duration and requirement of the design storm duration. In

some cases, where movement of the storm over elongated catchments is involved, daily isohyetal maps will be required to be prepared.

7.2 DEPTH-AREA-DURATION (DAD) ANALYSIS

The World Meteorological Organization (WMO) guide to hydro-meteorological practices (1971) and the report on operational hydrology (Rep. No. 1, 1973) recommend depth-area-duration (DAD) technique for storm analysis, described below in components.

a) Depth-area analysis

Data on precipitation volume of severe storms are important to examine and study storms suitable for design purpose. Such information is generally presented in the form of tables of maximum average depth of storm precipitation for various standard areas, such as 100 km^2, 500 km^2, etc. Such information is known as depth-area relation and it can be presented in tabular as well as graphical form.

b) Depth-duration analysis

The depth-duration analysis implies development of a relation between storm depth and its duration for a catchment. This relation can utilize both in-situ and transposed storms. The average depths of precipitation for different durations, such as 1 day, 2 days, and 3 days are computed similar to that of the isohyetal method. The depths thus determined for different durations constitute the depth-duration data.

Where a good network of recording raingauges is available, the depth-duration analysis can also be performed for shorter duration (1 hr, 3 hr, 6 hr, 12 hr, 18 hr, etc.) rainfall. Alternatively, the analysis can be performed by distributing the rainfall observed at non-

recording stations using the data of nearby representative recording raingauge station using the procedure described above.

c) Depth-area-duration analysis

The depth-area-duration (DAD) analysis is carried out for storms occurred over different regions during principal flood seasons. The DAD analysis aims at determining the largest (maximum) depth of precipitation amounts for various sizes (areas) during storm periods of 1 day, 2 days, and 3 days etc. durations. Such values when determined for each transposable storm provide the requisite information (data) for estimation of the design storm for a basin in question. The steps involved in the DAD analysis are given below:

i) The selected storm is assigned a definite beginning and ending with the help of available rainfall records of stations in the region. Ideally, the storm is assigned a value for duration starting from a period of no rain to the next period of no rain. This ensures that the storm totals from all stations are for exactly the same time interval. The storm analysis is then carried out for the period of the principal burst.

ii) For each day of the total storm period, rainfall values of each raingauge station in the storm region are listed in a tabular form.

iii) Separate isohyetal maps for each duration, viz., maximum 1 day, 2 days, 3 days, etc. are drawn.

Since the greatest rainfall of 1-day may not occur over all the areas under consideration, it becomes necessary to analyze for two alternative days, for establishing maximum 1-day rainfall for areas of all sizes. Therefore, a 2-day duration map is also prepared. After the maps

for each duration are prepared, the DAD analysis is carried out as follows using catchment boundary or last closed isohyets as boundary.

a) The isohyetal maps are divided into zones to represent principal rainfall centres. Starting with the central isohyets in each zone, the area encompassed by each isohyets is planimetered and net area between isohyets is determined.
b) In case of storm centered DAD analysis, the last closed isohyet's is taken as boundary and in case of catchment centered analysis, the catchment boundary is considered.
c) The average isohyetal value is multiplied by the area to compute the volume. The volumes are cumulated for every successive isohyetal range and then divided by the total area encompassed by particular isohyets to determine the maximum storm depth for the corresponding area.
d) The analysis at steps (a) through (d) is repeated for other durations, if adequate data of recording stations are available. Incremental maps can also be prepared for shorter durations such as 3, 6, 9, 12, and 18 hours. For further details, WMO manual (1969) of depth-area-duration analysis may be referred.

The following discuss about the procedure to calculated the *depth-area-duration analysis* for 1-day storm occurred on 30th August 1982 as per IMD data on Mahanadi region (refer Table 7.2). For the depth-area-duration analysis, an isohyetal map for given 1-day storm is prepared, and other computations are given in the following Table 7.2. In this table, Col. 1 presents the isohyetal value, Col. 2 presents the planimetered area between a isohyets and the catchment boundary which is converted to actual area in Col. 3. Col. 4 presents the net area between the two consecutive isohyets. Col. 5 shows the average depth of precipitation occurring between the two consecutive isohyets and its volume is shown in Col. 6. Col. 7 cumulates the rainfall volume

and Col. 8 presents the ratio of the value in Col. 7 to that of Col. 3. Thus, the values of Col. 3 and Col. 8 are the data of depth-area relation for 1-day duration.

Table 7.2: Depth-Area-Duration analysis for a typical 1-day storm of Aug. 30, 1982

Point Isohyetal value	Planimetered area (Sq. cm)	Area (Sq. km)	Net area (Sq. km)	Average depth (mm)	Incremental volume= 4 x 5	Cumulative volume (mm x Sq.km)	Average rainfall (mm) 7/3
1	2	3	4	5	6	7	8
488	-	-	-	-	-	-	488.0
450	13.4	1340	1340	469	628460	628460	469.0
400	38.1	3810	2470	425	1049750	1678210	440.5
300	91.2	9120	5310	350	1858500	3536710	387.8
200	177.1	17710	8590	250	2147500	5684210	321.0

d) DAD curve

The purpose of preparing DAD curves is to determine maximum rainfall depth in part or all of drainage basin in different duration rainfall. Rainfall recording stations record the amount of point rain fall and this point may or may not be the centre of rainfall. Therefore the maximum rainfall is often not recorded by the stations. Design rainfall values are generally expressed in the form of point rainfall values which is the rainfall depth at a location. In order to obtain real average values for an area, hydrologists and engineers require techniques whereby point rainfall amounts can be transformed to average rainfall amounts over a specified area. For this reason DAD curves are prepared in various time durations.

The total storm depth-area curve for each duration is constructed by plotting the average depth against the accumulated area. The starting point on the curve is taken as the central storm precipitation (highest value). Because of the sparse network of precipitation stations, there exists remote probability of stations' recording the highest point precipitation. In the study of most large area storms, it is considered reasonable to assume that the highest station precipitation represents the average depth over an appreciable area rather than the maximum point precipitation. The depth-area curves are plotted for different durations for each of the major rain

storms. Conventionally, a logarithmic scale is used for area, and linear scale for precipitation depths.

7.3 INTENSITY-DURATION-FREQUENCY RELATIONSHIP

Storm intensity decreases with increase in storm duration. Further, a storm of any given duration will have a larger intensity if its return period is large. In other words, for a storm of given duration, storms of higher intensity in that duration are rarer than storms of smaller intensity. In many design problems related to watershed management, such as runoff disposal and erosion control, it is necessary to know the rainfall intensities of different durations and different return periods. The interdependency between intensity i (cm/h), duration D (h), and return period T (years) is commonly expressed in a general form as

$$i = \frac{KT^x}{(D + a)^n} \qquad (7.1)$$

where K, x, a, and n are constants for a given catchment. Typical values of these constants for a few places in India are given in Table 7.3. These values are based on the reported studies of the Central Soil and Water Conservation Research and Training Institute, Dehradun, Extreme point rainfall values of different durations and return periods have been evaluated by India Meteorological Department and iso-pluvial (lines connecting equal depths of rainfall) maps covering the entire country have been prepared.

Table 7.3 Typical values of constants in Eq. 1

Place	K	x	a	n
Champua	6.93	0.189	0.50	0.878
Angul	11.45	0.156	1.25	1.032
Cuttack	5.82	0.160	0.40	0.750
Bisra	6.16	0.694	0.50	0.972
Raipur	4.68	0.139	0.15	0.928

These are available for rainfall durations of 15 min, 30 min, 45 min, 1 h, 3 h, 6 h, 9 h, 12 h, 15 h and for return periods of 2, 5, 10, 25, 50, and 100 years.

Table 7.4(a): 31-card daily rainfall data format from IMD figures.

								DAILY RAINFALL (.01 INCHES)											
CATCHMENT NUMBER	SUB-DIV. NUMBER	LATITUDE	LONGITUDE	STATION NUMBER	HEIGHT OF STATION IN TENGTHS OF FEET	YEAR	DATE	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC
3	5	9	13	15	19	23	25	29	33	37	41	45	49	53	57	61	65	69	73

24-card daily rainfall data format.

								2nd CARD																
AS IN Ist CARD								17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	MONTHLY TOTAL	
								Ist CARD											DAILY RAINFALL (0.1 mm)					
CATCHMENT NUMBER	LATITUDE	LONGITUDE	STATION NUMBER	BLANK	YEAR	MONTH	CARD NO.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	BLANK
3	5	7	9	10	12	14	15	19	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79	

24-card daily rainfall data format.

																2nd CARD			MAX IN 1 HR. DURATION					
AS IN 1st CARD						16 17	17 18	18 19	19 20	20 21	21 22	22 23	23 24	am t	Time	Hr mts								
														Ist CARD					HOURLY RAINFALL (0.1 mm)					
ELEMENT CODE	INDEX NO. OF STATION	YEAR (OMIT 19)	MONTH	DATE	CARD NO.	0 1	1 2	2 3	3 4	4 5	5 6	6 7	7 8	8 9	9 10	10 11	11 12	12 13	13 14	14 15	15 16			
	6	8	10	12	13	17	21	25	29	33	37	41	45	49	53	57	61	65	69	73	77			

Table 7.4(b) Hourly rainfall data of Cuttack and Champua SRRG stations

Hour	CUTTACK			CHAMPUA		
	321	456	682	565	443	693
1	0.0	0.0	0.0	0.0	0.0	23.5
2	0.0	0.0	0.0	0.0	0.0	30.5
3	0.0	0.0	0.0	0.0	0.3	10.2
4	0.0	0.0	0.0	0.0	0.5	30.8
5	0.0	0.0	0.0	0.0	0.1	16.9
6	0.0	0.0	0.0	0.0	1.0	1.9
7	3.0	1.7	0.0	0.0	24.4	0.6
8	0.0	9.0	0.0	0.0	0.2	3.1
9	0.0	2.3	0.5	0.0	0.0	5.3
10	0.9	0.0	0.4	0.0	0.0	3.8
11	0.0	1.5	0.1	0.0	23.8	0.8
12	0.0	5.5	0.0	0.0	1.3	1.5
13	0.0	0.0	0.0	0.0	0.2	0.0
14	0.0	0.0	0.0	0.0	0.0	0.3
15	0.0	0.0	0.0	0.0	0.1	0.9
16	0.0	2.5	0.3	0.0	0.6	1.5
17	0.2	1.5	0.4	0.0	29.0	0.4
18	2.1	0.8	0.3	0.0	11.0	0.5
19	0.7	3.4	0.0	25.7	7.1	0.1
20	0.3	0.6	0.0	0.0	0.4	0.5
21	0.1	2.5	0.0	0.1	2.7	0.7
22	0.1	0.0	0.0	1.0	6.8	0.0
23	4.7	0.0	0.0	6.0	20.6	0.0
24	0.0	0.0	0.0	1.3	0.0	0.0

Table 7.4(c) Rainfall for past 24 hours ending at 08:30hrs on 16.09.2008 (Unit in mm.)

Station	Rainfall	Station	Rainfall
Cuttack	37.8	Khairmal	15.1
Bhubaneshwar	50.6	Naraj	43.0
Hirakud	14.6	Puri	48.8
Kantamal	22.0	Salebhata	16.2
Kendrapara	75.0	Tikarpada	53.0
Kesinga	12.6	Sundergarh	31.6
Jenapur	59.8	Talcher	20.6
Altuma	18.0		

Table 7.5 Daily rainfall distributed into hourly rainfall

Hour	Champua	Talcher
1	40.0	33.5
2	66.0	53.5
3	85.5	73.5
4	98.0	84.5
5	105.0	101.5
6	119.9	112.5
7	139.4	128.0
8	150.9	135.8
9	160.0	140.3
10	164.7	156.7
11	169.1	167.7
12	172.6	183.2
13	180.4	183.2
14	189.5	189.5
15	191.0	194.0
16	192.0	197.5
17	192.9	214.0
18	202.5	229.5
19	205.0	234.0
20	208.4	234.7
21	209.9	236.3
22	210.3	236.3
23	210.4	236.4
24	210.4	236.3

7.4 Probable Maximum Precipitation (PMP)

The Probable Maximum Precipitation (PMP) is used for deriving the design flood. For a given duration, PMP is theoretically the greatest depth of precipitation that is physically possible over a given area at a particular geographical location at a certain time of the year. Both PMP and standard projected storm (SPS) for the study area were supplied by IMD.

Table 7.6 The Standard Project Storm (SPS) and Probable Maximum Precipitation (PMP) values for Mahanadi

Duration in days	SPS Values (cm)	PMP Values (cm)	Duration in hours.
1-day	210	256	24
2-day	347	423	24
3-day	386	471	24

7.5 DESIGN STORM

As per the IMD guidelines, the 1-day PMP value is increased by 15% to convert them to any 24-hour values and no clock hour corrections are applied for 2- day PMP (Table 7.6). The time distributions of 24-hour and 48-hour storm rainfall were also supplied by IMD, and are given at Table 7.6 (b). These values shall be taken as the design storm depths for estimating the design flood. . Storm Isohyetal map of Mahanadi and adjoining river basins for 1Day, 2Day and 3Day are given in Figure7 (b), Figure7(c) and Figure7 (d) (Refer Appendix 7).

7.5.1 Duration of design storm

The duration of design rainfall is determined considering the size of the drainage basin, duration of the flood, and the type of hydraulic structures. It is necessary that the selected duration be at least as long as the supply duration. As per the CWC (1993) guidelines, the design duration of PMP is adopted as follows:

(i) 2.5 to 3 times the time of concentration for fan-shaped catchments whose length to width ratio is less than 1.5;

(ii) equal to time lag of UH for an elongated catchment;

(iii) In respect of elongated catchments and/or large catchments (where distributed models are used for modeling the response of the catchments) the storm duration for causing the PMF is to be equivalent to 2.5 times the time of concentration.

Table 7.6 (b): Time Distribution coefficient for the 24 Hrs Storm.

Duration	24-hour storm
1	17
2	28
3	38
4	45
5	52
6	57
7	61
8	66
9	70
10	74
11	78
12	82
13	85
14	87
15	90
16	92
17	94
18	95
19	96
20	97
21	98
22	99
23	99
24	100

Most widely used empirical formula for computing time of concentration (t_c) was proposed by *Ramser* [1927] [quoted in *Haan et al*, 1984] as follows:

$$t_c = 0.02 L_c^{0.77} S^{-0.385} \qquad (7.2)$$

where t_c is the time of concentration in minutes, S denotes the average slope of channel in m/m, and L_c is the length of main stream from outlet to centroid of the catchment in meters. The slope is taken to be the average of highest and lowest contours given in the toposheet. Using the relationship of eq. (7.2), t_c can be calculated. Therefore, following these guidelines, the storm duration of the PMP for this study is adopted as 24 hours. In the next step, critical sequencing of hyetograph peaks shall be done to arrive at the design flood.

7.5.2 Critical sequencing of the PMP

The critical sequencing of PMP hyetograph is carried as per the recommended guidelines given in manual of CWC (2001). The detailed calculation of the sequencing for the Mahanadi zone is discussed in the following.

Step -1: The 1-day and 2-day PMP depths are 483mm, and 806mm respectively, the incremental rainfall depths on the individual days 1st, and 2nd day, are taken to be 483 mm and 323mm respectively. Note that no clock hour corrections are included here.

Table 7.7: Incremental 1- and 2-day PMP depths- Mahanadi

Duration	PMP in mm	Increment depths in mm	
1-day	256	1st Day	256
2-day	423	2nd Day	167

Step-2: The rainfall depths are arranged in 2 bells per day. The storm depths for 12 hour blocks are determined using the 12-hour to 24-hour ratio of 72% as given in the IMD data. The calculations are shown in Table 7.8.

Step-3: Using the distribution percentages given in Table 7.6(b), cumulative depths and incremental depths for the 12-hours blocks were obtained and are given in the Table 7(a) and 7(b)(refer Appendix 7).

Table 7.8: Incremental 12 hour storm distribution ratios for 1-, and 2-day PMP depths - Mahanadi

1st Day Depth= 483 mm	
Fraction of one day rainfall occurring in each block	
1st 12 hrs	2nd 12 hrs
0.71	0.29
Rainfall occurring in each block	
181.8mm	74.2 mm

Step-4: The above incremental depths are arranged in critical order for each bell separately. For this the largest of the increments is placed against the peak of UH, then the next largest against

the next largest of the UH ordinate and so on until all rainfall increments get arranged Table7(c) (Refer Appendix 7). Now the rainfall sequence is reversed to get the critical sequence as shown in Table 7(d) (Refer Appendix 7)

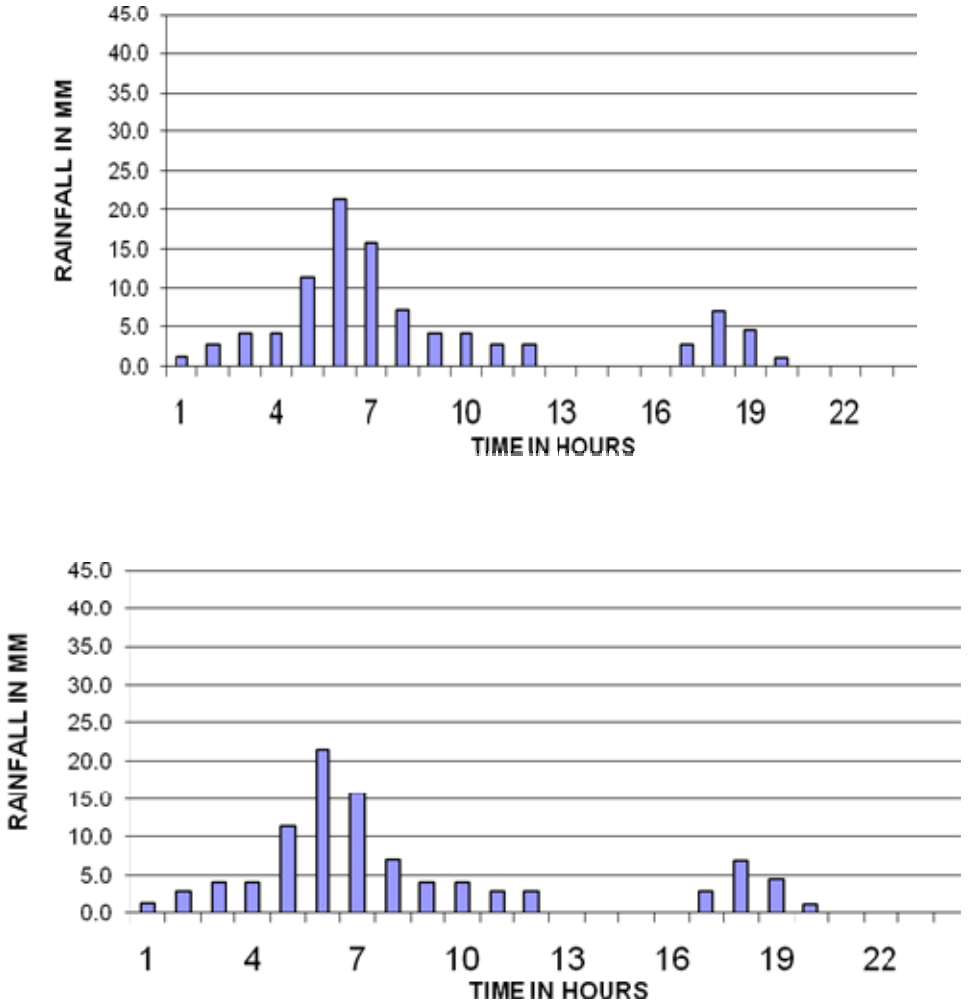

Figure 7.1 (a) 24-hours excess PMP hyetograph for Mahanadi zone

Step-5: Now, the 1st 12-hour block is interchanged with the 2nd 12-hour block to create the critical situation as shown in Table 7(e) (Refer Appendix 7).

The excess rainfall hyetograph of the PMP for Mahanadi zone is shown in Figure 6.6(a) and the values are reported in Table (f) (Refer Appendix 7).

CHAPTER-8
MODELING RAINFALL-RUNOFF PROCESS USING UNIT HYDROGRAPH APPROACHES

8.0 INTRODUCTION

When rain or snow falls onto the earth, most of it flows downhill as runoff, and a portion of the precipitation seeps into the ground to replenish groundwater. Surface runoff is that part of the precipitation which, during and immediately following a storm event, ultimately appears as flowing water in the drainage network of a watershed. An accurate understanding of hydrologic behavior of a watershed, which involves accurate estimation/ prediction of depth and rate of runoff, is the key parameter for effective watershed management.

Water yield refers to the volume of water available from a stream at a specified point over a specified period of time. Frequently, the point of determination is the watershed outlet (many times measured at a gauging site) and the period of time a day or longer. Here the emphasis is on the volume of flow rather than the instantaneous discharge. Therefore, the water yield is the integration of discharge, as a function of time, for a specified duration and reflects the volumetric relationship between rainfall and runoff. The total volume of runoff from a catchment is important when examining overall site water balances or storage capacities required for water supply dams etc, whereas, interception drainage, erosion protection, settling ponds and essential drainage infrastructure (e.g. culverts, spillways etc.) must all be carefully designed to suit the expected peak flow they are expected to experience. Design of structures controlling river flows considers hydrological extremes, viz., droughts and floods. The structural design includes sizing the capacity of outlet works, spillways, bypasses, and others, to cater floods. It requires quantitative information on both peak flows and time distribution of runoff or hydrographs. The unit hydrograph yields time distribution of runoff using rainfall excess. Thus, the estimation of flood hydrograph requires a unit hydrograph for the watershed.

The *unit hydrograph* theory proposed by Sherman in 1932 is primarily based on the principle of linearity and time and space invariance. For the gauged catchments, unit hydrographs can be derived from the analysis of available rainfall-runoff records. The procedures used to derive a unit hydrograph depend on the type of storm, whether a single-period storm or a multi-period storm. Thunder-storms generally being intense and of short duration are treated as single-period storms. On the other hand, frontal storms of usually longer duration are treated as multi-period storms. A unit hydrograph converts the rainfall (rainfall excess) to runoff (direct surface runoff). For its derivation, the duration of the unit hydrograph is generally taken as the duration of rainfall excess block. However, S-curve and superimposition methods can be utilized for deriving a unit hydrograph of desired duration. The former approach is more general than the latter, for the former changes the duration of available unit hydrograph to any degree. This chapter includes description of unit hydrograph theory, its assumptions, and limitations; derivation of unit hydrograph and factors influencing its derivation, instantaneous unit hydrograph (IUH), S-curve; change of unit hydrograph duration and derivation of average unit hydrograph.

The unit hydrograph derived from an event usually differs from that derived from another event. There are two possible approaches available for the derivation of the representative unit hydrograph for a watershed. The first approach averages the unit hydrographs derived for various events using conventional averaging method. The second approach, however, clubs first various events to one and then derives a representative single unit hydrograph for the watershed. In a nutshell, an *Unit hydrograph* is a *hydrograph of direct surface runoff* resulting *at a given location on a stream from a unit rainfall excess amount occurring in unit time uniformly over the catchment* area up to that location. The rainfall excess excludes losses (hydrological abstractions) from total rainfall and unit rainfall excess volume

equals 1 mm (may be considered as a standard value since rainfall is measured in mm). The selection of unit time depends on the duration of storm and size of the catchment area. For example, for small catchments, periods of 1 or 2 hours can be assumed and for larger catchments, 3, 4, 6, or even 12 hours can be adopted. Thus,

(i) A unit hydrograph is a flow hydrograph;

(ii) A unit hydrograph is a hydrograph of direct surface runoff (DSRO), not total runoff;

(iii) The hydrograph of surface runoff results from the rainfall excess;

(iv) The rainfall excess represents total rainfall minus losses (abstractions);

(v) During the unit time period, the rainfall excess is assumed to occur uniformly over the catchment;

(vi) Typical unit times used in unit hydrograph analyses are 1, 2, 3, 6, 8, and 12 hours. Time period is generally taken as an integer multiple of 24 hrs.

8.1 METHODS FOR DERIVING A UNIT HYDROGRAPH

A unit hydrograph can be interpreted as a multiplier that converts rainfall excess to direct surface runoff. The direct surface runoff (DSRO) is the streamflow hydrograph excluding base flow contribution. Since, a unit hydrograph depicts the time distribution of flows, its multiplying effect varies with time. In real-world application, the unit hydrograph is applied to each block of rainfall excess and the resulting hydrographs from each block are added for computing direct surface runoff hydrographs, to which base flows are further added to obtain total hydrographs. In general four methods are used for the derivation of the representative unit hydrograph for a watershed, and they are as follows:

(i) **Rainfall runoff series method:** Runoff is generated by rainstorms and its occurrence and quantity are dependent on the characteristics of the rainfall event, i.e. intensity, duration and distribution.

(ii) **Clark method:** The Clark method (1945) requires three parameters to calculate a unit hydrograph: Tc, the time of concentration for the basin, R, a storage coefficient, and a time-area curve. A time-area curve defines the cumulative area of the watershed contributing runoff to the sub catchment outlet as a function of time (expressed as a proportion of Tc).

(iii) **Synthetic method:** The synthetic unit hydrograph of *Snyder* is based on relationships found between three characteristics of a standard unit hydrograph and descriptors of basin morphology. The hydrograph characteristics are the effective rainfall duration, the peak direct runoff rate and the basin lag time.

(iv) **GIUH:** The GIUH approach has many advantages over the regionalization techniques as it avoids the requirement of flow data and computations for the neighboring gauged catchments in the region as well as updating of the parameters. Another advantage of this approach is the potential of deriving UH using only the information obtainable from topographic maps or remote sensing, possibly linked with GIS and digital elevation model (DEM).

(v) **SCS-Curve Number (CN) Method:** The SCS-CN method is based on the water balance equation and two fundamental hypotheses. The first hypothesis state that ratio of actual amount of direct runoff to maximum potential runoff is equal to the ratio of amount of actual infiltration to the amount of the potential maximum retention. For the given hydrologic group and various types of land use/ treatment and antecedent moisture conditions, CN- values are available. Knowing the CN,

the value of recharge capacity S is calculated and using this value of S, runoff is calculated.

(vi) Nash Model (1957): This is a conceptual model in which catchment impulse could be represented as the outflow obtained from routing the unit volume of the instantaneous rainfall-excess input through a series of *n* number of successive linear reservoirs having equal delay time. The output yields the ordinate of the instantaneous unit hydrograph. The model has two parameters viz, k (the storage coefficient) and n (number of reservoirs). The complete shape of the UH can be obtained by linking q_p and t_p of the GIUH with the scale (k) and shape (n) parameter of the Nash model.

8.2 FACTORS AFFECTING SURFACE RUNOFF

Surface runoff depends on a great number of factors, such as rainfall characteristics, watershed morphometric characteristics, soil physical characteristics (depth, texture, structure, and hydraulic conductivity), land cover, land use, and soil moisture conditions prior to rainfall events. Many factors affect water yield depending upon the period of its determination. Some of these factors are interdependent. These factors can be classified as (1) meteorological factors and (2) watershed factors. Space-time distribution of precipitation amount, intensity and duration, and space-time distribution of temperature are some of the most important meteorological factors. Some important watershed factors include surface vegetation, soil moisture, soil characteristics, surface topography, and drainage density. The physiographic features not only influence the occurrence and distribution of water resources within a watershed but these, particularly the orography, play a significant role in influencing rainfall and other climatic factors, such as temperature, humidity and wind. However, within a geographical location and physiographic framework, it is primarily the rainfall (its intensity, duration &

distribution) and the climatic factors affecting evapo-transpiration that determine the totality of water resources in the watershed.

8.2.1 Meteorological factors affecting runoff:

a) *Type of precipitation* (rain, snow, sleet, etc.): Types of precipitation have a great effect on the runoff. For example a precipitation which occurs in form of rainfall, starts immediately in from of surface flow over the land surface, depending upon its intensity as well as magnitude, while a precipitation which takes place in form of snow or hails the flow of water on ground surface will not take place immediately, but after melting of the same. During the time interval of their melting the melted water infiltrates into the soil and results a very little surface runoff generation.

b) *Rainfall intensity*: The intensity of rainfall has a dominating effect on runoff yield. If rainfall intensity is greater than infiltration rate of the soil the surface runoff takes place very shortly while in case of low intensity rainfall, where is found a reverse trend of the same. Thus high intensities rainfall yield higher runoff and vice-versa.

c) *Rainfall duration :* Rainfall duration is directly related to the volume of runoff due to the fact, that infiltration rate of the soil goeson decreasing with the duration of rainfall till it attains constant rate. As a result of this even a mild intensity rainfall lasting for longer duration may yield a coverderable amount of runoff.

d) *Distribution of rainfall over the watersheds:* Runoff them a water sheed depends very much on the distribution of rainfall. The rainfall distribution for this purpose can but expressed by a team "distribution coefficient which may be

defined as the ratio of maximum rainfall at a point to the mean rainfall of the watershed. For a given total rainfall, if all other conditions are the same, the greater the value of distribution coefficient, greater will be the peak runoff and vice - versa. However, for the same distribution coefficient, the peak runoff would be resulted from the storm, falling on the lower part of the basin i.e. near the outlet.

e) *Direction of storm movement*: The direction of prevailing wind, affected greatly the runoff flow. If the direction of prevailing wind is same as the drainage system then it has great influence on the resulting peak flow and also on the duration of surface flow, to reach at the outlet. A storm moving in the direction of stream slope produces a higher peak in shorter period of time than a storm moving in opposite direction.

f) Other meteorological and climatic conditions that affect evapotranspiration, such as temperature, wind, relative humidity, and season: The other climatic factors, such as temperature wind velocity, relative humidity, annual rainfall etc. affect the water losses from the watershed area to a great extent and thus the runoff is also affecter accordingly. If the losses are more the runoff will be less and vice - versa.

8.2.2 Physical characteristics affecting runoff:

a) Land use: Land use and land management practices have great effect on the runoff yield. E.g. an area with forest cover or thick layer of mulch of leaves and grasses contribute less runoff because water is absorbed more into soil.

b) *Vegetation: Vegetation* catches precipitation on its leaves, slowing it down or holding it for a while. This gives the ground more time to absorb the water, so there's less *runoff*.

c) *Soil type*: In filtration rate vary with type of soil. So runoff is great affected by soil type.

d) *Drainage area*: It is defined as the ratio of the total channel length [L] in the watershed to total watershed area [A]. Greater drainage density gives more runoff

Drainage density = L/A

e) *Basin shape*: Runoff is greatly affected by shape of watershed. Shape of watershed is generally expressed by the term "form factor" and "compactness coefficient".

Form Factor = Ratio of average width to axial length of watershed

= B/1 or A/1/1 = A/I2

f) Slope: It has complex effect. It controls the time of overland flow and time of concentration of rainfall. E.g. sloppy watershed results in greater runoff due to greater runoff velocity and vice-versa.

g) Topography: It includes those topographic features which affects the runoff. Undulate land has greater runoff than flat land because runoff water gets additional energy [velocity] due to slope and little time to infill rate.

h) Direction of orientation: This affects the evaporation and transpiration losses from the area. The north or south orientation, affects the time of melting of collected snow

i) Ponds, lakes, reservoirs, sink, etc. in the basin, which prevent or alter runoff from continuing downstream.

8.3 UNIT HYDROGRAPH THEORY AND ASSUMPTIONS

The unit hydrograph theory can be described using the following underlying propositions:

(a) Constant Base Length Proposition: The base length of direct surface runoff (t_b) corresponding to a rainfall amount of a given duration (T) is constant for a catchment and does not depend on the total runoff volume. For example, if T is the duration of rainfall excess, the base length of the DSRO for all rainfall events of this duration is the same for all events and is equal to t_b, as shown in Figure 8.1.

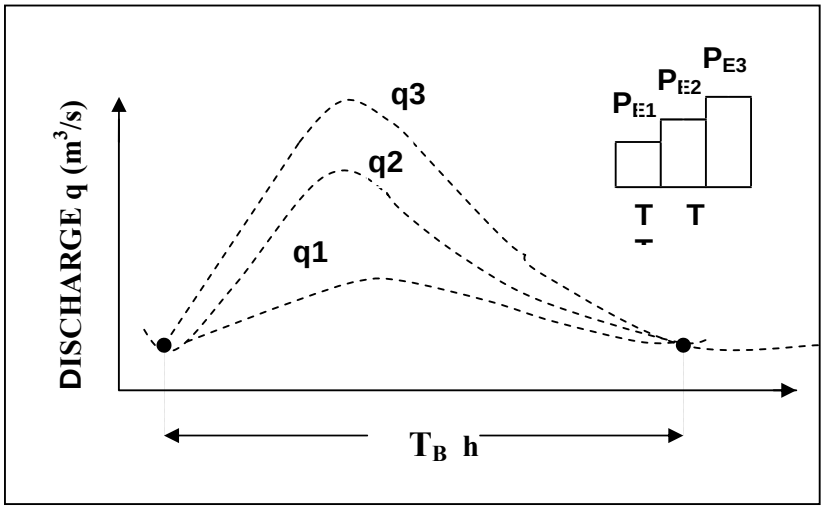

Figure 8.1 Three different unit hydrographs of same base but varying peak

(b) Proportional Ordinate Proposition: For storms of equal duration (unit periods), the ordinate of DSRO at a time is proportional to the total volume of rainfall excess. For example, the rainfall excess of an event, $P_{E1,}$ over time period T is equal to P_{e1} x t_p. It is to note that P with

$$\frac{q_1(t)}{q_2(t)} = \frac{P_{e1} xT}{P_{e2} xT} = \frac{P_{E1}}{P_{E2}}. \qquad (8.1)$$

subscript 'E' denotes volume in terms of depth and P with subscript 'e' denotes intensity of rainfall excess. Similarly, the rainfall excess of another event, P_{E2}, over time period T is equal to P_{e2} x T. Therefore, at a time t, the following proportionality holds as in the Eq. (8.1) given below:

Figure 8.1 shows the surface runoff hydrographs for three different rainfall excess volumes of 100 mm, 50 mm, and 25 mm of same unit period. The ratios of q100:q50:q25 = 100:50:25 = 4:2:1. Thus, the ordinates of the DSRO produced by 60 mm rainfall excess will be 0.6 times the available 100 mm unit hydrograph. Similarly, if the actual rainfall excess is 220 mm, the DSRO ordinates are 2.2 times the unit hydrograph ordinates and so on.

(c) Concurrent Flow Proposition: The hydrograph of surface runoff that results from a particular portion of storm rainfall is not affected by the concurrent runoff resulting from other portions of the storm. In other words, the total hydrograph of surface runoff is the sum of the surface runoff produced by the individual portions of the rainfall excess. Figure8.2 shows two hydrographs: one resulting from rainfall excess P_{E1} and the other from rainfall excess P_{E2}. The DSRO (q_1) from the first rainfall excess (P_{E1}) is not affected by the DSRO (q_2) from the second rainfall excess (P_{E2}). The total DSRO is, therefore, the sum of q_1 and q_2.

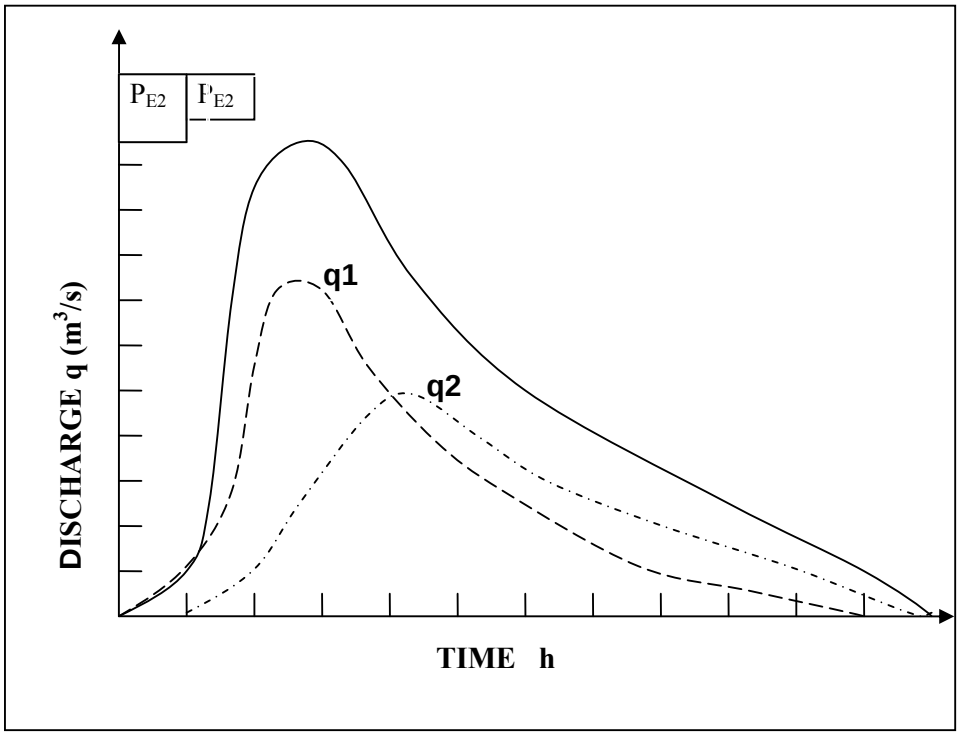

Figure 8.2 The total hydrograph resulting from consecutive storms.

For example of concurrent flow proposition, if A is the given hydrograph corresponding to a net rainfall depth of 5 cm of some duration. The hydrograph for a rain consisting of two consecutive periods, each of the same duration, but rainfall excess depths of 6 and 3 cms respectively, can be derived by summing the ordinates of the hydrographs B and C, resulting from rainfall excess values of 6 and 3 cms, respectively.

(d) Uniform Rainfall Excess in Time: It assumes that the rainfall excess is the result of constant intensity of rainfall excess, which is equal to the rainfall excess divided by time duration T. The assumption requires that

(i) the selected storm be of short duration and intense, and

(ii) the resulting hydrograph be of short time base and single peaked.

(e) Uniform Rainfall Excess in Space: It assumes that the intensity of rainfall excess is uniform over the whole catchment. For the validity of this assumption, the catchment considered should be small. Therefore, the unit hydrograph theory is not applicable to large watersheds (more than 5000 sq. km). It can, however, be applied in components by sub-dividing the large watersheds.

(f) Time Invariance Proposition: It assumes that the unit hydrograph derived for a watershed does not vary with time. This principle holds when

- the physical characteristics of the watershed do not change with time, and
- Storm pattern and its movement do not change with time.

In brief, the above can be summarized as follows:

The principle of proportionality and superposition and the assumption of uniform rainfall excess in time and space constitute the unit hydrograph theory.

- The principle of proportionality states that for a unit period of rainfall excess, the surface runoff hydrograph is directly proportional to the rainfall excess volume.
- The hydrographs resulting from the rainfall excess volumes are superimposed sequentially and added for computing total direct surface runoff.
- An isolated, intense, short duration storm that is distributed uniformly in space and time forms an ideal storm for unit hydrograph derivation. Since such storms are rare, complex storms are used for deriving the unit hydrograph.
- The shape of the unit hydrograph depends significantly on storm pattern and antecedent physical conditions of the watershed. Consequently, unit hydrographs derived from different storms of the same watershed may differ from one another.

8.4 SURFACE RUNOFF ESTIMATION

There are several approaches to estimate the surface runoff. 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 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.

Generally, for most of the basins in India, 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 along with other problems, always lack the true representation of natural behavior 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 neighboring 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. Depending upon the type of data available, the water availability can be computed from the following methods.

8.4.1 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, a flow duration curve is constructed.

The stream flow data is important to determine the extent and pattern of runoff and used in determining the reliable water availability for various purposes, which include domestic water supply, commercial and industrial use; irrigation, hydropower and transport channels etc. These

records are therefore very useful in planning and designing and later for operating and managing the surface water related projects. Apart from water resources projects the stream flow records are also utilized in designing the bridges, culverts, flood plain delineation and flood warming systems.

Stream flow is measured in units of discharge (m^3/s) occurring at a specified time and constitutes historical data. Stream flow is not directly recorded, even though this variable is perhaps the most important in hydrologic studies. Instead, water level is recorded and stream flow is deduced by means of a rating curve. The rating curve is developed using a set of measurements of discharge and gauge height in the stream, these measurements being made over a period of months or years so as to obtain an accurate relationship between the stream flow rate, or discharge, and the gauge height at the gauging site commonly known as stage-discharge relation.

8.4.2 Extension of Streamflow Record

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

8.4.3 Extension with long-term runoff data at neighboring site

Runoff record can be extended using the long-term runoff record of the neighboring 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. Correlation with catchment areas.
2. Regression analysis.

Correlation with catchment area:

To extend the stream flow record of a station the following relation is used:

For north Indian watersheds

$$Q_s = Q_i (A_s/A_i)^{3/4} \quad (8.2a)$$

For south Indian watersheds

$$Q_s = Q_i (A_s/A_i)^{2/3} \quad (8.2b)$$

where,

Q_s, A_s = Discharge and catchment area at the station under consideration for extension of its short term data

Q_i, A_i = Discharge and catchment area at the index station with long term data.

Regression analysis

A correlation between short term discharge data and long term discharge data can be established. A regression line of the form $Q_s = a + bQ_i$ can be fitted. Here a and b are regression coefficients which can be estimated by regression analysis as:

$$b = \frac{\sum_{i=1}^{n}(Q_i - \overline{Q}_i)(Q_s - \overline{Q}_s)}{\sum(Q_i - \overline{Q}_i)^2}$$

$$a = \overline{Q}_s - b\overline{Q}_i$$

where $\overline{Q}_s = \frac{1}{N}\sum Q_s \qquad \overline{Q}_i = \frac{1}{N}\sum Q_i \quad (8.3)$

In Langbein's Log-deviation method, 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.

8.5 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.

8.5.1 Statistical or Correlation Approach

If long term precipitation record along with a stream flow data for a few years at the site is available, the procedure to be adopted 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. If long term precipitation record is available at the site along with precipitation and stream flow data for a few years at a neighboring site is available, first the rainfall-runoff relationship at the neighboring 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. If only precipitation record at the site is available, one can make use of various empirical formulae to compute the runoff.

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 Table8.1.

Table 8.1 Strange Rainfall-runoff ratios

Daily rainfall (mm)	Runoff percentage yield when the original stage of ground is					
	Dry		Damp		Wet	
	Percentage	Yield	Percentage	Yield	Percentage	Yield
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.

Inglis and *De Souza's* formulae are applicable for the western ghats and plains of Maharashtra, India and are reproduced as follows:

For ghat (Bombay-Deccan) areas,

$$R = 0.85\,P - 30.5$$

For plains,

$$R = \frac{(P - 17.8)P}{254} \qquad (8.4)$$

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

Khosla's formula

$$R = P - \frac{T}{3.74} \qquad (8.5)$$

where, R = Runoff (cm), P = Precipitation (cm), and T = Mean temperature (ºC)

To evaluate monthly runoff Khosla gave the following relationship

$R_m = P_m - L_m$

$L_m = 5T_m$ when $T_m > 4.5$ºC

where R_m = monthly runoff (mm), P_m = monthly precipitation (mm), L_m = Monthly losses (mm) and T_m = Mean monthly temperature in ºC. For $T_m < 4.5$ºC, losses are taken from the Table 8.2

Table 8.2 Monthly losses as proposed by Khosla

T_m (in ºC)	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.

A *runoff hydrograph* is a continuous record of stream flow over time. A complete runoff hydrograph contains information on runoff volume as the area under the hydrograph and peak flow rates as the maximum flow or peak of the runoff hydrograph as well as complete time history of flow. The volume of runoff is equal to the volume of rainfall excess or effective precipitation. Thus runoff volume is rainfall minus abstractions. Any of the methods for estimating abstractions or rainfall losses can be used in the computation of the volume of runoff.

The unit hydrograph of a drainage basin is defined as a hydrograph of direct runoff resulting from one centimeter of rainfall excess of a specified duration generated uniformly over

the basin area at a uniform rate. The specified duration is the period within which the rainfall excess is considered to be uniformly distributed so that direct runoff originates at the beginning. First a unit hydrograph of suitable unit duration is derived from an observed hydrograph for the drainage basin due to known storm. The unit hydrograph so derived can be applied for any other storm occurring on the basin and the resulting flood hydrographs can be obtained. The details of the derivation of unit hydrographs and their application, given a unit hydrograph of one duration to derive a unit hydrograph of some other duration and derivation of unit hydrographs for meteorologically homogeneous catchments (i.e. from basin characteristics like shape, size, slope etc.) will be discussed in other lectures.

Determination of runoff from rainfall data involves three steps 1) deduct the storm loss from rainfall to get value of rainfall excess 2) apply a derived relationship (unit hydrograph approach) between the rainfall excess and the surface discharge, yielding the time pattern of surface runoff and 3) to add the base flow to determine the total discharge.

8.5.2 Watershed Simulation

Mathematical modeling of hydrological processes provides a most powerful technique for an accurate assessment of the available water in space and time considering the physical processes to a certain extent close to the reality and incorporating the various factors affecting the natural hydrologic cycle due to man's influence. 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 behavior is studied. The essential feature of a simulation model is that it produces an output or series of outputs in response to an input or series of inputs. In the case of a rainfall-runoff model, the inputs are characteristics of the watershed being modeled, such as drainage area and channel network geometry (size and length), topography, soil and land

use characteristics and a time series of precipitation as input. The output is a time series of stream flow at an outlet location. Digital simulation is needed in a watershed research because it is a complex system to be analyzed by exact mathematical techniques. In practice rainfall runoff models rely on numerical and conceptual representations of the physical rainfall – runoff processes to achieve continuous runoff generation simulations. 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 or a series of functions which attempt to reproduce all the steps in the runoff process. Lumped models treat a whole catchment, or a significant portion of it, as a single unit, with inputs, internal state variables and outputs representing the hydrologic processes over the catchment as a whole. Distributed models divide the catchment into a number of sub areas; simulate each of them, and the interactions between them separately, maintaining different state variables for each model element, then combine the outputs to obtain catchment response.

8.5.3 Clark Model for UH Derivation

The Clark model concept (Clark, 1945) suggests that the instantaneous unit hydrograph (or referred as IUH) can be derived by routing the unit inflow in the form of *time-area diagram*, which is prepared from the isochronal map, through a single reservoir. The time area diagram is developed for an instantaneous rainfall excess over a catchment and represents the relationship between the areas that contributes to the runoff at the outlet versus the time of travel. It is assumed that the rainfall excess first undergoes pure translation by a travel time-area histogram and the attenuation is attained by routing the results of the above through a linear reservoir at the outlet. The time of concentration is the time taken for a droplet of water to travel from outer of a

catchment up to the outlet point. This is generally computed using empirical formulae which relates the time of concentration to the *geomorphological parameters* of the catchment. It represents the maximum time of translation end of the surface runoff of the catchment. In gauged areas the time interval between the end of the rainfall excess and the point of inflection of the resulting surface runoff provides a good way of estimating t_c from known rainfall-runoff data. In simple terms the total catchment area drains into the outlet in t_c hours.

First a the catchment is drawn using a toposheet map and than the points on the map having equal time of travel,(say t_{1c} h where $t_{1c} < t_c$), are considered and located on a map of the catchment. t_{1c} is calculated using the length of stream, length of centroid of the are from the outlet, average slope and area (from empirical equations). Than a line joined to all such points having same t_{1c}, to get an isochrones (or runoff isochrones). A typical *time-area diagram* and isochrones in catchment is given in Figure 8.3.

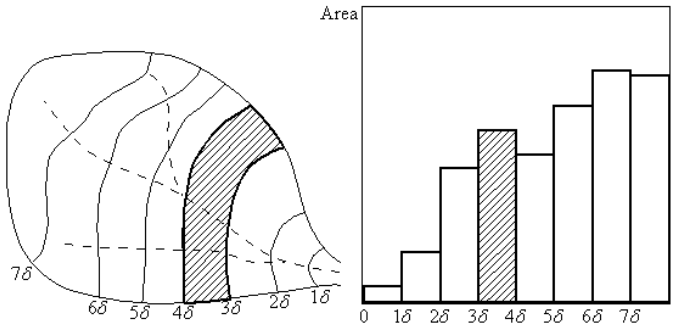

Figure 8.3 A typical *time-area diagram* and isochrones

Routing

The linear reservoir at the outlet is assumed to be described by S = KQ, where K is the storage time constant. The value of K can be estimated by considering the point of inflection P_i of a surface runoff hydrograph. At this point the inflow into the channel has ceased and beyond

this point the flow is entirely due to withdrawal from the channel storage. The continuity equation

$$I-Q = \frac{dS}{st}\text{; or } -Q = \frac{dS}{st} = K\,\frac{dQ}{dt}\text{ ; or } K = -Qi/(dQ/dt)i \qquad (8.6)$$

where suffix *i* refers to the point of inflection, and *K* can be estimated from a known surface runoff hydrograph of the catchment. The constant *K* can also be estimated from the data on the recession limb of a hydrograph. Knowing *K* of the linear reservoir, the inflows at various times are routed by the Muskingum method. Note that since a linear reservoir, the parameter *x* is known as *weighting factor* and takes a value of zero in 0 and the storage is a function of discharge only and the *Muskingum equation* reduces to

$$S = KQ \qquad (8.7)$$

Such storage is known as *linear storage* or *linear reservoir*. When $x = 0.5$ both the inflow and outflow are equally important in determining the storage. The coefficient *K* is known as *storage-time constant* and has the dimensions of time. It is a function of the flow and channel characteristics. It is approximately equal to the time of travel of a flood wave through the channel reach.

The outflow peak does not occur at the point of intersection of the inflow and outflow hydrographs. Using the continuity equation

$$(I_1 + I_2)\frac{\Delta t}{2} - (Q_1 + Q_2)\frac{\Delta t}{2} = \Delta S \qquad (8.8)$$

The increment in storage at any time *t* and time element Δt can be calculated. Summation of the various incremental storage values enables to find the channel storage *S* vs time relationship. If an inflow and outflow hydrograph set is available for a given reach, values of *S* at various time

intervals can be determined by the above equation. By choosing a trial value of *x*, values of *S* at any time *t* are plotted against the corresponding $[xI + (1-x)Q]$ values.

If the value of *x* is chosen correctly, a straight-line, however, if an incorrect value of *x* is used, the plotted points will trace a looping curve. By trail and error, a value of *x* is so chosen that the data describe a straight line, and the inverse slope of this straight line will give the value of *K*.

The total storage in the channel reach can then be expressed as

$$S = K [x I^m + (1-x) Q^m] \quad (8.9)$$

The inflow rate between an inter-isochrones area A_{1c} km^2 with a time interval Δt_{1c} (h) is

$$I = \frac{A_r \times 10^4}{3600 \Delta t_c} = 2.78 \frac{A_{1c}}{\Delta t_{1c}} \ (m^3/s)$$

The Muskingum routing equation would now be as follows:

$$Q_2 = C_0 I_2 + C_1 I_1 + C_2 Q_1$$

Where

$$C_0 = (0.5 \Delta t_c)/(K+0.5 \Delta t_c)$$

$$C_1 = (0.5 \Delta t_c)/(K+0.5 \Delta t_c)$$

$$C_2 = (K-0.5 \Delta t_c)/(K+ 0.5 \Delta t_c) \quad (8.10)$$

.e. $C_0 = C_1$. Also since the inflows are derived from the histogram $I_1 = I_2$ for each interval. Thus Eq. (8.10) becomes

$$Q_2 = 2C_1 I_1 + C_2 Q_1 \quad (8.11)$$

Routing of the time-area histogram by Eq. (8.11) gives the ordinates of IUH for the catchment. Using this IUH any other D-h unit hydrograph can be derived. So, in a gist the computational steps involved in Clark model are as follows:

(i) Make first estimate of Clark Model parameters, T_c and R, from the excess rainfall hyetograph and direct surface runoff hydrograph computed using the procedure described earlier. R is taken similar to constant *K* that can be estimated by considering the point of inflection P_i in the recession portion of the observed surface runoff hydrograph. At this point the inflow into the channel has ceased and beyond this point the flow is entirely due to withdrawal from the channel storage. In gauged areas the time interval between the end of the rainfall excess and the point of inflection of the resulting surface runoff provides a good way of estimating t_c

(ii) Construct the time-area curve, taking the T_c value obtained from step (i), using the procedure described in the early part of this chapter.

(iii) Measure the area between each pair of isochrones *planimeter*.

(iv) Plot the curve of time versus cumulative area. Note that the abscissa is expressed in percent of T_c. Tabulate increments between points that are at computational interval Δ t apart.

(v) Convert the units of inflow using the following relation

$$I = \frac{A_r \times 10^4}{3600 \Delta t_c} = 2.78 \frac{A_{1c}}{\Delta t_{1c}} \quad (m^3/s)$$

(vi) Route the inflow obtained from step (v) using the Eq. (8.18) and (8.19) to get IUH ordinates.

(vii) Compute the unit hydrograph of the excess rainfall duration using Eq. (8.10) and (8.11).

For the derivation of IUH the Clark model uses two parameters, time of concentration (T_c) in hours, which is the base length of the time-area diagram, and storage coefficient (R), in hours, of a single linear reservoir in addition to the time-area diagram. The governing equation of IUH using this model is given as:

$$u_i = C\, I_i + (1-C)\, u_{i-1} \qquad (8.12)$$

Using the derived IUH, a S-curve is obtained by lagging the IUH with unit time or for the Δt required i.e. 2-hr or 3- or 6- hrs UH. Then subtract the two S-curves that represents the unit hydrograph of duration Δt -h with a unit volume equal to 1 cm.

8.5.4 Synthetic Method for UH Derivation

The term "synthetic" in synthetic unit hydrograph (SUH) denotes the unit hydrograph derived from watershed characteristics rather than rainfall-runoff data. Their simplicity and ease in development can characterize these synthetic or artificial unit hydrographs, and requires less data and yield a smooth and single valued shape corresponding to one unit runoff volume, which is essential for unit hydrograph derivation. These methods utilize a set of empirical equations relating the physical characteristics of watershed to the few salient points of the hydrograph such as peak flow rate (Q_p), time to peak (t_p), time base (t_B), and UH width at 0.5 Q_p and 0.75 Q_p i.e. $W_{0.5}$ and $W_{0.75}$,respectively. However, in the SUH development a great degree of subjectivity is involved in fitting the remaining points on the SUH. In addition, simultaneous adjustments are required for the area under the SUH to be unity corresponding to unit rainfall-excess. The empirical equations involve certain constants, which vary over wide range, e.g. the Snyder's

proposed non-dimensional constants C_t and C_p which varies in the range of 1.01-4.33, 0.23-0.67, 0.4 - 2.26 , 0.31 -1.22 respectively. This is used to describe the shape of the UH, expressible as

$$t_p = C_t \left(LL_c\right)^{0.3} \tag{8.13}$$

$$Q_p = 640\left(\frac{AC_p}{t_p}\right) \tag{8.14}$$

$$t_b = 3 + 3\left(\frac{t_p}{24}\right) \tag{8.15}$$

where C_t and C_p are non dimensional constants and in general varies from 1.8 to 2.2 and 0.56 to 0.69, respectively. Equetions (1) to (3) hold good for rainfall-excess (RE) duration (or unit duration = t_r) as

$$t_r = \frac{t_p}{5.5} \tag{8.16}$$

If the duration of rainfall-excess (RE), say $t_{r1,}$ is different from t_r as defined above, a modified lag time t_{pm} can be estimated from

$$t_{pm} = t_p + \frac{\left(t_{r1} - t_r\right)}{4} \tag{8.17}$$

One can sketch many UHs through the three known characteristic points of the UH (i.e. t_p, Q_p, and t_b), with its specific criteria, i.e. area under the SUH, to be unity. To overcome with this ambiguity associated with the Snyder's method, the U.S. Army Corps of Engineers [26] developed empirical equations between widths of UH at 50% and 75% of Q_p i.e. W_{50} and W_{75} respectively as a function of Q_p per unit area (Q_p/A), expressible as

$$W_{50} = \frac{830}{Q_p^{1.1}} \tag{8.18a}$$

$$W_{75} = \frac{470}{Q_p^{1.1}} \tag{8.18b}$$

where W50 and W_{75} are in units of hour. Thus one can sketch a smooth curve through the seven points (t_p, t_b, Q_p, W_{50}, and W_{75}) relatively in an easier way with less degree of ambiguity and to have the area under the SUH that equal to unity.

8.5.5. GIUH Method for UH Derivation

Rodriguez-Iturbe and Valdes (1979) first introduced the concept of geomorphologic instantaneous unit hydrograph (GIUH) which led to the renewal of research in hydro geomorphology. The derived expression yields full analytical, but complicated expressions for the instantaneous unit hydrograph. Thus, they suggested that it is adequate to assume a triangular instantaneous unit hydrograph and only specify the expressions for the time to peak and peak value of the IUH. These expressions are obtained by regression of the peak as well as time to peak of IUH derived from the analytic solutions for a wide range of parameters with that of the geomorphologic characteristics and flow velocities. The expressions are given as:

$$q_p = 1.31\, R_L^{0.43}\, V / L_\Omega \qquad (8.19a)$$

$$t_p = 0.44\,(L_\Omega / V)(R_B / R_A)^{0.55}(R_L)^{-0.38} \qquad (8.19b)$$

where;

L_Ω = the length in kilometers of the stream of order Ω

V = the expected peak velocity, in m/sec.

q_p = the peak flow, in units of inverse hours

t_p = the time to peak, in hours

R_B,R_L,R_A = the bifurcation, length, and area ratios given by the Horton's laws of stream numbers, lengths, and areas, respectively.

Empirical results indicate that for natural basins the values for R_B normally range from 3 to 5, for R_L from 1.5 to 3.5 and for R_A from 3 to 6 (Smart, 1972).

On multiplying Esq. (8.19 a) and (8.19b) one get a non-dimensional term q_p t_p as:

$$q_p * t_p = 0.5764 \, (R_B / R_A)^{0.55} (R_L)^{0.05} \quad (8.20)$$

This term is independent of the velocity and, thereby, on the storm characteristics and hence is a function of only the catchment characteristics. The methodology adopted for estimation of these geomorphological parameters is described below:

Stream Ordering

Strahler's stream ordering system may be used for stream ordering (u). The principles of Strahler's method of stream ordering are mentioned below.

(i) Channels that originate at a source are defined to be first order streams.

(ii) When two streams of order u join, a stream of order u+1 is created.

(iii) When two streams of different orders join, the channel segment immediately downstream has the higher of the order of the two continuing streams.

(iv) The order of a basin is the order of the highest stream.

This ordering system can be applied through ILWIS over the entire drainage network. In ILWIS, length of each stream is stored in a table. Then, after adding length of each stream for an order, one can get the total stream lengths of each order. The total stream length divided by the number of stream segment (N_u) of that order gives the mean stream length L_u for that order.

Stream Number (N_u)

In ILWIS, the number of streams of each order can be stored in a table and for each order the total number of streams can be computed. Horton's law of stream numbers states that number of stream segments of each order is in inverse geometric sequence with order number i.e. $N_u = R_b^{u-k}$ where, k is the order of trunk segment, u is the stream order, N_u is the number of stream of order u and R_b is a constant called the bifurcation ratio.

Stream Length (L_u)

Length of each stream is stored in a table. Then after adding length of each stream for a given order, the total stream length of each order (L_u) may be computed. The total stream length divided by the number of stream segments (N_u) of that order gives the mean stream length L_u for that order. Length of the main stream from its origin to the gauging site is indicated as (L) while the length of the highest order stream is indicated as L_Ω.

Bifurcation Ratio (R_b)

Horton's law of stream numbers states that number of stream segments of each order is in inverse geometric sequence with order number i.e.

$$N_u = R_b^{u-k} \quad (8.21)$$

where, *k* is the order of trunk segment, u is the stream order, N_u is the number of stream of order u and R_b is a constant called the bifurcation ratio.

Bifurcation ratio (R_b) is defined as the ratio of stream segments of the given order N_u to the number of stream segments of the next higher order N_{u+1} i.e:

$$R_b = N_u / N_{u+1} \quad (8.22)$$

Length Ratio (R_l)

Length ratio is one of the important geomorphologic characteristics. Horton (1945) defined length ratio (R_l) as the ratio of mean stream length ($\bar{L}_u$) of segment of order u to mean stream segment length ($\bar{L}_{u-1}$) of the next lower order u-1, i.e.:

$$R_l = \bar{L}_u / \bar{L}_{u-1} \quad (8.23)$$

Area Ratio (R_a)

The area of streams of each order can be estimated using the area and length relationship (Strahler, 1964). Horton stated that mean drainage basin areas of progressively higher order streams should increase in a geometric sequence, as do stream lengths. The law of stream areas can be expressed as:

$$A_u = A_1 R_a^{u-1} \quad (8.24)$$

Here, A_u is the mean area of basin of order u, A_1 is the mean area of first order basin, and R_a is the area ratio. Areas for different order basins were estimated using the relationship between area of any order and area of highest order as given below:

$$A_u = A_1 R_b^{u-1} (R_{lb}^{u}-1)/ (R_{lb}-1) \quad (8.25)$$

Where, R_{lb} is the Horten's term for the length ratio to bifurcation ratio. In this relationship the only unknown is A_1 and it can be computed from physical characteristics. The mean areas are computed using the value of A_1.

Area ratio (R_a) is defined as the ratio of area streams ($A_{u)}$ of order u to the area of streams (A_{u-1}) of the order u-1, i.e.:

$$R_a = Au / A_{u-1} \quad (8.26)$$

Area ratio is one of the important geomorphologic characteristics. These require some of the important geomorphological parameters from Toposheets. From the above geomorphological data in a basin, the terms q_p, t_p and qp tp ca be computed. Using these values, an IUH may be computed for basin where the data is lacking. Section 8.5.3 discuss about converting an IUH to a he unit hydrograph of duration Δt -h with a unit volume equal to 1 cm.

8.5.6 SCS-Curve Number (CN) Method for UH Derivation

The curve number method is an empirical approach to estimate the direct runoff from the relationship between rainfall, land use and hydrologic soil group. Many hydrologic models use the curve number method to estimate direct runoff from fields or watersheds. The curve number method was initially developed from many small experimental watersheds. The SCS-CN method is based on the water balance equation and two fundamental hypotheses. The first hypothesis state that ratio of actual amount of direct runoff to maximum potential runoff is equal to the ratio of amount of actual infiltration to the amount of the potential maximum retention. The water balance equation and the two hypotheses are expressed, respectively as;

$$P = I_a + F + Q \tag{8.27}$$

$$\frac{\text{Actual retentation}}{\text{Potential maximum retentation}} = \frac{\text{Actual direct runoff}}{\text{Potential maximum runoff}}$$

$$\frac{F}{S} = \frac{Q}{P - I_a}, \quad I_a = 0.\lambda S \tag{8.28}$$

where, P is the potential precipitation (cm); I_a is the initial abstraction before pounding; Q is the actual direct runoff; F is the actual retention (cumulative infiltration) in the watershed and S is the potential maximum retention or infiltration. The current version of SCS-CN method assumes λ equal to 0.2 for usual practical applications although it may vary in the range $(0, \infty)$

The popular form f the method is derived by above equations a;

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \qquad \text{for } P > I_a = 0.2\ S$$

$$Q = 0.0 \qquad \text{for } P < I_a = 0.2\ S \tag{8.29}$$

Since the parameter S can vary in the range $0 \leq S \leq \infty$, it is mapped into a dimensionless parameter CN, the CN varying in a more appealing range $0 \leq CN \leq 100$, as below;

$$S = \frac{2540}{CN} - 25.4 \qquad (8.30)$$

For the given hydrologic group and various type of land use/ treatment and antecedent moisture conditions, CN- values are available. Knowing the CN, the value of recharge capacity S is calculated and using this value of S, runoff is calculated.

8.5.7 Nash Model for UH Derivation

Using the concept of cascade of n-linear reservoirs, Nash (1959) and Dooge (1959) derived the gamma-form expression for IUH as:

$$q = \frac{1}{K\Gamma n}\left(\frac{t}{K}\right)^{n-1} e^{\frac{-t}{K}} \qquad (8.31)$$

where n and K are the number of reservoirs and the storage coefficient, respectively, and these describe the shape of IUH; and q is the depth of runoff per unit time per unit effective rainfall. Eq. (8.31) is used for SUH derivation from known n and K

Defining q_p equal to the peak discharge expressed in mm/hr/mm and t_p as the time to peak in hr, a simple relation between n and β ($=q_p t_p$) of the gamma distribution for SUH derivation is given as follows [*Bhunya et al*, 2003]:

$n = 5.53\ \beta^{1.75} + 1.04 \qquad 0.01<\beta<0.35$

$n = 6.29\ \beta^{1.998} + 1.157 \qquad \beta \geq 0.35 \qquad (8.32)$

where and $\beta = q_p\, t_p$. Eq. (8.32) can be used to estimate the shape of the SUH for known values of q_p and t_p.

8.6 ESTIMATION OF PEAK RUNOFF RATE AND UH DISTRIBUTION GRAPH

8.6.1 Rational Formula Method

The Rational Formula for estimating peak runoff rate, introduced in the USA in 1889 has become widely used as a tool for drainage design, particularly for sizing water-conveyance

structures. It is an empirically developed model, with simplifying assumptions including uniform rainfall with uniform intensity over the entire watershed for duration equal to the time of concentration. The rational formula estimates the peak rate of runoff at any location in a watershed as a function of the drainage area, runoff coefficient, and mean rainfall intensity for a duration equal to the time of concentration (the time required for water to flow from the most remote point of the basin to the location being analyzed).The Rational Formula is expressed as

$$Q = \frac{1}{360} C.I.A \tag{8.33}$$

where Q is the peak runoff rate (m^3/s), C is the runoff coefficient (unit less, ranging from 0 to 1), I is the rainfall intensity (mm/hr), and A is the watershed area (hectares). The coefficient C is determined from a table, based on land-cover, topography, soil type, condition (management practice), and storm return period.

Use of the rational formula requires the time of concentration (t_c) for each design point within the drainage basin. Empirical formulae are available for determination the time of concentration. One such formula given by Kirpich is

$$t_c = 0.02 L^{0.77} S^{-0.385} \tag{8.34}$$

where, t_c is time of concentration (minutes);

L is length of channel reach (meters) and

S is slope of channel reach.

The duration of rainfall is then set equal to the time of concentration and is used to estimate the design average rainfall intensity (I). The rainfall intensity (I) is the average rainfall rate for duration equal to the time of concentration for a selected return period. Once a particular

return period has been selected for a design, a time of concentration is calculated for the drainage area, the rainfall intensity can be determined from Rainfall-Intensity-Duration data.

8.6.2. Features of the Unit Hydrograph and Distribution Graph

Distribution graph is a non-dimensional form of unit hydrograph, generally available in percent form. The sum of all percentage values should be equal to 100. The main features of a unit hydrograph or a distribution graph are:

1. It shows the time distribution of the discharge hydrograph of a watershed produced by a uniform rainfall excess of given depth on an area.
2. It transforms the rainfall excess to time-varying discharge at the outlet assuming a linear process.
3. It represents the watershed characteristic by way of exhibiting the integrated effect of the surface features on the routing of the rainfall excess through the catchment.

It is worth mentioning that the derivation of a unit hydrograph usually excludes records of moderate floods and the principles of unit hydrograph can be applied for estimating the design flood, supplementing missing flood records, and short-term flood forecasting that is based on recorded rainfall.

8.7 RESULTS AND DISCUSSION

8.7.1 Unit Hydrograph for Mahanadi basin based on Clark's Approach

The time of concentration t_c for the basin is estimated using the empirical formulae developed by Kirpich. A catchment for *Jira* catchment (Br-195 in *Sambalpur – Titlagarh* tributary) with a catchment area of 615 Sq. Km is considered for this case. The method is followed to estimate t_c and the points on the area with equal time of travel, (say t_1 h where T_1 <

t_c) was joined to get *isochrones* of one hour interval and shown in Figure 8.4. The figure shows the catchment being divided into N (=8) sub areas by isochrones having equal time interval. The inter-isochrones areas are used to construct a travel time-area histogram (Figure 8.5). The linear reservoir at the outlet is assumed to be described by S = KQ, where K is the storage time constant. The value of K is estimated as 0.33 Hr by considering the point of inflection of a surface runoff hydrograph. The parameters C_0, C_1 and C_2 are calculated using K and Δt and is given in Table 8.3. The derivation of unit hydrograph ordinates are shown last column in Table 8.3.

As an alternate approach, a unit hydrograph has been conducted for Mahanadi basin at National Institute of Hydrology, Roorkee. In this study the parameters, T_c and R, of Clark Model have been derived for the Mahanadi sub-basins using HEC-1 programme package. The value of (T_c + R) is equal to 10.3 Hour, and R/ (T_c+R) of this flood event equals to 0.417. The unit hydrograph is given in Figure 8.6. Using the same approach, the salient points of unit hydrograph ordinates for eleven small catchments are given in Table 8.3. These catchments were selected because of the event and data availability status.

The regional relationship has been presented in the graphical form where average of (T_c + R) for each sub basins have been plotted against their respective basin area is being discussed in chapter 9 and this plot along with the fixed value R/(T_c+R) has been used to estimate the regional parameters for ungauged basins.

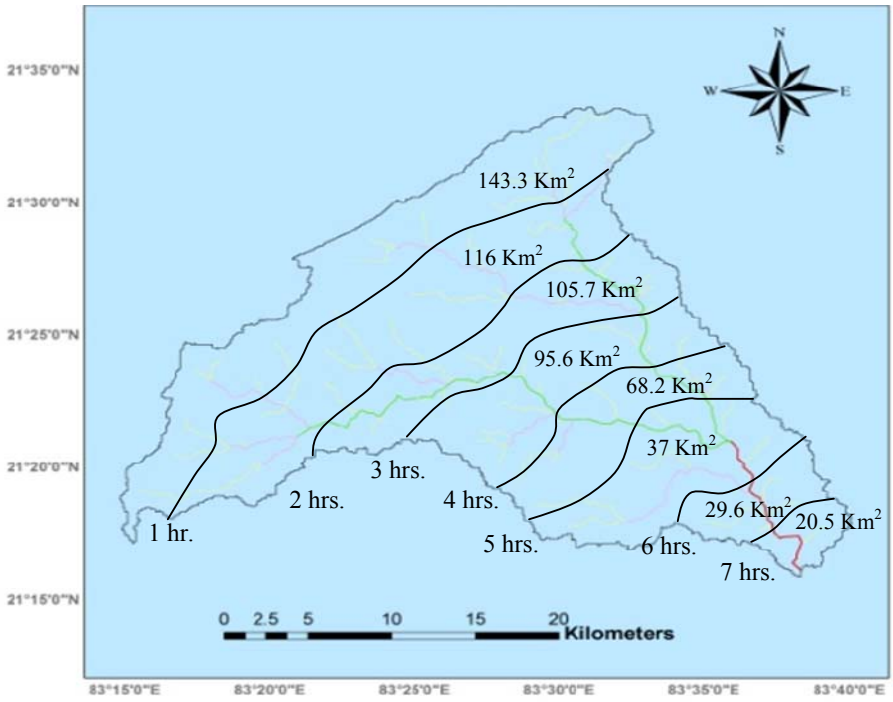

Figure 8.4 Isochrones of one hour interval for *Jira* catchment (Br-195) A= 615 Km^2

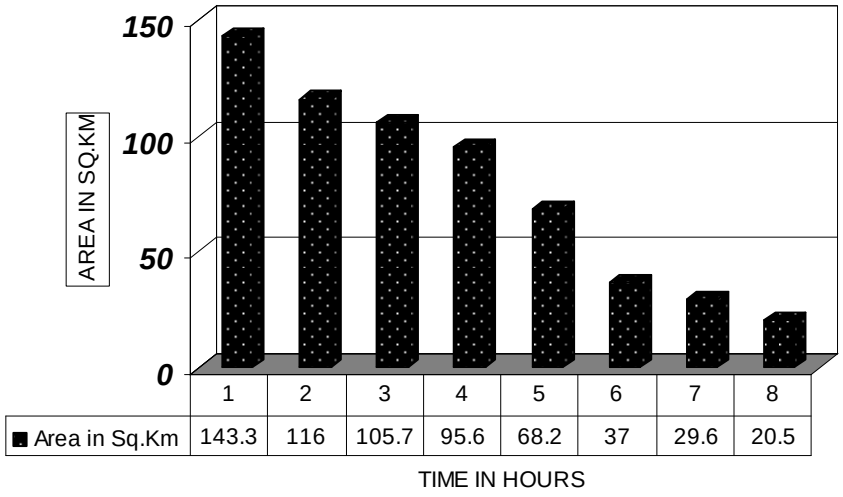

Figure 8.5 Time-area histogram for *Jira* catchment (Br-195) with catchment area= 615 Km^2

Table 8.3 Unit hydrograph ordinates for *Jira* catchment using Clark's Approach

T	Partial flow	C_0*I_2	C_1* I_1	C_2* O_1	Sum	Q
Hr	km^2-cm/h		km^2-cm/h			m^3/s
0	0.00	0.00	0		0.00	0.00
1	143.3	7.26	0.00	0	7.26	20.19
2	116	11.74	6.19	5.187262	23.11	64.25
3	105.7	7.66	10.00	16.50959	34.17	95.00
4	95.6	5.97	6.53	24.41057	36.91	102.61
5	68.2	4.73	5.09	26.36464	36.18	100.58
6	37	4.03	4.03	25.84329	33.90	94.23
7	29.6	4.96	3.43	24.21263	32.61	90.64
8	20.5	3.77	4.23	23.29013	31.29	86.99
9	0.00	0.00	3.21	22.35207	25.57	71.07
10	0.00	0.00	0.00	18.26204	18.26	50.77
11	0.00	0.00	0.00	13.04458	13.04	36.26
12			0.00	9.317741	9.32	25.90
13	Total=615..9			6.655663	6.66	18.50
14				4.75414	4.75	13.22
15				3.395882	3.40	9.44
16				2.425679	2.43	6.74
17				1.732662	1.73	4.82
18				1.237641	1.24	3.44
19				0.884047	0.88	2.46
20				0.631475	0.63	1.76
21				0.451062	0.45	1.25
22				0.322194	0.32	0.90
23				0.230143	0.23	0.64
24				0.164391	0.16	0.46
25	Clark's Unit	0.117425	0.12	0.33		
26				0.083876	0.08	0.23
27	Area	615 Km^2		0.059913	0.06	0.17
28	K	0.33 Hr		0.042796	0.04	0.12
29	Δt	1 Hr		0.030569	0.03	0.08
30	C	0.33333		0.021835	0.02	0.06
31	C0	0.14286		0.015597	0.02	0.04
32	C1	0.14286		0.011141	0.01	0.00
33	C2	0.71429		0.007958	0.01	0.00

Table 8.4 Unit hydrograph peak ordinates of small catchments using Clark's Approach

Sl.	Br-No	A	L	L_c	S	T_c	T_p	R	q_p
		km^2	km	km	m/ km	Hours	Hours	Hours	cumecs
1	7	3108	96.6	51.84	0.61	16.61	14.36	0.43	632.99
2	121	950	80.5	38.64	4.99	10.58	9.64	0.39	228.97
3	12	666	67.02	25.75	1.09	11.13	10.08	0.28	223.36
4	195	615	53.94	28.16	1.59	10.14	9.28	0.31	172.38
5	235	312	41.06	21.09	1.61	8.57	8.00	0.26	116.14
6	385	194	39.36	15.13	4.39	6.62	6.37	0.29	63.29
7	69	173	35.42	18.5	2.31	7.48	7.10	0.32	52.13
8	59B	136	28.18	9.26	5.87	4.97	4.94	0.27	63.38
9	40K	95	24.21	12.27	4.56	5.36	5.28	0.17	55.51
10	698	93	26.57	14.4	9.04	5.22	5.16	0.22	35.36
11	154	58	12.48	9.65	5.18	4.03	4.10	0.17	36.07

8.7.2 Unit Hydrograph Derivation for Mahanadi Basin Using Conventional Methods

A number of isolated storm hydrographs of short spells are selected among the short term data supplied by the agency. As an example in this section, the available data event for small catchment *Kelo* (Br-121) with an (A= 1150 Km^2 is taken. After estimating the total rainfall and the total flood volume, rainfall excess is derived. For each of these hydrographs the baseflow is separated by

adopting straight line method. The ordinates of the DRH's are divided by the ER to obtain the unit hydrograph. Selecting an event storm of 3.9.96 (9-10 pm), and 4.9.96 (11-2 am), the Unit Hydrographs derived are given in Table-8.5. Adopting such methods as discussed the mean of the unit hydrographs is adopted as the standard UHG for the basin for the given duration. While taking mean the peak and time to peak are first calculated. The base is than adjusted and checked for the UH volume. Figure 8.6 depicts the unit hydrograph for the same catchment using conventional methods and the Clark method.

Table 8.5 Hydrograph due to Rainfall as on 3.9.96 (9-10pm) at Kelo (Br-121) (A= 1150 Km^2)

T (Hr)	Rainfall	P (mm)	Q (mm)	DRH Cumecs	Time	P (mm)	ER(mm)	Q (mm))	UHO cumec
1	14	11.2	50.8	0	1	11.2	2.30	0	0.00
2	3.5	2.8	60.46	17.66	2			17.66	60.00
3			60.6	17.8	3			17.8	77.39
4			100.1	57.3	4			57.3	249.13
5			98.9	56.1	5			56.1	243.91
6			68.45	25.65	6			25.65	111.52
7			56.1	13.3	7			13.3	57.83
8			52.3	9.5	8			9.5	41.30
9			51.35	8.55	9			8.55	37.17
10			50.8	8	10			8	34.78
11			50.8	8	11			8	34.78
12			50.8	8	12			8	34.78
13			44.68	1.88	13			1.88	8.17
14			42.8	0	14			0	0.00
15			42.8	0					
Total Direct Runoff = 16 Cumecs-Hr = 810000 m3									
Total Rainfall = 14 mm over the basin = 4900000 m3									
Total Loss =11.70 mm									
Interception Rate = 6.70 mm/h (av)									
ER = 2.3 mm									

8.7.3 Synthetic Unit Hydrograph (SUH) Method for Mahanadi Basin

The synthetic unit hydrograph (SUH) concept proposed by Sherman (1932**)** for estimating the storm runoff hydrograph at the gauging site in a catchments corresponding to a rainfall hyetograph, is still a widely accepted and admired tool in hydrologic analysis and synthesis. One of the popular and frequently used empirical equations is Snyder's proposed non-dimensional approach that is used here in this section. This method utilize a set of empirical equations relating the physical characteristics of watershed to the few salient points of the hydrograph such as peak flow rate (Q_p), time to peak (t_p), time base (t_B), and UH width at 0.5 Q_p and 0.75 Q_p i.e. $W_{0.5}$ and $W_{0.75}$,respectively. This involve certain constants, which vary over wide range, e.g. proposed non-dimensional constants C_t and C_p varies in the range of 1.01-4.33,

0.23-0.67 (Miller AC, Kerr SN, Spaeder DJ, 1983), 0.4 - 2.26 , 0.31 -1.22 (Hudlow MD, Clark DM, 1969), respectively. Earlier, the direct surface runoff hydrographs are obtained using linear base flow separation, and adjustments of the UH properties by (Chow VT, Maidment DR, Mays LW, 1988). These relationships and empirical formulae are discussed in section 8.5.4.

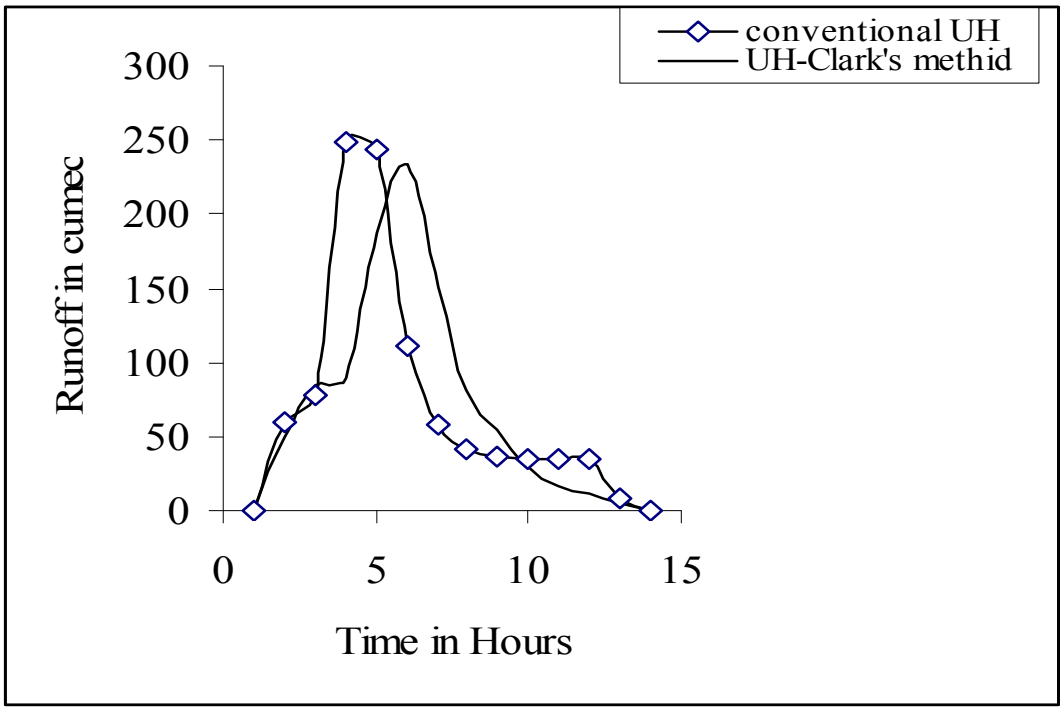

Figure 8.6 UH using conventional methods and the Clark method (*Kelo* Br-121) catchment

For the determination of q_p and t_p, for the SUH the formulae are given in Eqs. 8.13 and 8.14. The procedure described in the section 8.5.3 is taken for demonstration with geomorphological data for *Aherajhor* (Br-154). Here, C_p = 0.65, L_{CA} = 9.65 Kms. = 6.03 miles, and A = 58 km^2= 22.65 mi^2. The parameters of the Snyder method computed for SUH derivation are: t_p = 5 hr, q_p = 0.13 per hr, W_{50} = 6.42 hr, and W_{75} = 3.64 hr. Apparently, the result due to both the unit hydrograph using observed data with conventional method and the Snyder method

are close to each other as per the goodness of fit evaluated using the coefficient of determination, COD, (Mendenhall and Sincich, 1988) as

$$\text{COD} = 1 - \frac{\sum_{i=1}^{N}\left(Y_{observed} - Y_{computed}\right)^2}{\sum_{i=1}^{N}\left(Y_{observed} - Y_{average}\right)^2} \qquad (8.35)$$

where N is the number of observations. The estimated COD for the derived SUH is 0.68. In this case, the observed UH is difficult, so the ordinates of the UH computed from conventional method with unit effective rainfall is considered to be the observed one. The salient ordinates of the derived UH using Snyder's method is given in Table 8.6, and the unit hydrographs are given in Appendix-8. The average COD for the derived SUH compared to conventional methods is in the range of 0.63-0.83.

Table 8.6 Unit hydrograph ordinates derived using Snyder's method for 11 catchments in Mahanadi region.

Sites	A $miles^2$	L(Miles)	Lc(Miles)	Tp (Hr.)	Tb(Hr.)	Tr (Hr.)	Qp(m^3/s)	W50 (hr.)	W75(hr.)
7	1199.69	59.99	32.19	17.42	51.77	3.17	698.97	29.81	16.88
121	366.70	49.99	24.00	15.94	49.92	2.90	241.83	26.01	14.73
12	257.08	41.62	15.99	12.65	47.82	2.30	209.86	20.57	11.65
195	237.39	33.50	17.49	13.53	46.91	2.46	187.58	21.32	12.07
235	120.43	25.50	13.10	10.86	43.58	1.97	122.59	16.14	9.14
385	74.88	24.44	9.40	10.22	42.77	1.86	74.38	16.58	9.39
69	66.78	22.00	11.49	9.99	42.49	1.82	72.69	14.99	8.49
59B	52.50	17.50	5.75	8.38	40.47	1.52	71.56	11.70	6.63
40K	36.67	15.03	7.62	8.29	40.37	1.51	53.69	10.82	6.13
698	35.90	16.50	8.94	8.50	40.63	1.55	46.69	12.33	6.98
154	22.39	7.75	5.99	5.69	37.12	1.04	49.17	6.93	3.92

8.7.4 Unit Hydrograph for Mahanadi basin using GIUH method

Geomorphological characteristics, which are quite commonly used in geomorphological based instantaneous unit hydrograph (GIUH), are evaluated using a both toposheets and GIS package. Manual estimation of the geomorphological parameters is a tedious and cumbersome process and often discourages the field engineers from developing the regional methodologies

for solving various hydrological problems of the ungauged basins or in limited data situations. At times, it also leads to erroneous estimates. Salient basin topographical features like length of the stream for different orders and total length till outlet i.e. GD sites, slope computed by dividing length with maximum ground level difference and basin area computed from toposheets are shown in Table 8.6. . On the other hand, modern techniques like the GIS serve as an efficient approach for storage, processing and retrieval of large amount of database. Spatial modeling and tabular databases stored in GIS system are updated here. Some of the details of the works are given for *Jira* (Br-195) catchment in the following tables.

Table 8.7a Geomorphological characteristics of *Jira* River Catchment: Stream and Drainage Area Details.

Order of the stream	Total number of stream	Total length of stream	Drainage area
1	209	2,26,92.832	246.840
2	49	1,9,929.875	285.569
3	12	8,682.656	214.09
4	3	71,615.808	438.296
5	1	23,737.824	615.000

Table 8.7b Geomorphological characteristics of *Jira* River Catchment for different stream orders.

Stream order (w)	Total number of streams (m) (N_W)	Mean stream length (m) (L_W)	Mean stream area (km^2) (A_W)	Bifurcation ratio (R_W)	Stream length ratio (R_L)	Stream area ratio (R_A)
1	209	1081.88	1.181	4.265	-	-
2	49	2284.28	5.828	4.083	2.91	4.935
3	12	4056.89	17.834	4.000	1.776	3.060
4	3	23871.94	146.099	3.000	5.884	8.192
5	1	23737.82	615.000	-	0.994	4.390

Using the above geomorphological characteristics in sixteen small catchments in Mahanadi basin, two salient points i.e. q_p (the peak flow, in units of inverse hours), and time to peak of GIUH (t_p) are calculated from Eqs. (8.19). In this section, seven catchments out of total twenty three have not been considered due to difficulty in assessing these in toposheets and imageries.

Table 8.8 GIUH parameters for sixteen small catchments in Mahanadi basin.

Sr.	Bridge No.	A Km^2	Velocity m/s	R_B	R_L	R_A	L Km.	qp m^3/Sq.Km	tp hour	STDER
1	7	3108	2.5	3.832	2.147	4.155	239.00	0.18	16.5	15.6
2	121	1150	2.5	3.523	1.787	3.940	138.00	0.29	12.5	27.1
3	489	823	2.5	4.447	2.277	4.803	185.00	0.33	8.5	12.3
4	12	666	2.5	3.920	1.523	4.319	77.00	0.17	15.5	19.3
5	195	615	2.5	4.272	2.123	4.611	52.00	0.20	10.5	11.2
6	235	312	2.5	3.448	1.868	4.309	11.08	0.17	11.5	8.4
7	332(ii)	225	2.5	2.927	1.520	3.442	75.30	0.38	6.5	3.4
8	385	194	2.5	4.272	2.364	5.737	56.51	0.30	7.5	35.8
9	69	173	2.5	4.215	2.775	4.844	33.60	0.22	9.5	21.7
10	59(B)	136	2.5	4.282	1.907	5.553	35.42	0.30	6.5	6.4
11	698	113	2.5	4.364	1.711	4.365	64.25	0.46	4.5	5.2
12	58	109	2.5	4.201	2.375	4.199	82.18	0.19	7.5	2.2
13	79	67	2.5	3.708	1.672	4.126	54.00	0.41	4.5	31.2
14	37	64	2.5	3.697	2.337	3.983	63.02	0.38	5.5	25.3
15	154	58	2.5	4.139	1.758	4.179	73.67	0.4	5.5	2.7
16	59(S)	47	2.5	2.978	1.953	4.501	49.31	0.37	5.5	7.5

For the dynamic parameter velocity (V), Rodriguez and Valdes (1979) in their studies assumed that the flow velocity at any given moment during the storm can be taken as constant throughout the basin. The characteristic velocity for the basin as a whole changes throughout the storm duration. For the derivation of GIUH, this can be taken as the velocity at the peak discharge for a given rainfall-runoff event in a basin. However, for ungauged catchments, the peak discharge is not known and, therefore this criterion for estimation of velocity cannot be applied. In such a situation, the velocity may be estimated using the velocity-rainfall excess relationship. In this case, the velocity is assumed to be 2.5 m/s and computed peak discharge and

time to peak are for separate catchments are given in last columns. The GIUH are plotted and given in Appendix-8.

To check the performance of the methods, the goodness-of-fit is further evaluated using the ratio (STDER) of the absolute sum of non-matching areas to the total hydrograph area, expressed mathematically as [HEC-1, 1990]:

$$STDER = \left[\left(\sum_{i=1}^{N} (q_{oi} - q_{ci})^2 w_i \right) / N \right]^{1/2} \; ; \; w_i = (q_{oi} + q_{av})/(2q_{av}) \tag{8.36}$$

where q_{oi} = i^{th} ordinate of the observed (by conventional method as discussed above) hydrograph, q_{ci} = i^{th} ordinate of the computed GIUH, w_i = weighted value of i^{th} hydrograph ordinates, q_{av}= average of the observed hydrograph ordinates, and N = total number of hydrograph ordinates. The dimension of STDER is same as q_{av}, therefore, it should be compared to the same expression computed with q_{av} instead of q_{ci} to become a hint at a possible good fit. Since the computed w_i-values (Eq. 8.36) are larger for higher q-values, the resulting high STDER value signifies larger non-matching areas on the upper portion of the hydrograph to non-matching areas in the lower portion below q_{av}. A low value of STDER-value represents a good-fit, and vice versa; STDER equal to zero represents a perfect fit. The STDERs for GIUH methods for sixteen catchments are given in Table 8.8. It can be observed from Table 8.8 STDER values due to GIUH are less as compared to the fitness test of COD in Snyder's method. However, on the other hand, the STDER for the GIUH method is found lower for smaller size catchments. This cannot be inferred that GIUH fitting or driving UH by this method is better for smaller sized catchments, nevertheless, in the region considered for this study, more of the lesser size basins gives better fitting of UH compared to conventional methods.

8.7.5 Unit Hydrograph for Mahanadi basin using SCS-Curve Number (CN) Method:

The Soil Conservation Service method of U.S. Department of agriculture (USDA) uses a specific average dimensionless unit hydrograph. It is derived from the analysis of large number of natural UHs for the watersheds of varying size and geographic locations, to synthesize the UH (Singh VP, 1988). To enable defining time base, t_b, in terms of time to peak, t_p, and time to recession, t_{rc}, the SCS method represents the dimensionless UH as a triangular UH, which further facilitates the computation of runoff volume (V) and peak discharge q_p as

$$V = \frac{(q_p t_b)}{2} = \frac{1}{2} q_p (t_p + t_{rc}); \text{ where } t_{rc} = 1.67\ t_p \qquad (8.37a)$$

$$Q_p = 0.749\left(\frac{V}{t_p}\right) \qquad (8.37b)$$

where q_p is in mm/hr/mm; V is in mm; t_p and t_{rc} are in hrs.

To determine the complete shape of the SUH from the non-dimensional (q/q_p vs t/t_p) hydrograph, the time to peak is computed as

$$t_p = t_L + \left(\frac{t_r}{2}\right) \qquad (8.37c)$$

where t_L = lag time from centroid of excess-rainfall to peak discharge (q_p) in hour; and t_r = excess-rainfall duration for unit duration (hour).

The lag time (t_L) can be estimated from the watershed characteristics using curve number (CN) procedure as

$$t_L = \frac{L^{0.8}(2540 - 22.86CN)^{0.7}}{14104 CN^{0.7} Y^{0.5}} \qquad (8.37d)$$

Where t_L = in hours; L = hydraulic length of watershed (m); CN = curve number ($50 \leq 95$); and Y = average catchment slope in (m/m).

Alternatively Eq. (8.37b) can be expressed as

$$Q_p = 2.08\left(\frac{A}{t_p}\right) \tag{8.38}$$

where Q_p = peak discharge in m^3/s for one cm of excess-rainfall; A = area in km^2.

Thus with known q_p, t_p, and specified dimensionless UH, an SUH can be easily derived. Using the observed data of catchments as in the previous section for GIUH, the SCS method results are given in Table 8.9. The derived flood hydrograph salient points as observed from the results, the SCS method over-estimates the peak discharge, under-estimates the rising limb, and closely matches with the recession limb of the hydrograph. However, if the equation given i.e. empirical rational formulae at Eq (8.38) is multiplied by two, then the values of Q_p given in Table 8.9 is close to Snyder, Clark and GIUH methods as discussed in the nest section.

Table 8.9. UH using SCS-CN for eleven small catchments in Mahandi basin.

Sl.	**Br-No**	**A**	**L**	**S**	**CN**	**T_L**	**T_P**	**Q_p**
		km^2	Km	m/ km	65	Hours	Hours	cumecs
1	7	3108	96.6	0.61	35	19.63	20.13	642.24
2	121	950	80.5	4.99	65	12.99	13.49	292.86
3	12	666	67.02	1.09	45	10.96	11.46	241.72
4	195	615	53.94	1.59	45	12.70	13.20	193.83
5	235	312	41.06	1.61	30	10.15	10.65	121.92
6	385	194	39.36	4.39	45	9.10	9.60	84.04
7	69	173	35.42	2.31	25	7.53	8.03	89.67
8	59B	136	28.18	5.87	25	7.14	7.64	74.06
9	40K	95	24.21	4.56	20	7.17	7.67	51.50
10	698	93	26.57	9.04	35	6.67	7.17	53.92
11	154	58	12.48	5.18	65	2.87	3.37	71.58

However, the inconsistencies associated with the method can be enumerated as

(i) Since the curve number method is applicable to watersheds of areas ranging from 8 km^2 to 16 km^2, its application to large and mid size watersheds may lead to erroneous results.

(ii) Since the SCS method fixes the ratio of time base to time to peak (t_b/t_p) for triangular UH equal to 2.67 (or 8/3), ratios other than this may lead to the other shapes of the UH. In particular, larger ratio implies greater catchment storage. Therefore, since the SCS method fixes the ratio (t_b/t_p), it should be limited to midsize watersheds in the lower end of the spectrum (Ponce VM, 1989)

(iii) the SCS method is one of the popular methods for synthesizing the UH for only small watersheds of less than 500 square miles.

8.7.6 Unit Hydrograph for Mahanadi basin using Nash Model (1957)

Based on the concept of n-linear reservoirs of equal storage coefficient K, Nash (1959) and Dooge (1959) derived the instantaneous unit hydrograph (IUH) in the form of a gamma function as:

$$q(t) = \frac{1}{K\Gamma n}\left(\frac{t}{K}\right)^{n-1} e^{\frac{-t}{K}} \qquad (8.39)$$

Here, n and K define the shape of IUH and q is the depth of runoff per unit time per unit effective rainfall. Eq. (8.39) is used for deriving a SUH from parameters n and K.

Parameter Estimation

Assuming that Eq. 8.39 represents a UH of unit duration, the condition at peak ($t = t_p$) can be given by $dq_p/dt = 0$. Thus,

$$K = t_p / (n-1) \quad (8.40)$$

Defining $\beta = q_p t_p$, it follows from Eqs. (8.39) and (8.40) that

$$q_p t_p = \frac{(n-1)^{(n-1)} e^{-(n-1)}}{\Gamma(n-1)} = \beta \quad (8.41)$$

McCuen (1989) described n to vary with Q_p, t_p, and A as

$$n = 1.045 + 0.5\, f + 5.6\, f^2 + 0.3\, f^3 \quad (8.42)$$

where $f = Q_p t_p / A$, A is the area in acres, Q_p is in ft^3/s, and t_p is in hours.

A simple numerical procedure was used by Bhunya et al. (2003) to get an approximate solution of Eq. (8.41) given as:

$$\begin{aligned} n &= 5.53\, \beta^{1.75} + 1.04 && 0.01<\beta<0.35 \\ n &= 6.29\, \beta^{1.998} + 1.157 && \beta \geq 0.35 \end{aligned} \quad (8.43)$$

These equations were derived using numerical simulation and optimization. Nash model given in Eq. (8.39) has two parameters which can be readily obtained from the known values of q_p and t_p. Though exact solution of n in terms of β is difficult to evaluate, it is possible to derive a relatively simple expression from Eq. (8.43), and K using Eq. (8.40a). The present method is compared on basis of their ability to describe the SUH shape using a data of Mahandi as discussed with four methods. The results for eleven small catchments in Mahandi basin is given in Table 8.10.

In testing, all the four methods i.e. Clark's , GIUH and Nash methods were found to perform better than the existing synthetic methods, i.e. SCS (1957), and Snyder (1938), on basis

of their comparison with observed UH. Among all methods analyzed, the Clark's and GIUH based Nash

Table 8.10 Unit hydrograph peak ordinates of sixteen small catchments (applied with GIUH approach in Table 8.8) using Nash model

Sr.	Bridge No.	q_p in m^3/Sq.Km	t_p in hour	β (=q_p t_p)	n	K (hr.)	STDER
1	7	0.18	16.5	0.297	1.713	23.14	11.3
2	121	0.29	12.5	0.3625	1.985	12.69	27.4
3	489	0.33	8.5	0.2805	1.653	13.01	12.1
4	12	0.17	15.5	0.2635	1.595	26.06	19.0
5	195	0.20	10.5	0.21	1.435	24.12	11.4
6	235	0.17	11.5	0.1955	1.398	28.88	8.8
7	332(ii)	0.38	6.5	0.247	1.542	12.00	3.2
8	385	0.30	7.5	0.225	1.476	15.74	13.6
9	69	0.22	9.5	0.209	1.433	21.96	21.5
10	59(B)	0.30	6.5	0.195	1.397	16.37	16.7
11	698	0.46	4.5	0.207	1.427	10.53	5.8
12	58	0.19	7.5	0.1425	1.285	26.30	12.5
13	79	0.41	4.5	0.1845	1.372	12.10	3152
14	37	0.38	5.5	0.209	1.433	12.71	2.5
15	154	0.4	5.5	0.22	1.462	11.90	12.7
16	59(S)	0.37	5.5	0.2035	1.418	13.15	7.2

(as salient pints of UH i.e. peak runoff and time to peak were taken from GIUH) methods predicted the hydrograph more accurately than other distributions. The parameter sensitivity showed the Nash method (similar to Gamma distribution) to overestimate q_p highly for any overestimation of the *n* parameter, which is also true for the Clark method where the parameter K is sensitive than R. This observation comes from the calculation done through Systat and Microsoft Excel in the process of trial and error. The trend in the sensitivity graph could not be

attempted as it is difficult in this case with small number of catchments and event data.. The physical significance of the results derived from five methods and their properties of the UH are summarized in the following section.

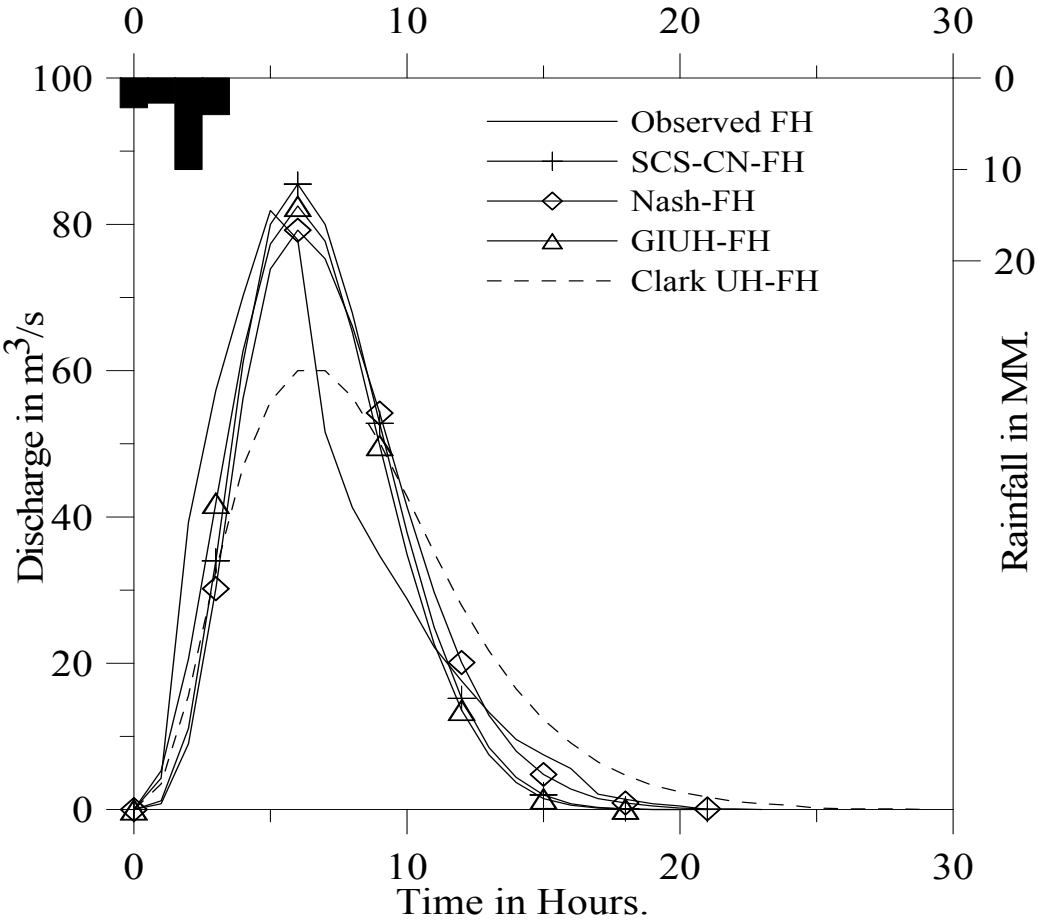

Figure 8.7 Flood hydrograph (FH) using four methods for Br-385 (*Sandu*-stream) catchment

8.8 RESULT AND DISCUSSION

Clark used only gauged basins in his original work and did not provide guidelines for the estimation or determination of the K value and time-area relationships for ungauged basins. Clark specifically states that "the determination of any kind of unit hydrograph for a stream without records is a hazardous task". In fact, Clark simply recommends that a stream gauge should be installed at the point of interest. Although there are a number of methods and references for estimating the Muskingum coefficients for hydrologic routing in a channel

system, the K value in the Clark IUH cannot be estimated in such a manner with any degree of confidence. It is possible to transfer the K value from one basin to another nearby basin. This should be done with a regression analysis, 3 of the FERC Guidelines. (Note that the Guidelines use the more recent "R" designation for the parameter called K in Clark's paper.) Parameters to be considered include (but are not limited to) drainage area, lengths, and slopes. The U.S. Army Corps of Engineers, Hydrologic Engineering Center, has noted that the ratio of K/(Tc + K) tends to remain constant for a region. (HEC-1 User's Manual 1990a) As given in section 8.6.1. in this case, most results have been obtained by computing Clark parameters from available storm and runoff events using HEC-1. Excel sheet and Matlab Programs. The inferences obtained from UH computations using the other models like Nash. GIUH, SCS-CN and Snyder's are given below:

(i) In case of Snyder's' SUH approach for UH derivation, the manual fitting of the characteristic points needed great degree of subjectivity and trial and error, and may involve error. The constants C_t and C_p vary over wide range from catchment to catchments, and may not be equally suitable for all the regions. The time base of Snyder's method is always reasonable greater for fairly large watersheds only.

(ii) Given two points on the UH, e.g., time to peak and peak flow, the Nash model can be used to describe the shape of the unit hydrograph, and they perform better than the existing synthetic methods, i.e., methods suggested by Snyder, SCS-CN, and GIUH models. For ungauged catchments, empirical or regional formulae can be used to estimate these salient points of the UH or even Snyder, SCS-CN and GIUH. The proposed parameter estimations methods of Nash model are simple to use, and gives accurate results of the actual (gamma) pdf parameters, as verified using fitting index measure.

(iii) Among the four pdfs analyzed in this study, the Clark's and GIUH models are more flexible in description of SUH shape as they skew on both sides similar to a UH, and on the basis of their application to field data. However, in case of GIUH, the salient points like time to peak and peak flow can be computed with a close degree to observed, but the fitting of UH has to be improved further. Same is the case with SCS-CN but the variable parameters like CN for a catchment has to be ascertain talking updated land use index or else there occurs a large error in determining the peak flow and time to peak along with W50 , W75 and time to base. Since the curve number method is applicable to watersheds of areas ranging from 8 km^2 to 16 km^2 (Ponce, 1989). In the present study some catchments are large and more are mid sized watersheds, thus SCS-CN method results as observed has erroneous results in such cases (with area more than 100 Sq.Km). In case of Nash model that is similar to a Gamma distribution (in a limiting case), and the results shows that it should be a preferred method for deriving SUH.

(iv) Parameters are highly sensitive to peak flow of the UH in case of Nash model (when β is low). Any overestimation in parameter estimates increases the peak flow of the UH and the trend is reverse at large β values.

(v) Assuming a triangular UH, relationships between the model parameters and statistical properties of the UH were derived and used to express the parameters with time to peak and time to base of the UH. The results showed close accuracy for parameter estimates when verified using field data. Further, the variance of a positively skewed UH was observed to be more sensitive to t_B than t_p, implying that the σ^2-ratio for any two catchments approximately varied in proportion to the corresponding t_B^2-ratio.

CHAPTER 9
REGIONAL UNIT HYDROGRAPH ANALYSIS

9.0 INRODUCTION

Whenever sufficient and reliable records on stream flow and rainfall are available the unit hydrograph for those basins can be derived from the rainfall-runoff data of storm events using one of the techniques. However, most of the small basins are generally not gauged and many water resources projects are being planned in those basins. Therefore, it becomes necessary to have the estimates of floods at the proposed sites in small ungauged basins. As we know, the unit hydrograph technique is one of the simple and most powerful techniques among other for the estimation of design flood. Therefore, the unit hydrographs for such basins have to be estimated by using data on climatological, physiographic and other factors of these basins.

The main purpose of the regional unit hydrograph study is to estimate the unit hydrograph ordinates or the unit hydrograph parameters for basins for which no gauge discharge data are available. The procedure involved in regional unit hydrograph analysis requires the evaluation of representative unit hydrograph parameters and pertinent physical characteristics for the gauged basins in the region. Then multiple linear regression analysis is performed, considering one of the unit hydrograph parameters at a time as a dependent variable and various basin characteristics as independent variables, in order to develop the regional relationship for the unit hydrograph derivation. Further, knowing the basin characteristic for an ungauged basin in the region from the available toposheet and climatological data, the unit hydrograph for that basin can be derived using the relationships developed for the region.

The conventional techniques of unit hydrograph (UH) derivation require historical rainfall-runoff data. Due to the various reasons, adequate runoff data are not generally available

for many of the small and medium size catchments. Indirect inferences through regionalization are sought for such types of ungauged catchments. A large number of regional relationships have been developed by many investigators relating either the parameters of unit hydrograph (UH) or instantaneous unit hydrograph (IUH) models with physiographic and climatologic characteristics. Regionalization of the parameters is, however, a very tedious task to accomplish since the hydrological behavior of many nearby catchments have to be ascertained before being confident about the values of the parameters. These conventional approaches which are in vogue for estimation of design floods require rainfall-runoff records and the model parameters need to be updated from time to time. Many times, this task of regionalization becomes very tedious, and in certain cases, even impossible. This study is concerned with the prediction of surface runoff discharge hydrographs, particularly for ungauged watersheds.

Nash, Clark's, GIUH and Snyder's methods for developing UH is explained in chapter 8. Among them, the GIUH approach has many advantages over the regionalization techniques as it avoids the requirement of flow data and computations for the neighboring gauged catchments in the region as well as updating of the parameters. An alternate approach which couples the parameters of the Clark model of instantaneous unit hydrograph derivation (Clark, 1945) with the GIUH approach has been developed at the National Institute of Hydrology.

In this Chapter, three steps are used to develop regional unit hydrographs: (i) Unit hydrograph derivation using Clark's and Nash model (ii) Unit hydrograph derivation using synthetic unit hydrograph and GIUH methods for all catchments in the region (iii) regionalizing the UH computed with available data base in the region. Various steps involved in developing the regional unit hydrograph relationships for a hydro meteorologically homogeneous region are described and discussed. Various regional unit hydrograph studies conducted in India as well as

abroad are also presented to provide the proper understanding to the participants about the different forms of relationships established.

9.1 BASIC STEPS INVOLVED IN DEVELOPING THE REGIONAL UNIT HYDROGRAPH

The following steps should be followed in executing a regional study to develop regional unit hydrograph relationships for a basin.

(i) *Choice of the basin:* In regional study, care should be taken to select those basins which are indeed similar in hydro-meteorological characteristics. The basins considered for developing the regional unit hydrograph should be able to represent the regional behavior as close as possible. Further, one should always try to include maximum no. of gauged basins in the regional study. However, minimum eight to ten basins are required for the regional study.

(ii) *Split sample tests for the region:* In order to test the performance of the developed regional relationships, the data of at least two to three basins should be kept independent. It means, those basins should be treated as ungauged basins and they should not be considered while developing the regional relationship.

(iii) *Rainfall-runoff data:* Rainfall-runoff data of different basins for each of the major past flood events should be considered for analysis. If the basin underwent to some major changes due to man's influence or landuse changes, then the rainfall-runoff data of only recent past flood events should be considered for analysis.

(iv) *Computation of excess rainfall*: A suitable technique should be adopted to separate the loss from total rainfall in order to get the excess rainfall hyetograph.

(v) *Base flow separation:* The base flow should be separated from the streamflow hydrograph using a consistent base flow separation technique, in order to get the direct surface runoff hydrograph.

(vi) *Derivation of Unit Hydrograph*: The unit hydrograph should be derived by analyzing the excess rainfall-direct surface runoff data for each event of different basin using a suitable unit hydrograph derivation technique.

(vii) *Derivation of representative unit hydrograph*: The representative unit hydrograph for each basin may be derived by averaging the unit hydrograph obtained from different events of the basin using standard averaging procedure. However, if considerable variations are observed in unit hydrographs derived from different events of a basin, then the unit hydrograph parameters of each event should be considered, alongwith the basin and storm characteristics, in the regional study.

(viii) *Split sample test for the storms:* The performance of the representative unit hydrograph of a basin should be tested by reproducing the two or three independent storms which are not to be used for deriving the representative unit hydrograph.

(ix) *Development of regional unit hydrograph relationship:* Step-wise multiple linear regression analysis can be performed, taking the unit hydrograph parameters of different basin as dependent variables, and/or climatic characteristics as independent variables to develop the optimal regional unit hydrograph relationships.

(x) *Representative Unit Hydrograph for ungauged basins:* The regional relationships developed at step (ix) are used for split sample test for the region as described in step (ii). Further the representative unit hydrograph for the ungauged basins of the hydrometeorolgoically homogeneous region can be derived using measurable basin and/or climatic characteristics in the generalized relationships developed in step (ix).

The basic procedure can be summarized as: (a) From records of gauged basin in a given region, derive the relations between characteristics of the unit hydrograph and the physical characteristics of the basin. These relations depend on the method used, some of which are described in next section, (b) Assume that these relationships apply to the ungauged basins in the region and use them to derive the synthetic unit hydrograph.

9.2 REGIONAL UNIT HYDROGRAPH FOR MAHANADI BASIN BASED ON CLARK'S APPROACH

To derive a regional unit hydrograph one simple approach is to compute the parameters, T_c and R, of Clark Model for the Mahanadi sub-basins using HEC-1 programme package. Next the value of $(T_c + R)$ and $R/(T_c+R)$ of each of the floods analyzed in each of the sub-basins is averaged for the respective sub-basins. The regional relationship can be developed by a showing graphically the relation between average of $(T_c + R)$ for each sub basins against their respective basin area. This is shown Figure 9.1. This plot along with the fixed value $R/(T_c+R)$ can be used to estimate the regional parameters for ungauged basins.

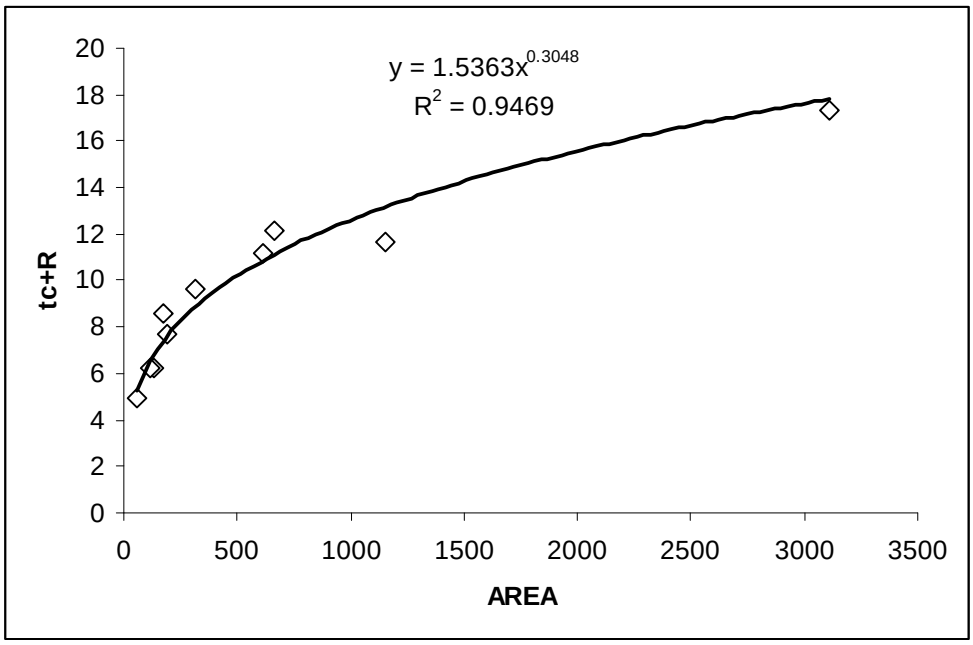

Figure 9.1 Plots between Area versus (T_c + R)

9.3 REGIONAL UNIT HYDROGRAPH FOR LOWER MAHANADI REGION 3 (D) BASED ON NASH AND CLARK APPROACH

To develop the regional unit hydrograph relationships, the physical parameters of Mahanadi Basin subzone 3 (d) is related with the average parameters of Nash Model and Clark model for those basins as given in Chapter-8 (Table 8.4 and 8.10). Figures. 9.2 and 9.3 illustrate the variation of n K with ($LL_{cc}/\sqrt{S}$) and K with main stream length L respectively. These two plots are used as the regional relationships based on Nash model. The Clark's model parameter T_c has been related with ($LL_{cc}/\sqrt{S}$) as shown in Figure 9.4. A fixed value of the ratio $R/(T_c+R)$ along with T_c vs $LL_c/\sqrt{S}$ plot is used to establish the regional unit hydrograph relationships based on Clark's model. Due to non-availability of much data for the other bridge basins of the subzone, only five basins have been considered in the regional study. Therefore the study has some what limited scope. However, it provides encouraging results for further investigation.

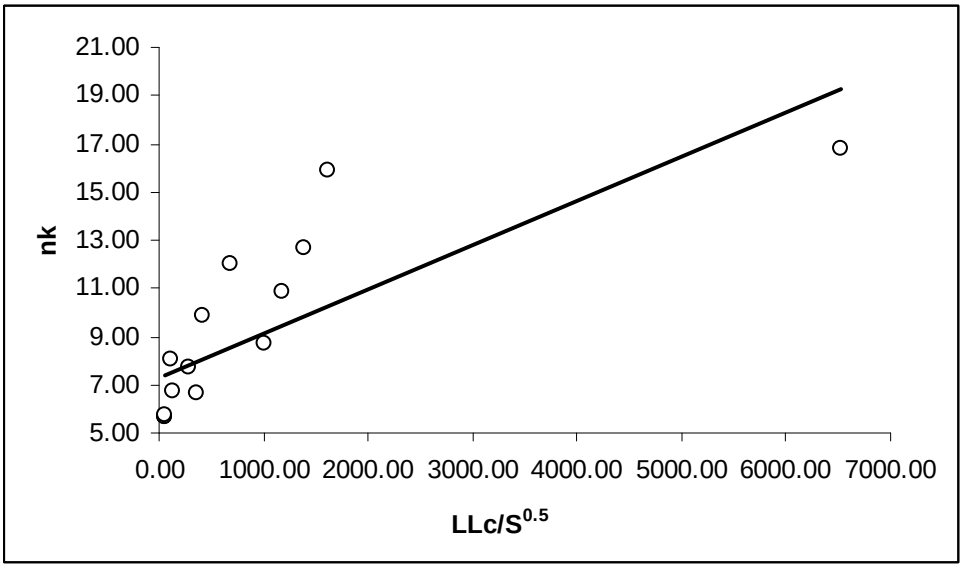

Figure 9.2 Plot between nk and $LL_c/S^{0.5}$ for Mahanadi basin ($R^2 = 0.57$)

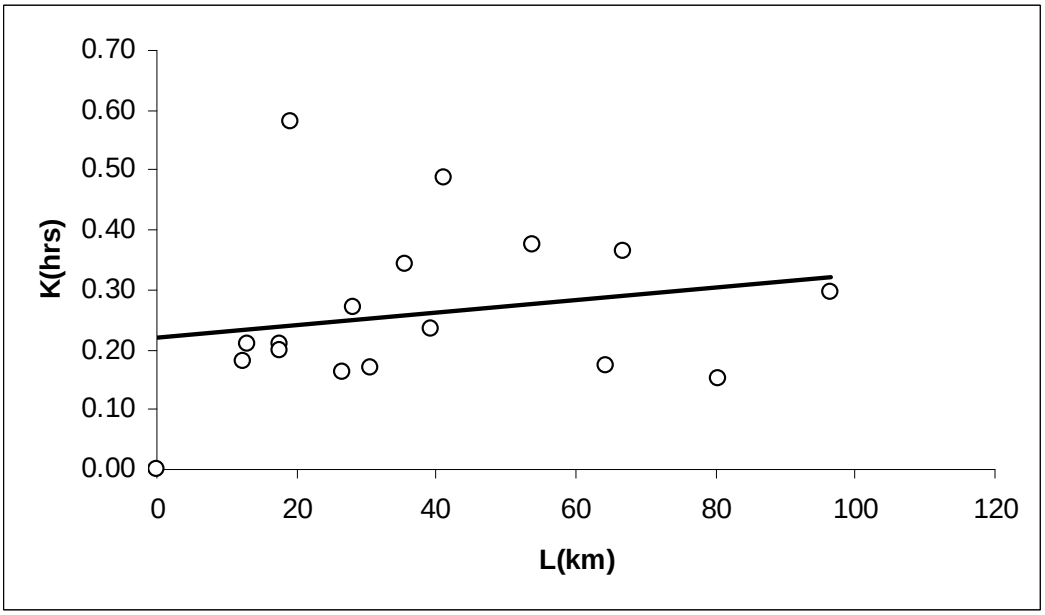

Figure 9. 3 Plot between K and L for Mahanadi basin ($R^2 = 0.63$)

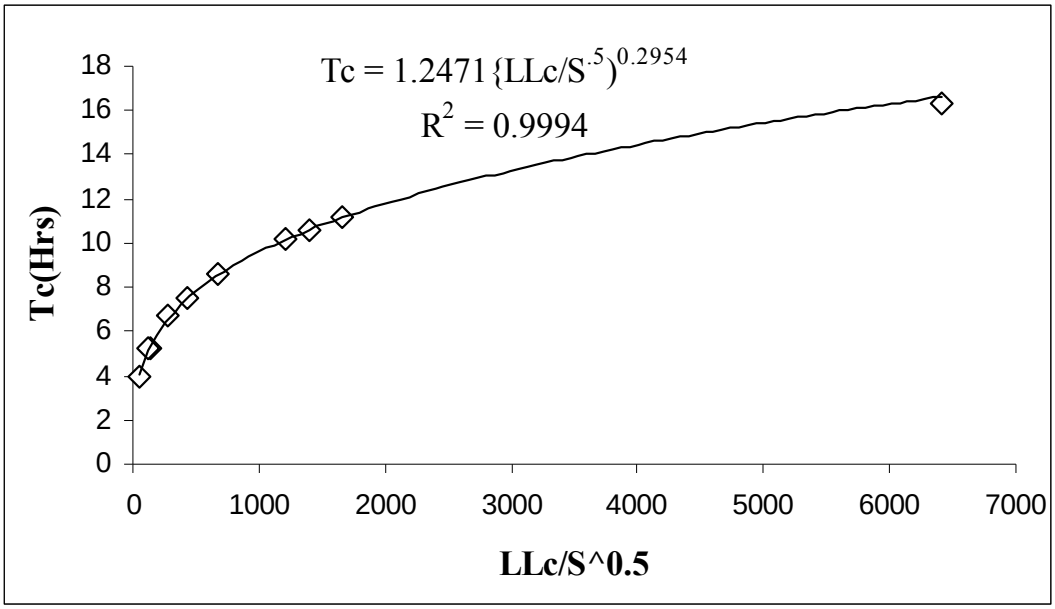

Figure 9.4 Plot between Tc and $LL_c/\sqrt{S}$ for Mahanadi basin

Taking the geometrical data from toposheet or imageries, one an get the Clark's model parameters from these figures and get the UH shape as discussed in Chapter 8. A similar step has to be followed for Nash model by approximating both the parameters from the regional relationship, and describe the UH shape as per the steps given in section 8.7.6**.**

9.4 REGIONAL UNIT HYDROGRAPH USING GIUH METHOD

Evaluation of geomorphologic characteristics involves preparation of a drainage map, ordering of the various streams, measurement of basin area, channel length and perimeter etc. and thereafter computing various geomorphological parameters such as bifurcation ratio, length ratio, area ratio, drainage density, drainage frequency, basin configuration and relief aspects etc. The geomorphological parameters viz. bifurcation ratio (R_b), length ratio (R_L), area ratio (R_A), length of the highest order stream ($L_ç$), length of the main stream (L) for Mahanadi Basin sub zone (3d) is given in Table 8.8l.

Rodriguez-Iturbe and Valdes (1979) suggested that it is adequate to assume a triangular instantaneous unit hydrograph and only specify the expressions for the time to peak and peak value of the IUH. These expressions are obtained by regression of the peak as well as time to peak of IUH derived from the analytic solutions for a wide range of parameters with that of the geomorphologic characteristics and flow velocities. The related expressions along with the way to derive the salient points of an UH are given in Eqs (8.19)-(8.20) at section 8.5.5 in Chapter 8. In developing an UH from these two salient points is somewhat difficult, so two alternate approaches are being adopted and they are discussed in the following sections.

9.4.4 GIUH Based Clark Model Approach

The Clark model concept (Clark, 1945) suggests that the IUH can be derived by routing the unit inflow in the form of time-area diagram, which is prepared from the isochronal map, through a single reservoir. For the derivation of IUH the Clark model uses two parameters, time of concentration (T_c) in hours, which is the base length of the time-area diagram, and storage coefficient (R), in hours, of a single linear reservoir in addition to the time-area diagram. The governing equation of IUH using this model is given as:

$$u_i = C\, I_i + (1-C)\, u_{i-1} \qquad (9.1)$$

where;

u_i = ith ordinate of the IUH; C & (1-C) = the routing coefficients.;

and $C = \Delta t / (R + 0.5\, \Delta t)$; Δt = computational interval in hours ; I_i = i^{th} ordinate of the time-area diagram

A unit hydrograph of the desired duration (D) may be derived using the following equation:

$$U_i = \frac{1}{n}(0.5\, u_{i-n} + u_{i-n} + u_{i-n+1} \ldots\ldots\ldots\ldots + u_{i-1} + 0.5\, u_i) \qquad (9.2)$$

where; U_i = ith ordinate of unit hydrograph of duration D-hour and at computational interval Δt hours, n = no. of computational intervals in duration D hrs = $D/\Delta t$, and u_i = ith ordinate of the IUH.

The step by step explanation of the procedure to derive unit hydrograph for a specific duration using the GIUH based Clark model approach is given here under:

(i) Rainfall-excess hyetograph is computed either by uniform loss rate procedure or by SCS curve number method or by any other suitable method.

(ii) For a given storm, the peak velocity V using the highest rainfall-excess is determined from the relationship between velocity and intensity of rainfall-excess.

(iii) Compute the time of concentration (T_c) using the equation :

$$T_c = 0.2778\, L / V \qquad (9.3)$$

where; L = length of the main channel, and V = the peak velocity in m/sec.

Considering this T_c as the largest time of travel, find the ordinates of cumulative isochronal areas corresponding to integral multiples of computational time interval with the help of the non-dimensional relation between cumulative isochronal area and the percent time of travel. This describes the ordinates of the time-area diagram at each computational time interval.

(iv) Compute the peak discharge (Q_{pg}) of IUH given by Eq. (8.20).

(v) Assume two trial values of the storage coefficient of GIUH based Clark model as R_1 and R_2. Compute the ordinates of two instantaneous unit hydrographs by Clark model using time of concentration T_c as obtained in step (iii) and two storage coefficients R_1 and R_2 respectively, (refer section 8.5.3). Compute the IUH ordinates at a very small time interval say 0.1 or 0.05 hrs so that a better estimate of peak value may be obtained.

(vi) Find out the peak discharges Q_{pc1} and Q_{pc2} of the instantaneous unit hydrographs obtained for Clark model for the storage coefficients R_1 and R_2, respectively, at step (v).

(vii) Find out the value of objective function, using the relation:

$$FCN1 = (Q_{pg} - Q_{pc1})^2 \quad (9.4)$$

$$FCN2 = (Q_{pg} - Q_{pc2})^2 \quad (9.5)$$

(viii) Compute the first numerical derivative FPN of the objective function FCN with respect to parameter R as:

$$FPN = \frac{FCN1 - FCN2}{R_1 - R_2} \quad (9.6)$$

(ix) Compute the next trial value of R using the following governing equations of Newton-Raphson's method:

$$\Delta R = \frac{FCN1}{FPN} \quad (9.7)$$

and

$$R_{NEW} = R_1 + \Delta R \quad (9.8)$$

(x) For the next trial, consider $R_1 = R_2$ and $R_2 = R_{NEW}$ and repeat steps (v) and (ix) till one of the following criteria of convergence is achieved.

(a) FCN2 = 0.000001

(b) No. of trials exceeds 200

(c) $ABS(\Delta R)/R_1 = 0.001$

(xi) The final value of storage coefficient (R_2) obtained as above is the required value of the parameter R corresponding to the value of time of concentration (T_c) for the Clark model.

(xii) Compute the instantaneous unit hydrograph (IUH) using the GIUH-based Clark Model with the help of the final value of the storage coefficient (R), time of concentration (T_c) as obtained in step (xi) and time-area diagram.

(xii) Compute the D-hour unit hydrograph (UH) using the relationship between IUH and UH.

9.4.5 GIUH Based Nash Model Approach

Nash (1957) proposed a conceptual model in which catchment impulse could be represented as the outflow obtained from routing the unit volume of the instantaneous rainfall-excess input through a series of *n* number of successive linear reservoirs having equal delay time. The equation for the instantaneous unit hydrograph for the Nash model is given as:

$$U(0,t) = \frac{1}{k} \cdot \frac{1}{\Gamma n} e^{-\frac{t}{k}} \cdot \left(\frac{t}{k}\right)^{n-1} \tag{9.9}$$

where, U(0,t) is the ordinate of the instantaneous unit hydrograph, k is the storage coefficient and n is the number of reservoirs.

The complete shape of the GIUH can be obtained by linking q_p and t_p of the GIUH with the scale (k) and shape (n) parameter of the Nash model. Now, by equating the first derivative (with respect to t) of the equation (9.9) to zero, t becomes the time to peak discharge, t_p. Further, peak discharge q_p and time to peak t_p can be linked as:

$$q_p \cdot t_p = \frac{(n-1)}{\Gamma n} e^{-(n-1)} \cdot (n-1)^{n-1} \tag{9.10}$$

The left side of the equation can be computed with basin characteristics using Eq. (8.20). All terms in the left hand side in equation (9.14) are known. Only unknown is Nash Model parameter *n*, which can be obtained by solving the equation (9.10) using Newton Rapshon method of nonlinear optimization.

The parameter n may be substituted in the following equation to determine the Nash Model parameter k for the given velocity V.

$$k = \frac{0.44 L_W}{V} \cdot \left[\frac{R_B}{R_A}\right]^{0.55} \cdot R_L{}^{-0.38} \cdot \frac{1}{(n-1)} \tag{9.11}$$

The derived value of n and k can be utilized for determination of the complete shape of IUH as per the steps given in section 8.7.6. Table 8.4 shows the unit hydrograph peak ordinates of small catchments using Clark's Approach where peak flow and time to peak is derived using

GIUH. Similarly Table 8.10 gives the unit hydrograph peak ordinates of sixteen small catchments (applied with GIUH approach in Table 8.8) using Nash model.

9.5 SYNTHETIC UNIT HYDROGRAPH METHODS USED FOR REGIONAL ANALYSIS

These methods utilize a set of empirical equations relating the physical characteristics of watershed to the few salient points of the hydrograph such as peak flow rate (Q_p), time to peak (t_p), time base (t_B), and UH width at 0.5 Q_p and 0.75 Q_p i.e. $W_{0.5}$ and $W_{0.75}$,respectively. This is explained with verification with eleven small catchments. The salient points of the hydrograph Q_p, t_p, W_{50} , W_{75}, Tr and T_r estimated (Table 8.6) for eleven catchments are regionalized with individual catchment characteristics like A, L , Lc and slope (S) that is given in Table 9.1. In brackets in col- 4 and 5 for L and Lc is checked from recent toposheet and imageries. The regional equations are given in Table 9.2.

Though the degree of calibration as given in Table 9.1 is low for W_{50} , T_B and time to peak, however, in the SUH development a great degree of subjectivity is involved in fitting the remaining points on the SUH. In addition, simultaneous adjustments are required for the area under the SUH to be unity corresponding to unit rainfall-excess. The empirical equations involve certain constants, which vary over wide range (Miller AC, Kerr SN, Spaeder DJ, 1983)

Snyder method slightly under-estimates the peak discharge and the discharge in the rising phase, increases the time to peak, over-estimates the recession part of the hydrograph, and consequently, increases the time base to conserve the flood hydrograph area. It is because the catchments are considered to be ungauged i.e. without any observed flood data, which might be advantageous in such cases.

Table 9.1 Physiographic characteristics of the Mahanadi Br-catchments

Br-No.	Name of Stream	A (Sq.Km.)	L (Km)	Lc (Km)	S(m/Km)
7	Bhedon	3108	96.60 (99.70)	51.84 (53.41)	0.61
121	Kelo	950	80.50 (83.33)	38.64 (39.37)	4.99
12	Lilagar	666	67.02 (66.82)	25.75 (26.41)	1.09
195	Jira	615	53.94 (49.58)	28.16 (27.44)	1.59
235	Ranjhor	312	41.06 (40.91)	21.09 (22.21)	1.61
385	Sandul	194	39.36 (38.68)	15.13 (15.29)	4.39
69	Borai	173	35.42 (35.76)	18.50 (19.21)	2.31
59B	Karwar	136	28.18 (26.35)	9.26 (9.15)	5.87
40K	Sargood	95	24.21 (24.21)	7.62 (12.27)	2.11
698	Bisra	93	26.57 (25.9)	14.40 (14.91)	9.04
154	Aherajhor	58	12.48 (12.45)	09.65 (07.71)	5.18

Table 9.2 Regional relationship for SUH derivation

Sl	UH parameters	Relationship	COD (R^2)
1	t_p	$1.723\ (LLc/\sqrt{S})^{0.25}$	0.61
2	q_p	$1.151\ (t_p)^{-0.731}$	0.72
3	W_{50}	$1.974\ (q_p)^{-1.105}$	0.57
4	W_{75}	$0.961\ (q_p)^{-1.113}$	0.62
5	T_B	$5.411\ (t_p)^{0.823}$	0.52
6	T_m	$t_p + t_r/2$	0.63
7	Q_p	$q_p * A$	0.71

9.6 CONCLUSIONS AND REMARKS

The study on SUHs review reported herein mainly concentrates on two aspects: (i) the traditional methods of SUH derivation, e.g., GIUH based Nash and Clark's approach, and Snyder's method, and (ii) SCS-CN and conventional methods using events data. . It can be inferred from the study that, albeit the methods of Snyder and GIUH are used widely in practical engineering problems, but the manual fitting of the characteristic points needed great degree of subjectivity and trial and error, and may involve error. Also each time it is not possible to get the area under the curve to be unity, which is the prerequisite for UH derivation. This questions the wider applicability and acceptability for practical field applications. On the other hand, the pdfs based SUH methods gives the complete shape of unit hydrograph, and the area under the curve is guaranteed to be unity. Some methods that those are frequently used is Nash model

which is similar to two parameter- Gamma distribution and Clark's model that has a conceptual basis on routing the hydrographs through isochronal areas. Well tested and applied relationships have been developed to estimate the parameters of pdfs. This strong mathematical perception and conceptual basis of pdfs successfully fills the technological niche for SUHs derivation. In a nutshell the summaries are briefed as below:

[1] Estimated constants in the regional relationships are transposed to the basin under study for the derivation of unit hydrograph parameters. Regional unit hydrograph for lower Mahanadi region based on Nash and Clark approach yielded the model parameters with the basin as follows: nk and $LL_c/S^{0.5}$ with $R^2 = 0.57$, K and L with $R^2 = 0.63$ and time of concentration (Tc) and $LL_c/\sqrt{S}$ with $R^2 = 0.99$.

[2] The regional unit hydrograph or synthetic unit hydrograph is a tool to overcome lack of stream flow data at the specific site under investigation. It is assumed that the unit hydrograph represents the physiographic characteristics of a basin, and multiple linear regression analysis is a most powerful tool for the regional unit hydrograph analysis. Data of unit hydrograph derived from stream flow and rainfall records in basins adjoining the basin under study should be used to estimate the constants for use to any regional unit hydrograph study. In this case, eleven catchments with relevant data are used for synthetic unit hydrograph methods used for regional analysis and the results given in Table 9.2 gives R^2 mostly more than 0.6 and a fair fit. In this relationships for Mahandi basin salient points in the unit hydrograph may be obtained with geo morphological basin characteristics that can be obtained either from SOI toposheet or RS imageries. The details of such cases are discussed in Chapter 8.

CHAPTER 10
FLOOD FREQUENCY ANALYSIS TECHNIQUES

10.0 INTRODUCTION

Hydrological processes are generally a chance, and time dependent processes. Probabilistic modeling considers only the probability of occurrence of an event with a given magnitude and uses probability theory for decision making. As the probabilistic modeling completely ignores the time dependence of the process, it can be used only for *design purposes* and *not for operational purposes*.

Probabilistic modeling or frequency analysis is one of the earliest and most frequently used applications of statistics in hydrology. Early applications of frequency analysis were largely in the area of flood flow estimation but today nearly every phase of hydrology is subjected to frequency analysis is to assume the specific probability distribution which the event is likely to follow and to proceed to evaluate the parameters of the distribution using the available data of the events to be modeled. Using the statistically derived sample estimates, probability levels can be assigned to any specific event and the prediction can be for the required probability event. Though the present work deals with the flood flow only, however, the concepts given can be extended for other hydrological variables also.

In fact, information on flood magnitudes and their frequencies is needed for design of hydraulic structures such as dams, spillways, road and railway bridges, culverts, urban drainage systems, flood plain zoning, economic evaluation of flood protection projects etc. Pilgrim and Cordery (1992) mention that estimation of peak flows on small to medium-sized rural drainage basins is probably the most common application of flood estimation as well as being of greatest overall economic importance. These estimates are required for the design of culverts, small to

medium-sized bridges, causeways and other drainage works, spillways of farm and other small dams and soil conservation works. For this purpose, statistical flood frequency analysis has been one of the most active areas of research since the last forty to fifty years. Flood frequency analysis is expected to provide solutions to some of the questions such as (i) which parent distribution the data may follow? (ii) What should be the most suitable parameter estimation technique? (iii) How to account for sampling variability while identifying the distributions? (iv) What should be the suitable measures for selecting the best fit distribution? (v) What criteria one should adopt for testing the regional homogeneity? The procedures of finding better solutions to the aforementioned questions have improved with the efforts made by the various investigators. The scope of frequency analysis would have been widened if the parameters of the distribution could have been related with the physical process governing floods. Such relationships, if established, would have been much useful for studying the effects of non-stationary and manmade changes in the physical process on frequency analysis. In spite of many limitations, the statistical flood frequency analysis remains the most important means of quantifying floods in systematic manner.

As such there are essentially two types of models adopted in flood frequency analysis literature: (i) *Annual Flood Series* (AFS) or generally used as *Annual Maximum Series* (AMS) models, and (ii) *Partial Duration Series* models (PDS). Maximum amount of efforts have been made for modeling of the annual flood series as compared to the partial duration series. In the majority of research projects attention has been confined to the AFS models. The main modeling problem is the selection of the probability distribution for the flood magnitudes coupled with the choice of estimation procedure. A large number of statistical distributions are available in literature. Among these are the *Gumbel Distribution* or *Extreme Value* –Type 1 (EV1), General *Extreme Value* (GEV), *Pearson Type* III (PT3), *Log Pearson Type* III (LP3),

Generalized Pareto (GP), *Generalized Logistic* (GL) and *Wakeby* (WAK) distributions those are being commonly used in most of the flood frequency studies. For the estimation of the various distributions parameters some commonly used methods are e.g. Graphical method, Method of least squares these are quite easier to apply, however, a slight recent i.e. for last ten years three methods are being used in may cases; they are as follows: Method of moments (MOM), Method of maximum likelihood (ML), method based on principle of maximum entropy (PME), Method of probability weighted moment (PWM). As said, PWM and ML are somewhat difficult for selected *pdf*s and are based on L-moment. These are some of the methods which have been most commonly used by many investigators in frequency analysis literature.

Once the parameters are estimated accurately for the assumed distribution, goodness of fit procedures then test whether or not the data do indeed fit the assumed distribution with a specified degree of confidence. Various goodness of fit criteria has been adopted by many investigators while identifying the best fit distribution from the various distributions fitted with the historical data. Some fitness criteria are based on bias or robust analysis (BIAS), root mean square error (RMSE) and standard error (SE). Among the three, bias criteria is better as it gives the tilt of the PDF graph both sides (+ve or –ve) from the normal centre point and also analyze with Monte Carlo simulations where the data are generated with the statistical parameters of the available data in the specified catchment. Step-wise procedure before taking up the flood frequency analysis in Mahandi data, a summarized is given in the next section for readers to understand, and for guide for new person to take up allied study.

The broad area of flood frequency analysis has been covered in the light of the following topics:

(i) Definitions

(ii) Assumptions and data requirement

(iii) Plotting positions

(iv) Commonly used distributions in analysis

(v) Parameter estimation techniques

(vi) Goodness of fit tests and

(vii) Estimation of T year flood and confidence limits

10.1 DEFINITIONS OF TERMS USED

a) *Peak annual discharge*: The peak annual discharge is defined as the highest instantaneous volumetric rate of discharge during a year.

b) *Annual flood series*: The annual flood series is the sequence of the peak annual discharges for each year of the record. This is also frequently used as Annual maximum Series (AMS)

c) *Design Flood*: Design flood is the maximum flood which any structure can safely pass. It is the adopted flood to control the design of a structure.

d) *Recurrence interval* or *return period*: The return period is the time that elapses on an average between two events that equal or exceed a particular level. For example, T year flood will be equaled or exceeded on an average once in T years.

e) *Partial flood series*: The partial flood series consists of all recorded floods above a particular threshold regardless of the number of such floods occurring each year. This is more a time used as Peak over Threshold (POT).

f) *Mean* (μ): Mean is a measure of central tendency. Other measures of central tendency are median and mode. Arithmetic mean is the most commonly used measure of central tendency and is given by

$$\bar{x} = \sum_{i=1}^{N} x_i / N \qquad (10.1)$$

Where xi is the i^{th} variants and N is the total number of observations.

g.) *Standard deviation* (σ): An unbiased estimate of standard deviation (Sx) is given by

$$S_X = (\sum_{i=1}^{N}(x_i - \bar{x})^2 / N-1)^{0.5} \quad (10.2)$$

Standard deviation is the measure of variability of a data set. The standard deviation divided by the mean is called the coefficient of variation (C_v), and is generally used as a regionalization parameter.

h) *Coefficient of skewness* (C_s): The coefficient of skewness measure the asymmetry of the frequency distribution of the data and an unbiased estimate of the Cs is given by

$$C_s = \frac{N\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)(N-2)S_x^{\ 3}} \quad (10.3)$$

i) *Coefficient of kurtosis* (C_k): The coefficient of kurtosis is C_k measures the *peakedness* or flatness of the frequency distribution near its center and an unbiased estimate of it is given by

$$C_k = \frac{N^2\sum_{i=1}^{N}(x_i - \bar{x})^4}{(N-1)(N-2)(N-3)S_x^{\ 4}} \quad (10.4)$$

j) *Probability paper*: A probability paper is a specially designed paper on which ordinate represents the magnitude of the variable and abscissa represent the probability of

exceedance or non-exceedance. Probability of exceedance, Pr ($X \geq x$), probability of non exceedance, Pr ($X \leq x$) and return period (T) are related as

$$P_r(X > x) = 1 - P_r(X \leq x)$$
$$P_r(X > x) = 1/T \qquad (10.5)$$

Plotting position formulae are used to assign probability of exceedance to a particular event.

10.1.1 Assumptions and Data Requirement:

The following three assumptions are implicit in frequency analysis.

(i) The data to be analyzed describe random events,

(ii) The natural process of the variable is stationary with respect to time and

(iii) The population parameters can be estimated from the sample data.

Data Requirement:

For flood frequency analysis (FFA) either annual flood series or partial duration flood series may be used. The requirements with regard to data are that

(i) It should be relevant,

(ii) It should be adequate and

(iii) It should be accurate.

The term relevant means that data must deal with problem. For example, if the problem is of duration of flooding then data series should represent the duration of flows in excess of some critical value. If the problem is of interior drainage of an area then data series must consist of the volume of water above a particular threshold. The term adequate primarily refers to length of data. The length of data primarily depends upon variability of data and hence there is no

guide line for the length of data to be used for frequency analysis. Generally a length of 30 – 35 years is considered adequate for flood frequency analysis. The term accurate refers primarily to the homogeneity of data and accuracy of the discharge figures. The data used for analysis should not have any effect of man made changes. Changes in the stage discharge relationship may render stage records no homogeneous and unsuitable for frequency analysis. It is therefore preferable to work with discharges and if stage frequencies are required then most recent rating curve is used.

Plotting Positions:

In order to see the fit of the distribution, the sample data is plotted on various probability papers. The plotting position formulae are required to assign probability of exceedance or non exceedance to a particular event. The general plotting position formula is given by

$$P(X \geq x) = \frac{m - a}{N + 1 - 2a} \qquad (10.6)$$

Where p(x>x) is the probability of exceedance, *m* is the rank of the event when arranged in descending order and N is the total number of observations. For the largest value m will naturally be 1 while for the lowest value it will be equal to N. Varies plotting position formulae have been proposed in the literature. Given below are some of the formulae and the values of *a* in the formula is given in the following table.

Table10.1a Plotting position formulae and the values of a

Formula	**Value of a**
Weibull	0
Blom	3/8
Gringortan	0.44

Table10.1b Plotting position formulae and the values of a

Distributions(s)	Recommended plotting position formula	Form of the plotting position formulae
Normal and Log Normal	Blom	F (X ≥ x) = (m-3/8)/ (N+1/4)
Gumbel EV-I and GEV	Gringorton and for GEV when for $\kappa \neq 0$,	F (X ≥ x) = (m-0.44)/ (N+0.12)
PT3 and Log PT3	Cunnane	F (X ≥ x) = (m-0.40)/ (N+0.20)
Generalized Pareto distribution	Cunnane	-do-
Generalized Logistic distribution	Cunnane	-do-
Wakeby distribution	Cunnane	-do-

Weibull is the most commonly used formula while *Blom* and *Gringortan* have been recommended for *Normal* and *Gumbel* distributions respectively. NERC (1975) recommends a = 0.4 for plotting position formula to be used for *Pearson type III* (PT3) and *log Pearson type III* (LP3) distributions. For GEV, GP, GL and Wakeby distributions Cunnane's plotting position formula is probably the most appropriate in that it is reported to be nearly unbiased with regard to flows for most distribution used in hydrologic applications (see Stedinger et al., 1992).

10.2 COMMONLY USED DISTRIBUTIONS

In flood frequency analysis (FFA) the sample data is used to fit probability distribution which in turn is used to extrapolate from recorded events to design events either graphically or analytically by estimating the parameters of the distribution. Some of the probability distributions which are commonly used in frequency analysis are explained in brief in subsequent sections.

Normal Distribution

The normal distribution is one of the most important distributions in statistical hydrology. This is a bell shaped symmetrical distribution having coefficient of skewness equal

to zero. The normal distribution enjoys unique position in the field of statistics due to central limit theorem. This theorem states that under certain very broad conditions, the distribution of sum of random variables tends to a normal distribution irrespective of the distribution of random variables, as the number of terms in the sum increases. The probability distribution functions *(pdf*) and cumulative distribution functions (*cdf*) for the frequently common distributions are given in Appendix 9(a).

Log Normal Distribution (LN)

The causative factors for many hydrologic variables act multiplicatively rather additively and so the logarithms of these variables which are the product of these causative factors follow the normal distribution If Y = loge(x) follows normal distribution, and then x is said to follow log normal distribution. If the variable x has a lower bound x0, different from zero and the variable Y = log(X-x0) follows normal distribution then X is log normally distributed with thee parameters.

Pearson Type III *Distribution (*PT3)

Pearson Type III distribution is a three parameter distribution. This is also known as Gamma distribution with three parameters. The *pdf* and *cdf* of the distribution are given in Appendix 9 (a).

Log Pearson Type III *Distribution* (LP3)

If $Y = \log_e(X)$ follows Pearson type III distribution then X is said to follow log Pearson type III distribution. In 1967, the U.S. Water Resources Council recommended that the log Pearson type III distribution should be adopted as the Standard flood frequency distribution by

all U.S. federal government agencies. The *pdf* and *cdf* of the distribution are given in Appendix 9 (a).

Gumbel (Extreme Value Type 1) Distribution (EV1)

One of the most commonly used distributions in flood frequency analysis is the double exponential distribution known as Gumbel distribution or extreme value type1 or Gumbel EV1 distribution. The *cdf* of EV-1 distribution is defined as

$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right] \qquad (10.7)$$

Where u and α are the location and scale parameters of the distribution. Using method of moments, u and α are obtained by following equation.

$$\bar{x} = u + 0.5772\alpha \text{ and } S_x^2 = \frac{\pi^2\alpha^2}{6} \qquad (10.8)$$

The above equations can be written in the reduced variate form as

$$F(x) = \exp[-\exp(-y)] \text{ , where, } y = (x-u)/\alpha \qquad (10..9)$$

The reduced variate *y* can be written in terms of return period (T), also by replacing (1-1/T) in place of F(*x*) giving the following relationship

Y = - ln(-ln (1-1/T)) = - ln (ln (T/(T-1))) or XT = u - α Ln [Ln (T/ (T-1))]

(10.10)

For Gumbel distribution or EV1, the coefficient of skewness is taken equal to 1.139. The *pdf* and *cdf* of this distribution and details are given in Appendix 14 (c).

General Extreme value Distribution (GEV)

The cumulative distribution function (*cdf*) for GEV distribution is given by Jenkinson (1969) as

$$F(q) = \exp\left[-\left(1-\kappa\frac{q-\xi}{\alpha}\right)^{\frac{1}{\kappa}}\right] \qquad \text{For } \kappa \neq 0 \qquad (10.11a)$$

$$F(q) = \exp\left[-\exp\left(-\frac{q-\xi}{\alpha}\right)\right] \qquad \text{For } \kappa = 0 \qquad (10.11b)$$

In the above equationξ, α and κ are location, scale and shape parameters respectively. The shape parameters (κ) and coefficient of skewness are interrelated. The notation q is the flood exceedance value, and the range of the variable *q* depends on the sign of κ. When κ is negative the variable *q* can take values in the range of $(\xi + \alpha/\kappa) \leq q < \infty$, making it suiTablefor flood frequency analysis; when κ is positive the variable *q* becomes upper bounded; and for κ = 0, the distribution reduces to a two-parameter extreme value type I (EV1) or Gumbel distribution. For EV1 distribution κ is 0, and coefficient of skewness is equal to 1.139. For EV2 distribution κ is negative and Cs is greater than 1.139 for EV3 distribution κ is positive and Cs is letter than 1.139.

Wakeby Distribution (WAK)

Houghton (1978) found that a significant majority of high quality flood records cannot be modeled adequately by conventional distributions and so advanced the Wakeby distribution for fitting food records.

$$X = \xi + (\alpha/\beta)\,[(1-(1-U)^{-\beta}] - (\gamma/\delta)[1-(1-U)^{-\delta}] \qquad (10.12)$$

Where U is a standard uniform random variable. That is, the above equation defines the percent point function for the Wakeby distribution. The parameters β, γ and δ are shape parameters. The parameter ξ is a location parameter and the parameter α is a scale parameter. The following restrictions apply to the parameters of this distribution:

Either $\beta + \delta \geq 0$

Either $\beta + \delta > 0$ or $\beta = \gamma = \delta = 0$

If $\gamma \geq 0$, then $\delta > 0$

$\gamma \geq 0$ and $\alpha+\gamma \geq 0$

The domain of the Wakeby distribution is

1. ξ to ∞
   if $\delta \geq 0$ and $\gamma > 0$
2. ξ to $\xi + (\alpha/\beta) - (\gamma/\delta)$
   if $\delta < 0$ or $\gamma = 0$

(10.13)

With three shape parameters, the Wakeby distribution can model a wide variety of shapes. The cumulative distribution function is computed by numerically inverting the percent point function given above. In a more simple way, the Wakeby distribution defined implicitly in inverse form as

$$x = m + a(1-Fb) - c(1-F-d) \qquad (10.14)$$

Where,

F is the non exceedance probability of x,

a, b, c and d are positive constants and

m is a location parameter.

The parameters a and b govern the left hand tail (drought) while parameter m, c and d govern flood flows. Being a five parameter distribution, the Wakeby distribution can span regions of distribution function space inaccessible to the conventional two or three parameter distributions, besides span regions occupied by the conventional distributions. This implies that

the Wakeby distribution can mimic the form of those two parameter distributions whose inverse form exists. Wakeby distribution is currently becoming popular in USA among researchers due to its capability to model both the tail ends of the data separately. Hydrologists and engineers in the past years have occasionally felt the need to go beyond three parameters but it was recognized that the use of moments higher than the third would introduce too much error into the estimation process. The estimation procedure development for the Wakeby distribution circumvents this problem.

Main features of the Wakeby distribution over the traditional distributions are

(i) In traditional estimation procedures, the smallest observation can have a substantial effect on the right hand side (large observations) of the distribution. But the left hand side/small observations do not necessarily add information to an estimation of quantile on the right hand side. Since floods are not known to follow any particular distribution, it seems intuitively better to model the left and right hand tails separately.

(ii) None of the traditional distributions have properties to reflect the nature of their left tails accurately. If in reality, the low observations follow the left hand tail of the low skewed distribution and highest observation follows the right hand tail of highly skewed distributions, then none of the conventional three parameter distributions will be able to model it accurately. They lack enough kurtosis for any given skew. Fitting a three parameter curve to a five parameter nature would distort the whole fit including the higher quintiles.

(iii) The separation effect which refer to the differences appear between samples of synthetic stream flow data and natural stream flow data when the standard deviation of skew is plotted versus the mean of skew for regional data, is explained by Wakeby distribution.

(iv) Given the correct choice of parameters, the Wakeby distribution can generate synthetic flows in the pattern of most of the conventional distribution

General Normal (GNO) distribution

The probability density function of the distribution and the cumulative density function of the three parameter GNO distribution as parameterized by Hosking and Wallis (1997) is given below.

$$F(x) = \phi\left[-k^{-1}\log\left(1-k(x-\xi)/\alpha\right)\right] \qquad k \neq 0$$

$$= \phi\left[(x-\xi)/\alpha\right] \qquad k = 0 \qquad (10.15)$$

Where,ξ, α and k are its location, scale and shape parameters respectively. When k = 0, it becomes normal distribution (NOR) with parameters ξ andα. This distribution has no explicit analytical inverse form. ϕ is the cumulative distribution function of the standard Normal distribution given by

$$\phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right)dx \qquad (10.16)$$

Generalized Pareto distribution (GP)

The GP distribution introduced by Pickands (1975), has the cdf

$$F(q) = 1 - \exp\left[-\frac{q-c}{\alpha}\right] \qquad ;\kappa = 0 \qquad (10.17a)$$

$$F(q) = 1 - \left[1 - \kappa\frac{q-c}{\alpha}\right]^{1/\kappa} \qquad ;\kappa \neq 0 \qquad (10.17b)$$

where α is the scale parameter, c is the location parameter and κ is the shape parameter, and q_0 is the threshold. For $\kappa = 0$, the exponential distribution (ED) is obtained as a special case.

Generalized logistic distribution (GL)

The flood magnitudes following a generalized logistic distribution is a special case of the Tukey's [1960; 1962] symmetric lambda distribution, and is given by

$$x = m + A\,F^{B} - C\,(1-F)^{D} \qquad (10.18)$$

where $F \equiv F(x) = P(X \le x)$, X is a variate, and A, B, C, D are parameters of the distribution. In the limit as $A \to \infty$ and $B \to 0$, (18) approaches the following form of GLD (Joiner et al., 1971; Ramberg 1975; Greenwood et al., 1979) given by

$$F = \left\{1 + \exp\left[-\frac{X-m}{\alpha}\right]\right\}^{-1} \qquad (10.19a)$$

where α ıs a scale parameter and m is a shape parameter of the distribution. *Dubey* [1969] derived the GLD as a mixture of EV1 (or Gumbel) distribution, and the general form of the distribution for using it as a flood exceedance model is given by *Ahmed et al*. [1988], *Hosking and Wallis* [1997], and Rao and Hamed (2000) as

$$F(q) = \left\{1 + \exp\left[-\frac{q-m}{\alpha}\right]\right\}^{-1} \qquad (10.19b)$$

In case of peak over threshold or partial duration series, the value of q is taken as a variable (q-q_0) with q_0 as threshold, and values are in the range $-\infty < (q - q_0) < \infty$, which means it does not have an upper bound. $q(T)$ can be obtained inserting the sample parameter estimates in (10.19b). The generalized logistic distribution given by *Hosking* [1986],

$$F = \left\{1 + \left[1 - k\frac{X - \varepsilon}{\alpha}\right]^{1/k}\right\}^{-1} \qquad (10.19c)$$

where α, is a scale parameter and k is a shape parameter of the distribution.

10.3 PARAMETER ESTIMATION TECHNIQUES

Various parameter estimation techniques which are in current use includes

(i) Graphical method

(ii) Least squares method

(iii) Method of moments (MOM)

(iv) Method of maximum likelihood (ML)

(v) Method based on probability weighted moments (PWM)

For further details lecture notes of Workshop on 'Flood Frequency Analysis', NIH (1987-88) may be consulted. For GEV and Wakeby distributions method based on probability weighted moments has proved to be the most robust method of parameter estimation.

10.3.1 Graphical Method

In graphical method of parameter estimation, the variate under consideration is regarded as a function of a reduced variate of a known distribution. The steps involved in the graphical method are as follows:

1. Arrange the variates of annual maximum flood series in ascending order and assign different ranks to the individual variates.
2. Assign the plotting positions to each of the variates. The plotting position formula may be used depending upon the type of distribution being fitted. The recommended plotting

position formulae for different distributions are given in Table. 10.1(b).

3. Estimate the reduced variates for the selected distribution corresponding to different plotting positions, which represent the probability of non-exceedance.

The reduced variates for normal and log normal distributions are computed with the help of Table10.1 (b). F (z) represents probability of non-exceedance which have been computed using the suitable plotting position formula, and z is the corresponding reduced variates. In case of Pearson type-III and Log Pearson type-III which are three parameter distributions different sets of reduced variates are obtained for different coefficients of skewness. At this stage it is essential to introduce the concept of frequency factor as it has been found easier to develop probability paper based on the frequency factor.

Frequency Factor

A general frequency equation proposed by Chow, applicable for different distributions are in the following form:

$$X = m + Ks \qquad (10.20)$$

Where, X = Magnitude of the different variates of the peak flood series, or the magnitude of the flood at required return period T. K = Frequency factor corresponding to X. For Pearson Type-III and Log Pearson Type-III distribution, the values of frequency factors are given in Appendix 9(c) at different probability of exceedance levels corresponding to the various coefficients of skewness values. Where, m and s = Mean and standard deviation of the population which would be replaced by the sample statistics.

The step by step procedure is as follows:

1. Plot the sample data as a series of discrete points on an ordinary graph paper with the

ordinate being the variate and the abscissa being the reduced variate or frequency factor. Such plot can be prepared for different distribution.

2. Draw a best-fit line through the plotted points. The slope and intercept of the line provide the estimates for the parameters of the respective distributions. This straight line can be projected to arrive at the flood magnitudes of desired return periods.

In graphical estimation procedure the line is subjectively placed and could vary with analyst. This subjectivity is regarded as a major drawback by hydrologists. The following example provides the procedure for estimating the reduced variate corresponding to a given probability level for Normal, EV1 and Pearson Type- III distributions:

10. 3.2 Least Squares Method

In the method of least squares for the parameter estimation the steps from (i) to (ii) of the Graphical method may be repeated. In this technique a simple linear regression equation is fitted between the variate under consideration and the corresponding reduced variate or frequency factor, rather than drawing a best fit line subjectively on a simple graph paper. This method has not been accepted as a standard method in practice as it involves the use of plotting position formula to determine the reduced variate or the frequency factor and due to the assumption that the error variance remains same for all observations. The defect due to the former assumption could be eliminated by using the appropriate plotting positions formula given in Table10.1 (b). However, the later assumption makes the method more defective as the higher events recorded have more error variance than the recorded lower events. All these assumptions affect the correct estimation of the slope and intercept of the line, which represent the parameters of the distribution.

10.3.3 Method of Moments (MOM)

The method of moments makes use of the fact that if all the moments of a distribution are known then everything about the distribution is known. For all the distributions in common usage, four moments or fewer are sufficient to specify all the moments. For instance, two moments, the first together with any moment of even order are sufficient to specify all the moments of the Normal Distribution and therefore the entire distribution. Similarly, in the Gumbel EV Type-I Distribution, the first two moments are sufficient to specify all the moments and hence the distribution. For Pearson Type III Distribution three moments, always taken as the first three required specifying all the moments. In these cases the number of moments needed to specify all the moments and hence the distribution equals the number of parameters.

The method of moment's estimation is dependent on the assumption that the distribution of variate values in the sample is representative of the population distribution. Therefore, a representation of the former provides an estimate of the later. Given that the form of the distribution is known or assumed, the distribution which the sample follows is specified by its first two or three moments calculated from the data.

10.3.4 Method of Maximum Likelihood (ML)

From a statistical point of view, the method of maximum likelihood estimation is, with some exceptions, considered to be the most robust of the parameter estimation techniques discussed here. This method is presented in this section for complete data, that is, data consisting only of single times-to-failure. As an example, consider the following data (-3, 0, 4) and assume that we are trying to estimate the mean of the data. Now, if one has to choose the most likely value for the mean from -5, 1 and 10, which one would be choosed? In this case, the most likely value is 1 (given the limit on choices). Similarly, under ML, one determines the most likely

value(s) for the parameter(s) of the assumed distribution. It is mathematically formulated as follows:

If x is a continuous random variable with *pdf:*

$$f(x;\theta_1,\theta_2,\ldots,\theta_k) \quad f(x;\theta_1,\theta_2,\ldots,\theta_k) \tag{10.21}$$

where θ_1, θ_2,..., θ_k are k unknown parameters which need to be estimated, with R independent observations, x_1, x_2,... x_R, which correspond in the case of life data analysis to failure times. The likelihood function is given by:

$$L(\theta_1,\theta_2,\ldots,\theta_k|x_1,x_2,\ldots,x_R) = L = \prod_{i=1}^{R} f(x_i;\theta_1,\theta_2,\ldots,\theta_k)$$

$$L(\theta_1,\theta_2,\ldots,\theta_k|x_1,x_2,\ldots,x_R) = L = \prod_{i=1}^{R} f(x_i;\theta_1,\theta_2,\ldots,\theta_k)$$

$i = 1, 2,..., R$ (10.22)

The logarithmic likelihood function is given by: (Note: Weibull provides a three-dimensional plot of this log-likelihood function.)

$$\Lambda = \ln L = \sum_{i=1}^{R} \ln f(x_i;\theta_1,\theta_2,\ldots,\theta_k) \quad \Lambda = \ln L = \sum_{i=1}^{R} \ln f(x_i;\theta_1,\theta_2,\ldots,\theta_k) \tag{10.23}$$

The maximum likelihood estimators (or parameter values) of θ_1, θ_2,..., θ_k are obtained by maximizing L or Λ. By maximizing Λ, which is much easier to work with than L, the maximum likelihood estimators (ML) of θ_1, θ_2,..., θ_k are the simultaneous solutions of k equations such that:

$$\frac{\partial \Lambda}{\partial \theta_j} = 0, \; j = 1,2,\ldots,k \quad \frac{\partial \Lambda}{\partial \theta_j} = 0, \; j = 1,2,\ldots,k \tag{10.24}$$

Even though it is common practice to plot the ML solutions using median ranks (points are plotted according to median ranks and the line according to the ML solutions), this is not completely representative. As can be seen from the equations above, the ML method is independent of any kind of ranks. For this reason, the ML solution often appears not to track the data on the probability plot. This is perfectly acceptable since the two methods are independent of each other, and in no way suggests that the solution is wrong. Example of The GEV distribution using ML is as follows

Defining $y = 1 + \exp\left(-\frac{q_0 - m}{\alpha}\right)$ the derivatives of the log-likelihood functions are given by

$$\frac{\partial \log(L)}{\partial m} = -\frac{N}{\alpha} + \frac{2}{\alpha}\sum_{i=1}^{N} y_i^{-1} = 0$$

$$\frac{\partial \log(L)}{\partial \alpha} = -\frac{N}{\alpha} - \frac{1}{\alpha^2}\sum_{i=1}^{N}(x_i - m) + \frac{2}{\alpha^2}\sum_{i=1}^{N}(x_i - m)y_i^{-1} = 0$$

Equation is solved numerically to obtain the values of α and m.

10.3.5 Probability Weighted Moments and L-Moments (PWM)

The L-moments are an alternative system of describing the shapes of probability distributions (Hosking and Wallis, 1997). They arose as modifications of probability weighted moments (PWMs) of Greenwood et al. (1979). Probability weighted moments is defined as:

$$M_{p,r,s} = E\left(x^p \{F\}^r \{1-F\}^s\right) = \int_0^1 \{x(F)\}^p F^r \{1-F\}^s \, dF \qquad (10.25)$$

where, F = F(x) is the cumulative distribution function (CDF) for x, x(F) is the inverse CDF of x evaluated at the probability F, and p, r and s are real numbers. If p is a nonnegative integer, $M_{p,0,0}$ represents the conventional moment of order p about the origin. If p = 1 and s = 0,

$$M_{1,r,0} = \beta_r = \int_0^1 x(F) F^r dF \qquad (10.26)$$

For an ordered sample $x_1 \le x_2 \ldots . \le x_N$, $N > r$, the unbiased sample PWM's are given by

$$\hat{\beta}_r = \frac{1}{N} \frac{\sum_{i=1}^{N} \binom{i-1}{r} x_i}{\binom{N-1}{r}} \qquad (10.27)$$

For any distribution the r^{th} L-moment λ_r is related to the r^{th} PWM (Hosking, 1990), through:

$$\lambda_{r+1} = \sum_{k=0}^{r} \beta_k (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \qquad (10.28)$$

These L-moments are linear functions of PWMs. For example, the first four L-moments are related to the PWMs using:

$$\lambda_1 = \beta_0$$

$$\lambda_2 = 2\beta_1 - \beta_0$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \qquad (10.29)$$

The L-moments are analogous to their conventional counterparts as they can be directly interpreted as measures of scale and shape of probability distributions and hence, are more convenient than the PWMs. Hosking (1990) defined L-moment ratios which are analogous to conventional moment ratios as:

L-coefficient of variation, L-CV: $\tau_2 = \lambda_2 / \lambda_1$

L-coefficient of skewness, L-skew: $\tau_3 = \lambda_3 / \lambda_2$ (10.30)

L-coefficient of kurtosis, L-kurtosis: $\tau_4 = \lambda_4 / \lambda_2$

Analogous to the conventional moment ratios, λ_1 is a measure of location, τ_2 is a measure of scale and dispersion, τ_3 is a measure of skewness and τ_4 is a measure of kurtosis. Hosking (1990) showed that for $x \geq 0$, the value of τ_2 lies between 0 and 1, while the absolute values of τ_3 and τ_4 lie between 0 and 1. This restriction in the values of the L-coefficients works out to be an advantage in their interpretation as opposed to the conventional moments which do not have any bounds (Rao and Hamed, 2000).

10.4 GOODNESS OF FIT TESTS

The validity of a probability distribution function proposed to fit the frequency distribution of a given sample may be tested graphically or by analytical methods. Graphical approaches are usually based on comparing visually the probability density function with the corresponding empirical density function of the sample under consideration. In other words model CDF is compared with empirical CDF. Often these CDF graphs are made on specifically designed paper such that the model CDF plots as a straight line. An example of this is the Gumbel paper. If empirical CDF plots as a straight line on the Gumbel paper it is an indication that the Gumbel distribution may be valid model for the data at hand. Often graphical approaches for judging how good a model is are quite subjective.

A number of analytical tests have been proposed for testing the goodness of fit of proposed models. Some of these tests are presented here under.

(a) *Coefficient of determination* (r^2): It describes the extent of best-fit and is expressed as:

$$r^2 = \left[\sum(x_i - \bar{x})\ (y_i - \bar{y})\right]^2 / [\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2] \quad (10.31)$$

(b) *Coefficient of correlation* (r): It is the square root of the coefficient of determination (r^2):

$$r = \sqrt{r^2} = \sum(x_i - \bar{x})(y_i - \bar{y}) / [\sum(x_i - \bar{x})^2 \sum(x_i - \bar{y})^2]^{1/2} \quad (10.32)$$

Where 'x' is the observed and 'y' is the computed value.

(c) Efficiency (EFF): The efficiency of the best-fit is given as

$$EFF = 1 - \frac{S}{S_y} \quad (10.33)$$

where,

$$S^2 = \sum(y_i - \hat{y}_i)^2 / (n - 2) \quad \text{and} \quad S_y^2 = \sum(y_i - \bar{y})^2 / (n - 1) \quad (10.34)$$

where 'n' is the total number of observations.

(d) *Standard error* (SE): Standard errors of the estimated regression coefficients a and b are computed, respectively as:

$$S_a = S\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right)^{1/2} \quad (10.35)$$

$$S_b = S / \sum(x_i - \bar{x}^2)^{1/2} \quad (10.36)$$

where S_a and S_b are standard errors of coefficients a and b, respectively.

(e) *Confidence interval*: The confidence interval for a is given as:

$$l_a = a - t_{(1-\alpha/2),\,(n-2)}\ S_a \quad (10.37)$$

$$u_a = a + t_{(1-\alpha/2),\,(n-2)}\ S_a \quad (10.38)$$

and for b, it is given as:

$$l_b = b - t_{(1-\alpha/2),\,(n-2)}\ S_b \quad (10.39)$$

$$u_b = b + t_{(1-\alpha/2),\,(n-2)}\, S_b \qquad (10.40)$$

where l_a, u_a and l_b, u_b denote lower and upper confidence limits of a and b, respectively; α is the confidence level; and $t_{(1-\alpha/2),\,(n-2)}$ represent t-values corresponding to (1- α/2) confidence limits and (n-2) degrees of freedom.

10.5 ESTIMATION OF T- YEAR FLOOD

T- Year flood estimated can be obtained either graphically or analytically. Graphically approach is applicable only for normal, log normal and Gumbel EV1 distribution as for other distributions probability papers are not readily available. The main drawback of graphical method is that different engineers will get different estimates for T year flood. Analytical approach of estimating T year flood for following cases is given in subsequent sections.

Generalized logistic distribution $Q(T) = \xi\left[1 + \frac{\beta}{\kappa}\left(1 - (T-1)^{-\kappa}\right)\right]$

In the above table, Z_T corresponds to the probability of non-exceedance. For the sample data series, $\overline{Y}$ is the mean, and standard deviation is denoted as S_Y.

Table 10.2 T- year flood using various distributions.

Sl.	cdf	Flood with return period (T) flood, Q(T)
1		$Y_T = \mu_Y + \sigma_Y Z_T$, $Q(T) = e^{Y_T}$
2	Pearson Type 3 Distribution (PT3)	$Q(T) = \overline{X} + S\,K_T$
3	Log Pearson Type 3 Distribution	As PT3 with Log transformed series.
4	Gumbel (EV1) Distribution	$Z_T = -\ln(-\ln F(Z))$, $Q(T) = u + \alpha Z_T$
5	General Extreme value Distribution	$Q(T) = \xi + \alpha/\kappa\left(1 - (-\log(F))^{\kappa}\right)$, F=1-1/T
6	Wakeby (WAK) distribution	$Q(T) = \xi + \alpha/\beta\left(1 - (1 - F)^{\beta}\right) - \gamma/\delta(1 - (1-F)^{-\delta})$
7		$F(x) = \phi\left[-k^{-1}\log(1 - k(x-\xi)/\alpha)\right]$, $k \neq 0$
8	Generalized logistic distribution	$Q(T) = 1 - ((1-F)/F)^{k}$, $k \neq 0$ $Q(T) = 1 - ((1-F)/F)$, k=0
9	Generalized Pareto distribution (GP)	$Q(T) = \xi + \alpha/\kappa\left(1 - (1-F)^{\kappa}\right)$ for c =0

The notation μ_Y and σ_Y are the mean, and standard deviation from the log transformed series. For the Gumbel (EV1) Distribution, $\alpha = 0.7797\ \sigma$ and $u = \mu - 0.45\ \sigma$. These are used to estimate the parameter α. For Pearson Type 3 Distribution (PT3), K_T from the Table corresponding to the computed coefficient of skewness g, and the probability of exceedence. Coefficient of skewness is determined as follows:

$$g = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \frac{(X_i - \overline{X})^3}{S} \qquad (10.41)$$

10.6 CONFIDENCE ITNERVALS

Hydrologic variables such as annual peak floods or rainfalls do not occur in a set pattern and are mostly random. In modeling these events, the help of frequency analysis is taken such that the estimate of these hydrological variables for a desired return period can be estimated with a reasonable accuracy. The estimates, usually, arrived from a single set of sample data are variable because of randomness associated with these events and the size of the sample used for arriving at the estimates.

Moreover, the sample under consideration is assumed to have resulted from a specific parent population and is random. This results in the fact that there are many equally likely possible samples that can originate from this assumed population. If estimates of the variables for all such samples for the desired return period are plotted against the return period, they seem to follow a normal or t-distribution with its mean as the expected value of the variables at that return period

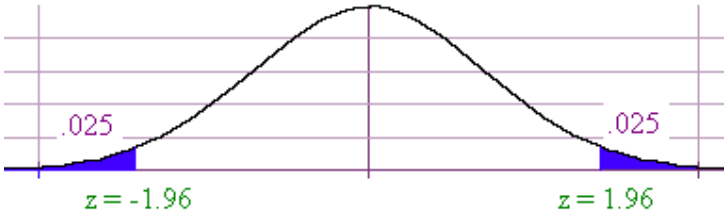

Figure 10.1 Normal distribution

(Figure 10.1) This therefore indicates that due to sampling variation there can be many estimates and therefore, should be defined through a continuous run of estimates rather than single or point value of the estimate. This range is defined as confidence interval and can written as

$$\text{Prob}\,[x(T)_L \le x(T) \le x(T)_U] = 1 - \alpha \qquad (10.42)$$

Where $x(T)_L$ and $x(T)_U$ are lower and upper confidence limits of the estimate x (T) so that the interval $x(T)_L$ to $x(T)_U$ is the confidence interval and $(1-\alpha)$ is the confidence level (α = significant level). However, the confidence level based on probability values give rise to the limits on either side of curve developed by frequency analysis to indicate the reliability of the estimates as well as the fit. This concept along with develop of confidence bands coupled with detail examples are given in detain in Chapter 12.

10.7 RESULTS AND DISCUSSION

10.7.1 Date Base and Homogeneity of the Observed Data

The overall plan of study involves a significant effort to develop the data and models needed for this evaluation. Briefly, this resulted the significant additional period of record available since the last study (approximately 30-years of additional record); the occurrence of the great flood of 1993; and finally the potential limitations of methods in CWC (1982, 1983),

the standard guidelines (as discussed in the first section of this chapter) for performing flood frequency analysis, for application to the small and large basins involved in the Mahanadi Basin Study.

This section describes the results of a study where the methods described in the guidelines are evaluated in comparison with other flood frequency estimation techniques for application to small catchments i.e. the bridge catchments considered in this study in addition to the CWC eight gauging sites engulfing the large watersheds. The study was deemed to be necessary because the earlier studies used to develop the methods focused on drainage areas significantly smaller than those of interest in the Mahanadi Basin. This limitation is apparent from the regional skew map
(Chapter-2 and 3) observed in the guidelines which is recommended for applications to drainage areas less than 1,000 square kilo meters.

The flood flow frequency estimation methods investigated do not ensure flood quantile estimates that consistently vary along the study Area Rivers. Chapter 3 describes investigations into algorithms appropriate for obtaining regularly varying flood quintiles. For this case F-test, t-test and other standard methods were used for testing the stationary or homogeneity of the observed data is a key assumption in a flood frequency analysis. The estimated unregulated flood record for the Mahanadi basin may deviate from this assumption either due to the influence of land use change, channel change or the climatic variability. Standard statistical tests were applied to the period of record to determine if any of these influences might cause a deviation from the standard assumption as is described in Chapter 3 that provides some concluding remarks.

10.7.2 Data Collection and Analysis

Before the estimation can be done by using the FFA, analyzing the rainfalls and stream flows data are important in order to obtain the probability distribution of flood and other phenomenon related to them. By knowing the probability distribution, prediction of flood events and their characteristics can be determined.

However, their used are often being complicated by certain characteristics of the data such as the skewness and range of variation. Daily maximum instantaneous flow data (in m^3/s) covering from 10 water years to 38 water years were recorded for 14 gauging sites of Mahanadi Basin (given in Appendix-2 and details in Table 10.7) and 23 Small Bridge catchments from CWC (1982) as given in Appendix -1. Flow data were expressed in terms of exceedance probabilities and recurrence intervals. Denoting *Qi* as the annual maximum flood in year *i*, the quantile Qi (F) is the value expected Qi to exceed with probability F, that is, P(Qi ≥ Qi (F)) = F during the year of interest. Thus, there is a *F*% chance that $Q \geq Q(F)$. Conversely, there is a (1−F) % chance that X < Q (F). The return period of a flood, *1/(F)* is the reciprocal of the probability of exceedance in one year (Haan, 1977; Shaw, 1983).

10.7.3 Estimation Methods and Distributions

Distributions selected for testing corresponded to the standard two and three parameter ones described in the literature (see Table 10.3). Additionally, the five parameter Wakeby distribution was selected because it has been often applied in combination with L-moment estimation procedures.

Method of moments (MOM)

Skewness (G) = $2(1-\kappa)(1+2\kappa)^{1/2} / (1+3\kappa)$

With κ calculated, > and α follow from equation as follows:

$S^2 = \alpha^2 / [(1 + \kappa)^2 (1 + 2 \kappa)]$

$\alpha = S (1 + \kappa) (1 + 2\kappa)^{1/2}$ and $\zeta = Mean - \alpha / (\kappa + \alpha)$

Table 10.3 Distributions used in comparisons

Distribution	**Cumulative distribution or density function**	**Transform**	**Parameters**
Gumbel (GEV k=0)	y=-log(-log(F))	y=(x-u)/□	u, ,
Generalized Extreme Value (GEV)	$y = 1 - \left(- (\log F^{\kappa})\right) / \kappa$ □≅	y=$\frac{x-\xi}{\alpha}$	, >, κ
Generalized Pareto (GP)	y=1-(1-F)^k k ≠ 0 y=-log(1-F) k=0	**y =(x-c)/□**	, c, κ
Generalized Logistic (GL)	y=1-{(1-F)/F}^{k} k ≠ 0 y=log{(1-F)/F} k=0	y=$\frac{x-\xi}{\alpha}$	, >, κ
log-Normal (LN)	f(y)= $\frac{1}{\sqrt{2\pi}} e^{-y^2/2}$	y=$\frac{\log(x)-\mu}{\sigma}$	μ, σ
General Normal (GNO) distribution	$F(x) = \phi\left[-k^{-1}\log\left(1-k(x-\xi)/\alpha\right)\right]$, k ≠ 0	--	α, ζ, κ
log-Pearson III (LP3)	f(y)=$\frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}$	y=$\frac{\log(x)-\xi}{\beta}$	
Wakeby (WAK)	x=$\xi+\frac{\alpha}{\beta}\{1-(1-F)^{\beta}\}-\frac{\gamma}{\delta}\{1-(1-F)^{-\delta}\}$	----------	

[1]F is the cumulative probability (non-exceedance probability), f(y) is probability density function

[2]x is the quantile or flow value of interest

The parameter estimates for Generalized Pareto distribution is given as follows:

Probability-weighted moments (PWM)

$\kappa = (W_0 - 8W_1 - 9W_2) / (-W_0 + 4W_1 - 3W_2)$

$\alpha = (W_0 - 2W_1) (W_0 - 3W_2) (-4W_1 - 6W_2) / (-W_0 + 4W_1 - 3W_2)^2$

$> = (2W_0W_1 - 6W_0W_2 + 6W_1W_2) / (-W_0 + 4W_1 - 3W_2)$

where the rth probability-weighted moment W*r* is given earlier in section 10.3.5

The distributions and estimation procedures were evaluated using two different criteria. The first criterion investigated how well the distribution fit the entire data set (full period of

record). Basically, this judges the "curve fitting" capability of the combined distribution and estimation procedure. A second criterion employing split sample testing was selected because of its use in the original investigation to select the methods. In split sample testing the distribution is estimated from half of the data and then evaluated based on a comparison with plotting positions for the remaining half of the data. The divided data sets were obtained by examining: 1) a forecast (estimate parameters for first half and compare predictions to second half); 2) a hind cast (estimate parameters for second half and compare predictions to first half); 3) an alternating odd sequence (estimate parameters from earliest first, third, etc. values and compare to remaining data); and finally, 4) an alternating even sequence (estimate parameters from earliest second, fourth, etc. values and compare to remaining data). The different methods for dividing the data sets were selected to ascertain if trends in the flow series influence the comparison results.

Another criteria that is followed here in this analysis is the goodness-of-fir using error criteria that are already discussed in Chapter 3. They are related with error measures used in comparing quantiles were average bias, average absolute relative error and mean square error. Comparisons were made for predicted flood quantiles and exceedances. The computation of the distribution predicted flow values was accomplished using a variety of different programs and statistical libraries. However, some approximations were needed to correct estimators for bias in comparing the various methods tested. Consequently, a comparison of quantile predictions was also made to attempt to avoid, or at least obtain a different perspective of estimator bias.

10.7.4 L-moment

In statistics, L-moments are statistics used to summarize the shape of a probability distribution. They are analogous to conventional moments in that they can be used to calculate

quantities analogous to standard deviation, skewness and kurtosis, termed the L-scale, L-skewness and L-kurtosis respectively (the L-mean is identical to the conventional mean). Standardized L-moments are called L-moment ratios and are analogous to standardized moments.

Table 10.4 Catchment area, sample statistics, sample size and discordancy measure for 14 gauging sites of Mahanadi Subzone 3(d)

Stream Gauging Site	Catchment Area (km^2)	L(Length of stream)	L-Ratio	L-CV (τ_2)	L-skew (τ_3)	L-kurtosis (τ_4)	Qav	Discordancy Measure (D_i
Altuma	830	265	1.182	0.250	0.085	0.133	486.10	1.046
Anandpur	8570	78	1.40861	0.317	0.110	0.116	3781.46	0.699
Champua	1710	76	0.29605	0.274	0.112	0.154	620.69	0.973
Gomlai	21950	52	3..1176	0.279	0.071	0.112	5672.83	0.978
Jaraikela	9160	48	3.97569	0.374	0.336	0.235	3471.5	0.143
Kantamal	19600	115	1.48204	0.228	0.077	0.016	6130.6	0.909
Jenapur	33955	278	0.43935	0.363	0.066	-0.016	8284.54	1.171
Kesinga	11960	94	1.35355	0.458	0.248	0.088	6991.7	1.050
Pamposh	19448	89	2.45525	0.250	0.106	-0.018	6863.8	1.270
Salebhata	4650	49	1.93669	0.456	0.367	0.315	2444.7	1.409
Sukuma	1365	37	0.99708	0.521	0.306	0.172	709.14	1.481
Sundergarh	5870	142	0.29111	0.410	0.482	0.352	2277.30	0.603
Talcher	29750	174	0.98263	0.186	0.328	0.291	4879.52	1.263

L-moments differ from conventional moments in that they are calculated using linear combinations of the ordered data; the "l" in "linear" is what leads to the name being "L-moments". Just as for conventional moments, a theoretical distribution has a set of population L-moments. Estimates of the population L-moments (sample L-moments) can be defined for a sample from the population. If the ordinary standard deviation of this data set is taken it will be highly influenced by this one point: however, if the L-scale is taken it will be far less sensitive to this data value. The detail of the catchment characteristics along with the standardized L-moments for fifteen gauging sites referred in this chapter is given ion Table 10.4.

A number of favorable comparisons of L-moments with ordinary moments have been reported in Chapter-14 dealing with regional flood frequency methods.. A graphical plot between L-Skew and L-Kurtosis and between L-Skew and L-Kurtosis for 23 Bridge Catchment of Mahanadi Sub zone is shown in Figure 10.2. These values are used in heterogeneity measure for the Mahandi zone and are discussed in Chapter-14.

10.7.5 Heterogeneity and Discordancy of the AMS Data in Catchments

The values of the heterogeneity measures H(1), H(2) and H(3) were computed utilizing the data of 14 gauging sites by generating many regions using the fitted Kappa distribution. Using the

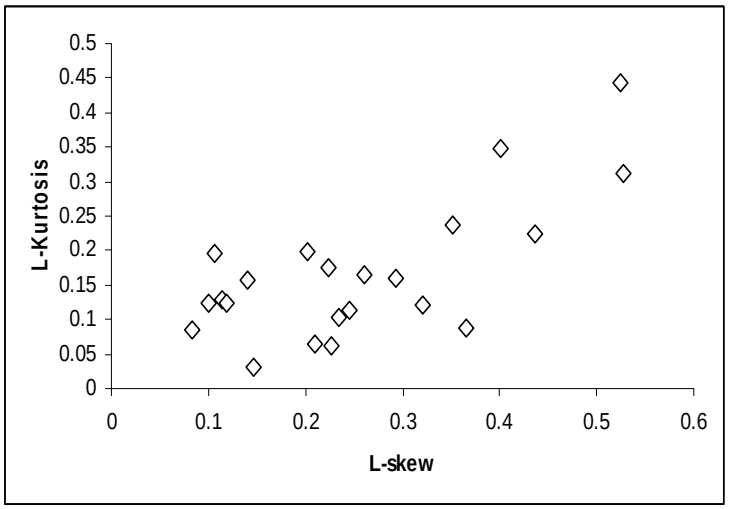

Figure 10.2a Graph between L-Skew and L-Kurtosis for 23 Bridge Catchment of Mahanadi Sub zone

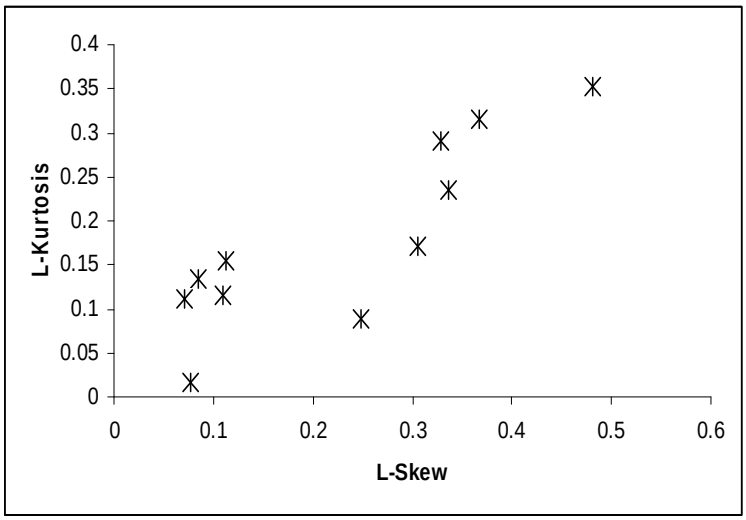

Figure 10.2b Graph between L-Skew and L-Kurtosis for 13 Gauging Sites of Mahanadi.

data of 14 sites, H(1), H(2) and H(3) values are computed as 1.68, -0.71 and -1.98, respectively. As H(1) is greater than 1.0 and H(2) is less than 1.0, the region defined by the 14 gauging sites is considered as heterogeneous as given in Table 10.5. Thus, based on the statistical properties of the data of the gauging sites, one by one seven sites of the region are excluded till H(1) value between 1.0 and 2.0; and both H(2) and H(3) values less than 1.0 is obtained. Further efforts to reduce the value of H(1) led to significant loss of data and hence, this was not attempted. The details of catchment data and statistical parameters including the discordancy measure, for the 14 gauging sites are given in Table 10.4. The values of heterogeneity measures computed by carrying out 500 simulations using the Kappa distribution based on the data of 14 sites are given in Table 10.5. Values of Discordancy statistic have been computed in terms of the L-moments for all the 14 gauging sites of the study area. It is observed that the D_i values for all the 14 sites vary from 0.143 to 1.481 (with 0.143 that is minimum for *Jariakela* and of 1.481 is computed for *Sukuma*.). Similarly, Di values estimated for 23 Bridge Catchment (Table 3.6 and Appendix-1) vary from 0.309 to 2.102.

Table 10.4 Heterogeneity measures for 14 gauging sites of Mahanadi Subzone 3(d)

Sl. No.	Heterogeneity measures	Values
1.	Heterogeneity measure H(1)	
	(a) Observed standard deviation of group L-CV	0.0615
	(b) Simulated mean of standard deviation of group L-CV	0.0472
	(c) Simulated standard deviation of standard deviation of group L-CV	0.0085
	(d) Standardized test value H(1)	1.68
2.	Heterogeneity measure H(2)	
	(a) Observed average of L-CV / L-Skewness distance	0.0861
	(b) Simulated mean of average L-CV / L-Skewness distance	0.0980
	(c) Simulated standard deviation of average L-CV / L-Skewness	0.0167
	(d) Standardized test value H(2)	-0.71
3.	Heterogeneity measure H(3)	
	(a) Observed average of L-Skewness/L-Kurtosis distance	0.0821
	(b) Simulated mean of average L-Skewness/L-Kurtosis distance	0.1198
	(c) Simulated standard deviation of average L-Skewness/L-Kurtosis distance	0.0191
	(d) Standardized test value H(3)	-1.98

10.7.6 Estimated Parameters

Before the analysis can be done, the parameter for each selected distribution needs to be estimated first. In this study, MOM, PWM and ML approaches were used to estimate the parameter of the selected distribution those considered are : GL, GP, GEV, EV1, PT3, LP3 and WAK distributions since they are commonly used in flood frequency analysis. In general, the number of moments needed equals the number of distribution parameters. Table 10.5 shows the parameter estimates for the selected distribution using MOM and PWM and ML approaches. The distribution functions F_X or f_X have parameters that need to be estimated based on observed data.

The table also gives the estimated parameters using MOM, ML and PWM approaches using Gumble (EV1) distribution for the 13 GD sites. For one catchment the length of data was not enough to get the proper parameter estimation and that is excluded from the table. Similar type of parameter estimation for different distribution is given in Appendix 9(b). The estimated parameters for all the considered seven distributions using the AMS data for 23 Bridge catchments are also given in Appendix 9(b).

Table 10.5 Estimated parameters using Gumbel (EV1) *d*istribution for 13 GD Sites

GAUGING SITES	Parameters of the Gumbel (EV1) Distribution					
	MOM		ML		PWM	
	α	u	α	u	α	u
Altuma	0.0053	347.9	0.0047	344.8	0.0051	347.12
Anandpur	0.0004	833.26	0.0005	1144.4	0.0005	987.99
Champua	0.0075	62.81	0.010	79.89	0.0087	70.89
Gomlai	0.0533	37.18	-	-	0.0267	18.01
Jaraikela	0.0005	641.13	0.0007	876.64	0.0006	759.05
Kantamal	0.0003	1015.0	0.0004	1399.0	0.0003	1206.55
Jenapur	0.0007	485.6	0.0010	660.09	0.0008	571.99
Kesinga	0.0152	52.63	0.0202	61.06	0.0178	57.55
Pamposh	0.0002	2012.5	0.0002	2791.9	0.0002	2401.87
Salebhata	0.0679	16.15	0.0899	18.04	0.0790	17.83
Sukuma	0.0067	63.9	0.0089	82.97	0.0079	72.97
Sundergarh	0.2591	33.27	-	-	0.1300	16.39
Talcher	0.0004	860.63	0.0005	1191.9	0.0004	1028.01

10.7.7 Return Period Flood

For GL, GP, GEV, EV1, PT3, LP3 and WAK respectively, results of the 14 Gauging Sites high flow frequency analyses are provided in Appendix 9(c), for recurrence intervals of 20, 50, 100 200, 300 and 500 years. For which the area are between the range of 830km^2 and 33955km^2.

For various GD site the corresponding return period flood are given in Appendix 9(c). In Table 12.6 we have taken return period flood for the data of Anandpur GD Site. In the above

Table we can see that the value of return period flood for 20 years for LP3 is 5409.70 and for 500 years the value is 17261.10 for MOM. Here you can see a great increase in the value of MOM similarly for ML and PWM same increase in value can be noticed. Similar type of calculation for Return Period Flood for different distribution can be find in Appendix 9(c). Similarly for 23 Bridge Catchment the estimated parameter for the similar distribution is also given in Appendix 9(c).

10.7.8 Goodness –of-Fit Test with Predictive Ability Test

Simulation is designed to compare quantiles from generated samples with those from known population. A flood frequency procedure, i.e. combined distribution/ parameter estimation, is considered robust if the difference between the average quantile estimated from generated data and that from known population is minimum. A predictive ability test (Bhunya, 2000) was employed for all the above regions with the above three distributions of which parameters were estimated using L-moments. The model yielding the lowest bias for high return periods is taken as the robust combination of parameter estimation and distribution for that region. On the basis of the Z-test, L-moment diagram, and bias, the procedures selected for the delineated region are summarized in Table 10.8.

Table10.6 Return Period Flood distributions using MOM, ML and PWM estimation methods for Anandpur GD site

T in years	Log Pearson Type 3 Distribution			General Extreme value Distribution		
	MOM	ML	PWM	MOM	ML	PWM
20	5049.70	8231.80	5090.10	8758.18	7885.32	8321.75
50	5831.00	13625.00	9020.10	9392.75	8610.75	9001.75
100	6998.60	19108.50	7880.90	10982.05	10319.22	10650.63
200	9602.10	8907.20	7992.40	12307.55	12856.50	12582.03
300	12155.10	11250.90	10984.40	14698.79	15421.30	15060.05
500	17261.10	15678.90	12004.90	18678.17	17875.30	18301.70
T in years	**Wakeby (WKB) distribution**			**Pearson Type 3 Distribution (PT3)**		
	MOM	ML	PWM	MOM	ML	PWM
20	6615.23	6885.33	7155.42	8090.60	8371.35	8652.09
50	7623.45	8010.75	8398.05	9708.30	10193.15	10678.00
100	8765.57	9319.22	9872.87	10889.00	10944.04	10999.07
300	12356.34	12821.52	13286.69	12423.00	12834.40	13245.80
500	16987.98	17421.31	17854.64	13754.00	14693.00	15632.00

T in years	Generalized Pareto Distribution			General Logistic Distribution		
	MOM	ML	PWM	MOM	ML	PWM
20	5805.206	6387.452	5966.79	8512.32	8663.46	8714.59
50	6653.724	8019.944	7580.984	9713.12	9960.07	10607.02
100	7678.342	8198.745	8050.484	10156.49	10102.72	11348.95
200	9506.891	9659.573	9603.584	11845.46	11718.23	12611.00
300	11709.819	11950.894	12102.677	12953.60	13567.93	14582.26
500	13459.714	13997.236	14362.238	13559.30	14631.95	15804.59
T in years	**EV1 or Gumbel Distribution**					
	MOM		ML		PWM	
20	8654.432		8259.425		8154.425	
50	9481.177		9949.131		9858.142	
100	11206.87		11197.79		11134.84	
300	13151.781		13163.800		13149.810	
500	14179.089		14134.088		14085.090	

R.M.S. Error. LP3	299.056
R.M.S. Error GEV Method	276.585
R.M.S. Error WAKEBY DISTRIBUTION	251.56
R.M.S. Error PE3	315.06
R.M.S. Error GPA	322.19
R.M.S. Error GL	279.013
R.M.S. Error EV1 Method	343.214

Table 10.8 Simulation Results for the Goodness of Fit Measure

Distribution	Z^{DIST} A	T4[B]	% times Accepted	RANK[C]	
				Z-Test	LM-Diagram
USING AM SERIES					
PT3	-22.45	0.129	0	3	2
GP	-3.67	0.129	31	1	3
GEV	4.11	0.129	21	2	1
USING POT SERIES					
PT3	-7.3	0.146	0	2	3
GP	-2.11	0.146	37	1	1
GEV	-0.62	0.146	29	3	2

Table 10.9 Simulation Results for the Whole Region (23 Catchments)

Distribution	Z^{DIST} *	T4**	% times Accepted	RANK	
				Z-test	LM-Diagram
PT3	-21.53	0.107	0	3	2
GP	-0.883	0.107	42	1	1
GEV	1.145	0.107	81	2	3

The fit is adequate if Z^{DIST} is close to 1.64; [B] T_4 = Regional average L-Kurtosis. Rank is evaluated on basis of closeness of Z^{DIST} to 1.64 in case of Z-test and regional parameters to theoretical curve in case of L-moment diagram

10.8 REMARKS AND CONCLUSION

Flood frequency estimation remains an important topic for design purposes, and flood data constitute the main source of information for this analysis. As such, the present study aimed at site a flood frequency analysis was carried out in the present study to determine the characteristics of high flow events with flood magnitudes and their frequencies. Single stations or at-site analyses or were carried out for 36 catchments located in Mahandi basin. *Maximum daily discharges* (m^3/s) or Annual Maximum Series (AMS) for 25 years for 23 small catchments and about 35 years for 14 gauging sites were analyzed using seven distributions viz. EV1, GEV,

PT3, LP3, GP, GL and WAK. Three parameter estimations methods were applied for this case and they are method of moments, method of maximum likelihood and method of probability weighted moment (PWM). A regional flood frequency analysis was also carried out for the same study area and is given in next chapter. In general, the results of the present study are consistent with those from early studies, although it can be seen that updating the flood information resulted, for many stations, in an improvement of flood estimates.

Characteristics of floods differ from one drainage basin to another and results of single station analysis are only applicable to the specific gauged streams or those streams near hydrometric stations. For EV1, GEV, PT3, LP3, GP,GL and WAK , respectively, results of the 36 single station high flow frequency analyses are provided in Appendix 9 (b) and (c) for the estimated parameters and the flood for recurrence intervals of 20, 50, 100, 200 , 300 and 500 years. For the *Kesinga* site, which has a comparatively large drainage area, the 500-year flood was estimated to be around 29866.9 m^3/s using WAK. This corresponds to the highest estimated 500-year flood in Kantamal with LP3 distribution. Conversely, using the GEV distribution, the 2-year flood was estimated at 20895.70 m^3/s and 32609.83 m^3/s using EV1; being quite similar to those obtained using LP3. Notably greater differences among distributions were noted at higher recurrence intervals. For all 36 stations, the maximum likelihood parameter estimates, as well as their corresponding MOM and PWM statistics, are presented in Appendix 9(b).

From the single station results (Appendix 9 b and c) it is clear that, for the majority of the single analysis, both the LP3 and GEV fitted the data almost exactly the same (especially in the central portion of the plot). For the upper right portion of this plot (at high recurrence intervals), it is evident that GEV adjusts better to the observational data than LP3 and GP. Results of the R.M.S. error statistics favor GEV over LP3 approximately 64% of the time. For the 36 GD sites,

the corresponding return period flood are estimated separately and the goodness-of-fit assessments suggested the GEV model to be the overall most appropriate distribution function. These findings were also strengthened by extreme value theory, which suggests that the annual maxima (such as flood data) be modeled as realizations of random variables distributed according to a member of the *generalized extreme value* (GEV) family of distributions. Based on such considerations, the GEV was therefore be accepted for the estimation of at-site floods as a function of annual maximum runoff for various recurrence intervals (i.e., 20, 50, 100, 200, 300 and 500). In all cases, the fitted model was consistent with the calculated T-year flood events, such that they could be applied to predict floods for ungauged basins (within their range of application).

CHAPTER 11
FUNDAMENTALS OF REMOTE SENSING AND ITS APPLICATION IN HYDROLOGY

11.0 GENERAL

Remote sensing can be broadly defined as the collection and interpretation of information about an object, area, or even without being in physical contact with the object. Aircraft and satellites are the common platforms for remote sensing of the earth and its natural resources. Aerial photography in the visible portion of the electromagnetic wavelength was the original form of remote sensing but technological developments has enabled the acquisition of information at other wavelengths including near infrared, thermal infrared and microwave. Collection of information over a large numbers of wavelength bands is referred to as multispectral or hyperspectral data. The development and deployment of manned and unmanned satellites has enhanced the collection of remotely sensed data and offers an inexpensive way to obtain information over large areas. The capacity of remote sensing to identify and monitor land surfaces and environmental conditions has expanded greatly over the last few years and remotely sensed data will be an essential tool in natural resource management.

India has made tremendous progress in the aerospace technology in the last two decades. We are one, among the world's best in satellite and the remote sensing technology. We are also one of the developing countries facing the natural disasters and trying level best to mitigate their ill effects. At the best we can minimize their effect. To achieve this, we need to have a comprehensive management program dealing with mapping and monitoring. The remote sensing technology with the fieldwork provides an edge to characterize the geomorphologic aspects. The present study reviews the application of the remote sensing tools in the application of geomorphology. A few case studies from the different part of the country are presented in the

context of geomorphology of the foothills of Siwaliks, Central India and the eastern coast and use of remote sensing.

11.1 APPLICATIONS

Remote sensing as described earlier is sensing from a distance. The electromagnetic energy after interacting with matter can be recorded by a sensor and the recorded data can be seen an image or photo depending upon the sensor. The electromagnetic spectrum is a continuum of energy that ranges from meter to nanometers in wavelength and travels at the speed of light.

The spectrum has several windows through which reflected or emitted radiation is recorded in sensors (Sabins 1997). Among the sensors are aerial cameras, which employ a film for recording in visible (B&W and color) and near infrared band. Data from visible, near infrared and other optical windows are recorded by multispectral scanners. Thermal scanner records thermal (emitted) radiation. Active microwave sensors (SLAR or SAR) record data from microwave region of spectrum and have the advantage of recording day and night as well as through clouds. Thus technologically it is possible to record terrain information by various sensors

11.2 STREAM MORPHOLOGY

Hydrological and land forming processes impacts stream morphology in a basin. Thus, stream morphological parameters have been used in hydrological studies. Conventionally, the parameters were derived from measurements from hardcopy topographic maps. However with developments in GIS technology, automatic and computer based basin characterization has become possible. The technique requires digital elevation model as input. Many satellite remote

sensing based DEM are now available. With availability of both GIS technology and input data, study of stream morphology has become easy.

11.2.1 Stream Order

Basin area, channel dimension and the discharge are related to the order of the stream. Stream order is used in classifying the streams. Streams of larger orders have larger basin, channel size and the discharge. Stream order can be assigned as per Horton, Strahler, Shreve or Scheidegger methods of stream ordering (Figure 13a). As per Horton method of stream ordering, all fingertip streams having no branched are classified as first order streams. Steams having braches of only first order are classified as second order streams. Likewise, all higher order streams are classified. A parent stream is assigned same order up to the headwaters. When a stream joins at larger angle or when angles are equal and the stream is shorter, the steam is termed as tributary stream. Strahler modified the Horton method of stream ordering. This method is widely used and is less subjective. All fingertip streams are classified as first order streams. These streams normally carry wet weather discharge and are normally dry. Second order streams form when two first order stream confluences. Likewise other order streams are classified. Thus, second and higher order streams do not extend up to the headwaters. This ordering system has disadvantage is trellis pattern where large number of stream join a higher order stream without change of order of that stream. Thus though discharge will increase in that stream, its order remains same. Shreve ordering scheme improves the Strahler scheme. In this scheme, order of a stream downstream of a junction is a sum of the order of the streams upstream of the junction. Scheidegger further modified the Shreve scheme. In Shreve scheme, the order of stream increases rapidly. However, Scheidegger scheme allows for slower increase in the order of higher order streams. Initially streams are ordered as per Shreve scheme except

the order of all fingertip stream is two to start with. After ordering is complete, a logarithm to base two is taken of the stream orders, resulting in Scheidegger stream order.

11.2.2 Laws of Basin Geometry

Stream order has power relationship with several of the basin geometric measurements. The relationships form the laws of basin geometry. As per these laws, the stream number, average length and basin area are related to the stream order. As the stream order increase number of streams, average area and slope decreases, whereas average stream length increases.

$$N_u = R_b^{s-u}$$

$$\overline{L}_u = \overline{L}_1 R_l^{u-1}$$

$$\overline{A}_u = \overline{A}_1 R_a^{u-1}$$

Where,

N_u= Number of streams of order u

s=Highest order of the basin

L_u= Average stream length of order u

A_u= Average basin area of order u

R_b= Bifurcation ratio

R_l= Stream length ratio

R_a=Basin area ratio

11.2.3 Bifurcation Ratio

Bifurcation ratio is a ratio of number of stream in a given order to that in a higher order. Since number of streams decreases with increasing order, bifurcation ratio is always higher than unity. Its value ranges from 2 to 5. Elongated basins in general have higher bifurcation ratio.

11.2.4 Stream Length Ratio

It is a ratio of average length of streams of given order to that of a stream of next lower order.

Table 11.1: Measured Length, Length of Centroid and Slope of Small catchments at different locations using available toposheets/imageries

SL	Bridge No.	Catch Area (Sq. Km.)	Name of Stream	Location						Length (Km)	Lc (Km)	Slope (m/Km)
				Latitude			Longitude					
				Deg	Min	Sec	Deg	Min	Sec			
1	7	3108	Bhedon	21	47	06	84	02	00	96.60 (99.70)	51.84 (53.41)	061
2	121	1150	Kelo	21	57	45	83	18	18	80.50 (83.33)	38.64 (39.37)	4.99
3	12	666	Lilagar	22	02	00	82	20	00	67.02 (66.82)	25.75 (26.41)	1.09
4	195	615	Jira	21	16	00	83	38	18	53.94 (49.58)	28.16 (27.44)	1.59
5	235	312	Ranjhor	21	07	48	83	35	12	41.06 (40.91)	21.09 (22.21)	1.61
6	385	194	Sandul	20	04	36	83	21	00	39.36 (38.68)	15.13 (15.29)	4.39
7	69	173	Borai	22	00	36	82	58	42	35.42 (35.76)	18.50 (19.21)	2.31
8	59B	136	Karwar	22	01	00	82	53	42	28.18 (26.35)	11.26 (11.15)	5.87
9	698	113	Bisra	22	15	18	84	57	12	26.57 (25.10)	14.40 (14.91)	9.04
10	154	58	Aherajhor	21	51	38	83	41	35	11.48 (11.45)	09.65 (07.71)	5.18

11.2.5 Basin Area Ratio

It is a ratio of average area of basins of given order to that of the basins of next higher order.

11.2.6 MISCELLANEOUS MORPHOMETRIC MEASURES

11.2.6.1 Drainage Density

Drainage density is a ratio of total stream length to the basin area. It has unit of inverse of length. Small drainage density signifies higher infiltration in a basin. It also reflects climate

pattern, geology, soil and vegetation cover. Areas devoid of vegetation and with semi arid climate characterized by intense thunder storm possess high value of the drainage density.

11.2.6.2 Constant of Channel Maintenance

It is a ratio of basin area to total stream length. It has unit of length.

11.2.6.3 Relief Ratio

Relief ratio is a ratio of difference of the elevation at the basin outlet and highest point in the basin and basin length. The basin length is measured along the main stream. Sediment yield increases with increase in the relief ratio.

11.2.6.4 Mean Stream Slope

Mean stream slope is ratio of difference in elevations at the outlet and source and the length of the stream.

11.2.6.5 Mean Catchment Slope

Mean catchment slope is a ratio of difference in elevation at 85% and 10% of the maximum length of the basin measured from the basin outlet and the basin length (75% of the total length of the basin) for which the elevations are taken in account for obtaining the elevation difference. Average slope of the basin is also obtained by averaging the slopes of the grid points in the basin.

11.2.6.6 Longitudinal Profile

Longitudinal profile is a plot of elevation v/s distance. The distance is measured from a downstream point. In general, the shape of profile is concave. Local topography, bed rock

features and bed material also influence the profile. For erodible bed material, the bed slope is flatter.

11.2.6.7 Hypsometric Curve

It is a plot of cumulative area above a given elevation v/s the elevation. Elevation is plotted on y- axis and area is plotted on abscissa. Both elevation and area may also be expressed in percentages. To obtain the hypsometric curve, area enclosed by successive contour lines is measured. The cumulative area is obtained from these measured values to obtain the hypsometric curve. The curve is useful in describing the hydrological variables, e.g. rainfall and snow cover, that are altitude dependent.

11.2.6.8 Basin Shape

Several measures are used to describe the basin shape. These are form factor, elongation ratio, circularity ratio. Except for the compactness coefficient, which has values greater than unity, all other shape factors have values less than unity.

Form factor: Form factor is a ratio of basin area and square of the basin length. Basin length is measured parallel to the main stream in the basin. It is a dimensionless ratio. Its value is in general less than unity.

$$F_f = \frac{A}{L^2}$$

Where F_f= Form factor

A= Basin area

L= Basin length measured parallel to the main stream

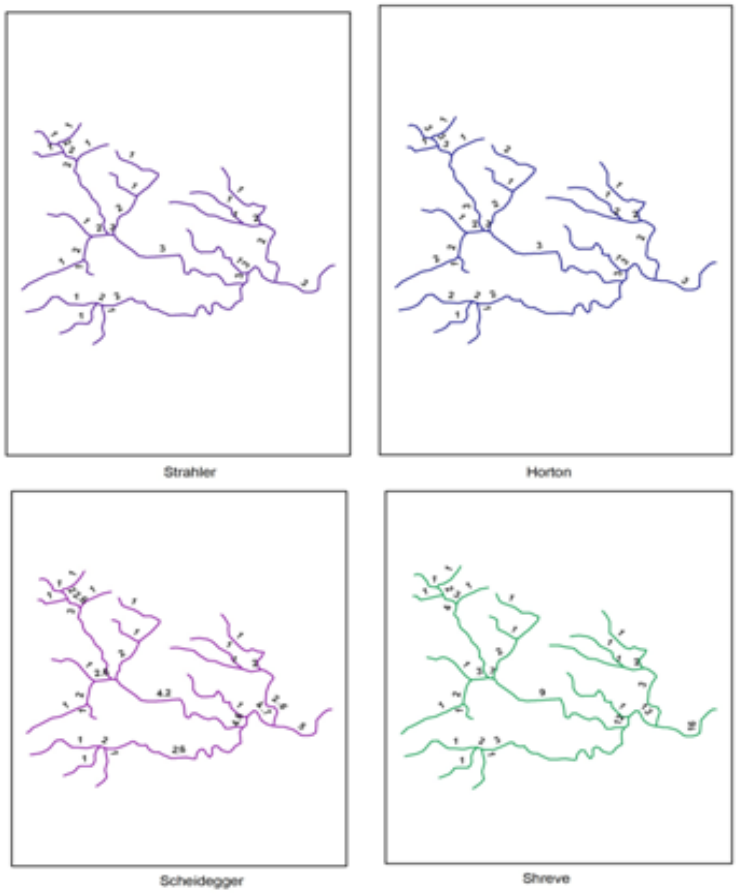

Figure 11.1 Methods of Stream Ordering.

11.2.6.9 Elongation Ratio

It is a ratio of diameter of circle having same area as the area of the basin to the length of the basin. It is a dimensionless ratio. Its value is in general less than unity.

$$E_r = \left(\frac{A}{0.786}\right)^{0.5} \frac{1}{L}$$

Where E_r = Elongation ratio

A= Basin area

L= Basin length measured parallel to the main stream

Compactness coefficient: The coefficient is a ratio of perimeter of the basin to the circumference of a circle with area equal to the area of the basin.

$$C_C = 0.2821\left(\frac{P}{A^{0.5}}\right)$$

Where C_c= Compactness ratio

P= Perimeter of the basin

A=Area of the basin

Circularity ratio: The ratio is obtained by dividing the area of basin to the area of the circle having same perimeter as the perimeter of the basin.

$$C_r = 12.57\left(\frac{A}{P^2}\right)$$

Where C_r=Circularity ratio

A= Basin area

P= Basin perimeter

11.3 DEM PROCESSING

DEM are excellent source of data for automatically extraction of basin morphology and characteristics. DEM processing tools are available in modern GIS software e.g. ILWIS, ArcGIS etc. Hydrological model interfaces also provide these capabilities. In particular many interfaces are built on Arc View and ArcGIS software for hydrological models. These interfaces exist for many models e.g. HEC HMS, HEC RAS, Mike-11, MikeBasin, SWAT etc. HEC software has interfaces based on Arc View 3.1, an older version of ArcGIS. Other software has ArcGIS based interfaces. These interfaces are available from organizations/ firms who are the developer of the model.

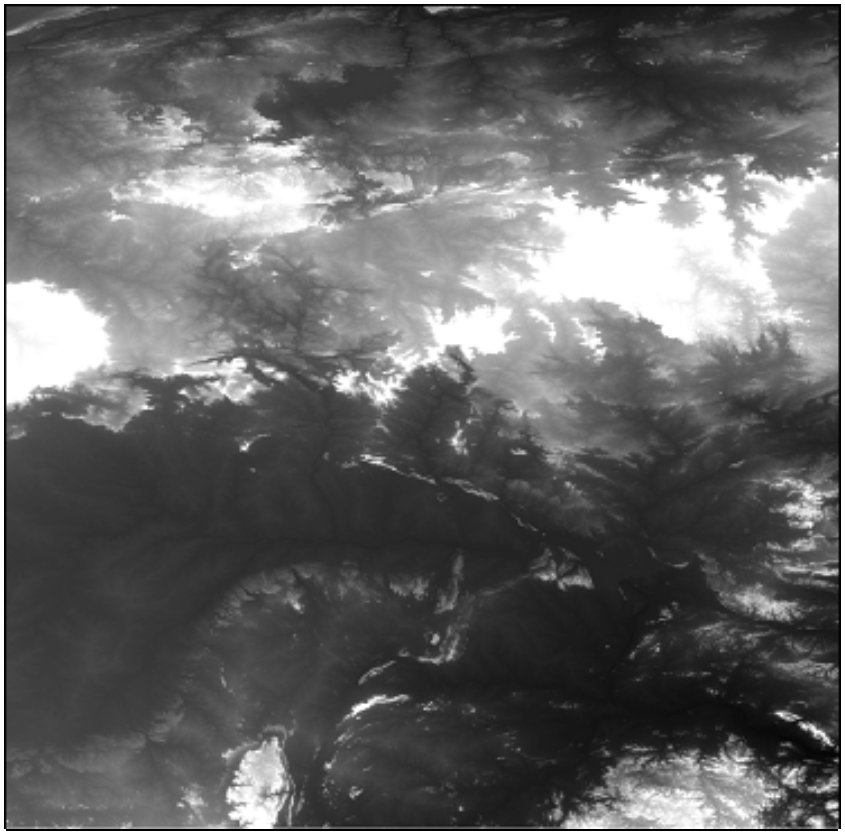

Figure 11.2 DEM Downloaded from Internet

11.4 HEC-HMS

In HEC-HMS first the DEM is preprocessed and thereafter model set up and basin characteristics are extracted. In terrain preprocessing, river network and sub basins are automatically delineated from digital elevation model. Various steps performed in the preprocessing are described below.

Depressionless DEM: Many of the depression or pits are artifacts results from DEM creation processes. Depression can also be natural. In the process, depressions are filled up to create a DEM. This Depressionless DEM is suitable for hydrological processing.

DEM reconditioning: Due to artifacts in DEM, the flow direction can arbitrarily change away from actual stream course. In DEM reconditioning, the actual river course can be burned in to the DEM to avoid this problem. In addition to burning in, in a buffer around the burn in vector, smoothing of the terrain can also be done. Further, sharp drop can be imposed from this smoothened burned in area. Alternately, terrain can be raised (fenced) in a process similar to this.

Flow direction: The method is based on eight- flow direction approach. In this approach, flow from a cell can be directed towards one of the eight flow directions, namely E, SE, S, SW, W, NW, N, NE from the central cell.

Flow accumulation: Flow accumulation depicts the count of upstream cells that drain into a cell. Stream network: Cells with flow accumulation greater than a threshold are assumed to be the cells falling on the stream network. The threshold indicates minimum flow accumulation area necessary for initiation of a channel. Default value is 1% of the largest basin area.

Stream segmentation: The process assigns unique value to each stream segment. The segments can be converted into a vector map.

Watershed delineation: Watershed corresponding to each stream segment is delineated. The watersheds can be converted into a vector map. Watershed aggregation: All the watersheds up to lowest junction point are aggregated. Additional features, namely finding catchment area at a stream location, flow tracing downstream a user defined point are also available. Former feature is particularly useful in locating a gauge point in the network from the information on catchment area up the gauge point. Rivers can be extended or merged. River elevation profile can also be plotted. Further, the basin and river network can be modified. Basins can be merged,

subdivided or split at the confluence. This will lead to a desired subbasins configuration for the model.

For all stream segments, their lengths and slopes are computed. Slope is computed as elevation difference at the end point of the stream segments divided by the length of the segment. The elevations at the end points are extracted from DEM automatically. For watersheds, their area and centroid locations are determined. The centroid are determined using one of the four methods, namely centre of the bounding box, centre of bounding ellipse, midpoint of longest flow path and user defined location. A centroid shape file is created in one of the first three methods. For assigning user defined centroid, the centroid shape file obtained from one of the other three methods is edited to move the centroid to a location of once choice. The elevation of the centroid is automatically placed extracted from DEM and placed in the centroid and watershed shape file attribute tables.

Length of the longest flow path and flow path up to point on the longest flow path nearest to the centroid of the basin are useful parameters in determining synthetic unit hydrographs. In the step for determining the longest flow path, shape file is created. The lengths, elevations of the end points, slope between points at 10% and 85% from the downstream location of the segment, slope between end points are populated automatically in the longest flow path and watershed shape file attribute tables. For obtaining centroidal flow path, the centroid location file is used. In the method the point of the longest flow path nearest to the centroid is determined. A shape file of the segments up to this point is created. The length of the segments is populated in both centroidal flow path and watershed shape file attribute tables. Basin slope can be computed by intersecting the slope raster and watershed shape file. In the operation, average is determined at all the cells of the slope raster falling in a watershed.

Additionally, curve number raster intersected with the watershed shape file to obtain average curve number for each watershed. Basin slope, curve number and basin length values are used to estimate the lag time using NRCS basin lag formula.

11.5 SWAT (SOIL AND WATER ASSESSMENT TOOL)

SWAT was developed by Jeff Arnold for USDA to predict the impact of management practices on water, sediment and chemical yield from large watershed. The model is physical based and thus suitable for studying the impact of weather, vegetation and management practices. The model may be used for both gauged and ungauged watersheds. The model is continuous time model and is not designed for detailed single event flood routing. ArcSWAT is a ArcGIS extension and provides a GUI for the SWAT model. The extension is evolved from an earlier version, namely AVSWAT which was developed for ArcView 3.1. The interface allows for automatic delineation of basins/ sub basins, river network, computation of catchment and river characteristics, delineation of hydrologic response units based on land use, soil and topography, input of weather data and running of the model etc. Reach database fields sub basin numbers (current and downstream), cumulative drainage area, reach length and slope (in percent), minimum and maximum elevations are filled through DEM processing. Sub basin properties namely area, slope, centroid location and elevation, minimum and maximum elevation, longest flow path length, reach length and slope are also computed.

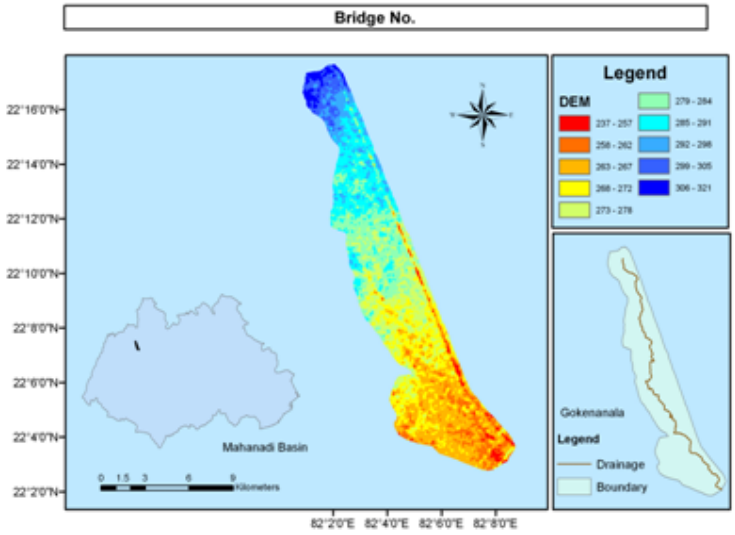

Figure 11.3 Small Catchment *Gokina nala*

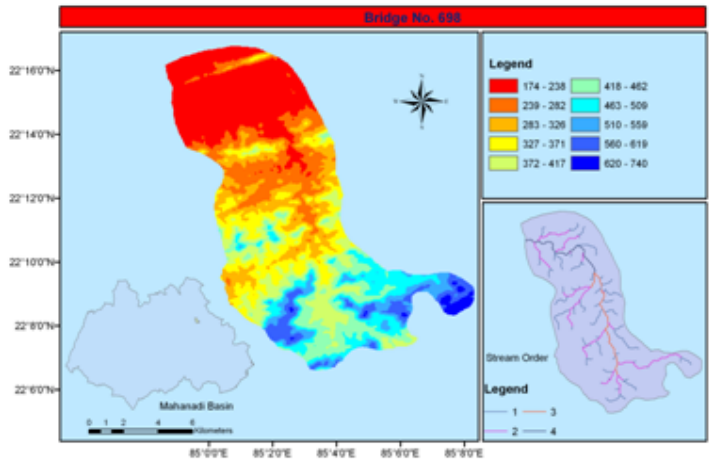

Figure 11.4 Small Catchment *Gokina nala*

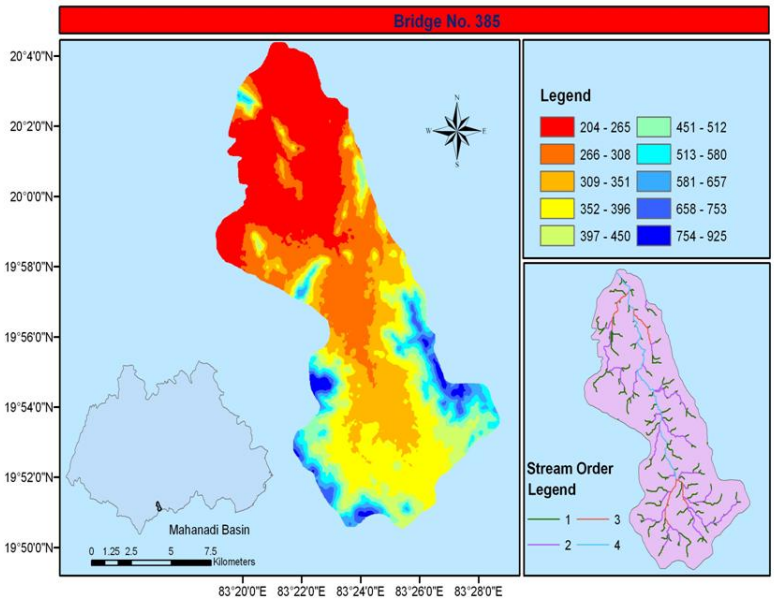

Figure 11.5 Small Catchment *Gokina nal*

Figure 11.6 Small Catchment *Lilagar*

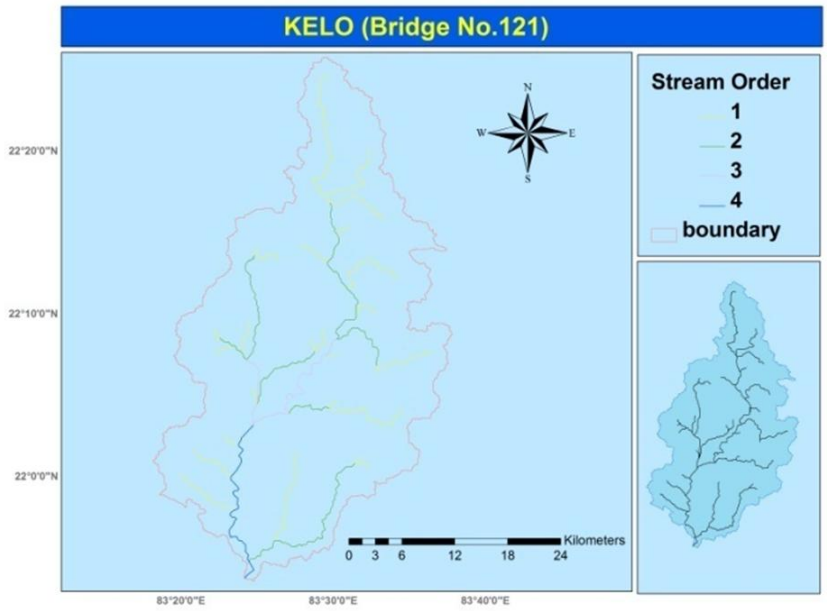

Figure 11.7 Small Catchment *Kelo*

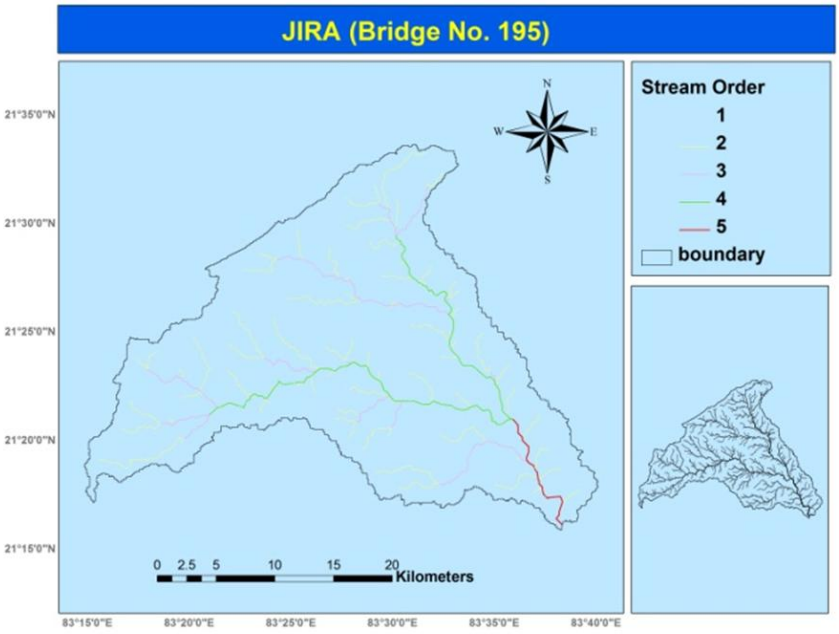

Figure 11.8 Small Catchment *Jira*

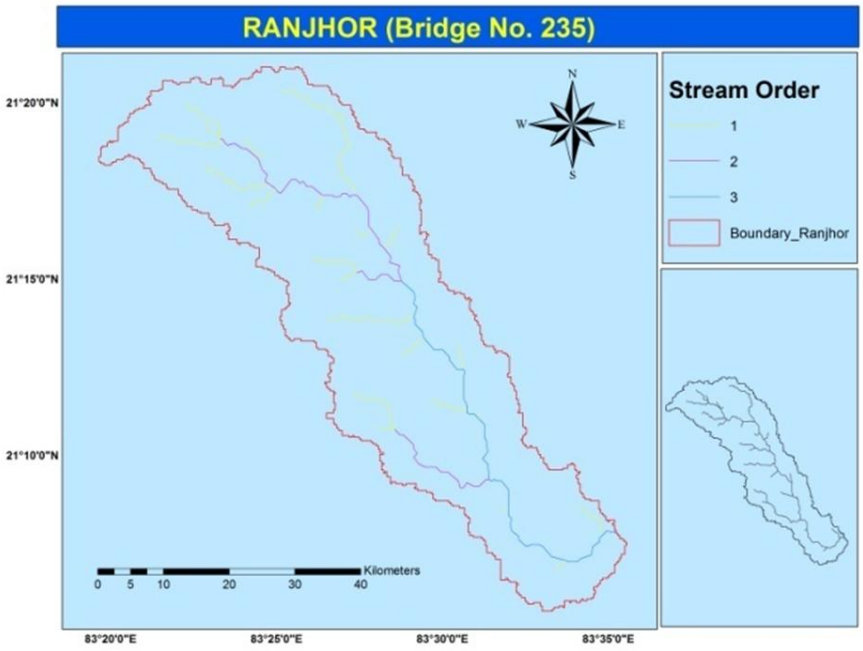

Figure 11.9 Small Catchment *Ranjhor*

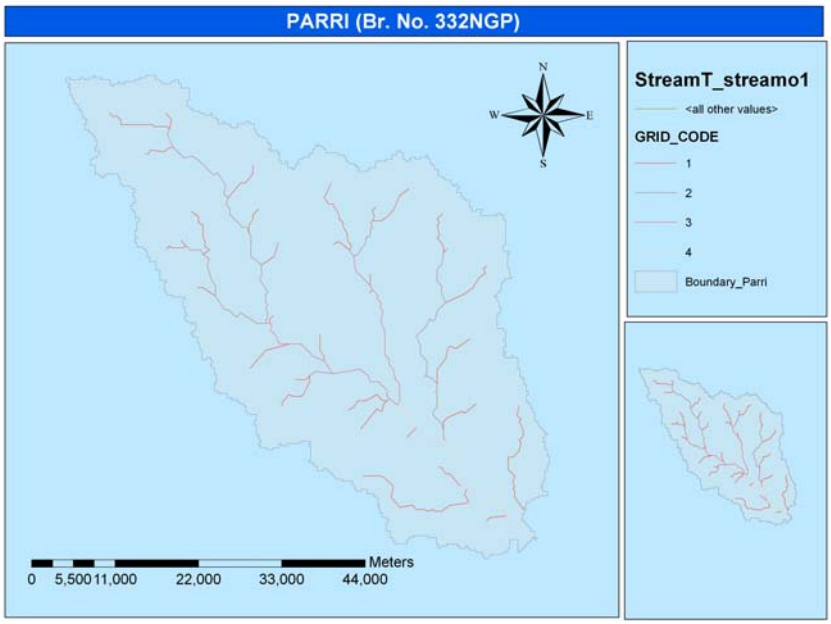

Figure 11.10 Small Catchment *Parri*

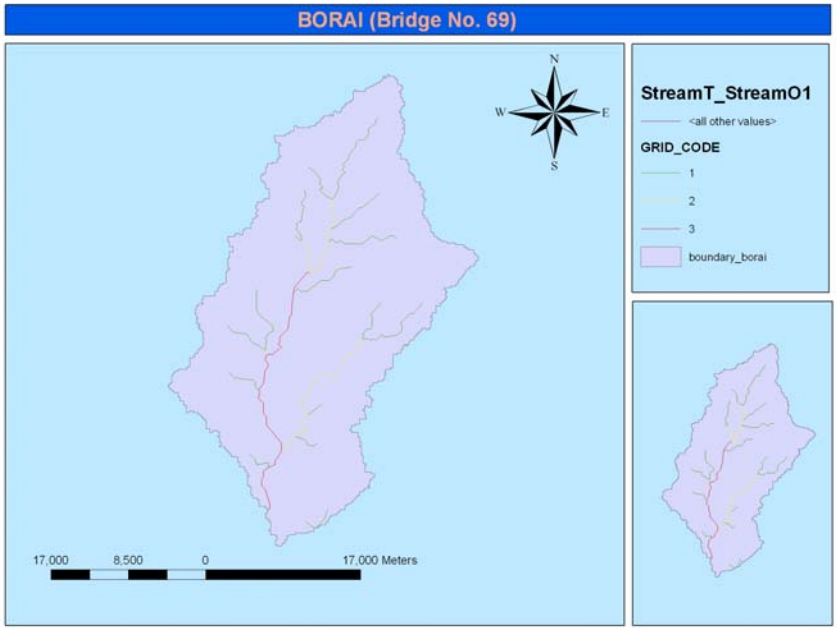

Figure 11.11 Small Catchment *Borai*

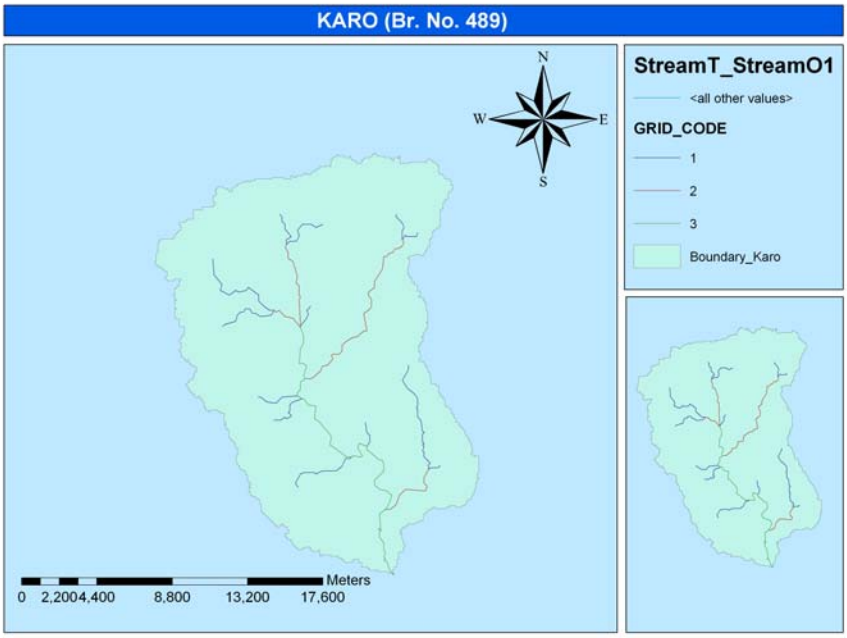

Figure l1.12 Small Catchment *Karo*

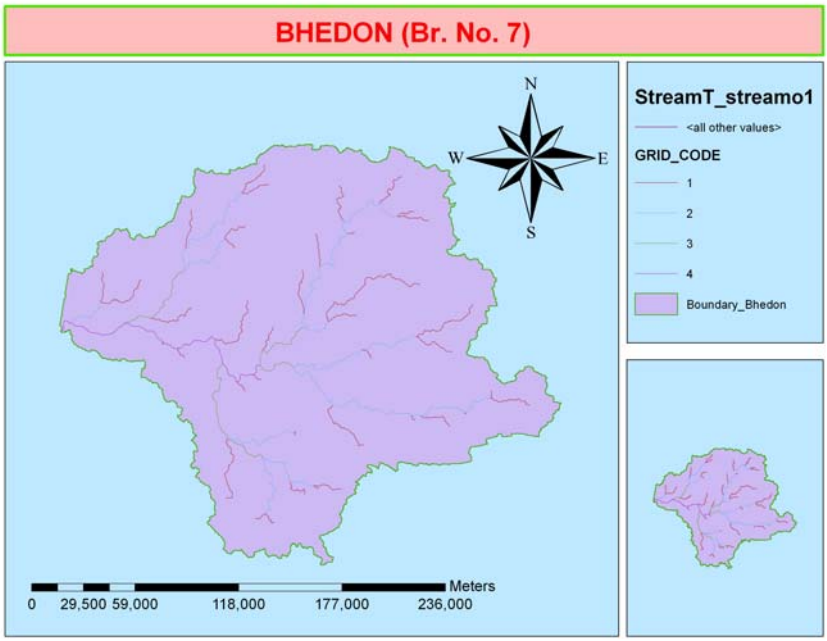

Figure 11.13 Small Catchment *Bhedon*

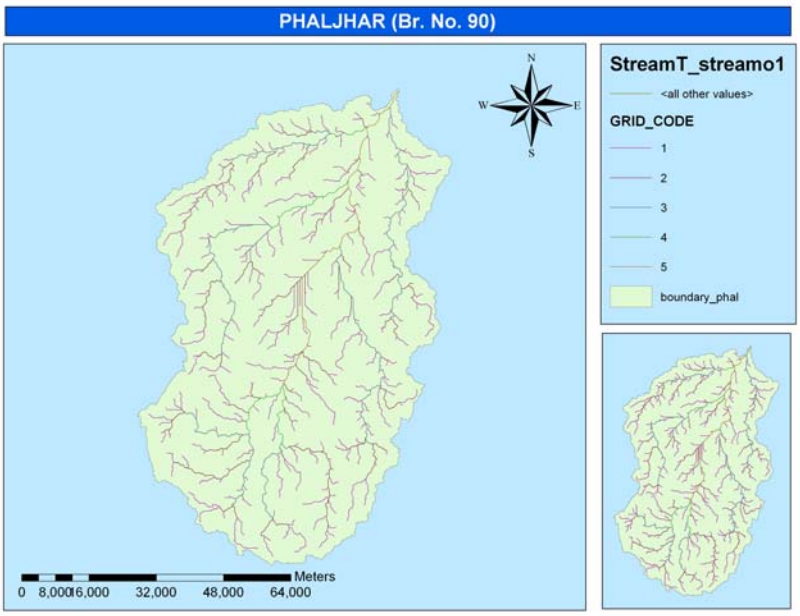

Figure 11.14 Small Catchment *Phaljhar*

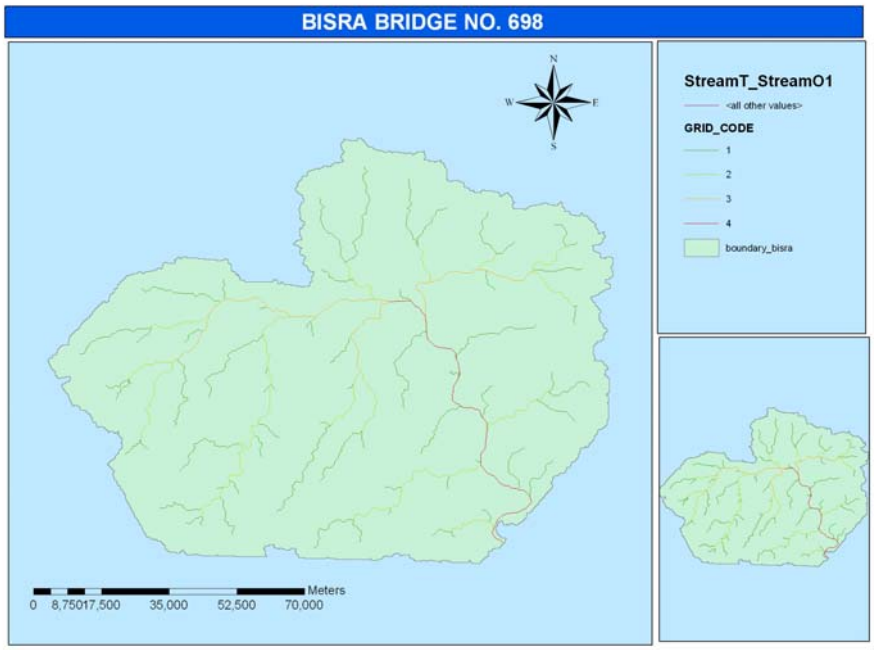

Figure 11.15 Small Catchment *Bisra*

11.6 REMOTE SENSING TECHNIQUE

Remote Sensing is the science of acquiring, processing and interpreting images that record the interaction between electromagnetic energy and matter. As per General Assembly Resolutions and International Treaties Pertaining to the Peaceful Uses of Outer Space of United Nations, the term Remote Sensing means the sensing of the Earth's surface from space by making use of the properties of electromagnetic waves emitted, reflected or diffracted by the sensed objects, for the purpose of improving natural resources management, land use and the protection of the environment. Remotely sensing data are important data source in hydrological modeling. Land use and its management practices affect runoff from basins. Land use and cover maps, up to level IV, may be prepared using remotely sensed data. Soil texture and types may also be interpreted. SRTM (Shuttle Radar Topographic Mapping Mission) Digital elevation models (DEM) are available at 90 m spatial resolution and 1 m least count for the world.

Rainfall data are available at 15' geographic and 3 hourly temporal resolution from TRMM (Tropical Rainfall Measuring Mission). Water and land temperature may be monitored using satellite data.

11.6.1 Land Use and Land Cover Classification

Information on existing land use and their changes over time is needed in water resources planning and management, amelioration of land degradation etc. Information is available from past land use maps, topographic maps etc. Due to time variant nature of this geographical data, satellite remote sensing is very useful.

Land use refers to, man's activities on land which are directly related to the land. Land cover, on the other hand, describe, the vegetation and artificial constructions covering the land surface. USGS (United States Geological Survey) has developed a land use and land cover classification system for use with remotely sensed data. The system has four hierarchical levels, namely, I to IV. Levels I and II are useful for national, statewide classification. Other two levels are useful for intrastate, regional, county and municipal level classification. Latter levels classes are developed by users. First two levels have standard and well defined classes. Level I and II classes can be delineated using medium resolution satellite data, e.g. IRS LISS- III, Landsat TM etc. Level III and IV classes can be delineated using fine resolution data, e.g. IRS LISS- IV, Panchromatic, aerial photographs etc.

Multi date satellite remotely sensed data in single season may be used in crop type discrimination. In particular, satellite data may be very useful in discrimination in paddy (seasonal) and sugarcane (perennial) crops. Staggered sowing and different duration in crops pose problem is accurate classification of crop types. Horticulture areas have different signature than crops.

11. 7 NRCS CURVE NUMBER TECHNIQUE

NRCS curve number technique was developed by Ogrosky and Mockus in 1957. The method is used to estimate direct runoff volume from a catchment using information on land use and cover, soil, cover treatment practices and hydrologic condition and antecedent rainfall condition and precipitation in a catchment. The technique operates at daily time step. However, event modeling can also be done by applying the technique to cumulative precipitation at different time.

Curve number is determined for the watershed for antecedent moisture condition II from known land use and cover, cover treatment practices and hydrologic condition and hydrologic soil group. For these basin characteristics, the curve number can be read from tabulated values for cultivated and other agricultural land and urban area. For an urban area, the impervious area is assumed to be directly connected to the drains and has CN of 98. The pervious area is assumed to be open land with good hydrologic condition. The curve number can be derived as area weighted sum of CN for impervious area directly connected to drain (98) and CN for open land with good hydrologic condition. Antecedent condition is obtained from cumulative precipitation for antecedent five days.

Good hydrologic condition indicates that the land favors greater infiltration and thus less runoff is generated from the area. Factors namely density and canopy of vegetative areas, amount of year-round cover, amount of grass or closed-seeded legumes in rotations, percent of residue cover on land surface (good > 20 percent), and degree of roughness are considered in determining the hydrologic condition. Crop residue cover applies only if residue is on at least 5 percent of the surface throughout the year. Treatment and hydrologic condition are often difficult to obtain from remotely sensed data. Thus, general condition and treatment are obtained

from field knowledge. For cultivated area, multi date satellite data are needed to obtain crop types. The crop types can often be distinguished based on differences in the sowing dates, difference in crop vigor etc. In general, from single date data, the crop type mapping is difficult.

Impervious area is seen in cyan color on standard FCC. Thus, impervious area can be mapped from satellite data. Sometimes fallow land also shows similar signature. These are may be located in the periphery of the urban area. Developing urban areas are difficult to delineate in the satellite images. It is difficult to distinguish impervious area directly connect to drain from that draining in to open spaces from satellite data. This information has to be obtained from ground truth data.

11.8 FLOOD PLAIN MANAGEMENT

Reliable data on river morphology, river meandering, extent of flooding and duration is required for proper planning of flood control projects. In the conventional methods of flood risk zoning, the flood discharge is routed through the river reach to estimate the likely inundation due to spilling over the banks/embankments based on topographic contour maps and configuration of the river geometry obtained through land surveys. Continuous availability of satellite-based remote sensing data has made the understanding of dynamics of flood events much easier. The satellite remote sensing technique provides a wide area synoptic coverage, repetitiveness and consistency, which enable the collection of information on all major flood events on a reliable basis. Remote sensing can provide information on flood-inundated areas for different magnitude of floods so that the extent of flooding in the flood plains can be related to the flood magnitude. Duration of flooding can be estimated with the help of multiple coverage satellite imagery of the same area within 2/3 days. High-resolution satellite data provides information on the floodplain and effectiveness of flood control works. Extent of inundation for

specific flood return periods can be estimated. Using close contour information, inundation extent for given elevation, can be estimated which is a vital input for risk zone mapping. GIS provides a broad range of tools for determining area affected by floods and for forecasting areas that are likely to be flooded due to high water level in a river. Spatial data stored in the digital database of the GIS, such as a digital elevation model (DEM), can be used to predict the future flood events. The GIS database may also contain agriculture, socio-economic, communication, population and infrastructure data. This can be used, in conjunction with the flooding data to adopt an evacuation strategy, rehabilitation planning and damage assessment in case of a critical flood situation.

11.9 MAPPING AND MONITORING OF WATERSHEDS

Proper planning of watersheds is essential for the conservation of water and land resources and their productivity. Characterization and analysis of watersheds are a pre-requisite for this. Watershed characterization involves measurement of parameters of geological, hydro-geological, geomorphological, hydrological, soil, land cover/land use etc. Remote sensing via aerial and space borne sensors can be gainfully used for watershed characterization and assessing management requirements and periodic monitoring. The various physiographic measurements that could be obtained from remotely sensed data are size, shape, topography, drainage pattern and landforms. The quantitative analysis of drainage networks enables relationships between different aspects of the drainage pattern of the basin to be formulated as general laws and to define certain useful properties/indices of drainage basin in numerical terms. The laws of stream numbers, stream length and stream slopes can be derived from measurements made in the drainage basin. Remote sensing along with ground based information in GIS mode can be used for broad and reconnaissance level interpretations for land capability classes, irrigation suitability classes, potential land users, responsive water harvesting areas and

monitoring the effects of watershed conservation measures. Correlations for runoff and sediment yields from different watersheds versus land use changes and land degradation could also be established.

CHAPTER 12
ESTIMATION OF CONFIDENCE INTERVALS AND UNCERTAINTY OF HYDROLOGICAL PREDICTIONS

12.0 INTRODUCTION

Hydrological processes are highly uncertain and subject to variability. The main reason for this uncertainty is that the intrinsic dynamics of many hydrological processes are still unknown leading to lack of perfect understanding of hydrologic phenomena and processes involved. The hydrological observations of rainfall, river flows and pollution loads in streams vary from one case to the next on account of factors that cannot be assessed to any degree of accuracy. In addition, the hydrologists are typically working under conditions of data scarcity, which limit the efficiency of an inductive approach for tackling the above problems. However, Uncertainty estimation in hydrological modeling is receiving increasing attention as uncertainty assessment is one of the main goals of the Prediction in Ungauged Basins (PUB) initiative promoted by the International Association of Hydrological Sciences.

The traditional way of dealing with uncertainty in hydrological science is through statistics and probability. The quantification of uncertainty and the assessment of its effects on design and implementation generally include concepts and methods of probability. When sufficient information is available to support statistical hypotheses with appropriate statistical tests, a probabilistic statistical approach is preferable as a way to efficiently summarize the information content of the data. Therefore the objective of this chapter is to (a) first define some of the important terms of statistical and probabilistic hydrology; (b) describe measures of location, dispersion, and symmetry of both grouped and ungrouped data; (c) describe standard errors of mean, standard deviation, and coefficient of skewness; and (d) describe performance evaluation measures including normal reduced variates, chi-square, student t-distribution, and

F-distribution, (e) estimation of confidence intervals, and (f) uncertainty in a hydrologic predictions. These are described and explained in the subsequent sections.

12.1 BASICS OF STATISTICAL ANALYSES

Statistical analyses generally dealt in hydrology can be categorized as univariate and multivariate analyses. Problems dealing with the data of a single random variable fall under the category of univariate analyses. Such studies, for example, include the studies of rainfall analysis. The univariate analysis helps identify the statistical population, from which the sample measurements are derived, and predict future occurrences of the random variable based on the assumed population. In analyses of multivariable data with more than one variable, the functional form of the inherent relationship is derived. The linear regression analysis is generally used to develop a suitable linear form of multiple variable models. Here, a dependent variable takes the values caused by variations in one or more independent or predictor variables. Such models are found to be of great use in hydrological prediction.

12.2 SAMPLE STATISTICS

In any analysis of statistical data in general and hydrologic data in particular, it is important to determine their basic properties. The sample mean and variance are the most important statistical characteristics of a data set. The sample statistics provides in general the basic information about the variability of a given data set about its mean. The commonly used sample characteristics include (i) the central tendency or value around which all other values are clustered, (ii) the spread of sample values about mean, (iii) the asymmetry or skewness of the frequency distribution, and (iv) the flatness of the frequency distribution.

12.2.1 Measures of central tendency or location

Some of the important statistical measures of location include the following:

(i) *Mid range*

It is the average of the minimum and maximum values of the sample (or population):

Mid range = (minimum value + maximum value)/2. (12.1)

(ii) *Symmetry*

If data are displaced about the mean symmetrically, the measure of symmetry is equal to zero. On the other hand, if these are displaced more to the right of the mean than those on the left, then, by convention, the asymmetry is positive, and it is negative otherwise. The commonly used measures of asymmetry are:

Inter-quartile measure of asymmetry (I_{as}): It is defined as:

$$I_{as} = |I_3 - I_2| - |I_1 - I_2| \quad (12.2)$$

where, I_1, I_2, and I_3 are the lower, median, and upper quartiles, respectively.

Third central moment (M_3): The third moment of sample data about the mean is given by

$$M_3 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^3 \quad (12.3)$$

If data are symmetrical, $M_3 = 0$; it is positive or negative otherwise.

(iii) *Skewness coefficient (C_s)*

It is a non-dimensional measure of asymmetry of the frequency distribution fitted to the data. Its unbiased estimate is given by

$$C_s = \frac{N\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)(N-2)S^3} \quad (12.4)$$

Skewness coefficient is an important indicator of symmetry of the distribution fitted to data points. Symmetrical frequency distributions exhibit a very low value of sample skewness coefficient C_s whereas asymmetrical frequency distributions show a positive or negative value. Smaller value of C_s implies more closeness of the distribution to the normal distribution since C_s = 0 for normal distribution. The third central moment having the dimension of the cube of the variable is not of direct use, as described above. On the other hand, C_s does not suffer from this limitation and is, therefore, generally preferred.

(iv) *Peakedness or flatness*

The peakedness or the flatness of the frequency distribution near its center is measured by

kurtosis coefficient (C_k), which is expressed as

$$C_k = \frac{N^2 \sum_{i=1}^{N} (x_i - \bar{x})^4}{(N-1)(N-2)(N-3) S^4} \qquad (12.5)$$

The excess coefficient (E) defined as: $E = C_k - 3$, is a variant of C_k. Positive E-values indicate a frequency distribution to be more peaked around its center than the normal distribution. Such a frequency distribution is known as leptokurtic. On the other hand, negative values of E indicate that a given frequency distribution is more flat around its center than the normal. These are called *platykurtic distributions*. Normal distribution is known as *masochistic*. The use of C_k is, however, not much popular in statistical hydrology.

(v) *Maxima or minima*

Maxima or minima refer to data extremes or extreme values of data, whether high or low. These data are commonly used in frequency analyses, for they are taken to be independent

of each other or random data. It is also for the reason that data observed for an event sometimes exhibit dependence. Examples of sample data extremes include annual maximum peak flood series, annual low flow series for a specific duration, and so on.

12.3 STANDARD ERRORS OF SAMPLE STATISTICS

Statistical characteristics, such as mean and standard deviation, derived from a short period of sample record represent only an estimate of the true or population statistic. The reliability of such estimates can be evaluated from their standard errors (SE). It is of common experience that there exists a probability of about 68% that the true or population statistic lies within one standard error of the value estimated from the available (or sample) data. The standard errors of commonly used mean, SE ($\overline{x}$); standard deviation, SE(S); and coefficient of skewness, SE (C_k), are derived as below,

$$SE(\overline{x}) = S/\sqrt{N}$$

$$SE(S) = S/\sqrt{(2N)}$$

$$SE(C_s) = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}} \qquad (12.6\ a, b, c)$$

12.4 PROBABILITY FUNCTIONS FOR HYPOTHESIS TESTING

Probability functions frequently used in hydrologic analyses include the normal, student t, chi-square, and F distributions. The last three are primarily used for testing statistically a hypothesis of concern. The normal distribution is however used for both prediction and testing

of the variable. The testing is based on the derivation of critical values derivable from generally available tables.

12.4.1 Normal Distribution

For normal distribution, the probability density function (PDF) and cumulative density function (CDF) has been discussed earlier. It is noted that μ and σ represent the population parameters of the distribution. The best estimates of μ and σ for sample data are taken as the sample mean ($\overline{X}$) and standard deviation (S), respectively. Thus, for computing the probability based on sample statistics, $\overline{X}$ and S are substituted for μ and σ, respectively. It is noted that it is possible to compute probabilities by integrating PDF over a range of x-values. However, since there exist an infinite number of values of $\overline{X}$ and S for the infinite number of samples in the population, it is rather impossible to integrate numerically. It can however be circumvented by transformation of the random variable x, as follows. If the random variable x has a normal distribution, a new random variable z can be defined as:

$$z = \frac{x - \mu}{\sigma} \qquad (12.7)$$

Where z is the standardized variates or normal reduced variates with mean and standard deviation of 0 and 1, respectively. This transformation enables PDF to express as a function of z, which is called the standard normal distribution. Probabilities can be computed using this PDF, and tabulated as a function of z as shown in Appendix I. In this table, z-values down the left margin are incremented by 0.1 and by 0.01 across the top. The corresponding cumulative probabilities are also given from -4 to the desired value of z. For example, the probability of z being less than 0.23 is 0.5910. The probability that z lies between -0.47 and 0.23 equals 0.2718

(= 0.5910-0.3192). The figure given at the top of this table makes this relationship self explanatory. Inversely, this table can also be used to find the value of z that corresponds to a certain probability. For example, for a probability of 0.05 on the left tail, the corresponding value of z is equal to -1.645. Since the distribution is symmetric about the ordinate, the value of z equal to 1.645 corresponds to 5% area under the curve f (z), but on the right tail.

12.4.2 Student t-Distribution

The student t or t_1 distribution is also a symmetric distribution similar to the normal distribution. It however has only one parameter υ, which is known as degrees of freedom. This parameter controls the spread of the t-distribution. Thus, the structure of the table of t-values slightly differs from the normal table. In the table in book (refer-i), the value of υ is given at the left margin, and the probability at the right tail of the distribution is given across the top. For example, for υ = 7 and a probability of 0.05, the t-value is 1.895. Since the t-distribution is symmetric, 5% of the area at the left tail is to the left of a t-value of -1.895 for υ = 7.

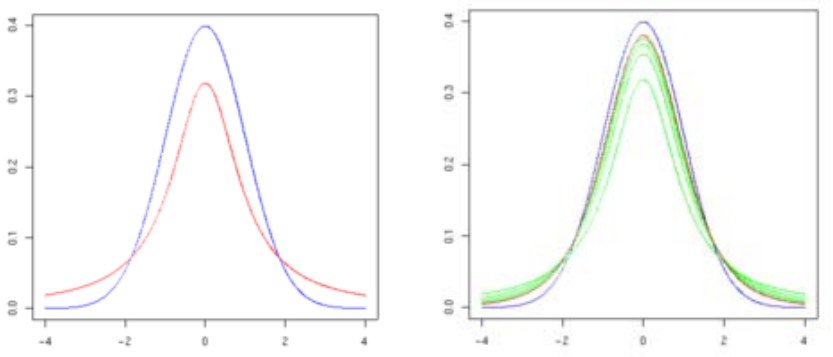

Figure 12.1 Density function of normal distribution (blue) and t-distribution (pink) for 1 degree of freedom (left) and 5 degrees of freedom (right).

As a corollary, for 5% of the area (2.5% at each tail), critical t-values are -2.365 and 2.365 for υ = 7. It is noted that for υ >30, the normal distribution is generally used in place of t-distribution.

12.4.3 Chi-Square Distribution

The Chi-square (χ^2) distribution is similar to the t-distribution in that it is a function of a single parameter n, the degrees of freedom. It however differs from the t-distribution in that the distribution is not symmetric. The table of chi-square values is identical in structure to the table for t-distribution, with n down the left margin, the probability across the top, and χ^2-values in the table. As an example, χ^2 varies from 0 to 4.575 if degrees of freedom = 11 and area at the left tail = 5%. For 17 degrees of freedom, there is a probability of 0.025 that χ^2 will lay between 30.191 and 4.

12.4.4 F-Distribution

The F-distribution is a function of two parameters, m and n, and these are tabulated as standard text books. The value of F corresponding to the appropriate probability in the right tail of the distribution can be derived from this table. For example, if m = 10 and n = 20, there is a 5% chance that F will be greater than 2.35. It is noted that the table is not symmetric. For m = 20 and n = 10, the critical value for 5% area in the right tail is 2.77. Thus the values of m and n must not be switched.

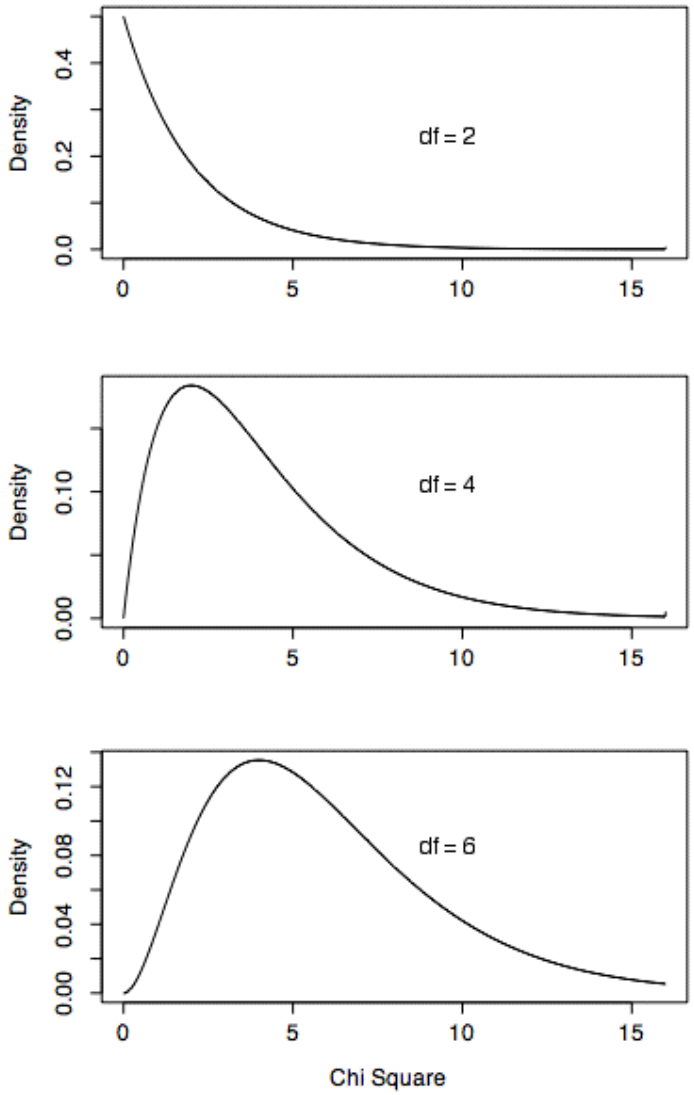

Figure 12.2 Chi Square Distributions with 2, 4, and 6 degrees of freedom.

12.5 ESTIMATION AND CONFIDENCE INTERVALS

12.5.1 Point and Interval Estimates

If one recall of earlier chapters, how the critical value(s)/outliers delineated our region of rejection. For a two-tailed test the distance to these critical values is also called the *margin of error* and the region between critical values is called the *confidence interval*. Such a confidence interval is commonly formed when we want to estimate a population parameter, rather than test a hypothesis. This process of estimating a population parameter from a sample statistic (or

observed statistic) is called statistical estimation. We can either form a point estimate or an interval estimate, where the interval estimate contains a range of reasonable or tenable values with the point estimate our "best guess." When a null hypothesis is rejected, this procedure can give us more information about the variable under investigation. It can also test many hypotheses simultaneously. Although common in science, this use of statistics may be underutilized in the behavioral sciences.

12.5.2 Confidence Interval or Margin of Error

The value $\sigma_{\bar{x}} = \sigma / \sqrt{n}$ is often termed the *standard error of the mean*. It is used extensively to calculate the margin of error which in turn is used to calculate confidence intervals. If we sample enough times, we will obtain a very reasonable estimate of both the population mean and population standard deviation. This is true whether or not the population is normally distributed. However, normally distributed populations are very common. Populations which are not normal are often "heap-shaped" or "mound-shaped". Some skewness might be involved (mean left or right of median due to a "tail") or those dreaded outliers may be present. It is good practice to check these concerns before trying to infer anything about your population from your sample.

Since 95.0% of a normally distributed population is within 1.96 (95% is within about 2) standard deviations of the mean, we can often calculate an interval around the statistic of interest which for 95% of all possible samples would contain the population parameter of interest. We will assume for the sake of this discussion that this statistic/parameter is the mean. We can say we are 95% confident that the unknown population parameter lies within our given range. This is to say, the method we use will generate an interval containing the parameter of interest 95% of the time. For life-and-death situations, 99% or higher confidence intervals may quite

appropriately be chosen. The population parameter either is or is not within the confidence interval so we must be careful to say we have 95% confidence that it is within, not that there is a 95% probability that it is. Since we expect it to 95% of the time, this can be a point of confusion.

12.5.3 Confidence Interval of Mean

Standard Deviation is known

We are not only interested in finding the point estimate for the mean, but also determining how accurate the point estimate is. The Central Limit Theorem plays a key role here. We assume that the sample standard deviation is close to the population standard deviation (which will almost always be true for large samples). Then the Central Limit Theorem tells us that the standard deviation of the sampling distribution is

(12.8)

We will be interested in finding an interval around x such that there is a large probability that the actual mean falls inside of this interval. This interval is called a *confidence interval* and the large probability is called the *confidence level*.

Example 12.1: Assume the population of annual maximum floods in Mahanadi river at a specific gauging site has a mean of 100 TCM (thousand cubic meter) per hour i.e. 100 x 1000 m^3/hr.; and standard deviation of 15 TCM per hour. Assume further that we draw a sample of n = 5 with the following values: 100, 100, 100, 100, and 150. The sample mean is then 110 and the sample standard deviation is easily calculated as

Standard deviation $(S) = [\sum_{i=1}^{N} (x_i - \bar{x})^2 / (N-1)]^{1/2}$

$= [\sum_{i=1}^{5} (10^2 + 10^2 + 10^2 + 10^2 + 40^2) / (5-1)]^{1/2}$ = 22.4 TCM per hour

The standard error of the mean = 22.4 / $(5)^{1/2}$ = 10 TCM per hour

Since 95.0% of a normally distributed population is within 1.96 (95% is within about 2) standard deviations of the mean. Our 95% confidence intervals are then formed with z = +/-1.96. Thus based on this sample we can be 95% confident that the populations mean lies between. Using Eq. (10.6), $z = \frac{x - \mu}{\sigma_x}$ or $\mu = \bar{x} - z\sigma_x$ (refer figure 10.1)

Or, 95% confident band of the populations mean = 110 +/- (1.96 x 10)

= (110 - 1.96 x 10) and (110 + 1.96 x 10) i.e. band is (90.4, 129.6).

Suppose, however, that you did not know the population standard deviation. Then since this is also a small sample you would use the *t*-statistics. The *t*-value of 2.776 would give you a margin of error of 27.8 and a corresponding confidence interval of (82.2, 137.8).

Example 12.2: Suppose that we check for flood depth near a gauging site in *Brahmani* River at 50 locations at the upstream of the gage during heavy monsoon. The measures were taken within an hour. And the average depth is 14 feet. Suppose that we know that the standard deviation for the entire depth measurements is 2 feet. What can we conclude about the average flooding depth with a 95% confidence level?

Solution: We can use x to provide a point estimate for μ. How accurate is x as a point estimate? We construct a *95% confidence interval for* μ as follows. We draw the picture and realize that we need to use the table to find the z-score associated to the probability of .025 (there is .025 to the left and .025 to the right).

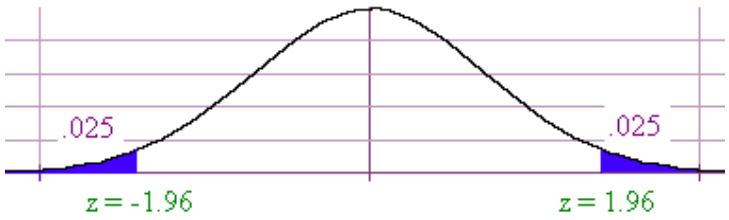

Figure 12.3 Normal distribution

We arrive at z = -1.96. Now we solve for *x*:

Using Eq. (10.6), $z = \frac{x - \mu}{\sigma_x}$ or where $\sigma_x \approx s/\sqrt{N}$, Thus,

$-1.96 = (x - 14)/ [2/ 50^{1/2}] = (x - 14) / 0.28$

Hence, x – 14 = -0.55

We say that $\pm$0.55 is the *margin of error*. We have that a 95% confidence interval for the mean clarity is (13.45, 14.55).

In other words there is a 95% chance that the mean depth is between 13.45 and 14.55. In general if z_c is the z value associated with c% then a c% confidence interval for the mean is

$$x \pm \frac{z_c \sigma}{\sqrt{n}} \qquad (12.9)$$

Standard Deviation is Unknown

When the population is normal or if the sample size is large, then the sampling distribution will also be normal, but the use of s to replace □ is not that accurate. The smaller the sample size the worse the approximation will be. Hence we can expect that some adjustment will be made based on the sample size. The adjustment we make is that we do not use the normal curve for this approximation. Instead, we use the *Student t* distribution that is based on the sample size. We proceed as before, but we change the table that we use. This distribution looks like the normal distribution, but as the sample size decreases it spreads out. For large *n* it nearly matches the normal curve. We say that the distribution has (*n* – 1) *degrees of freedom*.

Example 12.3:Suppose that we conduct a survey of 19 recordings of flooding volume at a gauging site in *Brahmani* River. To find out what percent of the observations is closes to actual the flood average. We found that the mean is 15 cumecs with a standard deviation of 5 cumecs. Find a 95% confidence interval for the mean percent. Assume that the distribution of all the observations is approximately normal.

Solution: We use the formula:

$$x \pm \frac{t_c s}{\sqrt{n}}$$ (Notice the t instead of the z and s instead of □in Eq.10.8)

We get

$15 \pm t_c\, 5 / \sqrt{19}$

Since n = 19, there are 18 degrees of freedom. Using the *t*-table in the book (Ref- i), we have that

$t_c = 2.10$

Hence the margin of error is

$\pm 2.10\ (5) / \sqrt{19} = \pm 2.4$

We can conclude with 95% confidence that the average flood at the gauging site is between 12.6 *and* 17.4 cumecs.

12.6 UNCERTAINTY OF HYDROLOGICAL ESTIMATES OR PREDICTIONS

12.6.1 Statistical Analysis for Linear Regression

The simple linear form of the equation is

$Y = a + bX$

where a and b are regression coefficients computed, respectively, as

$$b = \frac{\sum_{i=1}^{n}(X_i - \bar{x})(Y_i - \bar{y})}{\sum_{i=1}^{n}(X_i - \bar{x})^2} \qquad (12.10)$$

$$\text{And } a = \bar{y} - b\,\bar{x} \qquad (12.11)$$

Best-fit evaluation criteria are in common use as follows:

(a) Coefficient of determination $r^2 = \left[\sum(x_i - \bar{x})\ (y_i - \bar{y})\right]^2 / [\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2]$

(b) Coefficient of correlation (r): $r = \sqrt{r^2} = \sum(x_i - \bar{x})(y_i - \bar{y}) / [\sum(x_i - \bar{x})^2 \sum(x_i - \bar{y})^2]^{1/2}$

(c) Efficiency (EFF): $EFF = 1 - \frac{S}{S_y}$

where, $S^2 = \sum(y_i - \hat{y}_i)^2/(n-2)$; $S_y^2 = \sum(y_i - \bar{y})^2/(n-1)$

(d) Standard error (SE): $S_a = S\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right)^{1/2}$; $S_b = S/\sum(x_i - \bar{x}^2)^{1/2}$

where S_a and S_b are standard errors of coefficients a and b, respectively.

(e) Confidence interval: The confidence interval for *a* is given as:

$$l_a = a - t_{(1-\alpha/2),\,(n-2)}\, S_a \qquad (12.12a)$$

$$u_a = a + t_{(1-\alpha/2),\,(n-2)}\, S_a \qquad (12.12b)$$

and for b, it is given as:

$$l_b = b - t_{(1-\alpha/2),\,(n-2)}\, S_b \qquad (12.13a)$$

$$u_b = b + t_{(1-\alpha/2),\,(n-2)}\, S_b \qquad (12.13b)$$

where l_a, u_a and l_b, u_b denote lower and upper confidence limits of *a* and *b*, respectively; α is the confidence level; and $t_{(1-\alpha/2),\,(n-2)}$ represent t-values corresponding to (1- α/2) confidence limits and (n-2) degrees of freedom. These t-values can be obtained from the book given in reference - (i).

12.6.2 Test of Hypothesis

Hypothesis H_o: $a = a_o$ versus H_a: $a \neq a_o$ is tested by computing t statistics as follows:

$$t = (a - a_o)/S_a \qquad (12.14)$$

H_o is rejected if $|t| > t_{(1-\square/2),\,(n-2)}$. The t-statistic for a is given by $t_a = (a - a_o)/S_a$, and for b, it is given by $t_b = (b - b_o)/S_b$.

Hypothesis H_o: $b = b_o$ versus H_b: $b \neq b_o$ is tested by computing

$$t = (b - b_o)/S_b \qquad (12.15)$$

H_o is rejected if $|t| > t_{(1-\alpha/2),\,(n-2)}$. Significance of the overall regression is evaluated as follows.

Hypothesis H_o: b = 0 is tested by computing

$$t = (b - 0)/ S_b \qquad (12.16)$$

H_o is rejected if $|t| > t_{(1-\alpha/2),\,(n-2)\,a}$ and the regression equation explaining a significant amount of variation in Y.

12.7 CONFIDENCE INTERVALS FOR LINE OF REGRESSION

The lower and upper confidence intervals are expressed, respectively, as

$$L = \hat{y}_k - S_{\hat{y}_k}\, t_{(1-\alpha/2),(n-2)} \qquad (12.17a)$$

$$U = \hat{y}_{y_k} + S_{\hat{y}k}\, t_{(1-\alpha/2),(n-2)} \qquad (12.17b)$$

where L and U represent lower and upper confidence limits, respectively;

$$\overline{Y}_k = a + bx_k \qquad (12.17c)$$

and

$$S_{\hat{y}_k} = S\left[\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]^{1/2} \qquad (12.18)$$

which is the standard error of $S_{\hat{y}_k}$.

12.8 CONFIDENCE INTERVALS OF INDIVIDUAL PREDICTED VALUES

Confidence intervals of individual predicted values of y are given as

$$L' = \hat{y}_k - S_{\hat{y}_k} t_{(1-\alpha/2),(n-2)} \quad (12.19a)$$

$$U' = \hat{y}_k + S_{\hat{y}} t_{(1-\alpha/2),(n-2)} \quad (12.19b)$$

$$S'_{\hat{y}_k} = S\left[1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]^{1/2} \quad (12.20)$$

The above statistical analysis for the linear regression is illustrated below.

Example 12.4: For a typical catchment, the following precipitation and corresponding runoff data were observed in the month of July:

1. Develop a linear rainfall-runoff relationship using the above data.
2. How much variation in runoff is accounted for by the developed relationship?
3. Compute 95% *confidence interval* for regression coefficients.
4. Test the hypotheses that (a) the intercept is equal to zero and (b) the tangent is equal to = 0.500.
5. Calculate 95% confidence limits for the *predicted* (or fitted) line
6. Calculate 95% confidence interval for an *individual predicted value* of the dependent variable.
7. Comment on the *runoff prediction* in the developed linear relationship

Solution: The solution is provided in sequence as follows:

1. Assume the form of linear equation is: Y = *a* + *b* (X), where Y is the runoff, X is the rainfall, and *a* and *b* are regression coefficients, which can be determined as (Eqs.10 a, b):

Average value of X and Y are 42.9 and 14.6 respectively. Therefore, *a* and *b* is calculated as

Table 12.1 Precipitation – Runoff data for the catchment

Year	Precipitation (mm)	Runoff (mm)
1953	42.39	13.26
1954	33.48	3.31
1955	47.67	15.17
1956	50.24	15.50
1957	43.28	14.22
1958	52.60	21.20
1959	31.06	7.70
1960	50.02	17.64
1961	47.08	22.91
1962	47.08	18.89
1963	40.89	12.82
1964	37.31	11.58
1965	37.15	15.17
1966	40.38	10.40
1967	45.39	18.02
1968	41.03	16.25

Follows:

$$b = \frac{\sum_{i=1}^{n}(X_i - \bar{x})(Y_i - \bar{y})}{\sum_{i=1}^{n}(X_i - \bar{x})^2} = \frac{369 \times 432}{570 \times 0.0559} = 0.648$$

And $a = \bar{y} - b\,\bar{x} = 14.63 - 0.648 \times 42.94 = -13.1951$

Thus, the regression equation is: Y = -13.1951 + 0.648 X.

2. The variation in Y accounted by the regression is determined using the coefficient of determination (r^2) as:

$$r^2 = \frac{\left[\sum(x_i - \bar{x})\ (y_i - \bar{y})\right]^2}{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2} = \frac{\left[\sum(x_i - \bar{x})\ (y_i - \bar{y})\right]}{\sum(x_i - \bar{x})^2} \times \frac{\left[\sum(x_i - \bar{x})\ (y_i - \bar{y})\right]}{\sum(y_i - \bar{y})^2}$$

$$= b \times \frac{\left[\sum(x_i - \bar{x})\ (y_i - \bar{y})\right]}{\sum(y_i - \bar{y})^2} = 0.648 \times \frac{369 \times 432}{363 \times 0.0714} = 0.66$$

$r^2 = 0.66$ also implies that the regression equation explains 66% of variation of Y.

The coefficient of correlation (r) can be given by the square root of (r^2) and can be computed as: $r = \sqrt{0.66} = 0.81$. It is noted that r^2 can vary between 0 and 1 whereas r can vary from –1 to +1.

3. The 95% confidence interval for coefficients *a* and *b* is determined as follows:

(i) Compute the standard error of the regression equation as

$S^2 = \sum(y_i - \hat{y}_i)^2 / (n - 2) = 123 \times 7 / (1-2) = 8.83$

Therefore, S = 2.97.

(ii) Compute standard error of a (S_a) as (Eq. 10.10):

$$S_a = S\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right)^{1/2} = 2.97\left(\frac{1}{16} + \frac{42.94 \times 42.94}{570 \times 0.0559}\right)^{1/2} = 5.39$$

(iii) Compute standard error of b (S_b) as (Eq. 10.10):

$S_b = S / \sum(x_i - \bar{x}^2)^{1/2} = 2.97/(570 \times 0.0559)^{1/2} = 0.125$

(iv) Compute $t_{(1-\alpha/2),\ (n-2)}$ from t-table in book (refer-i) for $\alpha = 0.05$ and n = 16. It is equal to 2.14.

(v) Compute 95% confidence interval for a as (Eq. 10.11):

$l_a = a - t_{(1-\alpha/2),\ (n-2)}\ S_a = -13.1951 - 2.14 \times 5.39 = 24.73$

$u_a = a + t_{(1-\alpha/2),\ (n-2)}\ S_a = -13.1951 + 2.14 \times 5.39 = -1.66$

(vi) Similarly, compute 95% confidence interval for b as (Eq. 10.12):

$l_b = b - t_{(1-\alpha/2),\ (n-2)}\ S_b = 0.648 - 2.14 \times 0.125 = 0.38$

$u_b = b + t_{(1-\alpha/2),\ (n-2)}\ S_b = 0.648 + 2.14 \times 0.125 = 0.92$

4. In the assumed linear relation: Y = *a* + *b*(X), and *a* is the intercept and *b* is the tangent. The first hypothesis is tested for case (a) that a = 0 versus a ≠ 0 (or H_o: a = 0 versus a ≠ 0). To this end, the t-statistic is computed as (Eq. 10.13):

   $t = (a-a_o)/S_a = (-13.1951 - 0.0)/5.39 = -2.44$

   Note that $a_o = 0$. Then, determine the critical t-value for $\alpha = 95\%$ and n = 16, which is equal to 2.14, as computed in step 3 (iv). Since $|t| > t_{(1-\alpha/2),\,(n-2)}$ (= $t_{0.9951,\,14}$), reject H_o: a = 0. It implies that a can not be equal to zero.

   Similarly, the hypothesis H_o: b = 0.5 versus a ≠ 0.5 can be tested as (Eqs. 10.14 and 10.15):

   $t = (b - b_o)/S_b = (0.648 - 0.5)/0.125 = 1.184$

   which is less than $t_{0.9951,\,14}$ (= 2.14). Therefore, H_o: b = 0.5 can not be rejected.

   The above tests imply that the intercept is significantly different from zero whereas b is not significantly different from 0.5. The significance of the overall regression can be evaluated by testing H_o: b = 0 by computing t for $b_o = 0$, which comes out to be 5.184. Since this value of 5.184 is greater than the above critical t-value of 2.14, the hypothesis H_o: b = 0 is rejected.

5. The 95% limits for the regression line are computed in the following steps:

   (i) Compute the standard error $\hat{Y}_k$ as (Eq. 10.17):

   $$S_{\hat{y}_k} = S\left[\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]^{1/2} = 2.97 \times \left[\frac{1}{16} + \frac{(x_k - 42.94)^2}{570 \times 0.0559}\right]^{1/2}$$

   Thus, the lower and upper limits can be given by Eqs. 10.18, respectively, as:

   $L = \hat{y}_k - S_{\hat{y}_k}\, t_{(1-\alpha/2),(n-2)}$ and $U = \hat{y}_{y_k} + S_{\hat{y}_k}\, t_{(1-\alpha/2),(n-2)}$

6. The computation of 95% confidence limits for individual predicted Y-values can be performed as follows:

(i) Compute the standard error of individual predicted values as (Eq. 10.19):

$$S'_{\hat{y}_k} = S\left[1+\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right]^{1/2} = 2.97\,x\left[1+\frac{1}{16}+\frac{(x_k-42.94)^2}{570\,x\,0.0559}\right]^{1/2}$$

Table 12.2 Computations of lower and upper limits for the fitted line

No.	X	Y	$\hat{Y}_i$	L	U
1	42.39	13.26	14.27	12.67	15.87
2	33.48	3.31	8.50	5.52	11.48
3	47.67	15.17	17.69	15.66	19.72
4	50.24	15.50	19.36	16.85	21.87
5	43.28	14.22	14.85	13.25	16.44
6	52.60	21.20	20.89	17.86	23.91
7	31.06	7.70	6.93	3.39	10.47
8	50.02	17.64	19.22	16.75	21.68
9	47.08	22.91	17.31	15.38	19.25
10	47.08	18.89	17.31	15.38	19.25
11	40.89	12.82	13.30	11.62	14.98
12	37.31	11.58	10.98	8.79	13.16
13	37.15	15.17	10.87	8.66	13.09
14	40.38	10.40	12.97	11.24	14.70
15	45.39	18.02	16.21	14.50	17.93
16	41.03	16.20	13.39	11.72	15.06

(ii) Compute 95% confidence limits for individual predicted Y-values as (Eqs. 10.17 and 10.18)

The numerical values of these limits for different values of X are in Table 10.2.

$$L' = \hat{y}_k - S_{\hat{y}_k} t_{(1-\alpha/2),(n-2)} = -13.1951 + 0.648\,x_k - 2.97\,x\left[1.0625+\frac{(x_k-42.94)^2}{570\,x\,0.0559}\right]^{1/2} \times 2.14$$

$$U' = \hat{y}_k + S_{\hat{y}} t_{(1-\alpha/2),(n-2)} = -13.1951 + 0.648\,x_k + 2.97\,x\left[1.0625+\frac{(x_k-42.94)^2}{570\,x\,0.0559}\right]^{1/2} \times 2.14$$

The computed numerical values for the lower and upper limits are shown in Table 10.3 and plotted in Figure 10.4.

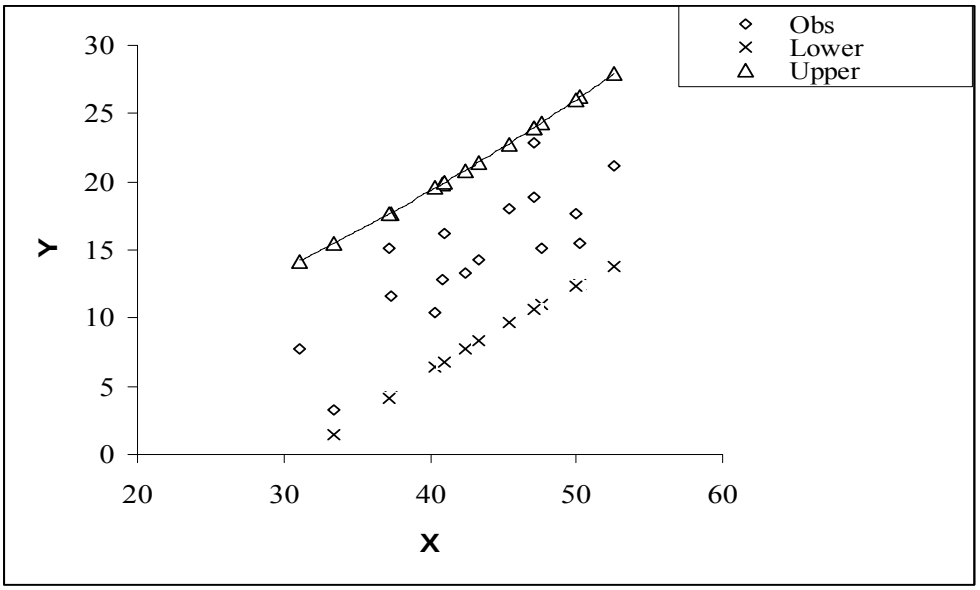

Figure 12.4 The computed numerical values for the lower and upper limits

7. The extrapolation of a regression equation beyond the range of x used in estimating a and b are generally discouraged for two reasons. First, the confidence interval of a fitted line increases as x-values increase beyond their mean, as shown in Figure 10.4. Second, the relation between Y and X may be non-linear whereas a linear relation is fitted for the prescribed range of X-values, as shown in Figure 10.4

12.9 UNCERTAINTY ANALYSIS OF FLOW DURATION CURVES

The flow duration curve for various return periods were developed using the characteristics of distribution of probability plots of stream calculated by four methods i.e. (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formula at monthly time intervals For stations of the thirteen gauged basins having several years of record, flow duration curve can be constructed using mean values of monthly mean values. For this case mean monthly data is used in developing a standard FDC that are independent. Here the monthly mean flows for this referred period (July-Sept) is first averaged for a years, by

taking the mean of all the thirteen gauging site in decreasing order. The FDC during monsoon period is developed using mean flow for the available period 1972 -2009.

Table 12.3 Computation of lower and upper limits for the fitted line

No.	X	Y	$\hat{Y}_k$	L′	U′
1	42.39	13.26	14.27	7.71	20.83
2	33.48	3.31	8.50	1.47	15.52
3	47.67	15.17	17.69	11.02	24.37
4	50.24	15.50	19.36	12.52	26.20
5	43.28	14.22	14.85	8.29	21.40
6	52.60	21.20	20.89	13.84	27.93
7	31.06	7.70	6.93	-0.35	14.21
8	50.02	17.64	19.22	12.39	26.04
9	47.08	22.91	17.31	10.66	23.96
10	47.08	18.89	17.31	10.66	23.96
11	40.89	12.82	13.30	6.72	19.88
12	37.31	11.58	10.98	4.25	17.70
13	37.15	15.17	10.87	4.14	17.61
14	40.38	10.40	12.97	6.38	19.56
15	45.39	18.02	16.21	9.63	22.80
16	41.03	16.20	13.39	6.81	19.97

. In chapter 5 (referred in Appendix 5) shows FDC using Weibull, Blom, Gringorton and Cunnane mean flow, respectively. Figures shows a clear picture of the variation for stations of the thirteen gauged basins having several years of record, flow duration curve can be constructed using mean values of monthly mean values. For this case mean monthly data is used in developing a standard FDC that are independent. Here the monthly mean flows for this referred period (July-Sept) is first averaged for a years, by taking the mean of all the thirteen gauging site in decreasing order. The FDC during monsoon period is developed using mean flow for the available period 1972 -2009

Table 12.4 Flow duration curve using different methods to get p(>Q)

Catchment Area in Sq km	Mean flow (m^3/s)	Flow duration curve using following methods to get p(>Q)			
		WEIBULL	BLOM	GRINGORTON	CUNNANE
Anandpur GD site(8570)	2070.26	3.1	2	1.8	1.92
	317.44	50.0	50	50.0	50.00
	118.07	96.9	98	98.2	98.08
Champua GD site(1710)	325.82	3.1	2	1.8	1.92
	72.85	50.0	50	50.0	50.00
	28.75	96.9	98	98.2	98.08
Gomlai GD site(21950)	3591.31	3.1	2	1.8	1.92
	836.12	50.0	50	50.0	50.00
	298.85	96.9	98	98.2	98.08
Jaraikela GD site(9160)	1981.686	3.1	2	1.8	1.92
	340.6172	50.0	50	50.0	50.00
	113.1675	96.9	98	98.2	98.08
Jenapur GD site(33955)	4040.29	3.1	2	1.8	1.92
	1231.21	50.0	50	50.0	50.00
	511.21	96.9	98	98.2	98.08
Kantamal GD site(19600)	39.81	3.1	2	1.8	1.92
	5.90	50.0	50	50.0	50.00
	2.74	96.9	98	98.2	98.08
Kesinga GD site(11960)	3258.50	3.1	2	1.8	1.92
	421.63	50.0	50	50.0	50.00
	73.11	96.9	98	98.2	98.08
Pamposh GD site(19448)	4043.829	3.1	2	1.8	1.92
	739.3026	50.0	50	50.0	50.00
	194.2735	96.9	98	98.2	98.08
Salebhatta GD site(4650)	1311.20	3.1	2	1.8	1.92
	119.37	50.0	50	50.0	50.00
	22.79	96.9	98	98.2	98.08
Sukma GD site(1365)	1981.686	3.1	2	1.8	1.92
	340.6172	50.0	50	50.0	50.00
	113.1675	96.9	98	98.2	98.08
Sundergarh GD site(5870)	1410.75	3.1	2	1.8	1.92
	239.91	50.0	50	50.0	50.00
	82.21	96.9	98	98.2	98.08
Talcher GD site(29750)	2515.78	3.1	2	1.8	1.92
	1171.17	46.9	46.8	46.8	46.79
	607.88	96.9	98	98.2	98.08
Tikerpada GD site	12384.56	3.1	2	1.8	1.92
	3602.97	50.0	50	50.0	50.00
	1352.31	96.9	98	98.2	98.08

In chapter 5 (referred in Appendix 5) shows FDC using Weibull, Blom, Gringorton and Cunnane mean flow, respectively. Figures shows a clear picture of the variation of the river

flow at this site during the monsoon period and also the probability of flow exceedance in X-axis P(>3,>50,and>97) % exceedance and the corresponding mean flow.

This table shows the relation ship between the catchment area and the mean flow of different thirteen gauging site of Mahanadi basin .For this we took the average of the mean flow of the monsoon period. The flow duration curve for various return periods calculated by four methods i.e. (i) Weibull (ii) Blom (iii) Cunnane and (iv) Gringorton plotting formula at monthly time intervals .The Flow duration curve were made by using the Weibull formula P = m / (N+1)*100 at various probability of flow i.e. P(>3,>50,and>97) % exceedance and the corresponding mean flow. In Blom method we use the formula P = ((m - 0.375) /(N + 0.25))*100 at various probability of flow i.e. P(>2,>50,and>98) % exceedance and the corresponding mean flow. In Gringorton and Cunnane method we use the formula P = ((I - 0.44) / (N + 0.12))*100 and P = ((I - 0.4) / (N + 0.2))*100 at various probability i.e. P(>1.8,>50,and>98.2) % and P(>1.92,>50,and>98.08) % exceedance respectively

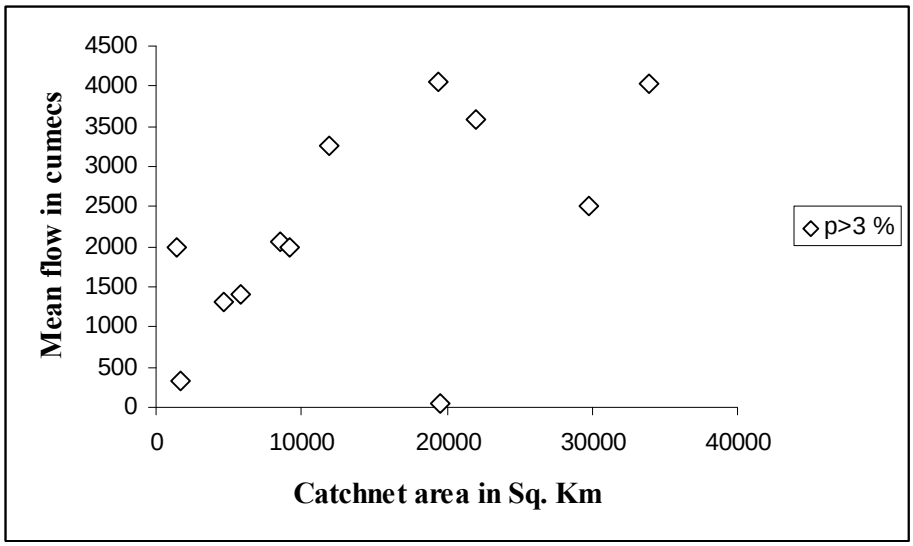

Figure 12.5(a). Flow exceeding (p>3%) using daily data of thirteen catchments in Mahanadibasin for Monsoon (July-Sept)

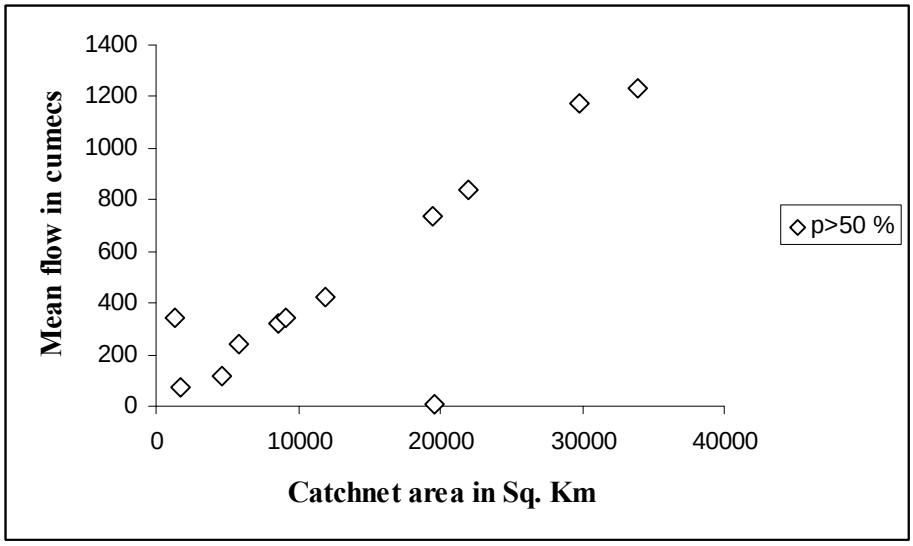

Figure 12.5(b). Flow exceeding (p>50%) using daily data of thirteen catchments in Mahanadi basin for Monsoon (July-Sept)

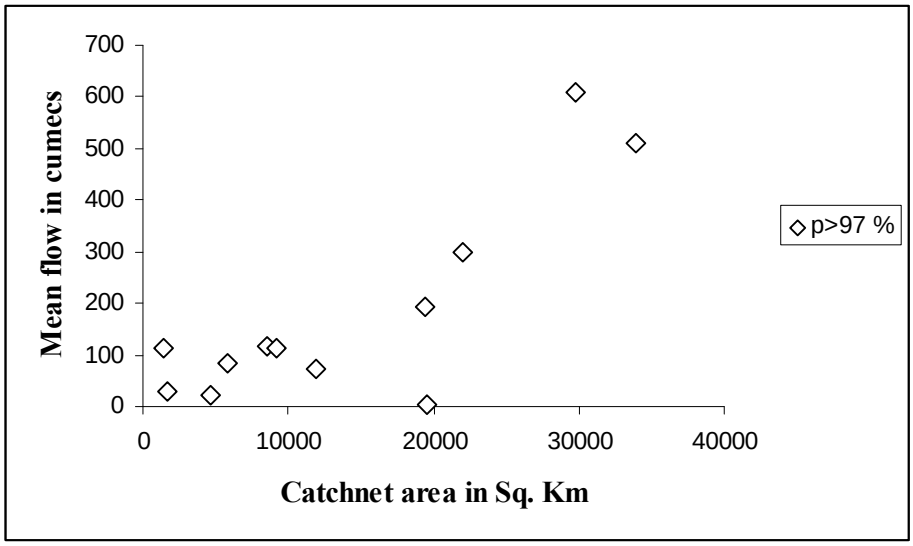

Figure 12.5 (c) Flow exceeding p (%) using daily data of thirteen catchments in Mahanadi basin for Monsoon (July-Sept)

12.10 UNCERTAINTY IN FLOOD HYDROGRAPHS USING UH

Statistical analyses generally dealt in hydrology can be categorized as univariate and multivariate analyses. Problems dealing with the data of a single random variable fall under the

category of univariate analyses. Such studies, for example, include the studies of rainfall analysis. In hydrograph computation analyses, in general more of the existing models that are being popularly used have multivariable data with more than one variable e.g. land use conditions and geomorphologic characteristics that change with time is indexed in SCS-CN method, Clark's model use the storage coefficient that relates to catchments characteristics, GIUH depends on geomorphologic characteristics and Nash has two parameters that are partially coherent to Clark's theory and has concept of statistical like Gamma pdf. They are the functional form of the inherent relationship and is generally derived. To find out what percent of the observations is closes to actual the flood i.e. in this particular case the unit hydrograph , we go for the mean, standard deviation and get a 95% confidence interval. Alternate to such approach, a fairly better approach would have to evaluated the ratio (STDER) of the absolute sum of non-matching areas to the total hydrograph area, expressed mathematically as [HEC-1, 1990] given in Eq. (8.36). In unit hydrograph, no such observed or real can be focused as it is a derivation from flood hydrograph with assumption of base flow separation, infiltration loss, direct runoff at gauging sites in case of runoff axis. More difficult is exact rainfall occurring in the basin that contributes to runoff at GD site. There we go for Thiessen polygon, arithmetic mean etc., but are just an approximate as some raingauge is not available or faulty readings and coverage area of a raingauge. So, in this case conventional method as described in Chapter-8 is taken. In this section the computed unit hydrographs and mean error compared with the conventional methods using the four different analysis for the eleven catchments shows different degree of errors that can also be seen from the earlier sections for a specified model. The mean error is considered here as it also has an inference on effect of two IUH parameters that comes under the model development and the approach for estimating them.

Table 12.5 Computed unit hydrograph (for 11 catchments) using four methods and comparison with conventional (observed) UH with error.

Sl	**Br-No**	**Peak flow in cumecs**				**Observed**	Mean STDER
		Clark's	Snyder's	SCS-CN	GIUH		
1	7	632.99	698.97	642.24	559.44	731.71	12.11
2	121	228.97	241.83	292.86	333.50	249.15	8.18
3	12	223.36	209.86	241.72	271.59	267.23	10.04
4	195	172.38	187.58	193.83	113.22	170.26	0.99
5	235	116.14	122.59	121.92	123.00	109.10	19.13
6	385	63.29	74.38	84.04	53.04	91.23	3.03
7	69	52.13	72.69	89.67	85.50	56.49	2.32
8	59B	63.38	71.56	74.06	58.20	63.47	30.56
9	40K	55.51	53.69	51.50	38.06	64.80	16.13
10	698	35.36	46.69	53.92	40.80	41.82	4.87
11	154	36.07	49.17	71.58	51.98	48.31	4.77

The STDER ranges from 2.32 the minimum to 30,56 as the maximum deviation. To have an glimpse in the confidence band, the following calculation shows for one catchment Br-No 7.

Table 12.6 Computations of lower and upper limits for the fitted line

No.	X	Y	$\hat{Y}_k$	L	U
1	742.39	713.26	731.71	72.67	735.87
2	293.48	253.31	249.15	235.52	261.48
3	307.67	295.17	267.23	255.66	279.72
4	250.24	215.50	170.26	162.85	181.87
5	143.28	114.22	109.10	103.25	116.44
6	122.60	121.20	91.23	87.86	93.91
7	91.06	67.70	56.49	53.39	60.47
8	90.02	87.64	63.47	56.75	71.68
9	97.08	82.91	64.80	61.38	69.25
10	77.08	68.89	41.82	39.38	52.25
11	70.89	62.82	48.31	41.62	49.98

The standard error for the observed i.e. the conventional case values, $\hat{Y}_k$ can be computed as follows (refer Eq. 10.20):

$$S_{\hat{y}_k} = S\left[\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]^{1/2} = 2.97 \times \left[\frac{1}{16} + \frac{(x_k - 42.94)^2}{570 \times 0.0559}\right]^{1/2}$$

Thus, the lower and upper limits can be given by Eqs. 18, respectively, as:

$L = \hat{y}_k - S_{\hat{y}_k}\ t_{(1-\alpha/2),(n-2)}$ and $U = \hat{y}_{y_k} + S_{\hat{y}k}\ t_{(1-\alpha/2),(n-2)}$

The numerical values of these limits for different values of X are in Table 10.6.

In hydrologic analysis e.g. for UH computation, one method deals with several hydrological characteristics simultaneously. For example, in defining the UH on a regional basis ,an engineer often relates the peak discharge, time –to-peak, and shape parameters to basin and meteorological characteristics. In fact, parameters describing a UH are often correlated. The conventional practice of univariate regression analysis treats each UH parameter separately and, hence ,does not take into account the correlation among the UH parameters.

For a given watershed, several methods, based on the statistical information of N and K, can e applied to estimate the statistical moments of Nash IUH ordinates and those of the corresponding DUH. Frequently, rainfall-runoff analysis and modeling have to be perfumed for watersheds in which data are not available or existing data are too scarce, as in the present study where it deals with watershed with paucity of flood data. Although there are various types of techniques that have been used in hydrological regionalization procedure, the primary objective of this study is to examine the performance of various regression and other available techniques applicable to hydrologic regionalization.

12.11 UNCERTAINTY IN FLOOD FREQUENCY ANALYSIS

Approximate expressions for the *T*-year event estimation corresponding to a probability of non-exceedance, and its Root Mean Square Error (RMSE) using PT3, LP3, GEV, WAKEBY, GP, GL and EV1 distributions are deduced in Appendix 9c allowing for a

comparison of different models. Furthermore Monte Carlo simulations are used here in this section in order to check the accuracy of the approximations. The combination of different models and estimation methods that yields the least variance of $\hat{q}(T)$ is selected to be the best method. For AMS model given in Chapter 9, the asymptotic variance of the *T*-year event estimate was given by Cunnane (1973) and Phien (1987) as follows:

$$Var_{AMS}[\hat{q}(T)] = \frac{\beta^2}{N}\left[1.168 + 0.192 y_T + 1.100 {y_T}^2\right] \quad (12.21a)$$

In case of method of maximum likelihood, the asymptotic variance for the *T*-year event estimator reads

$$Var_{AMS}[\hat{q}(T)] = \frac{\beta^2}{N}\left[1.109 + 0.514 y_T + 0.608 {y_T}^2\right] \quad (12.21b)$$

Similarly, the asymptotic variance in case of PWM estimation was given by Phien (1987) and later corrected by Madsen et al.(1997) as

$$Var_{AMS}[\hat{q}(T)] = \frac{\beta^2}{N}\left[1.113 + 0.457 y_T + 0.805 {y_T}^2\right] \quad (10.21c)$$

A comparison of variance of $\hat{q}(T)$ reveals that the variance ratios of two models are independent of sample size; hence this index is used in the present study to compare any two models. Before that the accuracy of the expressions of STDER –values given in Appendix 9c are checked using a simulation procedure using Monte Carlo experiments.

Simulation Algorithm

The Monte Carlo experiment uses the following simulation algorithm:

(i) Choose an initial value of the mean annual number of exceedances $[\mu]$ and variance $[\sigma^2]$ such that probability of no exceedance i.e. $e^{-\mu}$ and $(1-p)^k$, given by PD and NB distributions, respectively, are small.

(ii) Independent of μ and σ^2, initial values of q_0, m_0, and α for a PDS (or AMS) model are selected to generate the series.

(iii) The number of exceedances r in a single year are simulated from a parent PD distributions with the given μ, and for a given value of r exceedances q_r, r =1, 2, 3…n_r are drawn from logistic distribution with parameters m_0 and α_0. The annual maximum is then determined as $q_i = \max \{q_r\} + q_0$. This is repeated N times to get an AMS of length N years and an initial PDS sample with Σn_r exceedances.

(v) The T-year flood using AMS/PD-(with the five distributions given above) with all the three estimation methods, i.e., MOM, ML, and PWM are estimated.

(vi) Step (i)-(v) were repeated a large number of times [called number of simulations], and on the basis of the resulting samples of $q(T)_{AMS}$, the approximate $Var[\hat{q}(T)]$ for different parameter estimation is computed.

A brief result of the variances of LP3, GEV, WAKEBY, GP, GL and EV1 distributions for the catchments given in Appendix 9c catchment used in this study is given in Table 10.7.

Table 12.7 Variance of return period flood using seven pdfs using simulations.

T (yrs.)	PT3	LP3,	WAKEBY	GP	GL	GEV	EVI
10	45.22	62.38	57.56	61.46	51.53	42.16	60.95
50	67.4	83.2	89.6	101.3	95.3	85.2	79.8
100	97.7	137.5	104.3	116.8	173.5	143.2	102.3
500	122.31	142.65	117.69	139.16	212,94	193.11	128.64
1000	182.6	192.6	165.4	193.1	312.3	211.2	245.3

T= Return period in years.

For the PDS model, the results shows that variance of the flood estimate derived using PWM estimation method to be consistently small compared to MOM and ML method. The results of ratios of Var [q (T)] for all the catchments (Appendix 9c) lie in the range of 0.92 -0.94 for all estimation methods, which means that the PDS model perform marginally better than the

AMS model. Similarly Table 10.7 compares the seven AMS models on basis of the variance of the flood estimates, and as observed the GEV, LP3 and WAKEBY models yields less variance of q(T) than others for different return periods and is consistent.

12.12 RESULTS AND DISCUSSION

The chapter describes the concepts and methods of statistics and probability for the quantification of uncertainty in the hydrologic predictions. The available hydro-meteorological data in the Mahanadi basin is used to illustrate these concepts and methods through various examples. These methods are further extended to estimate the uncertainty in the regional Flow Duration Curves (FDC), flood hydrograph using Unit Hydrographs (UH) and recurrence flood using flood frequency analysis. A set of thirteen small watersheds distributed throughout Mahanadi basins are considered for the above purpose. These watersheds cover a range of watershed characteristics including soil type, topography and land use, as well as climate. Time series of daily precipitation, potential evapotranspiration and streamflow are available for a 10-35 year period.

Figure 12.5 is the type of diagram that would be produced if many hydrological ensembles were all subject to FDC analysis in the yield model – a process that would be very time consuming in all but simple systems. The pre-processing would be based on ranking the hydrology ensembles on the basis of the minimum flow volume over the critical duration for the system. Identifying the characteristics of the minimum flow duration distribution and deciding on appropriate uncertainty probabilities (the median plus the ensembles equaled or exceeded >3%, >50% and 97% of the time for example) would allow three flow ensembles to be selected for FDC analysis.

Adding hydrological uncertainty based on the uncertainty interpretation of the FDC results would have to be dealt with in a different way. In this all of the yield estimates from all of the FDC sequences run with each hydrological ensemble would be treated as uncertainty estimates of the 'real' yield and the yield probability curve would look something like that shown in Figure 12.5. Note that it is now no longer appropriate to associate the exceedence probabilities with recurrence intervals. The Figure 12.5a shows approximately >3% probability and the graph is exponential, the Figure 12.5b shows approximately >50% probability and the graph is straight line and The Figure 12.5c shows approximately >97% probability and the graph is Scattered based on FDC using the median hydrology ensemble. While the parts of the three curves that correspond with each other will clearly depend on the degree of uncertainty in the hydrology ensembles and this example has used hypothetical data, the result should be very similar for all situations. The conclusion is therefore that yield assessments based on a probabilities of exceedence of between about >3% and >97% would not be affected by including hydrological uncertainty. The results are based on the detailed analysis of Flow duration curve using four methods (Weibull, Blom, Gringorton and Cunnane) to get p(>Q) for the all 13 watersheds.

Figure 12.5a and Table 12.4 illustrate the results of applying the unstructured sampling program (no stochastic daily flow inputs) to the Mahandi basin data base and the FDC model. There are 13 sub-basins (small catchments) and about two uncertain parameters. It is very clear that the high flow exceedance probabilities i.e. < 3 % sampling has resulted in much lower uncertainty compared to low flow exceedance i.e. > 96 % in the basin. Figure 12.5 (a) and (c) present these two cases and Figure 12.5(b) is for middle of > 50 % exceedance. It is very clear from Figure 12.5 (a) – (c) that the structured samples have resulted in greatly increased uncertainty ranges for the data observations either through current meter or gauge sites at the

basin outlet. All of the flow data presented are given as a fraction of the mean monthly flow for the median simulations and are therefore comparable across different parts of the catchment. For the low flow group samples the equivalent values range are 29 – 2100 cumecs for high flows (> 3 % exceedance) and 23 – 1400 cumecs for low flows (> 96 % exceedance) 3, indicating the there is no bias between the two sampling schemes. It is also apparent that the P % exceedance and the corresponding mean flow uncertainty has increased by between 18 and 30% depending on the position on the FDC.

Similarly the Table 12.5 shows the computed unit hydrographs and mean error compared with the conventional methods using four different analysis for the eleven catchments. These show different degree of errors ranging from minimum 2.32 to maximum 30.56. Table 12.6 illustrates the calculation of confidence band, for one of the catchment (Br-No 7). The approximate expressions for the *T*-year event estimation corresponding to a probability of non-exceedance, and its Root Mean Square Error (RMSE) using seven different pdf has been attempted in section 12.11. Table 12.7 provides the variance of return period flood and show that GEV, LP3 and WAKEBY models yields less variance of $q(T)$ than others for different return periods and is consistent.

CHAPTER 13
PEAK OVER THRESHOLD TECHNIQUES

13.0 INRODUCTION

Two approaches commonly used for probabilistic analysis of extreme flood magnitudes are based on annual maximum series (AMS) and partial duration series (POT). The AMS model considers the annual maximum flood of each year in a series that has as many elements as there are years in the data record, notwithstanding that the secondary events in one year may exceed the annual maxima of other years. In some regions, annual maximum floods observed in dry years may be very small, and inclusion of these events can significantly alter the outcome of the extreme value analysis. In contrast, the POT model avoids all these drawbacks by considering all flood peaks above a certain threshold level; hence it is also referred to as peak over threshold (POT) method. Clearly in the POT model, more than one flood per year may be included, and exceptionally low annual maximum values that may cause problems in parameter estimation are not included. The POT models are often used for flood frequency analysis (FFA) when there is a paucity of data (Martins and Stedinger, 2001). In the limit, when the threshold is increased, recurrence interval of large events computed using AMS and POT models tend to converge.

Considering n to be the number of times the flood peaks are above a threshold (q_0), Shane and Lynn (1964) assumed this number to follow a Poisson distribution (PD) and the flood exceedance to be exponentially distributed, since many commonly used distributions are approximately exponential with a stretched upper tail. The dependence of flood peaks in POT may have an effect on the ability of a Poisson distribution to adequately describe the number of peaks above threshold, not in terms of the mean number but in terms of variance of the number. Although the exponential distribution (ED) is widely used for fitting observed flood frequency of annual number of exceedances, similar alternate frequency distributions have been proposed,

e.g., the generalized pareto distribution (GPD) which has the ED as a special case when the shape parameter (κ) is equal to zero (Martins and Stedinger,2001). The advantage of exponential distribution over other distributions in fitting the flood exceedances is that it has a single parameter.

With the above background, the present chapter discuss the partial duration series approach

to flood frequency analysis. The proposed model uses the ED and PD distribution to model the peaks exceeding a threshold, and the flood exceedances are fitted using generalized pareto distribution and compares the results with the three-parameter generalized extreme value (GEV) distribution. This chapter further checks the results with the variance of the *T*-year event estimates using method of moments (MOM), and probability weighted moment (PWM) estimation methods using Monte Carlo experiments.

13.1 DEVELOPING MODELS FOR POT ANALYSIS

13.1.1 Poisson Distribution

In the POT model, all peak events above a threshold level are considered and the number of peaks is assumed to follow a Poisson distribution having the probability mass function:

7410 k = 0, 1, 2 … (13.1)

where $P(r)$ is the probability of r events in N years, and μ is the expected number of threshold exceedances in a year. It is noted here that (13.1) is true as long as the flood peaks are independent (Langbein, 1949). For Poisson distribution, MOM and PWM estimations are identical and the estimator is given by (Cunnane, 1973)

$$\hat{\mu} = \frac{n}{N} \tag{13.2}$$

where $\hat{\mu}$ is the population estimate of μ and n is the number of observed exceedances in N years. For the μ estimator, the mean and variance are obtained as

$$E(\hat{\mu}) \approx \mu; Var(\hat{\mu}) \approx \mu / n \qquad \text{for} \quad n \geq 5 \tag{13.3}$$

Cunnane (1979) discussed ratio of variance to mean for some real catchments in detail and showed that for a sufficiently large μ (greater than 5), the distribution of r is symmetrical and it asymptotically approaches a normal distribution. For no flood exceedances to occur in a span of N years, substituting $r = 0$ in (13.1), the following is obtained

$$P(r = 0) = \frac{e^{-\mu N}(\mu N)^0}{0!} = e^{-\mu N} \tag{13.4}$$

which is the *pdf* of ED. The cumulative mass function is given by

$$F(x) = P(X \leq x) = 1 - \mu e^{-\mu N} \; ; \qquad N > 0 \tag{13.5}$$

where X is a variate.

13.1.2 Generalized Pareto Distribution

The Generalized Pareto (GP) distribution introduced by Pickands (1975), has the cdf

$$F(q) = 1 - \exp\left[-\frac{q - q_0}{\alpha}\right] \qquad ; \kappa = 0 \tag{13.6a}$$

$$F(q) = 1 - \left[1 - \kappa\frac{q - q_0}{\alpha}\right]^{1/\kappa} \qquad ; \kappa \neq 0 \tag{13.6b}$$

where α is the scale parameter, κ is the shape parameter, and q_0 is the threshold. For $\kappa = 0$, ED is obtained as a special case. For $\kappa \leq 0$, the range of q is $q_0 \leq q \leq \infty$ and for $\kappa \geq 0$ the upper bound exists: $q_0 \leq q \leq q_0 + \alpha/\kappa$. Assume that a threshold level q_0 is chosen, corresponding to a mean

annual number of exceedances of Poisson distribution μ; then at any higher threshold $q_1 > q_0$, the mean occurrence rate of exceedances over the new threshold is (Flood Studies Report, 1975)

$$\mu' = \mu\ [1 - F(q)]$$

Since F (q) ≤ 1, it follows that $\mu' \leq \mu$ i.e. the truncated mean of exceedances above the new threshold q_1 decreases. An important feature of the GP distribution in a POT context is that a truncated GP distribution remains a GP distribution, implying thereby that the choice of threshold is not critical for the assumption of the type of exceedance distribution (Madsen et al., 1997).

13.2 PEAK OVER THRESHOLD BASED FLOOD EXCEEDANCE MODELS

The AMS model used for determination of *T*-year return period flood [$q(T)$] requires identification of a statistical distribution so as to obtain the $q(T)$ - T relationship in explicit form. If n floods exceed q_0 in N years of record in POT, the average rate of occurrence is $\mu = n/N$. Consequently, the flood which has a return period T' sampling units in this POT has a return period of approximately $T = T'/\mu$ years. Thus, a POT method uses two probabilistic models: (a) one for the probability of occurrence (P_λ) for μ exceedances in each year, and (b) the statistical distribution $F_{\square,y}$ of the maximum of n exceedances, where $y = q - q_0$. The probability of annual maxima is obtained using the *cdf*, computed using the total probability theorem (Shane and Lynn, 1964) as

$$F(q) = \sum_{r=0}^{\infty} P_\mu(r)(F_{\mu,y})^r \qquad (13.7)$$

If P_μ follows a PD, and $F_{\square,y}$ follows an ED = 1- exp [-(q-q_0)/ β] with parameter β, then

$$F(q) = \sum\ (F_{\mu,y})^r \frac{1}{r!}(\mu^r e^{-\mu}) = \exp\left(-\exp\left(-\frac{1}{\beta}(q - q_0)\right) + Ln(\mu)\right) \qquad (13.8)$$

The $F(q)$ given by (13.7) is the Gumbel or EV1 distribution that is widely used in FFA. From (13.7) the following POT model can be obtained

$$q(T) = q_0 + \beta [\text{Ln}(\mu)] + \beta y_T \quad (13.9)$$

where $y_T = -\ln(-\ln(1-1/T))$ is the Gumbel reduced variate

The T-year event is obtained using (13.9) by inserting the estimated POT parameters. Madsen et al. (1997) denoted return period of the POT model as T_p in years; that is the event which is exceeded once in T_p years, is in POT context usually defined as the $(1- 1/\mu T_p)$ quantile in distribution of the exceedances. The model described by (13.9) is based on the hypothesis that: (i) the exceedance process is Poisson distribution, (ii) the exceedances are independent variates, and (iii) the exceedances follow a negative exponential law with a constant parameter. Similarly, using a negative binomial distribution for counting the exceedances over threshold and exponential distributed exceedances Önöz and Bayazit (2001) have derived the corresponding AMS/GP-ED expression for calculating the T – year flood exceedances as

$$q(T) = q_0 + \beta \ln\left(\frac{p}{1-p}\right) - \beta \ln\left(F^{-1/r} - 1\right) \quad (13.10)$$

13.3 RESULTS AND DISCUSSION

13.3.1 POT/PD-GP Model

The T-year event q (T), i.e. the level which on the average is exceeded once in T years, is in a POT context defined as the [1-1/ (μT)] quantile in the distribution of the exceedances. Therefore, in a POT/PD-GP model where occurrences of peak exceedances is Poisson distributed, a 1/Tp quantile of the fitted GP distribution is obtained by inserting [1-1/(μT)] for F (q) in Eq. (13.6) as follows:

$$1-\frac{1}{\mu T} = 1 - \exp\left[-\frac{q-q_0}{\alpha}\right], \text{ or } q(T) = q_0 + \alpha Ln(\mu T_p) \quad ; \kappa = 0 \quad (13.11a)$$

this is the POT/PD-ED model; similarly the POT/PD-GP model is deduced as follows:

$$q(T) = q_0 + \frac{\alpha}{\kappa}\left[1 - \left(\frac{1}{\mu T}\right)^{\kappa}\right] \quad ; \kappa \neq 0 \quad (13.11b)$$

In eq. (13.11), μ is the mean occurrences in PD given by Eq. (13.1) which can be estimated from the POT data.

13.3.2 AMS Model

The T-year return period is defined as the average interval of time within which the magnitude of q of the flood event will be equaled or exceeded once. By this definition, the POT model should be a more natural candidate than the AM model for estimating the q-T relationship. The first such comparison was made by Cunnane (1973) who compared the POT model described by the ED and Poisson arrival rate with AMS model described by the Gumbel distribution, and derived that for cases where the average number of peak exceedances per year was greater than 1.65, POT approach should be preferred.. Thus, a comparison of the AMS and POT model is of practical interest. The following section derives the AMS/PD-GP model.

13.3.3 AMS/PD-GP Model

For producing AMS samples consistent with POT samples which are used in simulations, as shall be discussed later in this study, the T-year flood for the annual maximum distribution corresponding to the parent POT model is derived as follows. Considering Poisson distributed rate of exceedances and a GP distributed flood exceedances, Madsen et al (1997) have expressed q (T) as follows:

$$q(T) = q_0 + \frac{\alpha}{\kappa}\left(1 - \mu^{-\kappa}\exp(-\kappa y_T)\right) \quad ; \kappa \neq 0 \quad (13.12a)$$

where, $y_T = -\ln(-\ln(1-1/T))$ is the Gumbel reduced variates, and is an equivalent to an AMS/PD-GP model. Similarly for $\kappa = 0$, expression for AMS/PD-ED model is

$$q(T) = q_0 + \alpha(\mathrm{Ln}(\mu y_T) \quad ; \kappa = 0 \qquad (13.12b)$$

13.3.4 Sampling Variance of the T-Year Flood

Expressions for the T-year event estimation corresponding to a probability of non-exceedances equal to 1-1/T, and its asymptotic sampling variance are calculated using a Taylor's expansion as follows:

$$\mathrm{Var}[\hat{q}(T)] = \left(\frac{\partial[q(T)]}{\partial\alpha}\right)^2 \mathrm{Var}(\hat{\alpha}) + \left(\frac{\partial[q(T)]}{\partial\kappa}\right)^2 \mathrm{Var}(\hat{\kappa}) + 2\left(\frac{\partial[q(T)]}{\partial\alpha}\right)\left(\frac{\partial[q(T)]}{\partial\kappa}\right)\mathrm{Cov}(\hat{\alpha},\hat{\kappa}) + \left(\frac{\partial[q(T)]}{\partial p}\right)^2 \mathrm{Var}(\hat{p})$$
$$+\left(\frac{\partial[q(T)]}{\partial k}\right)^2 \mathrm{Var}(\hat{k}) + 2\left(\frac{\partial[q(T)]}{\partial p}\right)\left(\frac{\partial[q(T)]}{\partial k}\right)\mathrm{Cov}(\hat{p},\hat{k}) + \left(\frac{\partial[q(T)]}{\partial\mu}\right)^2 \mathrm{Var}(\hat{\mu}) \qquad (13.13)$$

For the POT/PD-GP models, the partial derivatives of q (T) are derived as

$$\frac{\partial q(T)}{\partial\alpha} = \frac{1}{\kappa}\left[1 - \left(\frac{1}{\mu T}\right)^{\kappa}\right]; \quad \frac{\partial q(T)}{\partial\kappa} = \frac{\alpha}{\kappa^2}\left[\left(\frac{1}{\mu T}\right)^{\kappa}\left[1+\kappa\ln(\mu T)\right] - 1\right] \qquad (13.14a)$$

$$\frac{\partial q(T)}{\partial\mu} = \frac{\alpha}{\mu}\left(\frac{1}{\mu T}\right)^{\kappa}; \frac{\partial q(T)}{\partial p} = \frac{\kappa}{p(1-p)}\left(\frac{1}{\mu T}\right)^{\kappa}; \quad \frac{\partial q(T)}{\partial k} = \frac{\alpha}{k}\left(\frac{1}{\mu T}\right)^{\kappa} \qquad (13.14b)$$

Similarly, For AMS/PD-GP model, the partial derivatives of q (T) are derived as follows:

$$\frac{\partial q(T)}{\partial\mu} = \frac{\alpha\exp(-\kappa y_T)}{\mu^{\kappa+1}}; \frac{\partial q(T)}{\partial\alpha} = \frac{1}{\kappa}\left[1 - \mu^{-\kappa}\exp(-\kappa y_T)\right];$$

$$\frac{\partial q(T)}{\partial\kappa} = \frac{\alpha}{\kappa^2}\left[-1 + (1+\kappa y_T + \kappa\mathrm{Ln}(\mu))(\mu^{-\kappa}\exp(-\kappa y_T))\right] \qquad (13.15)$$

Substituting in Eq. (13.13) the partial derivatives for q (T), and variance and covariance of the parameters, the Var $[\hat{q}(T)]$ is obtained (details derivations given by ÖnÖz and Bayazit, 2001). Since the expression for Var $[\hat{q}(T)]$ is lengthy, in a simple way can be as expressed as follows:

$$\mathrm{Var}[\hat{q}(T)] = f_1(\mu_p, k, \alpha, \kappa, \nu, T) \quad (13.16)$$

13.3.5 Application Study Area

Fourteen test catchments used in this study are as are earlier given in Chapter 3 and Appendix-2. Theses catchments were selected as they contains the daily series for long period e.g. 30 years, and the similar short term series required for POT study was not available in bridge twenty three bridge catchments.

13.3.6 Checking for *Var* [$\hat{q}(T)$]

The objective here is to check the validity of expressions for variance of $\hat{q}(T)$ given in Eq. (13.16) . This was done by Monte Carlo simulations; for sample sizes, N=10, 20, 30, and 50 years. For this analysis, a ratio R_1 defined as

$$R_1 = \frac{Var[\hat{q}(T)]_{simulation}}{Var[\hat{q}(T)]_{analytical}} \quad (13.17)$$

is used; a value of R_1 equal to one means that the expressions are correct. The initial values for the simulation procedure were taken from the *Pamposh* GD site data (Table 3.6). For the POT series, an initial value of μ was taken as μ =3.7 and σ^2= 6.4, the probability of no exceedances e^{μ} is 0.06, which means that in the proportion 0.06 of N years the AMS is less than the threshold $q_{0,}$ for which the generated AMS samples cannot give any information. Since this quantity is very low, it is assumed that the generated sample retains the property of the real data to a fair extent. For the selected PD parameters, the corresponding GP distribution (PWM) parameters of the POT model are: q_0 = 250 m^3/s, α_0 = 120.3 m^3/s, and κ_0 = –0.0032. For each case 50,000 samples were generated and the corresponding Var $\hat{q}(T)$ at T = 10, 50, 100, 150,

300 and 500 years were computed. The present simulation was restricted to low recurrence intervals i.e. at 10 years as the study deals with the at-site analysis only.

For the AMS/PD-GP model, R_1 approaches one (Figure 13.1), especially when N=50. However, in case of smaller record length e.g. at N=10, the difference between the simulated and the analytical results becomes slightly significant with increase in return period e.g. R_1 is 0.92 for T=300 years.

Using POT/PD-GP models the results of R_1 are observed to be similar in falling trends, and they are reported in Figure 13.2. From Figure 13.1, where the accuracy of the $\mathrm{Var}\,\hat{q}(T)$ expressions for the AMS models is tested using ratio R_1, the analytical expressions are less than the corresponding simulation procedure at lower T till up to 250 years where the ratio is less than 1. Looking at results, the variances given by analytical expressions are overestimated in case of AMS models whereas they are underestimated for POT models (Figure 13.2), following the above said reason.

To examine the affect, if any, of the choice of initial simulation parameters on the outcome of R_1, a new simulation was carried out changing the initial value of μ to 3. Out of 50,000 samples that were generated taking N=50, about 18 samples with $\mu' \geq \mu$ were rejected, as q_0' cannot be calculated for such cases. The results for this simulation were observed to be very similar with the earlier one, which means that the choice of initial values does not affect the outcome of this analysis. Thus, the results so far infer that the variance of q (T) computed using simulated values comes close to the results obtained by approximate analytical expressions.

13.3.7 Comparison of AMS and POT Models

The performance of the POT and the corresponding annual maximum series models as a T–year flood estimator are compared in this section, based on the ratio of their variance of q (T) defined as

$$R_3 = \frac{\mathrm{Var}[\hat{q}(T)]_{PDS}}{\mathrm{Var}[\hat{q}(T)]_{AMS}} \qquad (13.18)$$

If R_3 is less than one it means that the POT model gives less variance of the q (T) than the respective AMS model; a value of one indicates that both the models give the same variance. It can be verified that R_3 is a function is independent of sample size. Figure 13.3 shows R_3 for T = 50,100,200,500 and 1000 years and μ in the range of 0.8-3.8. From the results it is observed that R_3 = 1 at μ = 1.4 to 1.65 for T = 50-1000 years i.e., AMS/PD-GP and POT/PD-GP models yields the same $\mathrm{Var}[\hat{q}(T)]$ for the specified value of μ. . The performance of these two models in the POT context is checked using real data in the subsequent section.

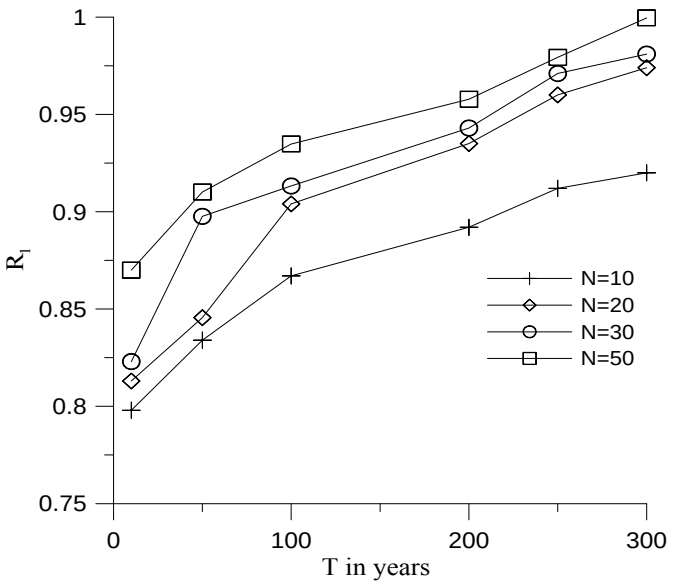

Figure 13.1 Ratio of variances R_1 values for AMS/PD-GP model. (N=Record length in years, μ =3.7 and σ^2= 6.4)

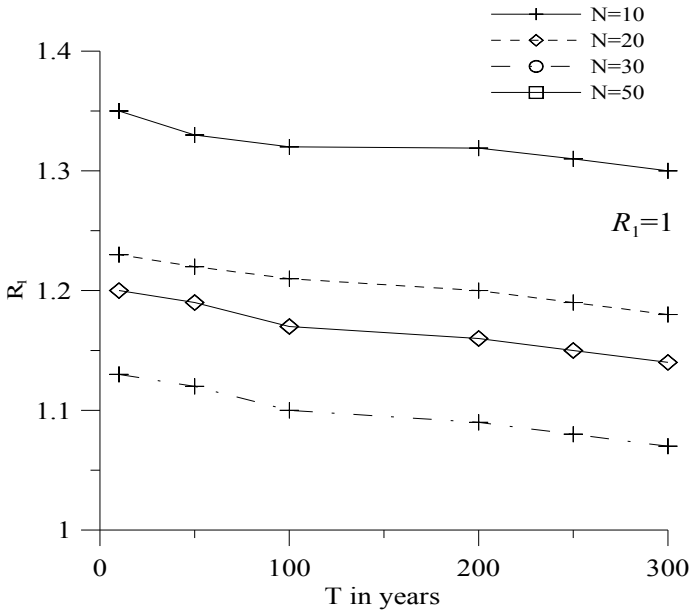

Figure 13.2 Ratio of variances R_1 values for POT/PD-GP models (N=Record length in years, μ =3.7 and σ^2= 6.4)

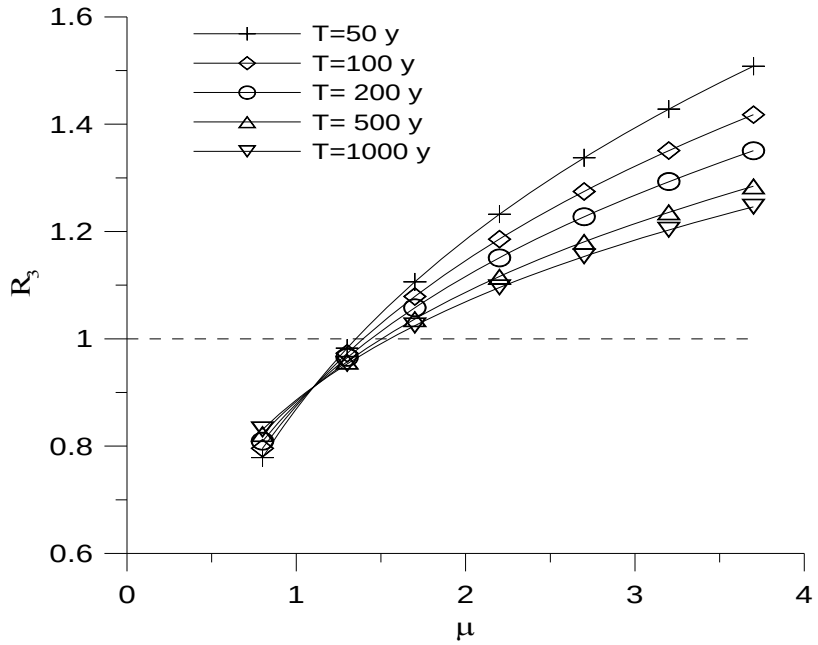

Figure 13.3 Ratio of variances of T-year estimates using POT/PD-GP and AMS/PD-GP models

13.4 DISCUSSION AND CONCLUSION

The analysis presented here focused on the choice of generalized Pareto distribution for computation of flood exceedance in PDS model, when the occurrences of peak exceedances are fitted with a PD distribution. The advantage of using GP distribution with PD in the PDS model has been examined in the past by many researchers, as referred in the introduction section, however, when choosing the most efficient T-year event estimator model, one should not only consider whether to use PDS or AMS but should also carefully choose the estimation method. Thus, one has to compare the models for PDS and AMS with the estimator method. The scope of this chapter however has been restricted to method of probability weighted moments since the present analysis solely focused on the advantage of negative binomial distribution in the PDS model. The same procedure could be used with other estimation methods in future studies. Discussions of the various issues dealt in this study are briefly presented below.

A Taylor series expansion was used to derive the variance of the T-year flood estimates for PWM estimation method. The reliability of the expressions for $\mathrm{Var}[\hat{q}(T)]$ derived using the analytical procedure was checked using the Monte Carlo simulation method. In selecting the initial parameters for the simulation, the PDS data of Mahanadi at *Tikaerpada* GD site were used. A ratio between the $\mathrm{Var}[\hat{q}(T)]$ of analytical and simulation procedures denoted as R_1 was used to check the reliability of these expressions. Results showed R_1 to be asymptotically reaching 1 for AMS and PDS models, especially for higher record length i.e., at N=50. The difference between the simulated and the analytical results becomes more prominent with increase in T years, with a significant difference e.g., $R_1 = 0.92$ for AMS/PD-GP model for a 300 years return period, when a 10 years sample was used. The difference is sometimes caused by an error in evaluation of the parameter estimates variances and covariance, when based on

the expected information matrix. The information matrix, i.e., the matrix evaluated for the estimated parameters, is a better estimator of the covariance matrix (Prescott and Walden, 1983).

For the AMS model the variance of quantiles derived using analytical expressions are overestimated while the PDS models underestimate the same. This is evident from Figure 13.2 and 13.3 for POT/PD-GP and AMS/PD-GP, respectively where the R_1 values lies below one for most of the cases. The above referred figure shows the change of R_1 values with T in case of PDS models, and R_1 lies over one for all cases in the range of 10 to300 years return period. Martins and Stedinger (2000, 2001) while discussing the use of generalized extreme value and generalized pareto distribution in POT models have suggested the use of certain restrictions to the estimated parameters to physically reasonable range. This was not tried in this study though the authors believe that incorporation of this restriction may improve the present results. However, to check the effect of choice of initial parameters selected for the simulation on the overall results of the ratios of variances, a different set of test was carried out with a new threshold. But the outcome was found to be similar. Hence, the choice of initial values of parameters in simulation was found not to affect the results of R_1.

The advantages of the Poisson distribution over negative binomial distributions have been dealt in the past by ÖnÖz and Bayazit (2001). On the basis of a detailed study they reported that the flood estimates and their corresponding variance based on the binomial and negative binomial distribution when combined with ED for the magnitudes of the peaks are almost identical to those obtained using PD. The above conclusions were in agreement with the findings of Kirby (1969) and Cunnane (1979). However, most of these studies focused on exponential distribution in POT models, and no such study have been reported in literature so far to where PD distribution is used with GP distribution for modeling the flood exceedances in

PDS. From the overall results (Table 2) it can be observed that the Poisson distribution performs slightly better when the difference between mean and variance is small e.g. when $\mu = 1$ and σ^2=1.2.

Table 13.1 Ratio of Var [$\hat{q}(T)$] of POT/PD-GP and AMS/PD-GP Models when mean of the POT is 1.

T in years	R_2 for κ= 0.2 for different variances			R_2 for κ= 0.1 for different variances			R_2 for κ= -0.2 for different variances		
	1.2	1.4	1.6	1.2	1.4	1.6	1.2	1.4	1.6
50	0.753	1.347	1.335	0.831	1.410	1.449	1.594	0.787	0.618
100	1.087	0.961	1.591	0.950	0.650	0.617	1.200	0.928	0.871
150	1.037	0.983	0.933	0.984	0.896	0.886	1.125	0.955	0.919
300	1.013	0.994	0.972	0.987	0.968	0.965	1.066	0.976	0.958
500	1.007	0.997	0.990	0.989	0.983	0.981	1.044	0.984	0.972

Table 13.2 Ratio of Var [$\hat{q}(T)$] of POT/PD-GP and AMS/PD-GP Models when mean of the POT is 2.

T in years	R_2 for κ= 0.2 for different variances			R_2 for κ= 0.1 for different variances			R_2 for κ= -0.2 for different variances		
	2.5	3.0	3.5	2.5	3.0	3.5	2.5	3.0	3.5
50	0.696	0.763	0.770	-	-	-	0.352	0.505	0.530
100	0.922	0.939	0.941	0.709	0.806	0.825	0.697	0.769	0.780
150	0.956	0.966	0.967	0.843	0.895	0.906	0.789	0.839	0.847
300	0.981	0.985	0.985	0.932	0.954	0.959	0.875	0.905	0.909
500	0.988	0.991	0.991	0.959	0.972	0.975	0.911	0.932	0.935

Table 13.3 Ratio of Var [$\hat{q}(T)$] of POT/PD-GP and AMS/PD-GP Models when mean of the POT is 3.

T in years	R_2 for κ= 0.2 for different variances			R_2 for κ= 0.1 for different variances			R_2 for κ= -0.2 for different variances		
	3.5	4.0	4.5	3.5	4.0	4.5	3.5	4.0	4.5
50	0.609	0.784	0.824	-0.446	0.263	0.428	-	0.311	0.447
100	0.864	0.925	0.939	0.554	0.773	0.824	0.342	0.641	0.712
150	0.918	0.954	0.963	0.732	0.864	0.894	0.522	0.739	0.791
300	0.961	0.978	0.982	0.870	0.934	0.949	0.702	0.837	0.869
500	0.976	0.987	0.989	0.918	0.958	0.967	0.780	0.880	0.904

CHAPTER-14
DEVELOPMENT OF REGIONAL FLOOD FREQUENCY RELATIONSHIPS

14.0 INTRODUCTION

Information on flood magnitudes and their frequencies is needed for design of various types of water resources projects/ hydraulic structures such as dams, spillways, road and railway bridges, culverts, urban drainage systems as well as for taking up various non-structural measures such as flood plain zoning, economic evaluation of flood protection projects etc. Since scientific hydrology began in the seventeenth century, one of the most difficult problems facing engineers and hydrologists is how to predict flow in basins with no records. Whenever rainfall or river flow records are not available at or near the site of interest, it is difficult for hydrologists or engineers to derive reliable design flood estimates directly. In such a situation, regional flood frequency relationships developed for the region are one of the alternative methods for prediction of design floods, especially for small to medium size catchments.

The approaches for design flood estimation may be broadly categorized as: (i) deterministic approach using design storm, and (ii) probabilistic approach involving flood frequency analysis. The deterministic and probabilistic methods, which have been used for design flood estimation, are: empirical methods, rational method, flood frequency analysis methods, unit hydrograph techniques, and watershed models. Pilgrim and Cordery (1992) mention that estimation of peak flows on small to medium-sized rural drainage basins is probably the most common application of flood estimation as well as being of greatest overall economic importance. In almost all cases, no observed data are available at the design site, and little time can be spent on the estimate, precluding use of other data in the region. The author's further state that hundreds of different methods have been used for estimating floods on small drainage basins, most involving arbitrary formulas. The three most widely used types of

methods are the rational method, the U.S. Soil Conservation Service method and regional flood frequency methods. Regional flood frequency analysis resolves the problem of short data records or unavailability of data by "trading space for time"; as the data from several sites are used in estimating flood frequencies at any site. The choice of method primarily depends on design criteria applicable to the structure and availability of data.

Considering the importance of prediction in ungauged catchments, the International Association of Hydrological Sciences (IAHS) launched "Prediction of Ungauged Basins (PUBs)" as one of its initiatives and declared the current decade as "Decade of PUBs". As per the Bureau of Indian Standards (BIS) hydrologic design criteria, frequency based floods find their applications in estimation of design floods for almost all the types of hydraulic structures viz. small size dams, barrages, weirs, road and railway bridges, cross drainage structures, flood control structures etc., excluding large and intermediate size dams. For design of large and intermediate size dams probable maximum flood and standard project flood are adopted, respectively. Most of the small size catchments are ungauged or sparsely gauged. To overcome the problems of prediction of floods of various return periods for ungauged and sparsely gauged catchments, a robust procedure of regional flood frequency estimation is required to be developed.

In this study, regional flood frequency relationships are developed based on the L-moments approach for the gauged and ungauged catchments of Mahanadi Subzone 3(d) in India. For this purpose, various frequency distributions namely Extreme value (EV1), General extreme value (GEV), Generalized logistic (GL), Generalized normal (GNO), Pearson Type-III (PT3), Log Pearson Type –III(LP3) , Exponential (EXP), Generalized Pareto (GP), Kappa (KAP) and five parameter Wakeby (WAK) are employed. Regional flood frequency relationship is

developed for gauged catchments based on the robust identified frequency distribution. This relationship is coupled with the regional relationship between mean annual peak flood and catchment area and a simple regional flood frequency relationship is also developed for ungauged catchments of the study area. The study presents regional flood frequency relationships for gauged catchments and in particular very simple and robust regional flood frequency relationships for ungauged catchments of the study area.

14.1 REGIONAL FLOOD FREQUENCY ANALYSIS

Cunnane (1988) mentions twelve different regional flood frequency analysis (RFFA) methods. Out of these methods, some of the commonly used methods are: (i) Dalrymple's Index Flood method, (ii) Natural Environmental Research Council (N.E.R.C.) method, (iii) United States Water Resources Council (USWRC) method, (iv) Bayesian method and (v) Regional Regression based methods.

Based on data availability and record length of the data the following three types of approaches may be adopted for developing the flood frequency relationships: (a) at-site flood frequency analysis, (b) at-site and regional flood frequency analysis, and (c) regional flood frequency analysis. The steps involved in carrying out regional flood frequency analysis based on the above approaches are mentioned below.

(i) Identify a hydro meteorologically homogeneous region.

(ii) Screen the observed annual maximum peak flood data of the stream flow gauging sites of the homogeneous region and test the regional homogeneity.

(iii) Develop flood frequency relationships for the region considering various frequency distributions.

(iv) Select the best fit distribution based on the goodness of fit criteria.

(v) Estimate the at-site mean annual peak flood.

(vi) Use the best fit regional flood frequency relationship for estimation of T-year flood.

Steps (i) to (iv) mentioned as above remain unchanged here. Subsequently, the following steps are followed:

(i) Develop a regional relationship between mean annual peak flood and catchment and physiographic characteristics for the region.

(ii) Estimate the mean annual peak flood using the developed relationship.

(iii) Use the best fit regional flood frequency relationship for estimation of T-year flood.

Zhang and Singh (2007) mention that using the Gumbel–Hougaard copula, trivariate distributions of flood peak, volume, and duration were derived, and then conditional return periods were obtained. The derived distributions were tested using flood data from the Amite River Basin in Louisiana. The authors state that a major advantage of the copula method is that marginal distributions of individual variables can be of any form and the variables can be correlated. Griffs and Stedinger (2007) presented evolution of flood frequency analysis with *Bulletin 17*. The current methodology recommended for flood-frequency analyses by U.S. Federal agencies is presented in *Bulletin 17B*. *Bulletin 17* was first published in 1976; minor corrections were made in 1977 resulting in *Bulletin 17A*, which was later succeeded by *Bulletin 17B* published in 1982. The authors mention that the fields of hydrology and flood frequency analysis have substantially evolved since *Bulletin 17* was first published. New techniques are now available which should become part of these standard procedures. A comparison is provided which demonstrates how the standard and weighted *Bulletin 17B* quantile estimators perform relative to alternative LP3 quantile estimators that also make use of regional information.

Chebana and Quarda (2007) presented a multivariate L-moments homogeneity test with the aim to extend the statistical homogeneity test of Hosking and Wallis (1993) to the multivariate case. The usefulness of the methodology is illustrated on flood events. Monte-Carlo simulations are also performed for a bivariate Gumbel logistic model with Gumbel marginal distributions. Results illustrate the power of the proposed multivariate *L*-moment homogeneity test to detect heterogeneity on the whole structure of the model and on the marginal distributions. In a bivariate flood setting, a comparison is carried out with the classical homogeneity test of Hosking and Wallis based on several types of regions. In the present study, regional flood frequency relationships have been developed based on the L-moments approach for estimation of floods of various return periods for gauged and ungauged catchments of the Mahanadi Subzone 3 (d) of India.

14.2 L-MOMENTS APPROACH

Some of the commonly used parameter estimation methods for most of the frequency distributions include: (i) method of least squares, (ii) method of moments, (iii) method of maximum likelihood, (iv) method of probability weighted moments, (v) method based on principle of maximum entropy, and (vi) method based on L-moments. L-moments are a recent development within statistics (Hosking, 1990). In a wide range of hydrologic applications, L-moments provide simple and reasonably efficient estimators of characteristics of hydrologic data and of a distribution's parameters (Stedinger et al., 1992). Like the ordinary product moments, L-moments summarize the characteristics or shapes of theoretical probability distributions and observed samples. Both moment types offer measures of distributional location (mean), scale (variance), skewness (shape), and kurtosis (peakedness).

Recently a number of regional flood frequency analysis studies have been carried out based on the L-moments approach. The L-moment methods are demonstrably superior to those that have been used previously, and are now being adopted by many organizations worldwide (Hosking and Wallis, 1997). The L-moments offer significant advantages over ordinary product moments, especially for environmental data sets, because of the following (Zafirakou-Koulouris et al., 1998).

i. L-moment ratio estimators of location, scale and shape are nearly unbiased, regardless of the probability distribution from which the observations arise (Hosking, 1990).

ii. L-moment ratio estimators such as L-coefficient of variation, L-skewness, and L-kurtosis can exhibit lower bias than conventional product moment ratios, especially for highly skewed samples.

iii. The L-moment ratio estimators of L- coefficient of variation and L-skewness do not have bounds which depend on sample size as do the ordinary product moment ratio estimators of coefficient of variation and skewness.

iv. L-moment estimators are linear combinations of the observations and thus are less sensitive to the largest observations in a sample than product moment estimators, which square or cube the observations.

v. L-moment ratio diagrams are particularly good at identifying the distributional properties of highly skewed data, whereas ordinary product moment diagrams are almost useless for this task (Vogel and Fennessey, 1993).

The linear functions of PWMs, the first four L-moments and L-coefficient of variation (L-CV denoted as τ_2), L-coefficient of skewness (L-skew denoted as τ_3) and L-coefficient of kurtosis, L-kurtosis denoted as τ_4) are earlier discussed in Chapter-10 at section 10.3.5.

14.3 SCREENING OF DATA USING DISCORDANCY MEASURE TEST

The objective of screening of data is to check that the data are appropriate for performing the regional flood frequency analysis. In this study, screening of the data was performed using the L-moments based Discordancy measure (D_i). Discordancy is measured in terms of the L-moments of the sites' data and the aim is to identify those sites that are grossly discordant with the group as a whole. The sample L-moment ratios (t_2, t_3 and t_4) of a site are considered as a point in a three-dimensional space. A group of sites form a cluster of such points in the three-dimensional space. A site is considered discordant if it is far from the centre of the cluster. Hosking and Wallis (1997) defined the Discordancy measure D_i for a site i in a group of N sites.

$$\bar{u} = N^{-1} \sum_{i=1}^{N} u_i \tag{14.1}$$

Let $u_i = [t_2^{(i)} \;\; t_3^{(i)} \;\; t_4^{(i)}]^T$ be a vector containing the sample L-moment ratios t_2, t_3 and t_4 values for site i, analogous to their regional values termed as τ_2, τ_3, and τ_4, expressed in Eqs. (10.29 and 10.30). T denotes transposition of a vector or matrix. Let

$$A_m = \sum_{i=1}^{N} (u_i - \bar{u})(u_i - \bar{u})^T \tag{14.2}$$

be the (un weighted) group average. The Discordancy (D_i) measure for site i is defined as:

$$D_i = \frac{1}{3} N (u_i - \bar{u})^T A_m^{-1} (u_i - \bar{u}) \tag{14.3}$$

The site i is declared to be discordant, if D_i is greater than the critical value of the Discordancy statistic D_i, given in a tabular form by Hosking and Wallis (1997).

14.3.1 Test of Regional Homogeneity

For testing regional homogeneity, a test statistic H, termed as heterogeneity measure was proposed by Hosking and Wallis (1993). It compares the "inter-site variations in sample L-moments for the group of sites" with "what would be expected of a homogeneous region". The inter-site variations in sample L-moments are evaluated based on any of the three measures of variability V_1 (based on L-CV), V_2 (based on L-CV and L-Skew) and V_3 (based on L- Skew and L-Kurtosis). These measures of variability are computed as follows:

(i) V_1 is the weighted standard deviation of at site L-CV's ($t_2^{(i)}$)

$$V_1 = \left[\sum_{i=1}^{N} n_i \ (t_2^{(i)} - t_2^R)^2 \Big/ \sum_{i=1}^{N} n_i \right]^{1/2} \qquad (14.4)$$

where, n_i is the record length at each site and t_2^R is the regional average L-CV weighted proportionally to the sites' record length as given below.

$$t_2^R = \sum_{i=1}^{N} n_i \ t_2^{(i)} \Big/ \sum_{i=1}^{N} n_i \qquad (14.5)$$

(ii) V_2 is the weighted average distance from the site to the group weighted mean on a graph of t_2 versus t_3

$$V_2 = \sum_{i=1}^{N} n_i \left\{(t_2^{(i)} - t_2^R)^2 + (t_3^{(i)} - t_3^R)^2\right\}^{1/2} \Big/ \sum_{i=1}^{N} n_i \qquad (14.6)$$

where, t_3^R is the regional average L-Skew weighted proportionally to the sites' record length.

(iii) V_3 is the weighted average distance from the site to the group weighted mean on a graph of t_3 versus t_4

$$V_3 = \sum_{i=1}^{N} n_i \left\{(t_3^{(i)} - t_3^R)^2 + (t_4^{(i)} - t_4^R)^2\right\}^{1/2} \Big/ \sum_{i=1}^{N} n_i \qquad (14.7)$$

Where, t_4^R is the regional average L-Kurtosis weighted proportionally to the sites' record length.

To establish "what would be expected of a homogeneous region", firstly simulations are used to generate homogeneous regions with sites having same record lengths as those of observed data. In order to generate the simulated data, a four parameter Kappa distribution is used. The four parameter Kappa distribution is chosen so as not to commit to a particular two or three parameter distribution. Further, the four parameter Kappa distribution includes as special cases the generalised logistic, generalised extreme value and generalised pareto distributions and hence, acts as a good representation of many of the probability distributions occurring in environmental sciences.

The parameters of the Kappa distribution are obtained using the regional average L-moment ratios t_2^R, t_3^R, t_4^R and mean = 1. A large number of data regions are generated (say N_{sim} = 500) based on this Kappa distribution. The simulated regions are homogeneous and have no cross-correlation or serial correlation. Further, the sites have the same record lengths as the

observed data. For each generated region, V_j (i.e. any of V_1, V_2 or V_3) is computed using Eqs. 14.4 to 14.7. Subsequently, their mean (μ_v) and standard deviation (σ_v) are calculated, and the heterogeneity measure H(j) (i.e. H(1), H(2) or H(3)) is computed as:

$$H(j) = \frac{V_j - \mu_v}{\sigma_v} \qquad (14.8)$$

If the heterogeneity measure is sufficiently large, the region is declared to be heterogeneous. Hosking and Wallis (1997) suggested the following criteria for assessing heterogeneity of a region: if $H(j) < 1$, the region is acceptably homogeneous; if $1 \leq H(j) < 2$, the region is possibly heterogeneous; and if $H(j) \geq 2$, the region is definitely heterogeneous. These boundary values of H(j) being 1 and 2 were determined by Hosking and Wallis (1997) by performing a series of Monte Carlo experiments in which the accuracy of quantile estimates corresponding to different values of H(j) were computed. The authors further observed that for both real world data and artificially simulated regions, H(1) has much better power to discriminate between homogeneous and heterogeneous regions as compared to H(2) and H(3).

14.3.2 Identification of Robust Regional Frequency Distribution

The best fit frequency distribution for a homogeneous region is determined by how well the L-skewness and L-kurtosis of the fitted distribution match the regional average L-skewness and L-kurtosis of the observed data (Hosking and Wallis, 1997). The procedure adopted (Hosking and Wallis, 1997) is briefly stated below.

Initially, several three parameter distributions are fitted to the regional average L-moments t_2^R, t_3^R and mean =1. Let τ_4^{Dist} be the L-kurtosis of the fitted distribution which may

be GEV, GL, GNO, and PT3 etc. Using the N_{sim} number of simulated regions of the Kappa distribution (as obtained for the heterogeneity measure described in section 14.3.1 above), the regional average L-kurtosis, t_4^m is computed for the m^{th} simulated region. The bias of t_4^R is computed as:

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} \left(t_4^m - t_4^R \right) \qquad (14.9)$$

The standard deviation of t_4^R is computed as:

$$\sigma_4 = \left[\left(N_{sim} - 1 \right)^{-1} \left\{ \sum_{m=1}^{N_{sim}} \left(t_4^m - t_4^R \right)^2 - N_{sim} B_4^2 \right\} \right]^{1/2} \qquad (14.10)$$

The goodness-of-fit measure for each distribution is computed as (Hosking and Wallis, 1997):

$$Z^{dist} = \left(\tau_4^{dist} - t_4^R + B_4 \right) / \sigma_4 \qquad (14.11)$$

The fit is considered to be adequate if $|Z^{dist}|$–statistic is sufficiently close to zero, a reasonable criterion being $|Z^{dist}|$ –statistic less than 1.64. Hosking and Wallis (1997) states that the $|Z^{dist}|$ –statistic has the form of a normal distribution under suitable assumptions. Thus the criterion $|Z^{dist}|$ –statistic less than 1.64 corresponds to acceptance of the hypothesized distribution at a confidence level of 90%.

14.4 STUDY AREA AND DATA AVAILABILITY

Mahanadi Subzone 3(d) in India comprises of Mahanadi, Brahmani and Baitarani basins (being regularly maintained by Director Hydrometry, Govt. of Orissa and CWC GD sites). It is located between longitudes of 80^0 25' to 87^0 east and latitudes 19^0 15' to 23^0 35' north. Its total drainage area is about 195 256 km^2. About 50% of the area of this Subzone is hilly varying from 300 m to 1350 m. Rest of the area lies in the elevation range of 0 to 300 m. The normal annual rainfall over the region varies from 1200 to 1600 mm. The Subzone receives about 75% to 80% of the annual rainfall from South-West monsoon during the monsoon season from June to September. The red and yellow soils cover major part of the Subzone. The red sandy, sandstone and coastal alluvial soils cover the remaining part of the Subzone. The Subzone has an extensive area under forest. Paddy is the main crop grown on the cultivable land.

The annual maximum or peak series of 23 Bridge sites lying in the hydro meteorologically homogeneous region) of Mahanadi Subzone 3 (d) is taken from CWC (1982) and discussed at section 10.7.2 in Chapter-10 (refer Appendix -1 and Table 14.1a). The data period was from 1957 to 1990. These data were observed during the period 1957 to 1990 with record lengths varying from 11 to 30 years. The data of 13 GD site were taken between a time intervals of 11 to 38 years as shown in table 14.1(b) is averagely have the duration of 1972-2009, and also discussed in Chapter-10 (Appendix-2).

14.5 ANALYSIS AND DISCUSSION OF RESULTS

Regional flood frequency analysis was performed using the various frequency distributions: viz. Extreme value (EV1), Generalized extreme value (GEV), Generalized logistic (GL), Generalized normal (GNO), Pearson Type-III (PT3), Log Pearson Type –III (LP3), Generalized Pareto (GP), and five parameter Wakeby (WAK). Screening of the data, testing of

regional homogeneity, identification of the regional distribution and development of regional flood frequency relationships are described below.

14.5.1 Screening of Data Using Discordancy Measure Test

The site discordancy (D_i) are computed for the study area following the procedure given by (Hosking and Wallis, 1997). Record lengths (Table 2.2a), catchment area, *L*-moment ratios, of six sites with high discordancy are given in Table 14.1 (a) and (b) for twenty three Br-catchment sites and thirteen GD sites. Values of discordancy statistic have been computed in terms of the L-moments for all the 23 gauging sites of the study area. It is observed that the D_i values for all the 23 sites vary from 0.04 to 3.06. As 3.06 is greater than the critical value of D_i i.e. 3.00 for a region defined by 15 or more sites (Hosking and Wallis, 1997); hence, the site having D_i value of 3.06 is discarded from the analysis. The D_i values for 22 sites vary from 0.04 to 2.32 and the data of 22 sites may be considered suitable for regional flood frequency analysis.

The critical value (=3) is not exceeded by any one site: site out of thirty six total catchments that was tested. Br-catchments No: 176 and 325 which has a higher value of D_i i.e. above 2 has the highest value of t, and t_3 with a relatively moderate t_4 discordant with the pattern found for other catchments in the region. Figure 14.1 (a)-(e) shows the *L*-moment diagrams (*L-CV – L-SK* and *L-SK-L-KR* plots) for the data for both 23 Br-catchments and 13 cGD sites separately, and is used for visual assessment of the dispersion in data. For this site (site 325) the deviation of the *L*-moment ratios from the group average is not in direction concordant with the deviation of the other sites both in *L-CV – L-SK* and *L-SK-L-KR* plots. Additionally, from Table 14.1a, it is observed that the mean at this site is comparatively low compared with other sites in the region with comparable catchment areas. The point that lies in the far right hand corner of the graph (Figure 14.1b) has a very high t_3, and t_4 and looks to be discordant, it has a D_I-value

of 2.101 (Br-176) however, the value D_i is less than the critical value of 3). For 13 GD sites the values of L-moment statistics and discordancy statistic (Table 14.1 b) varies from 0.143 that is for *Jariakela* to a highest of 1.481 that is computed for *Sukuma*.

Table 14.1(a) Catchment area, sample statistics, sample size and discordancy measure for 23 Bridge catchment of Mahanadi Subzone 3(d)

Stream Gauging Site	Catchment Area (km^2)	Mean Annual Peak Flood (m^3/s)	Sample Size (Years)	L-CV (τ_2)	L-skew (τ_3)	L-kurtosis (τ_4)	Discordancy Measure (D_i)
66k	154	260.32	28	0.386	0.437	0.224	0.401
48	109	103.90	30	0.406	0.293	0.160	0.410
176	66	81.48	31	0.512	0.525	0.444	**2.101**
12	666	509.97	30	0.195	0.114	0.129	1.243
93K	74	153.07	28	0.273	0.106	0.197	1.560
59KGP	30	75.33	27	0.410	0.260	0.165	0.589
308	19	41.22	27	0.348	0.227	0.061	0.847
332NGP	225	188.59	22	0.289	0.202	0.199	0.461
59BSP	136	196.23	22	0.412	0.352	0.237	0.223
698	113	247.00	25	0.429	0.321	0.120	0.933
37	64	25.09	23	0.434	0.366	0.088	1.808
325	26	50.00	19	0.482	0.234	0.102	**2.102**
121	1150	1003.86	21	0.267	0.146	0.031	1.054
385	194	115.40	20	0.355	0.083	0.085	1.375
332KGP	175	71.83	24	0.311	0.140	0.158	0.771
40K	115	260.67	21	0.349	0.224	0.175	0.309
154	58	160.16	19	0.432	0.401	0.349	0.916
489	823	1071.95	21	0.517	0.528	0.311	1.108
42	49	53.50	20	0.222	0.013	-0.01	1.464
69	173	238.00	18	0.360	0.246	0.114	0.399
90	190	130.73	11	0.363	0.118	0.124	1.248
195	615	963.77	13	0.234	0.101	0.123	0.924
235	312	176.14	14	0.314	0.209	0.065	0.745

Table 14.1(b) Catchment area, sample statistics and discordancy measure for 13 GD sites of Mahanadi Sub zone.

Stream Gauging Site	Catchment Area (km^2)	L(Length of stream)	L-Ratio	L-CV (τ_2)	L-skew (τ_3)	L-kurtosis (τ_4)	Qav	Discordancy Measure (D_i
Altuma	830	265	1182	0.250	0.085	0.133	486.10	1.046
Anandpu	8570	78	1.40861	0.317	0.110	0.116	3781.4	0.699
Champua	1710	76	0.29605	0.274	0.112	0.154	620.69	0.973
Gomlai	21950	52	8.1176	0.279	0.071	0.112	5672.8	0.978
Jaraikela	9160	48	3.97569	0.374	0.336	0.235	3471.5	0.143
Kantamal	19600	115	1.48204	0.228	0.077	0.016	6130.6	0.909
Jenapur	33955	278	0.43935	0.363	0.066	-0.016	8284.5	1.171
Kesinga	11960	94	1.35355	0.458	0.248	0.088	6991.7	1.050
Pamposh	19448	89	2.45525	0.250	0.106	-0.018	6863.8	1.270
Salebhata	4650	49	1.93669	0.456	0.367	0.315	2444.7	1.409
Sukuma	1365	37	0.99708	0.521	0.306	0.172	709.14	1.481
Sunderga	5870	142	0.29111	0.410	0.482	0.352	2277.3	0.603
Talcher	29750	174	0.98263	0.186	0.328	0.291	4879.5	1.263

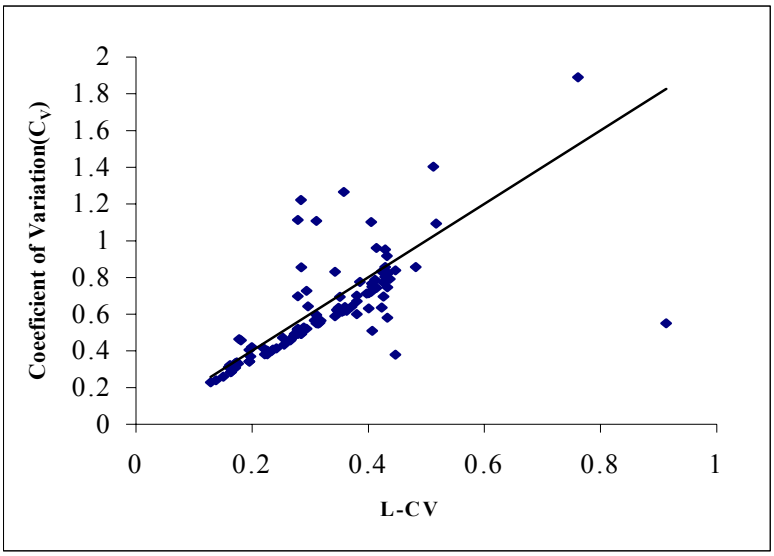

Figure 14.1a L-CV (τ) –Coefficient of variation (C_v) diagram for total 36 i.e. 23 and 13 catchments.

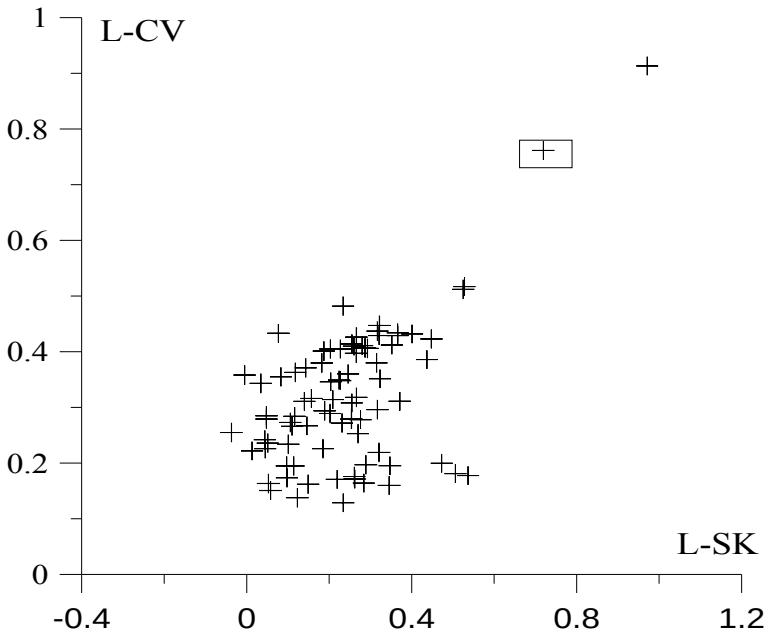

Figure 14.1b *L*-Moment ratio diagram for the study area (total 36 i.e. 23 and 13 catchments).

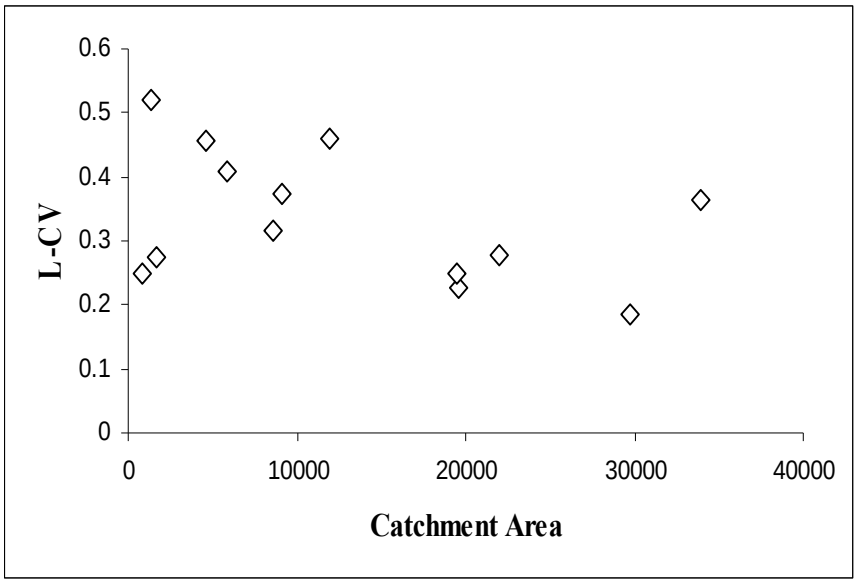

Figure 14.1c Graph between Catchment Area and L-CV for 13 GD sites of Mahanadi Sub zone.

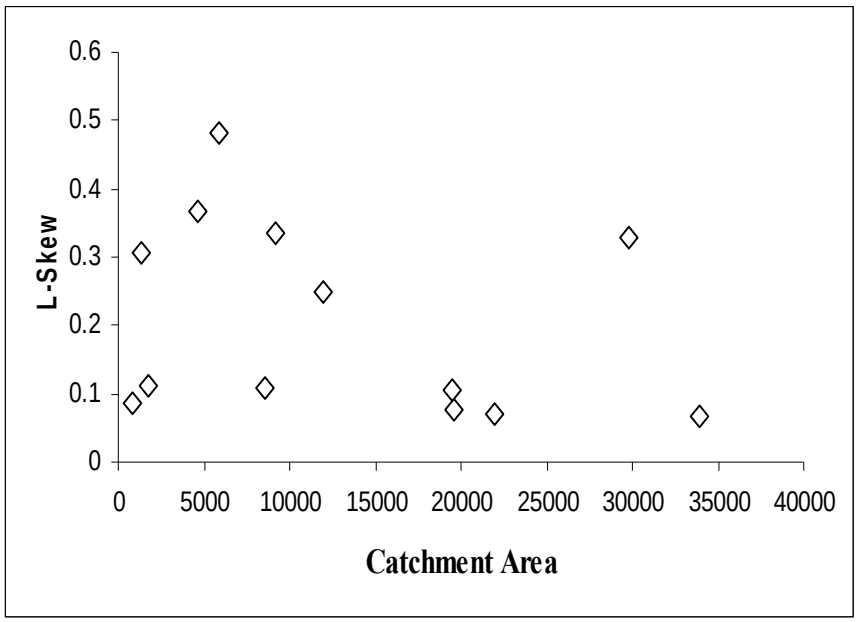

Figure 14.1d Graph between Catchment Area and L-Skew for 13 GD sites of Mahanadi Sub zone.

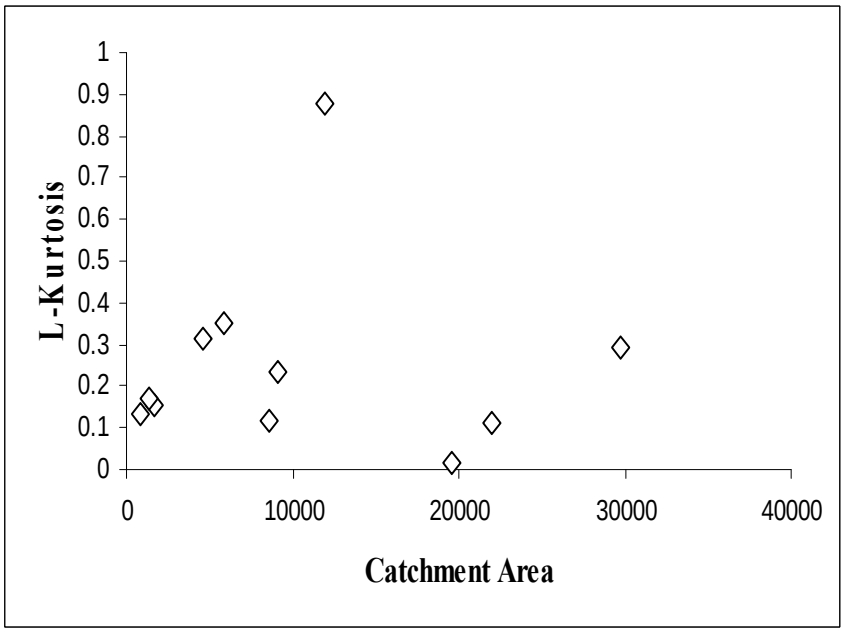

Figure 14.1 e Graph between Catchment Area and L-Kurtosis for 13 GD sites of Mahanadi Sub zone.

14.5.2 Test of Regional Homogeneity

Table 14.2 provides a summary of the range of annual peak flows and corresponding catchment area used for the present study. The coefficient of variation (C_v) and skewness (γ) for the annual peak flow data (Figure 14.1a) are observed to be varying in the range of 0.23 to 1.9 and –0.48 to 4.4 respectively. The variation of C_v means high dispersion of annual peak flood data in the region, and this can be judged from the deviation of the points from the straight line fit in the $\tau - C_v$ diagram given in Figure 14.1b. Four catchments on the $\tau - C_v$ are far away from the best-fit line ($C_v = 2\tau$) indicating heavy tailed distribution or highly skewed data, however, most of the catchments near the line indicate that the data follows a moderately skewed distribution. Hosking and Wallis (1997) provided an approximate relation between by γ and τ_3 to infer the symmetry of annual peak flood distribution, and for a near symmetric distribution γ is close to six times τ_3. A comparison of the values computed with skewness measure and *L*-skewness for the data of the region using $\gamma - \tau_3$ diagram (Figure 14.1a) shows that most of the stations lie on the best-fit line given by $\gamma = 4.2\ \tau_3$, and 15 stations are outside the range of best-fit line given by $\gamma = 6\ \tau_3$. Analysis gives a preliminary idea about the statistical properties of the annual peak flood data considered for the study; nearly 90% of the stations follow a symmetric distribution with moderate skewness.

The values of the heterogeneity measures H(1), H(2) and H(3) were computed utilizing the data of 36 (23 Br-catchments and 13 GD sites) gauging sites by generating 500 regions using the fitted Kappa distribution**.** Using the data of 36 sites, H(1), H(2) and H(3) values are computed as 4.21, 1.79 and 0.16, respectively. As H(1) is greater than 2.0 and H(2) is greater than 1.0, the region defined by the 22 gauging sites is considered as heterogeneous as given in Table 14.3(a). Thus, based on the statistical properties of the data of the gauging sites, one by one seven sites of the region are excluded till H(1) value between 1.0 and 2.0; and both H(2) and

H(3) values less than 1.0 is obtained. Further efforts to reduce the value of H(1) led to significant loss of data and hence, this

Table.14.2 Statistical properties of data used

Regions	No of Catchments	Mean annual peak flood ($m^3/s.$)			Catchment area (Km^2)		Length of main stream (Km.)	
		Mean	Max.	Min	Max.	Min	Max.	Min.
3D Sub zone	23	734	4550	72	1150	19	80.5	9.8
13 GD sites in Mahandi region	13	4047	8284	486	33955	830	278	37

Table 14.3(a) Heterogeneity measures for 36 gauging sites of Mahanadi basin

Sl. No.	Heterogeneity measures	Values
1.	Heterogeneity measure H(1)	
	(a) Observed standard deviation of group L-CV	0.0615
	(b) Simulated mean of standard deviation of group L-CV	0.0472
	(c) Simulated standard deviation of standard deviation of group L-CV	0.0085
	(d) Standardized test value H(1)	1.68
2.	Heterogeneity measure H(2)	
	(a) Observed average of L-CV / L-Skewness distance	0.0861
	(b) Simulated mean of average L-CV / L-Skewness distance	0.0980
	(c) Simulated standard deviation of average L-CV / L-Skewness	0.0167
	(d) Standardized test value H(2)	-0.71
3.	Heterogeneity measure H(3)	
	(a) Observed average of L-Skewness/L-Kurtosis distance	0.0821
	(b) Simulated mean of average L-Skewness/L-Kurtosis distance	0.1198
	(c) Simulated standard deviation of average L-Skewness/L-Kurtosis	0.0191
	(d) Standardized test value H(3)	-1.98

was not attempted. In a separate data sample comprising of 14 gauging sites and yielding H(1), H(2) and H(3) values as 2.27, 0.93 and -0.2 is considered as reasonably homogeneous. The details of catchment data and statistical parameters including the discordancy measure, for the 14 gauging sites are given in Table 14.2(b). The values of heterogeneity measures computed by carrying out 500 simulations using the Kappa distribution based on the data of 14 sites are given in Table 14.2(b).

Table 14.3(b). Discordancy, Outlier-test, and Heterogeneity Measures of 14 GD sites

Region	Range of Discordancy (D) (2)	Results of outlier test (Maidment,1993) (3)	Heterogeneity measure (H_I)					
			All stations			Discordant stations removed		
			H_1 (4)	H_2 (5)	H_3 (6)	H_1 (7)	H_2 (8)	H_3 (9)
3D Sub zone	0.22-2.01(3)	No outliers	2.27[c]	0.93[a]	-0.2[a]	1.6[b]	0.27[a]	-0.7[a]

* Numbers in parenthesis denote the number of basins in the region that are discordant with the rest, [a] homogeneous ($|H_i|<1$), [b] Possibly heterogeneous ($1<|H_i|<2$), [c] Heterogeneous ($|H_i|>2$)

14.5.3 Identification of Robust Regional Frequency Distribution

The L-moment ratio diagram and $|Z_i^{dist}|$ –statistic are used as the best fit criteria for identifying the robust distribution for the study area. The regional average values of L-skewness i.e. $\tau_3 = 0.2180$ and L-kurtosis i.e. $\tau_4 = 0.1510$ are obtained. the L-moments ratio diagram for the study area. The Z_i^{dist} –statistic for various three parameter distributions is given in Table 14.4. It is observed that the $|Z_i^{dist}|$ –statistic values are lower than 1.64 for the three distributions viz. GEV, GNO and PT3. Further, the $|Z_i^{dist}|$ –statistic is found to be the lowest for GNO distribution i.e. 0.22.

Table 14.4 Z_i^{dist} statistic for various distributions for Mahanadi Subzone 3 (d)

S. No.	Distribution	Z_i^{dist} –statistic
1	Generalized logistic (GL)	2.08
2	Generalized Extreme Value (GEV)	0.66
3	Generalized Normal (GNO)	0.22
4	Pearson Type III (PT3)	-0.62
5	Generalized Pareto (GP)	-2.68
6	Wakeby (WKB) distribution	1.21
7	Log Pearson Type 3 Distribution(LP3)	-2.45

Thus, based on the L-moment ratio diagram and $|Z_i^{dist}|$ –statistic criteria, the GNO distribution is identified as the robust distribution for the study area .

The values of regional parameters for the various distributions which have Z^{dist} –statistic value less than 1.64 (i.e. distributions accepted at the 90% confidence level) as well as the five parameter Wakeby distribution are given in Table 14.5. The regional parameters of the Wakeby distribution have been included in Table 14.4 because, the Wakeby distribution has five parameters, more than most of the common distributions and it can attain a wider range of distributional shapes than can the common distributions. This makes the Wakeby distribution particularly useful for simulating artificial data for use in studying the robustness, under changes in distributional form of methods of data analysis. It is preferred to use Wakeby distribution for heterogeneous regions.

Table 14.5 Regional parameters for various distributions for Mahanadi Subzone 3 (d)

Distribution	Parameters of the Distribution				
GNO	$\xi = 0.870$	$\alpha = 0.548$	$k = -0.451$		
PT3	$\mu = 1.000$	$\sigma = 0.629$	$\gamma = 1.316$		
GEV	$\xi = 0.704$	$\alpha = 0.452$	$K = -0.073$		
WAK	$\xi = 0.100$	$\alpha = 1.985$	$\beta = 6.486$	$\gamma = 0.684$	$\delta = -0.078$
GP	-	$\alpha = 0.812$	$k=1.436$	-	-
GL	$u=0.645$	$\alpha = 1.115$	$k=1.142$	-	-
LP3					

14.5.4 Regional Flood Frequency Relationship for Gauged Catchments

For estimation of floods of various return periods for gauged catchments regional flood frequency relationship has been developed based on the robust identified GNO distribution. The cumulative density function of the three parameter GNO distribution as parameterized by Hosking and Wallis (1997) is given below.

$$F(x) = \phi\left[-k^{-1}\log\{1-k(x-\xi)/\alpha\}\right] \qquad k \neq 0$$
$$= \phi\{(x-\xi)/\alpha\} \qquad k = 0 \qquad (14.12)$$

where, ξ, α and k are its location, scale and shape parameters respectively. When k = 0, it becomes normal distribution (NOR) with parameters ξ and α. This distribution has no explicit analytical inverse form. ϕ is the cumulative distribution function of the standard Normal distribution given by

$$\phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right) dx \qquad (14.13)$$

Floods of various return periods may be computed by multiplying mean annual peak flood of a catchment by the corresponding values of growth factors of GNO distribution given in Table 14.6. The growth factor or site-specific scale factor (Q_T / $\overline{Q}$) is computed by dividing flood quantile (Q_T) by the annual mean peak flood of a gauging site ($\overline{Q}$).

Table 14.6 Values of growth factors ($Q_T/\overline{Q}$) for Mahanadi Subzone 3 (d)

Distribution	Return period (Years)						
	2	10	25	50	100	200	1000
	Growth factors						
GNO	0.8	1.821	2.331	2.723	3.125	3.538	4.552
PT3	0.8	1.842	2.329	2.683	3.028	3.366	4.134
GEV	0.8	1.809	2.332	2.745	3.175	3.627	4.767
WAK	0.8	1.848	2.353	2.712	3.052	3.374	4.058
LP3	0.8	1.765	2.301	2.765	3.056	3.278	4.009
GP	0.8	1.654	2.846	2.669	3.201	3.806	4.120
GL	0.9	1.980	2.970	2.786	3.223	3.866	4.923

Monte Carlo simulation is used to estimate the accuracy of the quantiles of the GNO distribution. The region used in the simulation procedure contains 13 sites with record lengths

same as that for Mahanadi Subzone 3(d) data, the sites having GNO distributions with L-CV varying linearly from 0.2260 at site 1 to 0.4240 at site 15 and L-skewness 0.2180. The inter-site dependence is quantified in terms of an average correlation of about 0.32. 10,000 regions were generated and a GNO distribution was fitted to each generated region. The regional average relative RMSE of the estimated growth curve was computed from the simulations. The 90% error bounds for the growth curves are also computed are presented in Table 14.7. It can be seen from the above graph that the growth factor increases proportionally with the return period constantly and the lower error bounds is between the area covered by the growth factor and the return period similarly the upper error bound is outside this area.

14.5.5 Regional Flood Frequency Relationship for Ungauged Catchments

For ungauged catchments the at-site mean cannot be computed in absence of the observed flow data. Hence, a relationship between the mean annual peak flood of gauged catchments in the region and their pertinent physiographic and climatic characteristics is needed for estimation of the mean annual peak flood. Table 14.1 9a) and (b) shows the values of the mean annual peak flood versus catchment area for the 23 Br-catchments and 13 GD sites of Mahanadi basin. The regional relationship in the form of a power law ($Y = a X^b$) is developed for the region using Levenberg-Marquardt (LM) iteration on the data of the 36 gauging sites and is given below.

$$\overline{Q} = 2.519(A)^{0.863} \qquad (14.14)$$

Where, A is the catchment area, in km^2 and $\overline{Q}$ is the mean annual peak flood in m^3/s.

Table 14.7 Accuracy growth factors of GNO distribution for Mahanadi basin

Return period	Growth	RMSE	90% error bounds

(Years)	factors		Lower	Upper
2	0.870	0.035	0.829	0.889
10	1.821	0.076	1.789	1.905
25	2.331	0.101	2.248	2.517
50	2.723	0.116	2.587	3.008
100	3.125	0.128	2.924	3.526
200	3.538	0.140	3.267	4.075
1000	4.552	0.163	4.075	5.479

For Eq. (14.14), the coefficient of determination is, R^2 = 0.834 and the standard error of estimates is 127.7. For development of regional flood frequency relationship for ungauged catchments, the regional flood frequency relationship developed for gauged catchments is coupled with regional relationship between mean annual peak flood and catchment area, given in Eq. (14.14) and following regional frequency relationship is developed.

$$Q_T = C_T * A^{0.863} \qquad (14.15)$$

Where Q_T is the flood estimate in m^3/s for T year return period, and A is the catchment area in km^2 and C_T is a regional coefficient. Values of C_T for some of the commonly used return periods are given in Table 14.8.

Table 14.8 Values of regional coefficient C_T for Mahanadi Subzone 3(d)

Distribution	Return period (Years)						
	2	10	25	50	100	200	1000
	Growth factors						
GEV	0.8	1.80	2.332	2.745	3.175	3.627	4.767
GNO	2.1	4.58	5.872	6.859	7.872	8.912	11.466

The above regional flood formula, its tabular form (Table 14.8), or graphical representation may be used for estimation of floods of desired return periods for ungauged catchments of the Mahanadi basin. The conventional empirical flood formulae do not provide

floods of various return periods for the ungauged catchments. However, the regional flood formula developed in this study (Eq. 14.15) is capable of providing flood estimates for desired

Table 14.9 Floods of various return periods for different catchment areas of Mahanadi region

Catchment area (km^2)	Return period (years)						
	2	10	25	50	100	200	1000
	Floods of various return periods (m^3/s)						
10	16	33	43	50	57	65	84
20	29	61	78	91	104	118	152
30	41	86	111	129	148	168	216
40	53	111	142	166	190	215	277
50	64	134	172	201	230	261	335
60	75	157	201	235	270	305	393
70	86	179	230	268	308	349	448
80	96	201	258	301	346	391	503
90	107	223	285	333	382	433	557
100	117	244	312	365	419	474	610
150	166	346	443	518	594	673	866
200	212	444	568	664	762	863	1110
250	257	538	689	805	924	1046	1345
300	301	630	806	942	1081	1224	1575
350	344	720	921	1076	1235	1398	1799
400	386	807	1034	1207	1386	1569	2018
450	427	894	1144	1337	1534	1737	2234
500	468	979	1253	1464	1680	1902	2447
600	548	1146	1467	1713	1966	2226	2864
700	625	1309	1675	1957	2246	2543	3271
800	702	1469	1880	2196	2520	2853	3671
900	777	1626	2081	2431	2790	3159	4064
1000	851	1780	2279	2662	3056	3459	4451
1100	924	1933	2475	2891	3317	3756	4832
1200	996	2084	2668	3116	3576	4049	5209
1300	1067	2233	2858	3339	3832	4338	5581
1400	1137	2380	3047	3559	4085	4625	5950
1500	1207	2526	3234	3778	4336	4908	6315

return periods. The above table shows that the flood increases with the increase in return periods (years) and proportionally with the increase in catchment area. As you can see for return period

of 2 years the flood for a catchment area having area 10 km^2 is 16 m^3/s and for return period of 1000 years is 84 m^3/s similarly can be seen for 1500 km^2 catchment area

14.4 CONCLUSION

Screening of the data conducted using the annual maximum peak flood data of the Mahanadi Sub zone 3(d) employing the Ddiscordancy measure (D_i) test reveals that 22 out of the 23 Br-catchments (gauging sites) are suitable for regional flood frequency analysis. However, based on the heterogeneity measures, `H (j)'; the annual maximum peak flood data of other 13 GD sites are considered to constitute a homogeneous region. So, in a total, 36 catchments (with gauging sites) are used for regional flood frequency analysis. Various distributions viz. EV1, GEV, GL, PT3, LP3, GNO, GP, and WAK have been employed. Based on the L-moments ratio diagram and $|Z_i^{dist}|$ –statistic criteria, the GNO distribution has been identified as the robust distribution for the study area.

The developed regional flood frequency relationships may be used for estimation of floods of desired return periods for gauged and ungauged catchments of the study area. As the regional flood frequency relationships have been developed using the data of catchments varying from 19 to 1,150 km^2 in area; hence, these relationships are expected to provide estimates of floods of various return periods for catchments lying nearly in the same range of the areal extent. Further, the relationship between mean annual peak flood and catchment area is able to explain 83.4% of initial variance (R^2 = 0.834). Hence, in case of ungauged catchments the results of the study are subject to these limitations. However, the regional flood frequency relationships may be refined for obtaining more accurate flood frequency estimates; when the data for some more gauging sites become available and physiographic characteristics other than

catchment area as well as some of the pertinent climatic characteristics are also used for development of the regional flood frequency relationships.

SUMMARY AND CONCLUSION

For ungauged catchments, the need for developing a suitable methodology for realistic estimation of hydrological variables of interest is hardly over-emphasized. This study was taken up with the specific objective of design flood estimation in such catchments. Four different approaches were studied, and methodologies developed for both limited data and no data situations. The developed methods were checked for their workability taking two types of data i.e. short-term and annual maximum series data from real catchments. Thus, the work of this study is organized in six sections, it is accordingly summarized and concluded, and finally the limitations are discussed as follows:

Processing and Analysis of Hydro-Meteorological Data

The daily mean rainfall for the year 2008 is given in Table 3.1 in Chapter 3 and the mean monthly rainfall for few raingauges has been calculated using two methods: (i) Arithmetic mean (AM) and (ii) Thiessen Polygon (TP). In this study, three types of runoff data are used: (i) monthly runoff series, (ii) Annual maximum series (AMS), and (iii) short –term (hourly) data. The short term data, i.e., the hourly runoff data are used in this study are few and selected, and they have been used in unit hydrograph analysis. Referring in Chapter 3, the Appendix-1 shows the annual peak runoff of twenty five bridge catchments taken from CWC report and Appendix - 2 shows fourteen CWC gauging sites with daily flows and the period. Consistency of a rain gauge station and double mass curves analysis are used to cross check the significant change during the period of record and to check trends or non-homogeneity between flow records. The study area data has no such error. Both of these tests along with t- and F- distribution tests are done for the flow series that gives the crucial ratio between-group to the within-group variance estimates. From the flow data used in this study, it is noticed that between-group estimate of the population variance is sufficiently bigger than the within-group estimate of the population

variance which are always positive numbers. It means that the samples have been collected without bias. The performance of AR models are done for randomly selected paired stations i.e. for *Altuma and Anandpur, Kantamal and Salebhata, Pamposh and Champua* and *Gomlai and Jaraikela* and the results seen to be independent and normally distributed with a common variance. Lastly, residual flow series given in Annexure 3 is the simplest a means of identifying anomalies between stations is the plotting of comparative time series with respect to lag times from rainfall to runoff or the wave travel time in a channel. From the Remote Sensing data Imageries (Liss-III and Pan) which we taken from the NRSC data centre Hyderabad, measured the Sub basin properties namely area, slope, centroid location and elevation, minimum and maximum elevation, longest flow path length, reach length and slope are also computed.

Rating Curve (RC)

When developing rating curves based on existing data the traditional power curve model may not be the best choice. Depending on the situation, a polynomial or quadratic power curve model may provide a fit that statistically better than the power curve model. The power curve model tends to perform better when data points are closely spaced and no data points lay far from the main cluster. The results show that there was no back water correction for any sites and any steep or mild slope in the curve. The Rating curve for both Monsoon and Non-Monsoon were almost a straight line. The Value of R^2 for monsoon range from a maximum of 0.946 that is for *Tikerpada* and a minimum of 0.582 that was Talcher. Similarly for Non-Monsoon the value of R^2 range from 0.902 to 0.342. Regardless of the sign of the expected bias associated with traditional approaches (underestimation or overestimation) our study clearly points out that bias and overall uncertainty associated with rating-curves can be dramatically reduced by constraining the

identification of rating-curve with information resulting form simplified hydraulic modeling, with significant advantages for practical applications.

Regional Flow Duration Curves (RFDC)

Sometime, either adequate flow records for gauged catchments are not available or numbers of gauged catchments are limited. It makes the development of regional flow duration curve using method of regionalizing the parameters of chosen probability distribution sometimes erroneous. In the present study, the results presented in this section used two seasonal periods i.e. monsoon (July-Sept) and Non-Monsoon (Oct-June) periods to construct regional flow-duration curves. Regional flow duration curve in Mahandi basin were developed using flow data for fourteen catchments. Data in monthly steps were not available in twenty three catchments, so was not considered in this case. The mean monthly flows from both monsoon and non monsoon separate were fitted in flow exceedance points by a probable fitting equation (viz: second degree and power) which on logarithm tic transformation becomes linear that is popularly adopted. From the results as seen earlier, most ephemeral catchments have poorer RFDC fits than the perennial catchments, and during monsoon period, probability of more than ninety seven percent > 300 cumecs, flow exceeding 1000 cumecs has a probability of fifty percent and 3000 cumecs flow exceedance comes to around ten percent probability. The curve has a steep slope throughout and a flat slope towards the end that indicates a highly variable stream with flow largely from direct runoff and large amount of storage. In regards to uncertainty in RFDC results, the individual catchments are averaged for two bands i.e. lower (nearly 5 % exceedence) and upper (nearly 97 % exceedence) and the corresponding runoff values are related with catchment area in chapter 10 dealing with uncertainty. The RFDC as per standard procedure, and it can be seen from the results (Table 6.2 - 6.3 and figures in Annexure 6) with R^2 values exceeding 0.9382 (Weibull), 0.9182 (Blom), 0.913 (Gringorton) and 0.9163 (Cunnane) respectively. This is fairly a good fit as $R^2 > 0.9$ for monsoon and non-monsoon. In general for

the Mahanadi region, the curve has a steep slope throughout that denotes a highly variable stream with flow largely from direct runoff. The slope of the lower end of the duration curve shows a flat slope that indicates a large amount of storage; and at the upper end is a steep slope that indicates a negligible amount.

Synthetic Unit Hydrograph Methods

The probability distribution functions (*pdf*) that has similar shape to unit hydrograph (UH) are generally used for formulating Synthetic unit hydrograph (SUH) methods. Few methods frequently used in such cases are Nash model similar to two parameter- Gamma distribution Clark's and geomorphological instantaneous unit hydrograph (GIUH) methods. This study reported herein focused on two aspects i.e. the traditional methods of SUH derivation, e.g., GIUH based Nash and Clark's approach, Snyder's method, and SCS-curve number (CN) and conventional methods using events data. Among the four pdfs analyzed in this study, the Clark's and geomorphological instantaneous unit hydrograph models are more flexible in description of unit hydrograph SUH shape as they skew on both sides similar to a unit hydrograph, and on the basis of their application to field data. In case of SCS- Curve Number (CN) method for computing UH, the variable CN for a catchment has to be ascertain talking updated land use index or else there occurs a large error in determining the peak flow (q_p) and time to peak (t_p) along with W50, W75 (width of UH joining 0.5 and 0.75 peak flow ordinates) and time to base. In the present study some catchments with area more than 100 km^2 gives erroneous results. In case of Nash model that is similar to a Gamma distribution (in a limiting case), and the results shows that it should be a preferred method for deriving SUH. Parameters are highly sensitive to peak flow of the UH in case of Nash model (when β i.e. product of q_p t_p is low). Any overestimation in parameter estimates increases the peak flow of the UH and the trend is reverse at large β values. The variance of a positively skewed UH was observed to be

more sensitive to time to base (t_B) than t_p, implying that the σ^2-ratio for any two catchments approximately varied in proportion to the corresponding t_B^2-ratio. Figure 3 shows the results of flood hydrographs derived using four of the above methods in catchments in Mahanadi basin.

Regional Unit Hydrograph Analysis

In this case, eleven catchments in Mahanadi basin with relevant data are used for regional analysis, and in a nutshell is briefed hereby. Regional unit hydrograph for lower Mahanadi region based on Nash and Clark approach yielded the model parameters with the basin as follows: nk and $LL_c/S^{0.5}$ with $R^2 = 0.57$, K and L with $R^2 = 0.63$ and time of concentration (Tc) and $LL_c/\sqrt{S}$ with $R^2 = 0.99$. Data of unit hydrograph derived from stream flow and rainfall records in basins adjoining the basin under study should be used to estimate the constants for use to any regional unit hydrograph study. The results at Table 9.2 (in Chapter 9) gives R^2 mostly more than 0.6 and a fair fit.

Flood Frequency Analysis Techniques

The understanding of floods plays a key role in many hydrological studies, especially in the design of hydraulics structures such as dams, culverts, bridges and others. Extreme hydrological events are not only important in the design of water resource projects but also in the management of water resources. Flood frequency estimation remains an important topic for such design purposes, and flood data constitute the main source of information for this analysis. Single stations analyses were carried out for 37 hydrometric stations located in the Mahandi watershed. *Maximum daily discharges* (m^3/s) or Annual Maximum Series (AMS) for 25 years for 23 small catchments and about 35 years for 14 gauging sites were analyzed using seven distributions viz. EV1, GEV, PT3, LP3, GP, GL and WAK. Three parameter estimations methods were applied for this case and they are method of moments (MOM), method of

maximum likelihood (ML) and method of probability weighted moment (PWM). A regional flood frequency analysis was also carried out for the same study area and is given in next section. Since new data are available, the goal of the present study was to update those flood frequency analyses previously analyzed. As such, results presented in this document will better reflect our current state of knowledge regarding the high flow regimes throughout the Mahanadi region. In general, the results of the present study are consistent with those from early studies, although it can be seen that updating the flood information resulted, for many stations, in an improvement of flood estimates.

Results of the 36 single station high flow frequency analyses are provided in Appendix 9 (b) and (c) for the estimated parameters and the flood for recurrence intervals of 20, 50, 100, 200, 300 and 500 years. From the single station or at-site analysis results, it is clear that for the majority of the single analysis both the LP3 and GEV fitted the data almost similar. However, at the high recurrence intervals, it is evident that GEV adjusts better to the observational data than LP3 and GP. Results of the R.M.S. error statistics favor GEV over LP3 approximately 64% of the time. For the 36 GD sites, the corresponding return period flood are estimated separately and the goodness-of-fit assessments suggested the GEV model to be the overall most appropriate distribution function. These findings were also strengthened by extreme value theory, which suggests that the annual maxima (such as flood data) be modeled as realizations of random variables distributed according to a member of the *generalized extreme value* (GEV) family of distributions. Based on such considerations, the GEV was therefore be accepted for the estimation of at-site floods as a function of annual maximum runoff for various recurrence intervals (i.e., 20, 50, 100, 200, 300 and 500). In all cases, the fitted models were consistent with the calculated T-year flood events, such that they could be applied to predict floods for ungauged basins (within their range of application).

Regional Flood Frequency Analysis

Screening of the data conducted using the annual maximum peak flood data of the Mahanadi Sub zone 3(d) employing the Ddiscordancy measure (D_i) test reveals that 22 out of the 23 Br-catchments (gauging sites) are suitable for regional flood frequency analysis. However, based on the heterogeneity measures, `H(j)'; the annual maximum peak flood data of other 13 GD sites are considered to constitute a homogeneous region. So, in a total, 36 catchments (with gauging sites) are used for regional flood frequency analysis. Various distributions viz. EV1, GEV, GL, PT3, LP3, GNO, GP, and WAK have been employed. Based on the L-moments ratio diagram and $|Z_i^{dist}|$ –statistic criteria, the GNO distribution has been identified as the robust distribution for the study area.

The developed regional flood frequency relationships may be used for estimation of floods of desired return periods for gauged and ungauged catchments of the study area. As the regional flood frequency relationships have been developed using the data of catchments varying from 19 to 1,150 km^2 in area; hence, these relationships are expected to provide estimates of floods of various return periods for catchments lying nearly in the same range of the areal extent. Further, the relationship between mean annual peak flood and catchment area is able to explain 83.4% of initial variance (R^2 = 0.834). Hence, in case of ungauged catchments the results of the study are subject to these limitations. However, the regional flood frequency relationships may be refined for obtaining more accurate flood frequency estimates; when the data for some more gauging sites become available and physiographic characteristics other than catchment area as well as some of the pertinent climatic characteristics are also used for development of the regional flood frequency relationships.

Peak over Threshold Techniques

In some regions, annual maximum floods observed in dry years may be very small, and inclusion of these events can significantly alter the outcome of the extreme value analysis. In contrast, the peak over threshold (POT) model avoids all these drawbacks by considering all flood peaks above a certain threshold level; hence more than one flood per year may be included, and exceptionally low annual maximum values that may cause problems in parameter estimation are not included. POT model are often used for flood frequency analysis when there is a paucity of data. The analysis presented in this study focused on the choice of generalized Pareto (GP) distribution for computation of flood exceedance in POT model, when the occurrences of peak exceedances are fitted with a PD distribution. The reliability of the expressions for $\mathrm{Var}[\hat{q}(T)]$ derived using the analytical procedure was checked using the Monte Carlo simulation method. Next, a ratio between the $\mathrm{Var}[\hat{q}(T)]$ of analytical and simulation procedures denoted as R_1 was used to check the reliability of these expressions. Results showed R_1 to be asymptotically reaching 1 for AMS and PDS models, especially for higher record length i.e., at N=50. The difference between the simulated and the analytical results becomes more prominent with increase in T years, with a significant difference e.g., R_1 = 0.92 for AMS/PD-GP model for a 300 years return period, when a 10 years sample was used. The difference is sometimes caused by an error in evaluation of the parameter estimates variances and covariance, when based on the expected information matrix.

For the AMS model the variance of quantiles derived using analytical expressions are overestimated while the PDS models underestimate the same. This is evident from Figure 13.2 and 13.3 (given in Chapter 13) for POT/PD-GP and AMS/PD-GP, respectively where the R_1 values lies below one for most of the cases. The above referred figure shows the change of R_1 values with T in case of PDS models, and R_1 lies over one for all cases in the range of 10 to300

years return period. To check the effect of choice of initial parameters selected for the simulation on the overall results of the ratios of variances, a different set of test was carried out with a new threshold, but the outcome was found to be similar. Hence, the choice of initial values of parameters in simulation was found not to affect the results of R_1.

The advantages of the Poisson distribution over negative binomial distributions have been dealt in the past by ÖnÖz and Bayazit (2001). On the basis of a detailed study they reported that the flood estimates and their corresponding variance based on the binomial and negative binomial distribution when combined with ED for the magnitudes of the peaks are almost identical to those obtained using PD. The above conclusions were in agreement with the findings of Kirby (1969) and Cunnane (1979). However, most of these studies focused on exponential distribution in POT models, and no such study have been reported in literature so far as to where PD distribution is used with GP distribution for modeling the flood exceedances in POT. From the overall results it can be observed that the Poisson distribution performs slightly better when the difference between mean and variance is small e.g. when $\mu = 1$ and $\sigma^2=1.2$.

Confidence Intervals and Uncertainty of Predictions

To estimate the uncertainty in the regional Flow Duration Curves (RFDC), flood hydrograph using Unit Hydrographs (UH) and recurrence flood using flood frequency analysis, it needs the methods of statistics and probability for the quantification of uncertainty in the hydrologic predictions. A set of thirteen small watersheds distributed throughout Mahanadi basins are considered for this case. These watersheds cover a range of watershed characteristics including soil type, topography and land use, as well as time series of daily precipitation, and streamflow for about 10-35 years period.

Identifying the characteristics of the minimum flow duration distribution and deciding on appropriate uncertainty probabilities (the median plus the ensembles equaled or exceeded >3%, >50% and 97% of the time for example) shows approximately >3% probability and the graph is exponential, >50% probability graph is straight line and >97% probability graph is scattered based on FDC using the median hydrology ensemble. . The conclusion is therefore that yield assessments based on a probabilities of exceedence of between about >3% and >97% would not be affected by including hydrological uncertainty. Using the unstructured sampling program (no stochastic daily flow inputs), it is clear that the high flow exceedance probabilities i.e. < 3 % sampling has resulted in much lower uncertainty compared to low flow exceedance i.e. > 96 % in the basin.

Similarly the Table 12.5 (in Chapter 12) shows the mean errors estimates for computed unit hydrographs with the conventional methods using four different methods, and the results show different degree of errors ranging from minimum 2.32 to maximum 30.56. In similar way, Table 12.6 (Chapter 12) illustrates the calculation of confidence band, for one of the catchment (Br-No 7). The approximate expressions for the *T*-year event estimation corresponding to a probability of non-exceedance, and its Root Mean Square Error (RMSE) using seven different pdf provides the variance of return period flood and show that GEV, LP3 and WAKEBY models yields less variance of q(T) than others for different return periods and is consistent.

REFERENCES

1. Ahmad, A.I., C.D. Sinclair, and A. Werritty, Log-logistic flood frequency analysis, *J. Hydrol.,* 98, 204-225, 1988.
2. Anon. (1984); *The Macquarie Illustrated World Atlas*. McMahons Point NSW, Macquarie.
3. Aron. G, White, E.L., 1982. Fitting a gamma distribution over a synthetic unit hydrograph, Water Resources Bull. 18(1): 95-98.
4. Arthington, A. (1994) A holistic approach to water allocation to maintain the environmental values of Australian streams and rivers: a case history. *Mitt. Internat. Limnol*. 24, 165–177.
5. Bain, M.B., Finn, J.T. & Booke, H.E. (1988) Streamflow regulation and fish community structure. *Ecology* 69, 382–392.
6. Belbin, L. (1991a) *PATN Pattern Analysis Package*. CSIRO, Division of Wildlife and Ecology.
7. Belbin, L. (1991b) Semi-strong hybrid scaling, a new ordination algorithm. *Journal of Vegetation Science* 2, 491–496.
8. Bernard, M., (1935); An approach to determinate stream flow. Trans. ASCE, 100, 347-395.
9. Beumer, J.P. (1980) Hydrology and fish diversity of a North Queensland tropical stream. *Australian Journal of Ecology* 5, 159–186.
10. Bhunya, P K, R D Singh, R. Berndtsson, and S N Panda. Flood analysis using generalized logistic models in partial duration series, *J. Hydrol.,* HYDROL 17947, In press, 2012.

11. Bhunya, P.K, S. K. Jain , C. S. P. Ojha and A. Agarwal. A Simple estimation technique for three-parameter generalized extreme value (GEV) distribution, Vol-12(6), 682-689, Journal of hydrologic engg., ASCE

12. Bhunya, P. K., Berndtsson, R., Ojha, C. S. P., and Mishra, S. K. (2007) Suitability of Gamma, Chi-square, Weibull and Beta distributions as synthetic unit hydrographs. J. Hydrology 334, 28-38.

13. Bhunya, P. K., Ghosh, N. C., Mishra, S. K., Ojha, C. S. P. and Berndtsson, R., (2005) Hybrid Model for Derivation of Synthetic Unit Hydrograph. J. Hydrologic Eng., 10(6), 458-467.

14. Bhunya, P. K., Mishra, S. K., Ojha, C. S. P. and Berndtsson, R., (2004) Parameter Estimation of Beta Distribution for Unit Hydrograph Derivation. Hydrograph. J. Hydrologic Eng., 9(4), 325-332.

15. Bovee, K. (1982) A Guide to Stream Habitat Analysis using the Instream Flow Incremental Methodology. US Fish and Wildlife Service.

16. Brooks, S. & Lake, P. (1995) *Impacts of Hydrological Disturbance on Stream Communities* LWRRDC Milestone Report.

17. Central Water Commission (1982a), Flood estimation report for Mahanadi sub-zone-3d, Directorate of hydrology (small catchments), India.

18. Central Water Commission (1997), Flood estimation report for Mahanadi sub-zone-3d (Revised), M3 (d)/R-3/25-97, Directorate of hydrology (small catchments), India.

19. Characterisation of Flow in Regulated and Unregulated Streams in Eastern Australia Cooperative Research Centre for Freshwater Ecology.

20. Chow, V.T., 1964. Runoff Section 14, Handbook of Applied Hydrology. McGraw Hill Book Co., New York.

21. Chow, V.T, Discussion of annual floods and the partial duration flood series, Trans. Am. Geophys. Union, 31(6), 939 941, 1950.

22. Clarke, K.R. (1993) Non-parametric multivariate analyses of changes in community structure *Australian Journal of Ecology* 18, 117–143.

23. Clausen, B. & Biggs, B. (1997) Relationships between benthic biota and hydrological indices in New Zealand Streams.*Freshwater Biology* 38, 327–342.

24. *Continental Comparisons of Annual Flows and Peak Discharges*. Catena Verlag, Germany.

25. Croley II, T.E., 1977. Hydrologic and Hydraulic Computations on Small Programmable Calculators. Iowa Inst. Hydraul. Res., Univ. Iowa, Iowa City, Iowa, pp. 837.

26. Croley, T.E., II. , 1980. Gamma synthetic hydrographs, J. Amsterdam of Hydro. 47, Amsterdam: 41-52.

27. Cunnane, C., A Note on the Poisson assumption in partial duration series models, *Water Resour. Res*., 15(2), 489-493, 1979.

28. Cunnane, C., A particular comparison of annual maximum and partial duration series methods of flood frequency prediction', *J. Hydrol*., 18, 257-271, 1973.

29. Cunnane, C., Statistical distributions for flood frequency analysis, *World Meteorological Organization Operational Hydrology,* Report No-33, WMO No. 718, Geneva, 1989.

30. Dooge, J. C. I., (1959) A general theory of the unit hydrograph. J.Geophys. Res., 64(2), 241–256.

31. Dubey, S.D, A new derivation of the logistic distribution, *Naval Research Logistics Quarterly, 16,* 37-40, 1969

32. Ekanayake, S. T., and J. F. Cruise, Comparisons of Weibull- and exponential-based partial duration stochastic flood models, *Stochastic Hydrol. Hydraul.,* 7(4),283-297, 1993.

33. Espey, W. H. Jr., Winslow, D.E., 1974. Urban flood frequency characteristics, Proc. ASCE 100 HY2: 179-293.

34. Extence, C.A., Balbi, D.M. & Chadd, R.P. (1999) River flow indexing using British benthic macroinvertebrates: a framework for setting hydroecological objectives. *Regulated Rivers: Research and Management* 15, 543–574.

35. Gentilli, J. (1986) Climate. In *Australia — A Geography. Volume 1 The Natural Environment* (ed. D.N. Jeans) Sydney University Press, pp. 14–48.

36. Gray, D.M (1961a),"Interrelationship of watershed characterstics. ‚J. of Geophysical Research, 66, 1215-1223.

37. Gray, D.M (1961b),"Synthetic hydrographs for small drainage areas, Proc., ASCE, 87, HY4, 33-54.

38. Grayson, R.B., Argent, R.M., Nathan, R.J., McMahon, T.A. & Mein, R.G. (1996) *Hydrological Recipes: Estimation Techniques in Australian Hydrology*. Cooperative Research Centre for Catchment Hydrology, Clayton, Victoria.

39. Greenwood, J. A., J. M. Landwehr, N. C. Matalas, andJ, R, Probability-weighted moments definition and relation to estimators of several distributions expressible in inverse form, *Water Resour. Res.,* 15, 1049-1054, 1979.

40. Growns, I. & Growns, J. (1997) The relationship between biota and hydrology in the Hawkesbury–Nepean River system and their implications for environmental flow allocation. In *Science and Technology in the Environmental Management of the Hawkesbury–Nepean Catchment*. Australian Institute of Engineers, University of Western Sydney, pp. 54–60.

41. Gupta, H.V., Hsu, K-L., and Sorooshian, S, Effective and efficient modeling for streamflow forecasting' in Artificial Neural Networks in Hydrology, edited by R.S. Govindaraju and A.R. Rao, *Kluwer* Academic Publishers, Dordrecht, 2000.

42. Gupta, H.V., Hsu, K-L., and Sorooshian, S, Effective and efficient modeling for streamflow forecasting' in Artificial Neural Networks in Hydrology, edited by R.S. Govindaraju and A.R. Rao, *Kluwer* Academic Publishers, Dordrecht, 2000.

43. Haktanir, T and Nurullah Sezen., Suitability of two-parameter gamma distribution and three-parameter beta distribution as synthetic hydrographs in Anatolia, *Hydrlo. Sc. Journal*, 35(2), 167-184, 1990.

44. Hosking, J.R.M., and J. R. Wallis, Regional frequency analysis: An approach based on L-moments, Cambridge University Press, UK, 1997.

45. Hosking, J.R.M., The theory of probability weighted moments, *Res. Rep*., RC 12210, IBM, Yorktown Heights, NY, 1986.

46. Hudlow, M.D., and Clark, D.M., (1969) Hydrological synthesis by digital computers. J. Hydr. Div. ASCE, 95(3), 839-860.

47. Hughes, J. & James, B. (1989) A hydrological regionalization of streams in Victoria, Australia, with implications for stream ecology. *Australian Journal of Marine and Freshwater Research* 40, 303–326.

48. Institution of Engineers Australia (1987) *Australian Rainfall and Runoff: a Guide to Flood Estimation*. Revised edition. Institution of Engineers Australia.

49. Jowet, I. & Duncan, M. (1990) Flow variability in New Zealand rivers and its relationship to instream biota. *New Zealand Journal of Marine and Freshwater Research* 24, 305–317.

50. Jowet, I. (1982); The incremental approach to instream flow needs. New Zealand case studies.

51. Kafarov, V., (1976), Cybernetic methods in Chemistry and Chemical Engineering. English translation, Mir Publishers, Moscow, 483p.

52. Kafarov, V., (1976), Cybernetic methods in Chemistry and Chemical Engineering. English translation, Mir Publishers, Moscow, 483p.

53. Kimball, B. F., An approximation to the sampling variances of an estimated maximum value of given frequency based on the fit of double exponential distribution of maximum values, *Ann. Math. Sta/., 20,* 110-113, 1949.

54. Kite G.W., 1977, Frequency and Risk Analysis in Hydrology, Water Resources Publications, Fort Collins, Colorado, USA.

55. Kjeldsen, T. R. and D. A. Jones, Sampling variance of flood quantiles from the generalised logistic distribution estimated using the method of L-moments, *Hydrol. Earth System Sci.*, 8(2), 183-190, 2004.

56. Kottegoda N. T. and Rossa R., 1998, Statistics, Probability and Reliability for Civil and Environmental Engineers, McGraw-Hill.

57. Kuchment, L.S., 1967. Solution of inverse problems for linear flow models. Soviet Hydrology, Selected Papers, Vol. 2, p.94.

58. Landwehr, J.M., N.C.Matalas, and J.R.Wallis, Probability weighted moments compared with some traditional techniques estimating gumbel parameters and quantiles, *Water Resour. Res*., 15, pp 1055-1064, 1979.

59. Landwehr, J.M., N.C.Matalas, and J.R.Wallis, Probability weighted moments compared with some traditional techniques estimating gumbel parameters and quantiles, *Water Resour. Res*., 15, pp 1055-1064, 1979.

60. Lang, M, T. B. M. J. Quarda and B. Bobee, Towards operational guidelines for over threshold modeling, *J. Hydrolo*., 225, 103-117, 1999b.

61. Lang, M., Theoretical discussion and monte-carlo simulations for a negative binomial distribution paradox, *Stoch. Env. Res. Risk Assessment*, 13, 183-200, 1999a.

62. Langbein, W. B., Annual floods and the partial duration flood series, Transactions, AGU, 30(6), 879-881, 1949.

63. Langbein, W. B., et al. (1947) Topographic characteristics of drainage basins." Water supply paper 968-C, U.S. Geological Survey, 125–155.

64. Library, Division of Lands Mapping, Canberra, Department of Lands and Survey, Wellington.

65. Linsley, R.K., Kohler, M.A., and Paulhus, J.L.H. (1975) Hydrology for Engineers. McCraw-Hill, New York.

66. Madsen, H., P.F.Pearson, and D.Rosbjerg, Comparison of annual maximum series and partial duration series methods for modeling extreme hydrological events 2. At-site modeling, *Water Resour. Res.*, 33(4), 747-757, 1997.

67. Martins, E. S. and J. R. Stedinger, Generalized maximum-likelihood generalized extreme-value quantiles for hydrological data, *Water Resour. Res*., 36, 737-744, 2000.

68. Martins, E. S. and J. R. Stedinger, Generalized maximum-likelihood Pareto-Poisson estimators for partial duration series, *Water Resour. Res*., 37, 2551-2557, 2001.

69. McCuen, R.H., and Bondelid, T.R., (1983) Estimating unit hydrograph peak rate factors. J. of Irrigation and Drainage Engineering, ASCE 110(7): 887-904.

70. McMahon, T.A., Finlayson, B.L., Haines, A.T. & Srikanthan, R. (1992).

71. Miller, A.C., Kerr, S.N., and Spaeder, D.J., (1983) Calibration of Snyder coefficients for Pennsylvania. Water Resources Bulletin, 19(4), 625-630.

72. Mishra, S K, and Behera K. Development and Management of Water and Energy Resources. 7th International R& D Conference, 4-6 Feb, 2009, Bhubaneswar, Orissa.

73. Montgomery, D.C., and G.C. Runger, Applied statistics and probability for engineers, John Wiley Sons Inc., NY, 1994.

74. Nash, J.E., 1959. Synthetic determination of unit hydrograph parameters, J. of Geophysics Res., 64(1):111-115.

75. Nathan, R. & Weinmann, P. (1993) *Low Flow Atlas for Victorian Streams*. Department of Conservation and Natural Resources, Victoria.

76. National Environmental Research Council, *Flood studies report*, Vol-1, NERC, UK., 1975.

77. ÖnÖz, B., and M.Bayazit, Effect of occurrence process of the peaks over threshold on the flood estimates, *J. of Hydrol.,* 244, 86-96, 2001.

78. Parhi. P K, S K Mishra, R. Singh and V K Tripathi. Floods in Mahanadi River Basin, Orissa (India): A Critical Review. *India Water Week 2012 – Water, Energy and Food Security: Call for Solutions, 10-14 April 2012,* New Delhi.

79. Phien, H. N., A review of methods of parameter estimation for the extreme value type-1 distribution, *J. Hydrol., 90, 251-268,* 1987.

80. Poff, N. & Allan, J. (1995) Functional organization of stream fish assemblages in relation to hydrological variability. *Ecology* 76, 606–627.

81. Poff, N. & Ward, J. (1989) Implications of streamflow variability and predictability for lotic Community structure: a regional analysis of streamflow patterns. *Canadian Journal for and Aquatic Science* 46, 1805–1818.

82. Poff, N., Allan, J., Bain, M., Karr, J., Prestegaard, K., Richter, B., Sparks, R. & Stomberg, J.

83. Ponce, V. M. (1989) Engineering hydrology: Principles and practice, Prentice-Hall, Englewood Cliffs, N.J., 175–182.

84. Puckridge, J., Sheldon, F., Walker, K. & Boulton, A. (1998) Flow variability and the ecology of large rivers. *Marine and Freshwater Research* 49, 55–72.

85. Ramberg, A.S., A probability distribution with application to Monte Carlo simulation studies, *in Statistical Distributions in Scientific Work,* Vol-2, Ed. By G.P. Patil, S. Kotz, and J.K. Ord, pp. 51-64, D. Reidel, Hingham, Mass, 1975.

86. Rao, A.R and K.H. Hamed, Flood frequency analysis, CRC Press, Boca Raton, Florida, US, 2000.

87. Richter, B., Baumgartner, J., Powell, J. & Braun, D. (1996) A method for assessing hydrological alteration within ecosystems. *Conservation Biology* 10, 1163–1174.

88. Richter, B.D., Baumgartner, J.V., Wigington, R. & Braun, D.P. (1997) How much water does a river need? *Freshwater Biology* 37, 231–249.

89. *River Low Flows: Conflicts of Water Use*. R.S. McColl, Water and Soil Miscellaneous Publications 47, 9–15.

90. Rosbjerg, D., H. Madsen, and P.F. Rasmussen, Prediction in partial duration series with generalized pareto distributed exceedances, *Water Resour. Res*., 28(11), 3001-3010, 1992.

91. Rosso, R., (1984) Nash Model Relation to Horton Order Ratios, *Water Resour. Res*., 20(7), 914-920.

92. SCS. 1957. Use of storm and watershed characteristics in synthetic hydrograph analysis and application:V. Mockus, U.S. Dept. of Agriculture, Soil Conservation Service, Washington, D.C.

93. Searcy, J.K (1959). Flow duration curves. Water Supply Paper 1542-A, U.S., Geological Survey, Reston, Virginia.

94. Shane, R.M., and W.R.Lynn, Mathematical model for flood risk evaluation', *J. Hydraulics Div*., ASCE, 90 (HY6), 1-20, 1964.

95. Sherman, L.K., 1932. Streamflow from rainfall by the unit hydrograph method, Engrg. News Rec., 108, 501-505.

96. Singh, V.P., 1991. Elementary Hydrology. Prentice Hall, Englewood Cliffs, New Jersey.

97. Singh, V.P., 1988. Hydrologic systems: Rainfall-runoff modeling Vol-1, Prentice Hall, Englewood, New Jersey.

98. Snyder, F.F., 1938. Synthetic unit hydrographs, Trans. Am. Geophysics Union, 19: 447-454.

99. Sokolov, A.A, Rantz, S.E., and Roche, M., 1976. Methods of developing design flood hydrographs: Flood computation methods compiled from world experience, UNESCO, Paris, 1976.

100. Spiegel, M.R., Theory and problems of mathematical statistics, McGraw-Hill, NY, 118, 1987.

101. Subrahmanyam, V., A.S. Subrahmanyam, G.P.S. Murty and K.S.R. Murthy. 2008. Marine Geology, Vol.253; 63-72p.

102. Takeuchi, K., Annual floods and the partial duration flood series-Evaluation of Langbein's formula and Chow's discussion, *J. Hydrol.,* 68, 275-284,1984.

103. Taylor, A.B and H.E. Schwarz., 1952. Unit hydrograph lag and peak flow related to basin characteristics, Trans. of the American Geophysical Union 33:235-46.

104. Todini E. 1988. Rainfall-runoff modeling—past, present and future. Journal of Hydrology 100: 341–352.

105. Toner, M. & Keddy, P. (1998) River hydrology and riparian wetlands: a predictive model for ecological assembly. *Ecological Applications* 7, 236–246.

106. Tuckey, J.W., The practical relationship between the common transformations of percentages or fractions and of amounts, *Tech. Rep. 36,* Statists. Tech. Res. Group, Princeton Univ., Princeton, N.J, 1960.

107. U. S. Army Corps of Engineers (USACE). (1990) Flood hydraulics package. User's Manual for HEC-1, CPD-1A, Version 4.0, USACE, Washington, D. C.

108. U.S. Army Corps of Engineers, 1973. Hydrograph anlaysis. Vol. 4, Oct.

109. Verhulst, P. F., *Recherches Mathematiques sur La Loi D'Accroissement de la Population*, (Mathematical Researches into the Law of Population Growth Increase)Nouveaux Memoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles, 18, Art. 1, 1-45, 1845.

110. Vogel, R. M., and Fennessey, N. M. (1993). "L-moment diagrams should replace product-moment diagrams." Water Resour. Res., 29(6), 1745-1752.

111. W. M. O., 1968. Data processing for Climatological Purposes. Technical Note No. 100, Proceeding of the W. M. O. Symposium, Ashevillr, N.C.

112. W. M. O., 1981. Case Studies of National Hydrlogical Data Banks (Planning, Development and Organisation). Operational Hydrology Report No. 17.

113. W. M. O., 1981. Guide to Hydrological Practices, Volume 1, Data Acquisition and Processing No. 168.

114. Wang, R.Y., and Wu, I.P., (1972) Characteristics of short duration unit hydrograph. Transaction of ASAE, 15, 452-456.

115. Willson, E.M., 1969. Engineering Hydrology. Mac Millon, London.

116. Wu, I.P., (1969) Flood hydrology of small watersheds: evaluation of time parameters and determination of peak discharge. Transactions of ASAE, 12, 655-660.

117. Yeh, K C., Tung, Y.K., Yang, J. C., Zhao, B. 1993: Uncertainty analysis of hydrologic models and its implications on reliability of hydraulic structures (2). Technical Report, Agricultural Council, Executive Yuan, Taiwan.

118. Yue,S, Taha, B.M.J., Bernard Bobee, Legendre, P., and Pierre Bruneau., 2002. Approach for describing statistical properties of flood hydrograph, J. hydrologic engineering ASCE 7(2): 147-153.

119. Zelenhasic, E., Theoretical probability distributions for flood peaks, *Hydrol. Pap*42, Colo. State Univ., Fort Collins, 1970.

120. Zhao, B. 1992: Determinations of a unit hydrograph and its uncertainty applied to reliability analysis of hydraulic structures. M .S. thesis. Department of Statistics, University of Wyoming, Laramie, Wyoming. 352pp.

121. Zhao, B., Tung, Y. K., Yang, J .C.1994: Determination of nit hydrographs by multiple storm analysis. Journal of Stochastic Hydrology and Hydraulics 8(4), 269-280.