DEVELOPMENT OF FLOW DURATION CURVE FOR UNGAUGED BASINS USING VARIOUS REGIONALIZATION METHODS

A DISSERTATION

Submitted in partial fulfilment of the requirement

for the award of the degree

of

MASTER OF TECHNOLOGY

in

WATER RESOURCES DEVELOPMENT

by

BOLIWAL GOPALKUMAR ROSHANLAL

(Enrolment No. 17548002)

DEPARTMENT OF WATER RESOURCES DEVELOPMENT AND MANAGEMENT

INDIAN INSTITUTE OF TECHNOLOGY

ROORKEE- 247667 (INDIA)

May, 2019

CANDIDATE'S DECLARATION

I hereby declare that the work presented in the dissertation entitled "**DEVELOPMENT OF FLOW DURATION CURVE FOR UNGAUGED BASINS USING VARIOUS REGIONALIZATION METHODS**", in partial fulfilment of the requirement for the award of the degree of **Master of Technology in Water Resources Development**, submitted in the Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee is an authentic record of my own work carried out during the period from July 2018 to May 2019 under the supervision of **Dr. Deepak Khare, Professor,** Department of Water Resources Development and Management, IIT Roorkee and **Dr. Manohar Arora, Scientist D,** National Institute of Hydrology Roorkee , Roorkee.

I have not submitted the matter in this Dissertation for the award of any other Degree.

Date:

Roorkee, Uttarakhand, India

……………………………

Gopalkumar Boliwal

Enrolment No: 17548002

CERTIFICATE

This is to certify that the above declaration made by candidate is correct to the best of my knowledge.

……………………...

Dr. Manohar Arora

Scientist – D,

National Institute of Hydrology,

Roorkee

……………………...

Dr. Deepak Khare

Proffessor, WRD & M

Indian Institute of Technology,

Roorkee

ACNOWLEDGEMENT

It is my great pleasure in expressing my heartfelt gratitude to my supervisors **Dr. Deepak Khare, Professor**, Department of Water Resources Development and Management, Indian Institute of Technology Roorkee and **Dr. Manohar Arora, Scientist D, National Institute of Hydrology Roorkee,** Rorkee for their valuable guidance, direction and infilling support towards completion of this report. I am highly obliged to them for their keen interest, able guidance and encouragement throughout the writing. Working under their guidance is a privilege, great opportunity and an excellent learning experience that I will flourish in my life time.

I am extremely thankful to other professors, staffs and classmates from the department of Water Resources Development and Management, IIT Roorkee and the National Institute of Hydrology Roorkee, Roorkee for their kind support, advice, encouragement and cooperation.

…………………

(Gopalkumar Boliwal)

Enrolment no. 17548002

ABSTRACT

The planning and development of water resource structures of a region requires a proper understanding of the hydrological behaviour of the river basin. The analysis of available historical information is necessary to decide about water availability estimation of the available water resource. Water availability generally refers to the volume of water available from a basin or stream. The planning, development and operation of water resource project is very much depend upon the availability of hydrological data of desired quantity and quality. Generally, for most basin in the country, rainfall records of sufficient length are available. However, the available runoff data is either short or has gaps due to missing data. The shorter length of data along with other problems, always lack the true representation of natural behaviour of the time series. In the absence of data, the best suited alternative is to use the regionalisation method. Regionalisation techniques provide a mechanism to relate the hydrological behaviours of ungauged catchments in a region. This techniques can be helpful to utilise to derive stream flow characteristics at ungauged catchments of the region.

In the present study the data of gauged catchments of the upper Krishna basin has been used to find out the discharge of ungauged sites of the corresponding catchments. Data from eleven river gauging stations has first been taken to find out the flow in the upstream ungauged tributaries of these gauged catchments. Water availability with time variability at the proposed project site is essential to estimate the power potential and annual energy generation. The Flow Duration Curve (FDC) is a simple depiction of water availability with time variability. It shows a discharge which has equalled or exceeded certain percentage of time out of the total time period which is generally taken as one year. The FDC at the ungauged site has been developed using the regionalization of the parameters of chosen probability distribution for the gauged sites and concluded that which regionalisation method is best for the upper Krishna basin.

TABLE OF CONTENTS

List of Figures

List of Tables

LIST OF ABBREVIATIONS

C.W.C	Central Water Commission
M.C.M	Million Cubic Meters
C.G.W.B	Central Ground Water Board
U/S	Upstream
C.F.D	Cumulative Distribution Function
M.P	Madhya Pradesh
F.D.C	Flow Duration Curve
K1	Upper Krishna Basin
NIH	National Institute of Hydrology

CHAPTER 1

INTRODUCTION

1.1 General Background

A proper understanding of the hydrological behaviour of the river basin requires for the planning and development of any water resources structures. The analysis of available historical information is necessary to decide about water availability estimation of the available water resource. Volume of the water which is available from a stream or a basin is generally referred to as the water availability. According to the availability of the hydrological data of the required quantity and the quality, the planning, development and the operation of water resource projects are dependent. Rainfall records for most of the basins in our country of the required length is generally not available. Due to the missing data, the runoff data which is available are either having gaps or are short. In the absence of data, the best suited alternative is to use the regionalisation method. Regionalisation method is used to relate the physiographic characteristics of the basin with the flow of the ungauged catchment in a region. This technique can be helpful to utilise to derive stream flow data characteristics at ungauged catchment of the region.

Presently a standout amongst the most challenging tasks in surface water hydrology is flow simulation of ungauged catchment. Most of the ungauged catchment are located in the headwaters of the streams in hilly regions which has enormous potentials for sustainable water resources development. However, because of unavailability, rough and cold landscape, and chronicled absence of information they need these headwaters sufficiently checked, their potential isn't promptly feasible. Many places sufficient historical data are not available or not have sufficient gauged data at sites where most needed. As prescient devices for water assets, water quality, characteristic danger alleviation and water accessibility evaluation are for the most part information driven, the absence of sufficient hydrometric records presents troublesome issues for organizers, designers, farmers, and stack-holders alike. (M. Goswami et al. 2006)

Hydrological regimes can be described by the tool called Flow Duration (FD) curve. This curve shows the relationship between the frequency and the magnitude of the flow by defining the amount of time in proportion for which the discharge is equalled or exceeded (Vogel and Fennessey 1994). Flow-duration curve is the tool for engineers and hydrologists around the world which is used numerous applications related to designing of irrigation systems, water resources management, like hydropower generation and planning, management of stream-pollution, fluvial erosion and reservoir sedimentation (Castellarin et al. 2007).

Prediction in Ungauged Basins is an important task for water resources planning and management but remains a fundamental challenge for the hydrological community. During the previous decades, hydrologists have built up an assortment of prescient instruments (for example lumped and dispersed models) to appraise hydrological reactions to arrive use changes and atmosphere inconstancy. The parameters of these hydrologic prescient tools are by and generally obtained by hydrologic perceptions, which are inaccessible at ungauged destinations. The absence of information for model alignment and check in ungauged bowls requires the hydrological regionalization to exchange data (for example model parameters) from measured catchments.

The regionalization permits assessing parameter estimations of hydrological prescient instruments without calibration. Regionalization requires sound comprehension and learning of hydrological forms and spatio-fleeting heterogeneity of climatic and landscape properties present a challenge for expectation in ungauged catchments. Throughout the years, regionalization has gotten expanding consideration from the hydrological network. Various local models are presently accessible, including: (a) parametric regression (b) nearest neighbour technique and (c) hydrological comparability method. Among those, the parametric regression (ordinarily the numerous straight relapse) is one of the soonest and most broadly utilized strategies, in which the parameter esteems at checked locales are controlled by a parametric (for example straight) relapse condition between ideal parameters and catchment attributes set up in a lot of checked destinations. (Ming et al. 2010)

The regionalization allows estimating parameter values of hydrological predictive tools without calibration. Regionalization requires sound understanding and knowledge of hydrological processes and spatio-temporal heterogeneity of climatic and landscape properties present a challenge for prediction in ungauged catchments. Over the years, regionalization has received increasing attention from the hydrological community. A number of regional models are currently available, including: (a) parametric regression (b) nearest neighbour method and (c) hydrological similarity method. Among those, the parametric regression (typically the multiple linear regression) is one of the earliest and most widely used methods, in which the parameter values at gauged sites are determined by a parametric (e.g. linear) regression equation between optimal parameters and catchment characteristics established in a set of gauged sites. (Ming et al. 2010)

The parameters of the hydrological model for ungauged watersheds, parameter values were obtained from a selected from area having hydrometric stations, and process of transferring hydrological model parameters from neighbouring catchments area to the study is referred to as regionalisation (Bloschl et al. 1995). In the regionalisation methods the data itself check the randomness or hypotheses. The data that transposed from area having hydrometric stations were determined on the basis of the similar physiographical properties for the calibration of the model with validation of other gauged which we have taken as the ungauged catchment.

1.2 Research gap

The literature review shows that there are number of research and publication on hydrological modelling in watershed gauged area in different region of the world .but there is very limited works is found in urban watershed or ungauged area. The projected climate change and the growing population are likely to make the use of water in the ungauged area more difficult. By using regionalisation method we can easily understand the interface between water availability and hydro metrological parameters.

The requirement for strategies to manage ungauged catchments with the developing demand to harness undiscovered capability of waterway water assets in numerous parts of the world, the need to devise new methodologies and techniques for appraisal of water assets from these sources is additionally expanding. A typical serious issue regularly being looked by specialists to recreate the progression of the ungauged catchment. Typically such ungauged catchments are situated in headwater regions. Nonappearance of any historical information records or unimportant deficiency of whatever data is accessible for catchments in headwaters is by and large brought about by detachment, unwelcoming landscapes, and authentic absence of foreknowledge of organizers and designers to have potential locales measured for saddling the water asset of such catchments later on. Aside from the headwater regions, numerous potential destinations even in the downstream reaches additionally experience the ill effects of deficiency of site-explicit records of information. In certain nations, having wide and broad systems of gauge stations, data in many cases don't exist at areas where these are generally required. Absence of information, both subjective just as quantitative, regularly inhibits the undertaking of logical examinations for such catchments which are required for purposes, for example, surveying the water resource, guaranteeing its long haul accessibility, anticipating its event over short lead-times, predicting its future event, and building up its source.

1.3 Objective of the Study

Main objectives are:-

- To conduct trend analysis of the hydro metrological parameter
- Describe performance evaluation measures by hypothesis testing of the gauged data.
- Develop the Flow Duration Curve (FDC) at ungauged site for hydro power development from the data of gauged site.

CHAPTER 2

LITERATURE REVIEW

In Himalayan regions, it becomes very difficult to estimate rainfall- runoff catchment for given region because many physiographical and climate condition changes with time and space. At high elevation snow melting contributes the major parts to the runoff in stream. There may be a precipitation continuous increases with altitude, and it starts begin to decrease at a particular altitude (Singh et al. 1995, 1997).

Common streams speak to or describe the hydrological conduct of a basin and can be decided dependent on memorable hydrometric records and the mass preservation condition, through the use of precipitation overflow hydrological models or through measurable models (TNRCC, 1997). The assurance of the time arrangement of common moves through measured authentic streams comprises of evacuating the impacts of the human action that utilized water and of the water powered foundation that put away or inferred it; this incorporates extractions, return streams and capacity impacts (Wurbs, 2005).

Develop regional hydrologic model to estimate flow duration curve at gauged and ungauged catchments of massachusets regions using 23 sub basins. Regionalisation based on regression is used to relate log normal parameters of gauged catchments to their basin characteristics. Only basin relief and watershed area is used to develop relations and to develop regional flow duration model. Resulting regional flow duration curve seems to be satisfactorily at ungauged sites, considering the simplicity of model (Fennessey et al., 1990)

Ries and Friesz (2000) related physical attributes of basins to choose exceedance probabilities related with low streams. Streams at the 50, 60, 70, 75, 80, 85, 90, 95, 98, and 99 percent exceedance probabilities were relapsed against bowl qualities, for example, waste zone and percent of sand and rock stores in the bowl. The subsequent 10 conditions give a way to appraise stream flows not exactly or equivalent to the middle stream at ungauged regions in Massachusetts.

Dhar et al. (2000) studied on the regions of the Himalayas for high altitude and revealed that near the foot of the hills and at plain regions, precipitation is high. They showed that precipitation is start increasing with its altitude but at some elevation precipitation start decreasing with altitude. Thus the altitude or elevation characteristics play an important role in calculation of the flow data in the regional analysis.

See and Abrahart (2001) utilized an amalgamation of neural system, fluffy rationale, measurable, and ingenuity figures to create a solitary anticipated yield following a multi-model information combination way to deal with hydrological estimating. Working inside a multi-model system, Aspinall (2002) utilized numerous models of land use examples to draw deduction from a lot of time-variation models. With this developing proof of the benefits of the multi-model methodology, its application to stream catchment in an ungauged catchment is viewed as deserving of investigation.

Chiang et al. (2002) proposed hydrologic regionalisation term for classification catchment and applied a time series model for hydrologic regionalisation. Regionalised model for forecasting flow duration curves have been derived by them.

A FDC can be effectively gotten from measured stream flow information at an everyday or month to month timescale. The information are positioned in dropping request and each arranged esteem is associated10with an exceedance likelihood, for instance, through a plotting position recipe. The lack of stream gauges and the constrained measure of streamflow perceptions characterises several geological territories around the globe and, starting here of view, Sicily isn't an exception. This condition prompted the plan and proposition of various procedures for regionalizing FDC, whose point is the estimation of FDC at ungauged waterway basins15or the improvement of experimental FDC determined for stream gauges where just a limited amount of hydrometric data is accessible. An unpleasant classification of the available regionalization strategies recognizes two methodologies: factual and parametric. The first strategy considers FDC as the supplement of the aggregate frequency distribution of streamflows, while the second one doesn't make any association be-20tween FDC and the likelihood hypothesis (Castellarin et al., 2004).

Merz and Bloschl (2004) estimated the daily flow in an ungauged catchments of Austria region. 308 catchments are taken ranging from 3 to 5000 km^2. Multiple regression is estimate of the all model parameters couple with HBV hydrologic model. Uncertainty analysis is done using NSE which ranges from 0.63 to 0.67.

A.Bardossy (2007) used regionalisation methods for transferring hydrological model parameters from the gauged catchment to ungauged catchment. One can accept that watershed with alike characteristics display an alike hydrological behaviours and therefore, can be modelled using alike model parameters. Parameter sets could be considered as movable if the equivalent model performance (NSE) on the donor and study area catchments are good. Finally, the results indicates that the parameters transferred based on the above principles accomplish well on the target watersheds.

Xu et al. (2009) used swat hydrological model to model sediment yield and runoff in Miyun river basin, china. The model accurately estimate the sediment yield and runoff on monthly and daily basis with NSE value greater than 0.6. The sensitivity analysis performed to identify parameters which affect runoff and sediment yield from the water shed shows that runoff was sensitive to channel re-entrained linear parameters and curve number. The sensitivity parameter results is catchment specific and should not be applied directly to other catchments with different characteristics before conducting sensitivity analysis.

Margaret et al. (2010) applied two regionalisation methods (regression and global mean method) as mechanism for improving hydrological parameters then the model could be practically used in ungauged catchments in Arkansas. These two methods have been selected because they belong to the usual approaches. The resulting parameters were verified and the model performance was evaluated on three gauged watersheds. Finally the results indicate that the method was possible to gain transferred SWAT hydrological Model parameters sets for apply in watershed for which , calibration and validation due to missing monitoring data can't be performed.

Various power plants of various limits in INDIA are under development and their finishing will result in the expansion of the absolute level of tapped vitality. The estimation of limit of a hydropower plant is the aftereffect of definite examination and examination of the hydrological information like verifiable release time-series the place records of adequate length are accessible. however, these records are just accessible where the required information have been seen at measuring stations, and water asset appraisals are regularly required for ungauged locations. This makes a requirement for estimation of hydrological measurements (Flow Duration Curve, FDC) at ungauged sites (Booker and Sn elder, 2012).

Shi et al. (2013) used hydrological model to catchment of Xixian to determine water balance using three different uncertainty analysis and calibration methods and used to establish the model. Finally, the results displayed that hydrological mode; in the catchment are runs well, water balance evaluation of the basin indicates the base flow is an essential aspect of the total flow in the catchment. The result of hydrological model could be applied to further evaluation of the effects of land-use and climate variations and to examine the effects of the different management scenarios and cultivation styles on local water resources.

Singh et al. (2013) applied SWAT hydrological model in tungbhadra Catchment, India for stream flow determination. The model gave excellent results for monthly calibration time steps and good results for daily calibration time step between the observed and simulated data.

Emam et al. (2016) implemented SWAT model in the Aluoi district in central Vietnam to evaluate land resource management and planning as well as water. They applied a ratio method of regionalisation for transmission of recorded data obtain from the donor watershed to Aluoi. After the stream data in the gauged watershed determined and calibrates the model. According to the outcomes, the surface runoff is great due to the watershed.

Regional estimation of stream term qualities is significant for water assets advancement at the little catchment scale. Local examinations regularly neglect to sufficiently speak to the changeability of the stream routine in little catchments (<50 km ^{2}), particularly in remote mountains areas where the adjustment information are inadequate and allude to a lot bigger catchment scales. This examination recommends a methodology wherein territorial information are joined with real immediate stream information to develop an agent day by day stream term bend for little catchments. A provincial dimensionless stream length bend (FDC) is produced for a hydrologically homogeneous region in Western-Northwestern Greece and used to assess the FDC in two little sloping catchments inside the area. Various immediate stream estimations accessible at the two regions are utilized in a statically representation of the flow regime from which a gauge of the mean yearly stream at the destinations is made, permitting the development of the FDC from the local bend. Results got are in great concurrence with watched information and show huge estimation improvement over different techniques generally utilized in the investigation region. A sensitivity analysis utilizing Monte Carlo simulation is performed to set up sensible inspecting requirement for little ungauged catchments in the investigation region and similar mountainous territories in the Mediterranean region.

CHAPTER 3

STUDY AREA

3.1 Description of Study Area

The Upper Krishna sub-basin, also known as K1 sub-basin, lies between $18.06^{o}$to 16.04^{o} N latitude and $73.45^{o}$to 74.86^{o}E longitude. This sub-basin is surrounded by the Western Ghats in the west, the K5 sub-basin in the north and north-east, K3 sub-basin in the east, and K2 sub-basin in south. The total drainage area of the K1 sub-basin is 17912sq. km, of which about 95% falls in Maharashtra State and rest 5% falls in the Karnataka State. Upstream end of K1 sub-basin has the source of Krishna River and the downstream end has its confluence with Dudhganga. Koyana, Panchganga and Dudhganga are the major tributaries of the Krishna River that lie in the K1 sub-basin. Other major tributaries joining Krishna River in K1 sub-basin include Kudli, Vena, Urmodi, Tarali, Morna, Wang, Verna, Kadvi, Kasari, Dhamani, Tulsi, Bhogawati, Vedganga and Chikotra.

The average annual runoff in the K1 sub-basin is estimated to be 18.3 km^3. As per the 2011 Census, total population of the sub-basin is 7.92 million with a density of 440 persons per sq. km; and 23%of the population lives in urban areas. Location of G&D sites over line diagram of K1 sub-basin.

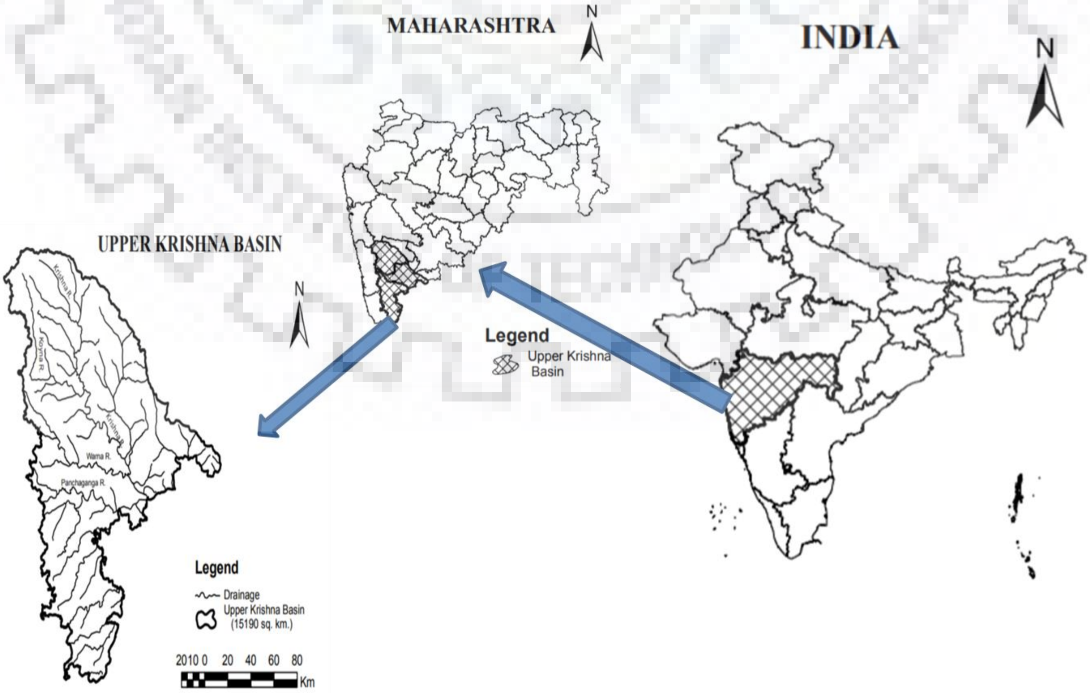


Figure 3.1 location of study area

3.2 Climate and hydro-meteorological network in K1 Sub-basin

Average annual rainfall in the sub-basin is 1584 mm. However, most of the area of K1 sub-basin falls in 1099 mm rainfall zone. Normally 85% of annual rainfall occurs during June to September due to south-west monsoon, 10% between September and December due to north-east monsoon and 5% between December and May. Rainfall decreases from west to east. Maximum temperature of the basin ranges between 40-42ºC while minimum temperature ranges between 6–8ºC.

There are 23 rain-gauge stations of IMD in/around the K1 sub-basin. Further, there are 10 Gauge-Discharge (GD) stations in the sub-basin: 5 GD stations have been set up by the Maharashtra State and rest 5 GD stations have been set up by the Central Water Commission (CWC) and their location is shown in Figs 3.2 -3.3

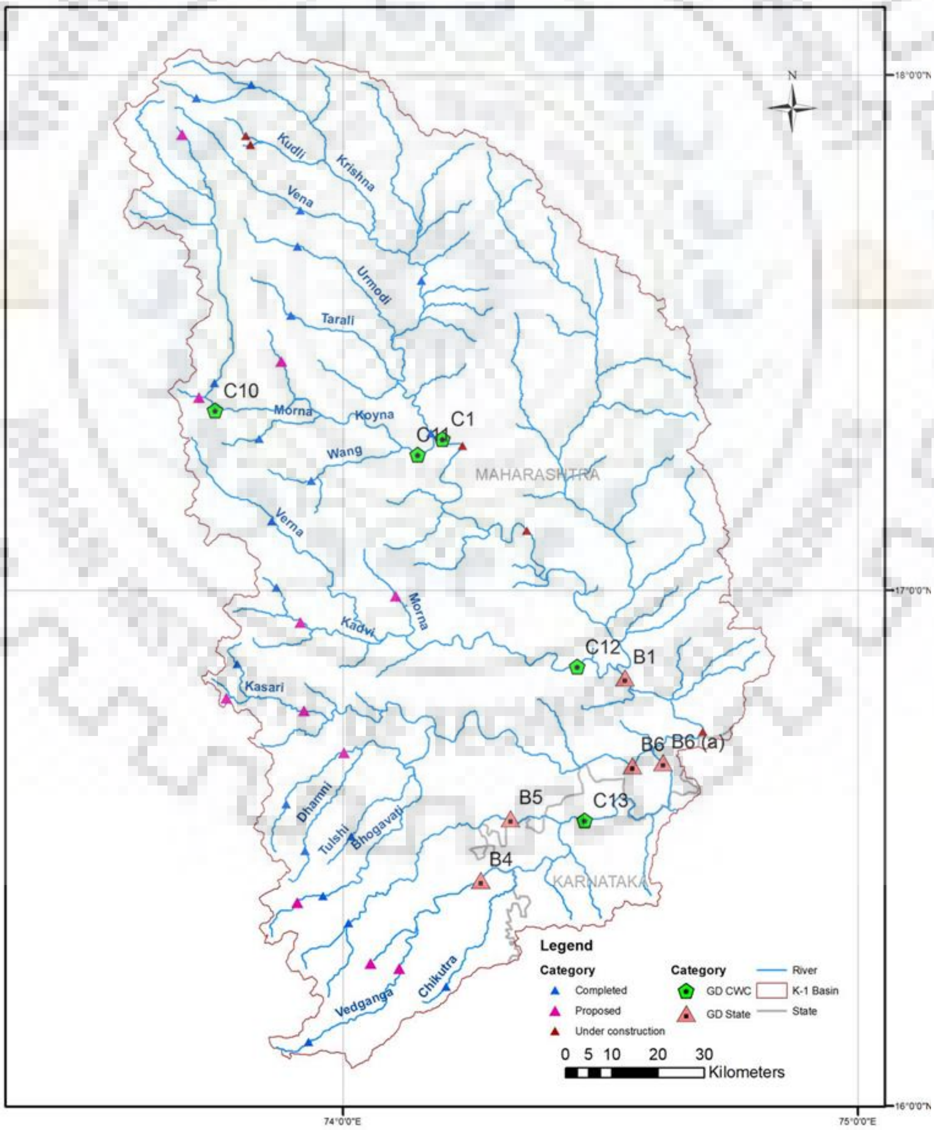


Figure 3.2 a map showing various gauging sites

Schematic line diagram of k1 sub basin is shown in fig 3.4 with its various lift irrigation schemes and gauging stations. Radhanagari, Koyna Hydel, Tulashi and Krishna Irrigation Project (including Dhom, Kanher and Jihe Kathapur) are the major irrigation projects in the K1 sub-basin. In addition, there are 13 medium irrigation projects in the sub-basin including Khodashi, Warna, Kadavi, Kasari, Kumbhi, Dudhganga, Vedganga, and Chikotra. There are four lift irrigation schemes (LIS), namely JiheKathapur, Tembu, Takri, and Mhaisal that are planned in the sub-basin is show in the Fig 3.3.

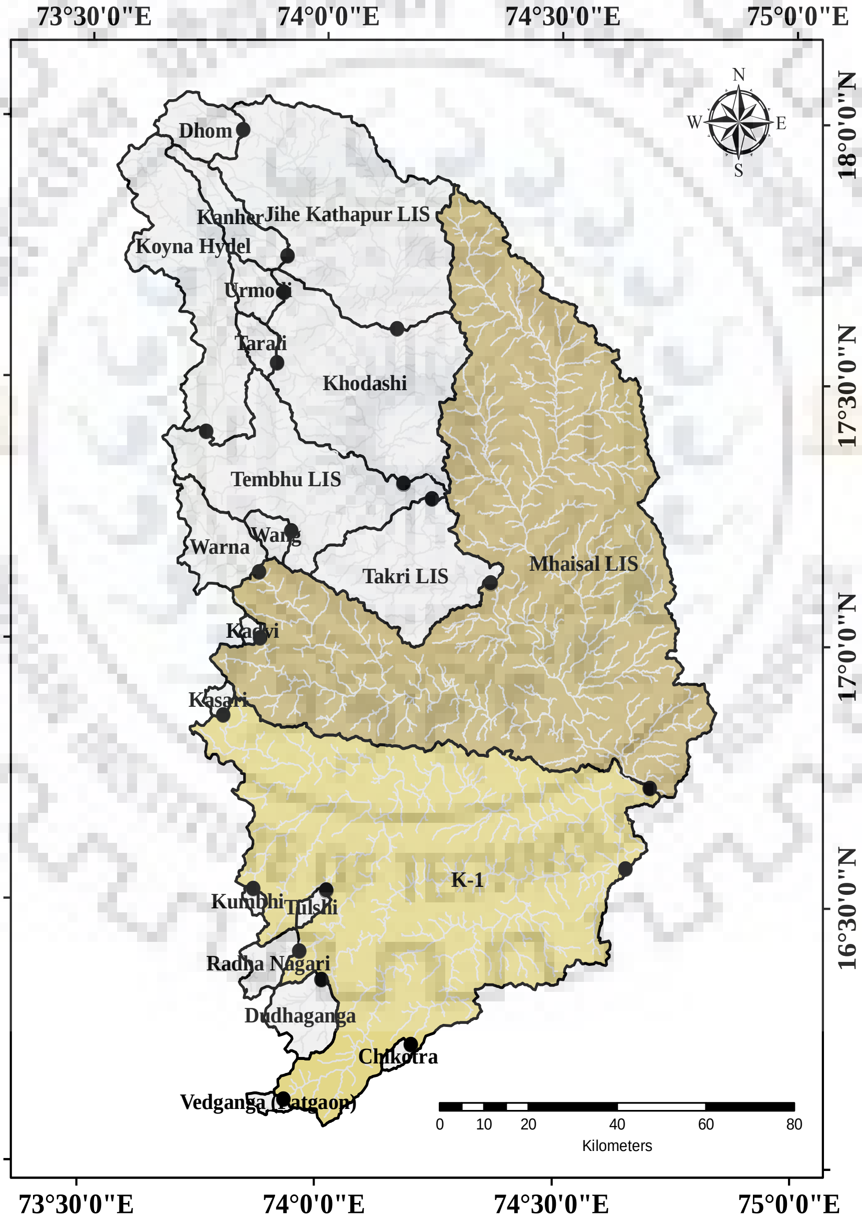


Figure 3.3 Location of various projects inK1 sub-basin and their catchments

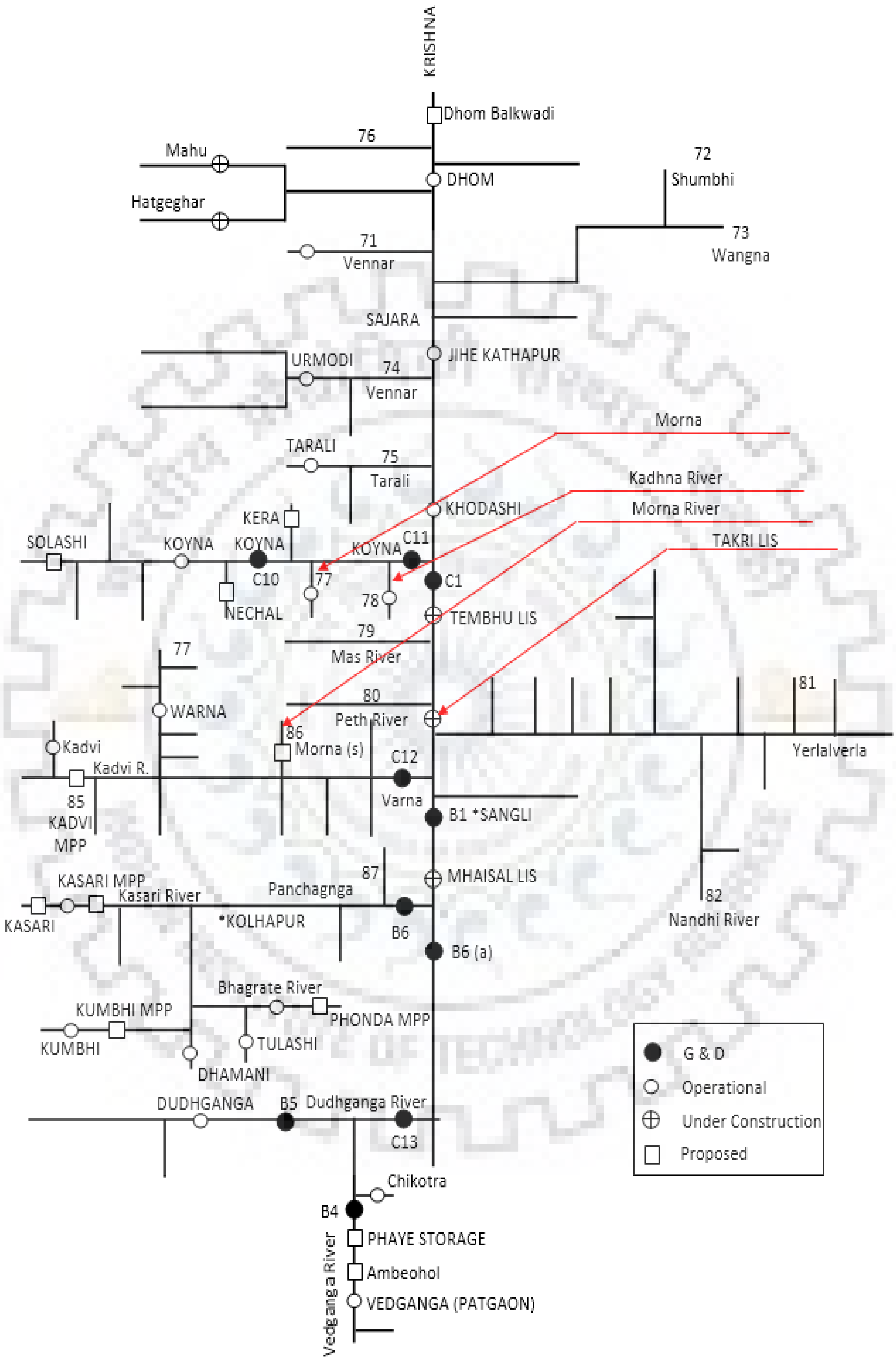


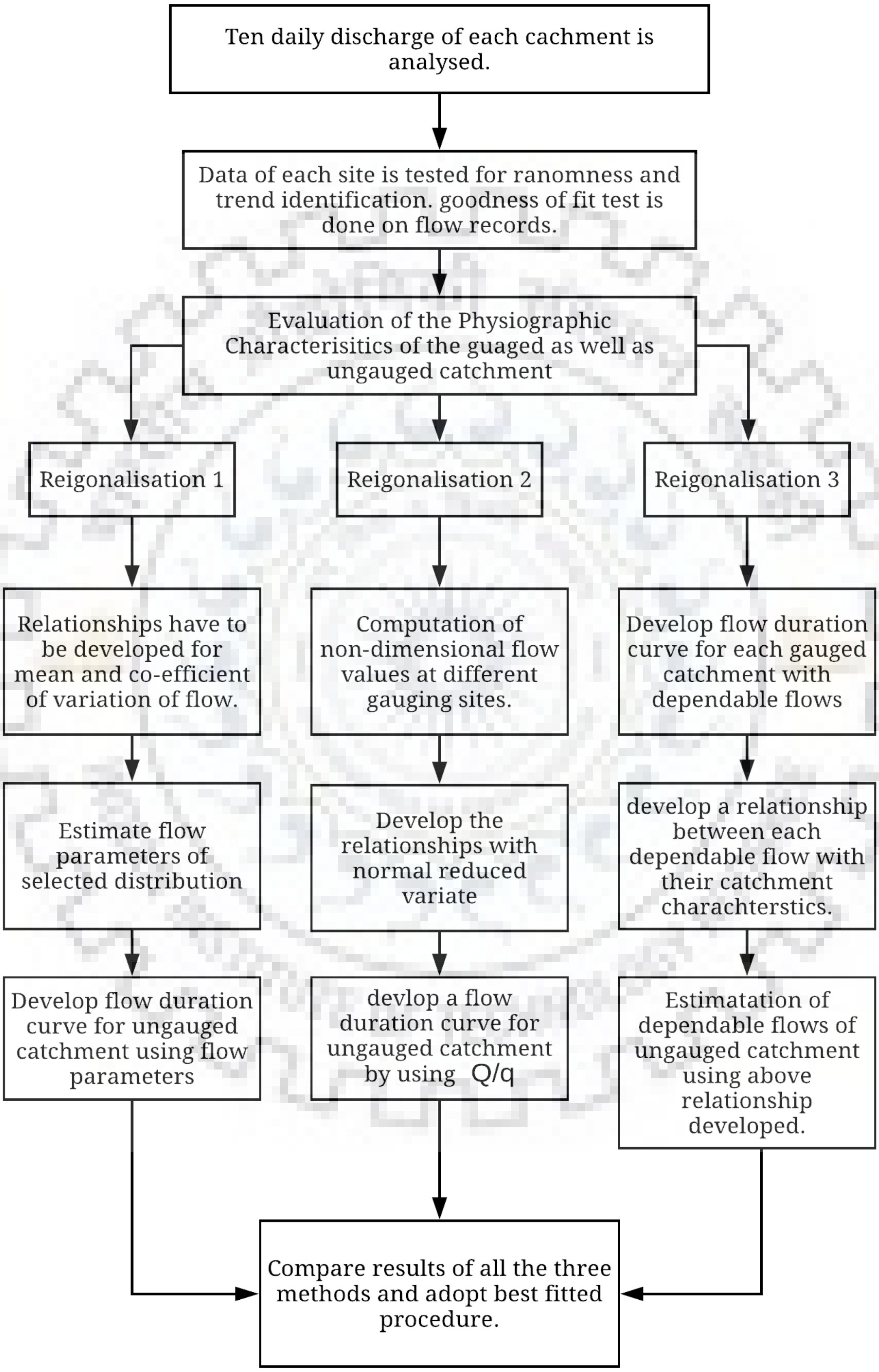
Figure 3.4 Line diagram of K1 sub basin

Table 3.1 Gauged Station with Their Catchment Area

Project name	Independent catchment area (sq. km)
Dhom	194.359
Kanher	211.915
Urmodi	112.497
JiheKathapur LIS	1632.625
Tarali	80.323
Koyna Hydel	898.366
Khodashi	1257.430
Wang	70.963
Tembhu LIS	1004.188
Warna	275.859
Kadvi	23.723
Takri LIS	709.001
Kasari	32.985
Mhaisal LIS	6208.819
Kumbhi	20.971
Tulshi	34.754
RadhaNagari	108.588
Dudhaganga	204.234
Chikotra	29.071
Vedganga	29.653
K-1	4771.571
Total	**17911.895**

Radhanagari, Koyna Hydel, Tulashi and Krishna Irrigation Project (including Dhom, Kanher and Jihe Kathapur) are the major irrigation projects in the K1 sub-basin. In addition, there are 13 medium irrigation projects in the sub-basin including Khodashi, Warna, Kadavi, Kasari, Kumbhi, Dudhganga, Vedganga, and Chikotra. There are four lift irrigation schemes (LIS), namely JiheKathapur, Tembu, Takri, and Mhaisal that are planned in the sub-basin.

3.3 Flowchart



CHAPTER 4

METHODS

4.1 Data Analysis

The analysis of annual data series for randomness and trend identification has been done using Turning point test and Kendall's rank correlation test respectively. The two tests used are given below:

4.1.1 Turning Point Test

This test counts the number of turning points (peaks and troughs) in a sequence. Let P be the total no of peaks and troughs. To calculate the test statistic the number of samples tested needs to be large. This allows for the assumption of a normal distribution with a mean of E (P):

$$E(p) = 2 \times \frac{(N\text{-}2)}{3} \tag{4.1}$$

And variance of

$$Var(p) = \frac{16N - 29}{90} \tag{4.2}$$

The test characteristic T can then be calculated with the following equation

$$T = \frac{\left(P - E(p)\right)}{Var(p)^{0.5}} \tag{4.3}$$

The hypothesis of independence is accepted at the (5%) probability level if *IZI* <1.96

4.1.2 Kendall's Rank Correlation Test

This test which is also known as T test is based on the proportionate umber o subsequent observations, which exceed a particular value. The test statistics T is computed by:

$$T = \left(\frac{4P}{N(N-1)}\right) - 1 \tag{4.4}$$

$$Var(T) = \left(\frac{2(2N+5)}{9N(N-1)}\right) \quad (4.5)$$

$$Z = \frac{T}{Var(T)^{0.5}} \quad (4.6)$$

The hypothesis of independence is accepted at the (5%) probability level if IZI <1.96

4.2 Goodness of Fit Test

The goodness of fit of a statistical model (distribution) describes how well it fits a set of observations. The tests used are given below:

4.2.1 Chi-Square Test

It is a goodness of fit test is useful in choosing about the applicable of a particular distribution. The chi-square formula used is

$$X^2_{computed} = \sum_{j=0}^{n} \frac{(O_j - Ej)^2}{E_j}$$

Where, n = number of classes, O_j= observed frequency of j_{th} class, and E_j = expected frequency of j_{th} class. The number of classes is selected in such a way as there are at least 4-6 observations in each class. The number of classes should not be more than 20 and lesser than 5. The classes can be divided in two ways

(i) Equal probability (ii) Equal interval.

In case of equal probabilities:

$$E_j = \frac{Number\ of\ observation\ (N)}{number\ of\ classes (n)}$$

$X^2_{computed}$ is compared with $X^2_{critical}$

$$X^2_{critical} = X^2_{(1-\alpha),(n-p-1)}$$

P= number of parameters of distribution, (m-p-1) = number of degrees of freedom. At 5% significance level:

$$X^2_{critical} = X^2_{(0.95),(n-p-1)}$$

If $X^2_{computed} < X^2_{critical}$ then the distribution can be assume to fit well.

For $X^2_{critical}$ standard tables are available.

STEPS:

1. Select no. of classes, N classes, and $5 \leq N$ classes' ≤ 20.
2. Compute E_j.
3. Compute probability limits of each class.
4. Compute reduced variates corresponding to these probabilities.
5. Compute X corresponding to these probabilities.
6. Compute O_j and X^2.
7. Apply the test.

4.2.2 Kolmogorov-Smirnov (KS) Test

The Kolmogorov-Smirnov test is intended to test the theory that a given informational index could have been drawn from a given appropriation. Not at all like the chi-square test, it is basically planned for use with persistent disseminations and is free of subjective computational decisions, for example, receptacle width. The Kolmogorov-Smirnov (K-S) test is based on the empirical distribution function (ECDF).Given N ordered data points Y_1, Y_2,..Y_n, the ECDF is defined as

$$Fe(Yi) = \frac{n(i)}{N}$$

Where *n(i)* is the number of points less than Y_i and the Y_i are ordered from smallest to largest value. This is a step function that increases by 1/N at the value of each ordered data point. The complete formula is as under:

$$D = Max\,\{Fe(Yi) - FD(Yi)\}$$

Where $F_D\,(Y_i)$ is the CDF from distribution. If D < d_α(N) then the distribution is selected. At 5% significance level:

$$d_\alpha = \frac{1.36}{\sqrt{N}}$$

4.2.3 D-Index Test

The D-Index for comparison of the fit of various distributions in the upper tail is given as:

$$D - index = \left(\frac{1}{x}\right)\sum_{i=1}^{6} Abs(x_i - \widehat{x_i})$$

Where x_i and $\widehat{x_i}$ are the ith highest observed and computed values for the distribution. The distribution giving the least index is considered to be the best fit distribution.

STEPS

1. Arrange the values in descending order.
2. Select the highest six values.
3. Compute the probability of exceedance.
4. $P = \frac{m}{n+1}$
5. Either compute of select from the table the value of frequency factor corresponding to probability of exceedance and coefficient of skewness.
6. Compute the value of $\widehat{x_i}$from the normal random variable equation of given distribution.
7. Compute the value of D-Index.
8. Select the distribution having the lowest value of D-Index.

The computer programmes of all these tests have been written in octave language, and the data is tested through those programmes.

CHAPTER 5

METHODS USED FOR DEVELOPING FLOW DURATION CURVE

5.1 DEVELOPMENT OF FLOW DURATION CURVE (FDC)

5.2 Without physiographic catchment parameters

5.2.1.1 Parametric flow duration curve

This method is used where stream flow data does not vary directly with catchment area of the region. The procedure is to make plots of FDC's for a gauged stream with in a rather homogenous drainage basin. From these flow duration curves are developed a family of parametric duration curves in which flow (q) is plotted against average annual runoff (Q) or average annual discharge (q), at the respective gauges for several exceedance percentage. A separate curve is developed for each exceedance percentage. After obtaining best fitting curve for the stream data from the historical records of the stream flow, a correlation analysis is done. The result is a parametric flow duration curve.

5.2.1.2 Dimensionless flow duration curves

In this method each dependable exceedance flow divided by the average annual discharge which is obtain by historical data to get dimensionless flow term. Then this term data is plotted against exceedance interval on log-scale to get a dimensionless FDC. At that point a best fitting curve is created for a specific region having homogeneous hydrology with the goal that a solitary curve results that relates a trademark dimensionless flow term to the exceedance rate. It is easy to perceive that every one of the breaking points of the curve leaves something to be desired in light of the fact that the quantity of qualities are the uncommon event of flash floods or low flows.

5.2.2 With physiographic catchment parameters

For ungauged catchments, regional FDC's are developed based on the available rainfall-runoff records for the gauged catchment of the region which is considered to be hydro meteorologically homogenous. The regional FDC is used to estimate the dependable flows for the ungauged catchment. Different methods for developing the regional FDC's have been developed by various investigators. Out of those three methods are outlined here under.

Method 1: regionalization of the parameters of chosen probability distribution for individual gauged sites.

The Flow duration curve for the measured catchment might be linearized by fitting either ordinary distribution or log normal distribution. Hypothetically any appropriation might be fitted to linearize the stream term curve and best fit distribution might be picked for further application. Let the parameters of the liquid dispersion are P1, P2... Pn. At that point these parameters might be assessed for the checked catchment of the locale and the local relationship might be created relating the parameters with the physiographic and climatological qualities.

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

Sometimes either sufficient stream records for checked catchments are restricted. It makes the advancement of regional FDC utilizing method 1. Under these conditions it is viewed as inexact to make all the stream information of individual site non dimensional by isolating it by the mean stream happening at the site. Presently the non-dimensionalised stream information arrangement of all the checked destinations are clubbed together to give a basic arrangement speaking to the region. While clubbing them together it is assumed that these non-dimensional stream arrangement for each measured site in the region is a basic drawn from a similar population, a solitary arrangement therefore got from the area might be examined and picked likelihood distribution might be fitted. It results in the parameter of the distribution which might be considered as the local parameters. Presently these parameters might be utilized to get the non-dimensionalised stream for any dimension of parameter. So as to build up the stream span bend for an ungauged catchment the gauge for mean stream is required. It might be duplicated by the non-dimensionalised stream esteems for the required FDC. It requires the improvement of a territorial relationship for the mean stream relating the estimations of the mean progression of the checked catchment with their catchment qualities.

Method 3: Regionalization of the dependable flows.

Sometimes one may interest in the evaluation of dependable flows corresponding to the limited number of probability of exceedance (say only for 50% and 90% dependable flows. In such a situation the dependable flows itself may be regionalized rather than regionalizing the FDC. For the accuracy of this method adequate number of gauging sites having the fellow series of the specific durations either method 1 or method 2, may be used to regionalize the FDC.

For the comparison of the results of the regionalisation method, percentage absolute error in dependable flows (PAEDF) are calculated for each of the three methods using the relationships developed.

$$PAEDF = \left|\frac{Q_{D} - \bar{Q}_{D}}{Q_{D}}\right| \times 100$$

Where Q_D represents the dependable flow corresponding to D% dependability computed from the historical daily flow data. $\bar{Q}_D$ Represents the percentage dependable flow calculated utilising any one of the three regionalisation methods.

The values of PAEDF are calculated for each of the 11 sub-basins corresponding to different dependability considered. And checked that which method has less percentages of error with observed flow has been taken as best suitable method for that particular region with their physiographic characteristics.

5.3 Methodology

The data of all the sites being analysed is ten daily. The steps described below were followed:

1. The first step is to analyse the data for its randomness and its existing trend. Turning point test and Kendall's rank correlation test is used to check the randomness and trend identification of annual data series respectively.
2. The next step is to find the most appropriate distribution fitting the data series through goodness of fit tests. The goodness of fit tests used is Chi-square test, Kolmogorov-Smirnov (KS) test and D-Index test. Some of the most common and important probability distributions used in hydrology are the normal, log-normal, gamma, Gumbel, and Weibull (AKSOY, 2000). Therefore the goodness of fit test is used to check the distribution against normal, log-normal, Gumbel, Pearson type-III and log Pearson type-III.
3. After fitting the appropriate distribution, regional relations for different parameters of the selected distribution are developed as a function of various catchment characteristics using linear regression approach. Then for the un-gauged sites the parameters of selected distribution are estimated using these regional relations. Estimated parameters are used to generate the flow duration curves for the un-gauged sites.
4. The power potential at different exceedance levels is estimated from the generated flow duration curve at each un-gauged site.

To do the data analysis and other tests computer programs were developed for use in octave programme and FORTRAN programme

CHAPTER 6

DATA ANALYSIS

The Turning point test and Kendall's rank correlation test are used to analyse the randomness and trend identification of annual data series respectively. The OCTAVE and FORTRAN program of these tests used to do the analysis is given in Appendix-I.

The goodness of test results shows that ten daily discharge and annually discharge data of the gauging stations follow 2 parameter log normal distribution like mean (μ) and standard deviation (σ) of the flow-data. By using this distribution we generated some equations and relations between flow and physical characteristics like catchments area (ca), altitude (A) and dimensionless parameter (Ca/A^2).

Analysis of regionalisation method -1

In this technique the accompanying connections have been created between the mean of stream data values in log space with the significant physiographic qualities of the ungauged catchment, for example, elevation (A), catchment territory (Ca) and Ca/A^2

$$\bar{Y} = 0.1019(Ca/A^2)^{0.2891} \qquad (R^2 = 0.9158) \qquad (6.1)$$

$$\bar{Y} = 0.0204(Ca/A)^{0.2721} \qquad (R^2 = 0.8871) \qquad (6.2)$$

$$\bar{Y} = 16.406(Ca)^{-0.149} \qquad (R^2 = 0.7877) \qquad (6.3)$$

These equations clearly shows that by the using of regionalization method 1, the most suited physiographic characteristics with flow is non dimensional parameter (Ca/A^2) with R^2 value more than 0.9. This indicates that most appropriate results we can get for ungauged catchment by this method is using equation 1.

Table 6.1 Basic statistics of ten daily mean flow data in Real Space

In real space				
Name	Mean (cumec)	S.D.	C.V.	Skewness
Koyna	109.47	212.19	1.94	2.86
Dhom	12.51	27.28	2.18	3.57
Kanher	10.60	24.81	2.34	3.91
Warna	35.53	76.17	2.14	3.10
Radhanagari	13.68	27.13	1.98	2.48

Dudhganga	24.29	52.48	2.16	2.88
Vedganga	4.59	9.25	2.01	2.88
Tulashi	2.42	5.20	2.15	3.15
Kumbhi	4.43	8.94	2.02	2.32
Kasari	5.40	12.46	2.31	4.94
Kadvi	3.83	8.98	2.35	4.53

Table 6.2 Basic statistics of ten daily mean flow data in log space

In log space				
Name	Mean (cumec)	S.D.	C.V.	Skewness
Koyna	3.40	1.72	0.51	0.27
Dhom	1.61	1.57	0.98	0.47
Kanher	1.74	1.58	0.91	0.24
Warna	2.58	1.91	0.74	0.07
RadhaNagari	2.67	1.47	0.55	-0.54
DudhGanga	2.91	1.74	0.60	-0.34
VedGanga	1.55	1.34	0.86	-0.10
Tulashi	1.14	1.15	1.00	0.20
Kumbhi	2.11	1.12	0.53	-0.58
Kasari	1.76	1.35	0.77	-0.09
Kadvi	1.77	1.21	0.68	-0.27

6.1 At Koyna

For the test of randomness and trend identification:-

6.1.1 Turning Point Test

Number of Peaks = 12

Number of Troughs = 11

Number of turning points = 23

The value of z is = 0.13521

Absolute z is lesser than 1.96, hence series is random at 5% significance level.

6.1.2 Kendall's rank correlation test

Value of p = 310

Test statistics = 0.13621

There is no trend in the data at 5% significance level.

6.2 At Dhom

For the test of randomness and trend identification:-

6.2.1 Turning Point Test

Number of Peaks = 12

Number of Troughs = 11

Number of turning points = 23

The value of z is = .13521

Absolute z is lesser than 1.96, hence series is random at 5% significance level.

6.2.2 Kendall's rank correlation test

Value of p = 365

Test statistics = 1.3621

There is no trend in the data at 5% significance level.

6.3 At Kanher

For the test of randomness and trend identification:-

6.3.1 Turning Point Test

Number of Peaks = 12

Number of Troughs = 11

Number of turning points = 23

The value of z is = 0.13521

Absolute z is lesser than 1.96, hence series is random at 5% significance level.

6.3.2 Kendall's rank correlation test

Value of p = 287

Test statistics = 0.76277

There is no trend in the data at 5% significance level.

Similar test results for other stations are as follows:

Table 6.3 Trend and Randomness test at other stations

Station	Turning point test	Kendall's correlation test
Warna	YES	No
RadhaNagari	YES	No
DudhGanga	YES	No
VedGanga	YES	No
Tulashi	YES	No
Kumbhi	YES	No
Kasari	YES	yes
kadvi	YES	No
Koyna	YES	No
Dhom	YES	No
Kanher	YES	No

6.4 Goodness of fit test

Detailed account of goodness of fit test at Koyna.

Expected frequency (E_J) = $N\ /no.of classes$ = 7.2.

Table 6.4 Chi-Square test at Koyna for normal distribution

Class	$P(X \leq x)$	χ_{T1}	χ_{T1}	χ_{T1}	χ_{T2}	O_j	$\chi_{computed}$
1	0.0-0.2	0.00	-0.476	0	2958	13	4.6722
2	0.2-0.4	-0.476	0.087	2958	3486.7	9	0.45
3	0.4-0.6	0.087	0.672	3486.7	4036.1	7	0.0055556
4	0.6-0.8	0.672	1.5	4036.1	4813.8	4	1.4222
5	0.8-1.0	1.5	0.00	4813.8	3405	3	2.45
						sum	9

At 95% confidence level the value of $\chi_{critical} = 5.99$. As $\chi_{computed} < \chi_{critical}$ hence log-normal distribution doesn't fits well.

6.4.1 Chi square Test

6.4.1.1 Chi Square Test for Log Normal Distribution at Koyna.

The mean of the series is = 8.0982

The standard deviation is = 0.2651

The skewness coefficient is = 0.28004

EJ = N/NO OF CLASSES = 7.2.

Table 6.5 Chi-Square test at Koyna for log-normal distribution

Class	$P(X \leq x)$	χ_{T1}	χ_{T1}	χ_{T1}	χ_{T2}	O_j	$\chi_{computed}$
1	0.0-0.2	0.00	-0.476	0	2958	8	0.089
2	0.2-0.4	-0.476	0.087	2958	3486.7	7	0.006
3	0.4-0.6	0.087	0.672	3486.7	4036.1	7	0.006
4	0.6-0.8	0.672	1.5	4036.1	4813.8	7	0.006
5	0.8-1.0	1.5	0.00	4813.8	3405	7	0.006
						sum	0.112

At 95% confidence level the value of $\chi_{critical} = 5.99$

As$\chi_{computed} < \chi_{critical}$, hence log-normal distribution is well fitted

6.4.1.2 Chi-square test for normal Pearson type-III distribution.

The mean of the series is = 3405

The standard deviation is = 939.16

The Skewness coefficient is = 0.8683

E_J = N/NO OF CLASSES = 7.2.

The result of the test is given in Table 6.6

Table 6.6 Chi-square test for normal Pearson type-III distribution.

Class	$P(X \leq x)$	χ_{T1}	χ_{T1}	χ_{T1}	χ_{T2}	O_j	$\chi_{computed}$
1	0.0-0.2	0	-0.838	0	2618	8	0.088889
2	0.2-0.4	-0.838	-0.429	2618	3002.1	7	0.0055556
3	0.4-0.6	-0.429	0.036	3002.1	3438.8	5	0.67222
4	0.6-0.8	0.036	0.719	3438.8	4080.3	9	0.45
5	0.8-1.0	0.719	0.00	4080.3	0	7	0.0055556
						sum	1.222223

At 95% confidence level the value of $\chi_{critical} = 3.8415$.

As $\chi_{computed} < \chi_{critical}$, hence normal pearson type-III distribution is well fitted.

6.4.1.3 Chi-square test for log -normal pearson type-III distribution.

The mean of the series is = 3405

The standard deviation is = 939.16

The Skewness coefficient is = 0.8683

EJ = N/NO OF CLASSES = 7.2.

Table 6.7 Chi-square test for log-normal Pearson type-III distribution.

Class	$P(X \leq x)$	χ_{T1}	χ_{T1}	χ_{T1}	χ_{T2}	O_j	$\chi_{computed}$
1	0.0-0.2	0.0	-0.838	0.0	7.876	8	0.088889
2	0.2-0.4	-0.838	-0.429	7.876	7.9844	5	0.67222
3	0.4-0.6	-0.429	0.036	7.9844	8.1077	5	0.67222
4	0.6-0.8	0.036	0.719	8.1077	8.2888	10	1.0889
5	0.8-1.0	0.719	0.00	8.2888	0.0	8	0.088889
						sum	2.6111

At 95% confidence level the value of $\chi_{critical} = 3.8415$

As $\chi_{computed} < \chi_{critical}$, hence log-normal pearson type-III distribution is well fitted.

Goodness of fit test for Gumbel distribution at Koyna

The mean of the series is = 3405

The standard deviation is = 939.16

EJ = N/NO OF CLASSES = 7.2.

Table 6.8 Goodness of fit test for Gumbel distribution at Koyna

Class	$P(X \leq x)$	χ_{T1}	χ_{T1}	χ_{T1}	χ_{T2}	O_j	$\chi_{computed}$
1	0.0-0.2	0	-0.476	0	1884	0	4.6722
2	0.2-0.4	-0.476	0.087	1884	2490.5	7	0.45
3	0.4-0.6	0.087	0.672	2490.5	2918.4	6	0.0055556
4	0.6-0.8	0.672	1.5	2918.4	3330.8	5	1.4222
5	0.8-1.0	1.5	0.00	3330.8	0.0	18	2.45
						sum	9

At 95% confidence level the value of $\chi_{critical} = 5.99$

As $\chi_{computed} > \chi_{critical}$,

Hence log-normal Pearson type-III distribution doesn't fits well.

6.4.2 Kolmogorov-Simonov (KS) test

K-S test for Gumbel distribution at Koyna

Table 6.9 k-s test for Gumbel distribution at Koyna

$X(i)$	Rank	$P(X \geq x)$	$FE(yi)$	$FD(yi)$	$D = FE(yi) - FD(yi)$
5718.90	1	0.03	0.97	0.92	0.05
5588.90	2	0.05	0.95	0.91	0.04
5500.10	3	0.08	0.92	0.90	0.02
4509.60	4	0.11	0.89	0.73	0.16
4356.10	5	0.14	0.86	0.70	0.17
4173.90	6	0.16	0.84	0.64	0.19
4157.10	7	0.19	0.81	0.64	0.17
3995.60	8	0.22	0.78	0.59	0.20
3973.40	9	0.24	0.76	0.58	0.18
3825.00	10	0.27	0.73	0.53	0.20
3753.60	11	0.30	0.70	0.50	0.20

3731.80	12	0.32	0.68	0.49	0.18
3711.20	13	0.35	0.65	0.49	0.16
3627.10	14	0.38	0.62	0.45	0.17
3476.40	15	0.41	0.59	0.40	0.20
3457.60	16	0.43	0.57	0.39	0.18
3435.50	17	0.46	0.54	0.38	0.16
3383.50	18	0.49	0.51	0.36	0.15
3319.40	19	0.51	0.49	0.33	0.15
3187.20	20	0.54	0.46	0.28	0.18
3183.70	21	0.57	0.43	0.28	0.15
2991.20	22	0.59	0.41	0.21	0.19
2962.60	23	0.62	0.38	0.20	0.18
2887.90	24	0.65	0.35	0.18	0.17
2839.90	25	0.68	0.32	0.16	0.16
2776.40	26	0.70	0.30	0.14	0.16
2760.10	27	0.73	0.27	0.14	0.13
2739.30	28	0.76	0.24	0.13	0.11
2579.00	29	0.78	0.22	0.09	0.13
2425.70	30	0.81	0.19	0.06	0.13
2347.30	31	0.84	0.16	0.05	0.12
2281.40	32	0.86	0.14	0.04	0.10
2261.00	33	0.89	0.11	0.03	0.07
2235.70	34	0.92	0.08	0.03	0.05
2232.80	35	0.95	0.05	0.03	0.02
2195.20	36	0.97	0.03	0.03	0.00

The highest value of D is = 0.20214

The value of $d(alpha)$ = 0.22667

As $d\ (max) < d(alpha)$, therefore the Gumbel distribution is fitted well.

K-S test for normal distribution at Koyna

Table 6.10 k-s Test for normal distribution at Koyna

$X(i)$	Rank	$P(X \geq x)$	$FE(yi)$	$FD(yi)$	$D = FD(yi) - FD(yi)$
5718.90	1.00	0.03	0.97	0.92	0.05
5588.90	2.00	0.05	0.95	0.91	0.04
5500.10	3.00	0.08	0.92	0.90	0.02
4509.60	4.00	0.11	0.89	0.73	0.16
4356.10	5.00	0.14	0.86	0.70	0.17
4173.90	6.00	0.16	0.84	0.64	0.19
4157.10	7.00	0.19	0.81	0.64	0.17
3995.60	8.00	0.22	0.78	0.59	0.20
3973.40	9.00	0.24	0.76	0.58	0.18
3825.00	10.00	0.27	0.73	0.53	0.20
3753.60	11.00	0.30	0.70	0.50	0.20
3731.80	12.00	0.32	0.68	0.49	0.18
3711.20	13.00	0.35	0.65	0.49	0.16
3627.10	14.00	0.38	0.62	0.45	0.17
3476.40	15.00	0.41	0.59	0.40	0.20
3457.60	16.00	0.43	0.57	0.39	0.18
3435.50	17.00	0.46	0.54	0.38	0.16
3383.50	18.00	0.49	0.51	0.36	0.15
3319.40	19.00	0.51	0.49	0.33	0.15
3187.20	20.00	0.54	0.46	0.28	0.18
3183.70	21.00	0.57	0.43	0.28	0.15
2991.20	22.00	0.59	0.41	0.21	0.19
2962.60	23.00	0.62	0.38	0.20	0.18
2887.90	24.00	0.65	0.35	0.18	0.17
2839.90	25.00	0.68	0.32	0.16	0.16
2776.40	26.00	0.70	0.30	0.14	0.16
2760.10	27.00	0.73	0.27	0.14	0.13
2739.30	28.00	0.76	0.24	0.13	0.11
2579.00	29.00	0.78	0.22	0.09	0.13

2425.70	30.00	0.81	0.19	0.06	0.13
2347.30	31.00	0.84	0.16	0.05	0.12
2281.40	32.00	0.86	0.14	0.04	0.10
2261.00	33.00	0.89	0.11	0.03	0.07
2235.70	34.00	0.92	0.08	0.03	0.05
2232.80	35.00	0.95	0.05	0.03	0.02
2195.20	36.00	0.97	0.03	0.03	0.00

The highest value of d is = 0.203

The value of d $_{(alpha)}$ = 0.227

6.4.3 D-index test

D index test for normal distribution.

The mean of the series is = 3405

The standard deviation is = 939.16

$$X(i) = EXP\ (mean + KT * standard\ deviation)$$

Table 6.11 D-index test for normal distribution at Koyna

$Rank$	X	$P(X >= x)$	K_T	$X(i)$	$ABS[X - X(i)]$
1	5718.9	0.027027	1.9264	5214.2	504.64
2	5588.9	0.054054	1.6068	4914	674.86
3	5500.1	0.081081	1.3978	4717.8	782.26
4	4509.6	0.10811	1.2367	4566.4	56.831
5	4356.1	0.13514	1.1024	4440.4	84.252
6	4173.9	0.16216	0.98561	4330.7	156.82
sum					**2259.7**
D- index = 0.66363					

D index test for log-normal distribution

The mean of the series is = 8.0982

The standard deviation is = 0.2651

$$X(i) = EXP\ (mean + K_T * standard\ deviation)$$

Table 6.12 D-index test for log-normal distribution at Koyna

$Rank$	X	$P(X \geq x)$	K_T	$X(i)$	$ABS[X - X(i)]$
1	8.6515	0.027027	1.9264	8.6089	0.042671
2	8.6285	0.054054	1.6068	8.5241	0.10442
3	8.6125	0.081081	1.3978	8.4687	0.14379
4	8.414	0.10811	1.2367	8.426	0.012037
5	8.3793	0.13514	1.1024	8.3904	0.011081
6	8.3366	0.16216	0.98561	8.3595	0.022858
				sum	**0.33685**
D- index = 0.66363					

D-index test for Gumbel distribution at Koyna

The mean of the series is = 3405

The standard deviation is = 939.16

$$X\ (i) = mean + alpha * X_T$$

Table 6.13 D-index tet for Gumbel distribution at Koyna

$Rank$	X	$P(X \geq x)$	K_T	$X(i)$	$ABS[X - X(i)]$
1	5718.9	0.027027	1.9264	5214.2	504.64
2	5588.9	0.054054	1.6068	4914	674.86
3	5500.1	0.081081	1.3978	4717.8	782.26
4	4509.6	0.10811	1.2367	4566.4	56.831
5	4356.1	0.13514	1.1024	4440.4	84.252
6	4173.9	0.16216	0.98561	4330.7	156.82
				sum	**2259.7**
D- index = 1.442127					

D-index test for Pearson type-III at Koyna

Table 6.14 D-index test for Pearson type-III at Koyna

$Rank$	X	$P(X \geq x)$	K_T	$X(i)$	$ABS[X - X(i)]$
1	5718.9	0.027027	2.2388	5507.591	211.3086
2	5588.9	0.054054	1.724	5024.112	564.7882
3	5500.1	0.081081	1.394	4714.189	785.911
4	4509.6	0.108108	0.8	4156.328	353.272
5	4356.1	0.135135	0.5718	3942.012	414.0883
6	4173.9	0.162162	0.5332	3905.76	268.1399
sum					**2597.608**
D- index = 0.765821					

D-index test for log-Pearson type-III at Koyna

Table 6.15 D-index test for log-Pearson type-III at Koyna

$Rank$	X	$P(X \geq x)$	K_T	$X(i)$	$ABS[X - X(i)]$
1	5718.9	0.027027	2.461	5716.273	2.62724
2	5588.9	0.054054	1.92	5208.187	380.7128
3	5500.1	0.081081	1.465	4780.869	719.2306
4	4509.6	0.108108	1.265	4593.037	83.4374
5	4356.1	0.135135	0.953	4300.019	56.08052
6	4173.9	0.162162	0.84	4193.894	19.9944
sum					**1262.083**
D- index = 15.58					

The Chi-Square Test and KS Test are not amazing tests as in their likelihood of tolerating the hypothesis when it is in actuality false is high when these tests are utilized. For this reason we take the results of D-Index test which indicates that Log-Normal distribution is the best suited distribution that can represent the data series of koyna site. As the data available for other sites is less than the minimum data needed for Chi-Square testing, so it has been omitted.

The goodness of fit tests for other sites are repeated in the same way as has been done for the first site. Only the final result for other sites is given in these Tables 6.16 to 6.25.

Brief account of Goodness of fit test at various gauged sites.

Table 6.16 Goodness of fit test at Koyna

Site at Koyna			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	NO
Log-Pearson type-III	YES	-	NO
Gumbel	NO	YES	NO

Table 6.17 Goodness of fit test at Dhom

Site at Dhom			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	YES
Log-Pearson type-III	YES	-	NO
Gumbel	NO	YES	NO

Table 6.18 Goodness of fit test at Kanher

Site at Kanher			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	YES
Log-Pearson type-III	YES	-	YES
Gumbel	NO	NO	NO

Table 6.19 Goodness of fit test at Warna

Site at Warna			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	NO
Log-Pearson type-III	YES	-	NO
Gumbel	NO	NO	NO

Table 6.20 Goodness of fit test at RadhaNagri

Site at RadhaNagari			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	YES
Log-Normal	NO	NO	YES
Pearson Type-III	NO	-	YES
Log-Pearson type-III	NO	-	NO
Gumbel	NO	YES	NO

Table 6.21 Goodness of fit test at Dudhganga

Site at Dudhganga			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	YES
Log-Normal	NO	NO	YES
Pearson Type-III	NO	-	NO
Log-Pearson type-III	NO	-	NO
Gumbel	NO	YES	NO

Table 6.22 Goodness of fit test at Vedganga

Site at Vedganga			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	NO
Log-Pearson type-III	YES	-	YES
Gumbel	NO	YES	NO

Table 6.23 Goodness of fit test at Kumbhi

Site at Kumbhi			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	NO
Log-Pearson type-III	YES	-	NO
Gumbel	NO	NO	NO

Table 6.24 Goodness of fit test at Kasari

Site at Kasari			
RESULTS	Chi-square test	KS test	D-index test
Normal	YES	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	YES	-	YES
Log-Pearson type-III	YES	-	YES
Gumbel	NO	NO	NO

Table 6.25 Goodness of fit test at Kadvi

Site at Kadvi			
RESULTS	Chi-square test	KS test	D-index test
Normal	NO	NO	NO
Log-Normal	YES	NO	YES
Pearson Type-III	NO	-	NO
Log-Pearson type-III	YES	-	NO
Gumbel	NO	NO	NO

CHAPTER 7

ANALYSIS AND RESULTS

The different physiographic characteristics of the distinctive checking destinations have been registered and dissected. The linear co-relation between relief (H) and altitude (A) is 0.84. It demonstrates a high co-relation between the elevation and relief. Catchment region, relief and elevation are considered for territorial investigation. A non-dimensional measure Ca/A^2 is determined for every one of the sub-basins. This non dimensional measure is considered as an autonomous variable for relapse. The qualities for checked catchment of Ca/A^2 is shift from 91 to 3150. The fundamental measurements, for example, standard deviation, mean, coefficient of variation and coefficient of skewness determined from the accessible chronicled information at individual checking regions in real as well in log space are given in table 1 and table 2. The standard deviation and mean of the stream information change from 9 to 212 and from 2.5 cumec to 109.5 cumec, separately. The estimations of the coefficient of variety and coefficient of skewness speaks to the non-dimensional proportion of symmetry. The estimations of the coefficient of variety and coefficient of skewness, determined from the day by day stream information at various measuring destinations, change from 1.94 - 2.35 and 2.48-4.94, separately. Since the estimations of the coefficient of skewness are high, the ordinary appropriation may not be utilized for, the typical dispersion may not be utilized for water accessibility investigation. The example measurements, for example, mean, standard deviation, coefficient of variety and coefficient of skewness are figured from the log changed estimations of the day by day stream (log base to e) for each checking site. The estimations of mean and standard deviation in log area change from 3.3–6.3 and 1.0–1.2, separately. The estimations of the coefficient of variety and coefficient of skewness in log area shift from 0.51–1.0 and 0.09–0.54, individually. The day by day stream information displayed high skewness. At the point when this arrangement was changed to a log changed arrangement the skewness was decreased and it is seen that the daily flow stream data are grouped around a straight line on log normal probability plots.

The goodness of fit test results shows that the annual data mostly follows the Log-Normal (2P) Distribution. The two parameters of the Log-Normal Distribution are Mean and Standard Deviation. The regional relations developed are used to estimate these parameters.

7.1 Results of Regionalisation Methods

The effect of urbanization on hydrological processes and water resources have been quantified and predicted by detailed characterization and physiographic characteristics of affected ungauged area Various relationships between coefficient of variation and physiographic characteristics like catchment area, altitude non-dimensional parameter (Ca/A^2) has been developed and which trend its follows has been shown in Figs. 7.1 to 7.5.

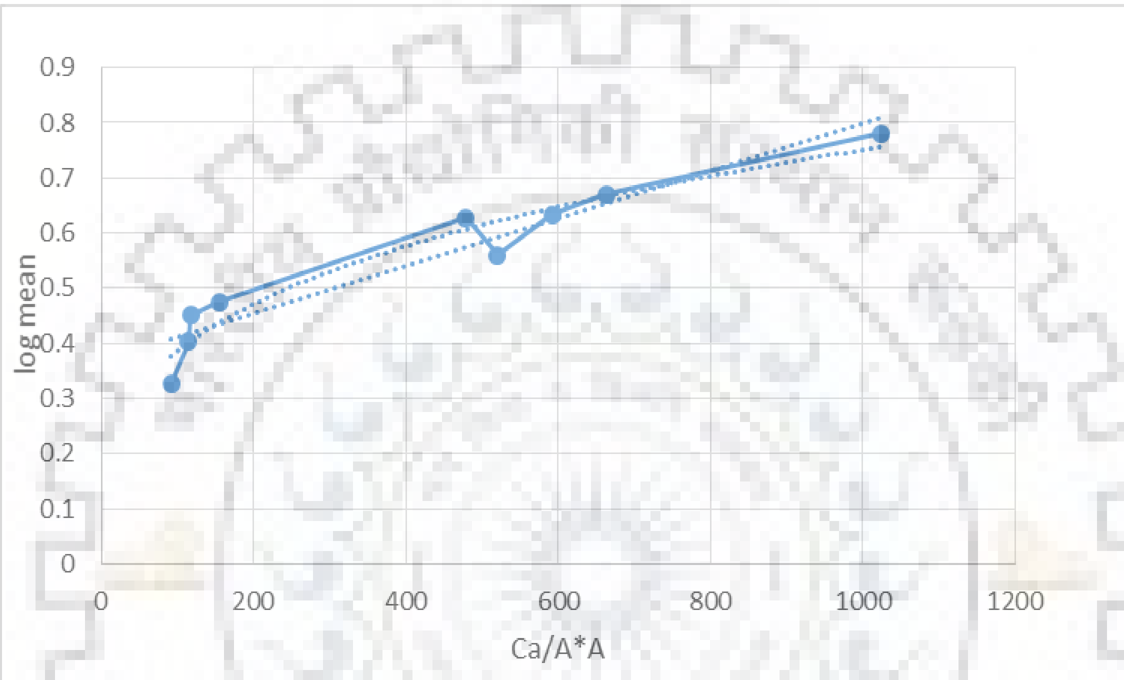


Figure 7.1 Mean vs (ca/A*A) graph in log scale

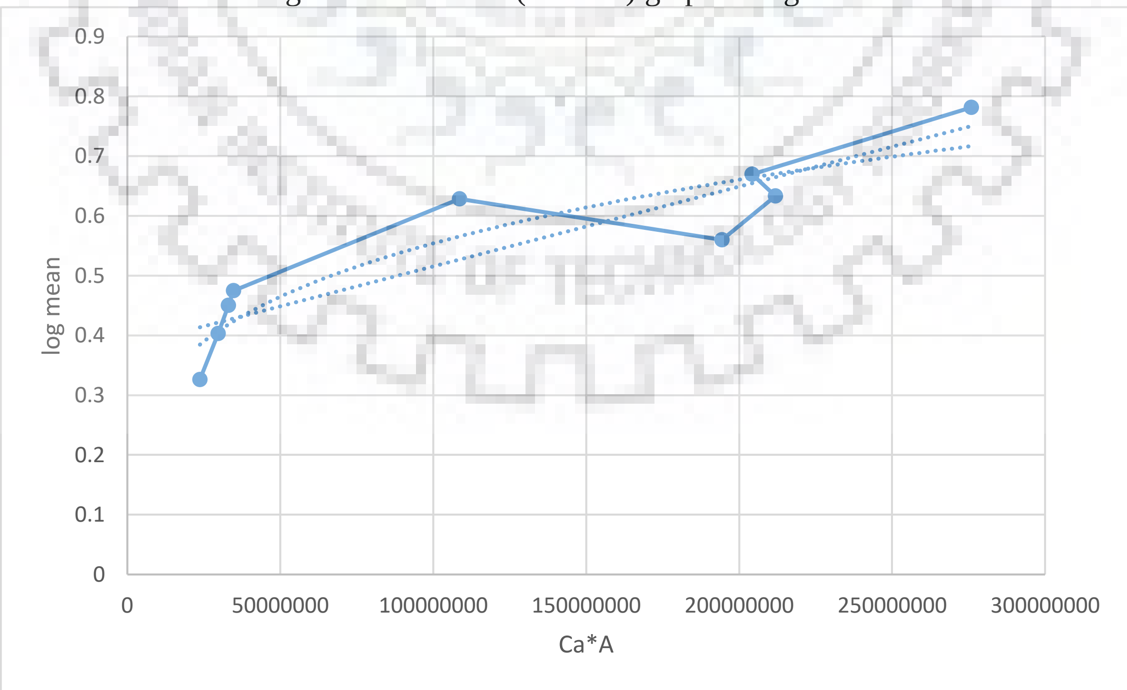


Figure 7.2 Mean vs (Ca/A) graph in log scale

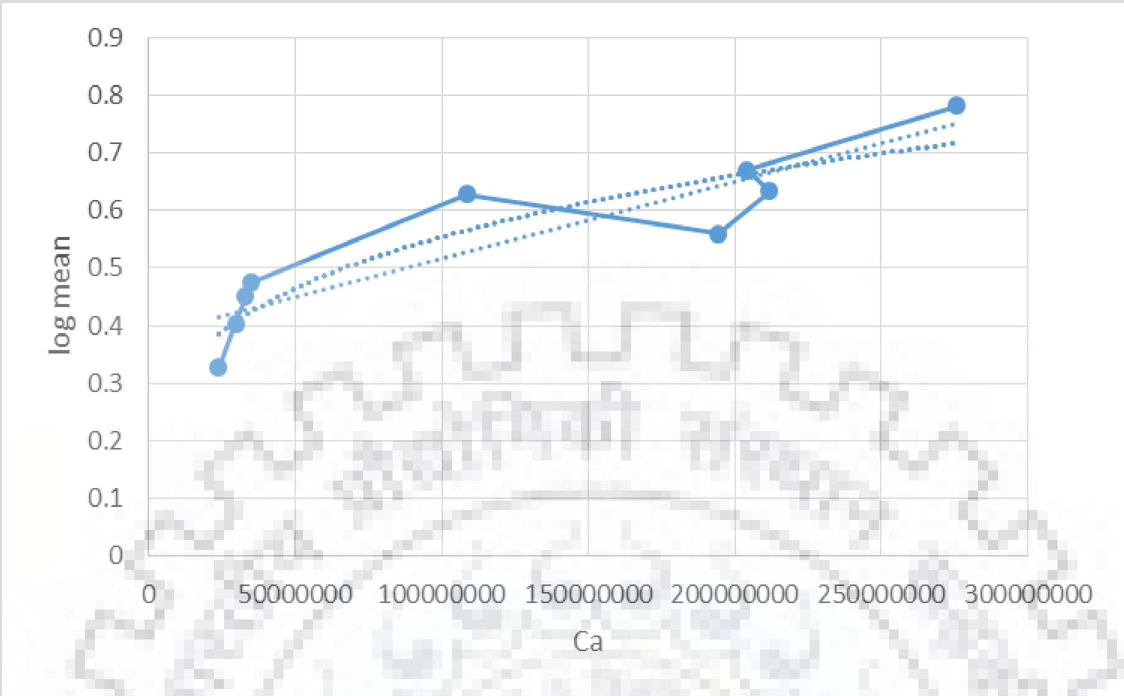


Figure 7.3 Mean vs catchment area graph in log scale

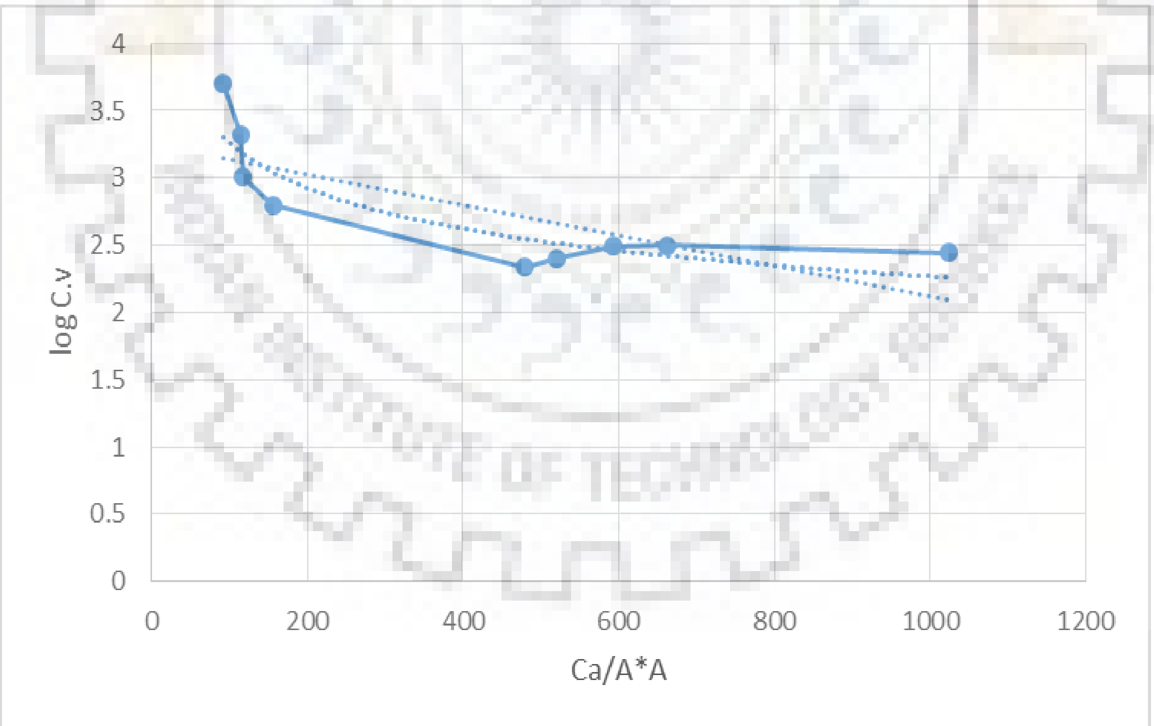


Figure 7.4 coefficient vs (Ca/A*A) graph in log scale

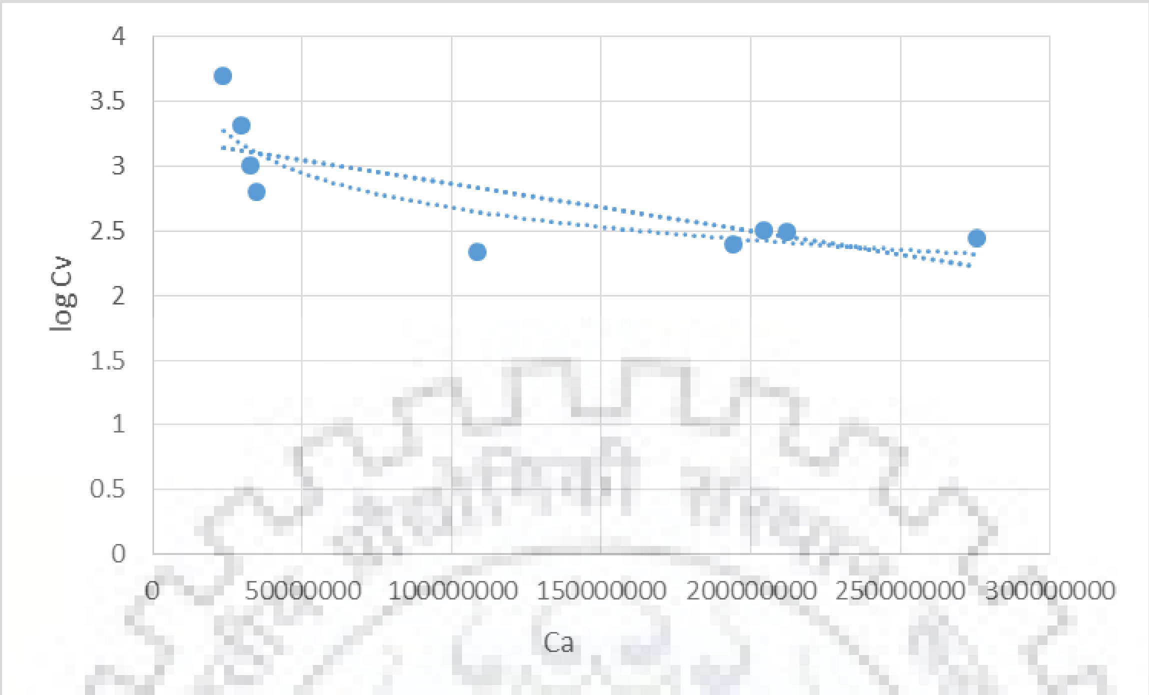


Figure 7.5 coefficient vs Catchment area graph in log scale

7.1.1 Analysis of regionalisation method -1

In this method, the following relationships have been generated between the mean of flow values in log space with the physiographic characteristics such as altitude (A), catchment area (Ca) and Ca/A^2 in logarithmic space of the ungauged catchment.

$$\bar{Y} = 0.1019(Ca/A^2)^{0.2891} \quad (R^2 = 0.9158) \quad (7.1)$$

$$\bar{Y} = 0.0204(Ca/A)^{0.2721} \quad (R^2 = 0.8871) \quad (7.2)$$

$$\bar{Y} = 16.406(Ca)^{-0.149} \quad (R^2 = 0.7877) \quad (7.3)$$

By using regionalization method 1 we calculated the various dependability flow for the gauged catchment for calibration and validated with ungauged catchment. This developed equation shows that best results can be get with (Ca/A^2) non dimensional characteristics. the results of the method 1 which we get are shown in Table 7.1

Table 7.1 dependable flow using Regionalisation method 1

Regionalisation method 1

Station	Q_{40}	Q_{50}	Q_{60}	Q_{70}	Q_{80}	Q_{90}
Kadvi	2.76	1.86	1.38	0.89	0.57	0.32
Ved Ganga	2.85	1.91	1.40	0.90	0.57	0.32
Kasari	2.93	1.95	1.43	0.91	0.57	0.32
Tulashi	2.00	1.46	1.15	0.81	0.57	0.36
Radhanagri	2.07	1.50	1.17	0.81	0.56	0.36
Dhom	2.71	1.83	1.36	0.88	0.57	0.33
Kanher	2.17	1.55	1.20	0.82	0.56	0.35
Dudhganga	2.07	1.49	1.17	0.81	0.56	0.36
Warna	3.28	2.13	1.53	0.95	0.58	0.32

7.1.2 Analysis of regionalisation method – 2

As same equations also generated by using regionalisation method 2 and the equations are given below.

$$\bar{Y} = 16.406(Ca)^{-0.149} \quad (R^2 = 0.7877) \qquad (7.4)$$

$$\bar{Y} = 0.0204(A)^{0.2721} \quad (R^2 = 0.8871) \qquad (7.5)$$

$$\bar{Y} = 0.1019(Ca/A^2)^{0.2891} \quad (R^2 = 0.9158) \qquad (7.6)$$

The form of the calculated relationships by the method-2 is tabulated below in Table 7.2. These results shows that the stream data or various dependability flow data generally varies with the non-dimensional physiographic characteristics (Ca/A^2) with linear co-relation value 0.70 to 0.85

Table 7.2 Relationship between physiographic characteristics and flow

Dependability	Relationship using Ca/A^2	R^2 value
40%	$\bar{Y} = 0.147(Ca/A^2)^{0.741}$	0.8771
50%	$\bar{Y} = 0.1076(Ca/A^2)^{0.6886}$	0.7164
60%	$\bar{Y} = 0.1019(Ca/A^2)^{0.7455}$	0.799
70%	$\bar{Y} = 0.0224(Ca/A^2)^{0.6874}$	0.5363
80%	$\bar{Y} = 0.009(Ca/A^2)^{0.71}$	0.8291
90%	$\bar{Y} = 0.0008(Ca/A^2)^{0.9504}$	0.8365

The results we get by using those equations for different dependability of various gauging stations are tabulated below in Table 7.3

Table 7.3 dependable flow using Regionalisation method 2

Regionalisation method 2						
Station	**Q40**	**Q50**	**Q60**	**Q70**	**Q80**	**Q90**
Kadvi	4.18	2.03	0.96	0.50	0.21	0.06
Ved Ganga	4.93	2.40	1.13	0.58	0.25	0.07
Kasari	5.02	2.44	1.15	0.59	0.25	0.07
Tulashi	6.18	3.01	1.42	0.72	0.31	0.10
Radhanagri	14.24	6.96	3.29	1.56	0.67	0.28
Dhom	15.14	7.40	3.50	1.65	0.71	0.31
Kanher	16.67	8.15	3.85	1.80	0.78	0.35
Dudhganga	18.12	8.86	4.19	1.95	0.84	0.38
Warna	25.00	12.25	5.79	2.63	1.14	0.58

7.1.3 Analysis of regionalisation method – 3

In regionalisation method 3, Ratios of the daily flow data to mean of the gauging site are computed to get non-dimensional data series at various gauging sites. Then all the gauging site’s data series pooled together to get a population of non- dimensional flow for the region and relationship is developed for the given region and results also given in Table 7.4

$$\log\frac{Q_D}{\bar{Q}} = \mu + \sigma \times Z_D \qquad (7.7)$$

$$\bar{Y} = 16.406(Ca)^{0.86} \qquad (R^2 = 0.9599) \qquad (7.8)$$

$$\bar{Y} = 0.0204(A)^{0.28} \qquad (R^2 = 0.8871) \qquad (7.9)$$

$$\bar{Y} = 0.1132(Ca/A^2)^{0.92} \qquad (R^2 = 0.9136) \qquad (7.\ 10)$$

Where μ = mean of the non-dimensional series in log space

σ = standard deviation of the non-dimensional series in log space

Q_D = various dependable flow

$\bar{Q}$ = mean flow

Table 7.4 computed dependable flow using Regionalisation method 3

Regionalisation method 3						
Station	Q_{40}	Q_{50}	Q_{60}	Q_{70}	Q_{80}	Q_{90}
Kadvi	2.43	1.37	0.89	0.47	0.25	0.11
Ved Ganga	2.98	1.69	1.09	0.58	0.31	0.14
Kasari	3.05	1.73	1.12	0.59	0.31	0.14
Tulashi	3.95	2.23	1.45	0.77	0.40	0.18
Radhanagri	11.13	6.30	4.08	2.16	1.14	0.51
Dhom	12.01	6.79	4.40	2.33	1.23	0.55
Kanher	13.53	7.65	4.96	2.63	1.39	0.62
Dudhganga	15.00	8.49	5.50	2.91	1.54	0.68
Warna	22.37	12.66	8.20	4.35	2.29	1.02

The results of the all the 3 methods or relationship between flow and physiographic characteristics in log scale shows that best physiographic characteristics suited with flow data is non- dimensional parameter (Ca/A^2) with co-relation R^2 value 0.80 to 0.95

7.2 Comparison of Different Regionalisation Methods

The dependable flow generated or calculated from analysing the available flow data at various basins, are compared with these regionalisation methods. The percentage absolute error in Dependable flows values obtain and graphs obtain for different dependability of flows in Figure 7.6 to Figure 7.10.

It is observed from the figures that the PAEDF values for regionalisation method 2 is lowest for most of the catchments that indicates that method 2 is best suitable for this catchment. However, this method is not beneficial having Ca/A^2 in the extrapolation range. So it is avoided to use this methods which catchment having Ca/A^2 value more than maximum value of gauged catchment.

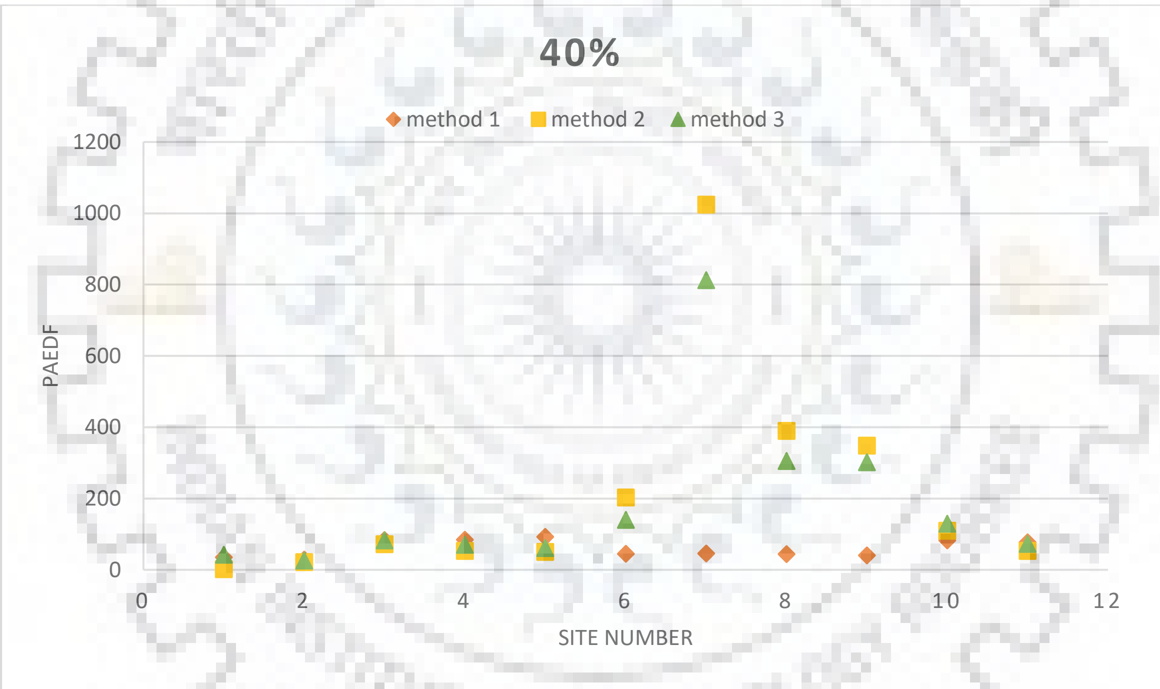


Figure 7.6 PAEDF for 40% dependability flow

Comparison of the all the results data shows that at site number 7 has largest PAEDF value. Which means at that station, regionalisation method is not giving better result and results were found to be inconsistent with respect to data of other sites. The reason of this error is having low altitude value of tulashi gauging site.

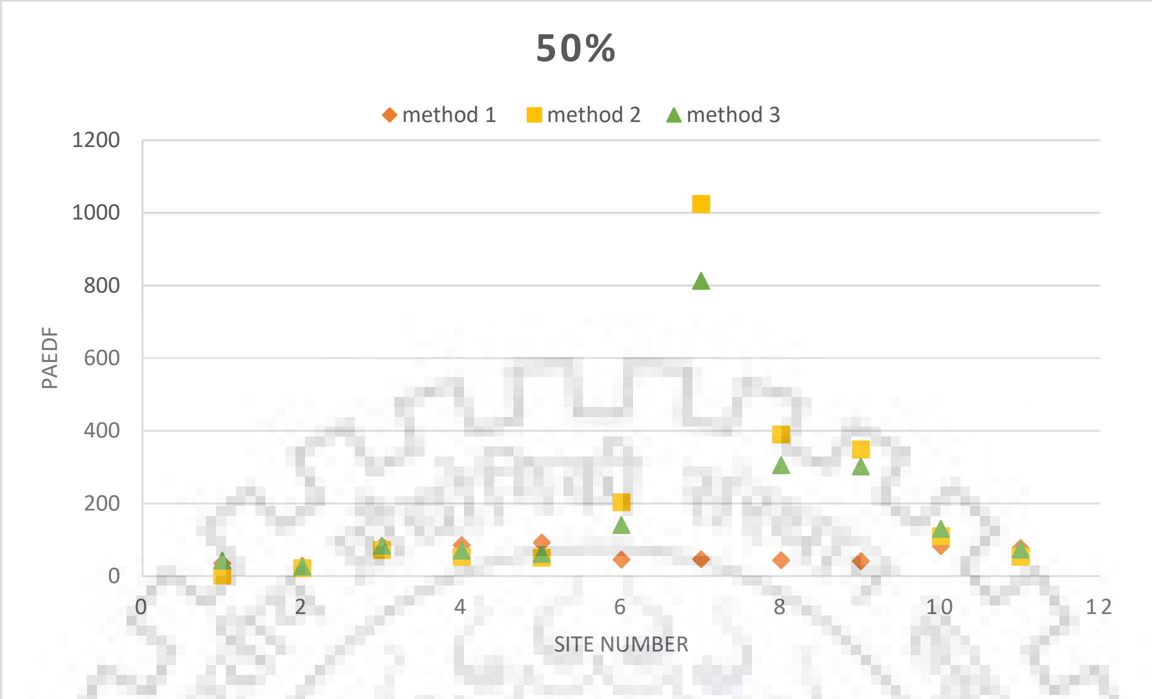


Figure 7.7 PAEDF for 50% dependability flow

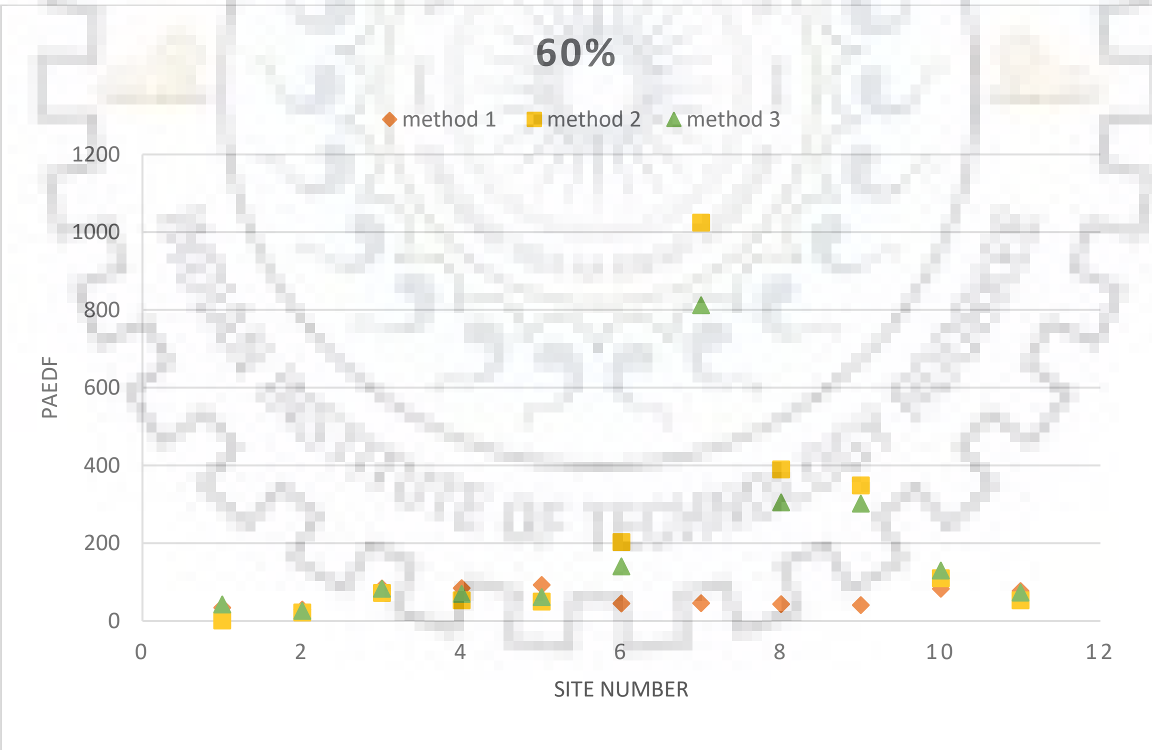


Figure 7.8 PAEDF for 60% dependability flow

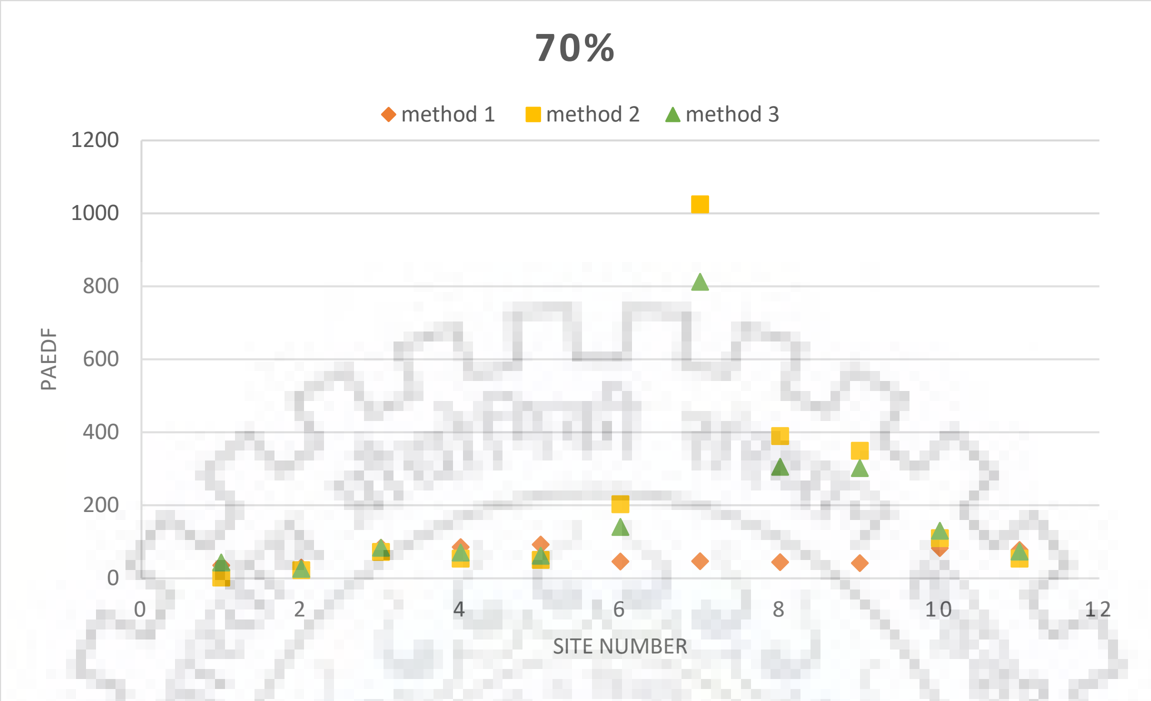


Figure 7.9 PAEDF for 70% dependability flow

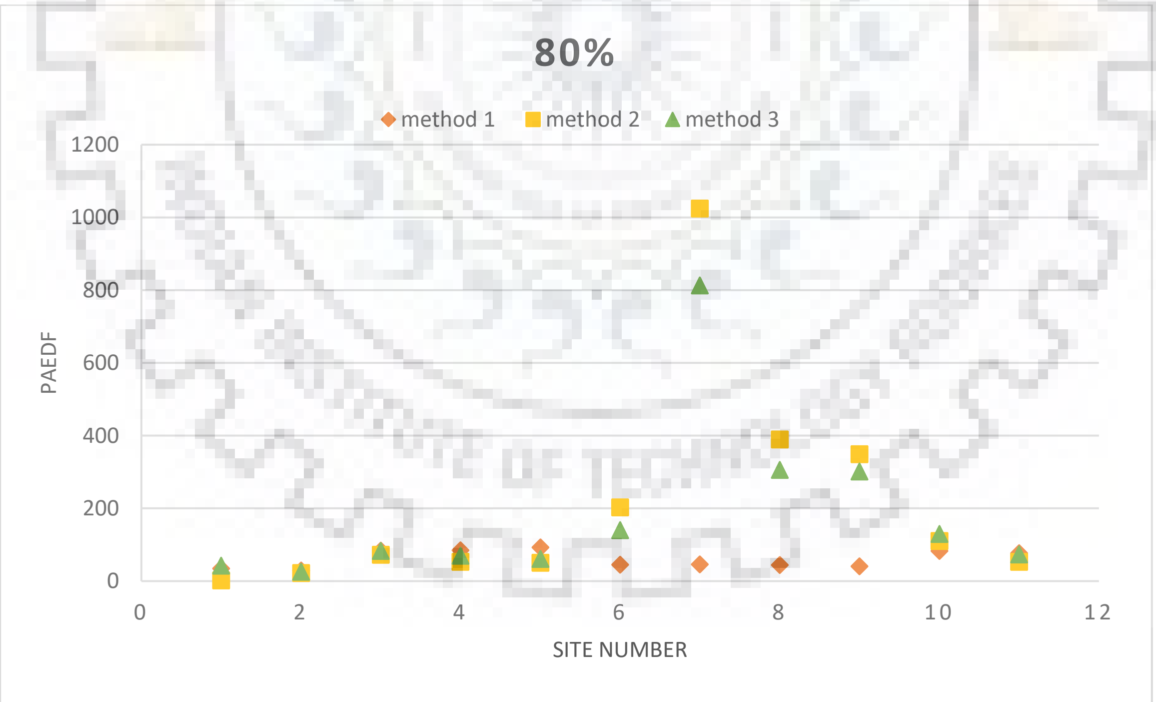


Figure 7.10 PAEDF for 80% dependability flow

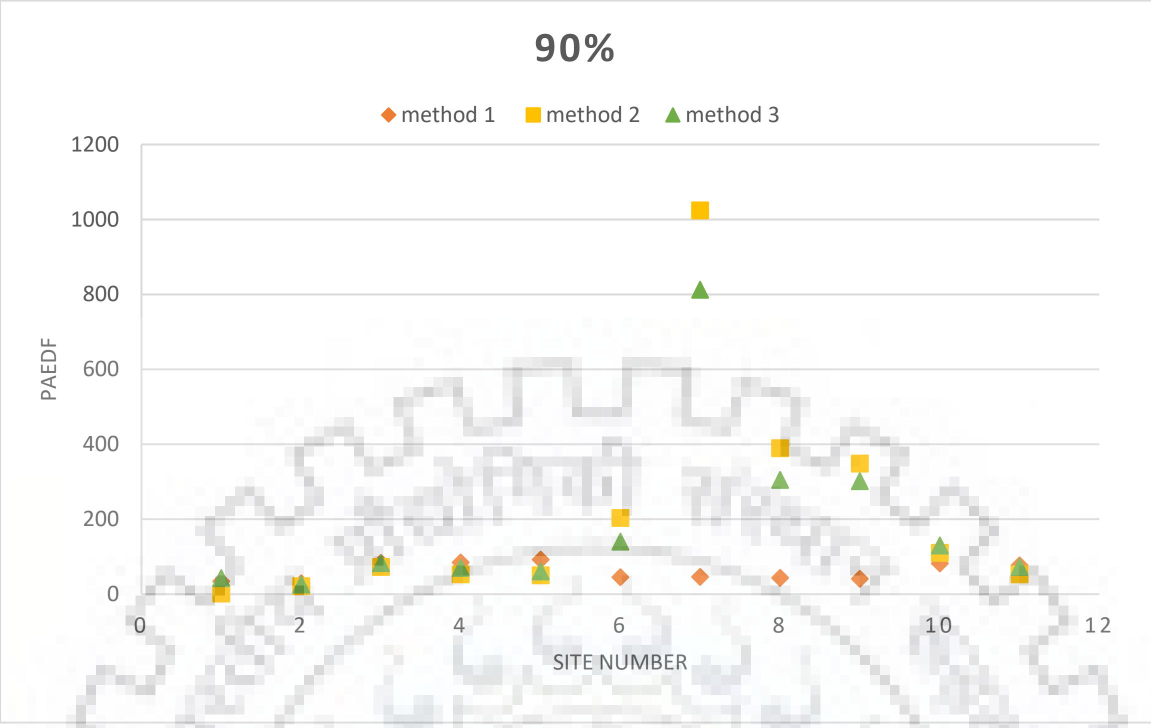


Figure 7.11 PAEDF for 90% dependability flow

7.3 Flow Duration Curve of Various Gauging Sites

For gauging sites with the available ten daily data, flow duration curve had been generated and shown in Figs. 7.12 to 7.21.

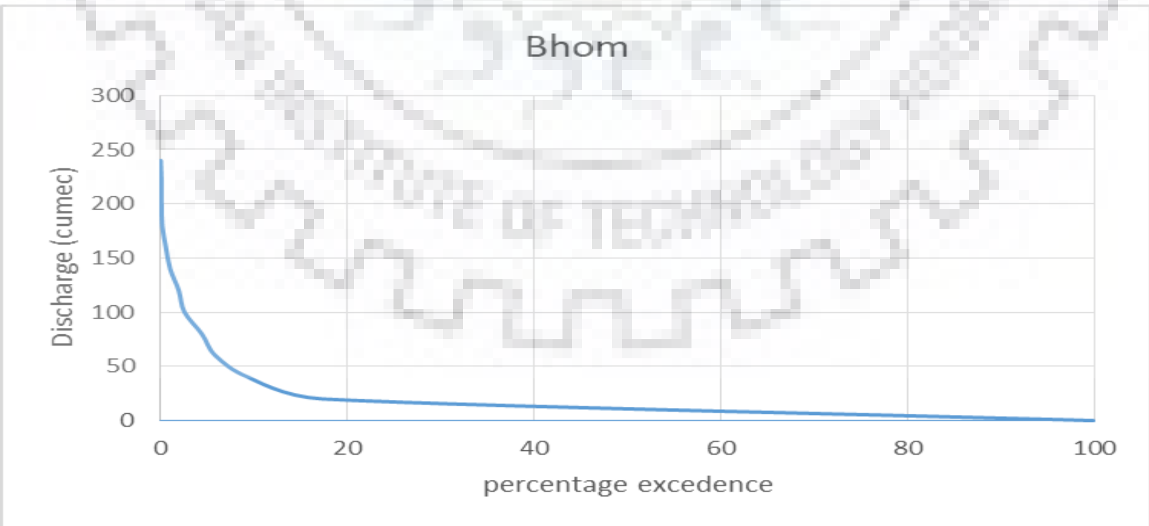


Figure 7.12 Flow Duration Curve at Dhom

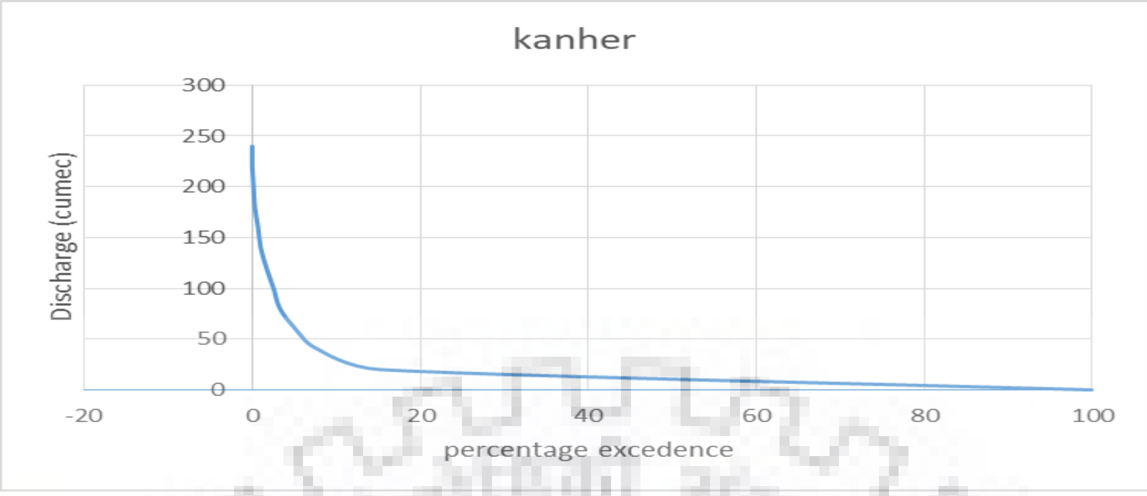


Figure 7.13 Flow Duration Curve at Kanher Catchment

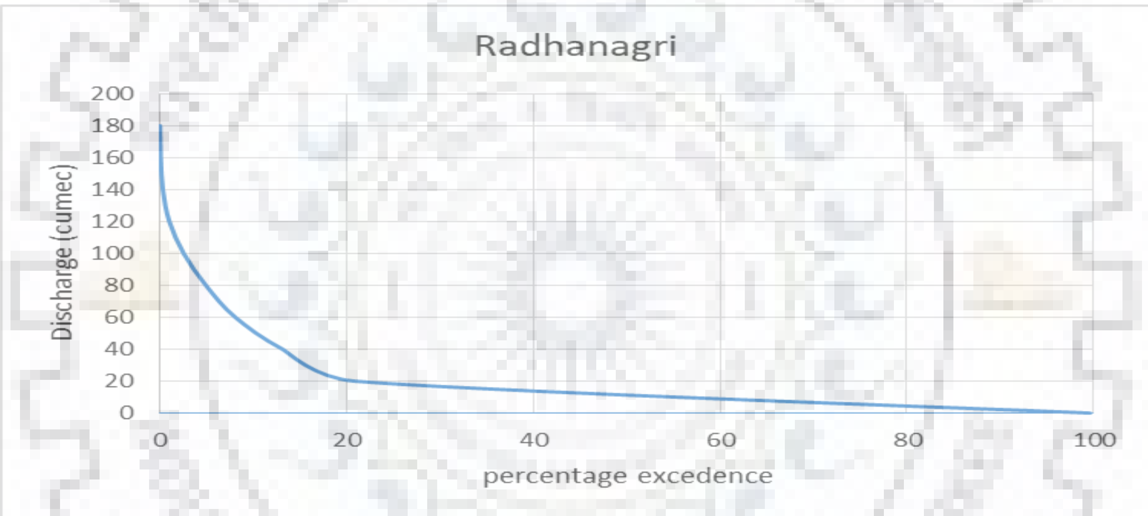


Figure 7.14 Flow Duration Curve at Radhanagri

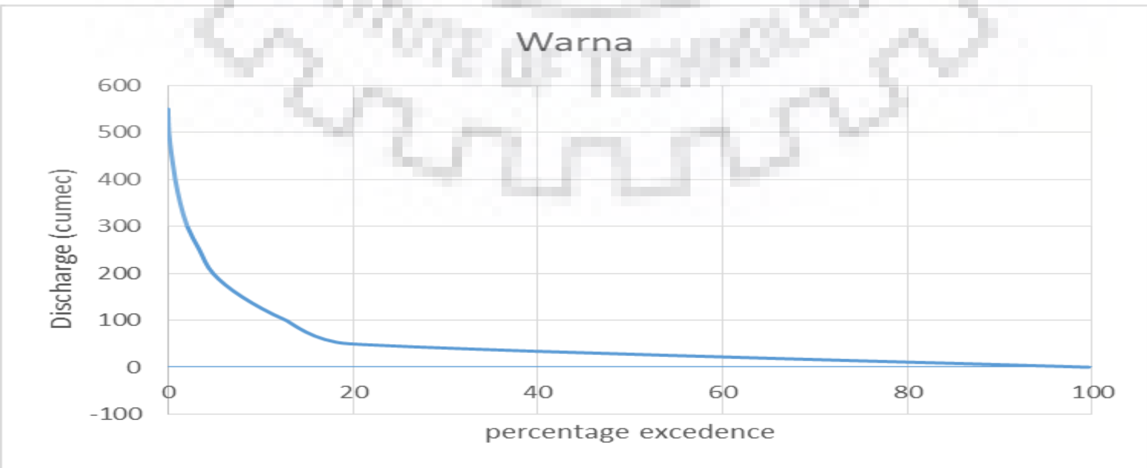


Figure 7.15 Flow Duration Curve at Warna

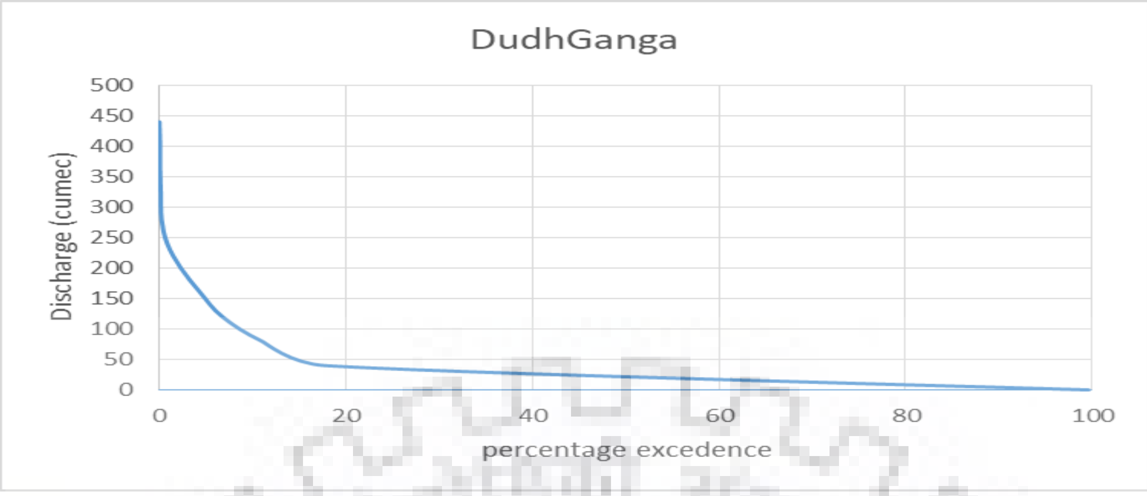


Figure 7.16 Flow Duration Curve at DudhGanga

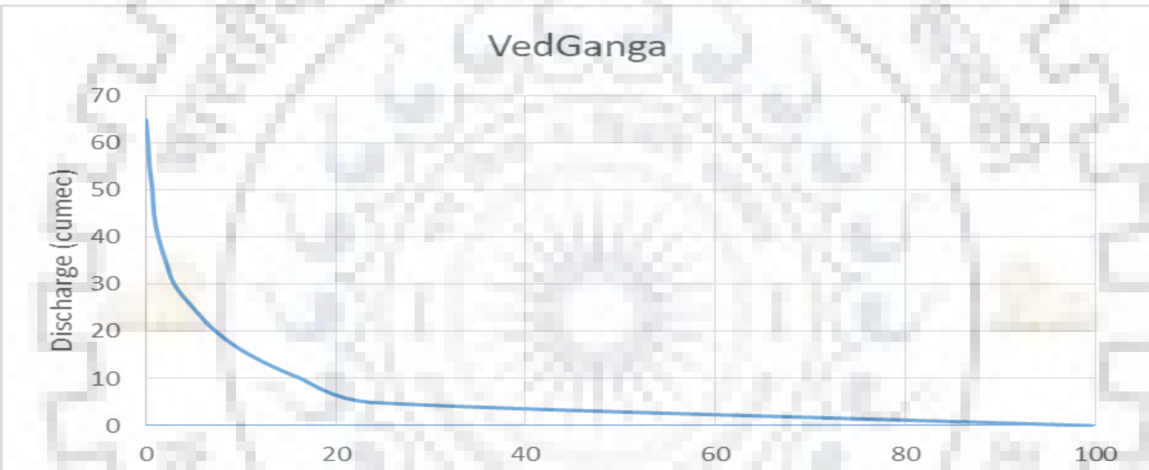


Figure 7.17 Flow Duration Curve at VedGanga

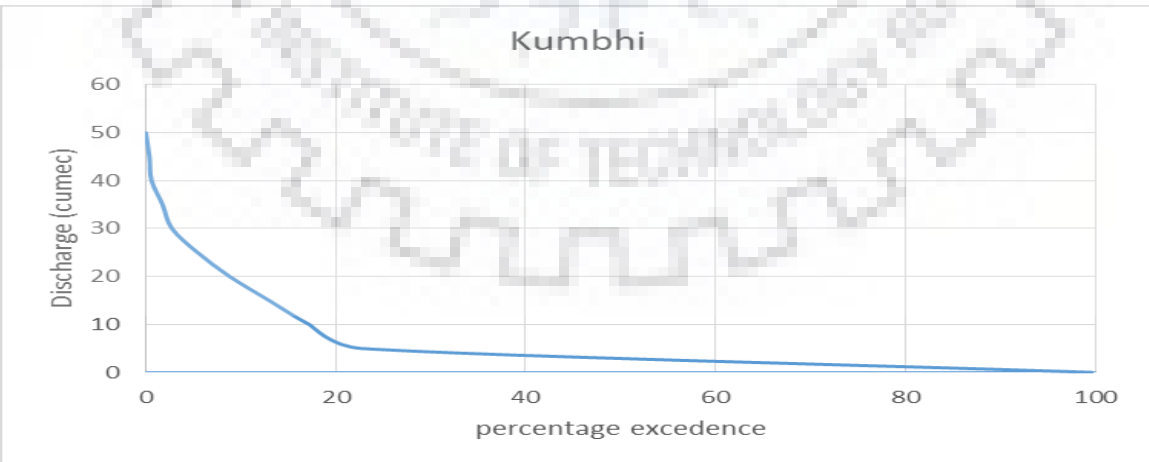


Figure 7.18 Flow Duration Curve at Kumbhi

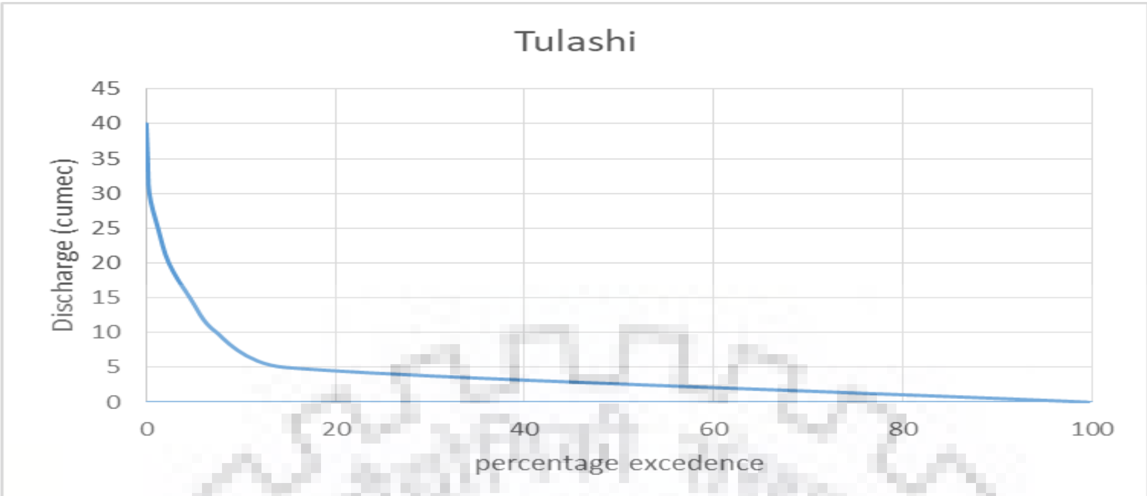


Figure 7.19 Flow Duration Curve at Tulashi

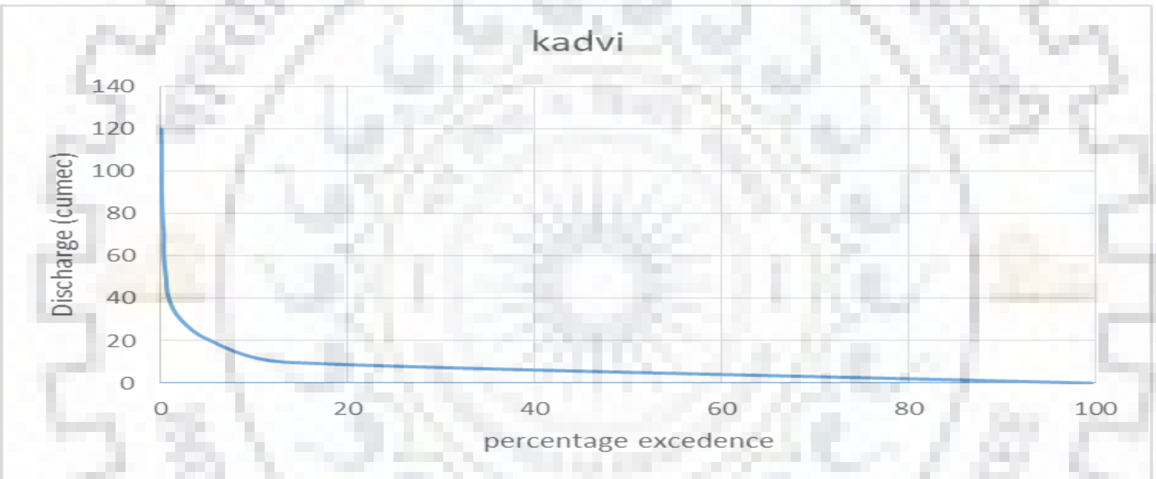


Figure 7.20 Flow Duration Curve at Kadvi

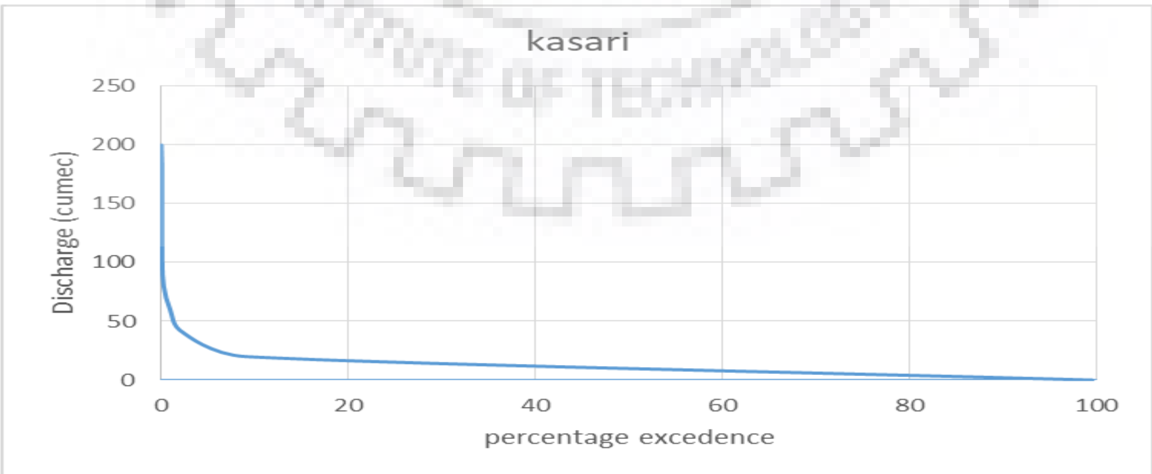


Figure 7.21Flow Duration Curve at kasari

CHAPTER 8

CONCLUSIONS

8.1 General

Discharge estimation at ungauged catchment is one of the important issues challenging water resource planners, and decision makers. Flow duration curve are commonly use to find discharge where flow records are not available like un-gauged sites in water resource engineering. Numerous methods of regional FDCs have been developed to provide synthetic FDCs at un-gauged sites. In the current study FDC was derived using the physiographic characteristics of the catchments. The method used was the regionalization of the parameters of chosen probability distribution for the gauged sites.

8.2 Conclusions

The following conclusions are drawn from the study:

1. Most appropriate common distribution fitting the annual data series at all the gauged sites through goodness of fit tests was found to be Log-Normal distribution. The parameters of Log-Normal distribution, Q (mean) and σ (standard deviation) in the developed equation were regionalized by using the physiographic characteristics of the catchments.

2. It was evaluated that the 10-daily flow series would also follow the same probability distribution as that of the followed by annual flow series.

3. The 10-daily discharge values at the un-gauged sites found using cumulative distribution function (CDF) were quite satisfactory between the limits Q95% and Q 10%. The extrapolation of generated flow duration curves (FDC) beyond Q95% and Q10% is not recommended since the CDF of Log-Normal distribution may produce unrealistic values.

4. The flow duration curves have been developed for k1 sub Basin of Krishna River. In regionalisation method-1 data was taken for each sites in log-normal distribution, in method-2 different dependable flows was taken in account and in method-3 region is taken as a whole of log- normal distribution was taken. In this study 11 gauging sites were taken in account 9 for calibration and 2 gauging sites for validation. The effect of the altitude (A), catchment area (Ca) and (Ca/A^2) were checked and found that the relationship of flow is best suited with (Ca/A^2) parameter.

5. The results of the all the methods were compared and conclude that, the results of the regionalisation method-2 was better than other regionalisation methods for whole catchment. So for the calculation of the dependable flows for the k1 basin, it is good to use regionalisation method-2.

REFERENCES

AniShabri, 2002. A comparison of plotting formulas for the Pearson type III distribution. JurnalTeknologi, 36(C) Jun. 2002: 61-74.

Arun Kumar, M. K. Singhal, 1999.Hydro power assessment for small ungauged catchments in Himalayan region using GIS techniques. Alternate Hydro Energy Centre, University of Roorkee, (Map India 1999)

Ashok K. Singh, Anita Singh, and Max Engelhard, 1997. The Lognormal Distribution in Environmental Applications. EPA/600/S-97/006, Dec. 1997.

Ashwani Kumar, Kapil Kumar, Naresh Kaushik, Satyawati Sharma, Saroj Mishra, 2010. Renewable energy in India: Current status and future potentials. Renewable and Sustainable Energy Reviews, 14(2010)2434-2442.

Attilio Castellarin, Giorgio Galeati, Luigia Brandimarte, Alberto Montanan; Armando Brath, 2004. Regional flow-duration curves: reliability for ungauged basins. Advances in Water Resources, 27 (2004), 953-965.

Attilio Castellarin, Giorgio Camorani, Armando Brath, 2007. Predicting annual and long-term flow-duration curves in ungauged basins. Advances in Water Resources, 30 (2007), 937-953.

By Richard M. Vogel, Member, ASCE, and Neil M. Fennessey, flow-duration curves. I: New interpretation and confidence intervals.

By Stacey A. Archfield, Richard M. Vogel, and Sara L. Brandt, 2007. Estimation of flow-duration curves at ungauged sites in southern New England. World Environmental and Water Resources Congress 2007.

B.C. Kusre, D.C. Baruah, P.K. Bordoloi, S.C. Patra, 2010. Assessment of hydropower potential using GIS and hydrological modelling technique in Kopili River basin in Assam (India). Applied Energy 87 (2010) 298-309.

Curves-Pao-Shan Yu, Tao-Chang Yang, and Yu-Chi Wang, 2002. Uncertainty Analysis of Regional Flow Duration. ASCE, 0733-9496 (2002) 128:6(424).

D.J. Booker, T.H. Snelder., 2012. Comparing methods for estimating flow duration curves at ungauged sites. Journal of Hydrology, (14 February 2012)

F. Viola, L. V. Noto, M. Cannarozzo, and G. La Loggia, 2010. Regional flow duration curves for ungauged sites in Sicily. Journal Hydrology and Earth System Sciences (HESS), Discuss. 7, 7059-7078, (2010) Gary W. Frey, Deborah M. Linke, 2002. Hydropower as a renewable and sustainable energy resource meeting global energy challenges in a reasonable way. Energy Policy 30 (2002), 1261-1265.

Geeta S. Joshi, Kapil Gupta, 2009.A simulation model for the operation of multipurpose multi reservoir system for River Narmada, India. Journal of Hydro-environment Research, 3 (2009), 96-108.

Goel N.K., 2011. Stochastic Hydrology. Lecture notes, Dept. of Hydrology, IIT Roorkee.

Hafzullah AKSOY, 2000. Use of Gamma Distribution in Hydrological Analysis. Turk J Engin Environ Sci, 24 (2000), 419- 428.

Humberto Silva-Hidalgo, Ignacio R. Martin-Dominguez, Maria Teresa Alarcon-Herrera and Alfredo Granados-Olivas. Natural flows determination in gauged hydrological basins. Part I: Alternative method for irrigation return flows estimation.

loannis A. Niadas, 2005. Regional flow duration curve estimation in small ungauged catchments using instantaneous flow measurements and a censored data approach. Journal of Hydrology, 314 (2005), 48-66.

Ming Li., Quanxi Shao, Lu Zhang, Francis H.S. Chiew, 2012. A new regionalization approach and its application to predict flow duration curve in ungauged basins, Journal of Hydrology, 389 (23 May 2010), 137-145.

Richard M. Vogel and Neil M. Fennessey, 1995. Flow duration curves 11: A review of applications in water resources planning. American water resource engineering vol. 3, 1995, No. 6.

R.A.J. Cole, H. T. Johnston & D. J. Robinson, 2003. The use of flow duration curves as a data quality tool. Hydrological Sciences—Journal-des Sciences Hydrologiques, (December 2003, 48(6)).

R. Halliday& Associates Saskatoon, .SK, 2010. Determination of natural flow for apportionment of the red river. 2010, 05 30.

Srivastava D. K., 2006. Flow duration curve and water availability. NTPC sponsored training course on `Hydrological Analysis and Planning for Hydro Power Project', (June 5-10, 2006).

Website: Engineering Statistics Handbook.

httR://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm

APPENDIX-1

OCTAVE PROGRAMMING

```
% READING AND WRITING DATA
a = load('Book1.txt');
[p,q] = size(a);
c = 1; d = 36;
for k = 1:11
for i = 1:36
b(i,k) = 0;
for j = c:d
b(i,k) = b(i,k) + a(j,k);
end
c = c + 36;
d = d + 36;
end
c = 1; d = 36;
end
b  = b*0.864;
year(1,1) = 1973;
for i = 2:36
year(i,1) = year(i-1,1) + 1;
end
% GOODNESS OF FIT TEST FOR NORMAL DISTRIBUTION
for j = 1:11
crest(1,j) = 1;
trough(1,j) = -1;
for i = 2:35
if b(i-1,j) > b(i,j) && b(i+1,j) > b(i,j)
crest(1,j) = crest(1,j) + 1;
end
if b(i-1,j) < b(i,j) && b(i+1,j) < b(i,j)
trough(1,j) = trough(1,j) + 1;
```

```
end
end
end

%Turning Point Test
ntp = trough + crest;
N = 36;
mean = 2*(N - 2)/3;
variance = (16*N - 29)/90;
variate = abs((ntp - mean)/sqrt(variance));

% Kendall's Rank Test
krt = zeros(36,11);
for j = 1:11
for i = 1:36
for k = i:36
if b(i,j) < b(k,j)
krt(i,j) = krt(i,j) + 1;
end
end
end
end
sum_krt = zeros(1,11);
for j = 1:11
for i = 1:36
sum_krt(1,j) = krt(i,j) + sum_krt(1,j);
end
end

for j = 1:11
tau(1,j) = 4*sum_krt(1,j)/(N*(N-1)) - 1;
end
```

```
var_tau = 2*(2*N + 5)/(9*N*(N-1));
for j = 1:11
variate_krt(1,j) = abs(tau(1,j)/sqrt(var_tau));
end
average = zeros(1,11);
for j = 1:11
for i = 1:36
average(1,j) = average(1,j) + b(i,j);
end
end
average = average/36;
G = zeros(1,11);
for j = 1:11
for i = 1:36
G(1,j) = (b(i,j) - average(1,j))^2 + G(1,j);
end
end
std_dev = sqrt(G/(N-1));

%Skewness
skewness = zeros(1,11);
for j = 1:11
for i = 1:36
skewness(1,j) = skewness(1,j) + ((b(i,j) - average(1,j))/std_dev(1,j))^3;
end
end
skewness = skewness*N/((N-1)*(N-2));

B = sort(b, 'descend');
kt1 = [0.0, -0.476, 0.087, 0.672, 1.5]';
kt2 = [-0.476, 0.087, 0.672, 1.5, 0.0]';
for j = 1:11
for i = 1:5
```

```
xt1(i,j) = average(1,j) + kt1(i,1)*std_dev(1,j);
xt2(i,j) = average(1,j) + kt2(i,1)*std_dev(1,j);
end
end
oj = zeros(5,11);
xt1(1,:)=0;
%xt2(5,:)=0;
for k = 1:5
for j = 1:11
for i = 1:36
if b(i,j) >= xt1(k,j) && b(i,j) < xt2(k,j)
oj(k,j) = oj(k,j) + 1;
elseif b(i,j) >= xt1(k,j) && k == 5
oj(k,j) = oj(k,j) + 1;
end
end
end
end
ej = N/5;
chi_square = ((oj - ej).^2)/ej;
chi_sum = zeros(1,11);
for j = 1:11
for i = 1:5
chi_sum(1,j) = chi_sum(1,j) + chi_square(i,j);
end
end

%Gumbel distribution for Godness of fit test
alpha = sqrt(6*(std_dev.^2)/pi^2);
U = average - 0.5772*alpha;
P1 = [0, 0.2, 0.4, 0.6, 0.8];
P2 = [0.2,0.4,0.6,0.8,1];
Gumbel_kt1 = (-log(-log(1-P1)))';
```

```
Gumbel_kt2 = (-log(-log(1-P2)))';

for j = 1:11
for i = 1:5
Gumbel_xt1(i,j) = U(1,j) - alpha(1,j)*Gumbel_kt1(i,1);
Gumbel_xt2(i,j) = U(1,j) - alpha(1,j)*Gumbel_kt2(i,1);
end
end
Gumbel_xt1(1,:) = 0;
Gumbel_oj = zeros(5,11);
Gumbel_xt1(1,:)=0;
%Gumbel_xt2(5,:)=0;
for k = 1:5
for j = 1:11
for i = 1:36
if b(i,j) >= Gumbel_xt1(k,j) && b(i,j) < Gumbel_xt2(k,j)
Gumbel_oj(k,j) = Gumbel_oj(k,j) + 1;
elseif b(i,j) >= Gumbel_xt1(k,j) && k == 5
Gumbel_oj(k,j) = Gumbel_oj(k,j) + 1;
end
end
end
end

Gumbel_chi_square = ((Gumbel_oj - ej).^2)/ej;
Gumbel_chi_sum = zeros(1,11);
for j = 1:11
for i = 1:5
Gumbel_chi_sum(1,j) = Gumbel_chi_sum(1,j) + Gumbel_chi_square(i,j);
end
end
for j = 1:11
for i=1:36
```

```
feyi(i,j) = i/37;
FD(i,j)=exp(-exp(-(B(i,j)-average(1,j))/std_dev(1,j)));
end
end
feyi = sort(feyi,'descend');

D = feyi - FD;
highest_D = max(D);
```

APPENDIX-2

FORTRAN PROGRAMMING

```
THIS PROGRAME IS FOR COMPUTING THE AVAILABILITY OF SURFACE
C  WATER FOR GIVEN DEPENDABILTIES FITTING LOG NORMAL
DISTRIBUTION
C   USING LEAST SQUARE METHOD
C   PROBABILITIES ARE ASSIGNED USING BLOM'S PLOTTING POSITION
C   FORMULA
C   X = VECTOR CONTAINING FLOW VALUES
C   F = VECTOR CONTAINING EXCEDENCE PROBABILITY
    CHARACTER*30 FYLE,FYLEN
    DIMENSION X(11200),ca(12),h(12),y(12),z(10)
   1  ,caa(12),hh(12),xnd(83400),qm(60),XX(31),MON(12)
    data MON/31,28,31,30,31,30,31,31,30,31,30,31/
    WRITE(*,*)'SUPPLY INPUT FILE NAME?'
    READ(*,234)FYLE
    WRITE(*,*)'SUPPLY OUTPUT FILE NAME?'
    READ(*,234)FYLEN
234         FORMAT(A)
    OPEN(UNIT=1,FILE=FYLE,STATUS='OLD')
    OPEN(UNIT=2,FILE=FYLEN,STATUS='unknown')
    open(unit=3,file='ungauge.dat',status='old')
    read(3,*)nug,m
    read(3,*)(y(i),i=1,m)
    READ(1,*)ns
    kk=0
    do jj=1,ns
            write(2,1245)jj
1245        format(10x,'gauging site no:-',i5)
            READ(1,*)nsn
C           WRITE(*,*)NSN
```

```
         READ(1,*)npP,ca(jj),rr,h(jj)
C        WRITE(*,*)NPP,CA(JJ),RR,H(JJ)
C        PAUSE
         NP=0
         DO JK=1,40
         DO KJ=1,12
         READ (1,111,END=999,ERR=999)icode,iy,im,icn,(XX(i),i=1,8)
C        WRITE(*,111)ICODE,IY,IM,ICN,(XX(I),I=1,8)
111          Format(i5,i6,2i2,8f7.1)
         READ (1,111)ICODE,IY,IM,ICN,(XX(I),I=9,16)
C        WRITE(*,111)ICODE,IY,IM,ICN,(XX(I),I=9,16)
         READ (1,111)ICODE,IY,IM,ICN,(XX(I),I=17,24)
C        WRITE(*,111)ICODE,IY,IM,ICN,(XX(I),I=17,24)
         IF(MOD(IY,4).EQ.0)THEN
         MON(2)=29
         ELSE
         MON(2)=28
         END IF
         READ (1,111)ICODE,IY,IM,ICN,(XX(I),I=25,MON(IM))
C        WRITE(*,111)ICODE,IY,IM,ICN,(XX(I),I=25,MON(IM))
         DO KK1=1,MON(IM)
         NP=NP+1
C        WRITE(*,*)NP
         X(NP)=XX(KK1)
         END DO
         IF(NP.EQ.NPP)GO TO 999
         END DO
         END DO
999          CONTINUE
         WRITE(*,*)NP,NPP
   call  stat(x,np,QMm,SDQq,SKEW,cv)
   qm(jj)=qmm
   write(2,580)
```

```
580         format(10x,'mean and standard deviatin in real space')
      write(2,581)qmm,sdqq
581         format(4x,'mean=',f10.2/4x,'st. dev.=',f10.2)
            do i=1,np
            kk=kk+1
            xnd(kk)=x(i)/qmm
            end do
      ca(jj)=alog(ca(jj))
      end do
      do ii=1,nug
      read(3,*)caa(ii),hh(ii)
      END DO
20    CONTINUE
      write(2,246)
246         format(4x,'method no:-3')
      AN=kk
      SUM1=0.0
      SUM2=0.0
      DO II=1,kk
      SUM1=SUM1+xnd(II)
      SUM2=SUM2+xnd(II)**2
      END DO
      XXM=SUM1/AN
      VARQ=(1./(AN-1.0))*(SUM2-AN*(XXM**2))
      SDQ=SQRT(VARQ)
      write(2,680)
680         format(10x,'mean and stand. dev. of non-dimen. in real space')
      write(2,681)XXM,SDQ
681         format(4x,'mean=',f10.2/4x,'st. dev.=',f10.2)
      call ln2(xnd,kk,qm2,sdq2)
      write(2,586)
586         format(10x,'relationship between flows and catchment area in the
     1 form log(qm) =a+b*log(ca)')
```

```
    do i=1,ns
    qm(i)=alog(qm(i))
    end do
      call REG(qm,ca,Ns,A,B)
    do jj=1,nug
    ym=a+b*alog(caa(jj))
    ym=exp(ym)
            DO I=1,M
            YY=Y(I)/100.0
            YY=1.-YY
            CALL NDTRI(YY,YX,D,IE)
            Z(I)=qm2+sdq2*YX
            Z(I)=EXP(Z(I))*ym
            end do
    write(2,585)jj
    WRITE(2,245)
    WRITE(2,145)(Y(I),Z(I),I=1,M)
    end do
585         format(4x,'ungauged catchment no.:',i4)
    WRITE(2,145)(Y(I),Z(I),I=1,M)
145         FORMAT(4X,F10.3,F14.3)
245         FORMAT(2X,//'% DEPENDABILITY',2X,'CORRESPONDING FLOWS'/)
    stop
    end
    subroutine ln2(x,n,QM,SDQ)
    dimension x(83400),t(83400)
    WRITE(2,579)
579         FORMAT(20X,'WATER AVAILABILITY STUDY -LOG NORMAL
DISTRIBUTION')
    WRITE(2,580)
580         FORMAT(20X,48('*')//)
    CALL RANK (X,N)
    AN=N
```

```
      DO 10 I=1,N
      AI=I
      F=(AI-0.375)/(AN+0.25)
      P=F
      CALL NDTRI(P,XY,D,IE)
      F=F*100.0
      T(I)=XY
10    CONTINUE
      DO 237 I=1,N
237         X(I)=ALOG(X(I))
      write(*,*)'we are in ln2'
      call  stat(x,N,QM,SDQ,SKEW,cv)
      WRITE(2,578)QM,SDQ,CV,SKEW
578         FORMAT(10X,'MEAN =',9X, F10.2/10X,'STAND. DEV.=',3X,F10.2/10X,
     1  'COEFF. OF VAR.=',F10.2/10X,'SKEWNESS=',6X,F10.2//)
C     WRITE(2,400)
C     WRITE(2,300)(X(I),F(I),T(I),I=1,N)
        CALL REG(X,T,N,C,B)
c     DO 20 I=1,M
c     YY=Y(I)/100.0
c     YY=1.-YY
c     CALL NDTRI(YY,YX,D,IE)
c     Z(I)=C+B*YX
c     Z(I)=EXP(Z(I))
c20CONTINUE
c     WRITE(2,200)
c     WRITE(2,100)(Y(I),Z(I),I=1,M)
c100         FORMAT(4X,F10.3,F14.3)
c200         FORMAT(2X,//'% DEPENDABILITY',2X,'CORRESPONDING FLOWS'/)
C300         FORMAT(4X,F10.3,7X,F10.3,20X,F10.3)
C400         FORMAT(5X,'LOG OF FLOWS',3X,'NON EXCEEDENCE PROBABILITY
(%)',3X
C      1,' NORMAL REDUCED VARIATE')
```

```
      return
      END
C
      **********************************************************************
**
      subroutine rank(y,n)
c     This subroutine is used for arranging the flood values in
c     ascending order
$large :y
      DIMENSION Y(83400)
      N1=N-1
      DO 3 I=1,N1
      K=N-I
      DO 3 J=1,K
      IF(Y(J)-Y(J+1))3,3,2
2       SAVE=Y(J)
      Y(J)=Y(J+1)
      Y(J+1)=SAVE
3       CONTINUE
      RETURN
      END
C     *****************************************************
        SUBROUTINE NDTRI(P,X,D,IE)
        IE=0
        X=.99999E+37
        D=X
        IF(P)1,4,2
1        IE=-1
        GO TO 12
2        IF(P-1.0)7,5,1
4        X=-0.999999E+37
5        D=0.0
        GO TO 12
```

```
7       D=P
        IF(D-0.5)9,9,8
8       D=1.0-D
9       T2=ALOG(1.0/(D*D))
        T=SQRT(T2)
        X=T-(2.515517+0.802853*T+0.010328*T2)/(1.0+1.432788*T+
     1  0.189269*T2+0.001308*T*T2)
        IF(P-0.5)10,10,11
10      X=-X
11      D=0.3989423*EXP(-X*X/2.0)
12      RETURN
        END
C     *************************************************************
       SUBROUTINE REG(Y,X,N,A,B)
        DIMENSION X(83400),Y(83400),SE(2),T(2),REGG(2)
$large :x,y
       AN=N
       SUM1=0.0
       SUM2=0.0
       DO 100 I=1,N
       SUM1=SUM1+X(I)
100    SUM2=SUM2+Y(I)
       XM=SUM1/AN
       YM=SUM2/AN
       SUMXY=0.0
       SUMX2=0.0
       DO 101 I=1,N
       SUMXY=SUMXY+(X(I)-XM)*(Y(I)-YM)
101    SUMX2=SUMX2+(X(I)-XM)*(X(I)-XM)
       B=SUMXY/SUMX2
       A=YM-B*XM
       WRITE(2,1)
1       FORMAT(4X,65('*'))
```

```
      WRITE(2,2)
2     FORMAT(10X,'LEAST SQUARE FITTING OF A STRAIGHT LINE')
      WRITE(2,3)
3     FORMAT(15X,'EQUATION OF THE FITTED LINE')
      IF(B.GT.0)WRITE(2,4)A,B
      IF(B.LT.0)WRITE(2,478)A,B
4     FORMAT(5X,'LN(X)=',F10.4,4X,'+',F10.4,'*Y')
478   FORMAT(5X,'LN(X)=',F10.4,F10.4,'*Y')
      SUMS=0.0
      DO 102 I=1,N
      YC=A+B*X(I)
102   SUMS=SUMS+(YC-Y(I))*(YC-Y(I))
      SUMS=SUMS/(AN-2.0)
      S=SQRT(SUMS)
      SUMY=0.0
      DO104 I=1,N
104   SUMY=SUMY+(Y(I)-YM)*(Y(I)-YM)
      SUMY=SUMY/(AN-1.0)
      R=1.0-(SUMS/SUMY)
      R=SQRT(R)
      WRITE(2,25)R
25    FORMAT(4X,'COEFFICIENT OF CORRELATION=',F10.3)
      SEB=S/(SQRT(SUMX2))
      SEA=S*SQRT(1.0/AN+(XM*XM)/SUMX2)
      TB=B/SEB
      TA=A/SEA
      REGG(1)=A
      REGG(2)=B
      SE(1)=SEA
      SE(2)=SEB
      T(1)=TA
      T(2)=TB
      WRITE(2,6)
```

```
6     FORMAT(4X,'REGR. COEFF.',10X,'STAND. ERR.',10X,'T-VALUES')
      WRITE(2,7)(REGG(I),SE(I),T(I),I=1,2)
7     FORMAT(1X,F11.3,10X,F11.3,10X,F11.3)
      WRITE(2,8) S
8     FORMAT(4X,'STANDARD ERROR OF REGRESSION EQUATION=',F11.2)
      WRITE(2,1)
      RETURN
      END
C     *********************************************************
      SUBROUTINE stat(Q,N,QM,SDQ,SKEW,cv)
c     This subroutine is used for computing the basic sample
c     statistics
      DIMENSION Q(83400)
      AN=N
      SUM1=0.0
      SUM2=0.0
      SUM3=0.0
      DO 20 I=1,N
      SUM1=SUM1+Q(I)
20    SUM2=SUM2+Q(I)**2
      QM=SUM1/AN
      VARQ=(1./(AN-1.0))*(SUM2-AN*(QM**2))
      SDQ=SQRT(VARQ)
      DO 30 I=1,N
30    SUM3=SUM3+(Q(I)-QM)**3
      SK=((AN)/((AN-1.0)*(AN-2.0)))*SUM3
      SKEW=SK/(SDQ**3)
      cv=sdq/qm
      RETURN
      END
```