

## **PROCESSING OF STREAM FLOW DATA AND ESTABLISHMENT OF RATING CURVE**

**A.K. LOHANI  
SCIENTIST F**

### **GENERAL**

The rapid increase in the population has resulted in the developments in various sectors. It leads to the increase in water demands for different sectors such as agriculture, energy, industry, domestic etc. For effective and efficient water resource assessment and management and, for proper river basin planning and development of flood forecasting system, reliable, accurate, processed and easily approachable data base containing all the data of hydrologic variables recorded, is very much needed for proper planning and management of water resources. Moreover several types of statistical information is required for the analysis of hydrological systems i.e. flood frequency estimation, design flood estimation, reservoir capacity estimation and its operation, low flow analysis and for various other purposes. In our country various organizations are involved in collection of hydrologic and other related data. Data collected on hydrologic process are generally raw data and can not be used directly in most hydrologic analysis work. Therefore, processing of such data is essential to make them usable for various studies.

Stream flow records are primarily continuous records of flow passing through a particular section on the stream. The measurement made at the gauging site may subject to various random, systemic and spurious errors. Therefore data processing is required to transform the raw data into their most usable forms through a variety of quality checks at appropriate stages to ensure data quality and reliability. Processing of stream flow data is not a single step process. Several processing steps are required to produce reliable stream flow data. Automated data processing using high-speed computers has immense potential for handling large volume of hydrologic data in a quick and economic way. Most of the data processing activities are accomplished with the help of computers using hydrological data processing software.

There is a shared need for methods of statistical and hydrological analysis of flow data. Various types of analysis are normally required for flow data these are: computation of basic statistics,

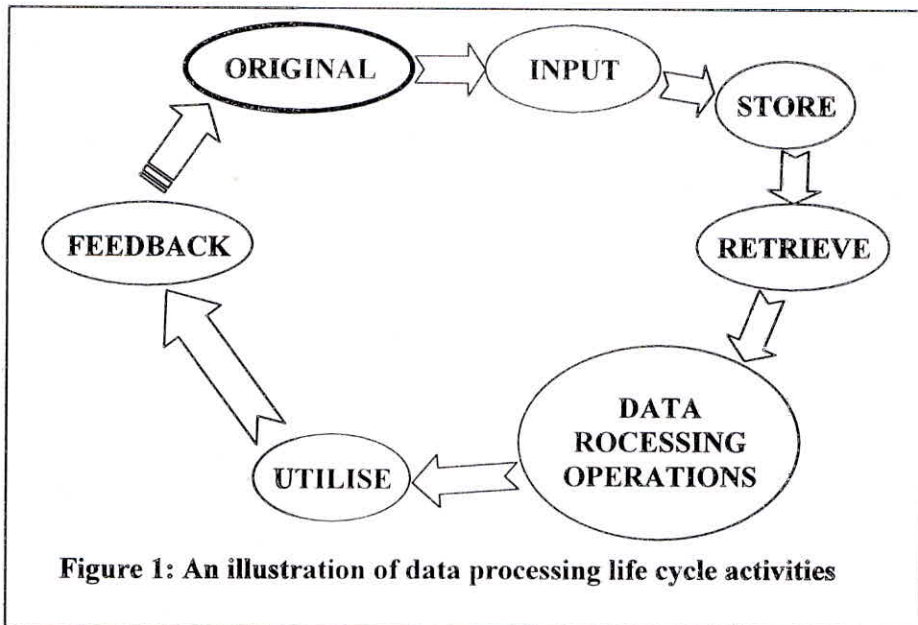
National Institute of Hydrology, Roorkee-247667 (Uttarakhand)

empirical frequency distributions and cumulative frequency distributions, establishment of rating curve, flow duration curves, fitting of theoretical frequency distributions, time series analysis, mass curves, residual mass curves, regression/relation curves etc. Needs and methodology of these analyses are explained with illustrative example to enable participants to achieve the capability of carrying out the analysis themselves.

### **HYDROLOGICAL DATA PROCESSING**

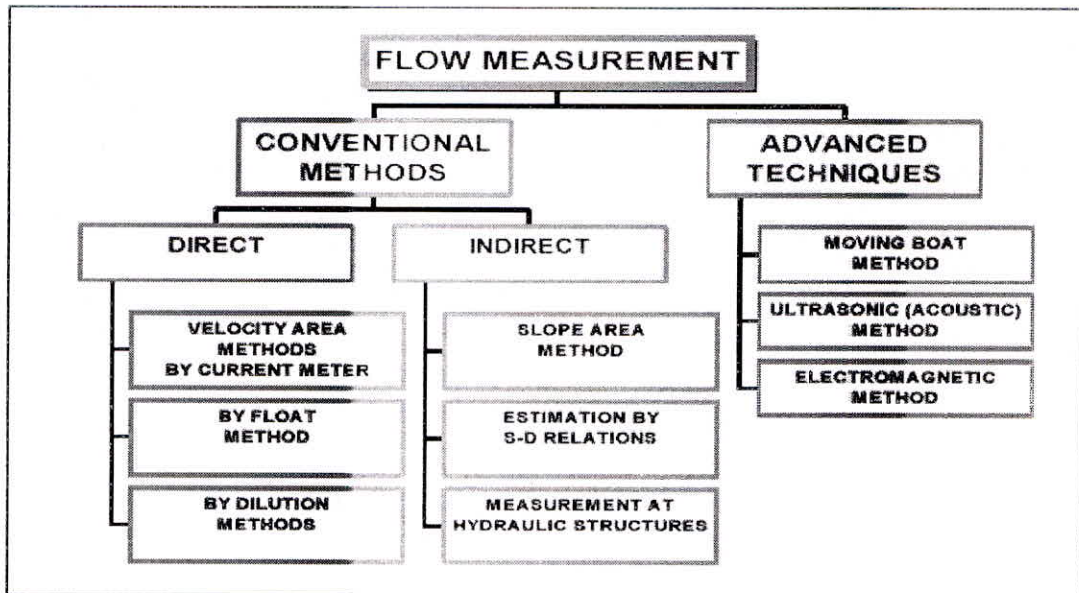
The term hydrological data processing refers to the recording and handling of hydrological data that are necessary to convert it into more refined or useful form of information. The volume of data has grown to such proportions that data processing has become a major activity attracting a great deal of interest. The data processed should not be taken as an end in itself. It is rather the beginning step of achieving objectives which can vary as the nature of data. The objectives of data processing now extend far beyond the routine handling of transaction documents and records of other types, providing timely information to facilitate greater control and improved decisions.

The data processing activities which come in cyclic way is called data processing cycle. In the data processing cycle first the information are gathered for data processing work. This is known as origination and collection. Then the collected data is considered as input for the processing. Subsequently it is stored for further processing and retrieval. The result and information obtained, as the products of data processing are subjected to feed back, which can be either negative or positive. Depending upon the type of feed back, corrections or modification are incorporated to get the new system. The obtained new system may be run for the subsequent operations. The data processing life cycle activities are illustrated in Figure 1.



### Flow measurement Methods

The following chart shows the classification of various methods of discharge measurement.



**Water stage:** Vertical distance of the water surface of a stream relative to a gauge datum. Stage or water level is the elevation of water surface above an established datum; it is the basic measure representing the state of a water body. Records of stage are used with a stage-discharge relation in computing the record of stream discharge. The reliability of the discharge record is dependent on the reliability of the stage record and on the stage discharge relation.

**Gauge datum:** Vertical distance of the zero of a gauge referred to a certain datum level. The datum is sometimes taken as a mean sea level but more often is slightly below the point of zero flow in the stream.

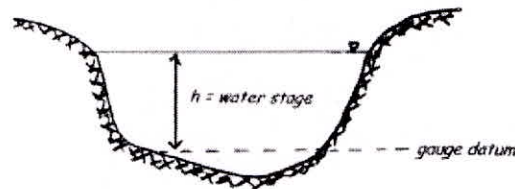


Figure 2: Cross section of a river with gauge datum and water stage

**Measurement devices:** (a) Non-recording, manual gauges, (b) Recording gauges

Discharge measurement

**Methods based on velocity measurements**

Discharge is the product of cross-sectional area of flowing water and its velocity.

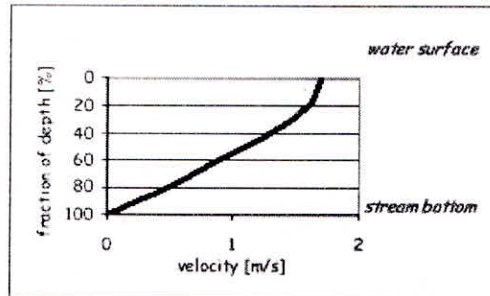
$$Q = v * A$$

Q = discharge [m<sup>3</sup>/s]

v = velocity [m/s]

A = cross-section of flow [m<sup>2</sup>]

**Velocity profile:** The water molecules in a stream are traveling at different speed. The water molecules in a stream are subject to friction as they come into contact with the sides and bottom



of the channel. Due to these frictional effects, water flows fastest at the surface and center of the channel (away from the immediate frictional influences). A typical velocity profile is shown below (Figure 3) where the velocity varies as a parabola from zero at the channel bottom to a maximum at (or near) the surface. Flow velocity can be measured by (a) Floats or (b) Current Meters

Figure 3: Typical velocity profile

**Isovel:** Line of equal velocity in a cross section of a watercourse.

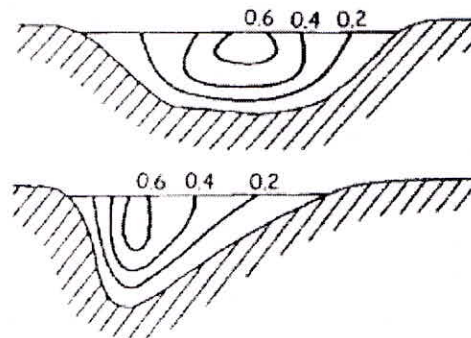


Figure 4: Velocity distribution in channel cross section (isovels in m/s)

## PROCESSING AND ANALYSIS OF FLOW DATA

### **Preliminary Processing and Scrutiny**

Preliminary Processing and scrutiny of the data are essential before the observed data is stored on computer. The preliminary processing is involved with comparisons within a single data series and is concerned with making comparisons between observations and pre-set limits and/or statistical range of a variable or with the expected hydrological behaviour of a hydrological phenomenon. Preliminary data processing highlights those data, which are not within the expected range or are not hydrologically consistent. To ensure data quality following information should also be furnished:

- i. Methods of measurement/observation of hydrological data, standards followed, instruments used, frequency of observation etc. shall be discussed item wise.
- ii. Details of history of station, shift in the location, shift in the rating curves should be identified. Sample calculation for discharge should be furnished. Mention shall be made as to whether discharge data is observed or estimated from the rating curve. Indicate the methods of estimation.
- iii. Discuss development of stage discharge curves at discharge site bringing out the extrapolations shall be verified by other methods such as hydraulic calculations etc.

**Filling up of short Data gaps:** The following techniques can be used for data gap filling:

- i. Random choice from values observed for that period.
- ii. Interpolation from adjoining values by plotting a smooth hydrograph (for runoff alone).
- iii. Double mass curve techniques.
- iv. Correlation with adjoining station either of the same hydrologic element or different hydrologic elements.
- v. Auto correlation with earlier period at the same station.

**Consistency of data:** The consistency of the data should also be checked to ensure

**Internal consistency check:** The study of consistency of the observed data at specific control points and corrections if any made shall be checked and discussed. The check can be done by study of stage discharge relationship for different periods. Large variations if any should be investigated, corrected and explained suitably if required. Trend analysis should be performed

- To detect a slow continuous variation of meteorological conditions or a long periodic variation of the climate.

- To observe the modification of catchment physiography especially through human activity.

**External consistency check:** The consistency of the observed stream flow data should be discussed with reference to the rainfall in the project catchment and observed data in adjacent locations/basins. The consistency can be checked by:

- Comparing monthly and annual rainfall with corresponding runoff.
- Comparing average annual specific flow expressed in depth unit with corresponding figures at other sites of the same river or adjacent basins.
- Comparing the hydrograph or daily discharge at the control point with adjacent sites etc.

**Quality control procedures:** Some of the methods for quality control are

- Testing the stage or discharge of a given day within a year against the highest and lowest value of the same date in all the previous years.
- Apply the same test on the difference between the value on the day and the day before.
- Comparing observed data with estimates based on data from adjacent stations.
- Comparing the observed data with estimates based on a precipitation runoff.
- Checking for negative values during the computation of inflow to a reservoir when the stage storage relationship and the outflow are known.
- Comparing the runoff at a station with runoff at upstream stations.
- Applying double mass curve analysis to identify shift in control.
- Applying time series analysis to detect changes in the homogeneity in time series. This is a valuable supplement to double mass analysis.
- Plotting a graph of the points at which measurements are made and comparison with the original cross section.
- Plotting the graph of the annual regime of specific discharges and regional comparison.
- Regional comparisons of monthly and annual streamflow deficits.

#### **Secondary Validation of Discharge Data**

The quality and reliability of a discharge series depends primarily on the quality of the stage measurements and the stage discharge relationship from which it has been derived. In spite of their validation, errors may still occur which show up in discharge validation. Discharge errors  
National Institute of Hydrology, Roorkee-247667 (Uttarakhand)

may also arise from the use of the wrong stage discharge relationship, causing discontinuities in the discharge series, or in the use of the wrong stage series. Validation of discharge is designed to identify such problems. The principal emphasis is in the comparison of the time series with neighbouring stations but preliminary validation of a single series is also carried out against data limits and expected hydrological behaviour. Various methods for secondary data processing of flow data are:

- Single station validation
- Multiple station validation
- Comparison of streamflow and rainfall

### ***Single Station Validation***

Single station validation is carried out by the inspection of the data in tabular and graphical form. Validation provides a means of identifying errors for correcting and completing the series. The following methods are available under single station validation.

#### ***Validation against data limits***

Data are checked numerically against, absolute boundaries, relative boundaries and acceptable rates of change. Further, individual values in the time series are flagged for inspection. The following data limits are assigned:

#### ***Absolute boundaries***

Values which exceed a maximum specified or fall below a specified minimum are flagged. The specified values are considered as the absolute values of the historic series. The object is to screen out spurious extremes, but care is taken not to remove or correct true extreme values as these are generally the most important values in the hydrological time series.

#### ***Relative boundaries:***

A larger number of values are flagged by specifying boundaries in relation to departures ( $a$  and  $b$ ) from the mean of the series ( $Q_{mean}$ ) by some multiple of the standard deviation ( $s_x$ ), i.e.

$$\text{Upper boundary } Q_u = Q_{mean} + a s_x$$

$$\text{Lower boundary } Q_l = Q_{mean} - b s_x$$



Whilst  $Q_{mean}$  and  $s_N$  are computed by the program, the multipliers  $a$  and  $b$  are inserted with values appropriate to the river basin being validated. The object is to set limits in such a way so that it can screen out a manageable number of outliers for inspection whilst giving reasonable confidence that all suspect values are flagged. This test is normally used with respect to aggregated data of a month or greater.

### ***Rates of change***

Values are flagged where the difference between successive observations exceeds a specified value. The specified value is greater for large basins in arid zones than for small basins in humid zones. Acceptable rates of rise and fall are specified separately. Generally an allowable rate of rise is greater than allowable rate of fall. For looking at the possible inconsistencies a listing of only those data points, which are beyond certain boundaries, are obtained.

### ***Graphical Validation***

Graphical inspection of the plot of a time series provides a very rapid and effective technique for detecting anomalies in the data. The discharge may be displayed alone or with the associated stage measurement. Stage and discharge time series plot can reveal any discontinuity which may appear between successive monthly updates of the data series. The main purpose of graphical inspection is to identify any abrupt discontinuities in the data or the existence of positive or negative 'spikes' which do not conform with expected hydrological behaviour.

### ***Multiple Station Validation***

#### ***Comparison plots***

The simplest and often the most helpful means of identifying anomalies between stations are in 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. Fig 5 presents a comparative plot of multiple discharge series of adjacent stations.

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.

## Rainfall-Runoff Modelling (March 11-15, 2013)

When two or more stations display the same behaviour there is strong evidence to suggest that the values are correct.

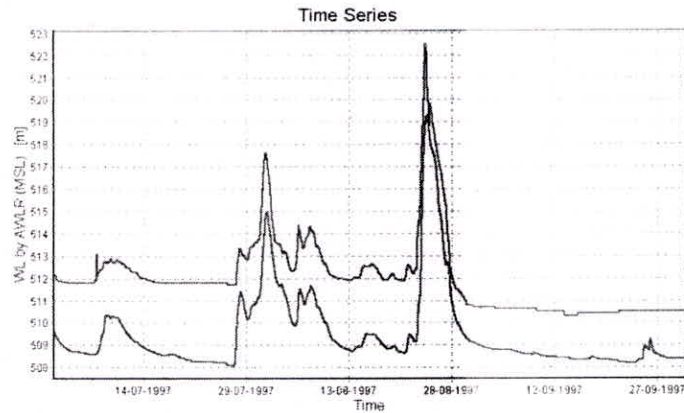


Figure 5: Plot of multiple discharge series of adjacent stations

### Residual series

An alternative way of displaying comparative time series is to plot their differences. This procedure may be applied to river flows along a channel to detect anomalies in the water balance. Both the original time series and their residuals can be plotted in the same figure. Any anomalous behaviour 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 water balance (Figure 6) are computed for the series presented in Figure 5.

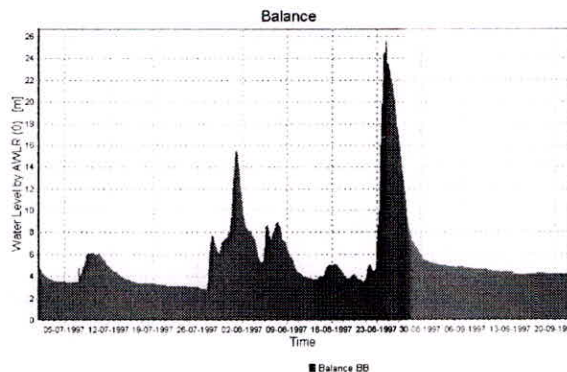


Figure 6: Water balance between two stations

### Double mass curves

Double mass curve analysis can also be used to show trends or inhomogeneities between flow records at neighbouring stations and is normally used with aggregated series. A difficulty in double mass curves with streamflow is in the identification of which if any station is at fault; this may require intercomparisons of several neighbouring stations.

### Comparison of Streamflow and Rainfall

The principal comparison of streamflow and rainfall is done through hydrological modelling. However, a quick insight into the consistency of the data can be made by graphical (Fig 7) and tabular comparison of areal rainfall and runoff. Basically the basin rainfall over an extended period such as a month or year should exceed the runoff (in mm) over the same period by the amount of evaporation and changes in storage in soil and groundwater. Tabular comparisons should be consistent with such physical changes.

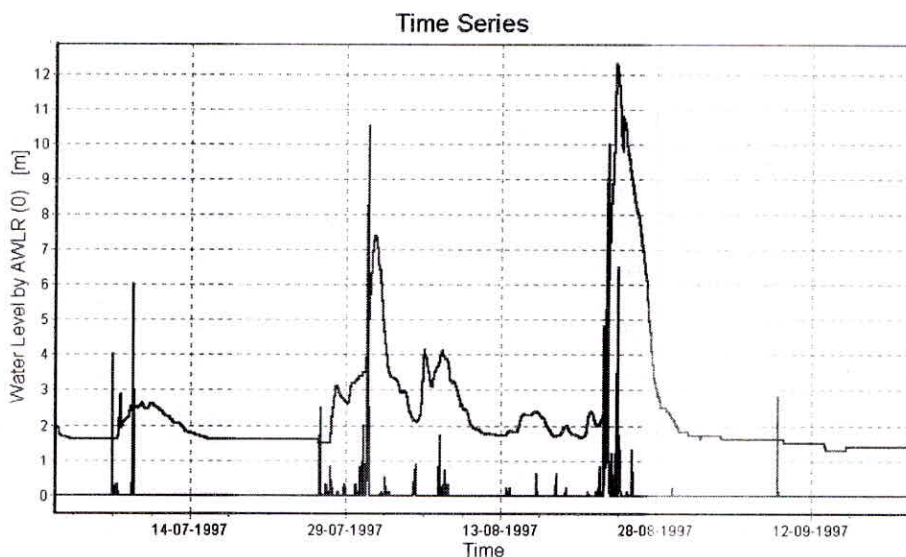


Figure 7: Comparison of streamflow and rainfall

### *Transformation of Stage to Discharge*

#### **Establishment of Rating Curve**

Quantitative records of time series of water level (stage) and flow in rivers constitute the bulk of hydrological data. Direct measurement of water level data is made using a variety of equipment. Measurement of flows is comparatively difficult to accomplish and is therefore estimated indirectly by using the stage-discharge relationship. However, intermittent

measurements of flow are made for the establishment of the stage-discharge relationship. The direct observation of discharge is made only for the purpose of establishing the stage-discharge relationship. Once the relationship is established, the same can be used to transform observed stages into a derived time series of discharge. The procedure used to transform stage to discharge depends on physical conditions at the station and in the river reach downstream. Generally, the following alternatives are considered:

- single channel rating curve
- compound channel rating curve
- rating curves with unsteady flow correction
- rating curves with constant fall backwater correction
- rating curves with normal fall backwater correction

Figure 8 presents a Rating curve established for a typical gauging site.

#### **Development of Rating Curve Under Permanent Control**

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:

$$Q=a(H-H_0)^b \quad (1)$$

in which  $Q$  = stream discharge,  $H$  = gauge height (stage),  $H_0$  = 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 an arithmetic or logarithmic plot). Logarithmic plotting is advantageous as Eqn. (1) plots as a straight line in logarithmic coordinates. The advantage of using the double logarithmic plot is two fold. Firstly, the plot would produce a straight line since the general form of rating curve is parabolic. Secondly, different straight lines allow to further grouping of data. 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  $H_0$  is necessary. As a first approximation the value of  $H_0$  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  $H_0$  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.

### Least square method

The best values of  $a$  and  $b$  in Eqn. (1) for a given range of stage can be obtained by the least-square-error method. Thus by taking logarithms, Eqn.(1) may be represented as

$$\log Q = \log a + b \log (H - H_0)$$

$$\text{or } Y = mX + c \quad (2)$$

in which the dependent variable  $Y = \log Q$ , independent variable  $X = \log (H-H_0)$  and  $c = \log a$ . The values of the coefficients for the best-fit straight line using data of  $N$  observations of  $X$  and  $Y$  are :

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

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

The Eqn. (2) 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  $H_0$ , plot  $\log Q$  vs  $\log (H - H_0)$  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  $H_0$  which gives a straight line plot of  $\log Q$  vs  $\log (H-H_0)$ .

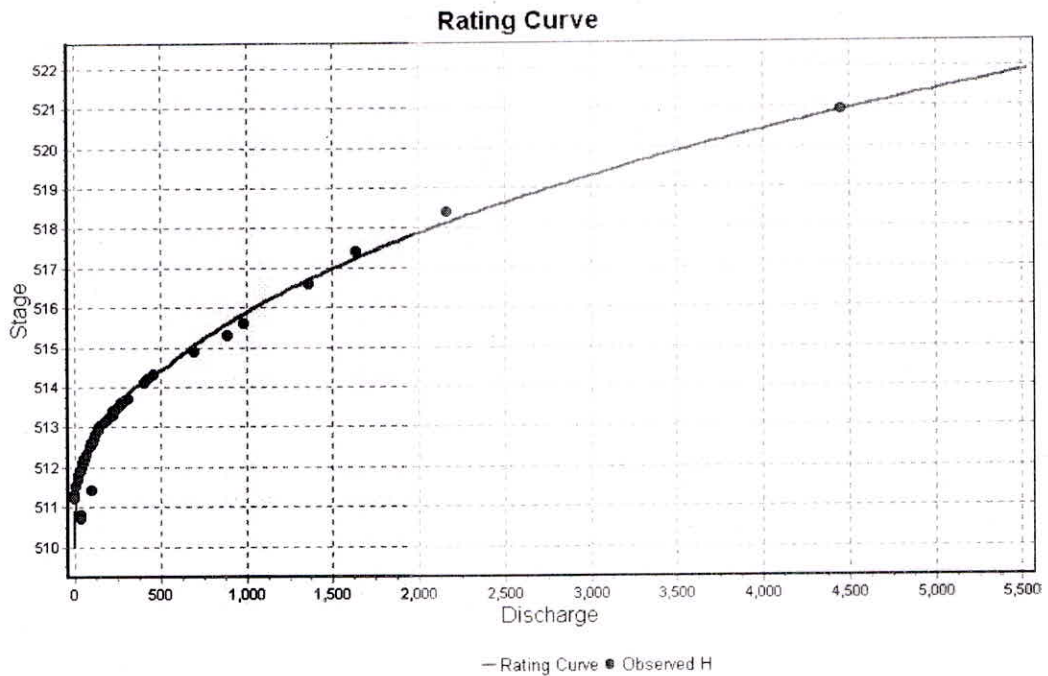


Figure 8: Stage Discharge Rating curve

### Estimation of discharge

The rating relationship is used to transform the observed stages into the corresponding discharges. In its simplest form, a rating curve can be illustrated graphically, as shown in Figure 8, by the average curve fitting the scatter plot between water level (as ordinate) and discharge (as abscissa) at any river section. In general the discharge data derived for shorter time interval are used for frequency analysis and rainfall runoff modeling.

### Extrapolation of rating curve

It is difficult to measure flow at very high and low stages due to their infrequent occurrence and also to the inherent difficulty of such measurements, extrapolation is required to cover the full range of flows. The rating curve may fall short at both the lower and the upper end. Extreme flows are often the most important for design and planning and it is important that the best possible estimates are made. Extrapolation is not simply a question of extending the rating curve from existing gaugings to extreme levels (although in some cases this may be acceptable); a different control may apply, the channel geometry may change, flow may occur

National Institute of Hydrology, Roorkee-247667 (Uttarakhand)

over the floodplain and form and vegetation roughness coefficients may change. Applicable methods of extrapolation depend on the physical condition of the channel, whether inbank or overbank and whether it has fixed or shifting controls. Consideration must also be given to the phenomenon of the kinematic effect of open channel flow when there may be reduction in the mean velocity in the main channel during inundation of the flood plain.

### **High flow extrapolation**

The following methods are considered:

- double log plot method
- stage-area / stage-velocity method
- the Manning's equation method
- the conveyance slope method

### **Low flow extrapolation**

Manual low flow extrapolation is best performed on natural graph paper rather than on logarithmic graph paper because the co-ordinates of zero flow can not be plotted on such paper. An eye-guided curve is drawn between the lowest point of the known rating to the known point of zero flow, obtained by observation or by survey of the low point of the control. There is no assurance that the extrapolation is precise but improvement can only come from further low flow discharge measurements. However low flows persist for a sufficient period for gaugings to be carried out and there is little physical difficulty in obtaining such measurements.

### ***Analyses of Discharge Data***

The purpose of hydrological data processing is not primarily hydrological analysis. However, various kinds of analysis are required for data validation and further analysis may be required for data presentation and reporting. These can be categorized as:

- aggregation of data
- computation of basic statistics

## Rainfall-Runoff Modelling (March 11-15, 2013)

- empirical frequency distributions and cumulative frequency distributions (flow duration curves)
- fitting of theoretical frequency distributions
- Time series analysis
  - moving averages
  - mass curves
  - residual mass curves
  - balances
- regression/relation curves
- double mass analysis
- series homogeneity tests
- rainfall runoff simulation

### **Aggregation of data**

Aggregation is the process by which discharge at its observational or recorded time interval and units is transformed to higher time interval. Aggregation is required for validation, analysis and reporting of discharge data. Discharge and its precursor water level is observed at different time intervals, but these are generally one day or less. Manual observation may be daily, hourly for part of the day during selected seasons, or some other multiple of an hour. For automatic water level recorders a continuous trace is produced from which hourly level and hence discharge is extracted. From digital water level recorders, level is usually recorded at hourly intervals though for some small basins the selected interval may be 15 or 30 minutes. Sub-hourly, hourly and sub-daily discharges, computed from levels, are typically aggregated to daily mean. Daily data are typically averaged over weekly, ten daily, 15 daily, monthly, seasonally or yearly time intervals. Averaging over longer time intervals is required for validation and analysis. For validation small persistent errors may not be detected at the small time interval of observation but may more readily be detected at longer time intervals.

### **Computation of Basic Statistics**

Basic statistics are widely required for validation and reporting. The following statistical properties can directly be computed:

National Institute of Hydrology, Roorkee-247667 (Uttarakhand)



- arithmetic mean
- median
- mode
- standard deviation
- skewness or the extent to which the data deviate from a symmetrical distribution
- kurtosis or peakedness of a distribution

### **Empirical Frequency Distributions (Flow Duration Curves)**

A popular method of studying the variability of streamflow is through flow duration curves which can be regarded as a standard reporting output from hydrological data processing. Some of their uses are:

- evaluation of dependable flows in the planning of water resources engineering projects
- evaluation of the characteristics of the hydropower potential of a river
- assessment of the effects of river regulation and abstractions on river ecology
- design of drainage systems
- flood control studies
- computation of the sediment load and dissolved solids load of a river
- computation of flows with adjacent catchments.

A flow-duration curve is a plot of discharge against the percentage of time the flow has equaled or exceeded. This may also be referred to as a cumulative discharge frequency curve and it is usually applied to daily mean discharges.

### **Fitting of Frequency Distributions**

The fitting of frequency distributions to time sequences of streamflow data is widespread whether for annual or monthly means or for extreme values of annual maxima or minima. The principle of such fitting is that the parameters of the distribution are estimated from the available sample of data, which is assumed to be representative of the population of such data. These parameters can then be used to generate a theoretical frequency curve from which discharges with given probability of occurrence (exceedence or non-exceedence) can be computed.

Generally, the parameters are known as location, scale and shape parameters which are equivalent for the normal distribution to:

- location parameter mean (first moment)
- scale parameter standard deviation (second moment)
- shape parameter skewness (third moment)

In other distributions parameters different from mean, standard deviation and skewness are used. Frequency distributions for data averaged over long periods such as annual are often normally distributed and can be fitted with a symmetrical normal distribution, using just the mean and standard deviation to define the distribution. Data become increasingly skewed with shorter duration and need a third parameter to define the relationship. Even so, the relationship tends to fit poorly at the extremes of the data which are often of greatest interest. This may imply that the chosen frequency distribution does not perfectly represent the population of data and that the resulting estimates may be biased.

#### **Frequency Distributions of Extremes**

In case of extremes of time series, either of floods or droughts, theoretical frequency distributions are most commonly applied. For fitting a distribution the following series are required:

- maximum of a series: The maximum instantaneous discharge value of an annual series or of a month or season may be selected. All values (peaks) over a specified threshold may also be selected. In addition to instantaneous values maximum daily means may also be used for analysis.
- minimum of a series: With respect to minimum the daily mean or period mean is usually selected rather than an instantaneous value which may be unduly influenced by data error or a short lived regulation effect.

The objective of flood frequency analysis is to assess the magnitude of a flood of given probability or return period of occurrence. Return period is the reciprocal of probability of exceedence and may also be defined as the average interval between floods of a specified magnitude.

A large number of different or related flood frequency distributions have been devised for extreme value analysis. These include:

- Normal and log-normal distributions and 3-parameter log-normal
- Pearson Type III or Gamma distribution
- Log-Pearson Type III
- Extreme Value type I (Gumbel), II, or III and General extreme value (GEV)
- Logistic and General logistic
- Goodrich/Weibull distribution
- Exponential distribution
- Pareto distribution

Different distributions provide fit best to different individual data sets but if it is assumed that the parent population is of single distribution of all stations, then a regional best distribution may be recommended.

There is no single distribution that represents equally the population of annual floods at all stations, and one has to use judgement as to which to use in a particular location depending on experience of flood frequency distributions in the surrounding region and the physical characteristics of the catchment.

A standard statistic which characterises the flood potential of a catchment and has been used as an 'index flood' in regional analysis is the mean annual flood, which is simply the mean of the maximum instantaneous floods in each year. This can be derived from the data or from distribution fitting. An alternative index flood is the median annual maximum, similarly derived. These may be used in reporting of general catchment data. Flood frequency analysis may be considered a specialist application required for project design and is not a standard part of data processing or validation.

### **Time Series Analysis**

Time series analysis may be used to test the variability, homogeneity or trend of a streamflow series or simply to give an insight into the characteristics of the series as graphically displayed. The following analysis are included:

National Institute of Hydrology, Roorkee-247667 (Uttarakhand)

## Rainfall-Runoff Modelling (March 11-15, 2013)

- moving averages
- residual series
- residual mass curves
- Run length and run sum characteristics
- Storage analysis
- Balances

Residual mass curve and moving average curve for a typical gauging site are given in figure 9 and figure 10 respectively.

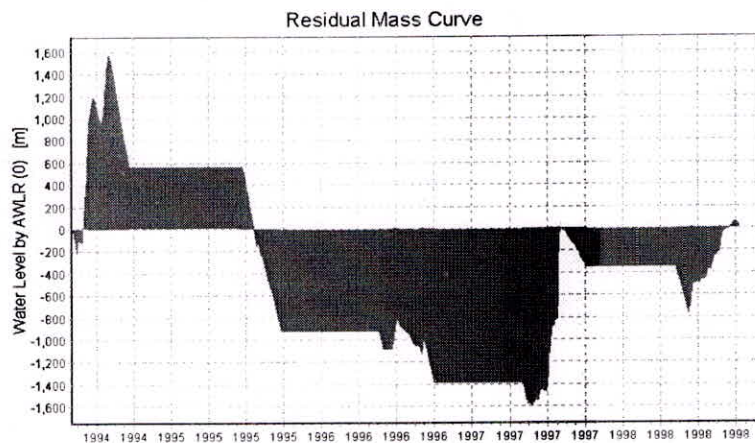


Figure 9: Residual mass curve

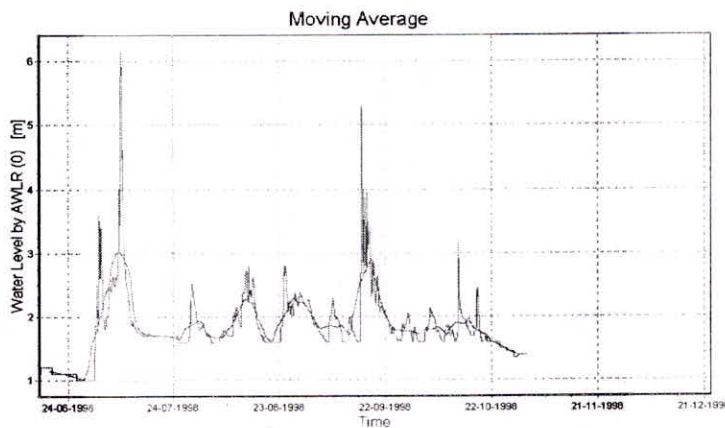


Figure 10: Moving average of daily discharge data  
National Institute of Hydrology, Roorkee-247667 (Uttarakhand)

### **Regression /Relation Curves**

Regression analysis and relation curves are widely used in validation and for the extension of records by the comparison of the relationship between neighbouring stations. Further more, regression analysis has wide applications in hydrology for developing the relationships in different forms.

### **Series Homogeneity Tests**

The following series homogeneity tests are commonly used:

- Student's t test for the stability of the mean
- Wilcoxon-W test on the difference in the means
- Wilcoxon-Mann-Whitney U-test

Series homogeneity tests may be applied to streamflow but it should also be recognised that inhomogeneity of streamflow records can arise from a variety of sources including:

- data error
- climatic change
- changes in land use in the catchment
- changes in abstractions and river regulation

### **Rainfall Runoff Simulation**

The uses of a rainfall-runoff model are much wider than data validation and include the following:

- filling in and extension of discharge series
- generation of discharges from synthetic rainfall
- real time forecasting of flood waves
- determination of the influence of changing landuse on the catchment (urbanisation, afforestation) or the influence of water use (abstractions, dam construction, etc.)

### **REMARKS**

Hydrological data processing is a multiple step process. It has to be carried out in a series of stages, starting with preliminary checking in the manuscript, through entry of raw data and

successively higher levels of validation, before it is accepted as fully validated data. Automated data processing using computer can help in repetitive computation with such high speed and accuracy that one can now afford to process extensive hydrological data.

**REFERENCE**

1. Chowdhary H. (1992). Training report on "Hydrological Data Processing", March 1992.
2. *HYMOS (Version 4.01) Manual*. (1999). Delft Hydraulics, p.o. box 177, 2600 M.H. Delft.
3. Lohani A.K. (1993). *Hydrological data collection processing and analysis*, Report No. SR-27, National Institute of Hydrology, Roorkee.
4. Lohani A.K., Chowdhary H, Singh R.D. (2000). Use of GIS based HYMOS for surface water data processing,, Spatial Information Technology-Remote Sensing and Geographical Information Systems, BS Publications, Hyderabad, India.
5. Lohani A.K., R. Kumar and C. Chatterjee, (2001). Computer based surface water data processing techniques and tools, Impact of Information Technology on Quality of Life-Future Strategy, Proceedings of the 16<sup>th</sup> National Convention of Computer Engineers, The Institute of Engineers (India) Bihar State Centre. Dec 15-16, pp137-143.
6. Lohani A.K., R.D. Singh, R. Kumar (2004). Application of HYMOS for processing gauge and discharge data, Workshop on Establishment of Water Resources Data Processing and Data Storage Centres, 17-19 Feb, 2004, State Water Resources Agency, Govt. of Uttar Pradesh, Lucknow, pp. 73-84.
7. Lohani A.K., R.D. Singh, R. Kumar (2004). Rainfall data processing using HYMOS, Workshop on Establishment of Water Resources Data Processing and Data Storage Centres, 17-19 Feb, 2004, State Water Resources Agency, Govt. of Uttar Pradesh, Lucknow, pp. 85-91.