भारत 2023 INDIA
वसुधैव कुटुम्बकम्
ONE EARTH • ONE FAMILY • ONE FUTURE

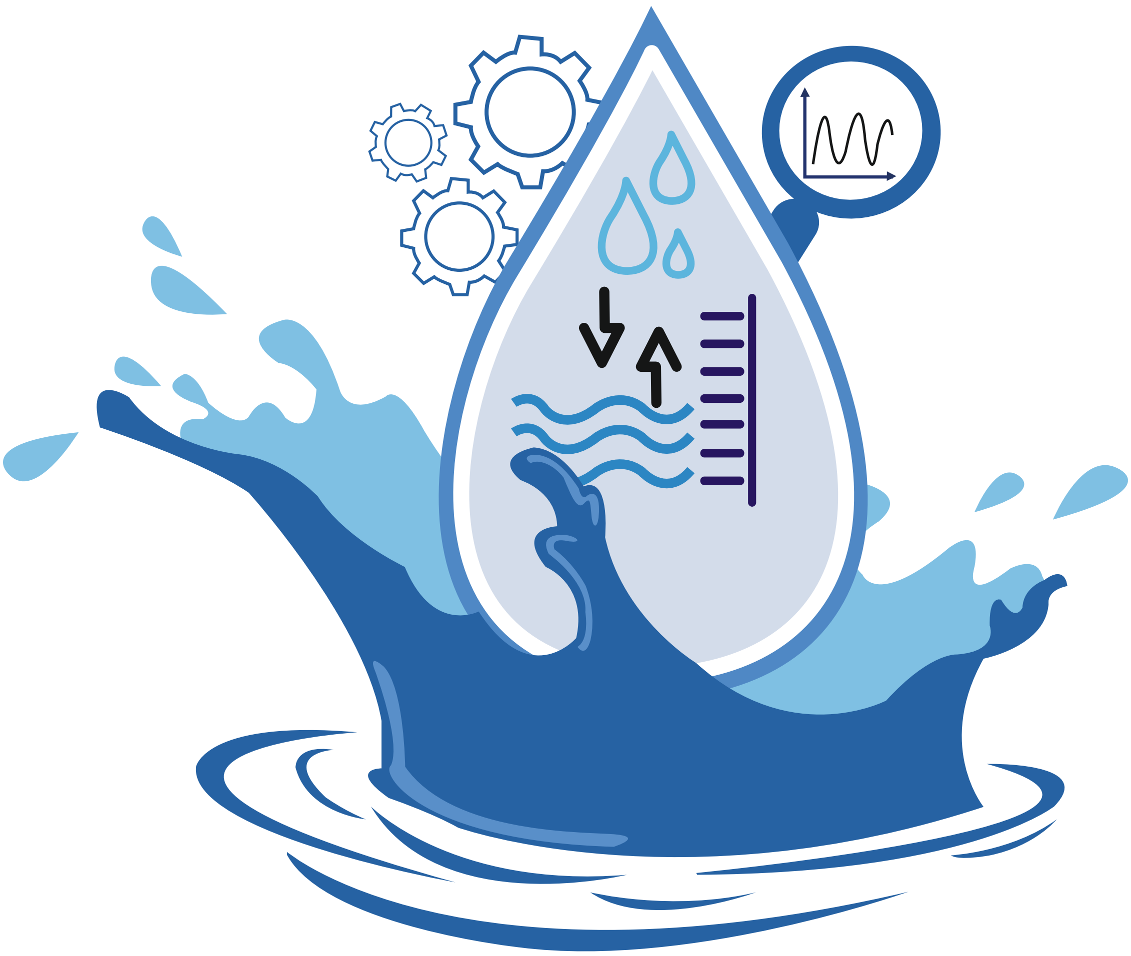
सत्यमेव जयते

Government of India

Azadi Ka Amrit Mahotsav

MANUAL ON

WATER LEVEL AND DISCHARGE DATA: VALIDATION, ANALYSES, PROCESSING AND MODELLING



NATIONAL HYDROLOGY PROJECT

DEPARTMENT OF WATER RESOURCES, RD & GR
MINISTRY OF JAL SHAKTI

Government of India

Ministry of Jal Shakti, Department of Water Resources, River Development and Ganga Rejuvenation

Manual on

Water Level and Discharge Data: Validation, Analyses, Processing and Modelling

National Hydrology Project

December 2022

Ministry of Jal Shakti, Department of Water Resources, RD & GR
National Hydrology Project
2nd Floor, Rear wing, MDSS (MTNL) Building, 9, CGO Complex,
Lodhi Rd, New Delhi 110003
E-mail: nhp-mowr@gov.in
Website: https://nhp.mowr.gov.in

Title: Manual on Water Level and Discharge Data: Validation & Analysis, Processing and Modelling

Summary:

This manual has been prepared by the Technical Assistance and Management Consultancy (TAMC) under the National Hydrology Project for the Department of Water Resources, River Development and Ganga Rejuvenation. It presents a compilation of standard methods used for the validation of water level and discharge data. The manual updates and further develops the manuals prepared under the previous HP-I and HP-II projects.

Date: December 2022

Sh. Anand Mohan
Joint Secretary (RD &PP)
Department of Water Resources, RD&GR,
Ministry of Jal Shakti, Government of India

Preface

Hydrological observations of water level and discharge are the primary inputs for water resource assessment and management. They play a crucial role in supporting the informed decision-making process in water resources planning and demand management. Telemetry-basedwater level data monitoring integrated with real-time Decision Support Systems (DSS) has a huge potential in flood forecasting and its mitigation through the operation of water storage structures during flood events.

The concept of Hydrological Information Systems was introducedin India through the Hydrology Project Phase-I.This included trainings forthe officers of the Implementing Agencies and preparing manuals on the analysis and multi-stage validation of hydrological data. Phase-II of the Hydrology Project focused on the analytical framework for state-of-art flood forecasting and water resources planning and management.

Under the National Hydrology Project (NHP), the monitoring networks and the analytical tools for water resources planning and management are beingenhanced further. It includes developing web-based Water Resources Information Systems comprising time series data, geographical databases, alongwith various analytical tools and applications to benefit the stakeholders. Dissemination portals and mobile applications would also be developed.

Water level and Discharge data can have errors ingested at various levels, including erroneous field measurements, data entry and transfer of information. Data validation ensures that the data reaching water resource planners, designers and managers arefree from errorsandcomplete. This Manual describes the checks to be applied for detection of errors and techniques of validation for the Water level and Discharge data. The current Manual is an effort to provide a ready reference for a variety of users, including water resources planners, hydrologists, site and field engineers, designers, and water systems operators.

The Manual, consisting of 15 chapters, goes into length about the principles of primary and secondary data validation, correction, compilation, completion, analysis, and report generation. The manual also addresses the critical subject of extrapolating rating curves, focusing on peak discharge predictions, the most crucial input in flood studies. Subsequently, an effort has been made to introduce the concept of basin modelling principles.

I am confident that this document will be useful to a wide range of water professionals at different levels, not only from Implementing Agencies of NHP but toa larger audience in the water sector including the students- our water managers of the future. It will bring us one step ahead to address the challenges of water resources management in India.

Sh. Anand Mohan
Joint Secretary (RD &PP)
DoWR, RD & GR
Ministry of Jal Shakti

Contents

List of Figures

List of Tables

Abbreviations and Acronyms

ARG	Automatic Rain Gauge
AWLR	Automatic Water Level Recorder
cm	Centimetre
CWC	Central Water Commission
H	Height
HP	Hydrology Project
IMD	India Meteorological Department
km	Kilometre
m	Metre
mm	Millimetres
N	Number
NHP	National Hydrology Project
P	Precipitation
SRG	Standard rain gauge
Stdev.	Standard deviation
TBRG	Tipping Bucket Rain Gauge
WMO	The World Meteorological Organisation

1 INTRODUCTION

1.1 Background

The importance of water is felt more strongly whenever there is an excess or a shortage. Though globally rivers hold less than 0.5% of the world's fresh water, they constitute an important source of water for human consumption. In countries like India where development has been taking place at a rapid rate, accurate measurement and assessment of the resources is an absolute necessity. This is intensified with the threat of climate change.

The measurement of water levels in the country has been carried out at the national scale by the Central Water Commission, and at the local scale by the Water Resources and Irrigation Departments of the States and Union Territories. Manual observations are most common with data recorded into registers. This often leads to errors in reading or data entry. Due to shortage of manpower or due to lack of dedication, data are sometimes not entered or the entries contain guessed values. Also, some erroneous data may be generated due to the malfunctioning of equipment. Data validation forms a crucial part of any exercise in the field of water resources – from yield to flood estimates. The procedures to check and validate the water level and discharge records are not very well studied in the academic institutions at the undergraduate engineering programmes, nor are they easily available as a single organised document. This Manual on Procedures for Validation of Water Level and Discharge Data is designed to help practitioners deal with the issues related to data maintenance, validation and processing. It contains a compilation of the most commonly available procedures in brief to accomplish those tasks.

1.2 The Need for the Manual

The primary goal of the Ministry of Jal Shakti, Department of Water Resources, River Development and Ganga Rejuvenation is to ensure optimal sustainable development, maintenance of quality and efficient use of water resources to match the growing water demands of the country. The Ministry is responsible for laying down policy guidelines and programmes for the development and regulation of the country's water resources. This includes providing technical guidance, scrutiny, clearance and monitoring of all aspects of water use.

Individual module reports were published under the earlier hydrology projects (HP-I and HP-II), which catered to the specific objectives of meeting the training needs on correction and completion of water level data, making data entry for flow measurement data, carrying out primary validation of stage-discharge data, the establishment, validation and extrapolation of stage-discharge rating curves, and carrying out secondary validation of stage and discharge data. To make it compatible with the slow internet speed available in those days, there were restrictions on the file size to ensure its successful download. Subsequent developments of hardware and software and the wide range availability of freeware have simplified many cumbersome tasks. The availability of data for hydrologic analyses has also significantly improved with the additional developments of the India WRIS/ WIMS website and other websites in the public domain.

The National Hydrology Project has been approved by the Cabinet on 6.4.2016 as a central sector scheme, with a further objective to improve the extent, quality, and accessibility of water resources information, decision support systems for floods and basin level resource assessment and planning, and to strengthen the capacity of targeted water resources professionals and management institutions in India. It includes the development of a series of new and revised manuals and guidelines. These include guidance documents such as this one that are applicable nationwide.

1.3 Purpose and Scope of the Manual

Recently, the Ministry has approved sharing of restricted data of the Ganga Basin among the concerned states by the users with due administrative privileges. This is expected to open up possibilities for studies related to conception, planning and optimisation of future and existing projects, ensuring betterment of future water resource management. The goal of this manual is to compile the available techniques of processing, validation and analysis of water level and discharge data under a single volume that is available free of charge. Apart from professional practitioners, it is also expected to benefit the research community and the students in India and other countries. Even though it is advisable to follow the procedures described here, most of which are commonly accepted among practitioners, neither the authors of this manual nor the Ministry accept explicitly or implicitly any responsibility resulting from errors or erroneous use of these methods.

1.4 Arrangement

The manual is arranged to suit the need of the practitioners at the field with the least possible investment of time. For the sake of completeness of the chapter concerned, some repetitions have been allowed to remain, so that it is not necessary to go searching through the entire text in search for the solution to a particular issue. The first chapter sets the background for its preparation. The next chapter describes the sequences of processing water level and discharge data followed by the government departments responsible for its collection and storage in India. The third chapter describes the validation of water level data, the hydrological variable which is most commonly observed. In the subsequent chapter, correction and completion of the water level data has been dealt with in details. Chapter 5 introduces the methods for primary validation of stage-discharge data, step-by-step. The procedures adopted for secondary validation of stage-discharge data have been described in Chapter 6. Chapter 7 provides information on the ways for computation of discharge data, as the normal practice is to measure the water levels regularly and calculate discharge on the basis of water level records. The techniques for validation of discharge data have been explained in Chapter 8. Chapter 9 details the procedures adopted for correction and completion of discharge data. The step that immediately follows deals with the compilation of discharge data, following the methodologies suggested in Chapter 10. Chapter 11 attempts to present a few ways of analysing discharge data for the purpose of presenting and reporting. The specialised science and art of developing the rating curve or stage-discharge curve, which provides discharge from water level measurements, has been briefly introduced in Chapter 12. Ways for validation of the rating curve have been dealt with in Chapter 13. Chapter 14 takes up the crucial topic of extrapolating rating curves. It describes the procedure for preparing estimates of the peak flood discharge, which is the most important input in flood studies but seldom available in the historical records. The final chapter, Chapter 15, introduces the concept of river basin modelling.

1.5 Publication and Contact Information

This document is available on the website for the National Hydrology Project

https://www.nhp.mowr.gov.in/

For any further information contact

National Project Monitoring Unit (NPMU),

National Hydrology Project,

Department of Water Resources, RD & GR, Ministry of Jal Shakti,

2nd and 3rd Floor, Rear Wing,

9, CGO Complex,

MTNL Building, Lodhi Road,

New Delhi – 110003.

Email: nhp-mowr@gov.in

1.6 Acknowledgements

Preparation of this manual relied on the work of others in India and abroad. Grateful appreciation is extended to the following organizations whose publications and websites provided valuable information:

- National Hydrology Project (NHP)
- The World Bank
- The World Meteorological Organization (WMO)
- The original authors of the Hymos software

2 DATA PROCESSING SEQUENCES

2.1 General

Before carrying out any analysis / modelling with observed data, it is essential to carry out data validation to ensure data quality and consistency. The procedure includes flagging the suspect values and completing the series using standard procedures, which are described in the sections that follow.

2.2 Data Validation

The statistics on hydrological data is the basis of the water management policies and practices of the water resource initiatives of a nation. However, the observations are subject to errors arising at various levels from field measurement, data entry, data computation, transfer or correction. Data validation is a process that ensures that the stored values are reliable and the best possible representation of true values of the measurements. The processes under data validation are multi-level and parameter specific, broadly covered under a series of steps depicted in Figure 2.1.

Data Validation is carried out mainly for three reasons:

1. To correct errors in the recorded data wherever possible,
2. To assess the reliability of records where it is not possible to correct errors
3. To identify the source of errors and to ensure that such errors are not repeated in the future.

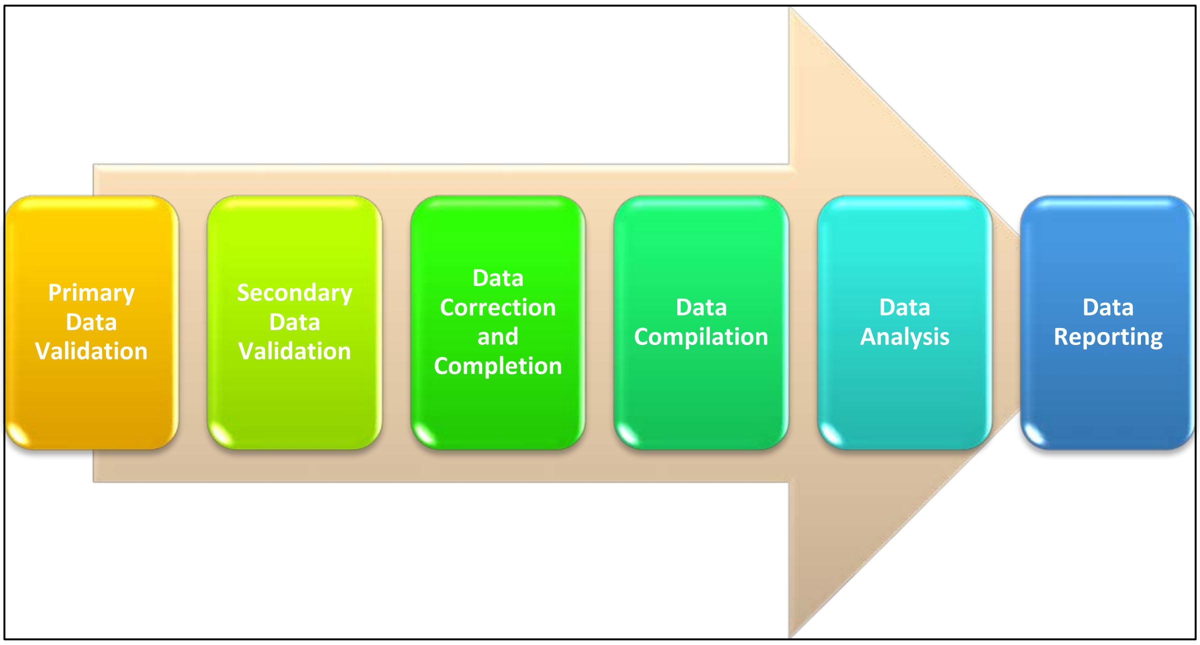


Figure 2.1: Multi-level Processes in Data Validation

By their nature, errors can be classified as random, systematic or spurious:

- **Random errors** are sometimes referred to as experimental errors and are equally distributed about the mean or 'true' value. The errors of individual readings may be large or small, e.g., the error in a gauge/ discharge reading where the water surface is subject to wind and wave action, but they tend to compensate with time or by taking a sufficient number of measurements.
- **Systematic errors** imply the existence of a systematic difference, either positive or negative, between the measured value and the true value, where the situation is not improved by increasing the number of observations. Hydrometric field measurements are often subject to systematic errors as well (e.g., error arising out of the erroneous setting of staff or recorder). Systematic errors are generally more serious. These are what the validation process is designed to detect and if possible, to correct. For example, water level values that do not fall below a particular threshold found in the data described by a number that is frequently found in the records indicates the limitation of the device. This could happen when the length of the rope that is attached to the float is insufficient to measure water levels during extreme droughts.
- **Spurious errors** are sometimes distinguished from random and systematic errors as arising due to some abnormal external factor. Such errors may be readily recognized but cannot easily be statistically analysed and the measurements are usually discarded.

2.2.1 Levels of validation

It is desirable to carry out data validation as soon as the data is observed. However, carrying out complete validation close to observation sites is impractical both in terms of technical support related to equipment and the available staff. The sequence of validation process has therefore been divided so that those which primarily require interaction with the observation station, are carried out nearby (i.e., at State Sub-divisional office) whereas the more complex validation procedure is carried out at higher levels. Essentially, data validation is a multi-stage process and sometimes a two-way approach.

Based on the sequence and level, data validation can be grouped into two major categories: Primary data validation and Secondary data validation.

2.2.1.1 Primary data validation for manual data observations

Primary data validation is presumed to be carried out immediately after the observations are made, data are extracted from charts or downloaded from loggers. This ensures that any obvious errors coming from the observer or instrument are spotted at the earliest and resolved. Primary validation checks the observation records within a single data series with pre-set limits, statistical range, to ensure conformance with the expected hydrological behaviour. Data from stations nearby may also sometimes be available and this may be used in primary validation.

Primary data validation highlights those data which are not within the expected range or are not hydrologically consistent. These data are then revisited in the datasheets or analogue records to see if there were any errors while making computations in the field or during manual data entry. If it is found that the entered values are different than the recorded ones, such entries should be immediately corrected. Where such data values are found to have been correctly entered, they are then flagged as doubtful with a remark against the value in the computer file, indicating the reason for such doubt.

Apart from data entry errors, the suspected values are identified and flagged but not amended at the Sub-divisional level. However, the flag and remarks provide a basis for further consideration of action at the time of secondary data validation.

2.2.1.2 Primary data validation for data observations using intelligent sensors

Some sensors contain software that can perform primary data validation procedures. Examples are alerts for values that are beyond the extreme limits set for the station based on an anticipated range, values that exceed the maximum anticipated rate of change compared to the previous observation.

2.2.1.3 Secondary data validation

Secondary data validation consists of a comparison of the variables at two or more stations or comparison of more than one variable. This is carried out to test the data against the expected behaviour of the system on a spatial scale. The underlying assumption is that the variables under consideration have adequate spatial correlation. This correlation is derived based on historical records and the statistics are utilized to validate the data. For certain hydrological variables like water level and discharge, which bear a very high degree of statistical dependence between adjacent stations, the inter-relationship can be established with a comparatively higher level of confidence. However, it is difficult to ascertain the behaviour with the desired level of confidence for some variables which lack good spatial correlation to variables from other stations and show significant variability. In such circumstances, it becomes very difficult, if not impossible, to detect errors.

While validating data based on a group of surrounding stations, the strategy must always be to rely on certain key stations known to be of good quality. If all the observation stations are given the status of being equally reliable, data validation will become comparatively more difficult. Field experience shows that the quality of data received from some stations is better than that received from other stations. Also, the process of allocating greater weights to a limited number of stations makes the data validation procedure simpler and faster to carry out. This may be due to physical conditions at the station, quality of instruments, or reliability of the staff. It must always be remembered that these key or reliable stations can also report incorrect data as there is no guarantee that any data series would be perfect.

Similar to primary data validation, the guiding factor for secondary data validation is that none of the test procedures should be considered as an objective on their own. They must always be taken as tools to screen out suspect data values. The validity of each of these suspect values is then confirmed based on other tests and corroborative facts based on the information received from all stations or other secondary information. Once it is clear that a certain value is incorrect and an alternative value provides a more reliable indication of the true value of the variable, a suitable correction should be applied and the value should be flagged as corrected.

If it is not possible to confidently conclude that the suspect value is incorrect, then such values should be left as recorded with a proper flag indicating doubt. All data which have been identified as suspicious at the level of primary validation are to be validated again based on additional information available from a larger surrounding area. If such data are supported by additional spatial information, they should be accepted as correct. Accordingly, the flags indicating them as doubtful must be removed at this stage.

2.3 Data In-filling (Completion) and Correction

Raw observed data may have missing values or sequence of missing values due to factors like equipment malfunction, observer absence, etc. These gaps should, where possible, be filled to make the series complete. Also, all values flagged as doubtful while carrying out the validation must be reviewed to decide whether they should be replaced by corrected values or whether doubt remains but a more reliable correction is not possible and the original value remains with a flag.

In-filling or completion of a data series is done in a variety of ways depending on the length of the gap, nature of the variable and availability of suitable records for estimation. The simplest case is where variables are observed with more than one instrument at the same site: the data from one gauge can be used to complete the data from the other gauge. For gap of a single value or short gaps in a series with high serial correlation, simple linear interpolation between known values or values filled with the graphical plot of the series may be acceptable. Gaps in series with high random component and poor serial correlation cannot be filled in this way and must be completed considering the neighbouring stations. It is important to take into account the effects of regulation and realize that the data from a station upstream of the reservoir cannot be used to fill the missing data located on a station downstream of the reservoir, unless the data at the downstream station have been processed into a naturalized flow series obtained by adjusting the downstream records by the amount of storage change and upstream water use.

Data correction is to be done using similar procedures as for completing the data series. There can be a shift in recorded values. The possible reasons can be due to identified systematic error or due to the change in control of the station. The data correction can involve techniques like Double Mass curve to adjust the portion of shift for the record to be consistent with the present and continuing data.

2.4 Data Compilation

Compilation refers primarily to the transformation of data observed at a certain time interval to a different interval, e.g., hourly to daily, daily to weekly, weekly to monthly, etc. This is done by a process of aggregation. Occasionally, disaggregation, or a conversion from longer to shorter time steps, may also be required, but it is usually not recommended due to the loss of accuracy of the resulting disaggregated data.

Various statistically derived series can also be created from the raw data. The examples of this include the maximum, minimum and mean for selected time intervals, or a listing of peaks over thresholds, to which a variety of hydrological analyses may be applied.

2.5 Data Analysis

Procedures used in data validation and reporting have wide analytical use. The following are examples of the available techniques:

1) Basic statistics (e.g., mean, standard deviations, etc.)
2) Statistical tests
3) Fitting of frequency distributions
4) Flow duration series
5) Regression analysis
6) Checking discharge with rainfall
7) Procedure to convert regulated flows into unregulated (natural) flow series.

Most of the above statistical analyses should be conducted on natural flow series.

3 VALIDATION OF WATER LEVEL DATA

3.1 Introduction to Primary Validation of Water Level Data

Stage or water level is the elevation of the water surface above an established datum. The water level records observed at a given location on a stream with a developed stage-discharge relationship at that particular site facilitates conversion of observed water levels into discharge estimates. The reliability of the discharge estimate is dependent on the reliability of the stage record and the stage-discharge relationship. The Stage or water level data is also used to define the state of water bodies for water management decisions that may involve reservoir operation, navigation, flood inundation and mitigation etc. The stage is usually expressed in metres.

Under the Hydrological Information System, Primary validation is proposed to be carried out at the sub-divisional level using the basic module of data processing software (formerly e-SWIS, now India-WRIS). It involves analyses of data at a single station by:

a) checking the data between individual observations and pre-set physical limits;

b) comparing the measurements of water levels at a single station, taken by staff gauge and by an automatic or digital water level recorder.

Before Primary Validation, data entry checks should be carried out to ensure that there have been no transcription errors from the field sheets to the database. This may flag some doubtful values. Where a doubtful or incorrect value is identified during Primary Validation, that value should be marked with a 'flag' to indicate that it is 'suspect'. In some instances, it may be possible to replace this value with the corrected value, in which case the replacement value is flagged 'corrected'. Otherwise 'suspect' values remain flagged for possible rectification during secondary validation. The missing values may be interpolated from stage records or discharge records depending on the nature and duration of missing data period or faulty records and the availability of other supported records.

3.2 Typical Instruments and Methods of Observation

Data validation must never be considered a purely statistical or mathematical exercise. Rather, it should be understood in the context of field practices. Three basic instruments are in use at river gauging stations for measurement of water level.

- Staff gauges or Manual Gauge
- Autographic water level recorders (AWLR) based on mechanical equipment

- Automatic/ Digital water level recorders (AWLR/ DWLR)

3.2.1 Staff or manual gauges

3.2.1.1 Instrument and Procedure

The staff gauge is the primary means of measurement at a gauging station, the zero of which depends on the geodetic elevation of the station usually referred to the mean sea level. It is read manually, and other recording gauges which may exist at a station are calibrated and checked against the staff gauge level. Staff gauges are located directly in the river. When the staff gauge is the only means of measurement at a station, observations are generally made once a day in the lean season and at multiple times a day during a flood period - even at hourly intervals during flood season on flashy rivers.

Figure 3.1: Staff Gauge or Manual Gauge at Chel, West Bengal

3.2.1.2 Typical measurement errors

Like other manual measurements, staff gauge readings are dependent on the observer's ability and reliability and it must not be assumed that these are flawless. Competency of the observer can be checked by the field supervisor, but the data processor must also be aware of typical errors made by observers.

A common problem to note is the misplacement of the decimal point for readings in the range 0.01 to 0.10. For example, a sequence of level readings on the falling limb of a hydrograph:

4.12, 4.10, 4.9, 4.6, 4.3, 4.1, 3.99 - should be interpreted as:

4.12, 4.10, 4.09, 4.06, 4.03, 4.01, 3.99.

Experience suggests that the records maintained by a single observer left unsupervised for extended periods may contain some 'estimated' readings, fabricated without visiting the station. Typical indicators of such 'estimates' show some sequences without any hydrological consistency such as,

- Abrupt falls or a sudden change in the slope of a recession curve.
- Long periods of uniform level followed by a distinct fall.
- Uniform mathematical sequences of observations (usually caused by copying previous records). For example, where the level falls regularly by the increments of 0.05 m or 0.10 m between readings for extended periods. Natural hydrographs have slightly irregular differences between successive readings and the differences decline as the recession progresses.

Besides, water level measurement may be difficult in high flows due to poor access to the gauge site, hazardous weather and wave action. Therefore, during the flood, it is difficult to get a good match between staff gauges and recording gauges. Quality of gauge observations is also affected if the gauge is damaged, bent or washed away. The station record book should be inspected for evidence of such a problem.

3.2.2 Autographic Water Level recorder (AWLR)

3.2.2.1 Instrument and Procedure

The vast majority of water level recorders used in India use float and pulley arrangement in a stilling well to record water level as a continuous pen trace on a chart. The chart is changed daily or weekly and the recorder level is set to the current level on the staff gauge, which is also written on the chart at the time of putting on and taking off. This kind of AWLR is known as stilling well with float and encoder gauge. Such recorders have typically been used by the CWC.

3.2.2.2 Typical measurement errors

Automatic water level recorders are subject to errors resulting from the malfunction of the instrument or the stilling well in which it is located. The following are typical malfunctions noted on charts and possible sources of the problems are described below.

a) Chart trace goes up when the river goes down

  - Float and counterweight reversed on float pulley

b) Chart trace does not go down when the river goes down

- Tangling of float and counterweight wires or insufficient length of the wire
- Backlash or friction in the gear arrangements
- Blockage of the intake pipe by silt or float resting on silt

c) Flood hydrograph truncated

- The top of the well may be of insufficient height for flood flows and float sticks on floorboards of gauging hut or recorder box.
- Insufficient damping of waves causing float tape to jump or slip on the pulley.

d) Hydrograph appears correct but the staff gauge reading and chart levels are different. There are few possible sources of such errors including problems of float system, recorder mechanism or the operation of the stilling well.

In addition to the errors noted above, the following errors may be considered.

Operator Problems

- Chart originally set at the wrong level

Float system problems

- Submergence of the float and counterweight (in floods)
- Inadequate float diameter and badly matched float and counterweight
- Kinks in float suspension cables
- Siltation on the pulley hoisting the float affecting the fit of the float tape perforations in the sprockets

Recorder problems

- Improper setting of the chart on the recorder drum
- Distortion and/ or movement of the chart paper (humidity)
- Distortion or misalignment of the chart drum
- Faulty operation of the pen or pen carriage

Stilling well problems

- Blockage of the intake pipe by silt
- Lag of water level in the stilling well behind that in the river due to an insufficient diameter of the intake pipe compared to the diameter of the well

Chart time and clock time disagree

- Chart clock in error and requires adjustments

To eliminate such kind of errors, Ultrasonic sensor and Radar sensors have been introduced these days to measure the water level of reservoirs and rivers.

3.2.3 Automatic/ Digital Water Level Recorders (AWLR/DWLR)

3.2.3.1 Instrument and procedures

Like the chart recorder, many DWLRs are based on a sensor operating in a stilling well. One significant improvement is that the mechanical linkage from the pulley system to the chart is replaced by the shaft encoder which eliminates mechanical linkage errors and the imprecision of a pen trace. The signal from the shaft encoder is logged as the level at a selected time interval on a digital logger and the information is downloaded from the logger at regular intervals and submitted for processing. The level is set and checked against the staff gauge.

Sensors for the measurement of water level do not require to be placed in still water. Loggers based on such sensors have the advantage that they do not need to be placed in a stilling well and thus can avoid the cost and problems commonly associated with stilling wells like silting.

3.2.3.2 Water Level Measurement Sensors

An attempt has been made to discuss the equipment most commonly used in the country:

- Stilling well with float and encoder gauge

  The most common method of measuring water level in a stilling well equipped with a float and shaft encoder. The components of this type of gauge include a stilling well, inlet pipes from the water, float, tape, wheel, and shaft encoder which electronically sends signals to the data collection platform. It has been described in Section 3.2.2 above.

- Gas-purge system (bubblers)

  Another common method of stage measurement is the bubbler system equipped with a non-submersible pressure sensor. This is also known as a gas-purge system. A small quantity of air or inert gas (for example, nitrogen) can bleed through a pipe or tubing to an orifice in the stream. The pressure of the gas that displaces the liquid in the orifice is then measured by a pressure sensor.

- Submersible pressure transducer

  Another method to measure the water level is to use a submersible transducer. In this case, a transducer is installed in a pipe below the minimum water line. The pressure exerted on the sensor by the head of the water above the sensor is converted to depth.

  The discussion on the details of the above instruments and their comparative advantages and shortcomings is beyond the scope of the Manual and the readers are encouraged to refer to the Instrumentations manuals to know further details about them.

- Ultrasonic sensors

  The ultrasonic measurement of water-level is a noncontact method of water-level measurement, which implies that silt load in water or water pollution will not interfere with the function of the sensor. The limitation of this measurement method is that the equipment needs to be installed directly over the body of water, which is not practical in reservoirs or rivers with long slopes. The ultrasonic measurement sensor has a narrow range, limited to 10 m in most applications. It is ideally suited for canal measurements with an ultrasonic sensor mounted on a boom over a canal. The accuracy of measurement is generally sufficient for small bodies of water such as creeks and small canals. Figure 3.2 shows the installation of the ultrasonic sensor on a bridge in Haryana that captures the water level of a reservoir, having a limited range of water surface variation.

- Radar type sensors

  In cases where a larger distance needs to be measured (large rivers) or greater accuracy is desired (large dams), a radar sensor offers a more practical approach. The radar sensor offers an accuracy of approximately 0.02 percent of full scale and has a maximum range of up to 100 m to the target. Figure 3.3 shows an example of radar sensor mounted on the side of a bridge in Gujarat to capture the river water level. Apart from the non-contact nature of the measurement, the major advantage of using radar is the high accuracy along with the extended range of measurement over the ultrasonic type. The radar is also relatively easy to install. The disadvantages include the high cost of radar, which can easily exceed US$3,000/INR 1,95,000 along with the need for some hydraulic structure to mount the radar, such as a bridge.

Figure 3.2: An Ultrasonic Sensor with a Boom Mount at a Canal in Haryana

- Future measurement techniques

  Innovative techniques like smart sensors, imaging video camera, satellite-based measurements of water level and more are being developed to bridge the wide gap between measurements required and that available. It is beyond the scope of the current work to present on these in detail.

Figure 3.3: A Radar Type sensor installed on a reservoir in Gujarat

3.2.3.3 Typical measurement errors

Measurements of the water level recorder with float system and stilling basins are prone to errors while the pressure sensor type does not suffer from such errors. Equivalent checks as mentioned in sections before are therefore necessary to ensure the continuity and accuracy of records. In particular, the comparison and noting of staff gauge and logger water levels (and clock time and logger time) in the Field Record Book at take-off and resetting are essential for the interpretation of the record in the office.

Typical errors for DWLR devices may include:

- Fouling or corrosion with direct exposure to the water, that affects the readings
- Readings affected by changes in the density of the water column
- Sometimes susceptible to flow (the velocity head) and electrical noise effects
- Liable to drift over relatively short time scales (less than one year)
- Damage to sensor head by human touch or other objects like boulders during a flood

Ultrasonic sensors:

Ultrasonic sensors are mounted above the water surface and measure the distance from the sensor to the water surface by emitting ultrasounds and recording the time it takes to record the bounced echo from the water surface. Typical errors may include (USDI, 2017; https://stevenswater.com):

- Power failure at remote sites (common to all electronic sensors)
- Corrosion and plugging of sensor ports
- Spider webs blocking ultrasonic signals from reaching the intended target
- The drift of the sensor when left in place over extended periods
- Overheating causing sensor malfunction
- Erroneous recording because of temperature fluctuations: The speed of sound through air varies with the temperature of the air. The transducer may have a temperature sensor to compensate for changes in operating temperature, but this only takes into account the temperature at the sensor, which may be different from that near the water. Also, there may be poor temperature compensation resulting in a non-linear calibration with temperature fluctuations.
- Presence of debris, extreme turbulence or wave action of the water causing fluctuating readings. A damping adjustment in the instrument or a response delay may help in overcoming this problem.
- Maximum distance from the water level surface is typically 9 m (30 feet) or less, which may not be sufficient to capture the entire range of variation at some places.
- Very high concentrations of fine sediment in suspension (as expected during high flood flows) can scatter and absorb the sonic pulse, preventing reflection of a detectable echo, thus causing a miss of the most pertinent information.

- Condensation on the sensor head can cause problems with the operation of the sensor.

Radar type sensors

Typical errors for the radar type water level sensors may include (Fulford and Davies, 2005):

- Erroneous results in the presence of ice or debris
- Accuracy may be affected by oscillator sensitivity to temperature changes
- Negative bias when surface waves are present. Wave troughs focus energy back at the radar and wave crests scatter energy away from the radar, causing bias in reading.
- Sensitivity due to temperature changes (in a limited manner)
- Horizontal structural surface such as beams, brackets and side wall joints near the sensor affecting the signal.
- Erroneous reflections due to obstruction in the beam signal.

Procedures in the office for checking the reliability of the record will depend on the associated data logger software but should include graphical inspection of the hydrograph for indications of any present malfunction (e.g., flat, stepped or truncated trace). Comparisons of the recorded graph should be made with the observer's readings and any bad or missing data may be replaced by manual observations.

3.3 Primary Validation of Water Level Data Through the Scrutiny of Tabular and Graphical Data from a Single Set of Record

3.3.1 Background

The first step in carrying out validation is the inspection of individual records from a single record or manual measurements for violations of some pre-set physical limits. Alternatively, data may be checked for the occurrence of sequences which show unacceptable hydrological behaviour.

Screening of some unacceptable values might already have been carried out at the data entry stage to eliminate incorrectly entered values. Numerical tests for the physical limits may be considered against three categories:

- Absolute maximum and minimum limits
- Upper and lower warning limits
- Acceptable rates of rising and falling

3.3.2 Absolute maximum and minimum limits

Checking against maximum and minimum limits is carried out automatically and values violating the limits are flagged and listed. The values of absolute maximum and minimum levels at a particular station are set by the data processor such that values outside these pre-set limits are incorrect. Generally, these values are set for the full year and do not vary with month or season. However, the maximum limit for monsoon season may vary from the maximum limit in the dry season.

The cross-section plot of the river gauging line in conjunction with the cross-section of the control section at higher flow depths provide a basis for setting these minimum and maximum limits. For stage records, at many stations, the absolute minimum level can be set at the zero of the gauges (the level at which flow is zero). However, for some natural channels and controls, negative stage values may be acceptable if the channel is subject to scour such that flow continues below the zero gauge. Such conditions must be confirmed by inspection of the Field Record Book.

Similarly, absolute maximum is set at a value after considering the topography of the flood plains around the control section and also the previously observed maximum at the station. If long term data on water levels are already available (say for 15 –20 years) then the maximum value attained in the past can be taken as an appropriate maximum limit.

3.3.3 Upper and lower warning limits

Validation of stage data against absolute maximum limit does not discriminate those unusually high or low values which are less than the maximum limit but which may be incorrect. For this, less extreme upper and lower warning limits are therefore set such that values outside the warning range are flagged for subsequent scrutiny. The upper and lower warning levels must be such that these limits are violated 1–2 times every year by an extreme event. This would ensure that on an average, one or two of the highest peaks or the deepest troughs are scrutinized more closely for their correctness at a site. These limits need to be worked out using suitable statistics but care must be taken for the time interval and the length of data series under consideration. Statistics like 50th percentile value of the collection of peaks over the lowest maximum annual value used to set the upper warning level for the hourly data series of say 15-20 years may be used for this purpose. Of course, such statistics will also be subject to the nature or shape of the hydrograph that the station under consideration experiences. The appropriateness of such limits has to be verified before adopting them.

3.3.4 Limits of rising and falling rates

The method of comparing each data value with immediately preceding and following observation values are of particular relevance to water level records as they exhibit significant serial correlation which can be very useful for checking water level observations. A limit can be set numerically as the maximum acceptable positive or negative change between successive observations. It should be also noted that an acceptable change in level during a rising flood hydrograph in the monsoon season may be unacceptable during the dry season. Violations of rising and fall limits are more readily identified from graphical plots of the hydrograph.

3.3.5 Listing of data

An essential requirement for an organized data processing activity is the listing of the entered data. It will help in validation as follows:

- In checking against various data limits
- In recording remarks/ comments of the data processing personnel while validating the data

Example 3-1

Daily water level data for 2013 from a gauging station on the Ravi River in Chamba district of Himachal Pradesh (slightly modified for demonstration purposes in this example) with a catchment area of 9,025 square kilometres has been considered for the current analysis. It is required to set the following data limits:

- Maximum water Level
- Minimum Water Level
- Upper Warning Level
- Lower Warning Level
- The Maximum and Minimum Rate of Change of Water level

The cross-section at the gauging section is given in Figure 3.4, from which it is apparent that the lowest bed level is about 1957.1 m and the top of the bank is about 1964 m. If the zero of the gauge is set at 1957.1 m, then the minimum data limit/ minimum water level can be easily set at 1957.1 m as it is not expected for the top of flowing water to reach this level, even after moderate scouring of the bed.

The gauging sites are usually located at a bridge pier for the sake of convenience. It is therefore not appropriate to consider the cross-section details of the same for setting up the maximum limit since the control section may be located somewhere downstream of the bridge and may have different levels at the flood plain. Such flood plain levels and the topography around the control section would be governing the levels at the gauging cross-section. As a first estimate, the top of the cross-section at the control section can be taken as the maximum limit - which is 1964 m. An alternative way to fix the maximum water level is to use the highest water level ever recorded at the site which is about 1962.5 m. Keeping in view the bank position and inspection road at level 1963, the maximum water level is kept at 1963 for this location.

Considering that the objective of setting the upper warning level is to flag 1-2 flood events each year for closer scrutiny, a limit of 1961.0 m as the upper warning level will be effective. Such limits can also be arrived at using suitable statistics on the data.

To establish the limits on the maximum rate of rising and rate of falling, it is best to use the historical data and obtain the derivative of the hourly water level series. The maximum limits of the rate of rising and falling can be obtained after calculating the rate of change of water levels from all the hourly data available. In this case, limits of 0.5 and -0.2 m/hr are set for the maximum rate of rising and rate of falling respectively.

The above example is shown in Figure 3.5 and Figure 3.6.

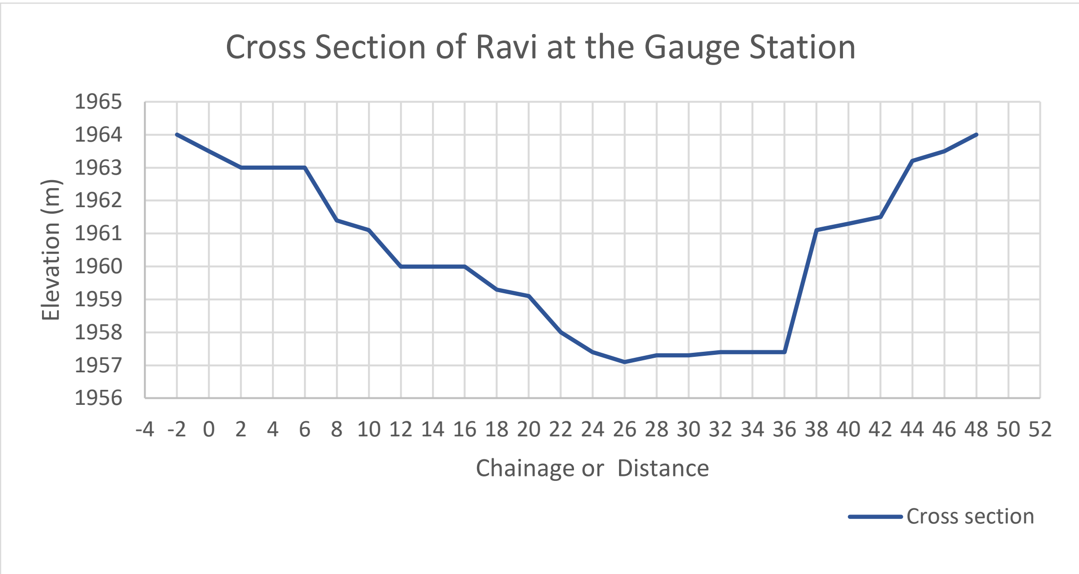


Figure 3.4: Cross Section of the River Ravi at Gauge Station

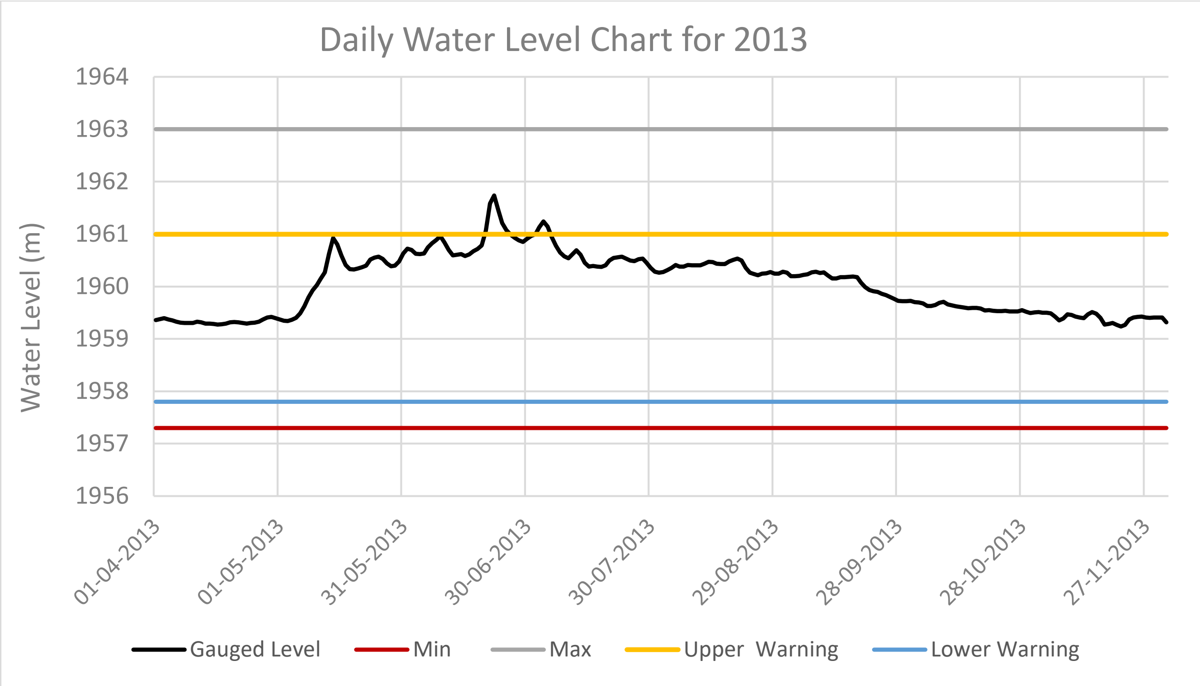


Figure 3.5: Hourly Water Level Data with Data Limits for the Year 2013

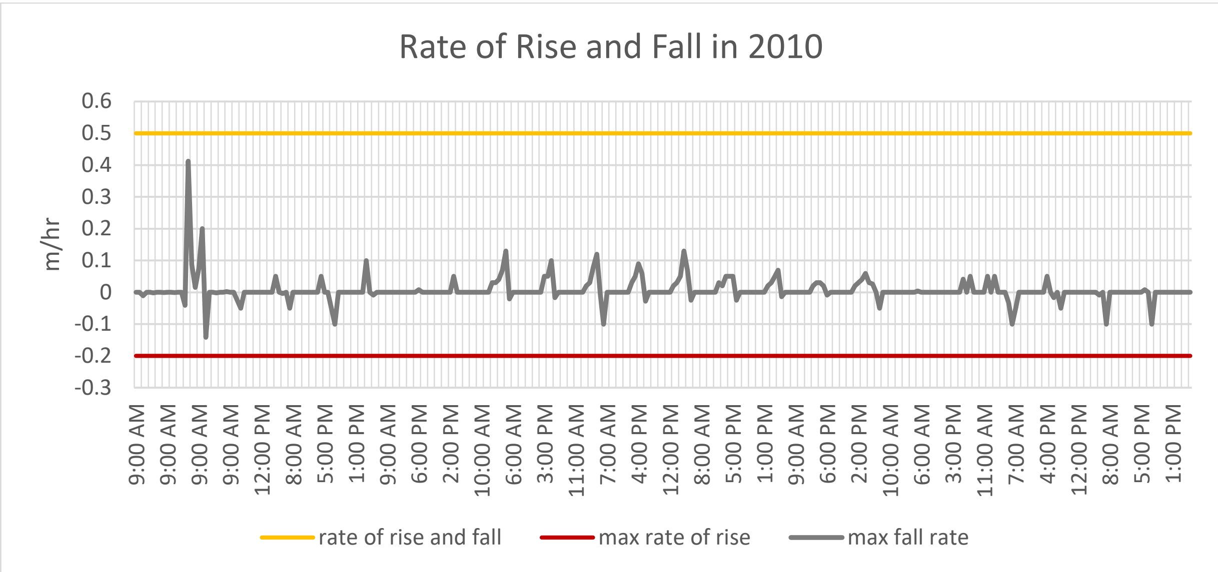


Figure 3.6: Rate of Change in Hourly Water Levels with Data Limits

3.3.6 Graphical inspection of hydrographs

Visual checking of time series data is often a more efficient technique for detecting data anomalies than numerical checking, and must be applied to every data set with an inspection of the stage hydrograph. Screen/ visual display will also show the maximum and minimum limits and the upper and lower warning levels. Potential problems identified using numerical tests will be inspected. They will then be flagged as spurious or doubtful and corrected wherever possible. An attempt must be made to interpret identified anomalies in terms of the performance of the observer, instruments or station, and recorded where it has been possible to communicate this information to the field staff for field inspection and correction. A few examples/ cases of inspection of hydrographs are discussed below:

Case 1

Case 1 represents a false recording of a recession curve (Figure 3.7) caused possibly by:

- An obstruction causing the float to remain hanging
- Blockage of the intake pipe
- Siltation of the stilling well

This also shows the time when the obstruction was cleared. It may be possible to interpolate a true recession curve as shown below

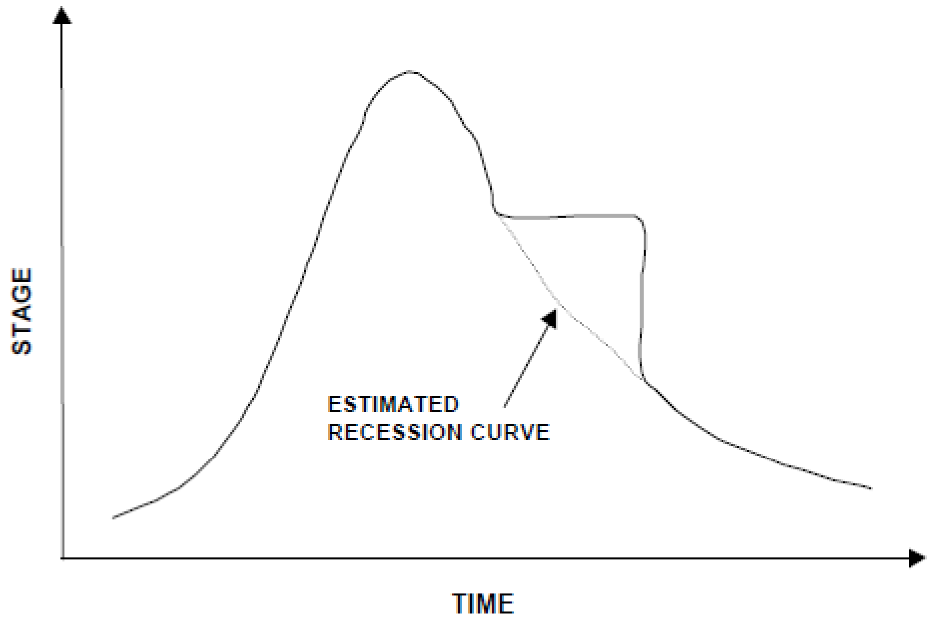


Figure 3.7: False Recording of Recession Curve

Case 2

Case 2 involves steps in the stage recording because of the temporary hanging of the float tape or counterweight as the result of a malfunction in mechanical linkages in the recorder (Figure 3.8). Such deviations can be easily identified graphically and true values can be interpolated.

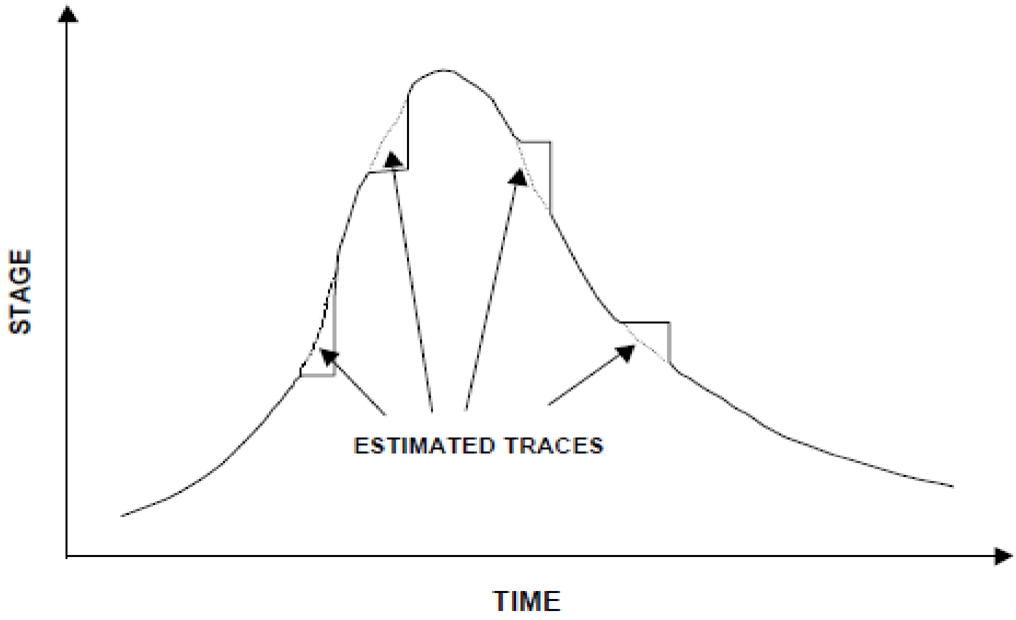


Figure 3.8: Temporary Hanging of the Float

Case 3

Spurious peaks and troughs (spikes) in the hydrograph may be generated by observer error or occasionally by electronic malfunction of transducers or shaft encoders. However, it should be noted that in some instances real variations of similar nature may also be generated:

- In rivers with small flow, switching on and switching off of pumps immediately upstream of the observation station will generate rapid changes
- The building of a bund or obstruction upstream of a station and its subsequent failure or release will generate first a negative spike followed by a positive one. In this case, the levels as observed would be correct.

3.3.7 Validation of water level in the regulated rivers

Rivers which are not affected by river regulation and abstraction have a flow pattern which can be described using suitable statistical functions. The hydrograph at stations in such basins follows a natural pattern on which errors and inconsistencies can be identified more easily. However, recording natural flow series in Indian rivers is not common since the river flows are affected by man-made structures to a greater or lesser extent. Generally, on regulated rivers, the natural pattern is disrupted by reservoir releases which may have abrupt onset and termination, combined with multiple abstractions and return flows. These influences are most clearly seen in low to medium flows where in some rivers, the hydrograph appears entirely artificial; high flows may still observe a pattern resembling natural conditions but low flows are significantly altered. In such cases, the validation becomes more difficult and the application of objective rules may result in the listing of many queries where the observations are correct. It is therefore recommended that the emphasis of validation for regulated rivers should be on graphical screening by which data entry and observation errors may still be readily recognized.

The officer performing validation should be aware of the principal artificial influences within the basin, the location of those influences, their magnitude, their frequency and seasonal timing, to provide a better basis for identifying values or sequences of values which are suspect.

3.3.8 Comparison of daily time series for manual and autographic or digital data

At stations where the water level is measured using an autographic or a digital recorder, a staff gauge reading is usually also available as a backup. Thus, at such observation stations, water level data at daily time interval are available from at least two independent sources. Discrepancies between readings may arise either from the staff gauge readings, the recorder readings or from both. The typical errors in field measurement have been described above and these should be considered in interpreting the discrepancies. Errors arising from the tabulation of levels at hourly intervals from chart records or during data entry are also possible.

3.3.9 Data validation procedure

Two or more series representing the same level at a site are plotted on a single graph, where the two lines should correspond. A residual series may also be plotted showing the difference between the two methods of measurement. The following, in particular, should be noted:

- If there is a systematic difference between staff gauge and recorder, the recorder has probably been set up at the wrong level. Check chart annotations and the field record book. Check for steps in the hydrograph at the time of chart changing. The data record should be adjusted by a constant difference from the staff gauge.
- If the comparison is generally good but there are occasional discrepancies, it is probably the result of an error in the staff gauge observations by the observer or incorrect extraction from the chart.
- In case of missing records during the flood, a failure associated with the stilling well or the recorder should be suspected.
- A gradual increase in the error may result simply from the recorder clock running fast or slow. This can be easily observed from the graphical plot and the recorder record should be adjusted.

Setting minimum and maximum limits ensures filtering of values outside the specified limits. Such values are considered suspect. They are first checked against manual entries and corrected if necessary. If both readings agree and fall outside prescribed limits, the value is flagged as doubtful. Where there are some other corroborative facts about such incidents, available in the manuscript or notes of the observer or supervisor, they must then be incorporated with the primary data validation report. This value has to be probed further at the time of secondary data validation when more data from adjoining stations are available.

When the data being entered exceed the prescribed limits while high rainfall events have been experienced by the staff and recorded by other nearby stations, the maximum limit is reset to a suitable higher value. If there is no justifiable basis for setting the new maximum value, the new value is reported in the form of a remark which can be reviewed at the secondary validation stage.

3.3.10 Multiple graphs of water levels at adjacent stations

Comparison of records between stations will normally be carried out as part of secondary validation at Divisional level and will not only be done for discharge but water level/ stages as well. An initial inspection may be done at Sub-divisional level where records for a few neighbouring stations are available. Such stations, if on the same river (sometimes on different rivers nearby with a similar type of catchment) will show a similarity in their stage plot and inspection of these plots may help in the screening of the outliers.

3.4 Secondary Validation of Water Level Data

3.4.1 Background

Water level data received at Divisional offices have already received primary validation based on knowledge of instrumentation and methods of measurement at the field station and information contained in Field Record Books. Primary validation includes comparisons between different instruments and methods of observation at the same site.

The data processor must be aware of field practice and instrumentation and the associated errors which can arise in the data.

Secondary validation at Division Office now puts most emphasis on comparisons with neighbouring stations to identify the suspect values. Special attention will be given to records already identified as a suspect in the primary validation.

The assumption, while carrying out secondary validation, is that the variable under consideration has an adequate spatial correlation. Since the actual value of the water level is controlled by specific physical conditions at the station, the amount of secondary validation of level is limited. Most of the comparisons with neighbouring stations must await transformation from level to discharge through the use of stage-discharge relationships. Only as discharge can volumetric comparisons be made. However, validation of level will identify serious errors in recorded time.

Secondary validation of level is used to identify suspect values or sequences of values but not usually to correct the record, except where this involves a simple shift (time or reference level) of a portion of a record.

The main comparisons are between water level series at successive points on the same river channel. Where only two stations are involved, the existence of an anomaly does not necessarily identify which station is at fault. Reference will be made to the historic reliability of the stations.

Comparisons are also to be made between incident rainfall and level hydrographs.

3.4.2 Scrutiny of multiple hydrograph plots

Graphical inspection of comparative plots of time series provides a very rapid and effective technique for detecting timing anomalies and shifts in reference level. Such graphical inspection is the most widely applied validation procedure.

For a given period, several time-series of water levels for neighbouring stations are plotted in one graph. For routine monthly validation, the plot should include the time series of at least the previous month to ensure that there are no discontinuities between one batch of data received from the station and the next. The time interval of observation rather than averaged values should be displayed. In general, peaks and troughs are expected to be replicated at several stations with the earlier occurrence at upstream stations and the lag between peaks, based on the travel time of the flood wave. The travel time will vary the magnitude of the events, the flood with a larger magnitude travelling fester. It should be noted that level fluctuations noted at the downstream stations should not necessarily be higher than that observed at the upstream stations - the actual value would depend on physical conditions at the stations.

An error may be suspected where the peaks occur at one station but not at its neighbouring stations or where the lag time between stations is widely different than expected. However, it must be recognised that the quality of the relationship between hydrographs from neighbouring stations depends not only on the accuracy of the records but also on a variety of other factors including:

- rainfall and inflow in the intervening reach between stations. If the intervening catchment is large or the rainfall is higher in comparison to that over the upstream basin, a very poor relationship may result.

- river regulation and abstractions between the stations may obscure natural variations, though high flows are usually less affected than low or medium flows.
- time of travel versus flow relationship should be developed for river reaches that are between the hydrometric stations and those should be used to assess the lag, which reduces with an increase in flow for a given river reach. Since this relationship is not routinely developed and maintained in India, and since the local runoff between the stations and the effects of regulation can alter the peak flows at the downstream station, comparing the peak flow lags should be used with extreme caution.
- one station may suffer from backwater effects on the stage hydrograph but not the another, obscuring the effects of the differences between their flow records. Where such effects are known to occur, a comparison should await the transformation of water level to discharge. In general, locations for hydrometric stations should be selected in places where backwater effects are not likely to occur.

Anomalies identified from comparative hydrograph plots are flagged for further stage validation or await validation as discharge.

Example 3-2

Application of the above-described technique is demonstrated for the stations Mahemdabad and NSB00I7 on Watrak river, a tributary of the Sabarmati River in Gujarat. The stations are 33 km apart (Mahemdabad d/s of NSB00I7) and the lateral inflow in between the sites is small compared to the river flow. The hydrographs of hourly water levels for September and October 1998 are shown in Figure 3.9. When carrying out such analysis, it should be ensured that a tabulated output of the water level observations is available to note down possible anomalies.

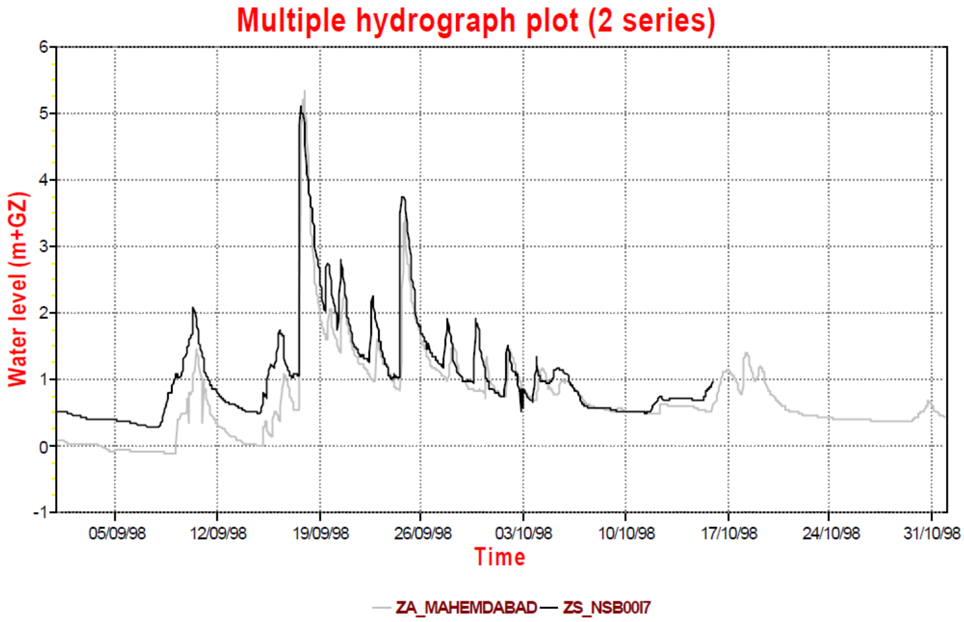


Figure 3.9: Multiple Hydrograph Plot

To get a better view one should zoom in. A detail of the hydrograph, shown in Figure 3.10 poses several questions:

- possible malfunction of the equipment at the upstream station, showing two sharp drops between 2nd and 3rd of October, 1998;
- a large variation of lag times for what appear to be similar upstream hydrologic events; and,
- a general anomaly where the upstream hydrograph seems to depict more flow than the downstream (although in strictest terms this comparison should be applied on the flow hydrographs, not the level hydrographs as is the case, for reasons that are explained further below). Assuming no regulation effects between the two stations, this should not be happening.

The best way to compare hydrographs is not be comparing water levels, as is done in this example, but rather by comparing flows, since each hydrometric station has a rating curve with a unique shape. Once water levels are converted to flows, the area under flow hydrograph represents the total flow volume that passed through a given station, and if the total volume at an upstream station is visibly higher than that at a downstream station for the same period, this would typically be a cause for concern, either due to data errors, unaccounted effects of regulation, or excessive unreported water abstractions from the river.

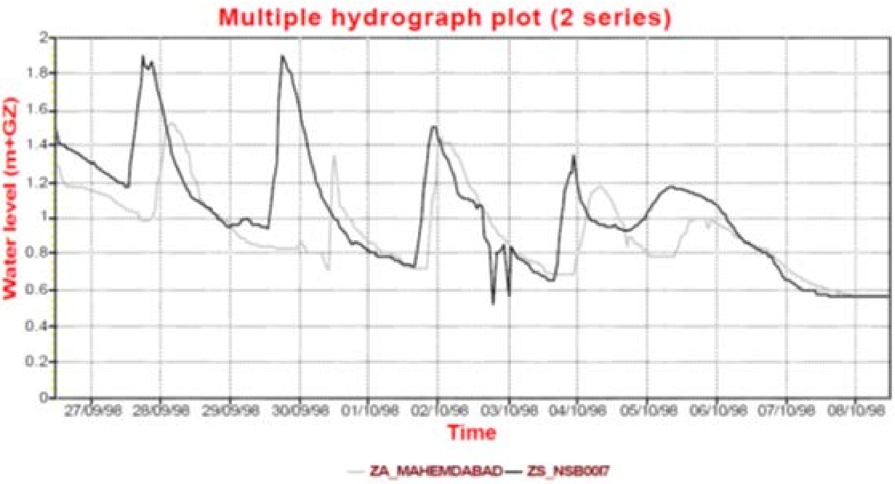


Figure 3.10: Details of Multiple Hydrograph Plot

3.4.3 Combined hydrograph and rainfall plots

At the outset, it should be emphasized that rainfall data should be the last resort for analysing the flow data, since there are many factors that may affect the rainfall-runoff relationship as explained below. Comparative plots may be useful in some instances, but other sources of information such as the flow records at nearby stations should be given more importance in the overall assessment of surface water data.

The addition of rainfall to the comparative plots provides a means of assessing the timing errors and of investigating the effects of inflow into the intervening catchment between stations. The comparison should be made using an average rainfall determined using Thiessen Polygons or other methods over the entire basin or for the intervening sub-basin for the various gauging stations. Extreme caution should be used for small basins with individual rainfall records due to the random distribution pattern of rainfall depth over the entire basin.

In general, a rise in river level should be preceded by a rainfall event in the basin (assuming the additional flow did not originate from reservoir releases). Conversely, it is expected that rainfall over the basin will be followed by a rise in the surface water levels. There must be a time lag between the occurrence of rainfall and the rise in water levels. Where these conditions are violated, an error in rainfall or the level hydrograph may be suspected. However, the above conditions do not apply universally and the assumption of an error is not always justified especially for isolated storms in arid areas. For example, an isolated storm recorded at a single rain gauge may be unrepresentative and much higher than the average basin rainfall. The resulting runoff may be negligible or even absent if the storm happens in dry season and the soil is dry. Where storm rainfall is spatially variable, it may be heavy and widespread but miss the rain gauges, thus resulting in a rise in river level without preceding measured rainfall. The amount of runoff resulting from a given rainfall varies significantly with the antecedent catchment conditions. Rainfall at the onset of the monsoon on a very dry catchment may be largely absorbed in soil storage without reaching the river channel. The situation may be reversed during the end of the monsoon period when the storages and depressions would mostly be full with little capacity left for absorption.

All of the above issues make the process of calibration and validation difficult, often resulting in large discrepancies between the observed and simulated records which serve as a warning that this approach is often not very useful in the assessment of water levels and flow data. An example of the comparison between observed and simulated flows (which should coincide for a perfect mode) is given in Figure 3.11 below.

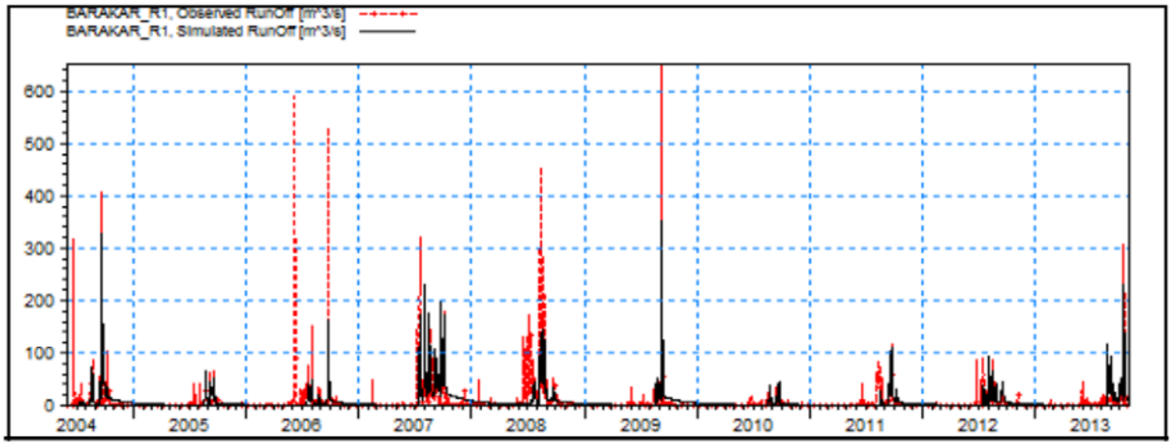


Figure 3.11: Example of Typical Validation Results of a Rainfall-Runoff Model

It should be obvious that the simulated flows shown using the solid black line should not be used to replace the observed flows that are shown with the red line for any purpose, be it for using the peak flows for frequency analyses studies or for using the time series of flows as an alternative to the historic flow series.

The use of comparative plots of rainfall and level is therefore qualitative but it may sometimes provide valuable ancillary information.

Example 3-3

An example of a combined hydrograph and rainfall plot is presented in Figure 3.12, which displays the water level record of station Ambaliyara on Mazam river together with the rainfall records of stations Rahol and Vadagam.

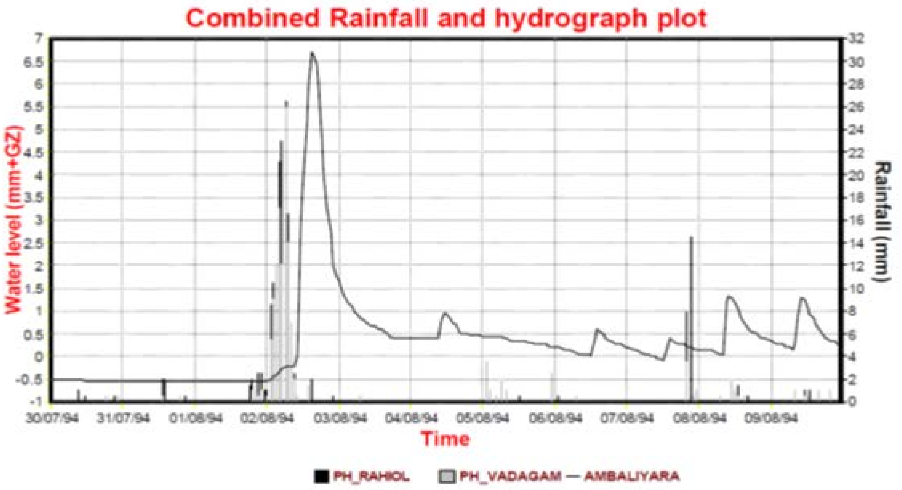


Figure 3.12: Combined Hydrograph and Rainfall Plot

From the graph, it is observed that the first peak is preceded by substantial rainfall. The remainder, however, shows suspect combinations of hydrographs of stage and rainfall, where hydrograph appears to occur before the rainfall and where the response to the rainfall is delayed. One may be tempted to doubt the rest of the record. However, the peaks may have been caused by runoff events on tributaries where there are no rain gauges. The peaks may also be attributed to reservoir spills if the reservoir was full, however, such events are usually marked by controlled releases at a flat rate set by the reservoir operator, which is not the likely cause- given the shape of the hydrograph. It should also be noted that the distance from the centroid of the rainfall and water level hydrograph is approximately the same (one day) for both rainfall events shown in the graph, which represents the time of concentration for the upper catchment located above the water level gauge. The graph illustrates the difficulties of using the rainfall information for analysing recorded water levels, and outlines the importance of using this approach with caution.

3.5 Relation Curves for Water Level

3.5.1 Background

A relation curve gives a functional relationship between two series of the form:

$$Y_t = F\left(X_{t+t_1}\right)$$

To account for the lag between level changes at one station and the next downstream, it may be necessary to introduce a time shift (t_1) between the two time series.

Relation curves are normally applied to water level data rather than discharge. This results in prediction of the water levels at a downstream station using real time water level measurements at an upstream hydrometric station. It is a convenient and quick way, that uses the established relational curve between the two stations. However, it is appropriate to also develop the relational curves using discharge data at both stations, since this allows users to conduct additional checks. For this purpose, it is recommended to select hydrometric stations where flows are very unlikely to be affected by backwater conditions.

If there is a distinct one to one relationship between the two series, random errors will be shown in a relation curve plot as outliers.

By comparing two relation curves, or data of one period with that of another period, shifts in the relationship can be detected. A shift is said to occur when water levels in one series changes due to changes in the settings of gauge zero.

3.5.2 Application of relation curves to water level

If two water level stations are located on the same river and no major tributary joins the mainstream between the two locations, a relation can be expected between the recordings at the two locations. With the help of this relation, the stage at a particular downstream station can be derived from the available data series of the upstream station.

Two important conditions need to be satisfied to obtain a high degree of relationship between the stage data of adjacent stations. These are:

- No major tributary joins the main stream in between the two adjacent stations.
- Time of travel of the flood wave between the two stations is taken into consideration.

The importance of travel time can be ascertained by inspecting the distance between the two stations. If the distance is close (i.e., 70 km or less), there is usually no need to account for travel time in the development of the statistical relationship of water levels between the stations. Two sample plots examine a relationship between the two stations that are close to each other without the lag and with one day lag, as shown in Figure 3.13. Rajegaon is the upstream gauge station on Bagh River, a tributary of Wainganga, and Kumhari the downstream station on Wainganga River.

It is obvious from the above graph that adding a 1-day lag does not improve the regression. Rather, it reduces the coefficient of determination R^2 to 0.67 from 0.76 which was obtained assuming no lag.

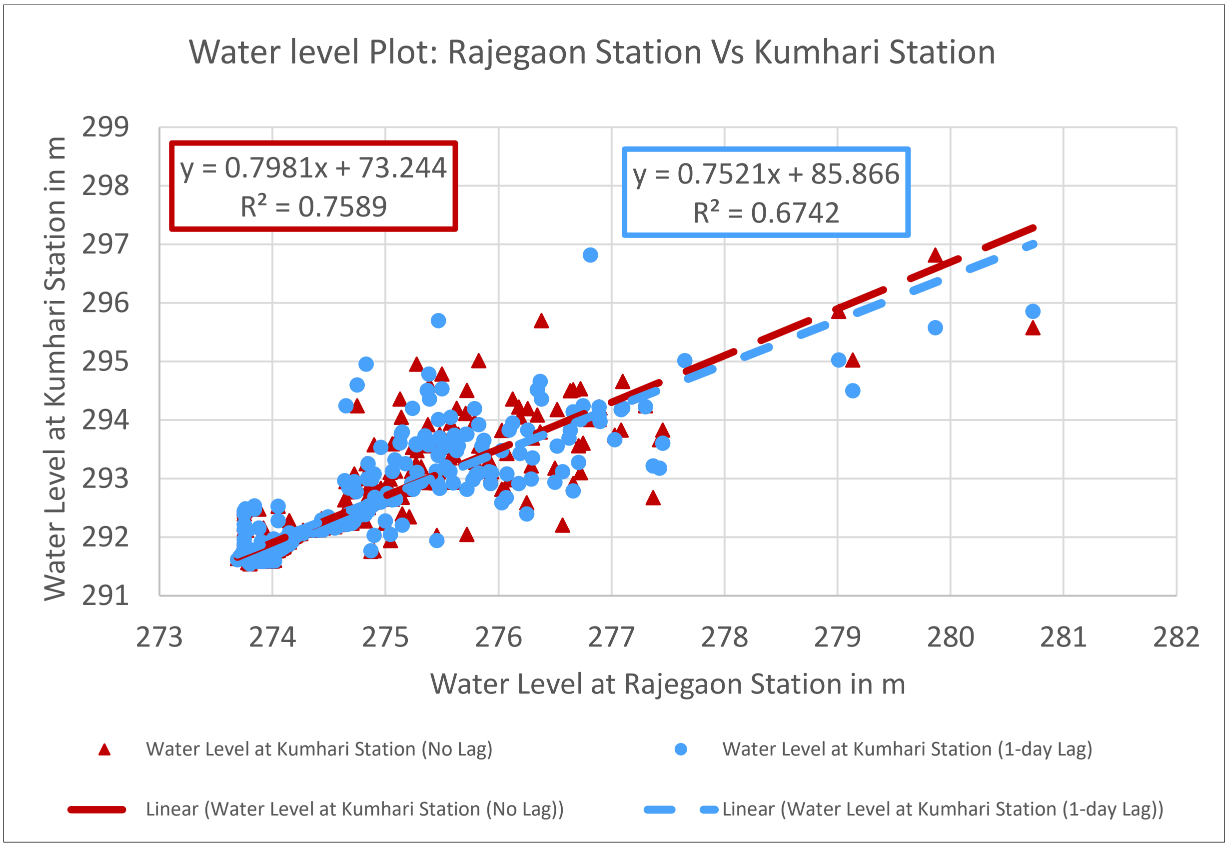


Figure 3.13: Example of Relation Curve (Hourly Water Level)

As mentioned for comparative hydrograph plots above, the occurrence of lateral inflow between stations limits the quality of the relationship between neighbouring stations. The lateral inflow may occur as the main tributary inflow or as distributed inflow over the reach. In either case, if it is a significant proportion of the downstream flow or variable, the quality of the correlation fit may be reduced.

3.5.3 Determination of travel time

For the second condition, the relationship between the two station time series must incorporate a time shift, representing the mean travel time of a flood wave between the stations. The time shift may be assessed using:

i. physical reasoning, or
ii. from an analysis of the time series

A cursory way to determine travel times is to observe the time of the arrival of the peak flow of an hourly hydrograph at both the upstream and the downstream stations. However, most of the times it is difficult to obtain hourly data. Stations that are relatively close to each other (distance of less than 60 – 70 km) may not show any impact of travel time on a daily basis, since peak flow would pass through both stations within 24 hours.

3.5.3.1 From physical reasoning

The time of travel of a flood wave can be approximately determined by the division of the interstation distance by the estimated mean water velocity. One basis to determine travel time during low and medium flows is to develop travel time versus flow table by using the mean velocity obtained from discharge measurements and the length of the river reach. It is done by dividing the length of the river reach with the mean velocity associated with the measured flow rate. An alternative to this is to get the time of travel for a given flow rate and for a defined river reach is to run the flow rate through the calibrated HEC-RAS model which provides travel time as one of the output options. In either case, the resulting table of flows vs travel time should have decreasing travel times along the reach with the increase of channel flow, as shown in an example below:

Flow (m3/s)	Travel Time (hrs)
0	262.86
0.5	142.32
1	109.28
3	71.89
6	55.19
10	45.43
30	29.88
60	22.95
100	18.89
300	15.63
600	13.42
1000	12.42
3000	9.54
6000	7.85
10000	7.27
20000	6.54
30000	5.9

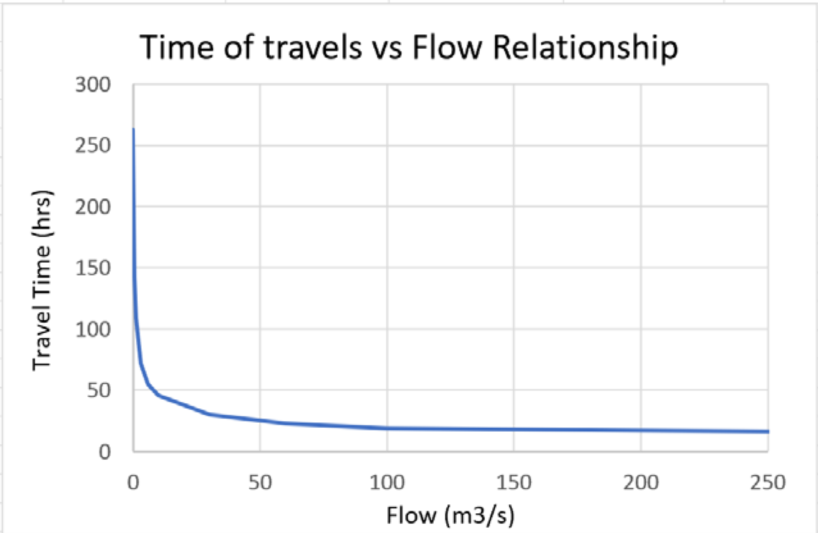


Figure 3.14: Typical Travel Time vs Flow Relationship

Travel times for high flow rates are typically determined by observing hydrograph peaks at the two adjacent stations on the same river. The reason for this is that flow measurements are usually not conducted during very high flows due to safety concerns. A fitted relationship between travel time and flows has the following form:

$$TT = \frac{A}{Q^n}$$

Where parameter A is associated with the length of the river reach while the exponent n implicitly contains the information related to the mean gradient of the river reach. TT is usually calculated in hours while Q is given in m^3/s. The above relationship is empirical and it is usually based on the observed flow and mean cross sectional velocity data points. The advantage of the above relationship is that it allows easy assessment of travel times as a function of the average flow along the river reach.

The most expedient way to determine coefficients A and n is to click on the data of the point graph of travel time vs flow plot in Excel, and select an option to add the trendline. At the bottom of the trendline option parameters, there is a check box that needs to be selected to show equation on the plot. Once this option is selected, the above plot for data range from 0.5 to 1000 m^3/s would result in the values of coefficients A and n of 95.077 and 0.299, respectively. The equation would be shown in the form $y=96.077x^{-0.299}$.

An additional option to determine travel time for a given flow rate along a river reach is to use dye studies, where a dye is injected at the upstream end at a recorded time. Thereafter, its arrival is then recorded at the downstream end of the river reach. Such measurements are sometimes conducted for specific studies where higher accuracy of travel time estimates is desirable.

3.5.4 Using the relation curve for data validation

In general, relation curves are statistical curves which are developed with varying levels of success. There is usually significant data dispersion between the water level data at the upstream and downstream station, which calls for more refined development of statistical functions for various parts of the years (i.e., separating data from the monsoon season and dry season). Consequently, statistical curves such as those shown in Figure 3.13 should not be used for validation of the data at the downstream station. It is always better to plot the actual time series of both hydrographs and analyse them together, as depicted in Figure 3.15.

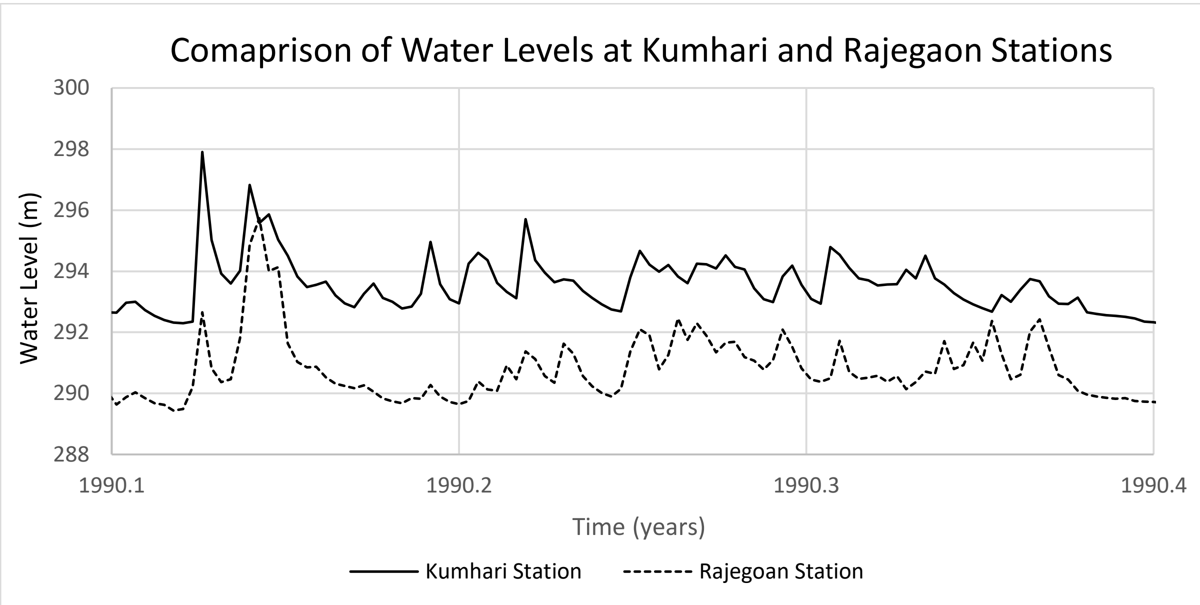


Figure 3.15: Hydrograph Plot of the Next Flood Wave

4 CORRECTION AND COMPLETION OF WATER LEVEL DATA

4.1 General

After validation of water level data, several values are likely to be flagged as incorrect or doubtful. Some records may be missing due to non-observation or loss of data in transmission. Wherever possible, the incorrect and missing values will be replaced by estimated values. The process of filling in missing values is generally referred to as 'completion'. As part of the Hydrological Information System, correction and completion of water levels are proposed to be carried out at the Divisional offices, as a continuous process with validation.

The records identified as suspect by numerical validation tests are inspected and corrected if necessary. Alternatively, if the values are found to be acceptable, the flag is removed. Numerical test of records for maximum, minimum, warning limits and rates of rising will have identified suspect values, which are flagged during primary validation. Unless these are due to a data entry error, they will not be corrected at this stage and will undergo further inspection, correction and completion.

Where multiple level records at the same station are flagged, but the observations agree, the records may be assumed to be correct. Other suspect values outside warning limits are inspected for violations of typical hydrological behaviour but are also checked against neighbouring stations before correction or acceptance.

During validation, it is important to recognize that values estimated from other gauges are inherently less reliable than the values properly measured. Doubtful original values will, therefore, be given the benefit of the doubt and will be retained in the record with a flag. Where no suitable neighbouring observations or stations are available, missing values will be left as 'missing' and incorrect values will be set to 'missing'. Often secondary information like published news for the date may help in explaining some extraordinary values, marked as a suspect through numerical analysis.

4.2 Correction Using River Level or Discharge

Correction and completion may be carried out for the water level series or it may await transformation to discharge using a stage-discharge relationship. The choice of water level or discharge for correction depends on the type of error, the duration of missing or faulty records and the availability of suitable records at other nearby stations. Correction of water level data has the advantage that it is the primary measurement whereas errors in discharge may result either from errors in the recorded levels or the errors in the stage-discharge relationship. However, it has the disadvantage that it provides no volumetric water balance checks.

4.2.1 Correction and completion usually carried out for water level data:

- where the level record is complete but the recorder has gone out of adjustment and periodic check observations are available
- where the level record is correct but shifted in time
- where the primary record (e.g., data from a digital water level recorder) is missing but an alternative level record (say manual gauge observations) of acceptable quality is available at the same station
- where the record is missing but the duration is short during a period of low flow or recession.

4.2.2 Correction and completion of water level data may be carried out:

- where record from a neighbouring station is available with little lateral inflow or abstraction between the stations.

4.2.3 Correction and completion carried out for discharge:

- where record is available only from a neighbouring station without much lateral inflow or abstraction
- where the only available means of infilling is from catchment rainfall and the use of a rainfall-runoff model.

Records completed as the stage will receive further validation as discharge and may require further correction.

4.3 Comparison of Staff Gauge and Autographic or Digital Records

Where two or more measurements of the same variable are made at a station, one record may be used to correct or replace the other where one is missing. Where more than one record exists but they differ, the problem in the first instance is to determine which record is at fault. Typical measurement errors from each source have been described in brief under 'primary validation' in section 3.2 and guidelines provided for identifying which record is at fault. Suspect values are flagged during the validation.

Errors related to mechanical water level recorders installed in a stilling well, and their correction may be classified as follows:

- observer errors
- recorder timing errors
- pen level errors
- errors arising from stilling well and intake problems

- miscellaneous instrument failures

4.3.1 Observer errors

Staff gauge and autographic or digital records can be displayed together graphically as multiple time series plots. Differences can also be displayed. Simple and isolated errors in reading and transcription by the observer (e.g., 6.57 for 5.67) can be identified and replaced by the concurrent measurement at the recording gauge. Persistent and erratic differences from the recording gauge (negative and positive) indicate a problem with the observer's ability or a possible tempering with the records. If detected, the appropriate Sub-division should be notified for corrective action; the full staff gauge record for the period should be flagged as doubtful, left uncorrected until the recording gauge record is modified accordingly.

4.3.2 Recorder timing errors

When the clock of the recording gauge runs fast or slow, the rate at which the recorder chart moves with time under the pen will also be fast or slow. This can be detected by comparing with staff gauge readings. For example, if observations are taken daily at 0800 hours and the clock of the recording instrument is running slower, then the observer's stage record at 0800 will correspond to the same observation in the recording gauge before 0800, say 0700 hours. Clock times and recorder times annotated on the chart or recorded in the Field Record Book at the time of putting on or taking off the chart can be used to determine the time slippage during the recording period.

4.3.2.1 Correction procedure

For time corrections, it is assumed that a clock runs fast or slow at a constant rate. Where a digital record is produced from an analogue record using a pen-follower digitizer, the annotated clock and recorder time and level can be fed into the digitizing program. Then the level record can be expanded or contracted as required to match the clock duration.

Where a digital record is extracted manually at a fixed interval from a chart, it will result in extra records for a fast clock and deficient records for a slow clock. This can be expediently corrected by removing or inserting (interpolating) records at appropriate intervals, e.g., if the clock runs 4 hours fast in eight days, and hourly data have been extracted, then one data point should be removed at 2-day intervals.

4.3.3 Pen level errors

The pen of the autographic recorder may gradually drift from its true position. In this case, analogue observations may show deviation from the staff gauge observations. This deviation can be static or may increase gradually with time.

4.3.3.1 Correction procedure

Where a digital record is produced from an analogue record using a pen-follower digitiser, the annotated clock and recorder time and level can be fed into the digitizing program. Thereafter, an accumulative adjustment spread over the level record from the time the error is thought to have commenced till the error was detected or the chart removed can be introduced. However, such a procedure is not recommended as the actual reasons for the shift may still be unknown at the time of digitizing the charts. It is always appropriate to tabulate/ digitize the chart record as it is in the first instance, and then apply corrections thereafter.

The data validation procedure involves correcting the gradual spread of error in digital records extracted from a chart recorder, with a growing adjustment from the commencement of the error until error detection. Let the error be X observed at time $t=i+k$ and assumed to have commenced at k intervals before, then the applied correction reads:

for j = i, i+1, i+k

$$Xcorr,j = Xmean,j - ((j - i)/k)\Delta X \quad \text{Equation 4.1}$$

Prepare the time-series plot of deviation of staff gauge observations from the recording gauge observations. If the deviation is static with time, then the difference must be settled (increased or decreased) directly from the analogue gauge observations. However, if the deviation increases gradually with time, then corrections for the difference between the pen observation and the staff gauge reading are made in the same way as time corrections. For example, assume that the pen trace record gradually drifted 0.08 m away (recording lower levels) from the corresponding staff gauge record in 10 days. This shows that the pen readings have an error which is increasing gradually from 0 to 8 centimetres in 10 days period. Errors in such data can be compensated by adding a proportionate amount of 8 mm per day from the starting point of the error.

4.3.4 Errors arising from stilling well and intake problems

Problems with stilling well or intake pipe may be intermittent or persistent and can be serious. In extreme floods, the hydrograph may be truncated due to inadequate height of the well restricting the travel of the float, or counterweight reaching the well bottom. Blockage of the intake pipe with silt will result in a lag between river level (as recorded by the staff gauge) and well level, or a flat trace.

4.3.4.1 Correction procedure

The recorder trace is replaced by the observer's staff gauge record if the time interval is sufficiently small for the changes in the water levels. If the staff gauge record is intermittent or frequent changes in the levels are expected to be present, then use of relation curves is to be preferred for correcting the water level record.

4.3.5 Miscellaneous instrument failures

Unacceptable recorder traces may result from a wide variety of instrument problems. These are often displayed as stepped or flat traces and may be corrected by interpolating a smooth curve on the hydrograph plot.

Figure 4.1 shows a false recording of the recession curve because of: a) silting of stilling well; or b) blocking of intakes or c) some obstruction causing the float to remain hung. The figure also shows the time when the obstruction is cleared. The correct curve can be estimated by reading the reconstructed smooth curve that joins the first and last reading during the period of obstruction.

Figure 4.2 shows small steps in the stage records because of the temporary hanging of the float tape or counterweight, or kinks in the float tape. Such deviations can be easily identified and true values can be interpreted by readings from the smooth curve that was reconstructed.

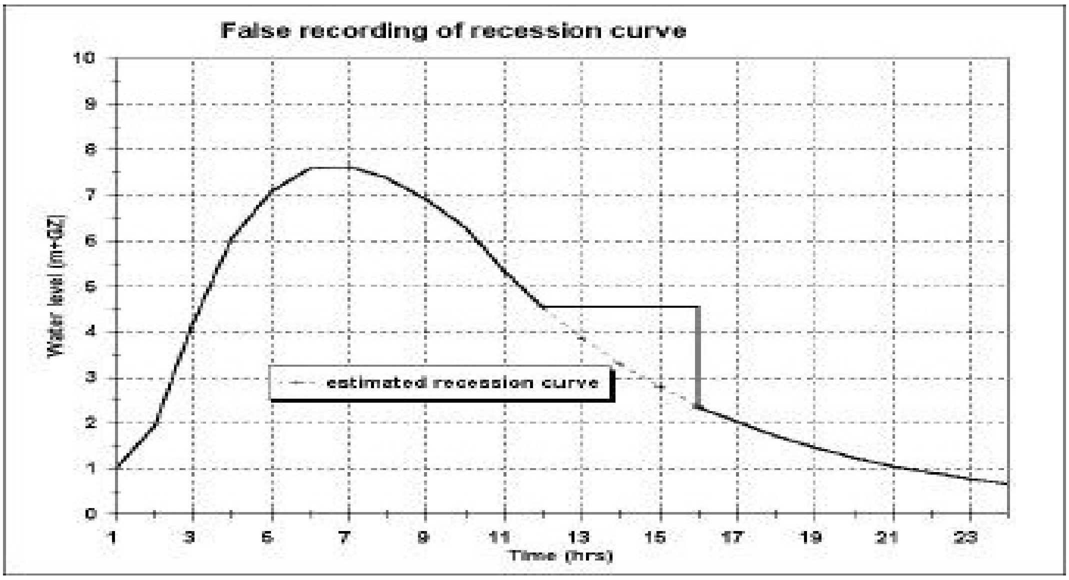


Figure 4.1: False Recording of Recession Curve

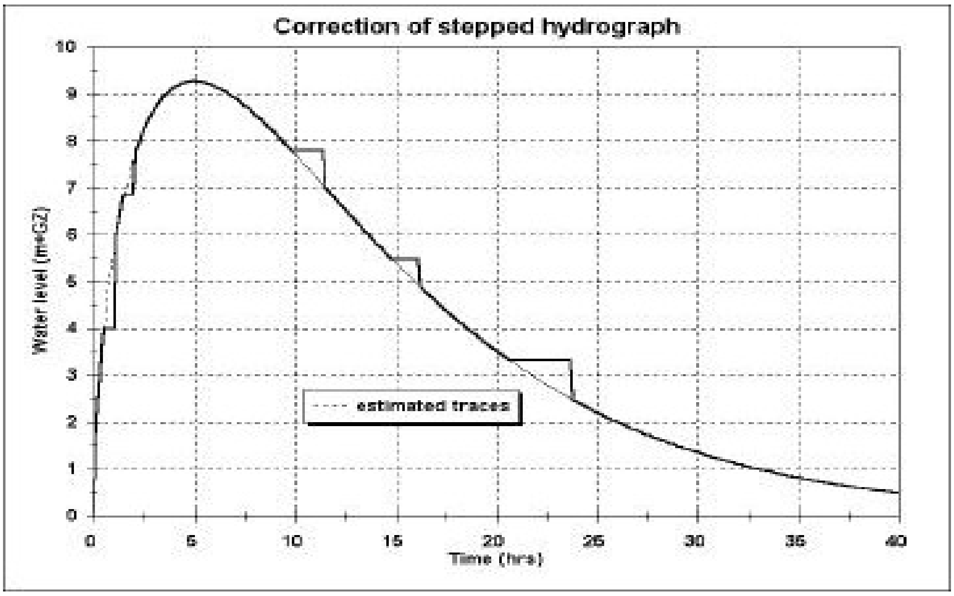


Figure 4.2: Correction of a Stepped Hydrograph

4.4 Correction for Bubble Gauge Recorders

Errors in the use of bubble gages for stage sensing include datum corrections, sediment deposition on the bubble orifice, and leaks in the system (Source: https://kacv.net/brad/nws/lesson4.html). The sources of error in a bubble gas record include the variations in gas friction and required bubble feed rate with increasing water levels and variation in the weight of the gas column with the stage. These can be rectified by calibration at the site.

Additionally, the conduits can be dislodged during high flow by debris or moving rocks. As a result, the sensor elevation can be moved up or down. If it has moved up, it will experience reduced hydrostatic pressure from lower depths of water, and consequently, record lower water depths. This can be rectified noting the time when the sudden change of elevation took place.

Also, the orifice may be blocked by sediment, generating oscillations in the stage-discharge plot as the cycle of the build-up of gas pressure followed by its sudden release continues. Such data is suspect and should not be used. Frequently, the systems are equipped with a purge system to send high-pressure gas for clearing it permanently.

Oscillation may also be generated due to turbulence around the gauge orifice, where alternating high and low pressures are reflected as changes in depth of water.

If the depth of water exceeds the limiting water depth for the pressure transducer, records are truncated. Operations resume after water level falls below the limiting range.

4.5 Correction for Ultrasonic Gauges

Ultrasonic sensors calculate the distance of the object using the speed of propagation of sound waves. In the air, the speed of sound is about 343 m/s at a temperature of 20°C. It is temperature-dependent, and changes by approximately 0.17% with each degree Celsius, affecting the transit time and distorting the calculated distance. So, most of the ultrasonic sensors are equipped with temperature probes to measure the temperature and use it to correct the measured distances. The measurement is also affected due to changes in relative humidity, but the effects are smaller than those due to temperature.

Turbulence, foam, steam, mists (vapours), and changes in the concentration also affect the performance of the ultrasonic sensors. Turbulence and foam prevent the sound waves from being reflected to the sensor properly. Steam, mists and vapours distort or absorb the sound waves. Variations in concentration cause changes in the amount of energy in the sound waves that is reflected to the sensor. Waveguides are used to prevent errors caused by these factors. The data capture program associated with the sensor can be manipulated to take averages of multiple instantaneous readings (say 20), and avoid erroneous measurements using false echo settings.

4.6 Correction for Radar Gauges

Unlike their ultrasonic counterparts, the radar or microwave-based water level sensors are largely unaffected by high temperature, pressure, vacuum or vibration. They allow operation under the condition of high pressure and vacuum, high temperatures, dust, temperature and vapour layers.

The main factors affecting the accuracy of non-contact radar are the relative permittivity of the medium (also known as the dielectric constant) through which the microwave radio signal must propagate (and off which it must reflect), multi-path interference from metal obstructions and the signal loss due to signal dispersion or other factors such as foam.

4.7 Linear Interpolation of Short Gaps

Where only a single record is available at a station, gaps may occur due to instrument failure, observer sickness, station maintenance, etc. Gaps may be infilled by simple linear interpolation where they occur during periods of low flow or during the recession and the difference between the level at the beginning and end of the gap is small. During periods of low flow, gaps of one to several days may be infilled in this way. However, it is recommended that infilling by linear interpolation during the monsoon or on a heavily regulated river should not exceed 6 hours.

For longer periods of missing data during a recession when the runoff is resulting only from the outflow of a groundwater aquifer, the flow will show exponential decay. When plotted as discharge on a semi-logarithmic scale, this will plot as a straight line. Using the stage-discharge relationship, it is possible to infill the series as water level rather than a flow, but infilling as the flow is conceptually simpler. Gaps of a month or more may be filled in this way but only for hydrometric stations that record unregulated (natural) flows.

4.8 Use of Data from Adjacent Stations

4.8.1 Background

Data from other stations may be statistically related to the data at a particular station where missing records need to be filled or the suspected records need to be corrected. This is especially true for sequential stations on a river with little lateral inflow from the intermediate catchment. The following are the typical uses:

- infilling of missing records
- identifying and correcting errors in one series
- identifying and correcting shift in gauge zero or change in cross-section

4.8.1.1 Infilling of missing records

In general, there are two standard ways to fill the missing data by using data available from other nearby stations:

- For stations with both flow and water level data, transpose the flow data from another station by adjusting the transposed data using the effective catchment area ratio. The catchment area ratio is obtained by dividing the catchment area of the station whose data are being filled with the catchment area of the station from which the data are borrowed to perform the in-filling. The transposed flows can then be converted to water levels by using the established rating curve for the station with missing water levels.
- Develop regression equations between the station with complete data, which acts as an independent variable in the regression, and the station that is being filled, which is a dependent variable. When using regression approach, the following general rules should be observed:
    - Regressions change depending on the season, so there should be separate regression equations for monsoon seasons and dry seasons. Sometimes it may be a good approach to develop separate regression equations for each month.
    - Regressions can sometimes result in negative values or impose a lower limit on the generated data that is equal to the regression constant. Either of those instances will require manual corrections that will require professional judgment.

- No more than 10% of all data should be filled by using the regression approach. Since the typical regression equation introduces linear dependence between the independent and dependent data set, the resulting long-term cross correlation coefficient between the two stations is increased. Results of data in-filing should not alter any important historical statistics related to the data. This includes the autocorrelation coefficients for various lags and cross-correlation coefficients with other adjacent stations, along with monthly means and standard deviations. All of these should remain within a close range before and after in-filling of missing data.

An example given in Figure 4.3 shows an ideal situation with a smooth hydrograph, where missing data can be filled by developing a non-linear function that fits a few data points before and after the gap. This situation may be more straight forward than the use of relational curve based on another station, which may result in an offset as shown in this example. Therefore, one should always visually verify the correctness of the infilled data.

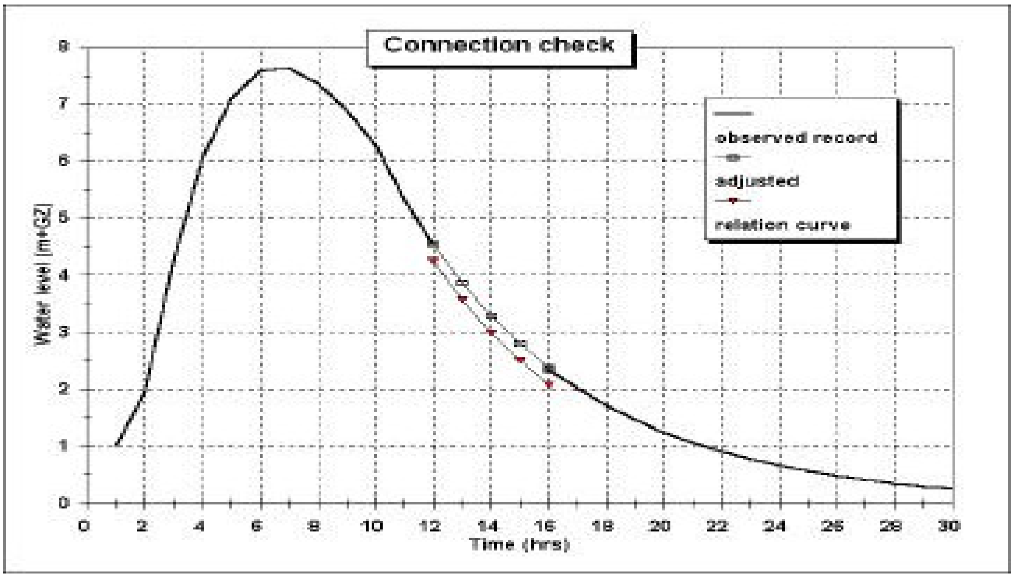


Figure 4.3: Infilling of Missing Data with the Relation Curve

4.8.2 Identifying and correcting shift in gauge zero or change in cross-section

Shifts in water level observations can happen due to change in gauge zero or changes in cross-section conditions. For routine validation and completion, gauge zero level changes should be confirmed with an urgent cross-sectional survey. It should not be determined on the basis of the flows at another station.

On the identification of such changes, consultation should be made with the sub-divisional staff and the Field Record Book should be inspected. If the conditions of change had been previously recognised in the field and adjustments made to the rating curve to account for the shift in gauge zero (or change in station location) due to altered cross-section, then no further action needs to be taken. If the change had not been recognised in the field, then since the analysis does not indicate which station is in error, further action is necessary along the following lines:

- The field staff are requested to re-survey gauges and the cross-section at all stations possibly affected with the shift
- If, after the survey, the gauge zero at one station is found to have been altered without justification, then it should be reset to its former level. The stage level during the period between gauge shift and resetting should be corrected by the deviation shown by the survey (and confirmed by the constant difference in relation curves, provided a meaningful relationship is found between the data from both the stations).
- If no change in gauge zero is found but the cross-section at one station has been altered, then the field staff are requested to intensify current meter gauging to update the stage-discharge relationship. Usually, the stage record is not changed, but the revised rating curve applied over the period from the occurrence of the change in cross-section (usually during a flood).

4.9 Concluding Remarks

Various primary and secondary validation techniques create data outputs which can be used to flag the suspect values. Some records may also be missing due to non-observation or loss during recording or transmission. This identifies the need to fill data gaps and correct the errors. The process of filling the missing data by estimated values based on other observations is referred to as "Data Completion".

5 PRIMARY VALIDATION OF STAGE-DISCHARGE DATA

5.1 General

Flow measurement in this chapter primarily refers to individual measurements of discharge made by a current meter. Thereafter, this is used in the plotting and fitting of a stage-discharge relationship known as the "rating curve". This is subsequently used to convert recorded water levels to flow estimates. This has by far been the most commonly followed technique all over the world. There are a host of other methodologies available, like the dilution technique or the tracer method for the fast streams, measurements at flow control structures, slope area method, the magnetic method, and the modern methods like use of ADCP, use of video camera, or using the satellite image processing techniques. Measurements at flow control structures like weirs are used at some places and measurements using ADCP is gradually gaining popularity in the country. Therefore, a brief discussion on these has also been included in this manual.

The initial calculation is carried out in the field and the completed field sheets are sent monthly to the Sub-divisional office or data centre. The data is entered using the Primary module of dedicated hydrological data processing system (WIMS at present) and the discharge is recomputed there.

Primary validation consists of:

- inspection of field sheets and Field Record Book
- comparison of discharge calculated at the field with that at the office
- comparison of computed discharge with an existing rating curve
- comparison of cross-sectional and velocity profiles

5.2 Inspection of Field Sheets and Field Record Book

Each current meter measurement of discharge contains multiple observations or calculations of width, depth, velocities, slope, areas, flows, etc. The information is entered into the standard "Discharge Measurement Sheet" (Figure 5.1). Before checking the arithmetic calculations of the discharge, it is necessary to check for additional information on the form and the Field Record Book. This is to ensure that the information is complete. The check is also done to understand whether any change has occurred at the station which may have influenced the relationship between stage and discharge. The possible reasons for a change in the stage-discharge relation at the station may include the following:

- high rates of rising and falling water level during the measurements (possibly due to unsteady flow effects)
- backwater due to very high stages (i.e., flooding) in the main river or contributing tributary downstream of a gauging station
- flood causing deposition or scour of the channel at the gauge site or at the downstream control, based on observations
- gravel or sand extraction at the station or downstream
- bunding or blockage in the downstream channel
- weed growth in the channel
- change in the datum at the station, adjustment or replacement of staff gauges.

The stage recorded at the beginning and the end of the current meter gauging must be compared with the hourly or other stage observation by the automatic or manual record. Any discrepancy must be investigated by cross-checking with the field supervisor. The errors may occur in the continuous record or the observation during the current meter gauging. The mean stage in the summary form for the current meter measurement must be amended in that case. The Central Water Commission maintains a comprehensive sample flow monitoring measurement sheet as shown in Figure 5.1 below. A similar form (based on an earlier version of this sheet) exists in e-SWIS (now India-WRIS) data platform. This form allows both data entry and automated recalculation of the mean river flow along a cross section. This is achieved by summing up the measured discharges that are obtained by multiplying the sectional area of the part of the cross section under consideration with the respective mean velocity. Velocities are typically measured at 0.6 of the water depth at the section measured from the water surface, if a single measurement is made. Otherwise, it is measured at at 0.2 and 0.8 of the water depths measured from the water surface, where two measurements are taken. These velocities are averaged to provide a representative mean velocity of the given section.

के.ज.आ./ आर.डी.-1/ CWC/RD-1

भारत सरकार / Government of India

केन्द्रीय जल आयोग / Central Water Commission

दैनिक निस्सरण आँकड़े / Daily Discharge Data

प्रेक्षण न. / Observation No...

नदी/River...................स्थल/ Site....................कोड सं0/ Code No. ....दिनांक/ Date....................समय/Time from................बजे से/to............बजे तक

Mode of Crossing:- By Wading/Boat/Cable Way/Bridge/Boat with OBE/ Boat With IBE/ Boat with Cable Way/Bank Operated Cable Way/Cable Way with Trolley...

वेग अवलोकन की विधि/Method of Velocity Observation:-Floats/ Current Meter/Slope Area/Dilution Method/ Floats with Respect to RDs/ ADCP...................

Location of Discharge Site/ निस्सरण स्थल का स्थान :

A) स्थायी स्थल / Permanent Site...

B) अस्थायी स्थल / Temporary Site...

स्थायी स्थल की दूरी / Distance of Permanent Site ....

U/S/ अनुप्रवाह / D/S/ प्रतिप्रवाह ....

Sounding taken weight:- भार का प्रकार Wading Rod/Sounding Pole/Metallic Reel/ Echo Sounder/Cross Section.

Sounding weight use: भार किलोग्राम/ (Kg/Lb)...

जल अवस्था / Condition of Water काफी साफ / Fairly Clear..

सामान्यतया गदला / Ordinary Silty...

बहुत गदला / Intensely Silty...

नदी जल का ताप (सें0ग्रे0) / River Water Temperature (^{o}C)...

औसत जल तल (मानक तट)/Mean Water Level (Standard Bank)...मी0/Meter

वायुमंडलीय ताप / Atmospheric Temperature -अधिकतम/ Max...(^{o}से0ग्रे0)/^{o}C...न्यूनतम/

Min..(^{o}से0ग्रे0)/^{o}C औसत जल तक (मानकतट)/Mean Water Level (Standard Bank)... मी0/Meter

मौसम की दशा / Weather Condition...

वायु की दिशा प्रवाह के साथ / Direction of Wind w.r.t. ....

वायु की शक्ति: बहुत भट्ठा / थोड़ा सा / मजबूत / बहुत मजबूत/Strength of Wind: Very Slit / Slight / Strong / Very Strong...

वायु का वेग/Velocity of Wind: कि० मी०/घ० Km / Hr ....

गेज सूचना/Gauge Information

शून्य सापेक्ष तल (जी.टी.एस)/Zero RL (GTS) .... (मी0)/m Separately entered in RL of gauge zero

गेज/ Gauge	स्थायी / Permanent	अस्थायी / Temporary
प्रारम्भिक / Beginning		
अंतिम / End		
औसत / Mean		

वर्तमान मीटर अवलोकन / Current Meter Observation

माप से पहले चक्रण / Spin Before Measurement...

माप के बाद चक्रण/ Spin After Measurement...

निलंबित मीटर लेबल की विधि / Method of Suspending Meter Label:...

उपयोग किए गए मीटर स्तर के साथ वजन / Weight with used meter level:...

मीटर नo / Meter No...

खंड संख्या/SECTION NO.	dkV ij 1kis{k nwjh ¼vkj-Mh-½ RD OF SECTION (m)	ty dh xgjkbZ ¼ehå½ WATER DEPTH (m)	m/okZ/kj dks.k (fMxzh½ VERTICAL ANGLE (Deg.)	,;j ykbu xgjkbZ ¼ehå½ AIRLINE DEPTH (m)	,;j ykbu करेक्शन¼ehå½ AIRLINE CORRECTION(m)	वेटलाइन करेक्शन ¼ehå½ WETLINE CORRECTION (m)	कुल संशुद्धि ¼ehå½ (6+7)Total Correction(m)	ty dh संशुद्धि xgjkbZ ¼ehå½ CORRECTED WAT ER DEPTH	खंड क्षेत्रफल ¼eh-^{2}½ AREA SECTION (m2)	वेग अवलोकन संख्या/ NO OF VELOCITY OBSERVATION	परिकलित वेग CALCULATE D VELOCITY	औसत वेग (मी./से.) MEAN VELOCITY (m/s)	dks.k ANGLE OBLQ	संशोधित औसत वेग (मी./से.) CORR. MEAN VELOCITY(m/s)	अपवाहन दूरी (मी.) DRIFT DISTANCE (m)	समय अपवाहन (सै०) TIME DRIFT (S)	अपवाहन संशोधन DRIFT CORRECTION	अंतिम औसत वेग FINAL MEAN VELOCITY	खंड निस्सरण (मी.^{3}/से.) DISCH SECTION (m3/s)	टिप्पणी REMARKS
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

परिकलित Computed:	शीर्ष चौड़ाई Top Width:	भीगा परिमाप W. Perimeter:	कुल क्षेत्रफल Total Area:	कुल निस्सरण Total Discharge:	मुख्य गेज रेखा पर जल विज्ञान संबंधी कारक Hydrological Factors at Main Gauge Line	Top Width Parameter	भीगा परिमाप Wetted Perimeter
In Form:	शीर्ष चौड़ाई Top Width:	भीगा परिमाप W. Perimeter:	कुल क्षेत्रफल Total Area:	कुल निस्सरण Total Discharge:			

Location of Level Observations स्तर के अवलोकन का स्थान	Right Bank Details/ दाये किनारे का विवरण		
	मध्य गेज लाइन से दूरी Distance from CGL	स्तर रीडिंग /Level Reading	मध्य गेज लाइन से दूरी Distance from CGL
U/S Gauge Line: अनुप्रवाह गेज लाइन:			
Central Gauge Line: मध्य गेज लाइन:			
D/S Gauge Line: प्रतिप्रवाह गेज लाइन:			

ढाल / Slope:

औसत ढाल / Mean Slope:

Figure 5.1: Specimen Discharge Measurement Sheet

5.3 Comparison of Discharge Computed at Field and at Office

The calculation of discharge from current meter measurements is initially carried out in the field by the gauging team. On receipt of the discharge measurement sheet in the sub-divisional office or data centre, individual observations made during the field measurement are entered in the computer and the discharge is re-computed with the relevant software. If the total discharges resulting from the two calculations differ, the source of the difference must be identified and the necessary corrections made. In particular, line by line comparison of the two calculations should be made to identify errors of data entry into the computer. If none are found, arithmetic errors should be checked in the field calculation. Other potential sources of discrepancy can be in the use of the wrong current meter reading in one of the calculations or incorrect entry of current meter rating parameters to the rating datafile. Any errors in the field computation should be communicated to the field supervisor.

5.4 Comparison of Computed Discharge with an Existing Rating Curve

The newly computed discharge can be compared graphically by plotting with the discharge obtained using the existing and previously validated rating curve. A table of the actual and percentage deviation of the gauging from the previously established rating curve can also be prepared.

Deviations may be due to:

- the reliability of individual gauging
- the general accuracy with which measurements can be made at a station
- actual changes in the river cross-section due to erosion or sediment deposits

Early identification of such deviations is necessary so that the gauging practices can be re-established and readjusted, in case there is a change in the rating relationship between flows and water levels. The percentage deviation of a gauging which requires further action will depend on the physical characteristics of the station and the assumed accuracy of individual measurements taken. For example, in a station with sensitive control and a regular gauging section, an error of 5% may only be acceptable but at irregular sections, with erratic velocity distribution, an error of 10% may have to be tolerated. In general, the individual gauging should be investigated and possibly the measurements should be repeated, if necessary, if the deviation from the computations using the previous rating curve exceeds 10% or if a sequence of gauging shows persistent positive or negative deviations from the established rating.

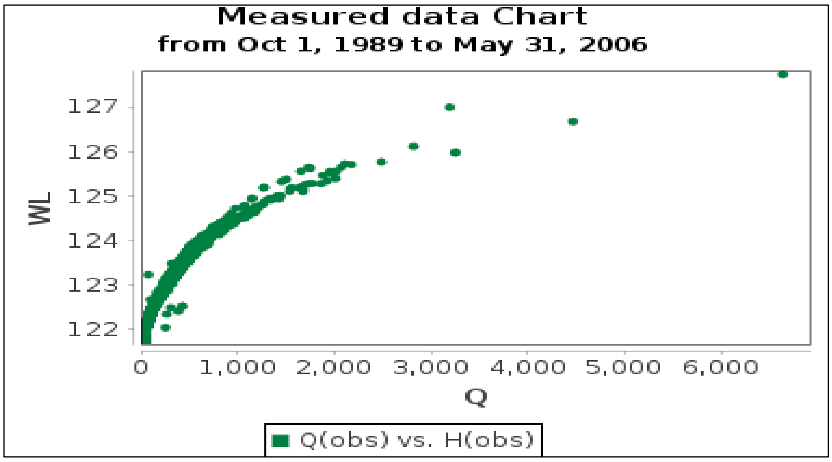


Figure 5.2: Scatter Plot of Stage Discharge Data

5.4.1 Deviations due to variation in the reliability of individual gauging

The individual gauging may be unreliable due to the following reasons

a) an inadequate number of vertical subsections taken to define the total area and mean velocity
b) very low velocities in the section not measured accurately by the equipment used
c) no wind /wet line corrections made to the depth measurement in high flow
d) no angle correction made for gauging taken oblique to the flow (cross sections should be laid out in such a way that they are perpendicular to the direction of the flow)
e) a faulty current meter

Items under the serial number (a) to (d) can be identified from the tabulated gauging. The use of a faulty current meter (e) cannot be identified so easily but it may be identified from field inspection or by persistent differences between the results from the specified meter and other meters at the same station. A plot of the cross-sectional velocity profile can be made for individual gauging and a comparison made between the gauging done at the same stage.

5.4.2 Deviations due to physical properties of selected river cross- section

The general accuracy with which gauging can be made at a station depends to a large extent on the regularity of the river bed and river banks at the gauging cross-section and also on the approach conditions. Both the roughness of the river bed and the curvature of a bend are subject to change. These control the velocity distribution across the cross-section, which differs from that of a smooth trapezoidal channel. Irregularities may result in deviation from a typical logarithmic vertical velocity profile, so that neither a depth 0.6d nor (0.2d + 0.8d)/2 measured from the water surface represent the mean flow, which are the commonly used is the formula. The irregularities may cause rapid velocity variations across the cross-section, and sometimes the number of verticals chosen may not adequately represent the mean flow.

The velocity distribution over the cross-section may be investigated by plotting velocity contours or velocity vectors across the cross-section with a sufficient number of observations taken.

5.4.3 Deviations due to actual changes in the stage-discharge relationship

Deviation from a simple power relationship at a gauging station may arise due to many reasons including the following:

- Unsteady flow causing hysteresis during the rising and falling floods. This can be identified by plotting the rate of rising (+) and of falling (-) during gauging, alongside the plotted point on the stage-discharge graph. Higher flow for a given level may be expected in rising flows when the energy slope is greater but this is generally only evident in reaches with low channel slope. For falling levels, the flows will be lower for the same level.
- Changes may occur in the cross-section at the control section due to natural scour or sedimentation. Such changes may be identified by plotting sequential cross-sections for the control section where available. Occasionally it is due to the changes in the control section which are critical, but sometimes these are accompanied by changes in cross-section at the gauging station. These can indicate the existence of scouring or deposition at the station. At least two cross-section surveys are to be conducted each year before and after the monsoon period, and both should be compared. Besides, the cross-section profile available from each current meter gauging may also be compared to analyse the presence of scour or deposition at the station. Reference should be made to gauging notes and to the Field Record Book for observations of the field staff. Introduction of a new rating curve or the use of the shift procedure should be considered under such a situation.
- Discharge for the given water level may also be affected by downstream river bed changes even if no change is found at the gauging station. In channels with low slope, the control may extend to many kilometres downstream for which no cross-section information exists.
- Comparison of mean velocities between sequential gauging across the width of the channel at the gauging section will help to identify such changes (though backwater may exhibit the same effect). Scour or gravel extraction downstream will result in increased velocity for given gauge level; bunding and the blockage will result in decreased velocity. Reference should again be made to gauging notes. In such cases, the introduction of a new rating curve or the use of the shift procedure should be considered.
- Similarly, the discharge for the given level may be affected by downstream backwater conditions caused by a confluence or by tidal effects (for example). These effects may also be seen during a comparison of velocity profiles. Unlike the effects of the downstream bed changes, these effects may not persist from one gauging operation to the next. For occasional effects caused by backwater, rating curves with backwater corrections should be applied.

- Weed growth downstream of the station as well as at the station may also be identified through changes in the mean velocity profile across the section. Weed growth decreases the velocity for a given level. Reference should be made to gauging notes. If weed growth causes significant variation from the mean rating, the introduction of the shift procedure should be considered.

Where bed profile and mean velocity profiles remain sensibly constant from one gauging operation to another but the plotted points deviate from the previous rating, then a change in the datum or a shift in the staff gauges should be suspected. Reference should be made to the Field Record Book and gauging notes. Field staff should be requested to carry out a check survey of the staff gauges.

5.5 Discharge Measurements at Measuring Structures

On small rivers it is often convenient to measure flows using a weir or a flume. Such structures have the advantage that they are less sensitive to the downstream conditions, the channel roughness and the influence of backwater compared to the velocity-area method applied directly to a river channel. The relation of discharge to water level measured at a prescribed distance upstream of the structure is found empirically or is based on physical principles. The height of the structure is chosen in a way that the downstream water level does not affect the flow. Typical structures may be classified as:

(a) thin plate weirs
(b) broad-crested weirs
(c) flumes
(d) compound measuring structures.

Further details on derivation of formulas for flows over weirs can be found in the standard text books on hydraulics.

5.6 Discharge Measurements Using ADCP

Measurement of discharge using Acoustic Doppler Current Profiler (ADCP) has become very common these days. It is a device that uses sound waves to measure the speed and direction of currents throughout the water column. It uses a series of acoustic transducers that emit and receive pings from different directions. Higher frequencies like 300 kilo Hertz (kHz) are used to provide high-resolution data near the surface, up to a depth of about 70 meters. Lower frequencies like 38 kHz, can be used to provide lower-resolution data to a depth range of up to about 1,300 meters.

The equipment is towed across the river with the help of a boat. The path need not be straight or perpendicular to the bank. The velocity is measured by the Doppler principle (change in frequency of the sound waves reflected by moving sediment particles or air bubbles in water). The area is measured by tracking the bed to provide the river depth and boat position. With specialised software, the discharge is directly available as output, eliminating the need of carrying out computations. The uncertainty of measurement by this method has been reported to be limited to only about ±5%. A photograph showing ADCP in operation on the River Brahmaputra in Assam is shown in Figure 5.3.

The principle of operation is given below in brief:

- An ADCP is a cylinder with transducer heads on the end. It has three or four transducers with their faces at angles to the horizontal and at right angles to each other.
- It subdivides the water column being sampled by each of three beams into depth cells or bins, ranging from 0.01m to 1m or more.
- A three-dimensional water velocity is determined therefrom, and assigned to a given depth cell in the water column.
- A single traverse/ transect may contain thousands of velocity measurements collected continuously across the river width, compared to a conventional current meter measurement of some 20 verticals.
- If the measured discharge on any given transect differs by more than 5% of the mean discharge, a further set of transects is made.

Figure 5.3: Discharge Measurement Using ADCP

5.6.1 Sources of errors

The measurement of discharge using ADCP may be affected by the following:

- The number of suspended particles in the water where too many particles (as encountered during the monsoon high flows) may reduce the penetration.

- Rivers with high bed-load, or high sediment concentrations near the bed, create additional problems of negative bias because the ADCP measures the speed of the moving sediments near the bed.
- External magnetic fields may influence the measurements.
- Errors are more frequent near the channel margins, presumably due to obstructions and multi-paths associated with riverbank vegetation and buildings.
- The discharges near the top, bottom and edge are not results of measurements but extrapolations by the software package.

5.6.2 Improvement of the of results obtained through ADCP

Even though the measurement errors are limited with ADCP, it can be further improved using

- Using DGPS for recording locations of discharge measurement (Rennie and Rainville, 2006).
- Loop correction method based on the closure error resulting from a two-way crossing of the river may be applied for moving bed corrections. This is particularly done for areas where DGPS cannot provide consistently accurate positions because of the multipath errors and satellite signal reception problems on waterways with dense tree canopy along the river banks, in deep valleys or canyons, and near bridges (Mueller and Wagner, 2006).

5.7 Selection of Discharge Measuring Method

The selection of the method for measuring discharge will depend on many factors including the

- accuracy required
- availability of equipment
- availability, skill and experience of personnel
- accessibility of the site and the stream
- costs
- width and depth of the stream
- range of flow velocities
- frequency of measurements

In general, the possibility of installing a velocity-area station is considered first, which is followed by the establishment of a relation between stage and discharge. Discharge measurements may be carried out using the current meter by wading (when the depth and velocity are small enough), by cableway (when the span permits its installation), by moving boat (if the river is wide enough), by floats (if the velocity is too low or too high to use a current meter or there is ice in the river), by slope-area (if no other method is suitable during floods) or from bridges (if a suitable one is found).

The ADCP is now being widely used for measurement of discharge, with very good results. Equipment to measure both large and small rivers and deep or shallow rivers when mounted on motor launches or small remote-controlled or tethered rafts or catamarans are available now. Another advantage of the method is its speed, as an ADCP measurement may be ten-times faster than the conventional method. The only apparent factor that currently works against ADCP is probably the cost of the equipment (around 20 to 25 Lakh Indian Rupees) and limited purchase and maintenance options available.

In small rivers that are less than 100 m in width, a measuring structure may be considered, particularly if backwater conditions are prevalent. The main factors to be assessed for a measuring structure are its cost, head loss (afflux) available, Froude number and bed conditions.

The ultrasonic method provides a continuous measurement of discharge for all designed stages of flow and continues to do so under backwater conditions even if the flow actually reverses due to tidal influence. The main restrictions for the ultrasonic method are its requirement of continuous electrical power, limiting river width of 300 m or less with suitable minimum depth and without weed growth or significant sediment transport.

Dilution technique is not in general use, as the technique requires specially trained staff. It is the most suitable method for discharge measurement in turbulent mountain streams, mainly used for spot measurements to calibrate the other methods. It is also the only fully direct method for the measurement of discharge since the velocity, depth or area does not enter into the computation.

Stage-fall-discharge and slope-area methods are indirect methods of measurement, but are used under conditions where other methods are not suitable or are unavailable. The stage-fall-discharge method is used under backwater conditions in large rivers, where it may be the only method suitable. The slope-area method is used for the measurement of flood discharges, either current or historical, the latter from flood marks.

5.8 Concluding Remarks

The errors/ suspect values of discharge that cannot be corrected through the application of primary validation techniques should be considered for rectification using the procedures of secondary validation.

6 SECONDARY VALIDATION OF STAGE-DISCHARGE DATA

6.1 General

Rating curves are usually developed and validated for the flows observed at individual stations. It is often necessary to extrapolate the relationship beyond the measured range.

One means of providing a further check on the reliability of the extrapolated rating curve is to make a comparison of the discharges from the neighbouring stations, computed using the stage-discharge relationships. The other option may be to calibrate the HEC-RAS model for high flow range and obtain simulated water levels for high flows that have not yet been encountered since the start of the observation. Both options only provide approximate results that should be in close agreement with the values on the extrapolated segment of the curve. A perfect match is not expected, due to various uncertainties and assumptions that have to be made.

6.2 Development of Rating Curves

Rating curve is obtained by finding the best fit line between the measured water levels and flows, where the measurements should come from both high and low flow seasons, thus covering an extensive range of flows. The curve typically displays water levels on the vertical axis and flows on the horizontal axis. This is a bit counter-intuitive since eventually the flows are obtained as the dependent variable, based on the water level readings from the field.

Rating curve is a non-linear function obtained as a result of finding the best fit line among the measured flows and the observed water level data points. In general, rating curve is a convex line on an elevation vs Q plot, with a gradient gradually reduced with the increasing flow values that are normally shown on the horizontal axis. The principal difficulty in constructing the rating curve is the fact that the [flow, elevation] data points assume fixed cross-sectional channel geometry, which is normally not the case due to river sedimentation. The selection of hydrometric station location should involve stable cross-sectional area, which is preferably reinforced by additional armoring (e.g., at bridge or downstream of weirs). However, sedimentation can still cause significant differences in the same cross-sectional areas between the monsoon and dry periods, which may result in the same flows being measured for different observed water levels at different times of the year, or in different years. Hence arises the need for finding the best fit line among the available points.

Typical functional fit of the [elevation, flow] data pairs is achieved by using the following options:

a) Parabolic, represented by a quadratic equation of the for $y = ax^2 + bx + c$, where parameters a, b and c are determined as part of the functional fitting procedure; or,
b) Power form, typically represented by the power equation of the form $y = ax^n + b$, where parameters a, n and b are determined as part of the fitting procedure.

In either case, the fitting procedure involves solution of the system of equations that define the fitting parameters such that the sum of deviations of observed y coordinates from the fitted curve is minimized.

Functional fits to the rating curve data should have two types of flexibilities resulting in possibly more than one functional fit when a single function does not fit the entire data range well. Two criteria should be available for splitting the set of [elevation, flow] data points:

a) split into two or more sub-sets based on the flow range; and,

b) split into two three sub-sets based on the season of the year.

Also, since a few statistical outliers can have significant effects on the fitted function, the fitting model should have a flexible filter that allows the user to remove the outliers. This document contains the instructions to implement a filter of this kind.

Some basic concepts are outlined below, which formed the basis of technical specifications for rating curve fitting written for an upgraded version of India-WRIS / WMIS. The two splits defined under a) and b) above are further discussed thereafter, and demonstrated through a numerical example.

6.2.1 Available data range

This includes all years of available [elevation, flow] data pairs. The range shows the starting and the ending date of the available record in the database. Users should be able to select to fit either Parabolic or Power function, and this selection will be in effect for all curve fitting until it is changed. Each time the new dataset is read, the fitted equation for the selected fit (Parabolic or Power) should be shown below the Parabolic / Power radial buttons. Both the parabolic (or polynomial functional fit of order 2) and the power fit are the available data fitting options in Excel. To use either option, it is necessary to create two data columns in excel, where the first data column represents recorded water levels and the second data column represents measured flows for each corresponding data level.

6.2.2 Custom data range

If users do not want to use the entire length of record, due to the older historic data being less reliable, or due to river channelization works that may have been completed in the past (which would also affect the shape of the rating curve), a shorter customized period of record should be used. Users should carefully select the starting and ending dates thus defining a customized data set, but they should ensure selecting a period for which the data are available.

6.2.3 Seasonal data plots (options 1 and 2)

Due to the changes of the river bed caused by the sediment flow regime, it is sometimes advisable to separate the data for monsoon (or non-monsoon) months and create separate rating curves for high and low flow seasons. This assumes selection of the starting and ending period for inclusion of data into the fitting procedure, based on the typical start and the end date of the monsoon season. This plot can be shown as a subset of the previous plots explained in the two previous paragraphs.

6.2.4 Compounded curve fit for braided channels or discontinuities in the data

Users should be able to select more than one flow data ranges (up to three) thus creating multiple functional fits based on the range of flows. This decision follows the initial visual data inspection on a scattered X – Y plot. The neighbouring flow ranges can overlap at the crossings, and it is up to the user to define the range such that there is a smooth transition achieved visually from one curve to the next. Once this is achieved, the actual point at which one equation is no longer valid and the next equation becomes valid is determined at the intersection point or in its close proximity, which the user can estimate visually from the graph (or solve analytically if necessary). The final rating curve may then be a composite curve consisting of three mathematical functions, each one valid within its own data range

6.2.5 Removal of statistical outliers

Users should be able to recognize and remove statistical outliers from the plots. This is achieved by identifying the points with the highest dispersion (distance from the fitted line) and their removal. User should select how many points to remove, based on the initial inspection of the default plot that contains all data points.

By default, all data points are initially included in the data fit. Any time the user wants to discard one or more data points, the total number of data points that should be removed is placed in the text box, and the graph is updated by pressing the graph button in the lower left corner of the form. A new updated equation fit will be shown automatically for the selected fitting option (parabolic or power). To remove one outlier, users should delete both [elevation, flow] coordinates. Excel will then update the fitted line plot and its equation. This process is demonstrated on a numerical example where the aim is to remove six points with the worst data fit. Figure 6.1 shows a sample with the calculated error term Er[i].

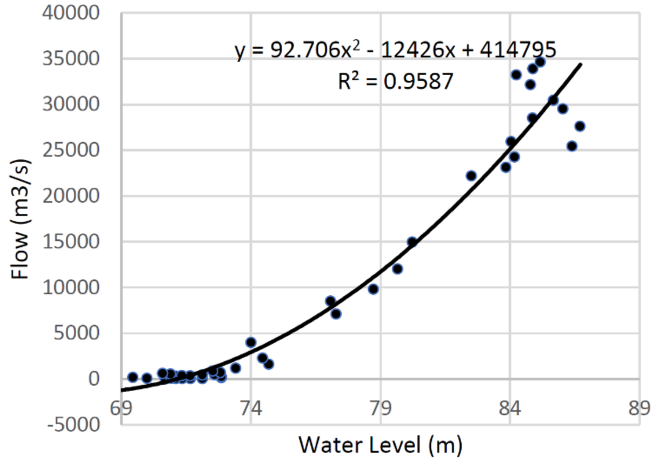


Figure 6.1: Initial Rating Curve Fit with All Available Data

The worst six data points are located at the upper end of the curve. Their error term can be calculated as the squared vertical distance $[f(x) - y_i]^2$ between the curve and the data point for the same X coordinate shared by the data point. The resulting fit is improved, as attested by the plot in Figure 6.2 below.

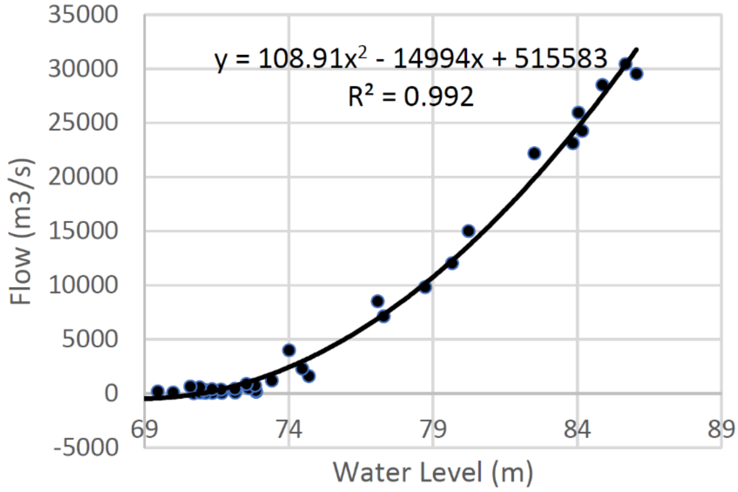


Figure 6.2 Improved Functional Fit due to the Removal of Outliers

The final fitted equation is in Figure 6.2, although the usual requirements is to plot this line with reversed axis (i.e., water levels on the vertical axis and flows on the horizontal axis. A common error is to plot the data with water levels on the Y axis and flows on the X axis and then enable functional fitting in Excel. This usually results in gross error, as shown in Figure 6.3, where the fitted line reaches a peak and then begins to fall, which should never happen.

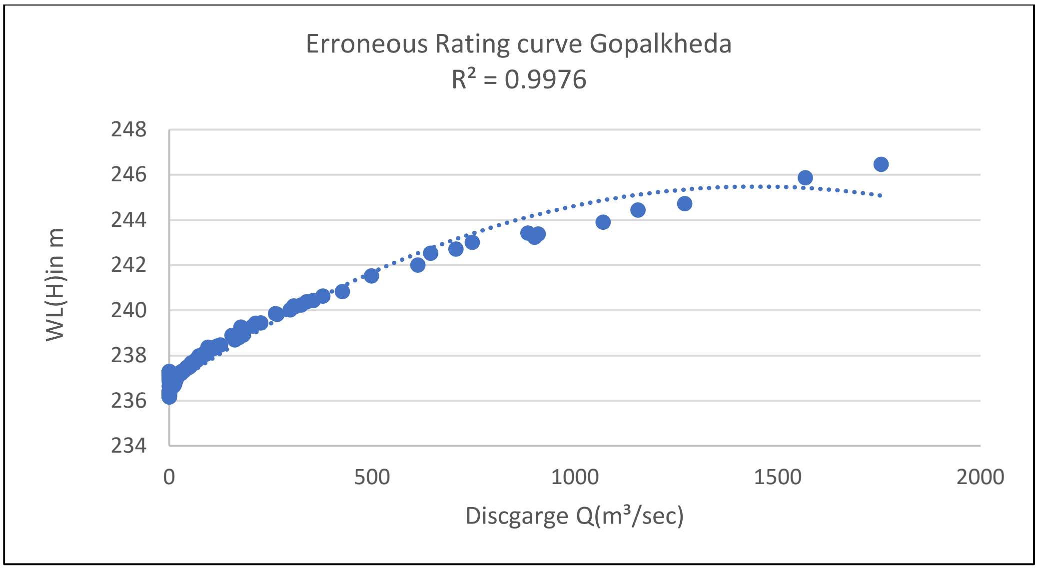


Figure 6.3 Stage-Discharge Rating Curve at Gopalkheda -Wrongly Fitted

7 COMPUTATION OF DISCHARGE DATA

7.1 General

With limited exceptions, continuous direct measurement of discharge is not common. Instead, measurements of the stage (or water level) are made continuously or at a specified time interval at a gauging station, and these are converted into discharge by using stage-discharge relationships later.

Computation of discharge is normally carried out every month based on the stage data of the previous month, but it should always be reviewed before transferring to the archive. Computation of discharge is carried out at Divisional offices and reviewed at the State Data Management Centre.

Computation of discharges by transforming the stage into discharges through various methods have been discussed in this chapter.

7.2 Station Review

Before computing the discharge, it is essential to have a summary of all the relevant information about the gauging station which includes the following:

- the stage records - to ensure that it is complete and without abrupt discontinuities.
- a listing of stage-discharge relationships to check that periods of application do not overlap or do not have gaps between ratings.
- Auxiliary information based on the field records (from the Field Record Book) or the information obtained from validation of stage or stage-discharge relationships. In particular, field information on datum changes, scour and deposition, blockage and backwater effects should be collected along with any adjustments or corrections applied during validation

7.3 Transformation of the Stage to Discharge

The procedure used to transform the stage into discharge depends on the physical conditions at the station and in the river reach downstream. The following alternatives are considered:

1. the single-channel rating curve
2. the compound channel rating curve
3. rating curves with unsteady flow correction
4. rating curves with constant fall backwater correction (the best way to permanently correct such errors is to move the hydrometric station to another location)

5. rating curves with normal fall backwater correction

7.3.1 The single-channel rating curve

When unsteady flow and backwater effects are negligibly small, the stage-discharge data are fitted by a single channel relationship, *valid for a given period* and *water level range*. Rating equations should be derived either as parabolic or power-law equations. It is assumed that in the majority of cases the power-law relationship can be applied. Equations for standard and non-standard gauging structures may also be re-computed in this form without much loss of accuracy.

The basic equations are as follows:

a) For the power type equation used for curve fitting.

Power form, typically represented by the power equation of the following form:

$$\boldsymbol{Q_t = c_{1i}(h_t + a_{1i})^{b_{1i}}} \qquad \text{Equation 7.1}$$

b) For the parabolic type equation used for curve fitting

$$\boldsymbol{Q_t = c_{2i}h_t^2 + b_{2i}h + a_{2i}} \qquad \text{Equation 7.2}$$

$\boldsymbol{Q_t}$ discharge at time t (m^3/sec)

$\boldsymbol{h_t}$ measured water level at time t (a.m.s.l.)

a_{1i},b_{1i},c_{1i} parameters of the power equation

a_{2i},b_{2i},c_{2i} parameters of the parabolic equation

i index for measured water level and the corresponding flow

The WIMS software uses two curve fitting options which can be developed separately for up to three sets of the data range.

Curve fitting procedure can be improved after the statistical outliers are removed, and the options for that are available in WIMS. The Power equation usually provides a good fit for Indian conditions. As mentioned previously, the data fit is automated nowadays by using the appropriate Excel function for parabolic or power data fit.

7.3.2 The compound channel rating curve

The compound channel rating curve is used to avoid large values of the parameter n and very low values of the a-parameter in the power equation at levels where the river begins to spill over from its channel into the floodplain.

When a compound channel rating curve is applied, the discharge will be computed as follows

$$Q_{Total} = Q_{River} + Q_{Floodplain} \qquad \text{Equation 7.3}$$

Q_{Total} total discharge

Q_{River} discharge flowing through the main river channel section up to the maximum water level

$Q_{Floodplain}$ discharge flowing through the flood plain section.

7.3.3 Rating curve with unsteady flow correction

Where an unsteady flow correction is required, the application of the simple rating curve first yields a discharge for the steady flow which must then be multiplied by the unsteady flow correction to generate the discharge for the required rate of change of water level. The relation between unsteady and steady flow discharge is as follows

$$Q_t = Q_{st} \sqrt{\left(1 + \frac{1}{C \times S_0} \frac{dh_t}{dt}\right)} \qquad \text{Equation 7.4}$$

And from the above equation, the following equation is derived

$$\frac{1}{CS_0} = \frac{(Q_t/Q_{st})^2 - 1}{dh_t/dt} \qquad \text{Equation 7.5}$$

where:

Q_t is the required discharge corresponding to the observed stage h_t and rate of change of stage (dh_t/dt) (+ for raising and – for falling)

Q_{st} is the steady-state discharge as obtained from the available steady-state rating curve.

S_0 is the energy slope

C is wave velocity (celerity)

Q_{st} is the steady-state discharge obtained by establishing a rating curve as a median curve through the uncorrected stage-discharge observations or using those observations for which the rate of change of stage had been negligible.

Care has to be taken to see that a sufficient number of gauging on rising and falling limbs are available, if the unsteady state observations are considered while establishing the steady-state rating curve. The expression $\frac{1}{CS_0}$ is expressed in the form of the parabolic equation as:

$$\frac{1}{CS_0} = a_3 + b_3 h_t + c_3 h_t^2 \qquad \text{Equation 7.6}$$

and $h_t > h_{min}$

a3, b3, c3 parameters of the parabolic equation

h_{min} is the lowest water level below which the correction is not to be applied. A maximum value of factor ($\frac{1}{CS_0}$) is also specified so that unacceptably high value can be avoided from taking part in the fitting of the parabola.

The rate of change of stage for time (dh_t/dt) at time t can be obtained from the stage time series as:

$$\frac{dh_t}{dt} = \frac{(h_{t+1} - h_{t-1})}{2\Delta t} \qquad \text{Equation 7.7}$$

where:

Δt time interval between two successive observations. If h_{t+1} or h_{t-1} does not exist, its value is replaced by h_t and the denominator by Δt.

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

1. Measured discharge is plotted against the stage and beside each plotted point the value of dh/dt for the measurement (+ or -) is noted
2. A trial Q_{st} rating curve representing the steady flow condition where dh/dt equals zero is fitted to the plotted discharge measurements.
3. A steady-state discharge Q_{st} is then estimated from the curve for each measured discharge and Q_t , Q_{st} and dh/dt are together used in the to compute corresponding values of the adjustment factor $\frac{1}{CS_0}$ from the above Equation 7.7
4. Computed values of $\frac{1}{CS_0}$ are then plotted against the stage and a smooth (parabolic) curve is fitted to the plotted points

7.3.4 Rating curve with constant fall backwater correction

The unique relationship between stage and discharge at the gauging station is not maintained if the gauging site is affected by backwater. Backwater is an important consideration in streamflow gauging site selection and sites having backwater effects should be avoided whenever possible. However, many existing stations in India are subject to variable backwater effects and require special methods of discharge determination.

When the backwater from the downstream control results in lowering the water surface slope, a smaller discharge passes through the gauging station for the same stage. On the other hand, if the surface slope increases, as in the case of sudden drawdown through a regulator downstream, a greater discharge passes for the same stage. The presence of backwater does not allow the use of a simple unique rating curve. Variable backwater causes a variable energy slope for the same stage. Discharge at such a station does not only depend on the stages, but also the fall (water surface slope). It is determined by taking measurements of stages at the main station and an auxiliary station downstream at the same time. It is called Stage -Fall-discharge relation.

An initial plotting of the stage-discharge relationship (either manual or by computer) with values of fall against each observation will show whether the relationship is affected by variable slope and whether this occurs at all stages or is affected only when the fall reduces below a particular value. In the absence of any channel control, the discharge would be affected by variable fall at all times and the correction is applied by the constant fall method. When the discharge is affected only when the fall reduces below a given value, the normal (or limiting) fall method is used.

The stage-discharge transformation with constant fall-back water method is carried out by the following procedure:

Step1: A constant fall of energy slope $\mathbf{F_r}$ is selected from the list of falls as most frequently observed fall or as an average fall.

Step 2: A rating Curve between stage h and the reference Discharge $\boldsymbol{Q_r}$ is fitted directly by estimating

$$\boldsymbol{Q_r} = \boldsymbol{Q_{measured}} \left({}^{\boldsymbol{F_r}}/_{\boldsymbol{F_m}} \right)^{\boldsymbol{p}} \qquad \text{Equation 7.8}$$

Exponent p is usually optimised between 0.4 to 0.6, to minimise the standard error.

Step 3: The discharge at any time corresponding to a stage h and the fall $\mathbf{F_m}$ is then calculated by first obtaining Q_r from the above relationship and the then calculating the discharge Q as

$$\boldsymbol{Q} = \boldsymbol{Q_r} \left(\frac{\boldsymbol{F_m}}{\boldsymbol{F_r}} \right)^{\boldsymbol{p}} \qquad \text{Equation 7.9}$$

As a special case, the constant fall method becomes the unit fall method when F_r is equal to unity.

7.3.5 Rating curve with normal fall backwater correction

When the discharge is affected only when the fall reduces below a given value, the normal (or limiting) fall method is used.

The computerised procedure considerably simplifies computation and is as follows:

1. Compute the backwater-free rating curve using selected current meter gauging (the Q_r -h relationship).
2. Using values of Q_r derived from (1) and Fr derived from:

$$F_r = F_m \left({}^{Q_r}/_{Q_{measured}} \right)^{1/p} \qquad \text{Equation 7.10}$$

3. A parabola is fitted to the reference fall with the stage (h) as:

$$F_r = a + b\,h + c\,h^2 \qquad \text{Equation 7.11}$$

4. The parameter p is optimised between 0.4 and 0.6, minimising the standard error

5. The discharge at any time, corresponding to the measured stage h and fall F_m , is then calculated by:

- obtaining $\mathbf{F_r}$ for the observed h from the parabolic relation between h and $\mathbf{F_r}$
- obtaining $\boldsymbol{Q_r}$ from the backwater free relationship established between h and $\boldsymbol{Q_r}$
- then calculating discharge corresponding to measured stage h as:

$$\boldsymbol{Q = Q_r \left({F_m}/{F_r} \right)^p} \qquad \text{Equation 7.12}$$

8 VALIDATION OF DISCHARGE DATA

8.1 General

After transformation of the stage to discharge data, secondary validation can be done for discharge data. The suspect (based on the assessment of stage), values should be corrected, or missing values need to be flagged and are required to be reviewed, corrected, or inserted.

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. Errors may show up during discharge validation. Validation flags which have been inserted in the validation of the stage record are transferred to the discharge time series. These include the data quality flags of 'good', 'doubtful' and 'poor' and the origin flags of 'original', 'corrected' and 'completed'. This transfer of flags is necessary so that stage values recognized as doubtful or poor can be corrected as discharge. The wrong stage-discharge relationship can give rise to discharge errors, causing discontinuities in the discharge series.

The secondary validation is carried out through comparison of the time series with neighbouring stations whereas preliminary validation of a single series is carried out through comparison of the values against data limits and expected hydrological behaviour.

8.2 Single Station Validation of Discharge Data

Single station validation is done by inspecting the data in tabular and graphical form. This will illustrate the status of the data concerning quality and origin, which may have been inserted during the stage validation or identified during discharge validation. Validation emphasises on identifying errors and it is followed by investigation for correcting and completing the series.

8.2.1 Validation against data limits

Data is checked numerically against absolute boundaries, relative boundaries and acceptable rates of change. The individual values in the time series are flagged for additional inspection.

- **Absolute boundaries**

Values which exceed a maximum specified value, or values that are smaller than a specified minimum value provided by the user may be flagged. The specified values may be the absolute values obtained from the historical series. The objective is to screen out spurious extremes, but care must be taken not to remove or correct true extreme values as these may be the most important values in the series.

- **Relative boundaries**

A large number of values may be flagged by specifying boundaries with departures from the mean of the series (Q_{Mean}) by some multiple of the standard deviation ($\boldsymbol{S_x}$), i.e.

Upper boundary $\boldsymbol{Q_u = Q_{mean} + \alpha S_x}$ Equation 8.1

Lower boundary $\boldsymbol{Q_l = Q_{mean} - \beta S_x}$ Equation 8.2

While Q_{mean} and the standard deviation S_x can be computed easily by WIMS or in a spreadsheet calculation, the multipliers α and β are inserted by the user with values between 2 and 3 (typically α =3 and β =2) and they can be calculated for seasonal or even monthly flows. This would help in creating a range of upper and lower boundaries that can be anticipated for various times of the year. The goal is to set limits which can allow screening of a manageable number of outliers for inspection while providing reasonable confidence that all suspect values are flagged. This test is normally only used for aggregated data of a month or longer periods, while the bounds should be calculated on the basis of all historic data that had previously been validated.

- **Rates of change**

Values will be flagged where the difference between successive observations exceeds a value specified by the user. Acceptable rates of rising and falling may be specified separately. Generally, the allowable rates of rising will be greater than the allowable rates of falling. This is a convenient way to identify possible inconsistencies, particularly since a listing of only those data points which are beyond certain acceptable boundaries can be obtained.

8.3 Graphical Validation

Graphical inspection of the plot of a time series provides a very rapid and effective technique for detecting anomalies. Graphical inspections are the most widely applied validation procedure, and they are carried out regularly for all discharge data sets.

The discharge may be displayed alone or with the associated stage measurement (Figure 8.1). In this example, the plot of 6 months reveals discontinuities which might have appeared between successive monthly updates of the data series.

The discharge plots may be displayed in the observed units or the values may be log-transformed where the data cover several orders of magnitude. This enables values near the maximum and minimum to be displayed with the same level of precision. Log-transformation is also a useful means of identifying anomalies in dry season recessions. Whereas the exponential decay of flow based on releases from natural storage are curved in natural units, they show as straight lines in log-transformed data.

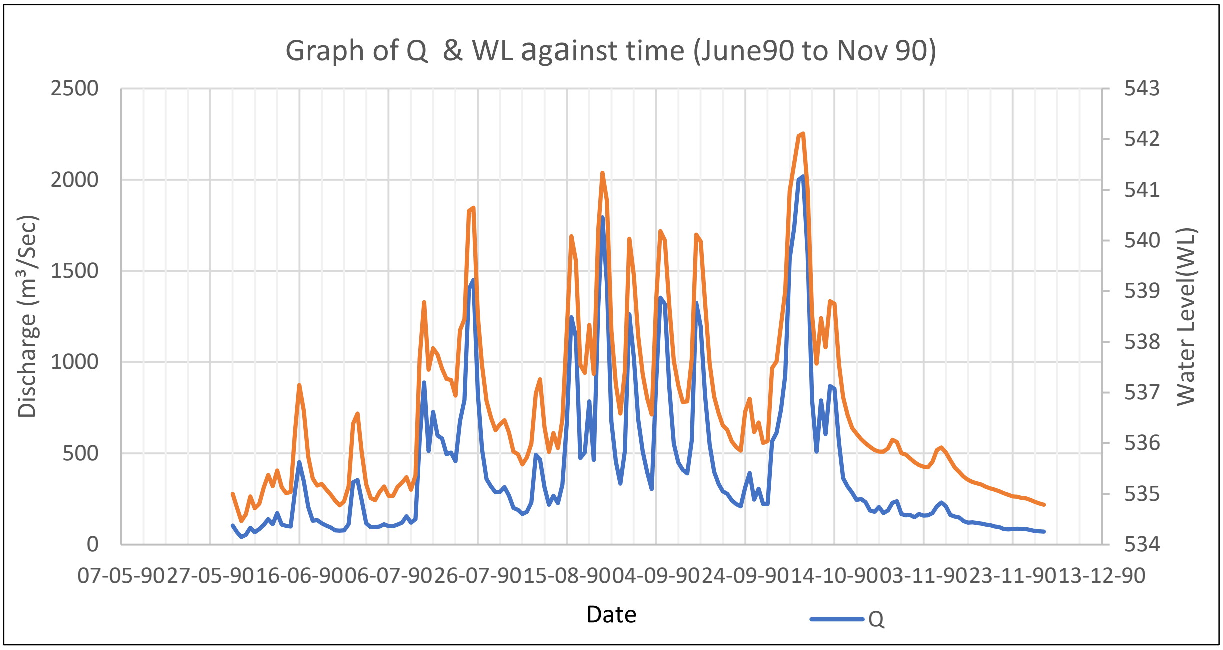


Figure 8.1: Q(t) and WL(t) Plots for Six Consecutive Months at Jagdalpur Station on the Godavari

The graphical display may also show the absolute and relative limits. The plots provide a better guide to the likely reliability of such observations, as compared to tabulations.

The main purpose of the graphical inspection is to identify any abrupt discontinuities in the data or the existence of positive or negative 'spikes' which do not conform to the expected hydrological behaviour. It is very convenient to apply this test graphically wherein the rate of change of flow together with the flow values are plotted against the expected limits of the rate of rising and falling in the flows.

Some of the reasons that may result in wrong values of discharge are:

- The use of the wrong stage-discharge relationship.
- The use of incorrect units (Figure 8.2). Please note that in this example, the discharge has been plotted at a logarithmic scale
- Abrupt discontinuity in a recession (Figure 8.3).
- Isolated highs and lows from unknown source (Figure 8.4) that may be due to recorder malfunction with stage readings.

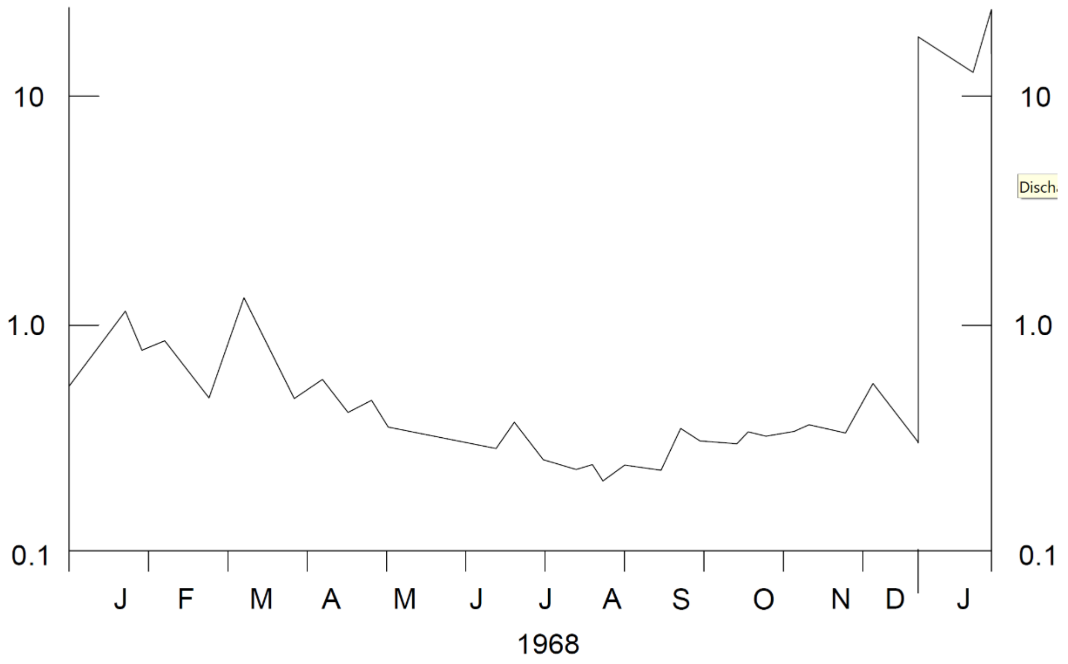


Figure 8.2: Use of Incorrect Unit

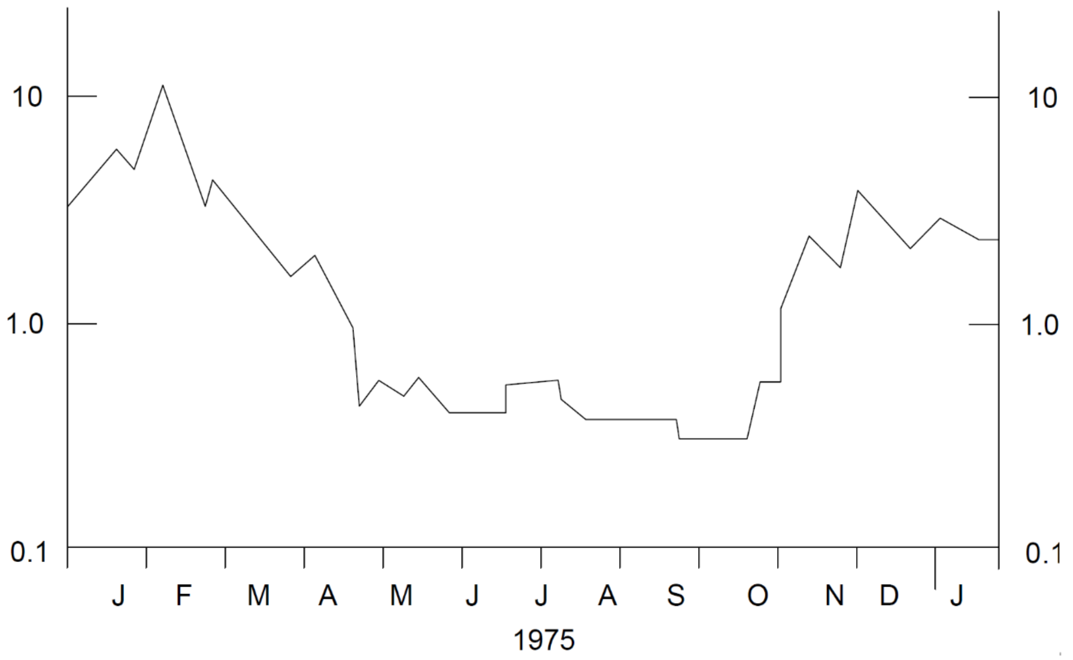


Figure 8.3: Unrealistic Recession

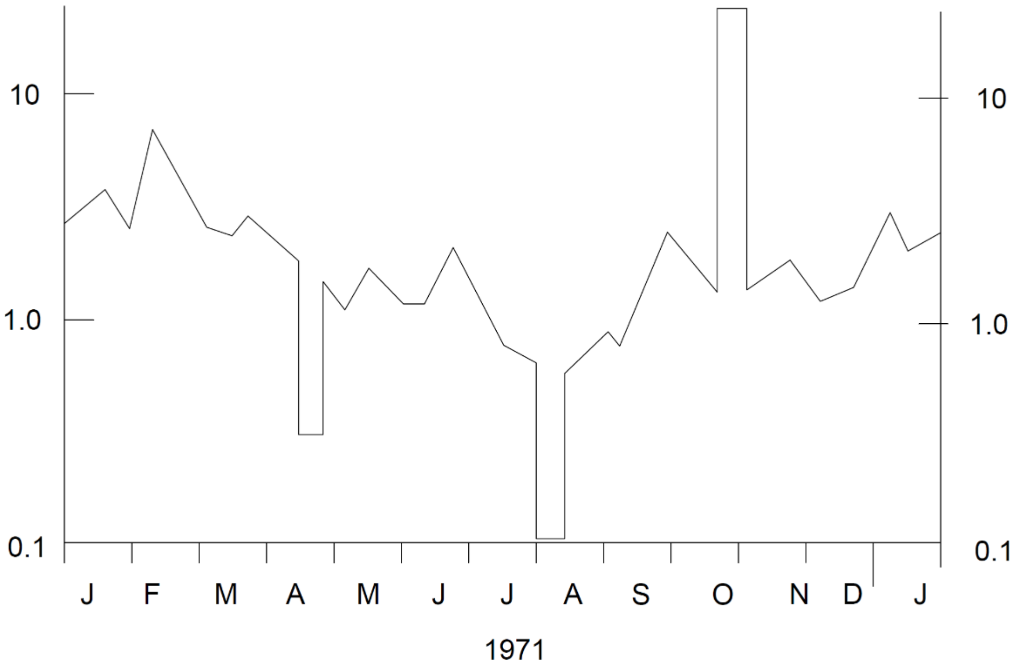


Figure 8.4: Isolated Highs and Lows

8.4 Validation of Data for Regulated Rivers

The problems of validating regulated rivers for stage data have already been mentioned and should also be considered while validating discharge data. Natural flow series are not common in India; most large rivers are regulated by the operation of reservoirs to a greater or lesser extent. The natural pattern is disrupted by reservoir releases which may have abrupt onset and termination, combined with multiple abstractions and return flows. These influences are most clearly seen in low to medium flows. In some rivers the hydrograph appears entirely artificial; although very high flows may still have a pattern similar to natural flows. The officers performing validation should be aware of the principal anthropogenic (man-made) influences within the basin, the location of those influences, their magnitude, their frequency and seasonal timing, to provide a better basis for identifying values or sequences of values which are suspect. A process of naturalization of flows is related to estimating and removing the effects of regulation, thus producing flows that would have happened under natural conditions (i.e., without reservoirs and other river basin infrastructure or abstractions). Natural flows are an essential input into water accounting and river basin planning studies, since they show the seasonal and inter-annual availability of water at all critical reaches of a river basin.

8.5 Multiple Station Validation

8.5.1 Comparison plots

The simplest and often the most helpful means of identifying anomalies between stations are in the plotting of comparative time series. The WIMS software permits the plotting of multiple time series for a given period in one graph. There will, of course, be differences in the plots depending on the contributing catchment area, differing rainfall over the basins and differing response to the rainfall. However, the differences should follow anticipated patterns.

The most helpful comparisons are between the sequential stations on the same river. The series may appear shifted relative to each other in time, to take into account different lag times due to the travel time in the channel.

Comparison of series may permit the acceptance of values which were earlier flagged as suspect because they fall outside of the warning ranges, when viewed as the stage or when validated as a single station. When two or more stations display the same behaviour there is strong evidence to suggest that the values are correct.

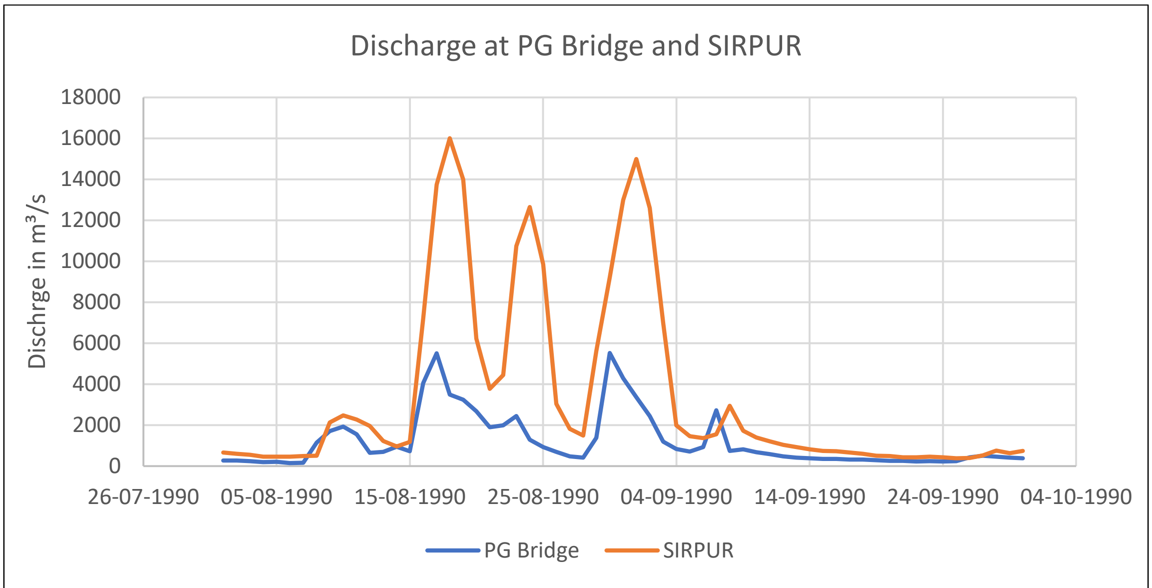


Figure 8.5 Plot of Discharge for Two stations on the Godavari (PG Bridge on Penganga at the upstream and Sirpur on Wardha at the downstream)

Comparison plots provides a simple means for identifying anomalies but not necessarily for correcting them. Correction may best be done through regression analysis or by transposing flows from the station without data issues.

8.6 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. The WIMS software provides a means for displaying residual series calculated on a steady state basis where the travel time between the control points is significantly below the length of the calculation time step which is used as a basis for comparison of the two or more series. Both the original time series and their residuals can be plotted in the same figure.

When the residual series are calculated using daily data, this typically requires that the upstream series be first subjected to hydrologic routing. More about hydrologic routing is available in subsequent chapters of this manual.

8.7 Comparison of Streamflow and Rainfall

This approach is a possibility, but its use should be subject to severe scrutiny. The reason why this approach should be the last option in general is due to the fact that relationship between rainfall and runoff being often sketchy at best. This explains why it is so difficult to calibrate rainfall runoff models and verify them for all the historical years available for verification. There are a lot of uncertainties involved in this relationship, due to the non-uniformity in temporal and spatial distributions of the rainfall, the variations of the ground (land use and land cover) and the antecedent moisture conditions.

A quick insight into the consistency of the data can be made by graphical and tabular comparison of historical areal rainfall and runoff. In particular, it is useful to track the value of runoff coefficient and see if it is consistent in wet seasons from year to year. The basin rainfall over an extended period such as a month or year should exceed the runoff (expressed in mm) over the same period. Tabular comparisons should be consistent with such physical understanding of the process. For example, an excess of runoff over rainfall either on an annual basis or for monthly periods during the monsoon will indicate an error. Also, it will cause suspicion if the ratio of runoff to rainfall exceeds the common threshold of 0.7.

Graphical comparison on a shorter time scale can be made by plotting rainfall and streamflow on the same axis.

In general, the occurrence of rainfall and its timing should be followed by the occurrence of runoff separated by a time lag. However, for a precise validation, this method should not be used due to imperfect assessment of areal rainfall, and also due to the variable proportion of rainfall that enters the river as discharge. The calculated runoff coefficients typically vary between 0.2 and 0.7, depending on the time of the year and the antecedent soil moisture conditions. Any value of runoff coefficient about or greater than 0.7, particularly over a period longer than a few hours to a few days, should be closely examined, since they are very rare.

9 CORRECTION AND COMPLETION OF DISCHARGE DATA

9.1 General

Data validation is the process to ensure that the final values stored in the database are the best possible representation of the true values at the measurement site at a given time or in a given interval of time. Validation recognizes that values observed or measured in the field are subject to errors which may be random, systematic or spurious. Incorrect and missing values will be replaced wherever possible by estimated values based on interpolation or other techniques that may rely on the observations at the same station or at neighbouring stations. The process of filling in the missing values is generally referred to as ‘completion’.

It must be recognised that values estimated from other gauges are inherently less reliable than the values properly measured. Doubtful original values may, therefore, be given the benefit of the doubt and retained in the record with a flag. Where no suitable neighbouring observations or stations are available, missing values may be left as ‘missing’ and incorrect values may be set to ‘missing’. Procedures for correction and completion depend on the type of error, its duration, and the availability of suitable source records on the basis of which new values are to be estimated.

9.2 Completion using Another Record from the Same Station

All streamflow stations equipped with autographic or digital recorders have manual observations made as a back-up. Where there is an equipment failure, the observer’s manual record is used to complete the instrumental record. This is normally done for the water level measurements, and the discharge is then estimated by converting the water level to flow using the appropriate rating curve.

9.3 Interpolating Discharge Gaps of Short Durations

Unlike rainfall, streamflow shows strong serial correlation; the value recorded on one day for a large catchment is closely related to the values observed on the previous and following days, especially during periods of low flow or recession.

Where gaps in the record are short, during periods of low flow (say, gaps less than 2 days), it may be acceptable to use linear (or non-linear) interpolation between the last value before the gap and the first value after it. To confirm that this is acceptable, a graphical display of the hydrograph at the station and one or more neighbouring stations is inspected to ensure that that there are no discontinuities in the flow sequence over the gap.

9.4 Interpolating Gaps During the Recession

During periods of recession when the flow is dependent on surface and sub-surface storage rather than rainfall, the flow exhibits a pattern of exponential decay. It appears as a curved trace on a simple plot of discharge versus time (shown in Figure 9.1) but will form a straight line on a logarithmic plot. During long recession periods, interpolation between the logarithmically transformed points before and after the gap will result in a more realistic recession than simple linear interpolation.

General recession equation can be used to determine the missing values of flow Q_t at times t is

$$Q_t = Q_{t_0}\, e^{(-kt)} \qquad \text{Equation 9.1}$$

Where time t is measured from the starting time t_0. which is arbitrarily selected by the user who is engaged in the exercise of in-filling the missing data. This approach is unfortunately not automated within Excel, which implies more work to fit the decaying recession constant k, but similar results can be obtained by using the power or logarithmic functional fit, as demonstrated through an example below.

Example 9-1

Daily data of a station with some gaps are plotted and the gap has been filled through the above method. The first two points have been filled with linear interpolation. The remaining 5 points have been filled by using the exponential fit that was obtained with fitting the exponential line between the following data points [1, 261.95]; [2, 261.488],[8,260.24] and [9, 259.987]. The fitted equation is

$$Y = 262.01388\ X^{-0.00343}$$

The missing values were calculated with the X coordinates for missing days 3, 4, 5, 6, and 7.

Date	Observed	Calculated
6/12/2004	260.175	260.175
6/13/2004	260.175	260.175
6/14/2004	260.215	260.215
6/15/2004	260.130	260.130
6/16/2004	260.105	260.105
6/17/2004	260.035	260.035
6/18/2004		259.951
6/19/2004		259.867
6/20/2004	259.784	259.784

Date	Observed	Calculated
6/21/2004	259.700	259.700
6/22/2004	259.900	259.900
6/23/2004	259.870	259.870
6/24/2004	259.830	259.830
6/25/2004	260.300	260.300
6/26/2004	260.170	260.170
6/27/2004	260.100	260.100
6/28/2004	261.950	261.950
6/29/2004	261.688	261.488
6/30/2004		261.028
7/1/2004		260.771
7/2/2004		260.571
7/3/2004		260.409
7/4/2004		260.271
7/5/2004	260.120	260.120
7/6/2004	260.050	260.050
7/7/2004	260.040	260.040
7/8/2004	260.020	260.020
7/9/2004	260.020	260.020
7/10/2004	260.020	260.020
7/11/2004	260.020	260.020
7/12/2004	260.020	260.020
7/13/2004	260.050	260.050
7/14/2004	260.050	260.050
7/15/2004	260.050	260.050

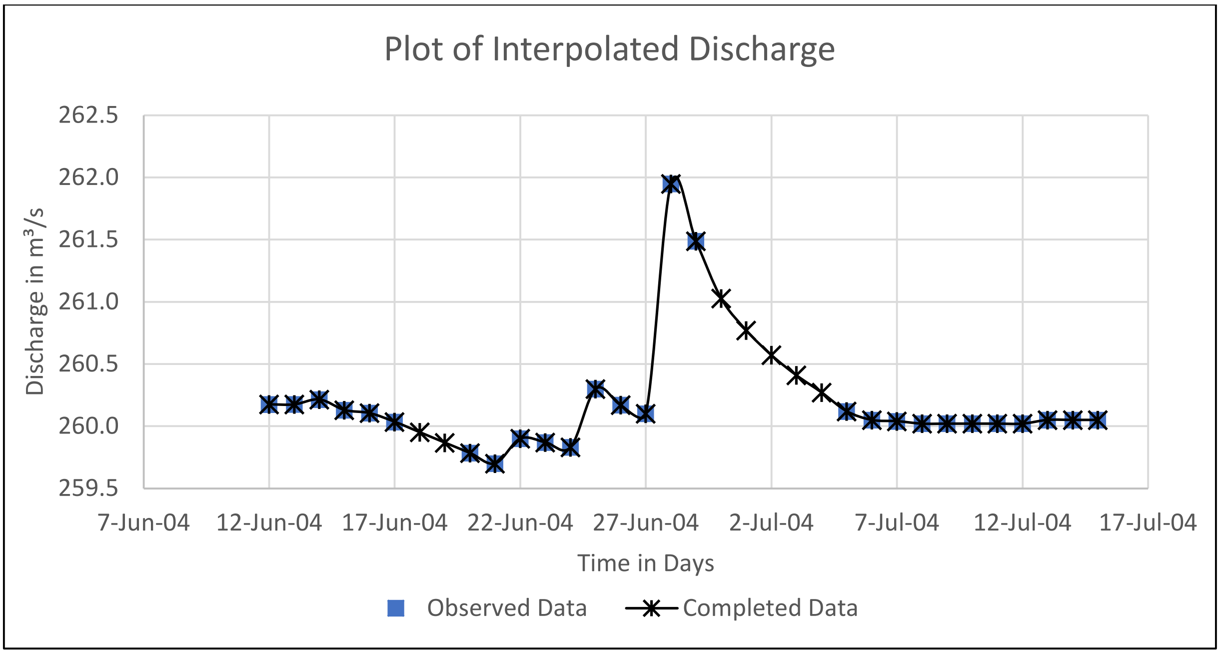


Figure 9.1 Gap-Filling During the Recession Using Interpolation

9.5 Interpolation Using Regression

During the periods of variable flow or in case of longer gaps, simple interpolation should not be used and the appropriate regression equation may be applied to fill in the missing data. If there are suitable stations on the same river or in a neighbouring catchment, regression relations may be developed for monthly, weekly or daily series based on the periods for which historical data are available at both stations.

Since the statistical dependence between the stations may change with the seasons, the option of fitting and applying the fitted equations should be limited only to a particular period of the year using daily, weekly or monthly values for limited periods within either wet or the dry season. Where two or more such relations are applied to fill up a single gap, the resulting interpolated hydrograph is inspected to ensure that there is no serious discontinuity at the junction between the periods of application.

For shorter daily time steps, a time shift is applied to a discharge series to allow for the average lag time between a station represented by the dependent series and the independent station series where they are on the same river. Otherwise, the average time shift is considered as the difference in travel time if they are on the neighbouring catchments. The time shift may be estimated by the user based on the historical propagation of peak flows, or based on the physical properties of the channel. Travel time can also be computed using calibrated HEC-RAS or other similar models.

For a user computed shift (t_1) between an upstream station X and a downstream one Y, spaced at a distance of s km, the following formula may be applied, given the average flow velocity of u (m/sec) which can be obtained from the gauged records):

$$t_1 = \frac{86.4s}{u}(\ days) \qquad \text{Equation 9.2}$$

Excel and other modern software with regression modules provide a range of regression relations as follows:

- Single independent variable
    - simple linear
    - polynomial
    - logarithmic
    - power
    - exponential
    - hyperbolic
- Multiple independent variables
    - linear multiple variables

To establish the relationship between X and Y and to view its functional form, a scatter graph should first be plotted. In general, a simple linear relationship should be tried first, then a polynomial, although using order 3 or higher is not recommended.

Multiple regression may be applied, for example, between:

- downstream station and two or more upstream stations
- downstream station and upstream stations and rainfall
- upstream station and downstream station and intervening tributaries

Irrespective of the scatter plot, regression analysis will produce a functional relationship, but if the relationship is poor, it should not be used to in-fill the missing values. The suggested acceptance criteria are that a correlation with the value of r equal to 0.85 may generally be considered the lower limit for acceptance. However, reference should also be made to the standard error of the estimate. Where no acceptable relationship is found, the missing values should be left 'missing' or an alternative method of in-filling may be used.

Application of regression analysis may also produce a discontinuity between the observed flows before and after the gap and the in-filled values due to error in the relationship. The in-filled hydrograph plot should be inspected for such discontinuities and suitable adjustment applied.

Regression analysis provides the ability to summarize a collection of sampled data by fitting it to a model that will accurately describe the data. This method fits a set of data points to a function, describing the dependent variable Y as a linear combination of one or more functions of the independent variable(s) X. They are broadly categorised into two classes: linear and non-linear regression. Linear regression models with more than one independent variable are referred to as multiple linear models, as opposed to simple linear models with only one independent variable.

9.5.1 Limitations

The variance of the regression estimate is always biased downward since regression estimates lie on the regression line, while the actual data are scattered about the regression line. Users should beware of two major pitfalls of using regression: (i) ability to produce negative values during low flow periods, and (ii) a propensity to introduce a positive artificial lower bound in the data if the regression constant is positive.

9.5.2 Effect of outliers in data

A single or individual observation that is substantially different from all other observations can make a large difference in the results of regression analysis. An outlier can exert undue influence on the coefficients. An outlier may either indicate a sample peculiarity, or it may indicate a data entry error or other problem.

9.5.3 Collinearity

The term collinearity implies that two variables are near perfect linear combinations of one another. Such relationships are never found between two or more flow monitoring stations, implying that if the R^2 coefficient equals 1, this is most likely due to the one of the flow records being manipulated manually based on the data in another data set.

When more than two variables are involved, this situation is called multicollinearity, and it should be treated as suspect, implying that the data at various stations are not genuine.

9.5.3.1 Assumptions

There are a few critical assumptions about the data set that must be judged to be true before it is decided to proceed with a regression analysis:

- The variables must be truly independent (using a Chi-square test).
- The data must not have different error variances across the range of values (i.e., heteroskedasticity).
- The error terms of each variable must be uncorrelated. If not, it means the variables are serially correlated.

9.6 Water Balance and Flow Routing Methods

Regression analysis may be used to estimate long periods of missing values or to extend a record. However, in-filling missing values by regression does not ensure water balance between the neighbouring stations. Thus, the application may result in significantly less volume of flow at a downstream station than that at the upstream station, which is uncommon when there are no abstractions or diversions in-between. If the balance conflicts with common sense, the functional relationship should be reviewed and if necessary, rejected. Alternatively, to achieve a satisfactory balance between stations, flow routing methods may be applied.

The mass-balance equation for a system state that the difference between the input and output is equal to the rate of change in storage. In flow routing using Muskingum method (most commonly used in India), the two routing parameters K and X are determined from measured hydrographs at upstream and downstream stations and applied to route the flow from upstream to a missing downstream station. Inflows and abstractions from the intervening reach can be incorporated to achieve a water balance. Flow routing is usually applied to floods but can be extended for use in low flows. An example of the classical Muskingum method of flow routing follows. More sophisticated alternatives are also available, which use a third parameter.

9.6.1.1 Muskingum method of flow routing

In the short term, inflow into a river reach is not equal to its outflow, since there is a constant change in the channel storage of the river reach. This is mathematically expressed as:

$$I - Q = \frac{dS}{dt} \qquad \text{Equation 9.3}$$

Storage is therefore a function of both inflow I and outflow Q, linked together by the channel storage change over the time step t, as per equation 9.3. This relationship can also be expressed by using the storage coefficient k and the weight factor X, as defined in equation 9.4 below.

$$S = k[XI - (1 - X)Q] \qquad \text{Equation 9.4}$$

where I = inflow, Q = outflow, S = storage, X = weighting factor, K = storage coefficient.

The values of storage at time t and t+1 can be written, respectively, as

$$S_t = K[XI_t + (1 - X)Q_t] \qquad \text{Equation 9.5}$$

$$\boldsymbol{S_{t+1} = K[XI_{t+1} + (1 - X)Q_{t+1}]} \qquad \text{Equation 9.6}$$

Using the equations (9.5) and (9.6), the change in the storage over time interval Δt is

$$S_{t+1} - S_t = K[XI_{t+1} + (1 - X)Q_{t+1}] - K[XI_t + (1 - X)Q_t] \qquad \text{Equation 9.7}$$

Considering that the variation of inflow and outflow over the interval is approximately linear, the change in storage can also be expressed as:

$$S_{t+1} - S_t = \frac{(I_{t+1}+I_t)}{2}\Delta t - \frac{(Q_{t+1}+Q_t)}{2}\Delta t$$ Equation 9.8

Coupling of the equations in finite difference form leads to

$$Q_{t+1} = [C_0 I_{t+1} + C_1 I_1 + C_2 Q_t]$$ Equation 9.9

where C_0, C_1, and C_2 are routing coefficients in terms of Δt, K, and X as follows:

$$C_0 = \frac{\left(\frac{\Delta t}{k}\right)-2X}{2(1-X)+\left(\frac{\Delta t}{k}\right)}$$ Equation 9.10

$$C_1 = \frac{\left(\frac{\Delta t}{k}\right)+2X}{2(1-X)+\left(\frac{\Delta t}{k}\right)}$$ Equation 9.11

$$C_2 = \frac{2(1-X)-\left(\frac{\Delta t}{k}\right)}{2(1-X)+\left(\frac{\Delta t}{k}\right)}$$ Equation 9.12

The above three coefficients sum up to 1, i.e., C_0+C_1+C_2=1. The routing coefficients can be interpreted as weighting coefficients.

If the observed inflow and outflow hydrographs are available for a river reach, the values of K and X can be determined. Assuming various values of X and using successive known values of the inflow and outflow, values of K can be derived as

$$K = \frac{0.5\,\Delta t[(I_{t+1}+I_t)-(Q_{t+1}+Q_t)]}{X(I_{t+1}-I_t)+(1-X)(Q_{t+1}-Q_t)}$$ Equation 9.13

The procedure for defining the values of K and X has been documented in most textbooks on hydrology. The computed values of the numerator and denominator are plotted for each time interval. Having the numerator on the vertical axis and the denominator on the horizontal axis usually produces a graph in the form of a loop. The value of X that produces a loop closest to a single line is taken to be the correct value for the reach. According to the equation mentioned above, K represents the slope of the line. Since K is the time required for the incremental flood wave to traverse the reach, its value may also be estimated as the observed time of travel of the peak flow through the reach. An example of this method has been included underneath.

The principal shortcoming of the Muskingum method is that the value of travel time K does not change with the changes of flow. Hence, for modelling single events where K has been determined based on the average flow of the event, this method may still provide acceptable results. However, when considering longer continuous periods that include multiple changes between the low and high flow range, the method no longer works properly. It becomes difficult (if not impossible) to reproduce the observed historical flows. Various methods have been introduced to improve the estimates of K dynamically as a function of flow, but most of them are not practical, since they require information that is difficult to obtain, such as the Manning's n coefficient, the slope and the channel width. All of this information has to be averaged over a river reach, in an effort to convert the mean channel flow into velocity and obtain the travel time on the basis of the known length of the reach. However, the channel slope, Manning's n and the channel width would normally vary widely along a 50 km river reach, which renders these methods impractical. One variant of the Muskingum method that solves this situation in an elegant manner has been incorporated in the SSARR model (US Corps of Engineers, 2021), based on an old publication of the procedure known as the William's equation (Williams, 1969). The only information required to determine the routing coefficients using this method that change dynamically with flow, is the travel time vs flow relationship for each river reach. This procedure is explained in more detail in the following section.

9.6.2 The SAARR routing method

The first significant application of the Williams routing equation was originally developed by the US Corps of engineers, the **S**tream **S**ynthesis **A**nd **R**eservoir **R**outing (SSARR) model offers probably the most user-friendly way to conduct hydrologic river routing with dynamic adjustments of the routing coefficients. It is based on routing coefficients which vary with channel flow, and the only input data requirement in addition to inflows at the upstream of the main stream and the tributaries are the travel time vs flow estimates, usually given in tabular form as shown previously in Figure 3.14. A major advantage of this model is that it does not need any channel geometry as input data, nor does it require the Manning's coefficients. Once the travel time vs flow relationship is available, the calibration consists of deciding how many sequential phases a given river reach should be divided into. This is conducted using repeated simulation trials until the observed downstream hydrograph closely matches the simulated channel outflow. As with the other river routing methods, the governing equation is related to channel storage change over a time step, which is a function of average inflow and outflow:

$$\frac{I_{t-1}+I_t}{2} - \frac{O_{t-1}+O_t}{2} = \frac{\Delta S}{t} \qquad \text{Equation 9.14}$$

By subtracting O_{t-1} from both the sides of the above equation, multiplying by $t/(O_t - O_{t-1})$ and by letting $\Delta S/(O_t - O_{t-1}) = TS$, the above equation becomes:

$$O_t = \frac{\left[\frac{I_{t-1}+I_t}{2} - O_{t-1}\right]\cdot t}{TS+\frac{t}{2}} + O_{t-1} \qquad \text{Equation 9.15}$$

where the term TS represents the average travel time along a river reach for a given flow condition, evaluated either by reading from the TS vs Q table or by using a functional form of the travel time vs flow curve as:

$$TS = \frac{Kts}{\left(\frac{O_{t-1}+O_t}{2}\right)^n} \qquad \text{Equation 9.16}$$

The routing coefficients *Kts* and *n* must previously be determined by finding the best fit curve for a given set of the available *(TS, Q)* coordinates. Alternatively, *TS* can be determined for any given flow rate by linear interpolation from a table of *(TS, Q)* points. In the above definition of *TS*, the base of the denominator:

$$\frac{O_{t-1}+O_t}{2} \qquad \text{Equation 9.17}$$

which is raised to the power exponent *n,* represents the estimate of the average outflow from a given reach during the time step *t*. For time steps sufficiently small, the variations of flow are also small, so it is common to assume $O_{t-1} = O_t$ in the first approximation. In his original publication by Williams (1969), the travel time along a reach is determined on the basis of the updated outflow O_t from the reach at the end of the current time step t, which better represents updated conditions in the basin.

The model typically conducts two to three iterations by updating O_t and recalculating the travel *Ts* time by using the updated coefficients before it converges to the final solution. Expression (6) can also be converted to the following form:

$$O_t = \frac{t}{2T_s+t} I_{t-1} + \frac{t}{2T_s+t} I_t + \frac{T_s - t/2}{T_s + t/2} O_{t-1} \qquad \text{Equation 9.18}$$

The above form is identical to the well-known Muskingum linear routing form:

$$O_t = C_1 I_{t-1} + C_2 I_t + C_3 O_{t-1} \qquad \text{Equation 9.19}$$

It can be noted that the SSARR routing coefficients listed in equation 9.18 sum up to 1.0 (i.e., $C_1 + C_2 + C_3 = 1$), which is also the condition for the Muskingum routing coefficients. In other words, the SSARR routing method uses identical formula as does the Muskingum routing procedure, except that the values of the routing coefficients *Ci* are determined in a different way, which has some obvious advantages:

a) The only required information for the values of routing coefficients is the time of travel vs flow relationship for a given river reach and the length of the calculation time step. No other data related to the channel geometry, gradient or roughness are required.

b) The values of routing coefficients undergo dynamic adjustments as the modeling moves through different flow regimes between dry seasons and wet seasons. This is a much more elegant and precise way than in the case of using the classical Muskingum method, which is most frequently used with fixed coefficients.

Implementations of the SSARR method may rely on different estimates of the average channel flow during a given time step. Input data requirements include time of travel versus flow table for a river reach, where time of travel is given in hours while flows are given in m^3/s. One condition that should be satisfied for a successful application of this channel routing technique is that the calculation time step is selected such that the travel time along the reach be at least more than twice the length of the calculation time step, i.e., $Ts \geq t/2$. If this condition is not satisfied, the terms that multiply I_{t-1} and I_t becomes greater than 0.5, and the mass conservation rule which requires that the sum of all three coefficients be equal to 1 can no longer be maintained. Similar conditions exist in the classical Muskingum approach.

Since the routed flows are not precisely known in advance, the SSARR routing method is iterative, requiring recalculation of the routing coefficients once the routed flows O_t have been set by the procedure. This makes the procedure more difficult to demonstrate on a numerical example. For examples of this procedure, readers can refer to the WEB.BM User Manual, which can be downloaded from www.riverbasinmanagement.com after logging in (the application is freely available). The numerical example presented in the last section of the WEB.BM User Manual shows the working of the SSARR routing procedure as part of solving a reservoir optimization problem, i.e., the SSARR routines have been incorporated into the model as a constraint for overall river basin optimization.

Example 9-2

The application of the Muskingum method is demonstrated with the help of the following example.

An inflow hydrograph to a channel is shown in Col. 2 of Table 9-1. Using the Muskingum method, this hydrograph is to be routed with K=2 days and X=0.1, to calculate an outflow hydrograph. The baseflow may be assumed as 352 m^3/s.

Table 9-1 shows the computation of outflow using the Muskingum method. For the given values of K=2 days and X=0.1 (derived from the historical record of inflow and outflow sets), C_0=0.1304, C_1=0.3044, and C_2=0.5652. It is noted that the sum of these routing coefficients is equal to 1.0. Taking Q_1 = 352 m^3 /s which is the baseflow, partial flows shown in columns 3 through col 5 of Table 9-1 can be computed and summed up to obtain total outflow given in Col. 6.

The principal difficulty in applying routing methods to infill the missing values is in the assessment of ungauged lateral inflows and outflows, and the method should not be used where the values are large and variable.

Table 9-1: Muskingum Routing

Time	Inflow	Partial flow			Total outflow
(day)	(m^3/s)		(m^3/s)		
		C_0I_2	C_1I_1	C_2Q_1	
Col 1	Col 2	Col 3	Col 4	Col 5	Col 6
0	352				352
1	587	76.6	107.1	199.0	382.7
2	1353	176.5	178.7	216.3	571.4
3	2725	355.4	411.8	323.0	1090.2
4	4408.5	575.0	829.3	616.2	2020.6
5	5987	780.9	1341.7	1142.1	3264.7
6	6704	874.4	1822.1	1845.3	4541.8
7	6951	906.7	2040.3	2567.1	5514.1
8	6839	892.0	2115.5	3116.7	6124.2
9	6207	809.6	2081.4	3461.5	6352.6
10	5346	697.3	1889.1	3590.6	6177.0
11	4560	594.8	1627.0	3491.3	5713.2
12	3861.5	503.7	1387.8	3229.2	5120.7
13	3007	392.2	1175.2	2894.3	4461.8
14	2357.5	307.5	915.2	2521.9	3744.5
15	1779	232.0	717.5	2116.5	3066.0
16	1405	183.3	541.4	1733.0	2457.7
17	1123	146.5	427.6	1389.1	1963.2
18	952.5	124.2	341.8	1109.6	1575.7
19	730	95.2	289.9	890.6	1275.7
20	605	78.9	222.2	721.0	1022.1

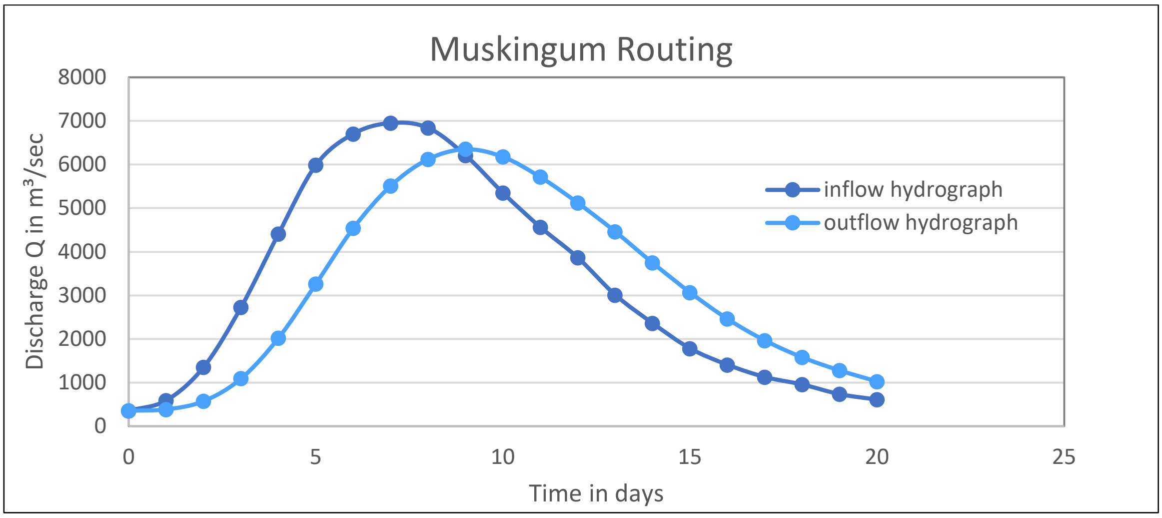


Figure 9.2: Example of Muskingum Routing

9.7 Using Rainfall-runoff Relationship

Hydrologic routing may be useful for data completion, especially during dry seasons when there is not much lateral flow contribution between the two stations on the same stream. However, in the absence of upstream flow records, if a sufficient number of rain gauges are present in the upstream catchment to calculate the average rainfall, a rainfall-runoff relationship may be developed for the monsoon season and used to infill missing records or to extend the record.

Much like regression analysis, the quality of the record generated through rainfall-runoff modelling cannot be guaranteed. It is limited by the reliability of the rainfall and flow records during the calibration period. Generally, a relationship having correlation coefficient of less than 0.70 between the observed and simulated flows for the calibration period should not be applied for infilling the missing data. Such records should be left missing. Hence, these methods are not recommended for data validation and calculation of missing values of discharges, and they may only be applicable in the monsoon season, since dry season flows are not driven by rainfall. Those interested the rainfall-runoff modelling may explore the same from other standard text books on hydrology.

10 COMPILATION OF DISCHARGE DATA

10.1 General

Discharge compilation is the process by which discharge at its observed or recorded time interval and units is in-filled (completed) and transformed, either:

- to another time interval base; or,
- from one unit of measurement, especially from discharge (a rate of flow) to water volume or runoff (water depth over the catchment).

Computations for aggregation of data from one-time interval to another depends on the data type. For flows, then the aggregation is done by computing the arithmetic average of the individual constituent data values. For water volume, the constituent values are summed up to obtain the aggregated value.

Discharge compilation is required for validation, analysis and reporting. The compilation is carried out at the Divisional offices before validation if required, but the final compilation is carried out after correction and 'completion'.

10.2 Aggregation of Data into Longer Duration

Discharge and its corresponding water level are observed at different time intervals, but intervals are generally one day or less. Manual observation may be daily, hourly for part of the day during selected seasons, or sometimes at multiples of an hour. For automatic water level recorders, a continuous trace is produced, from which the hourly levels and corresponding discharges are extracted. In the digital water level recorders, the level is usually recorded at the hourly interval, though for some small basins the selected interval may be 15 or 30 minutes. Sub-hourly, hourly and sub-daily discharges, computed from these levels, are typically aggregated to daily mean. For example, the daily mean discharge (Q_d) is computed from hourly values (Q_i) by:

$$\boldsymbol{Q_d = \frac{1}{24}\sum_{i=1}^{24} Q_i} \qquad \text{Equation 10.1}$$

For a given day, the mean is normally calculated for hours commencing 01:00 and finishing 24:00. Sometimes, daily discharge averages are calculated over the day from 09:00 am to 08:00 am the next day (i.e., for hourly measurements, the average of observations from 09:00 hours to 08:00 hours), to enable direct comparison with daily rainfall.

Daily data are typically averaged over weekly, ten dailies, monthly, seasonal or annual time intervals. In general,

$$Q_{Nd} = \frac{1}{Nd}\sum_{i=1}^{Nd} Q_i \qquad \text{Equation 10.2}$$

where,

Q_{Nd} = the discharge for the duration of Nd days,

Q_i = the discharge of the i^{th} day within the duration of Nd days.

Time intervals used while aggregating the data generally correspond to the month or year-end. For example, a ten daily data series corresponds to three parts of every month in which the first two parts are the 1-10 and 11-20 days of the month and the third part is the remaining days of the month. Thus, every third value in the series corresponds to either 8, 9, 10 or 11 days (the last part of the month) depending on the total days in the month. Similarly, weekly data, depending on its objective, is then taken in one of two different ways: (a) as four parts of the months where first three parts are of seven days each and the fourth part is of 7, 8, 9 or 10 days period (as per the total days in the month) or, which is much more typical in hydrology (b) as 52 parts of a year where first 51 weeks are of 7 days each and the last week is of 8 or 9 days depending upon whether the year is a non-leap or a leap year. Such culmination is often desirable for the operational purpose, as each time interval is related to January 1st of each simulated year.

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 over larger time intervals.

10.3 Comparison of Volume and Runoff Depth

To facilitate comparisons between rainfall and runoff, it is common to express values of rainfall and flow in similar terms. Both may be expressed as a total volume over a specified period (in m^3, thousand m^3, Hectare-metre (Ha m) or million m^3 i.e., MCM). Alternatively, the discharge may be expressed as an average depth in millimetres over the catchment.

Runoff depth is the runoff volume expressed as the depth of water over the specified catchment area adjusted into units of millimetres, which is typically used over longer durations d that represent months, seasons, or most typically years:

$$R_d(mm) = \frac{V_d(m^3)\times 10^3}{Area(\,Km^2)\times 10^6} = \frac{V_d(m^3)}{Area(\,Km^2)\times 10^3} \qquad \text{Equation 10.3}$$

Runoff depths not only provide a comparison with rainfalls, but they also provide a comparison with other catchments standardised by the area. Such comparisons may be made for monthly, seasonal and annual totals but are not generally helpful for daily or shorter duration depths, since the river basins respond to rainfall events at different time scales. It is usual to see higher values of runoff depth at the upstream parts of the catchment that are above the average catchment elevation.

For annual reporting, it is essential to compare the monthly and annual runoff from a station with the long-term average, maximum and minimum monthly runoff derived from the previous record. This requires the annual updating of runoff statistics, with the previous years used for computing the statistics.

Another unit which is sometimes used to standardise discharge against area is a *specific discharge* or *specific yield* which may be computed for the mean discharge over a specified duration as discharge over an area (m^3/sec / km^2).

Some imperial units and other units which were in vogue earlier are now regarded as obsolete and should no longer be used. These include Mgd (million gallons per day), acre-feet, ft^3/sec (cusecs), TAMC (thousand million acre-feet).

10.4 Compilation of Maximum and Minimum Series

The annual, seasonal or monthly maximum series are frequently required. Monthly minimum series may be required for drought analysis, while daily or instantaneous maximums are required for flood studies.

Annual maximum peaks are compiled as a separate series of annual maximums that is used as input into frequency analyses to help determine peak flows for design return periods (e.g., 10, 20, 50 or 100 years). In addition to the peak flows, peak hydrograph volumes can also be used for the computation of values corresponding to different return periods. When there are hourly data on the record, it is advised to collect both the hourly and daily peak flows, since historical data for all available years are usually not continuously available on an hourly basis. The common period when both daily and hourly data are available can then be used to develop a statistical relationship between the daily and hourly peak flows. This statistical relationship can then be used for the computation of hourly peak values for the years where only the daily peak flows are available.

11 ANALYSIS AND USE OF DISCHARGE DATA

11.1 General

The time series data refers to a dataset where each data point has a time coordinate. Statistical analyses of hydrological time-series data at various time scales are based on a set of fundamental assumptions -- that the series is homogenous, stationary, free from trends and periodicity, and with no persistence. Equally important is a need for statistical and hydrological analysis with rainfall and other climatic variables. Various kinds of analysis are required for data validation and reporting. Hydrological time series are typically used as input into water resource planning and management. The methods discussed in the current chapter will help to serve the objective of data presentation and reporting. The methods are**:**

- Computation of basic statistics
- Cumulative frequency distributions (flow duration curves)
- Fitting of theoretical frequency distributions
- Time series analysis
    - moving averages
    - check for trends
    - check for jumps
    - cross correlation with other series
    - auto-correlation
    - balances
- Regression/ relation curves
- Double mass analysis
- Series homogeneity tests

Analysis of hydrological data as mentioned above is carried out at Division level or State Data Processing Centres.

11.2 Computation of Basic Statistics

Basic statistics are required for data validation and reporting. The commonly used statistics are minimum, maximum, mean, median, mode, standard deviation, variance, coefficient of variation, skewness and kurtosis for a defined period of the time series say, annual and monthly.

For data vector X_i (i=1, N), the basic statistics are determined as follows:

- Minimum: $X_{min} = \min(X_1, X_2, X_3, ..., X_N)$
- Maximum: $X_{max} = \max(X_1, X_2, X_3, ..., X_N)$
- Arithmetic mean: $M_X = \frac{1}{N}\sum_1^N X_i$ Equation 11.1
- mode - the value of X which occurs with the greatest frequency or the middle value of the class with the greatest frequency
- median - the middle value of a ranked series X_i in a series, in case there are an odd number of entries, and the average of the two middle values of a ranked series in case there are even number of entries
- Standard deviation (sample): $S_X = \sqrt{\frac{\sum(X_i - M_x)^2}{N-1}}$ Equation 11.2
- Variance $V_x = S_x^2$ Equation 11.3
- Skewness or the extent to which the data deviate from a symmetrical distribution:

$$C_{SX} = \frac{N}{(N-1)(N-2)}\sum_1^N \frac{(X_i - M_x)^3}{S_X^3} \quad \text{Equation 11.4}$$

- Kurtosis or peakedness of a distribution:

$$K_X = \frac{(N^2 - 2N + 3)}{(N-1)(N-2)(N-3)}\sum_1^N \frac{(X_i - M_x)^4}{S_X^4} \quad \text{Equation 11.5}$$

11.3 Cumulative Frequency Distributions (Flow Duration Curves)

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

- in evaluating or computing the dependable flows for the planning of water resource engineering projects
- in evaluating the characteristics of the hydropower potential of a river
- in assessing the effects of river regulation and abstractions on river ecology
- in the design of drainage systems
- in flood control studies
- in computing the sediment load and dissolved solid load of a river
- in comparing with adjacent catchments

A flow duration curve is a plot of discharge against the percentage of time the flow was equalled or exceeded. This may also be referred to as a cumulative discharge frequency curve and it is usually applied to daily mean discharge, although it can also be developed using weekly, 10-daily, monthly, or even annual data, depending on the purpose. Another important feature of the flow duration curve is that it can contain data points from a selected period that may not include the entire year. Hence, flow duration curves for individual months or wet and dry seasons are usually plotted separately, due to a large difference in scale between monsoon and dry season flows.

The process of constructing a flow duration curve can be significantly simplified by using the tools available in excel. However, it is important to understand the concept of the positional probability formula. Let us assume that the given series of measurements are tabulated and ranked from the highest to the lowest value, and a variable called "Rank" is associated with the ordered numbers in the sorted list for each of them. While there are several variants, Weibull's positional probability formula is most commonly used, where the probability of exceeding the value in the sorted order is given as

$$P_i = \frac{Rank}{N+1} \quad \text{Equation 11.6}$$

Table 11-1 shows the construction of flow duration curve using the Weibull formula for monsoon flows. Plotting the last two columns as an X-Y plot produces the Flow-Duration curve.

Table 11-1: Derivation of Flow Duration Table

Original Data prior to sorting	Rank	Probability of Exceedance	Sorted values in descending order
428	1	0.0286	1660
875	2	0.0571	1470
250	3	0.0857	1140
733	4	0.1143	971
450	5	0.1429	887
971	6	0.1714	875
617	7	0.2	864
267	8	0.2286	841
184	9	0.2571	810
162	10	0.2857	767
767	11	0.3143	744
1140	12	0.3429	733
382	13	0.3714	697
1470	14	0.4	665
583	15	0.4286	617
569	16	0.4571	583
365	17	0.4857	569
442	18	0.5143	501
464	19	0.5429	464

Original Data prior to sorting	Rank	Probability of Exceedance	Sorted values in descending order
810	20	0.5714	450
864	21	0.6	447
697	22	0.6286	442
1660	23	0.6571	442
501	24	0.6857	428
442	25	0.7143	391
391	26	0.7429	382
841	27	0.7714	370
744	28	0.8	365
447	29	0.8286	267
370	30	0.8571	250
887	31	0.8857	244
244	32	0.9143	235
235	33	0.9429	184
665	34	0.9714	162

Normally, the flow duration curves should be plotted separately for periods that were not affected by regulation (i.e., prior to the construction of the dam) from the periods when significant regulation altered the previous flow regime. Flow duration curves for the natural streams are used in various ways, and they provide a quick assessment of the availability of flows in median, dry or wet years for a particular location on the river. It is common practice to include all available years of data that are considered homogeneous (i.e., they were not significantly altered by man-made activities or other natural occurrences that might have changed the characteristics of the catchment/ watershed). In some instances, the process of urbanization and land use change may call for separate flow duration curve analyses for periods that are affected by the changes and the previous periods.

Flow duration curves provide no representation of the chronological sequence. This important attribute, for example, the time of occurrence of flows below a specified magnitude, must be dealt with in other ways.

The reliability of flow duration curves depends on the length of the data record and the homogeneity of the data series. When the records are sufficiently long, it is possible to use all daily flows for certain periods within the year and construct customized flow duration curves. Examples of flow-duction curves for the month of July and December from a sample dataset with 67 years are presented below in Figure 11.1 and Figure 11.2.

It should also be noted that the excel function percentile.exc (data range, probability) can facilitate construction of flow duration curves relatively easily without having the rank for all the available data. This command is particularly convenient for very large data sets with thousands of data points that do not all need to be included in the plot. The examples provided below were generated using the percentile.exc function in excel.

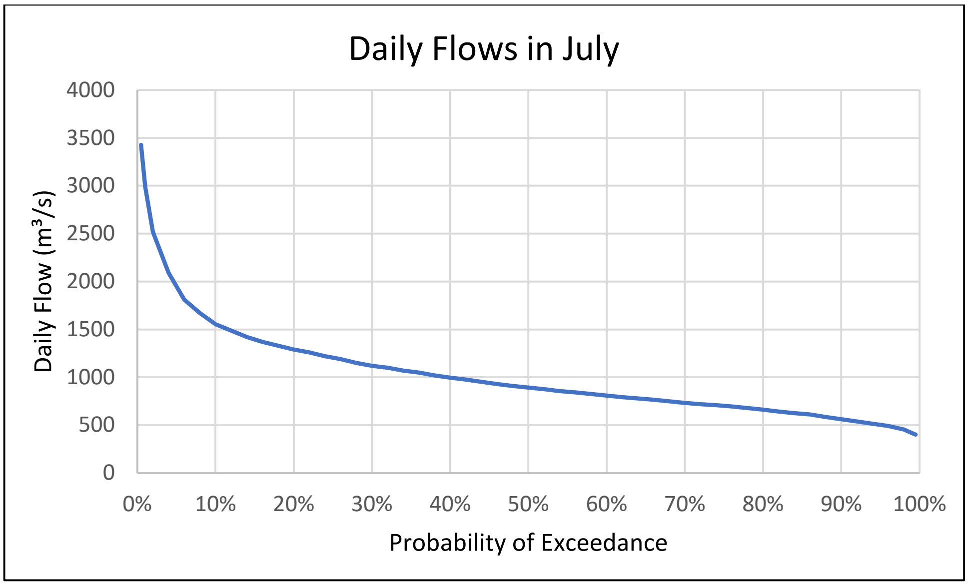


Figure 11.1: Representative Flow Duration Curve for Monsoon Season Flows

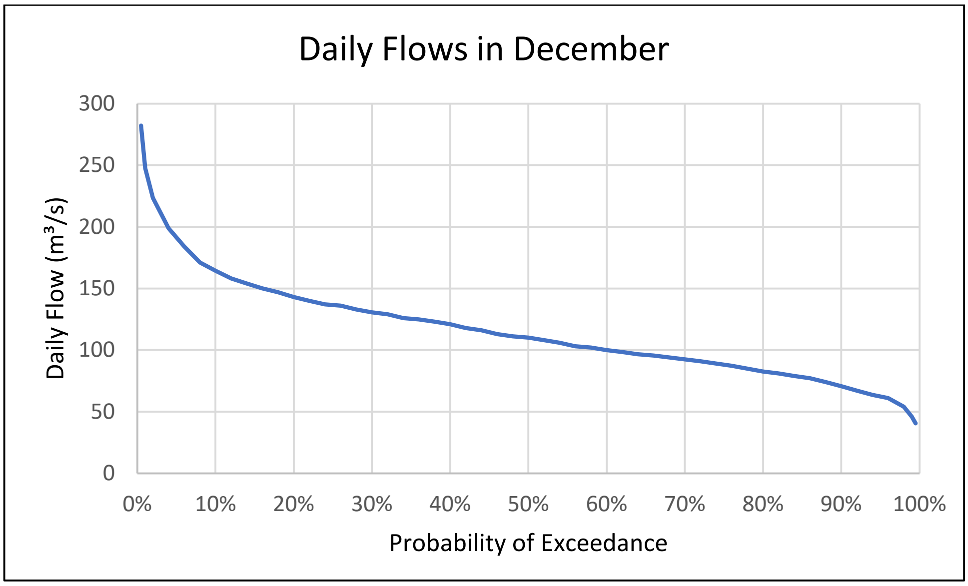


Figure 11.2: Representative Flow Duration Curve for Dry Season Flows

11.4 Fitting of Frequency Distributions

11.4.1 General description

The fitting of frequency distributions to time sequences of streamflow data is widespread, primarily to assess the rare events associated with annual maximums or minimums. The principle of such fitting is that the parameters of the distribution are estimated from the available data sample, 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 a given probability of occurrence (exceedance or non-exceedance) can be computed. Generically, 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)

Different parameters derived from mean, standard deviation and skewness are used in other distributions. Frequency distributions for data averaged over long periods such as annual are often normally distributed. These data can be fitted with the normal distribution, using just the mean and standard deviation to define the distribution. Data become increasingly skewed with shorter durations and need a third parameter to define the relationship. Even then, the fit of relationship at the extremes of the data is the worst. More often, these extremes are of the greatest interest. This may imply that the chosen frequency distribution does not perfectly represent the population of the data and that the resulting estimates may be biased.

Normal or log-normal distributions are recommended for distributions of mean annual flow.

11.4.2 Frequency distributions of extremes

Theoretical frequency distributions are most commonly applied to the extremes of time series, either of floods or droughts. The following series are required:

- maximum series: The annual series depicting the maximum instantaneous discharge value recorded in a year or a month or a 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 the analysis.
- minimum series: The annual series containing the minimum 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 object of flood frequency analysis is to assess the magnitude of a flood with a given probability or return period of occurrence. Return period is the reciprocal of probability, and may also be defined as the average interval between floods of a specified magnitude expressed in years, as may be expected over a large time period. It is to be noted that the recurrence interval does not convey about the time sequence of the occurrence of floods. Floods with magnitude equalling or exceeding that of the 100-year flood may occur at a place over two consecutive years, even though the probability for the same might be very small.

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

- Normal, log-normal 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)
- Goodrich/ Weibull distribution
- Exponential distribution
- Generalised Pareto distribution
- Kappa distribution
- Wakeby distribution

Different distributions fit best to different individual data sets but, there is no single distribution that represents the population of annual flood peaks equally well at all stations. One has to use judgement as to which one to use at a particular location, depending on the experience with flood frequency distributions in the surrounding region and the physical characteristics of the catchment. No recommendation can therefore be made with respect to the selection.

A standard statistic which characterises the flood potential of a catchment, and has been used as an 'index flood' in the regional analysis is the mean annual flood, which is the mean of the maximum instantaneous floods recorded in each year. This can be derived from the data or distribution fitting. An alternative index flood is the median annual maximum, similarly derived. These may be used in reporting 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. Detailed descriptions of the mathematical functions and application procedures have not been included here. These can be found in standard mathematical and hydrological texts.

However, there is a common tendency to apply the techniques of fitting distributions to carry out a flood frequency analysis of any extreme series, aided further because of its ease of execution. This is discouraged. The general approach presented here should be applied on natural flow series, not on the flow series that were significantly affected by regulation. A separate section within the manual deals with the process of naturalization of flows.

11.4.3 Check for outliers

Many statistical tests are available for the purpose. However, the removal of outliers simply based on the results of statistical analysis is not suggested in the hydrological analysis. There should be attempts to corroborate the extreme value with available secondary information (e.g., heavy flood causing life loss may speak for very high discharge, uncommon in the series otherwise). A final decision to exclude the outlier from the series will depend on the judgement.

11.5 Time Series Analysis

A time series is a collection of observations made sequentially in time, such that each data point is associated with its time coordinate. A typical time-series may be composed of four parts:

a) A trend or long-term movement

b) Sudden shift or jump

c) A seasonal or periodic/ cyclic effect

d) A "random", "unsystematic", or "irregular" component.

As a matter of mathematical description, a series can always be represented as a combination of one or more of these constituents.

(i) Trend

A trend is a long-term gradual change covering many years reflecting the general tendency of the series to increase or decrease.

(ii) Sudden shift or jump

Rather than gradual change, sometimes there may be a sudden shift or jump in the mean of the series.

(ii) Seasonal effect

This is a fluctuation imposed on the series by a cyclic phenomenon external to the main body of causal influences at work upon it. Seasonal effects are associated with an annual period of occurrence.

(iii) Random Component

These components represent the stochastic or random variations; the residuals left after trend and cyclic variations have been removed from a set of data.

Time series analysis may be used to test the variability, homogeneity or trend of a streamflow series or to provide an insight into the characteristics of the series as displayed graphically. The following time series statistics are described here:

- moving averages
- run sum and run length

11.5.1 Moving Averages

To investigate the long-term variability or trends in series, moving average curves are useful. A moving average series Y_i of series X_i is derived as follows:

$$Y_i = \frac{1}{2M+1} \sum_{j=i-M}^{j=1+M} X_i \qquad \text{Equation 11.7}$$

where averaging takes place over 2M+1 elements. A moving average is commonly plotted with time-series data to investigate the long-term variability or trends in series. The Moving Average smoothens out short-term fluctuations and highlights longer-term trends or cycles. The plot requires selection of the series with defined 'unit' for moving average. An example is shown in Table 11-2 and Figure 11.4.

Table 11-2: Computation of Moving Averages

Year	Annual Discharge (m^3/s)	3 year MA	5 Year MA
1	2	3 = 2(i)+2(i+1) +2(i+2)]/3	4 = [2(i)+2(i+1)+2(i+2) +2(i+3)+2(i+4)]/3
1990-91	534.73		
1991-92	290.75	383.28	
1992-93	324.37	277.43	362.91
1993-94	217.16	329.69	313.59
1994-95	447.55	317.60	294.13
1995-96	288.10	309.71	266.73
1996-97	193.49	222.98	249.64
1997-98	187.34	170.84	208.26
1998-99	131.70	186.57	183.97
1999-00	240.67	179.66	219.06
2000-01	166.62	258.75	202.63
2001-02	368.97	213.60	229.27
2002-03	105.21	246.36	236.14
2003-04	264.89	215.04	251.83
2004-05	275.02	261.66	274.85
2005-06	245.06	334.72	320.52
2006-07	484.07	354.23	309.43
2007-08	333.56	342.35	
2008-09	209.43		

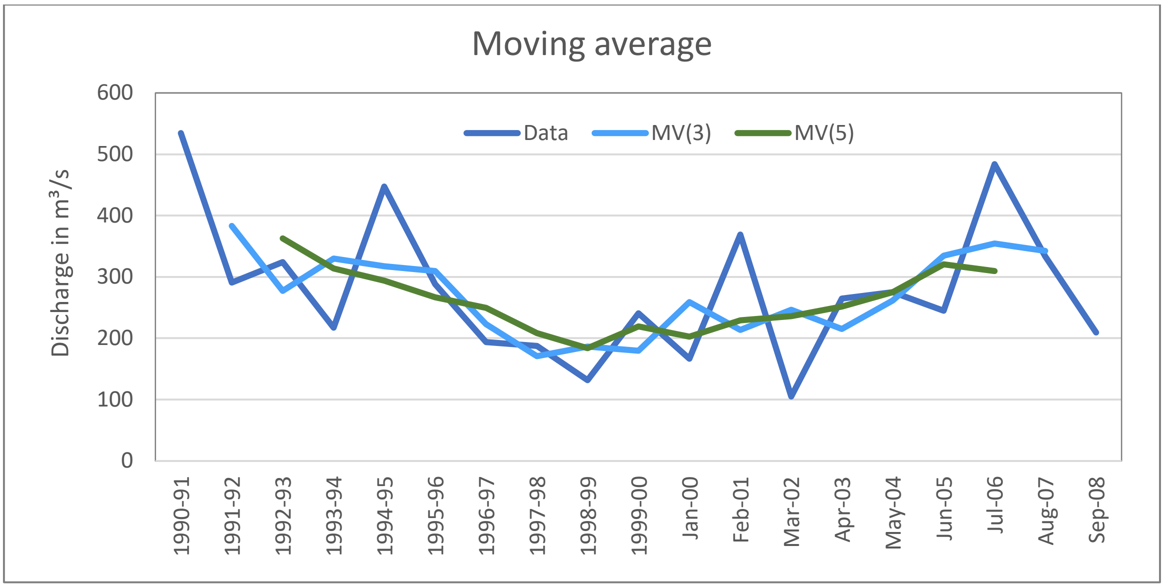


Figure 11.3: Moving Average

11.5.2 Check for trends

The Mann-Kendall It has been suggested for detection of a trend (CWC, 2001) in the annual flow series. Long term trends are studied using annual data (i.e., mean annual flow).

11.5.2.1 Mann-Kendall test

This test is based on the S statistic defined as (Yue et al., 2002):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i) \qquad \text{Equation 11.8}$$

where, x_j are the sequential data values, n is the length of the data set and

$$sgn(y) = \begin{Bmatrix} 1 \cdots if\,(y > 0) \\ 0 \cdots if\,(y = 0) \\ -1 \cdots if\,(y < 0) \end{Bmatrix} \qquad \text{Equation 11.9}$$

Mann and Kendall have documented that when n ≥ 8, the statistic S is approximately normally distributed with the mean:

$$E(S) = 0 \qquad \text{Equation 11.10}$$

and variance:

$$V(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{m} t_i(t_i-1)(2t_i+5)}{18} \qquad \text{Equation 11.11}$$

where m is the number of tied groups and t_i is the size of the i^{th} tied group. The standardised test statistic Z is computed by:

$$Z_{MK} = \begin{Bmatrix} \frac{S-1}{\sqrt{Var(S)}} \; when \; S > 0 \\ 0 \quad when \;\; S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} \; when \; S < 0 \end{Bmatrix}$$ Equation 11.12

The standardised Mann-Kendall statistic Z follows the standard normal distribution with a mean of zero and variance of one.

A numerical example that uses MS Excel to set up calculations for the Mann-Kendall statistic is shown in https://www.real-statistics.com/time-series-analysis/time-series-miscellaneous/mann-kendall-test/ and it is relatively easy to repeat it with hydrological data. Again, only the mean annual flows should be used in this analysis.

When the calculated Mann-Kendall statistics has a negative value, this indicates a declining trend present in data, which could be due to either a measurement error or what is usually more likely, an increase in unreported water use. It is often a good practice to plot the time series data (annual flows) against a moving average with a base between 6 and 10 years and compare the two plots, as demonstrated in Figure 11.3 below.

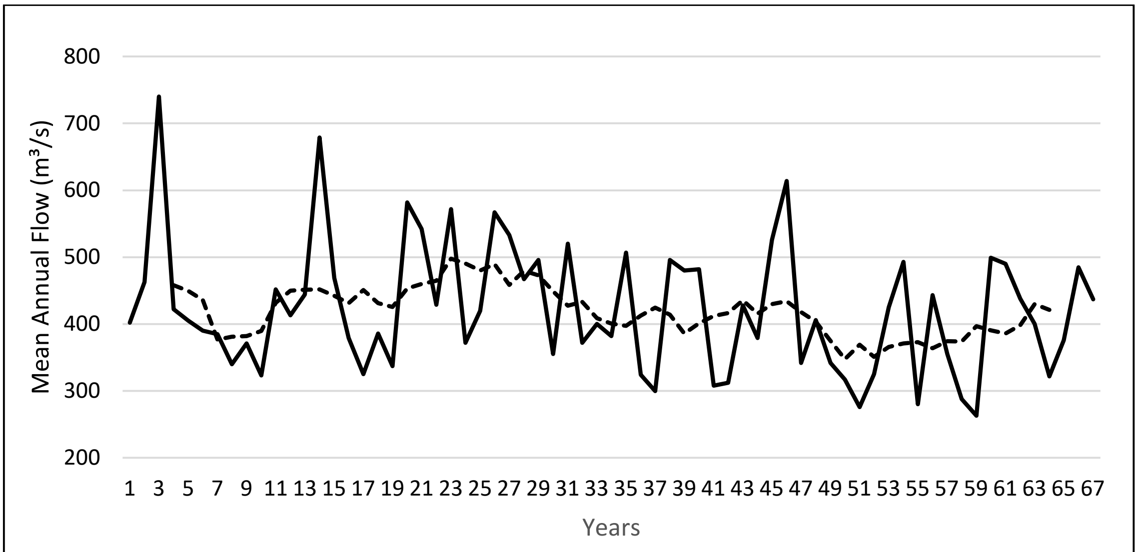


Figure 11.4: Example of Annual Flow Series with a 7-Year Moving Average

The existence of a small trend in a relatively short series of duration less than 40 years may be misleading. There are longer inter-annual climate cycles that are stationary in the long term, but that may be viewed as "trends" over shorter time periods. This is especially the case if similar trends are discovered in adjacent hydrometric stations within the same region or the same river basin. In such instances, no intervention with data modification is recommended. However, the possible reasons for the presence of the trend should be identified.

Fitting a regression line in MS Excel is one of the easiest ways to identify a trend in the data, and it is commonly used. However, the regression lines are influenced by the presence of outliers in the data. Even a single extreme data point may significantly influence the estimated slope, especially if it is located close to the beginning or the end of a series. To avoid this situation, it is suggested to use non-parametric methods like Theil and Sen's median slope estimation approach for the purpose, which is not affected by the presence of outliers or the probability distribution of the data. The hydrological data do not follow the normal distribution, an inherent assumption of their parametric counterparts like the regression line, thus affecting the results.

11.5.2.2 Theil and Sen's median slope

If a linear trend is present, this simple non-parametric procedure can be used to estimate the true slope (change per unit time). The methodology is described by Yu et.al. (1993) and Kahya and Kalayci (2004). The procedure is not greatly affected by gross data errors or outliers and can be used for records with missing values. In this approach, the slope estimates of N pairs of data are first computed by

$$Q_i = (x_j - x_k)/(j - k) \quad \text{for} \;\; i = 1, \dots N \qquad \text{Equation 11.13}$$

where, x_j and x_k are data values at times j and k, $(j > k)$ respectively. The median of these N values of Q_i is Sen's estimator of the slope. If there is only one data in each period, then

$$N = n(n-1)/2 \qquad \text{Equation 11.14}$$

where n is the number of periods. The median of the N estimated slopes is obtained in the usual way, i.e., the N values of Q_i are ranked by $Q_1 \le Q_2 \le \dots \le Q_{N-1} \le Q_N$ and

$$\text{Sen's estimator} = \begin{cases} Q_{(N+1)/2} & if\ N\ is\ odd \\ \frac{1}{2}\left(Q_{N/2} + Q_{(N+2)/2}\right) & if\ N\ is\ even \end{cases} \qquad \text{Equation 11.15}$$

11.5.3 Check for jumps

The t-test has been most commonly used for the detection of jump in a series, i.e., whether the series became statistically different from its original after some time, and continued to be the same thereafter. This might relate to some permanent change in the catchment. However, the application of t-test demands prior knowledge of the period of change. The Buishand's U test and Pettitt-Mann-Whitney test can be used for the detection of jump that is anticipated at an unknown period. The procedure may be the application of Buishand's U test to detect the presence of jump. If it is found to exist in the series, then application of Pettitt-Mann-Whitney test can be done to find out the most probable date of the occurrence of the jump. Thereafter, application of Student's t-test can be made to confirm its presence. Once the presence of jump is confirmed, the series can be adjusted by adding or subtracting the difference of the means for the two periods from one part. This random series should be used for the frequency analysis. After analysis, the effect of the jump should be added back to arrive at the final results.

11.5.3.1 Buishand's U test

This test, developed by Buishand (1984) is good for detecting a jump in the middle of the sequence. It is robust against departures from normality (Shahin et al., 1993).

The model with a single shift in the mean is

$$Y_i = \begin{cases} (\mu + \varepsilon_i) \forall i = 1, \dots, m \\ (\mu + \Delta + \varepsilon_i) \forall i = m+1, \dots, n \end{cases} \qquad \text{Equation 11.16}$$

ε_i = independent random variable with zero mean and variance σ^2

m = the change point

μ = mean (unknown)

Δ = shift in the mean (unknown)

Hypotheses

The null hypothesis H_0: Δ = 0

The alternative hypothesis H_1: Δ ≠ 0

Consider the adjusted partial sums or cumulative deviations from the mean:

$$S_k^* = \sum_{i=1}^{k} (Y_i - \overline{Y}) \forall k = 1, \dots . n \qquad \text{Equation 11.17}$$

Where:

$\bar{Y}$ = average of $Y_1, Y_2, \dots, Y_n$

D_Y = standard deviation, such that

$$D_Y^2 = \sum_{i=1}^{n} (Y_i - \overline{Y})^2 / n \qquad \text{Equation 11.18}$$

The test statistic U is defined as

$$U = [n(n+1)]^{-1} \sum_{k=1}^{n-1} \left(S_k^* / D_Y\right)^2 \qquad \text{Equation 11.19}$$

Critical values have been provided by Buishand (1984) and reproduced below:

Sample size n	α=0.10	α=0.05	α=0.01
10	0.333	0.416	0.574
20	0.340	0.440	0.659
30	0.343	0.447	0.688
40	0.344	0.451	0.702
50	0.345	0.453	0.710
100	0.346	0.457	0.727
∞	0.347	0.461	0.743

11.5.3.2 Pettitt-Mann-Whitney test

This non-parametric test aims to provide a more precise determination of change point (Kiely et al., 1998). Let T be the length of the time series and let t be the year of the most likely change point. The single time series is considered as two samples represented by $X_1 \ldots, X_t$ and $X_{t+1} \ldots, X_T$. The index is defined as:

$$V_t = \sum_{j=1}^{T} sgn(X_t - X_j) \text{ for any t} \qquad \text{Equation 11.20}$$

Where,

$$sgn(x) = \begin{Bmatrix} 1 \cdots for\forall x > 0 \\ 0 \cdots for\forall x = 0 \\ -1 \cdots for\forall x < 0 \end{Bmatrix} \qquad \text{Equation 11.21}$$

Let a further index U_t be defined as

$$U_t = \sum_{i=1}^{t} \sum_{j=1}^{T} sgn(X_i - X_j) \qquad \text{Equation 11.22}$$

A plot of U_t against t for a time series with no change point would result in a continually increasing value of $|U_t|$. However, if there is a change point (even a local change point) then $|U_t|$ would increase up to the change point and then begin to decrease. The most significant change point t can be identified as the point where the value of $|U_t|$ is maximum:

$$K_T = \max_{1 \le t \le T} |U_T| \qquad \text{Equation 11.23}$$

The probability of a change point being at the year where $|U_t|$ is the maximum, is approximated by

$$p = 1 - exp\left[\frac{-6K_T^2}{T^3+T^2}\right] \qquad \text{Equation 11.24}$$

If further, it is introduced, for $1 \le t \le T$, the series

$$\widehat{U}(t) = |U_t| \qquad \text{Equation 11.25}$$

and it is defined

$$P(t) = 1 - exp\left[\frac{-6\widehat{U}(t)^2}{T^3+T^2}\right] \qquad \text{Equation 11.26}$$

a series of probabilities of significance for each year can be obtained. The point with the highest probability is the desired change point.

11.5.4 Run Length and Run Sum characteristics

Run Analysis is particularly related to draughts. The properties of time series which are used in drought analysis are run-length and run- sum. A run can be determined by its length, its sum or its intensity. Consider the time series $X_1 \cdots X_n$ and a constant demand level X_C as shown in Figure 11.5. A negative run occurs when X_t is less than X_C consecutively during one or more time intervals. Similarly, a positive run occurs when X_t is consecutively greater than X_C. The means, the standard deviation and the maximum of run length and run sum are important characteristics of the time series.

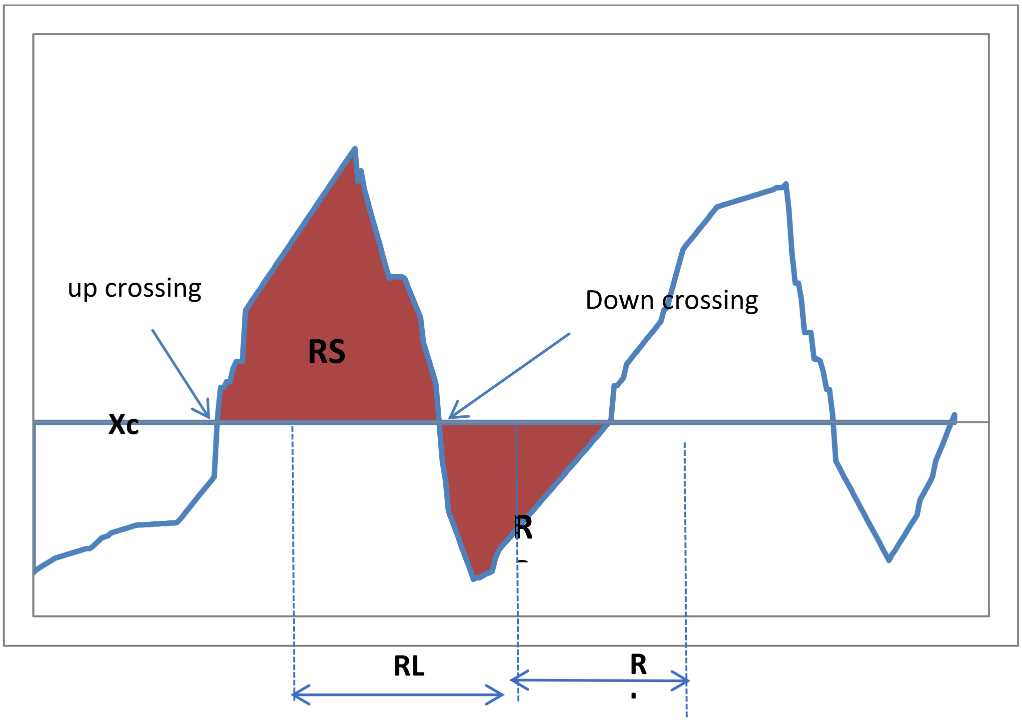


Figure 11.5: Statistics for Run Analysis

If X_c is the crossing level, the up crossing is defined by:

$X_{i+1} \geq X_C$ and $X_i < X_C$

Similarly, the down crossing is defined by:

$X_i \geq X_C$ and $X_{i+1} < X_C$

The statistics calculated under run analysis are:

Positive Run length (RL+): The period between an up crossing and a down crossing, given as several time intervals.

Negative Run length (RL -): The period between a down crossing and an up crossing given as several time intervals.

The positive run sum, negative run sum and total run length are calculated by:

Positive Run sum (RS +) = $\sum_{i=j}^{m}(X_i - X_c)\, C_f$

Where:

j = location of an up crossing

k = location of the next down crossing

C_f = conversion factor (= length of time steps per time interval) to transfer intensities into volumes

Negative Run sum (RS -) = $\sum_{i=k}^{m}(X_c - X_i)$ Cf

where:

k = location of the down crossing

m = location of the next up crossing

Total Run length is the sum of successive pairs of RL + and RL -.

The use of the Length Run and Run Sum is limited since it has been replaced by the use of modern river basin management models, in which water demand varies spatially and temporally throughout the basin, which has made the fixed parameter X_C obsolete.

Example 11-1

Table 11-3 and Figure 11.6 gives the Run-length and the Run sum of an annual flow series observed from 1976 to 1995

Table 11-3: Computation of Run Length and Run Sum

Date	Data (Unit of data)	Crossing level = 100	RS-	RS+
1976	0	-100		
1977	5	-95		
1978	26.2	-73.8		
1979	78.2	-21.8	-290.6	
1980	387.9	287.9		
1981	238.4	138.4		
1982	245	145		
1983	260.4	160.4		
1984	306.7	206.7		938.4
1985	56	-44		
1986	0	-100		
1987	0	-100		
1988	0	-100		
1989	3.9	-96.1		
1990	83.7	-16.3		
1991	67.8	-32.2	-488.6	

Date	Data (Unit of data)	Crossing level = 100	RS-	RS+
1992	279.9	179.9		
1993	206.6	106.6		
1994	232.7	132.7		
1995	217.5	117.5		536.7

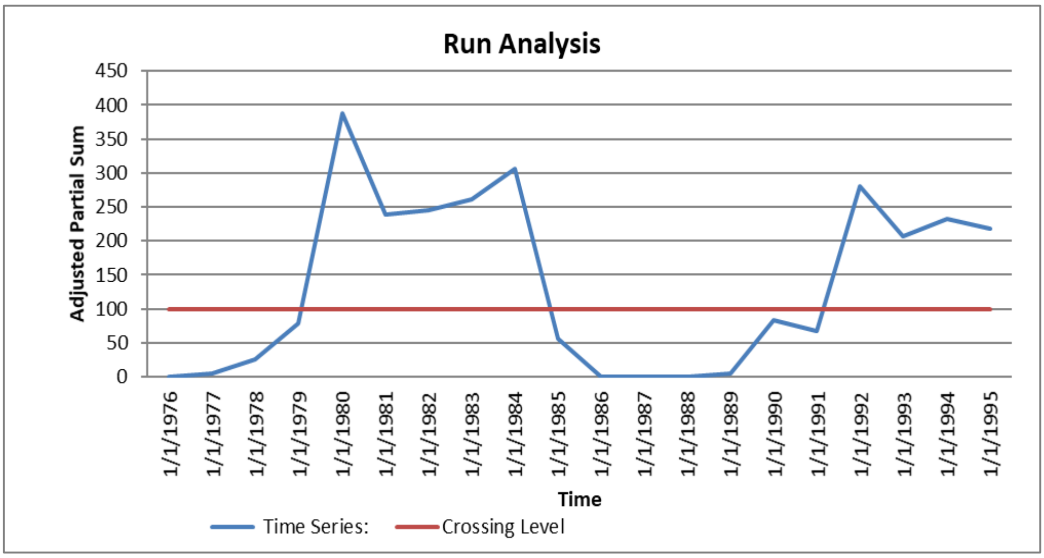


Figure 11.6: Run Analysis

Positive Runsum RS+	Negative Runsum RS-
938.400	290.600
536.700	488.600

Let us assume that m runs of run-lengths l(1), l(2),……l(m) and run-sums d(1), d(2),……d(m) occur. The mean, standard deviation and the maximum of run length and run sum are important characteristics describing the runs of a given time series. They can be used for comparison with the corresponding characteristics derived from mathematical models fitted to historical series.

11.5.5 Storage analysis

Time series of flows and water demands are jointly related to the needs for storage reservoirs. While river basin management models are used nowadays to address the storage requirements, the sequent peak algorithm was one of the first historical attempts to compute water shortage (or the equivalent storage requirements to eliminate water shortage) without running dry for various draft levels from the reservoir. The algorithm considers the following sequence of storages:

$$S_{\ i} = S_{i-1} + (X_i - D_x)C_f \qquad \text{for } i = 1, 2, N; \qquad \text{Equation 11.27}$$

$S_0 = 0$

where:

X_i = inflow

D_i = $D_L\ m_x$

m_x = average of x_i, i = 1, N

D_L = draft level as a fraction of m_x

C_f = multiplier to convert intensities into volumes (times units per time interval)

Since demands for irrigation in the dry season are met from the storage, the above analysis has to be conducted for each individual year, resulting in a series of annual maximum storage requirements, that can then be analysed statistically. This approach is no longer used today, having been replaced by water resources models that can use optimization over multiple time steps to arrive at the best solution in each simulated year, as outlined in the last section of this manual.

11.6 Homogeneity Tests of Time Series

For statistical analysis, time-series data from a single series should ideally possess the property of homogeneity. This implies that different sections taken out from the same long data series should be statistically similar, and their positional relationships should also remain preserved. Various methods for testing the Consistency and Homogeneity of flow series are:

- Student's 't' test for difference of means
- Fischer 'F' test for equality of variance
- Trend test

The homogeneity tests of time series are routinely applied to streamflow series. However, it should also be recognised that lack of homogeneity in streamflow series may arise from a variety of sources, including:

- data error
- changes in land use/ land cover in the catchment
- changes in abstractions and river regulation
- climate change

11.6.1 T-Test for equality of means

The T-test is a common application to check whether one dataset is different from the other dataset in a statistically significant manner. There are several variances of the t-test. The two-sample test (Shahin et.al., 1993) is used to test the hypothesis that two populations have the same mean. The test procedure for a two-sided test is summarised as follows:

(i) Assumptions: The two independent samples originate from normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$

(ii) Hypotheses: The null hypothesis reads$H_0: \mu_1 = \mu_2$. The alternative hypothesis reads$H_1: \mu_1 \neq \mu_2$.

(iii) The test statistic: The population variance is not known for either of the populations. There are two different test statistics, depending on the equality or inequality of the estimated variances. With the notation $\overline{x}_1$, $\overline{x}_2$ for the two sample means, s_1^2,s_2^2 for the two sample variances, n_1,n_2 for the sample sizes, and

$$s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} \qquad \text{Equation 11.28}$$

for the pooled sample variance (i.e., the weighted mean of the two sample variances), the test statistics may be written as

$$T_1 = \frac{\overline{x}_1-\overline{x}_2}{s_p\sqrt{1/n_1+1/n_2}} \qquad \text{Equation 11.29}$$

$$T_2 = \frac{\overline{x}_1-\overline{x}_2}{\sqrt{s_1^2/n_1+s_2^2/n_2}} \qquad \text{Equation 11.30}$$

Essentially, all test statistics are standardised differences of sample means. The first test statistic is used when the two population variances are equal as compared by the F test. Under the null hypothesis $T_1 \sim t(n_1 + n_2 - 2)$.

The second statistic can be applied in case the two population variances are unequal. Under the null hypothesis $T_2 \sim t(\nu)$, where

$$\nu = \frac{(s_1^2/n_1+s_2^2/n_2)^2}{s_1^4/(n_1^2(n_1+1))+s_2^4/(n_2^2(n_2+1))} - 2 \qquad \text{Equation 11.31}$$

(iv). Critical region: The test significance level is chosen first. In the two-sided case, the hypothesis is rejected if

$$|T_1| \geq t_{\alpha/2}(n_1 + n_2 - 2) \qquad \text{Equation 11.32}$$

$$\text{Or, } |T_2| \geq t_{\alpha/2}(\nu) \qquad \text{Equation 11.33}$$

The parameter ν has the same meaning as described in the previous item.

A detailed reference to T-Test is available in any Standard Text Book on Statistical Analysis.

11.6.2 F-Test for equality of variance

F-test (Snedecor and Cochran, 1994) is used to test if the variances of two populations are equal. This test can be one-tailed or two-tailed. The choice of the tests depends on the problem. The one-tailed test is in one direction i.e., the variance from the first population is either greater than or less than the second population variance. For instance, a new data set is tested, it is of interest to know if it is more variable than the old data, set requiring one-tailed test.

The details of the F-test procedure for testing the equality of two variances together with the necessary assumptions are as follows:

(i) Assumptions: Two independent samples originate from Normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. The population means and the population variances are unknown.

(ii) Hypothesis: The null hypothesis reads $H_0: \sigma_1^2 = \sigma_2^2$. The alternative hypothesis reads $H_1: \sigma_1^2 \neq \sigma_2^2$.

(iii) The test statistic: The test statistic is $F = \frac{s_1^2}{s_2^2}$ Equation 11.34

(iv) Here s_1^2 and s_2^2 are the two sample variances. Under the null hypothesis $F \sim F(n_1 - 1, n_2 - 1)$, where n_1 and n_2 are the sample sizes.

(v) Let α be chosen as the significance level of the test. In case of a two-sided test, the null hypothesis is rejected if $F \leq K_1$ or if $F \geq K_2$. The critical value K_2 is determined from

$Pr(F \geq K_2; F \sim F(n_1 - 1, n_2 - 1)) = \alpha/2.$ Equation 11.35

Hence $K_2 = F_{\alpha/2}(n_1 - 1, n_2 - 1)$ Equation 11.36

In the same way K_1 has to be determined from the equation $Pr(F \leq K_1; F \sim F(n_1 - 1, n_2 - 1)) = \alpha/2$ Equation 11.37

Hence $K_1 = F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$ Equation 11.38

Tables from which K_1 can be taken directly are not generally available. However, the relation

$F_{1-\alpha/2}(n_1 - 1, n_2 - 1) = 1/F_{\alpha/2}(n_2 - 1, n_1 - 1)$ Equation 11.39

may be used. The two critical values may be computed from the same table.

(vi) F is computed from the two samples and the null hypothesis is rejected if

$F \leq 1/F_{\alpha/2}(n_2 - 1, n_1 - 1)$ Equation 11.40

or if $F \geq F_{\alpha/2}(n_1 - 1, n_2 - 1)$ Equation 11.41

Further details are available in any standard textbook on statistics. This test can also be carried out on Microsoft Excel.

11.6.3 Trend line

A trend line represents the long-term movement in time series data. It indicates whether a particular data set has gradually increased or decreased over a lengthy period. Trend lines are typically estimated as straight lines, though its process of origin might be linear or non-linear. In the absence of a significant linear trend, a series can be considered homogenous.

T-test may be used for computing the significance (statistical significance = 100-confidence, refers to the probability of mistakenly rejecting the null hypothesis when it was true) of a trend line slope estimated using the regression approach. Student's 't' test against the slope of regression line different from zero can be adopted to identify a trend in a series.

Assumption: residuals ε_i are stationary and sequentially independent. The linear regression model is fitted given by:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad \text{Equation 11.42}$$

Where X_i, i = 1, ..., n are known, α and β are least-square estimators and ε_i are independent identically normally distributed random errors with expected value 0 and unknown variance σ^2, and Yi, i = 1, ..., n are the observed values.

It is desired to test the null hypothesis H_o that the slope β is equal to 0.

Then the Test Statistic

$$T = \frac{|\beta|}{SE_\beta} \qquad \text{Equation 11.43}$$

has a t-distribution with $n - 2$ degrees of freedom if the null hypothesis is true. The standard error SE_β of the slope coefficient is given by (Levine et al., 2010):

$$\boldsymbol{SE_\beta} = \frac{\sqrt{\frac{1}{n-2}\sum_1^n \left(Y_i - \widehat{Y_i}\right)^2}}{\sqrt{\sum_1^n (X_i - \bar{X})^2}} \qquad \text{Equation 11.44}$$

The Null Hypothesis is defined by H_0: $|\beta| = 0$

The null hypothesis that the series has zero trends is rejected if $|T| > t_{1-\alpha/2,\nu}$ where $t_{1-\alpha/2,\nu}$ is the critical value of the t distribution with ν degrees of freedom obtained from the statistical tables.

The presence of trend is accepted only when the test is carried out on long data series aggregated into the monthly or annual values for the trend analysis. However, performing operations for data correction in case of the trend observed in the series is not recommended without knowledge of underlying functions.

Example 11-2

To conduct a trend analysis on annual flow series given below (Please refer to Figure 11.7)

Consider the null Hypothesis H_o: The time series has zero trend

Intercept parameter $|\beta|$ = 0.0171

Slope parameter = 28.347

Table 11-4: Annual Flow Series for Trend Analysis

Year	Flow (m^3/s)
1994-95	52.77
1995-96	31.77
1996-97	24.14
1997-98	15.86
1998-99	12.37
1999-00	28.54
2000-01	15.86
2001-02	40.90
2002-03	12.6
2003-04	36.2
2004-05	29.5
2005-06	17.4
2006-07	47.9
2007-08	35.0
2008-09	22.4

SE_{β} = 0.7895

$T_{\text{Test Static}}$ = |-0.0171|/0.7895 = 0.0218

T_{critical} = 0.3918

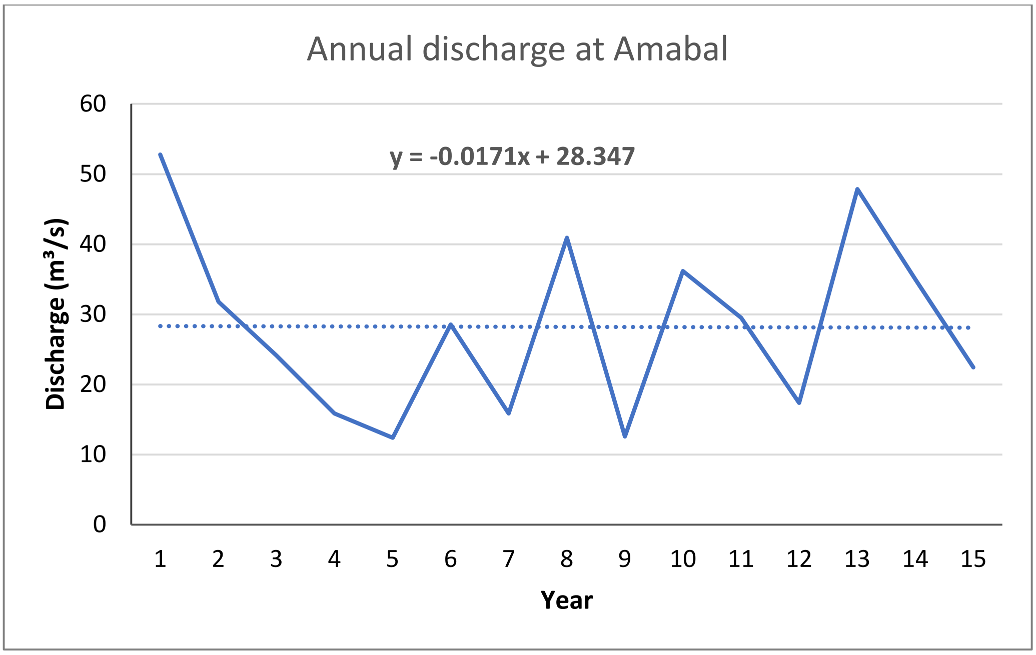


Figure 11.7: Annual Flow Series of Amabal

Referring to the annual flow series of Amabal, since T (=0.0218)<=T critical (0.3918), the Null Hypothesis H_0: The time series has zero trends is not rejected at 0.05 significance level

11.7 Rainfall-Runoff Simulation

The uses of Rainfall-Runoff simulation models are much wider than data validation and include the following:

- filling in and carrying out extension of discharge series
- generation of discharges from synthetic rainfall
- real-time forecasting of flood waves
- determination of the influence of changing land use/cover on the catchment

Rainfall-runoff modelling should be used as the last option. This is because the observed and the modelled data often reveal huge discrepancies, which can result in gross errors when this approach is used to in-fill the missing data. Whenever possible, filling up of missing data in gauged catchments should be conducted by using other regional analyses techniques that are commonly used by the hydrologists.

12 DEVELOPING THE RATING CURVE OR STAGE DISCHARGE CURVE

12.1 General

Hydrological analysis for undertaking any water resource or flood control initiative relies heavily on the availability of continuous streamflow data. However, measurement of flow past a river section continuously is usually impractical or prohibitively expensive. However, the stage can be observed continuously or at regular short time intervals with comparative ease and low cost. A relation exists between stage and the corresponding discharge at river section. This relation is termed a stage-discharge relationship or stage-discharge rating curve, or simply a "rating curve".

A rating curve is established by making several concurrent observations of stage and discharge over a period covering the expected data range at the river gauging section. At many locations, the discharge is not a unique function of the stage; variables such as surface slope or rate of change of stage against the time must also be known to obtain the complete relationship under such circumstances. The rating relationship thus established 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 12.1, by the average curve fitted on the scatter plot of water level (as ordinate) and discharge (as abscissa) at any river section. Even though the discharge is estimated from the observed water level (the primary variable) it is customary to plot the water level along the ordinate, rather than the abscissa.

If Q and h are the discharges and water levels respectively, then the relationship can be analytically expressed as:

$$Q = f(h) \qquad \text{Equation 12.1}$$

Where *f(h)* is an algebraic function of water level (typically exponential or quadratic). A graphical stage-discharge curve helps in visualizing the relationship and transforming the stage records manually to discharge whereas an algebraic relationship can be used for analytical transformation of observed water levels into the corresponding flow estimates.

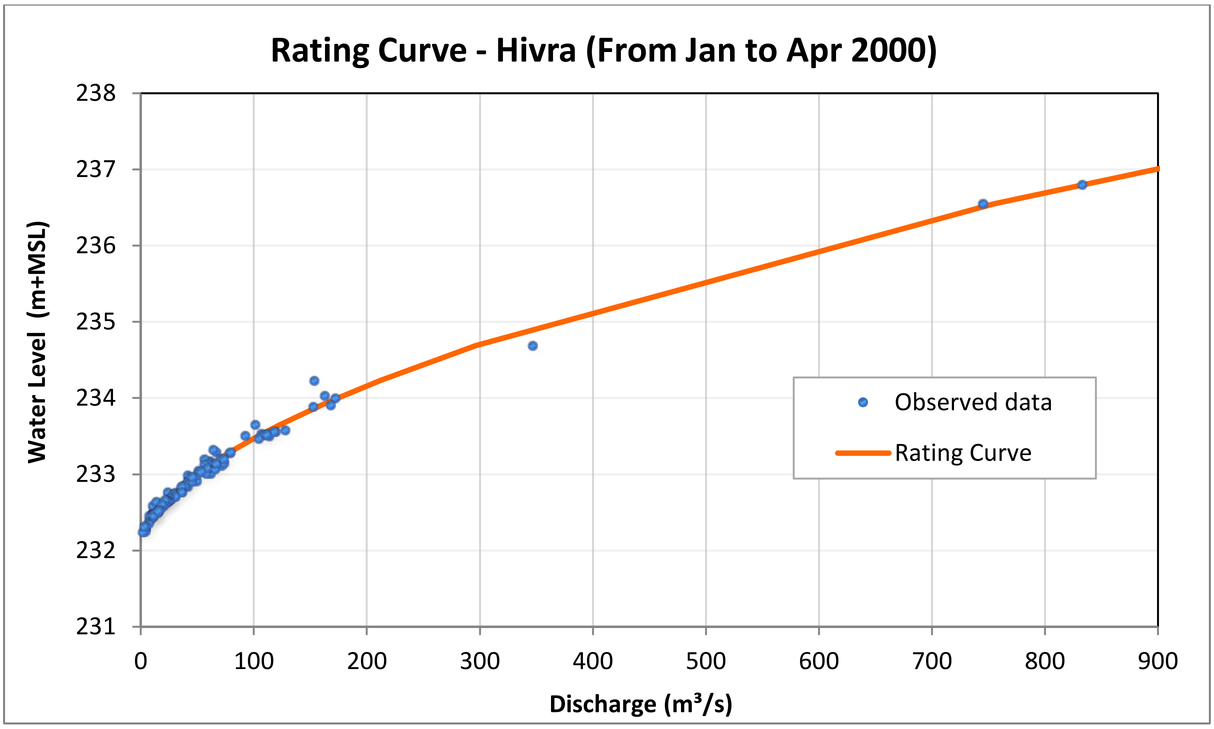


Figure 12.1: Example of Stage-Discharge Rating Curve

It is difficult to measure flow at very high and low stages due to their infrequent occurrence and the inherent difficulty of such measurements. For this, extrapolation is required to cover the full range of flows. Methods of extrapolation have been described in Chapter 14.

12.2 The Station Controls

A stage-discharge curve or rating curve is a graph of water surface elevation versus flow rate in a channel which is set up for a pre-selected cross-section referred to as 'control'. The shape, reliability and stability of the stage-discharge relation are controlled by a section or reach of the channel at or downstream from the gauging station and known as the station control. The establishment and interpretation of a stage-discharge relationship require an understanding of the nature of controls and the types of control at a particular station. Fitting of stage-discharge relationships must not be considered simply a mathematical exercise in curve fitting. Staff involved in fitting stage-discharge relationships should have familiarity and experience with field hydrometry. The channel characteristics forming the control include the cross-sectional area and shape of the stream channel, expansions and restrictions in the channel, channel sinuosity, the stability and roughness of the streambed, and the vegetation cover - all of which collectively constitute the factors determining the channel conveyance.

12.2.1 Types of station control

The character of the rating curve depends on the type of control which in turn is governed by the geometry of the cross-section and by the physical features of the river downstream of the section. Station controls are classified in many ways as:

- section and channel controls
- natural and artificial controls
- complete, compound and partial controls
- permanent and shifting controls

12.2.1.1 Section and channel controls

When the downstream control is such that any change in the physical characteristics of the channel downstream does not affect the flow at the gauging section itself, then such control is termed as section control. In other words, for any disturbance downstream, the section-control will not allow passing the disturbance in the upstream direction. Natural or artificial local narrowing of the cross-section creating a zone of acceleration are some examples of section controls (Figure 12.2). The section control necessarily has a critical flow section at a short distance downstream.

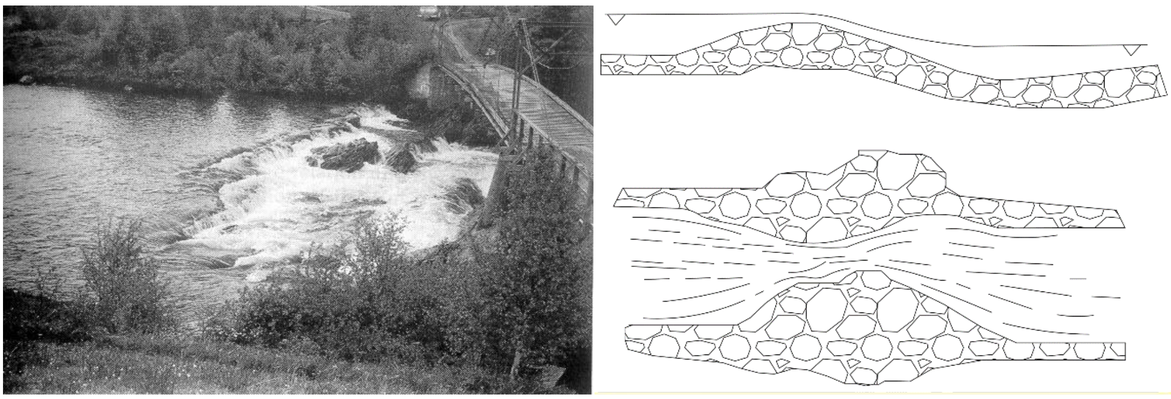


Figure 12.2: Example of Section Control [adopted from Herschy, 2009 and Boiten, 2005]

A section control can be a man-made feature such as a small dam, a weir, a flume, or an overflow spillway. Section controls can frequently be visually identified in the field by observing a riffle, i.e., pronounced drop in the water surface, as the flow passes over the control.

Frequently, as the gauge height increases because of higher flows, the section control will become submerged in such a way that it no longer controls the relation between gauge height and discharge. At this point, the rifle is no longer visible, and flow is then regulated either by another section control further downstream or by the hydraulic geometry and roughness of the channel downstream (i.e., channel control).

Figure 12.3: Example of Channel Control

A cross-section where no acceleration of flow occurs, or where the acceleration is not sufficient enough to prevent the passage of disturbances from the downstream to the upstream direction is called a channel control. A channel control consists of a combination of features throughout a reach downstream from a gauge. These features include channel size, shape, curvature, slope, and roughness. The rating curve in such case depends on the geometry and the roughness of the river downstream of the control (Figure 12.3). The length of the channel reach that controls a stage-discharge relationship can be extremely variable. The length of the reach affecting the rating curve depends on the normal or equilibrium depth h_e and on the energy slope S ($L \propto h_e/S$, where h_e follows from Manning's expression for flow $Q=K_m B h_e^{5/3} S^{1/2}$ (for a wide rectangular channel) so $h_e = (Q/K_m S^{1/2})^{3/5}$. The length of the channel that is effective as a control increases with discharge. Generally, the flatter the stream gradient, the longer the channel control reach.

At some stages, the stage-discharge relation may be governed by a combination of section and channel controls. This usually occurs for a short-range in the stage between section-controlled and channel-controlled segments of the rating.

Three types of controls can be recognized, depending on the channel and flow conditions:

- low flows are usually influenced by a section control;
- high flows are usually influenced by a channel control;
- medium flows can be influenced by both types of controls.

12.2.1.2 Artificial and natural controls

An artificial section control or structure control is one which has been specifically constructed to stabilize the relationship between stage and discharge and for which a theoretical relationship is available based on physical modelling. These include weirs and flumes discharging under free-flow conditions (Figure 12.4). Natural section controls include a ledge of rock across a channel, the brink of a waterfall, or a local constriction in width. All channel controls are 'natural'.

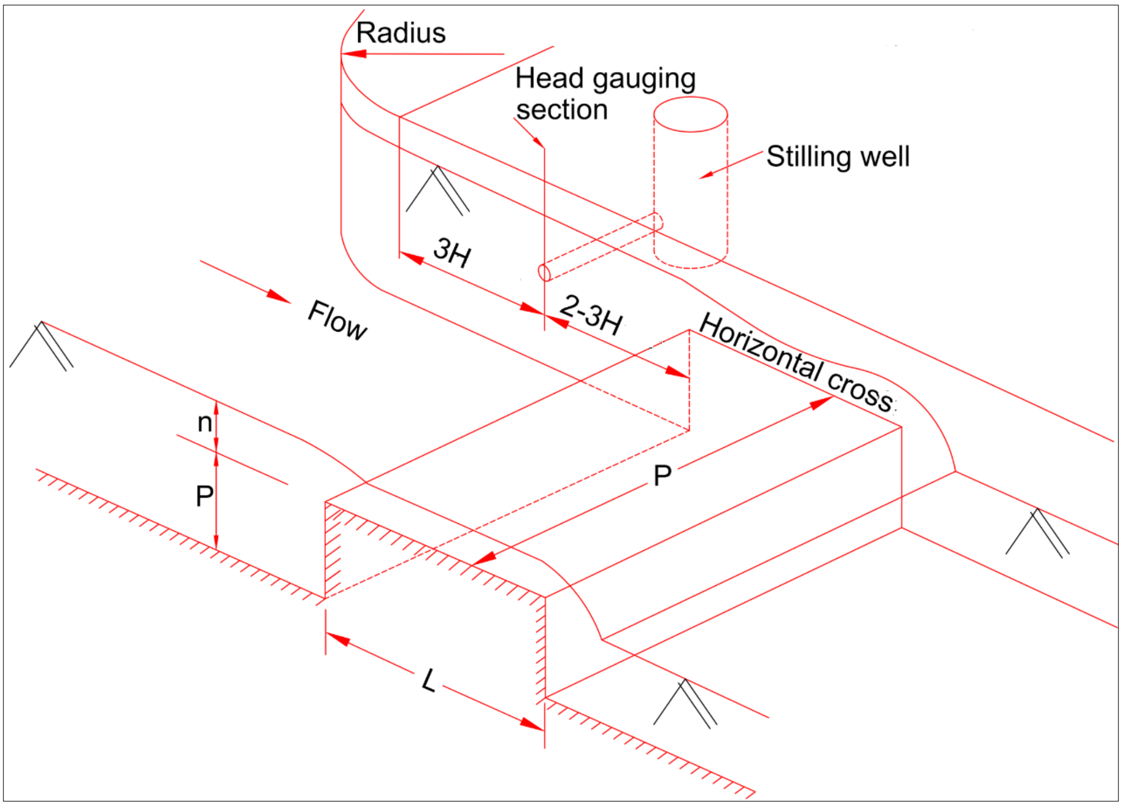


Figure 12.4: Example of an Artificial Control

12.2.1.3 Complete and compound controls

Y dependent both on the elevation and shape of the control and on the tailwater level.

12.2.1.4 Permanent and shifting controls

Where the geometry of a section and the resulting stage-discharge relationship does not change with time, it is described as a stable or permanent control (Figure 12.5). Shifting controls change with time. Section controls such as boulder, gravel or sand riffles which undergo periodic or near-continuous scour and deposition, or channel controls with erodible bed and banks comprise shifting controls. Shifting controls thus typically result from:

- scour and fill in an unstable channel;
- growth and decay of aquatic weeds; and,

- spilling over and ponding in areas adjoining the stream channel.

The amount of effort in gauging and maintenance cost to obtain a record of adequate quality is much greater for shifting controls than for permanent controls. Since rating curves for the unstable controls must be updated and/ or validated at frequent intervals, regular and frequent current meter measurements to estimate the discharge are required. In contrast, for stable controls, the rating curve can be established once and needs validation only occasionally. Since stage-discharge observations require significant effort and resources, it is always preferred to select a gauge site with a permanent section or structure control. However, this is not practicable in many cases and one has to be content with either channel control or a compound control, particularly for the alluvial rivers in the plains.

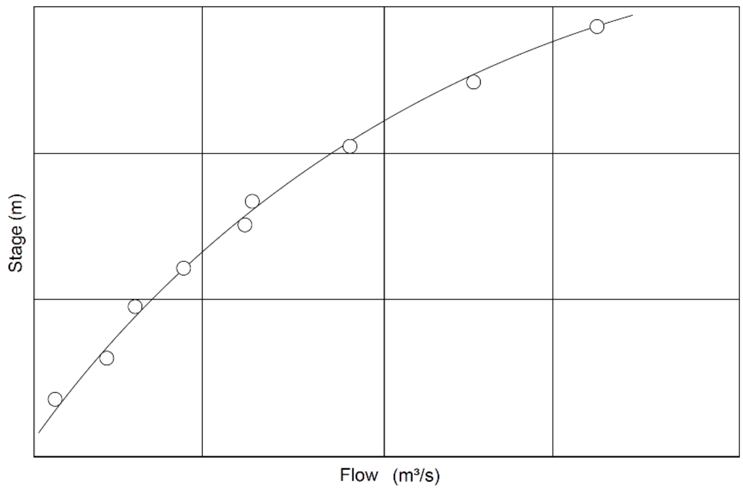


Figure 12.5: Permanent Control (adopted from Herschy, 2009)

12.3 Fitting of Rating Curves

12.3.1 General

A simple stage-discharge relation is one where discharge depends upon stage only. A complex rating curve occurs where additional variables such as the slope of the energy line or the rate of change of stage with time are required to define the relationship. The need for a particular type of rating curve can be ascertained by first plotting the observed stage and discharge data on a simple orthogonal plot. The scatter in the plot gives a fairly good assessment of the type of stage-discharge relationship required for the cross-section. Examples of the scatter plots obtained for various conditions have been described by Herschy (2009), and are shared below for better understanding. If there is negligible scatter in the plotted points and it is possible to draw a smooth single-valued curve through the plotted points, a simple rating curve is required. This is shown in Figure 12.5. However, if scatter is not negligible then it requires further probing to determine the cause of such high scatter. There are four distinct possibilities:

- **The station is affected by variable backwater conditions** due to tidal influences or high flows in a tributary joining downstream. In such cases, if the plotted points are annotated with the corresponding slope of energy line (surface slope for uniform flows) then a definite pattern can be observed. A smooth curve passing through those points having normal slopes at various depths is drawn first. It can then be seen that the points with greater variation in slopes from the corresponding normal slopes are located farther from the curve. This is shown in Figure 12.6 and Figure 12.7.

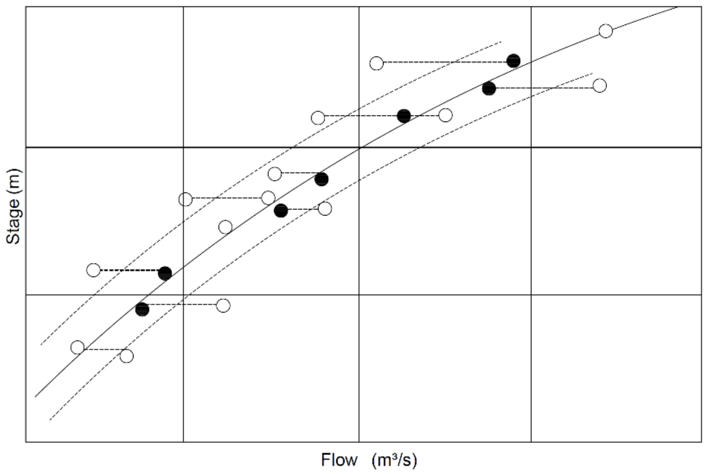


Figure 12.6: Rating Curve of Uniform Channel Affected by Variable Backwater (adopted from Herschy, 2009)

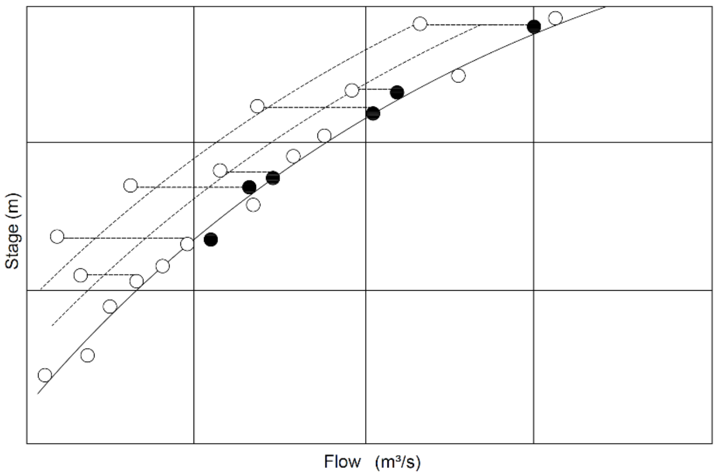


Figure 12.7: Rating Curve Affected by Variable Backwater (Submergence of Low-Water Control) (adopted from Herschy, 2009)

- **The stage-discharge rating is affected by variation in the local acceleration due to unsteady flow**. In such cases, the plotted points can be annotated with the corresponding rate of change of slope against time. A smooth curve (steady-state curve) passing through those points having the least values of rate of change of stage is drawn first. It can then be seen that all those points having positive values of rate of change of stage are towards the right side of the curve and those with negative values are towards the left of it. Also, the distance from the steady curve increases with the increase in the magnitude of the rate of change of stage. This is as shown in Figure 12.8.

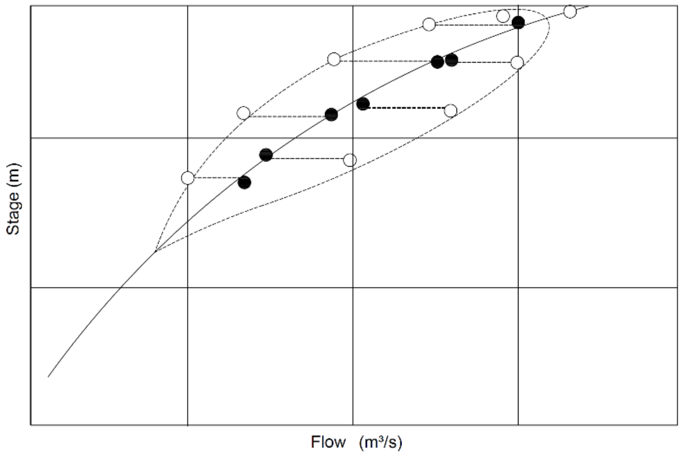


Figure 12.8: Rating Curve Affected by Unsteady Flow (adopted from Herschy, 2009)

- **The stage-discharge rating is affected by scouring of the bed or changes in vegetation characteristics.** Changes is sediment flow are usually the most common source of "noise" in the data, since the channel cross section is constantly changing between the wet and dry seasons. Shifting bed results in wide scatter of points on the graph. The changes are erratic and maybe progressive or may fluctuate from scouring in one event and deposition in another. A sample plot is shown in Figure 12.9.

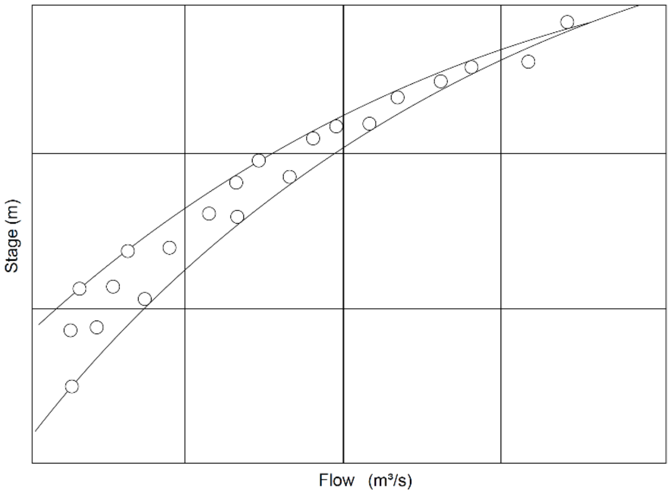


Figure 12.9: Stage-Discharge Relation Affected by Scouring and Fill (adopted from Herschy, 2009)

- **The growth of weeds** decreases the conveyance of the channel by increasing the roughness, with the result that the stage is increased for a given discharge.

The converse happens when the weeds die. Normally, the development of a family of stage–discharge curves corresponding to different conditions of weed growth presents the best means of gauging rivers with prolific weed growth, as shown in Figure 12.10.

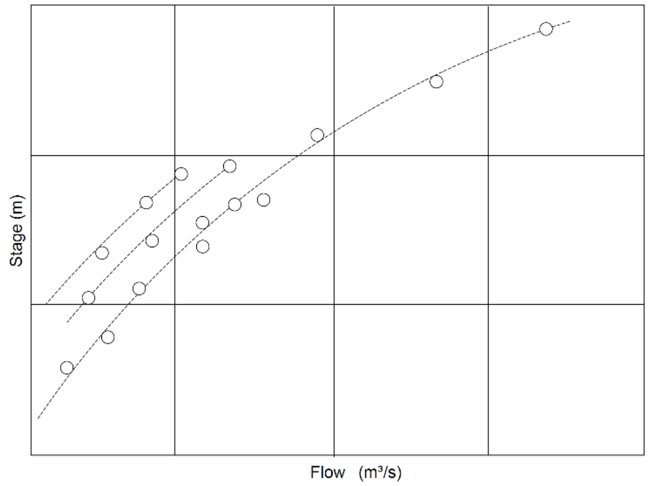


Figure 12.10: Stage-Discharge Relation Affected by Aquatic Vegetation Growth (adopted from Herschy, 2009)

In addition to the changes due to sediment regime, another likely source of the data scatter in the plots can be attributed to the observational errors. Such errors can occur due to adoption of non-standard procedures for stage-discharge observation.

The appropriate type of rating curve is fitted, based on the interpretation of scattering of the stage-discharge data. There are four main cases:

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

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

Rating curve with unsteady flow correction: If the flows are affected by the unsteady conditions, then the rating curve incorporating the unsteady flow effects is established. This requires information on the rate of change of stage against time corresponding to each stage-discharge data point.

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

12.3.2 Fitting of single-channel simple rating curve

A single-channel simple rating curve is fitted in those circumstances when the flow is contained in the main channel section and can be assumed to be fairly steady. There is no indication of any backwater affecting the relationship. The bed of the river also does not change significantly, so as create any shift in the stage-discharge relationship. The scatter plot of the stage and discharge data shows very little scatter if the observational errors are not significant. The scatter plot of stage-discharge data in such situations typically is as shown in Figure 12.1. The fitting of simple rating curves can conveniently be considered under the following headings:

- equations used and their physical basis
- determination of datum correction(s)
- number and range of rating curve segments
- determination of rating curve coefficients
- estimation of uncertainty in the stage-discharge relationship

12.3.2.1 Equations and their physical basis

The algebraic equation commonly fitted to the stage-discharge data in India is the **Power equation**:

$$\boldsymbol{Q = C(h + a)^b} \qquad \text{Equation 12.2}$$

where:

Q	discharge (m^3/sec)
h	measured water level (m)
a	water level (m) corresponding to Q = 0
C	coefficient derived for the relationship corresponding to the station characteristics
b	exponential constant

Taking logarithms of the power type equation results in a linear relationship of the form:

$$\log (Q) = \log (C) + b \log (h + a) \qquad \text{Equation 12.3}$$

or

$$Y = A + B X \qquad \text{Equation 12.4}$$

That is, if sets of discharge (Q) and the effective stage (h + a) are plotted on the double log scale, they should represent a straight line. Coefficients A and B of the straight-line fit are functions of a and b. Since values of a and b can vary at different depths owing to changes in physical characteristics (effective roughness and geometry) at different depths, one or more straight lines will fit the data on a double log plot. This is illustrated in Figure 12.11, which shows a distinct break like fit for the two water level ranges. A plot of the cross-section at the gauging section is also often helpful to interpret the changes in the characteristics at different levels.

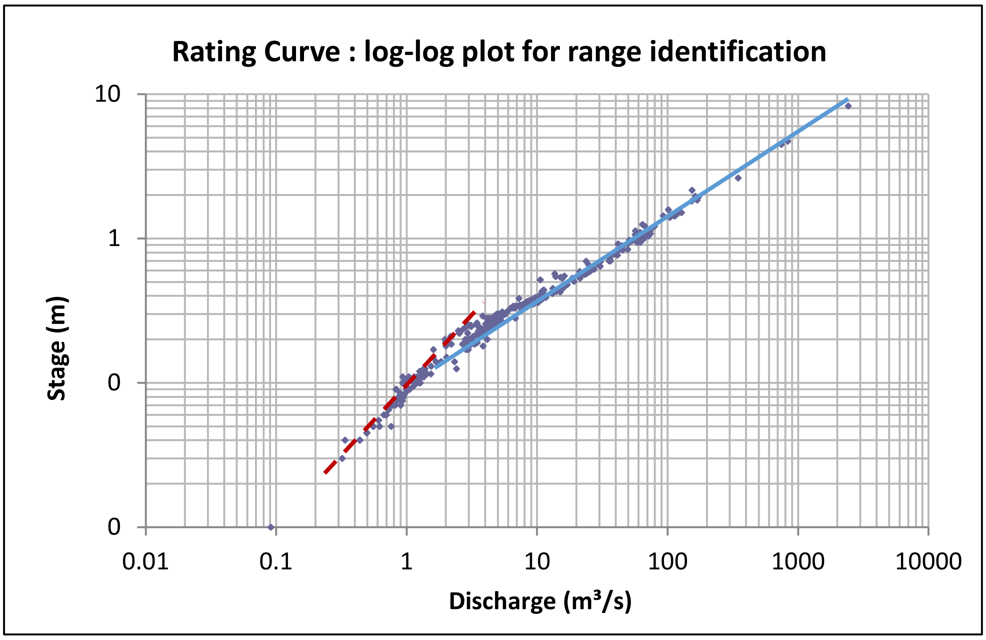


Figure 12.11: Double Logarithmic Plot of Rating Curve Showing a Distinct Break

The relationship between rating curve parameters and physical conditions is also evident if the power equation is compared with Manning's equation for determining discharge under steady flow situations. Manning's equation can be written as:

$$\boldsymbol{Q} = \frac{\mathbf{1}}{\boldsymbol{n}}\mathrm{A}\boldsymbol{R}^{\frac{2}{3}}\boldsymbol{S}^{\frac{1}{2}} = \left(\frac{\mathbf{1}}{\boldsymbol{n}}\boldsymbol{S}^{\frac{1}{2}}\right)\left(\mathrm{A}\boldsymbol{R}^{\frac{2}{3}}\right) \qquad \text{Equation 12.5}$$

= function of (roughness & slope) & (depth & geometry)

Hence, the coefficients a, c are some measures of roughness and geometry of the control and b is a measure of the geometry of the section at various depths. The value of coefficient b for various geometrical shapes are as follows:

For rectangular shape	:	about 1.5
For triangular or semi-circular shape	:	about 2.5
For parabolic shape	:	about 2.0
For irregular shape	:	1.6 to 1.9

Changes in the channel resistance and slope with stage, however, will affect exponent b. The net result of these factors is that the exponent for relatively wide rivers with channel control will vary between about 1.3 to 1.8. For relatively deep narrow rivers with section control, the exponent will commonly be greater than 2 and sometimes exceed a value of 3. Note that high values of the exponent are sometimes found (>5) for compound channels with the flow over the floodplain or braided channels over a limited range of water levels.

12.3.2.2 Determination of datum correction (a)

The datum correction (a) corresponds to that value of water level for which the flow is zero. From eq. (2) it can be seen that for Q = 0, (h + a) = 0 which means: a = -h.

Physically, this level corresponds to zero flow condition at the control effective at the measuring section. The exact location of effective control is easily determined for artificial controls or where the control is well defined by a rock ledge forming a section control. For the channel controlled gauging station, the level of the deepest point opposite the gauge may give a reasonable indication of datum correction. In some cases, identification of datum correction may be impractical especially where the control is compounded and it shifts progressively downstream at higher flows. Note that the datum correction may change between different controls and different segments of the rating curve. For upper segments the datum correction is effectively the level of zero flow had that control applied down to zero flow; it is thus a nominal value and not always possibly to ascertain physically.

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

- trial and error procedure
- arithmetic procedure
- computer-based optimization

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

Trial and error procedure

This method has been commonly used before the advent of the computer-based methods. In this method, the stage-discharge observations are plotted on double log plot and the best fit line is fitted through them. This fitted line is usually curved. However, as explained above, if the stages are adjusted for zero flow condition, i.e., datum correction a, then this line should be straight. This is achieved by taking a trial value of "a" and plotting (h + a), the adjusted stage, and discharge data on the same double log plot. It can be seen that if the unadjusted stage-discharge plot is concave downwards then a positive trial value of "a" is needed to make it a straight line. And conversely, a negative trial value is needed to make the line straight if the curve is concave upwards. A few values of "a" can be tried to attain a straight line fit for the plotted points of adjusted stage-discharge data. The procedure is illustrated in Figure 12.12. This procedure was slow but quite effective when done manually. However, making use of general spreadsheet software (having graphical provision) for such trial-and-error procedure can be very convenient and faster now.

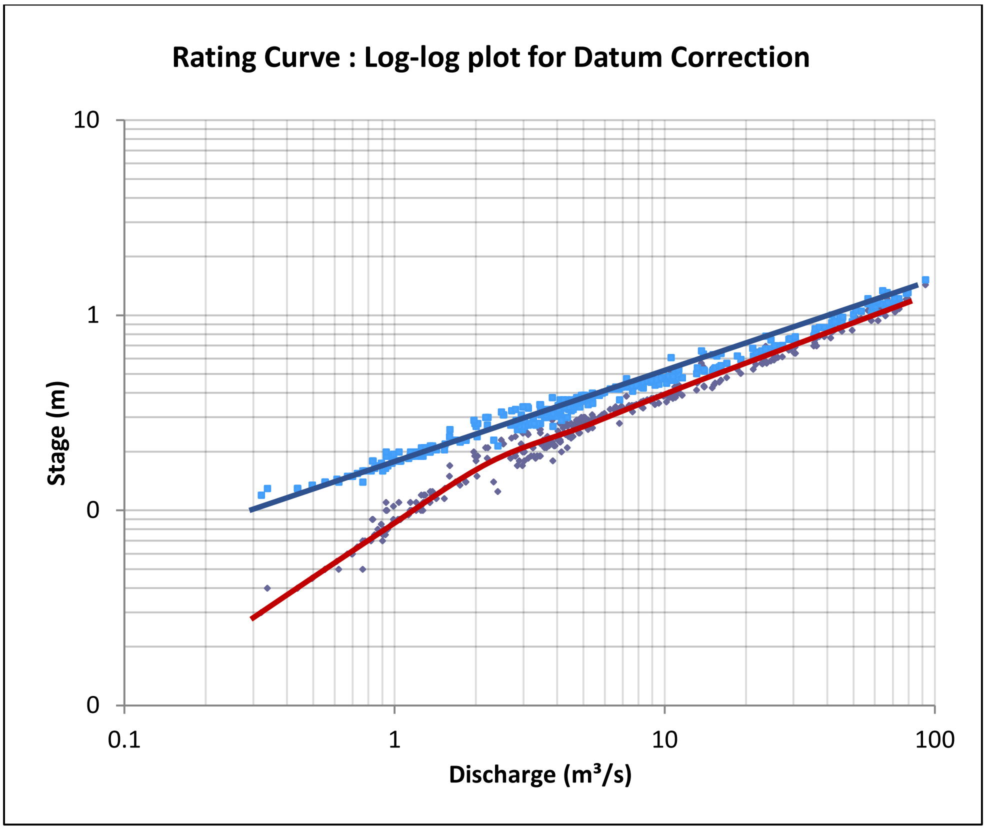


Figure 12.12: Determination of Datum Correction by Trial and Error

Arithmetic procedure:

This procedure is based on expressing the datum correction "a" in terms of observed water levels. This is possible by way of elimination of coefficients b and c from the power type equation between gauge and discharge using simple mathematical manipulation. From the median curve fitting the stage-discharge observations, two points are selected in the lower and upper range (Q_1 and Q_3) whereas the third point Q_2 is computed from $Q_2^2 = Q_1 \times Q_3$, such that:

$$\frac{Q_1}{Q_2} = \frac{Q_2}{Q_3} \qquad \text{Equation 12.6}$$

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

$$\frac{C\,(h_1+a)}{C\,(h_2+a)} = \frac{C\,(h_2+a)}{C\,(h_3+a)} \qquad \text{Equation12.7}$$

Which yields:

$$a = \frac{h_2^2 - h_1 h_3}{h_1 + h_3 - 2h_2} \qquad \text{Equation 12.8}$$

From this equation, an estimated value of "a" can be obtained directly.

Optimisation procedure:

This procedure is suitable for automatic data processing using a computer and "a" is obtained by an automated fitter procedure that minimizes the sum of squared errors. The first trial value of the datum correction "a" is either input by the user based on the field survey. Next, this first estimate of "a" is varied (usually within 2 m) to obtain a minimum mean square error in the fit. This is a purely mathematical procedure and probably gives the best results based on observed stage-discharge data but it is important to make sure that the result is confirmed wherever possible by physical explanation of the control at the gauging location. The procedure is repeated for each segment of the rating curve.

The datum for which the prediction efficiency is closest to 1 is to be chosen as the gauge-discharge relationship.

Efficiency = 1 - (Remaining variance)/ (Initial variance)

Or, $E = 1 - \frac{\sum(Y_i - \hat{Y_i})^2}{\sum(Y_i - \bar{Y})^2}$

where Y_i = observed values of discharge

$\hat{Y_i}$ = computed values of discharge

$\bar{Y}$ = mean value of discharge

12.3.2.3 Number and ranges of rating curve segments:

After the datum correction "a" has been established, the next step is to determine whether the rating curve is composed of one or more segments. Rating curves usually have a breakpoint, which is around the stage at which the river spreads out of its banks, or it could be at a lower stage if the river bed cross-section changes dramatically. Above that stage, the river does not rise as fast, given that other conditions remain constant. This can be seen by a change in slope in the rating curve. The 'h-a' and discharge data are plotted on a log-log scale as a scatter plot. The plot is inspected by the user for breaking points. The value of h at breaking point indicates a change in the rating curve. The number of water level ranges for which different rating curves are to be established is thus noted. The simple scatter plot between concurrent observations of stage and discharge over a period covering the defined range of stages can be used to ascertain the type of stage-discharge relationship.

A rating curve is established between concurrent observations of stage and discharge over a period covering the defined range of stages. The type of rating curve is ascertained by plotting the observed stage and discharge data on a simple plot. The scatter of the plot will indicate the type of stage-discharge relationship. If it is possible to draw a smooth curve through the plotted points and the scatter is little, a simple rating curve is fitted. More frequently, separate curves are developed for each range of observations. For example, Figure 12.11 shows that two separate rating curves are required for the two ranges of water level – one up to level "h_1" and second from "h_1" onwards. The rating equation for each of these segments is then established and the breaking points between segments are checked by computer analysis.

12.3.2.4 Determination of rating curve coefficients:

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

$$\widehat{Y} = \alpha - \beta X \qquad \text{Equation12.9}$$

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

$$\boldsymbol{E = \sum_{i=1}^{N}\left(Y_i - \widehat{Y_i}\right)^2 = \sum_{i=1}^{N}(Y_i - \alpha - \beta X_i)^2} \qquad \text{Equation 12.10}$$

Here i denotes the individual observed point, and N is the total number of observed stage-discharge data. Since this error is to be minimal, the slope of partial derivatives of this error with respect to the constants must be zero. In other words:

$$\frac{\delta E}{\delta \alpha} = \frac{\{\sum_{i=1}^{N}(Y_i - \alpha - \beta X_i)^2\}}{\delta \alpha} = 0 \qquad \text{Equation 12.11}$$

and

$$\frac{\delta E}{\delta \beta} = \frac{\{\sum_{i=1}^{N}(Y_i - \alpha - \beta X_i)^2\}}{\delta \beta} = 0 \qquad \text{Equation 12.12}$$

This results in two algebraic equations of the form:

$$\sum_{i=1}^{N} Y_i - \alpha N - \beta \sum_{i=1}^{N} X_i = 0 \qquad \text{Equation 12.13}$$

and

$$\sum_{i=1}^{N}(X_i Y_i) - \alpha \sum_{i=1}^{N} X_i - \beta \sum_{i=1}^{N}(X_i)^2 \qquad \text{Equation 12.14}$$

All the quantities in the above equations are known except α and β. Solving the two equations yield:

$$\beta = \frac{N\sum_{i=1}^{N}(X_i Y_i) - (\sum_{i=1}^{N} X_i)(\sum_{i=1}^{N} Y_i)}{N\sum_{i=1}^{N}(X_i)^2 - \left(\sum_{i=1}^{N} X_i\right)^2} \qquad \text{Equation 12.15}$$

and

$$\alpha = \frac{\sum_{i=1}^{N} Y_i - \beta \sum_{i=1}^{N} X_i}{N} \qquad \text{Equation 12.16}$$

The value of coefficients c and b of power type equation can then be finally obtained as:

$b = \beta$ and $c = 10^{\alpha}$ Equation 12.17

Reassessment of breaking points

The first estimate of the water level ranges for different segments of the rating curve is obtained by visual examination of the cross-section changes and the double log plot. However, exact limits of water levels for various segments are obtained by computerised analysis of the intersection of the fitted curves in the adjoining segments.

Considering the rating equations for two successive water level ranges be given as $Q = f_{i-1}(h)$ and $Q = f_i(h)$ respectively, let the upper boundary for the estimation of f_{i-1} be denoted by hu_{i-1} and the lower boundary of f_i by hl_i. To force the intersection between f_{i-1} and f_i to remain within certain limits, it is necessary to choose: $hu_{i-1} > hl_i$. That is, the intersection of the rating curves of the adjoining segments should be found to be situated within this overlap. This is illustrated in Figure 12.13. If the intersection falls outside the selected overlap, then the intersection is estimated for the least difference between $Q = f_{i-1}(h)$ and $Q = f_i(h)$. Preferably the boundary width between hu_{i-1} and hl_i is widened and the curves refitted.

A graphical plot of the fit of the derived equations to the data must be inspected before accepting them.

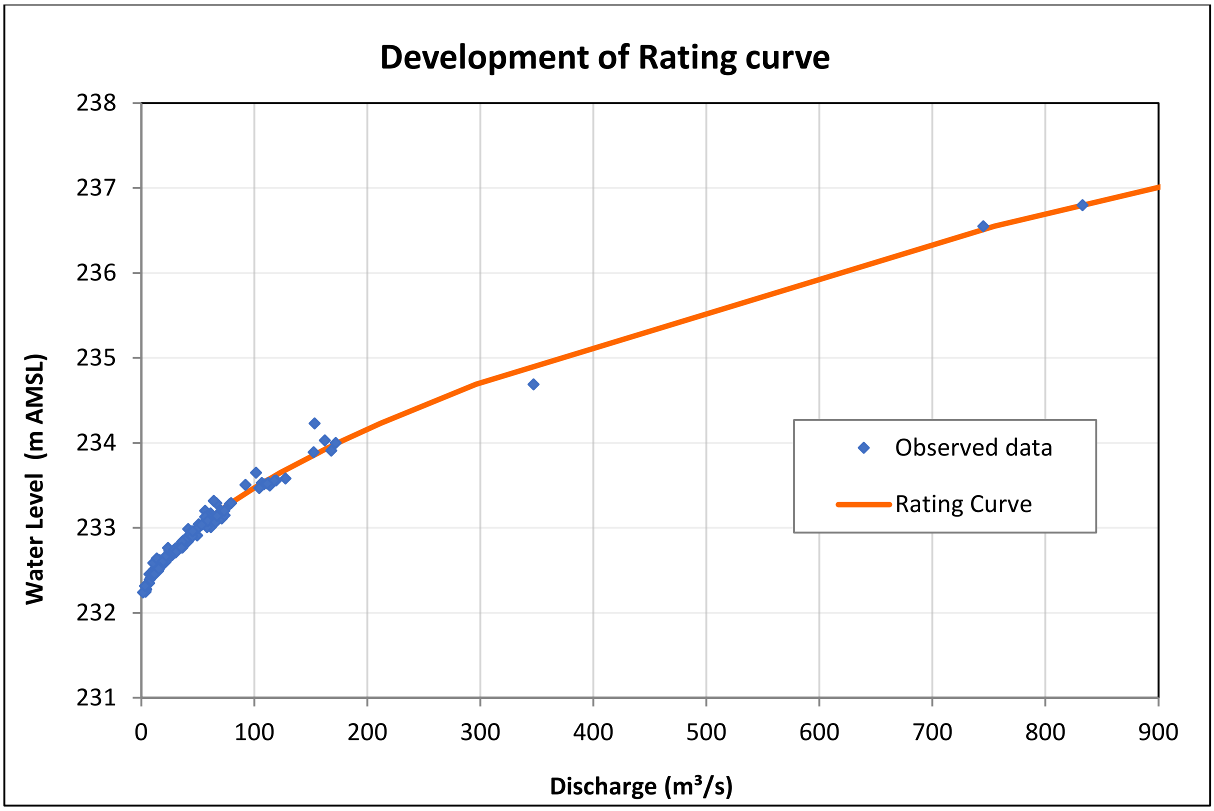


Figure 12.13: Rating Curve Fitted Using 2 Segments

12.3.2.5 Estimation of uncertainty in the stage-discharge relationship:

For the stage-discharge relationship, the standard error of estimate (S_e) is a measure of the dispersion of observations about the mean of the relationship. The standard error is expressed as:

$$S_e = \sqrt{\frac{\sum(\Delta Q_i - \overline{\Delta Q})^2}{N-2}} \quad \text{Equation 12.18}$$

Here, ΔQ_i is the measure of the difference between the observed (Q_i) and computed (Q_c) discharges and can be expressed in absolute and relative (percentage) terms respectively as:

$$\Delta Q_i = Q_i - Q_c \quad \text{Equation 12.19}$$

$$\Delta Q_i = \frac{Q_i - Q_c}{Q_i} * 100\% \quad \text{Equation 12.20}$$

Standard error expressed in relative terms helps in comparing the extent of fit between the rating curves for different ranges of discharges. The standard error for the rating curve can be derived for each segment separately as well as for the full range of data.

In general, an acceptable fit will have 95% of all observed stage-discharge data within t × S_e from the fitted line where:

Student's $t \cong 2$ where $n > 20$, but is increasingly large for smaller samples. Standard t tables available in statistical texts may be referred to for this purpose.

The stage-discharge relationship, being a best fit line, provides a better estimate of discharge than any of the individual observations, but the position of the line is also subject to uncertainty, expressed as the standard error of the mean relationship (S_{mr}) which is given by:

$$S_{mr} = S_e \sqrt{\left(\frac{1}{n} + \frac{(P_i - \bar{P})^2}{S_p^2}\right)} \quad \text{Equation 12.21}$$

And, $CL_{95\%} = \pm t S_{mr}$

where	t	=	Student's t-value at 95% probability
	P_i	=	ln (h_i + a)
	S_p^2	=	variance of P
	$CL_{95\%}$	=	95% confidence limits

This equation on standard error provides a single error value for the logarithmic relationship. The 95% confidence limits can be displayed as two parallel straight lines on either side of the relationship developed for the mean. By contrast, S_{mr} is calculated for each observation of *(h + a).* The limits are therefore curved on each side of the stage-discharge relationship and are minimum at the mean value of *ln(h + a)* where the S_{mr} relationship reduces to:

$$S_{mr} = \pm S_e / n^{1/2} \qquad \text{Equation 12.22}$$

Thus, with n = 25, the standard error S_{mr} of the mean relationship is approximately 20% of S_e, indicating the advantage of the fitted relationship over the use of the individual gauging.

12.3.3 The compounded channel rating curve

If the flood plains carry the river flow over the full cross-section, the discharge consists of two parts – flow through the river and flow through the flood plain. A compounded channel is a channel that has a flood plain section to accommodate the flood wave. The rating curve changes significantly as soon as the floodplain at level $(h - h_1)$ is flooded. The rating curve for this situation of a compound channel is determined by considering the flow through the floodplain portion separately.

$$Q = C_r(H - a)^{\beta} + C_f(H - a_1)^{\beta_1} \qquad \text{Equation 12.23}$$

Where, Q = stream discharge (m^3/s)

H = Gauge Elevation (m)

a = gauge reading corresponding to zero discharge / zero of gauge level (m)

a_1 = Flood Plain level (m)

C_r and β are rating curve coefficients for the river section; C_f and β_1 are rating curve coefficients for the flood plain section. This is illustrated in Figure 12.14. The rating curve changes significantly as soon as the flood plain at level $(h - h_1)$ is flooded, especially if the ratio of the floodplain storage width B to the width of the river bed B_r is large. The rating curve for this situation of a compound channel is determined by considering the flow through the floodplain portion separately. This is done to avoid large values of the exponent b and extremely low values for the parameter C in the power equation for the rating curve of the main channel portion.

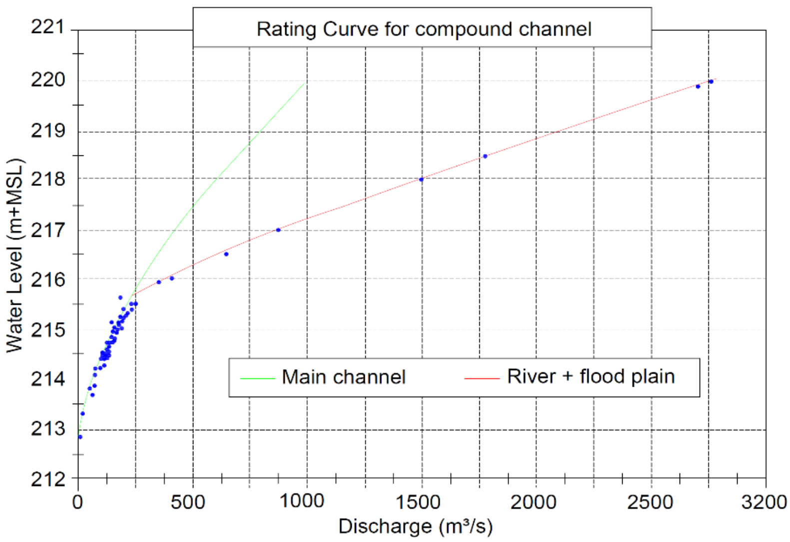


Figure 12.14: Example of a Rating Curve for Compound Cross-Section

The last water level range to be considered for fitting a rating curve is taken care of for the flood plain water levels. First, the river discharge Q_r is computed for this last interval by using the parameters computed for the last but one interval. Then a temporary flood plain discharge Q_f is computed by subtracting Q_r from the observed discharge (O_{obs}) for the last water level interval, i.e.

$$Q_f = Q_{obs} - Q_r \qquad \text{Equation 12.24}$$

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

The rating curve presented in Figure 12.14 for Jhelum River at Rasul reads:

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

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

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

12.3.4 Limitations of Simple and Compounded Rating Curves

To develop a rating curve, a series of streamflow measurements (current meter readings) are plotted versus the corresponding stage, and a smooth curve is drawn through the points. There can be significant scatter around this curve. Therefore, the discharge read from the rating curve is the most likely value, but it could be different from the measured value. Also, as the rating curves are usually developed with a limited number of stage/ discharge measurements where the measurements of high flows are rare. There can be significant errors in the rating curves at high levels, especially around record high level flows.

Stage discharge relationship for unstable controls poses problems when the segments of a stage-discharge relationship change position occasionally. Variable backwater caused mainly by downstream reservoirs, dams, dense vegetation, tides influence the flow at gauging-station control and affect the stage-discharge relation.

Defining the flood plain storage at times when rivers are flooded is a complex process. The interaction of flow between the main channel and flood plain often results in a flow pattern that is difficult to define at the gauging section. When water surface slope changes either due to rapidly rising or rapidly falling water levels in a channel control reach causing hysteresis in the rating curve, it adds to the complexities.

12.3.5 Rating curve with backwater correction

When the control at the gauging station is influenced by other controls downstream, the unique relationship between the stage and discharge at the gauging station may not be maintained. Backwater is an important consideration in streamflow site selection, and sites having backwater effects should be avoided wherever possible. However, backwater effects may emerge as a result of downstream development and urbanization that was initiated after the gauging station had already been commissioned. Typical examples of backwater effects on gauging stations and the rating curve are as follows:

- by regulation of watercourse downstream.
- level of water in the main river at the downstream confluence
- level of water in a downstream reservoir
- variable tidal effect occurring downstream of a gauging station
- downstream flow constriction with a variable-capacity at any level due to the growth of vegetation
- rivers with the return of overbank flow

Backwater effect due to variable controls downstream of the gauging station influences the water surface slope at the station for a given stage. When the backwater from the downstream control results in a lowering of the water surface slope, a smaller discharge passes through the gauging station for the same stage. On the other hand, if the surface slope increases, as in the case of sudden drawdown through a regulator downstream, a greater discharge passes for the same stage. The presence of backwater does not allow the use of a simple unique rating curve.

Variable backwater causes variation of the energy slope, which affects the flow for the same stage.

Discharge is a function of both stage and energy slope, and this relationship is usually termed as slope-stage-discharge relationship. The stage is measured continuously at the main gauging station. The slope is estimated by continuously observing the stage at an additional gauge station, called the auxiliary gauge station. The auxiliary gauge station is established some distance downstream of the main station. Time synchronization in the observations at the gauges is necessary for the precise estimation of the slope. The distance between these gauges is chosen in a way that it provides an adequate representation of the water surface slope at the main station, and at the same time, the uncertainty in the estimation is also reduced. When both the main and auxiliary gauges are set to the same datum, the difference between the two stages directly gives the fall in the water surface, which can be approximated as the measure of the surface slope. This fall is taken as the third parameter in the relationship, and the rating is therefore also called stage-fall-discharge relation.

Discharge expressed using the Manning's equation is:

$$Q = K_m R^{2/3} S^{1/2} A \qquad \text{Equation 12.25}$$

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

$$\frac{Q_m}{Q_r} = \left(\frac{S_m}{S_r}\right)^p = \left(\frac{F_m}{F_r}\right)^p \qquad \text{Equation 12.26}$$

Where,

Q_m	is the measured (backwater affected) discharge
Q_r	is a reference discharge
S_m	is the measured slope
S_r	is a reference slope
F_m	is the measured fall
F_r	is a reference fall
p	is a power parameter between 0.4 and 0.6

Based on the exponent used for slope in the Manning's equation given above, the exponent *p* is expected to be around ½. The fall (F) or the slope (S = F/L) is obtained by observing the water levels at the main and the auxiliary gauge. Since there is no assurance that the water surface profile between these gauges is a straight line, the effective value of the exponent can be different from ½, and must, therefore, be determined empirically.

An initial plotting of the stage-discharge relationship (prepared either manually or by computer) with values of fall against each observation will show whether the relationship is affected by variable slope. It will also show whether this occurs at all stages or only at stages when the fall reduces below a particular value. In the absence of any channel control, the discharge would be affected by the variable fall at all the times, and correction is applied by the constant fall method. When the discharge is affected only when the fall reduces below a given value, the normal (or limiting) fall method is used.

12.3.5.1 Constant fall method

The constant fall method is applied when the stage-discharge relationship is affected by variable fall at all times and for all stages. The fall applicable to each discharge measurement is determined, and plotted with each stage-discharge observation on the plot. If the observed falls do not vary too much, an average value (reference fall or constant fall) F_r is selected.

For computer computation, the procedure is simplified by mathematical fitting and optimization. First, as before, a reference (or constant) fall (F_r) is selected from amongst the most frequently observed falls.

A rating curve, between stage h and the reference discharge (Q_r), is then fitted directly by estimating:

$$Q_r = Q_m \left(\frac{F_r}{F_m}\right)^p \qquad \text{Equation 12.27}$$

where p is optimised between 0.4 and 0.6 based on minimization of standard errors.

The discharge at any time, corresponding to the measured stage h and fall F_m is then calculated by first obtaining Q_r from the above relationship and then calculating discharge as:

$$Q = Q_r \left(\frac{F_m}{F_r}\right)^p \qquad \text{Equation 12.28}$$

A special case of a constant fall method is the unit fall method, in which the reference fall is assumed to be unity. This simplifies the calculations and thus is suitable for manual method.

12.3.5.2 Normal Fall Method

The normal or limiting fall method is used when the backwater is not present at the station at times. Examples are when a downstream reservoir is drawn down or when there is low water in a downstream tributary or the main river.

The computerised procedure considerably simplifies computation, and is as follows:

- Prepare the backwater-free rating curve using selected current meter readings (the Q_r -h relationship).

- Using values of Q_r and F_r derived from:

$$F_r = F_m \left(\frac{Q_r}{Q_m}\right)^{\frac{1}{p}} \qquad \text{Equation 12.29}$$

The value of the parameter p is optimised between 0.4 and 0.6. A parabola is fitted to the reference fall with the stage *(h)* by using:

$$F_r = a + bh + ch^2 \qquad \text{Equation 12.30}$$

The discharge corresponding to the measured stage h and fall F_m is then calculated by:

- obtaining F_r for the observed h from the parabolic relation between h and F_r
- obtaining Q_r from the backwater free relationship established between h and Q_r
- then calculating discharge corresponding to measured stage h as:

$$Q = Q_r \left(\frac{F_m}{F_r}\right)^{p} \qquad \text{Equation 12.31}$$

12.3.6 Rating curve with unsteady flow correction

Gauging stations not subjected to variable slope due to backwater effects may still be affected by variations in the water surface slope due to high rates of change in the stage. This occurs when the flow is highly unsteady and the water level is changing rapidly. At stream gauging stations located in a reach where the slope is very flat, the stage-discharge relationship is frequently affected by the superimposed slope of the rising and falling limb of the passing flood wave. During the rising stage, the velocity and discharge are normally greater than they would be for the same stage under steady flow conditions. Similarly, during the falling stage, the discharge is normally lower compared to any given gauge height when the stage is constant. This is because the approaching velocities in the advancing portion of the wave are larger than that in a steady uniform flow at the corresponding stage. During the receding phase of the flood wave, the converse situation occurs with reduced approach velocities yielding lower discharges than that in the equivalent steady-state case.

Thus, the stage-discharge relationship for an unsteady flow will not be single-valued as in the case of steady flow, but it will be a looped curve as shown in the example below. The looping in the stage-discharge curve is also called hysteresis in the stage-discharge relationship. From the curve, it can be easily seen that at the same stage, more discharge passes through the river during rising stages than in the falling ones.

12.3.6.1 Application

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

$$Q_m = Q_r\sqrt{\left(1+\frac{1}{cS_0}\frac{dh}{dt}\right)}$$ Equation 12.32

Where Q_m = measured discharge

Q_r = estimated steady-state discharge from the rating curve wave velocity (celerity) measures

c = wave velocity (celerity)

S_0 = energy slope for steady-state flow

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

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

Rearranging the above equation:

$$\frac{1}{cS_0} = \frac{\left(\frac{Q_m}{Q_r}\right)^2 - 1}{\frac{dh}{dt}}$$ Equation 12.33

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

$$\frac{1}{cS_0} = a + bh + ch^2$$ Equation 12.34

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

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

- Measured discharge is plotted against the stage, and beside each plotted point the value of *dh/dt* for the measurement is noted (+ or -)
- A trial Q_S rating curve representing the steady flow condition where *dh/dt* equals zero is fitted to the plotted discharge measurements.

- A steady-state discharge Q_r is then estimated from the curve for each discharge measurement and Q_m, Q_r and dh/dt are used to compute the corresponding values of the adjustment factor $1/cS_0$
- Computed values of $1/cS_0$ are then plotted against the stage and a smooth (parabolic) curve is fitted through the plotted points

For obtaining unsteady flow discharge from the steady rating curve, the following steps are followed:

- obtaining the steady-state flow Q_r for the measured stage h
- obtaining factor ($1/cS_0$) by substituting the stage h in the parabolic relation between the two
- obtaining (dh/dt) from stage-discharge observation timings or continuous stage records
- substituting the above three quantities to obtain the correct unsteady flow discharge

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

12.3.7 Rating relationships for stations affected by shifting control

For site selection, it is a desirable property of a gauging station to have a stable control. However, no such ideal section may exist in the reach for which flow measurement is required. Therefore, the gauging station location may be selected at a place where it is subject to shifting control. Shifts in the control occur frequently in alluvial sand-bed streams. However, shift may occur even in stable stream channels, particularly at low flows because of weed growth in the channel, or as a result of debris caught in the control section.

In alluvial sand-bed streams, the stage-discharge relation usually changes with time, either gradually or abruptly. This is either due to scour and silting in the channel or because of the moving dunes and sand bars. The extent and frequency with which the changes occur depend on typical bed material size at the control and velocities at the station. In the case of controls consisting of cobble or boulder-sized alluvium, the control and hence the rating may change only during the highest floods. In contrast, in sand-bed rivers, the control may shift gradually even in low to moderate flows. Channels with intermediate conditions change frequently during the monsoon but remain stable during longer periods of recession.

For sand bed channels, the stage-discharge relationship varies not only because of the changing cross-section due to scouring or deposition but also because of changing roughness with different bedforms. Bed configurations occurring with increasing discharge are ripples, dunes, plane bed, standing waves, chute and pool (Figure 12.15). The resistance of flow is the greatest in the dunes range. When the dunes are washed out and the sand is rearranged to form a plane bed, there is a marked decrease in bed roughness and resistance to flow, causing an abrupt discontinuity in the stage-discharge relation. Fine sediment present in water also influences the configuration of sand-bed, and thus, the resistance to flow. Changes in water temperature in sand-bed channels may also alter the bed form, and hence the roughness and resistance to flow. The viscosity of water will increase with lower temperature and thereby mobility of the sand will increase.

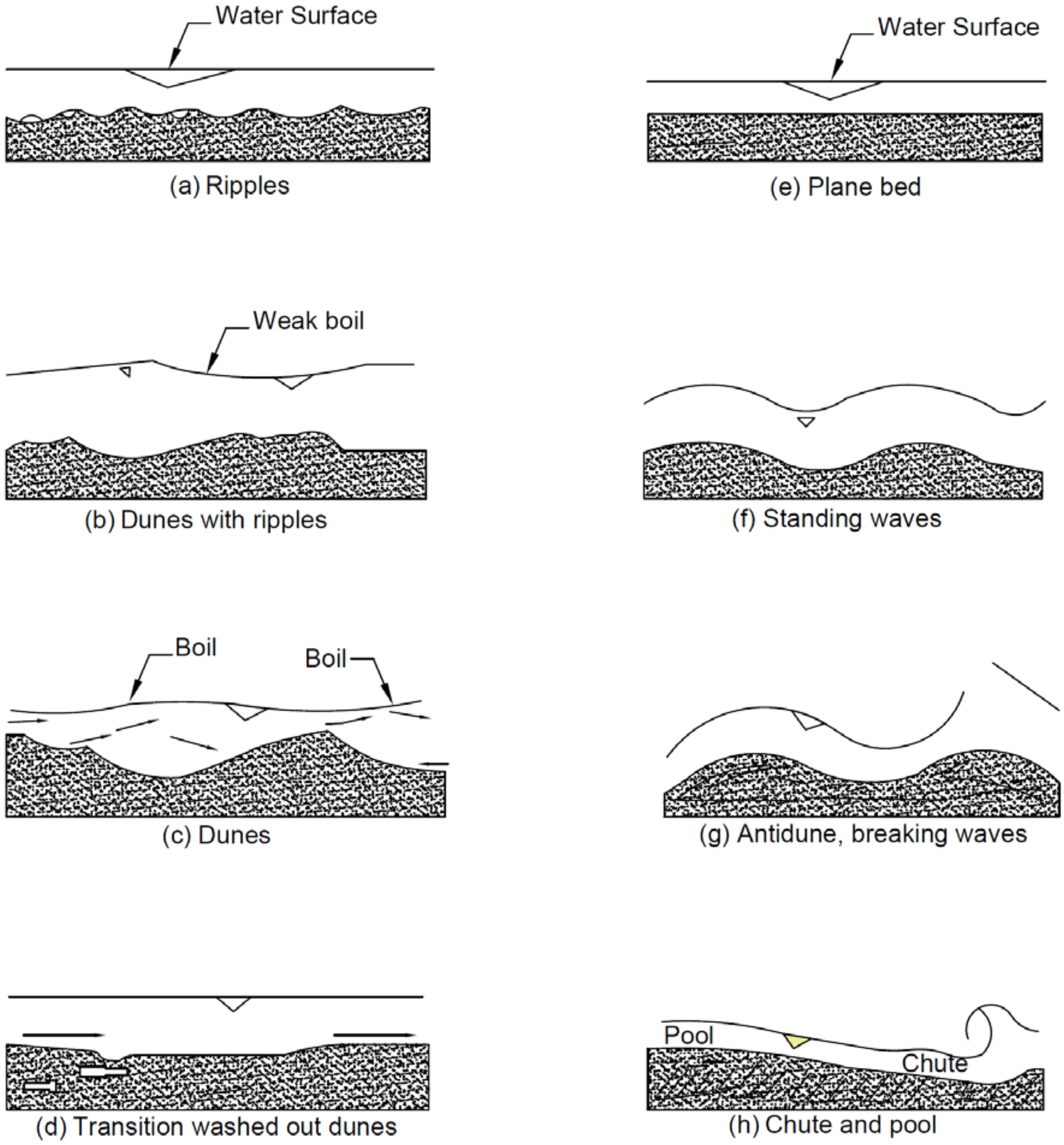


Figure 12.15: Bed and Surface Configurations in Sand-Bed Channels (from Herschy, 2009)

For alluvial streams where neither bottom nor sides are stable, a plot of the stage against discharge will very often scatter widely and thus be indeterminate (Figure 12.16). However, the hydraulic relationship becomes apparent by changing the variables. The effect of variation in bottom elevation and width is eliminated by replacing stage by mean depth (hydraulic radius) and discharge by mean velocity respectively. Plots of mean depth against mean velocity are useful in the analysis of stage-discharge relations, provided the measurements are referred to the same cross-section.

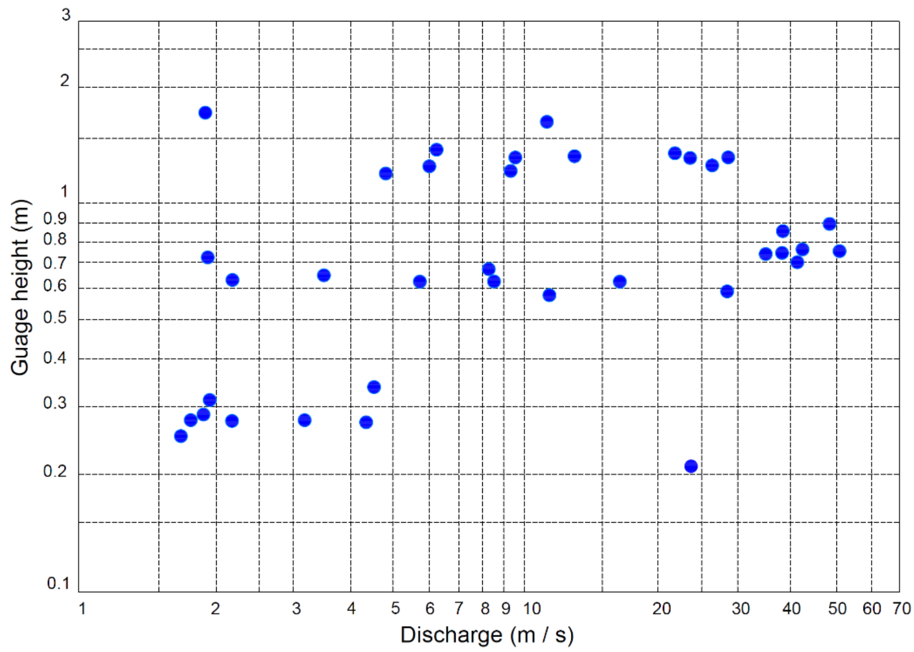


Figure 12.16: Plot of Discharge Against the Stage for a Sand-Bed Channel with Indeterminate Stage-Discharge Relationship

These plots will identify the bed-form regime associated with each discharge measurement (Figure 12.17). Thus, measurements associated with respective flow regimes, upper or lower, are considered for establishing separate rating curves. Information about bed-forms may be obtained by visual observation of water surfaces and noted for reference to develop discharge ratings.

Four possible approaches are available, depending on the severity of scouring and on the frequency of gauged records:

- Fitting a simple rating curve between scour events
- Varying the zero or shift parameter
- Applying Stout's shift method
- Estimating the flow by using daily records

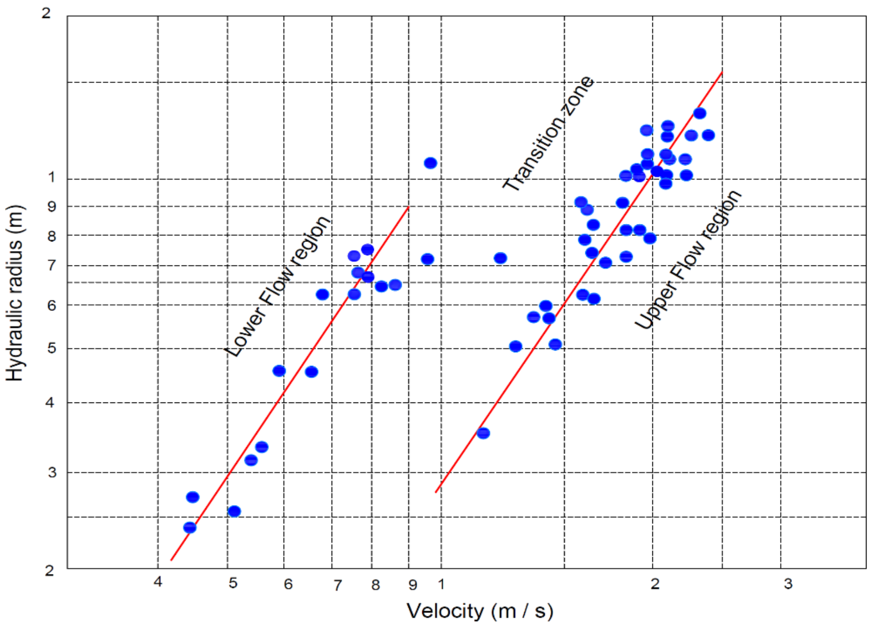


Figure 12.17: Relation of Mean Velocity to the Hydraulic Radius of the Channel

12.3.7.1 Fitting a simple rating curve between scour events

Where the plotted rating curve shows long periods of stability punctuated by infrequent flood events which cause channel adjustments, the standard procedure of fitting a simple logarithmic equation of the form $Q = c_1(h + a_1)^{b_1}$ should be applied to each stable period. This is possible only if there are sufficient gauge records in each period throughout the range of the stage.

To identify the date of change from one rating to the next, the records are plotted with their date or number sequence. The interval in which the change occurred is where the position of sequential plotted records moves from one line to the next. The analyst should then inspect the gauge observation record for a flood event during the period and apply the next rating from that date.

Notes from the Field Record Book or station log must be available while inspection and stage-discharge processing is carried out. This provides further information on the nature and timing of the event and confirms that the change was due to shifting control rather than to damage or adjustment to the staff gauge.

12.3.7.2 Varying the zero or shift parameter

Where the plotted rating curve shows periods of stability but the number of records are insufficient to define the new relationship over all or part of the range, then the parameter 'a' in the standard relationship $Q = c_1(h + a_1)^{b_1}$ may be adjusted. The parameter 'a_1' represents the datum correction between the zero of the gauges and the stage at zero flow. Scour or deposition causes a shift in this zero-flow stage and hence a change in the value 'a_1'.

The shift adjustment required can be determined by taking the average difference (Δa) between the rated stage (h_r) for measured flow (Q_m) and measured stage (h_m) using the previous rating. i.e.

$$\Delta a = \sum_{i=1}^{N} (h_r - h_m) / n \qquad \text{Equation 12.35}$$

The new rating over the specified range then becomes:

$$Q = c(h + a_1 + \Delta a)^{b_1} \qquad \text{Equation 12.36}$$

The judgement of the person processing the data is required to decide whether to apply the value of Δa over the full range of stage (given that the percentage effect will diminish with increasing stage) or only in the lower range for which the current meter gauging is available. If there is evidence that the rating changes from unstable section control at low flows to more stable channel control at higher flows, then the existing upper rating should continue to apply.

New stage ranges and limits between the rating segments need to be determined. The method assumes that the channel hydraulic properties remain unchanged, except for the level of the datum. Significant variation from this assumption will result in wide variation in (h_r - h_m) between included records. If this is the case, then the Stout's shift method should be used as an alternative.

12.3.7.3 Stout's shift method

This method has been described in details by Tilrem (1979). It is used for controls which are shifting continually or progressively. In such cases, the plotted current meter measurements show a widespread variation from the mean. They also show an insufficient number of sequential records with the same trend to split the simple rating into several periods. The procedure is as follows:

- Fit a mean relationship to (all) the points for the period in question.
- Determine h_r (the rated stage) from measured Q_m by reversing the power type rating curve or from the plot:

$$h_r = \left(\frac{Q_m}{Q_c}\right)^{\frac{1}{b}} - a \qquad \text{Equation 12.37}$$

Based on basic hydraulics, the exponent b is approximately 0.5, but this must be determined experimentally through a graphical plot of Q_m / Q_r against F_m / F_c where Q_m and Q_r are the measured and adjusted discharge and F_m and F_c are the measured and the selected constant fall on which the rating curve is based. Individual rating shifts (Δh), as shown in Figure 12.18, are then:

$$\Delta h = h_r - h_m \qquad \text{Equation 12.38}$$

- These Δh stage shifts are then plotted chronologically and interpolated between successive records (Figure 12.18) for any instant of time t as Δh_t
- These shifts Δh_t are used as a correction to the observed gauge height readings and the original rating applied to the corrected stages, i.e.

$$Q_t = C_f(h_t + \Delta h_t + a_t)^{b_1} \qquad \text{Equation 12.39}$$

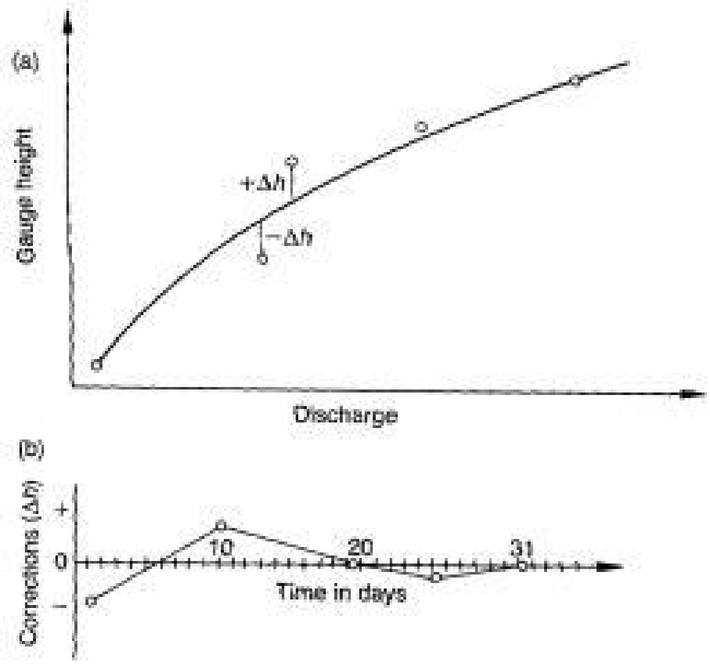


Figure 12.18: Stout's Method for Correcting Stage Readings when Control is Shifting (from Herschy, 2009)

Stout's method will yield acceptable results on the following conditions:

(a) There are ample flow measurements, preferably on a daily basis

(b) The mean rating is revised periodically - at least once a year.

The basic assumption in applying the Stout's method is that the deviations of the measured discharges from the established stage-discharge curve are due only to a change or shift in the station control and that the corrections applied to the observed gauge heights vary gradually and systematically between the days on which the check measurements are taken. However, the deviation of a discharge measurement from an established rating curve may be due to:

- gradual and systematic shifts in the control,
- abrupt random shifts in the control, and
- the error of observation and systematic errors caused by instrument and personnel both.

Stout's method is strictly appropriate for making adjustments only for the first type of error. If the check measurements are taken frequently enough, fair adjustments may also be made for the second type of error. The drawback of the Stout's method is that all the errors in current meter observation are mixed with the errors due to shifting in control and are thus incorporated in the final estimates of discharge. The Stout's method must therefore never be used where the rating is stable, or at least sufficiently stable to develop satisfactory rating curves between major shift events. Use of the Stout's method under such circumstances loses the advantage of a fitted line, where the standard error of the line S_{mr} is less than 20% of the standard error of individual records (S_e). Also, when significant observational errors exist, it is strongly recommended not to apply this method for establishing the rating curve.

12.3.7.4 Flow determined from daily gauging

Stations may be located at places where there is a very broad scatter in the rating relationship, which appears neither to result from backwater or scour and where the calculated shift is erratic. A cause may be the irregular opening and closure of a valve or a gate at a downstream flow regulating structure. Unless there is a desperate need for such data, it is recommended that the station be moved or closed. If the station is continued, daily measurement of the discharge may be adopted as the daily mean flow. This practice, however, eliminates the daily variations and peaks in the record.

It is emphasised that even when using the recommended methods, the accuracy of flow determination at stations which are affected by shifting control will be much reduced compared to the stations that are not affected. Additionally, the cost of obtaining worthwhile data will be considerably higher. At many such stations, uncertainties of 20 to 30% are the best that can be achieved, and consideration should be given to whether such accuracy meets the functional needs of the station.

12.4 Uncertainty in the Stage-Discharge Relationship

The uncertainty in stage-discharge is derived by statistical analysis of the scatter of the measurements around the rating curve. It is similar in concept to the standard error derived in regression analysis. However, unlike regression curve fitting, the stage-discharge curve is derived commonly using hydraulic reasoning as well as mathematical fitting. Therefore, the term uncertainty is used in preference to the more restrictively defined statistical standard error (IS 15119 (Part 2): 2002).

The uncertainty of a rating curve relationship is characterized by the standard error of estimate S, calculated from the dispersion of the stage-discharge data around the rating curve given by:

$$S = \left[\frac{\sum(\ln Q - \ln Q_c)^2}{N-2}\right]^{0.5} \qquad \text{Equation 12.40}$$

Where,

Q = Measured discharge

Q_c = Corresponding discharge calculated from the rating curve equation

N = Number of records in rating curve segment

The per cent uncertainty in the calculated value of $lnQ_c(2s_m r)$, at the point $\ln(h-e)$ may be found from the following equation

$$2s_{mr} = \pm t s_e \left\{ \frac{1}{N} + \frac{\left[\ln(h-e) - \overline{\ln(h-e)}\right]^2}{\sum\left[\ln(h-e) - \overline{\ln(h-e)}\right]^2} \right\}^{0.5} \times 100 \qquad \text{Equation 12.41}$$

Where t is the Student's t correction at the 95% confidence for N records and may be taken as 2 for 20 or more records.

To avoid systematic bias in the relation, it is recommended to use several current meters to establish the stage-discharge function. The uncertainty of $lnQ_c(2s_m r)$ should be computed for each observation of (h-e) related to the corresponding gauging. The value will be minimum at the mean value of $\ln(h-e)$.

Expanded uncertainty is derived by multiplying the standard uncertainty by a coverage factor k given by:

$$U(\ln Q_c(h)) = k\,u(\ln Q_c(h)) \qquad \text{Equation 12.42}$$

Where k is the coverage factor which provides a specified level of confidence. If the error distribution is assumed to be approximately normal (Gaussian), the coverage factors k of 1, 2 and 3 correspond to levels of confidence of about 68%, 95% and 99.8% respectively.

The expanded uncertainty defines the uncertainty interval around the computed value ln Qc(h) which is expected to encompass the specified fraction of the distribution of values that could reasonably be attributed to the discharge. The interval is thus expressed as $\ln Q_c(h) \pm U \ln Q_c(h)$. The corresponding uncertainty interval for discharges is found by taking anti-logarithm.

(Note: The 'expanded uncertainty' and 'level of confidence' are not to be confused with the statistical quantities 'confidence interval' and 'confidence level'.)

The expanded uncertainty U ln Qc(h) with coverage factor k=2 and the corresponding uncertainty limits on ln Qc(h) should be calculated for each observation of (h-e) related to the corresponding gauging. The limits will, therefore, take the form of curved lines on each side of the stage-discharge relationship and exhibit a minimum at the mean value of ln (h - e).

Example 12-1

Let us consider the test example of the site Hivra on river Wardha in the Pranhita sub-basin of the Godavari Basin. The flow and stage data from 2000 January to 2000 April measured on a daily basis have been used for developing the rating curve. The plot of flow and water level time series is shown in Figure 12.19 below:

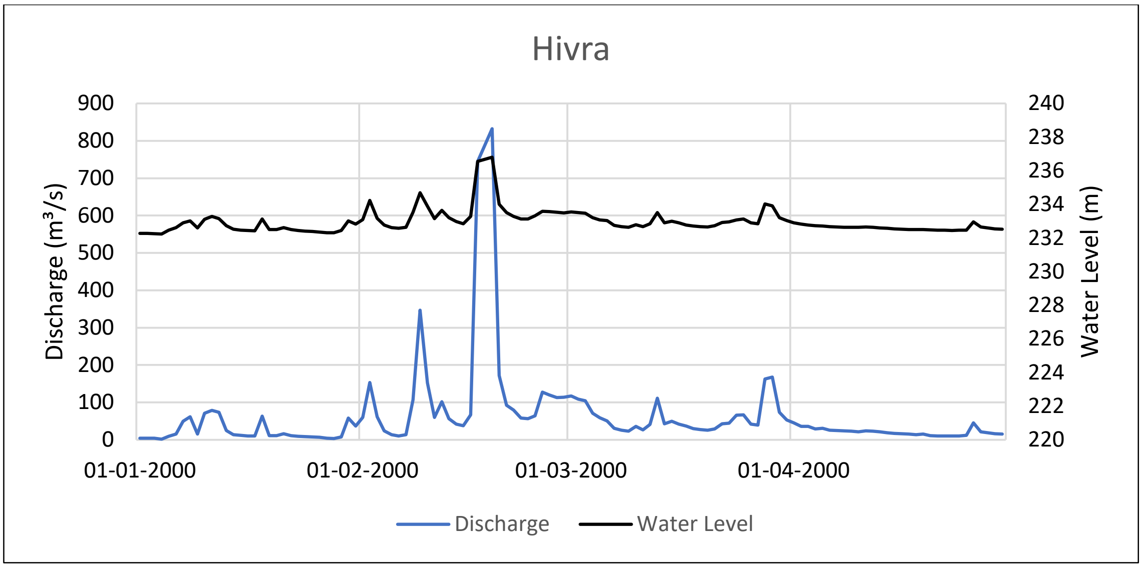


Figure 12.19: Plot of Flow and Water Level Time Series

Calculating the datum correction

The datum correction (a) is the water level for which the flow is zero. With the initial input of 230 m, the optimized value for minimum mean square error was arrived at as 232.06m. The corresponding rating curve is given below.

Zero Gauge level (a)
232.06 m

Rating Table

H (m)	Q (cumec)
232.06	0
233.06	54.93
234.06	189.24
235.06	390.17
236.06	651.95
237.06	970.86

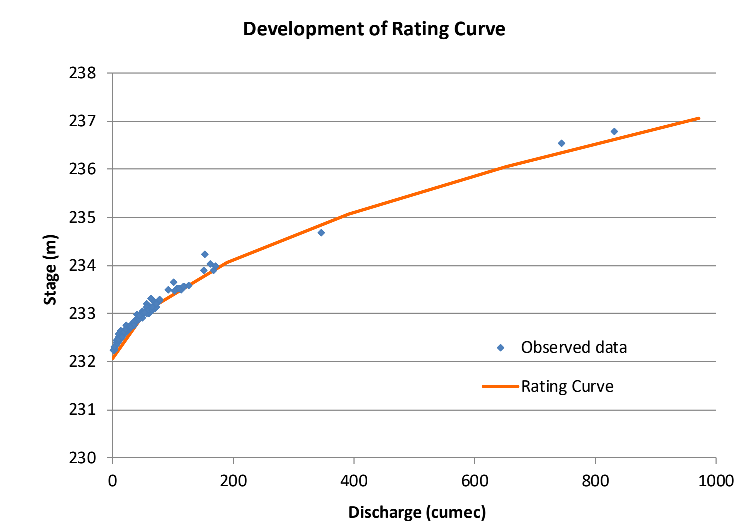


Figure 12.20: Development of Rating Curve

Range identification

The first estimate of water level range for different segments are identified by a log-log plot. Two segments have been identified through visual examination from the double log plot shown below.

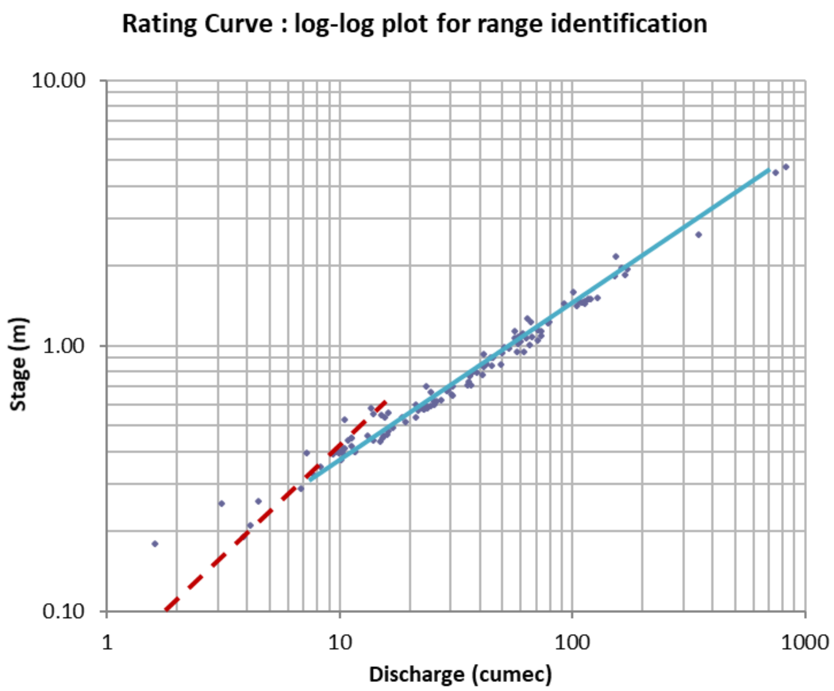


Figure 12.21: Log-Log Plot of Discharge with Stage

Identification of range

No of Range = 2	NZE = 232.06 m			
Segment	Boundary			
	Lower >	Upper <=	Lower >	Upper <=
Segment 1	0.00	0.49	232.06	232.55
Segment 2	0.49	8.29	232.55	240.35

The procedure is repeated for each segment of the rating curve. The NZE for segment 1 and segment 2 is revised. The details are below:

	N (No of obs.)	a = (m)	Cr = (10^b)	β	SE (standard error)	r^2
Segment 1	32	232.00	58.5	2.21	1.32	0.945
Segment 2	88	232.04	53.34	1.76	10.33	0.996

Table 12-1: Results of Rating Curve fitting

Station Name		Hivra		
Data	From	2000/1/1	To	2000/4/30
Number of data		120		

Observation	Parameter	Value (m or m^3/s)	Date
Gauge C_r	Minimum	232.24 m	2000/1/4
	Maximum	236.8 m	2000/2/19
Discharge Q	Minimum	1.60 m^3/s	2000/1/4
	Maximum	832.8 m^3/s	2000/2/19

Equation Type: Power $Q = C_r(H - a)^{\beta}$

Boundaries/ coefficients Gauge Zero: 232.06 m

Segment	Lower Bound	Upper Bound	a	C_r	β	s.e.	no. of data
Segment 1	232.06	232.55	232.00	58.50	2.21	1.32	32
Segment 2	232.55	240.35	232.04	53.34	1.76	10.33	88

Number	Discharge (observed)	Water Level	Discharge (computed)	Diff
	m^3/s	m	m^3/s	m^3/s
(1)	(2)	(3)	(4)	(2) – (4)
1	4.349	232.28	3.524	0.8
2	4.144	232.27	3.252	0.9
3	3.848	232.3	2.744	1.1
4	1.6	232.2	2.508	-0.9
5	9.686	232.5	10.541	-0.9
6	15.076	232.6	19.840	-4.8
7	49.407	232.9	41.747	7.7
8	61.581	233.0	50.554	11.0
9	15.575	232.6	19.232	-3.7
10	71.229	233.1	60.079	11.1
11	78.4	233.3	77.873	0.5
12	73.745	233.2	64.086	9.7
13	24.711	232.7	27.767	-3.1
14	13.18	232.5	13.817	-0.6
15	11.6	232.5	10.541	1.1
-	-	-	-	-
-	-	-	-	-
107	15.19	232.5	13.237	2.0
108	14	232.5	12.671	1.3
109	14.993	232.5	12.393	2.6
110	11.225	232.5	11.579	-0.4
111	10.543	232.5	11.054	-0.5
112	10.335	232.5	11.054	-0.7
113	9.891	232.5	10.290	-0.4
114	10.043	232.5	10.541	-0.5
115	11.6	232.5	10.541	1.1
116	45.448	233.0	46.060	-0.6
117	21.151	232.7	23.003	-1.9
118	18.554	232.6	19.232	-0.7
119	16.12	232.5	14.712	1.4
120	15.365	232.5	13.817	1.5

The plot is shown in Figure 12.22 below:

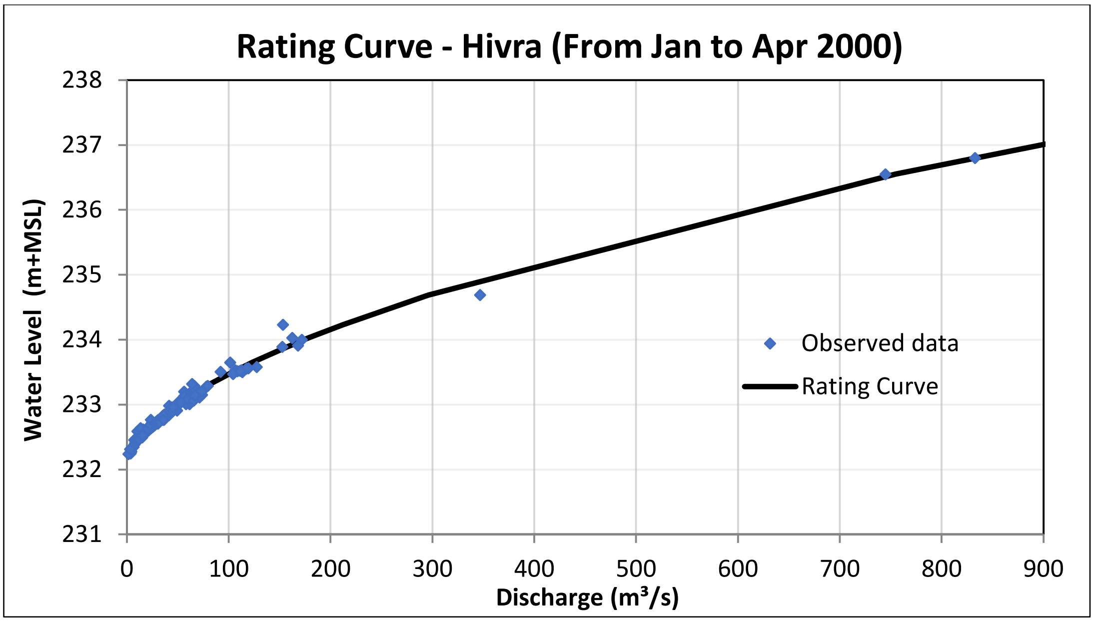


Figure 12.22: Rating Curve for Hivra GD Station

Computation of uncertainty limits (sample calculation)

Sl No	Q_{obs} (m^3/s)	H (m)	H-a (m)	log(H -a)=X	logQ= Y	XY	X^2	Y^2	Q_c (m^3/s)	$(Q_{obs}-Q_c)^2$
	1	2	3	4	5	6	7	8	9	10
1	1.6	232.24	0.24	-0.62	0.20	-0.13	0.38	0.04	2.51	0.82
2	3.848	232.25	0.25	-0.60	0.59	-0.35	0.36	0.34	2.74	1.22
3	4.144	232.27	0.27	-0.57	0.62	-0.35	0.32	0.38	3.25	0.79
4	4.349	232.28	0.28	-0.55	0.64	-0.35	0.31	0.41	3.52	0.68
5	3.121	232.315	0.31	-0.50	0.49	-0.25	0.25	0.24	4.57	2.10
6	4.488	232.32	0.32	-0.49	0.65	-0.32	0.24	0.43	4.73	0.06
7	6.793	232.35	0.35	-0.46	0.83	-0.38	0.21	0.69	5.77	1.05
8	7.6	232.39	0.39	-0.41	0.88	-0.36	0.17	0.78	7.32	0.08

Sl. Nc	Uncertainty				68% uncertainty limits			
	$(X-\bar{X})^2$	68% U(ln(Qc(h))	95% U(ln(Qc(h))	99.8% U(ln(Qc(h))	log(Qc)+ U	ln(Qc)- U	Upper Limit	Lower Limit
	11	12	13	14	15	16	17	18
1	0.063	0.037	0.074	0.111	0.44	0.36	2.73	2.30
2	0.054	0.035	0.070	0.104	0.47	0.40	2.97	2.53
3	0.040	0.031	0.061	0.092	0.54	0.48	3.49	3.03
4	0.034	0.029	0.057	0.086	0.58	0.52	3.77	3.30
5	0.017	0.023	0.046	0.068	0.68	0.64	4.82	4.34
6	0.016	0.022	0.044	0.066	0.70	0.65	4.98	4.50

7	0.007	0.018	0.036	0.055	0.78	0.74	6.01	5.53
8	0.002	0.015	0.030	0.044	0.88	0.85	7.58	7.08

Sl No.	95% uncertainty limits				99.8% uncertainty limits			
	log(Qc)+U	log(Qc) - U	Upper Limit	Lower Limit	log(Qc)+U	log(Qc) - U	Upper Limit	Lower Limit
	19	**20**	**21**	**22**	**23**	**24**	**25**	**26**
1	0.47	0.33	2.97	2.11	0.51	0.29	3.24	1.94
2	0.51	0.37	3.22	2.34	0.54	0.33	3.49	2.16
3	0.57	0.45	3.75	2.82	0.60	0.42	4.02	2.63
4	0.60	0.49	4.02	3.09	0.63	0.46	4.30	2.89
5	0.71	0.61	5.08	4.12	0.73	0.59	5.35	3.90
6	0.72	0.63	5.24	4.28	0.74	0.61	5.51	4.06
7	0.80	0.72	6.27	5.30	0.82	0.71	6.54	5.09
8	0.89	0.84	7.84	6.84	0.91	0.82	8.11	6.61

Where,

Parameter	Value
Standard Error of Estimate (S)	0.078
68% level of confidence (S)	0.078
95% level of confidence (2*S)	0.156
99.8% level of confidence (3*S)	0.233

$(12) = S\sqrt{1/n \ + ((11)/(\sum(11))}$

(18) = log (9) + (12)

(20) = 10^(18)

Where, (*) is col.(*)

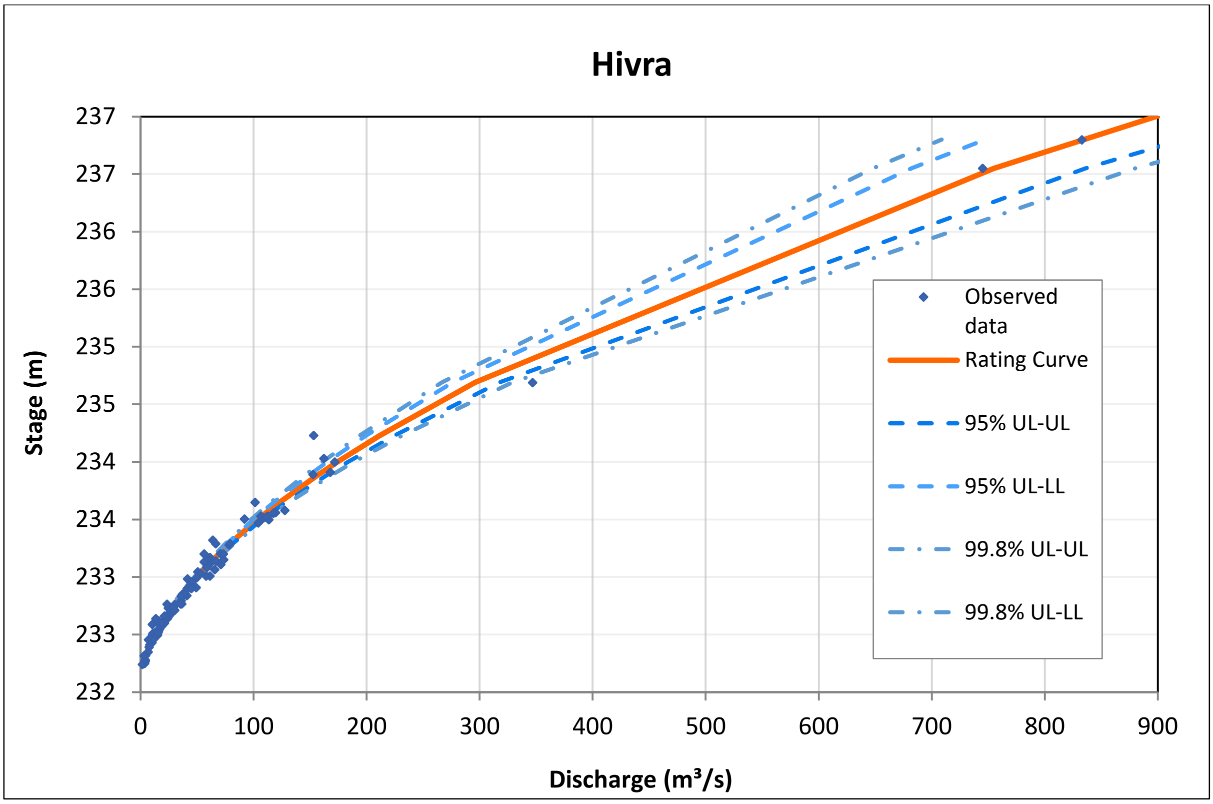


Figure 12.23: Rating Curve with Uncertainty

13 VALIDATION OF THE RATING CURVE

13.1 General

Validation of a rating curve is required after the stage-discharge relationship has first been fitted and when new records have been obtained, to assess whether the tests indicate a change in rating. Validation of the rating curve is also used to assess the reliability of historical ratings.

For discharge measurement, current meter gauging is carried out with variable frequency depending on previous experience of the stability of the control and the rating curve. As a minimum, it is recommended that six flow measurements per year should be carried out even for a station with a stable section that has been previously gauged over the full range of levels. More measurements are required at unstable cross sections. The deviation of such check measurements from the previously established relationship is computed, and any bias assessed to determine whether they belong to the same population as the previous stage-discharge relationship.

Graphical and numerical tests are designed to show whether the flow measurements fit the current relationship equally and without bias over the full range of flow and over the full-time period to which it has been applied. If they do not, then a new rating should be developed as described in Chapter 12, taking into account the deficiencies noted during the validation stage. Validation of the rating curves is to be carried out at the Divisional offices or the State Data Processing Centre.

13.2 Graphical Validation Tests

13.2.1 General

Graphical tests are often the most effective method of validation. These include the following:

- Stage/ discharge plot with the new flow measurements
- Period/ flow deviation scattergram
- Stage/ flow deviation scattergram
- Cumulative deviation plot of flow measurements
- Stage/ discharge plots with flow measurements distinguished by season

Judgements based on graphical displays are often indicative rather than prescriptive - a judgement on the part of the data analyst is still required.

13.2.2 Stage- discharge plot with new flow measurements

The simplest means of validating the rating curve for subsequent measurements is to plot the existing rating curve with the new check measurements. This is shown in the example for Station Hivra. A rating curve is established for the period 1/1 – 31/4/2000, as shown in Figure 13.1. It shows a proper fit of the data to the existing rating curve, of which the numerical results are shown in Table 12-1. New data are available for the period 1/8-30/9/2001. The plot of new data with the existing rating curve are shown in Figure 13.2. From this plot, it is observed that the new measurements do not match properly with the existing curve. In Figure 13.3 the new measurements are shown with the rating curve and the 95% confidence limits (derived as t-times the standard error S_e). From this plot, it can be judged whether most of the measurements lie inside the confidence limits, thus implying acceptable deviation. It is expected that 19 out of 20 observations will lie inside the limits if the standard error is considered at a 5% significance level. However, even though one can see whether all the new points lie above or below the previous regression line, the graph does not specifically address the problem of bias. For example, if 25 new measurements all lie scattered within 95% confidence limits, it does not show any significant change in behaviour. However, if these points are plotted with the time sequence of each observation, and from the plot if a certain pattern of deviation (against time) is perceivable and significant, then such a situation may warrant new rating curve to be developed for different periods of distinct behaviour. For the case of the station Hivra, the plot confirms that the new measurements differ significantly from the existing curve.

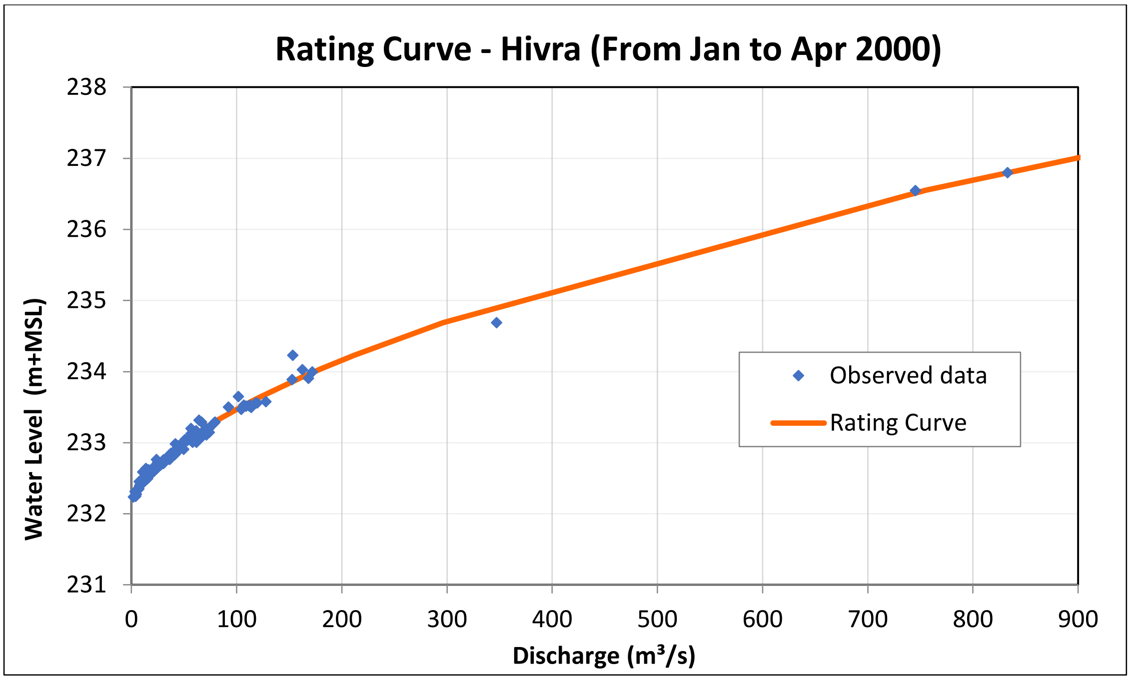


Figure 13.1: Rating Curve Based on Data for the Period 1/1/2000--30/04/2000

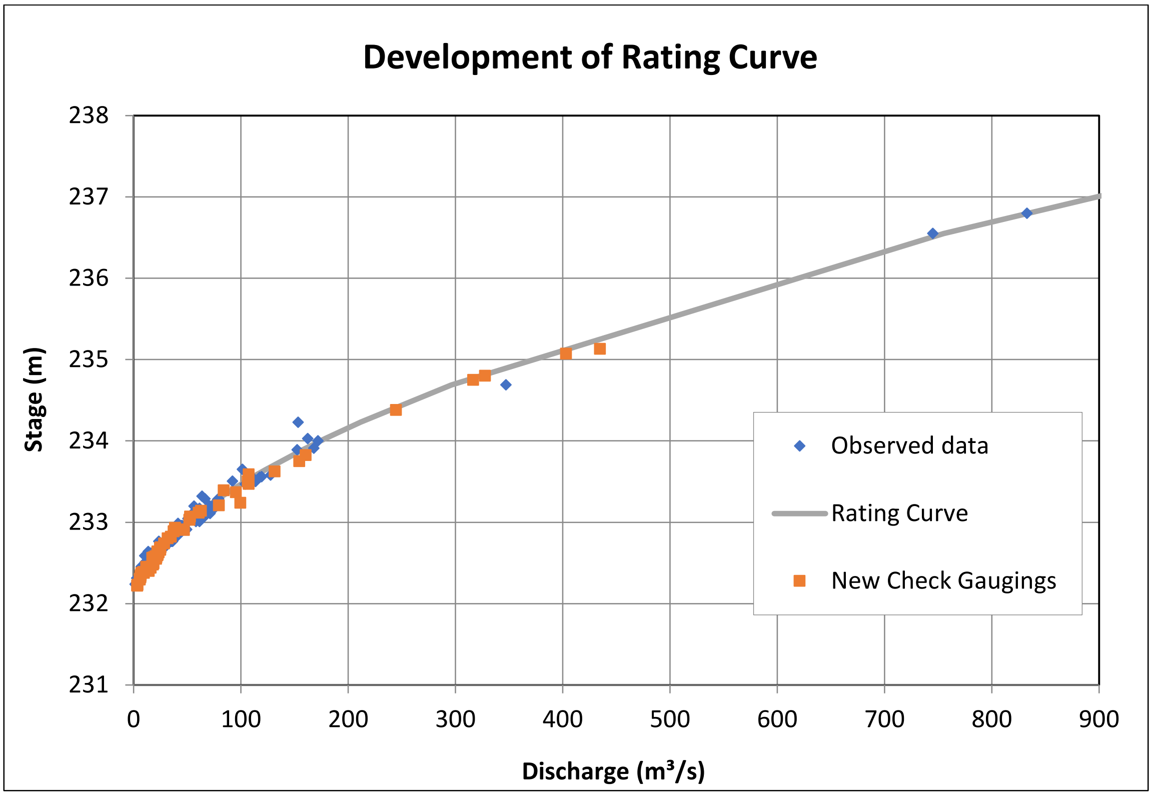


Figure 13.2: New Records at Hivra Station Plotted Against the Existing Rating Curve

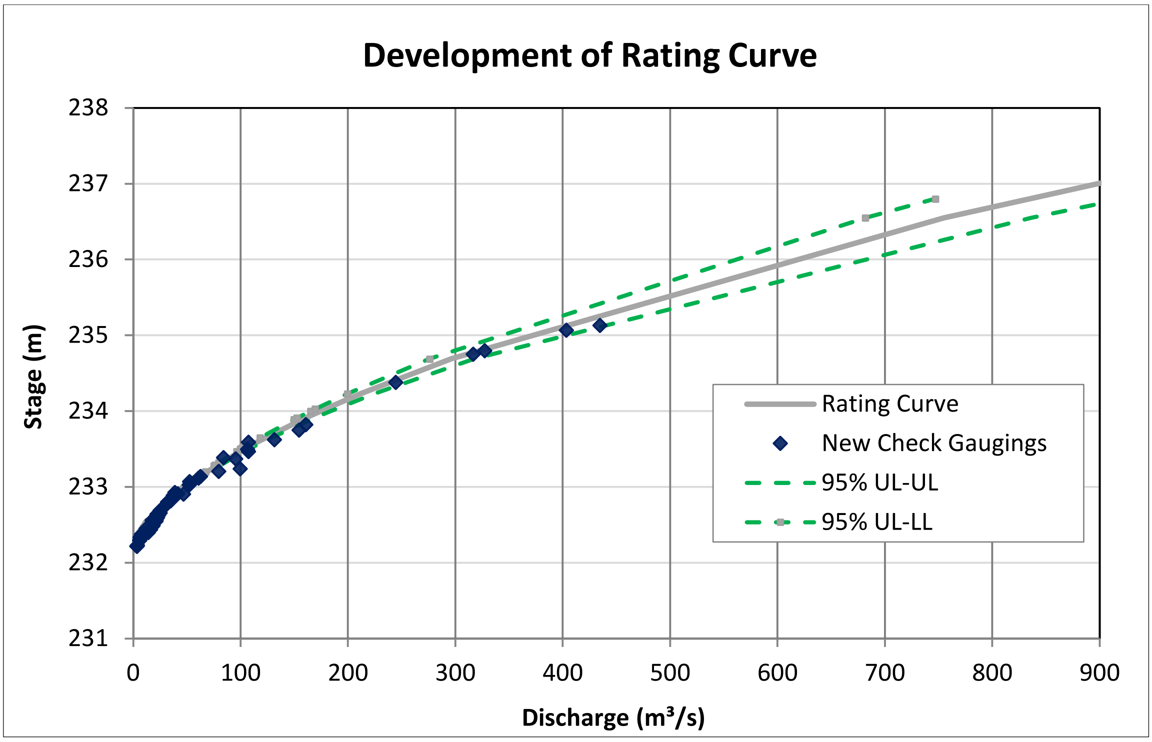


Figure 13.3: New Records at Hivra Station Plotted Against Existing Rating Curve With 95% Confidence Limits

13.2.3 Period/flow deviation scattergram

A period/ flow deviation scattergram shown in Figure 13.4 is a means of illustrating the negative and positive deviation of each current meter gauging from the present rating curve. It also depicts whether there has been a gradual or sudden shift in the direction of deviations within the period to which the rating has been applied. It also shows whether recent additional records show a deviation from the previous measurements. In the example shown in Figure 13.4, percentage deviations are very high; there are far more records with positive than with negative deviations. The rating is therefore biased and a revision of the rating is strongly recommended.

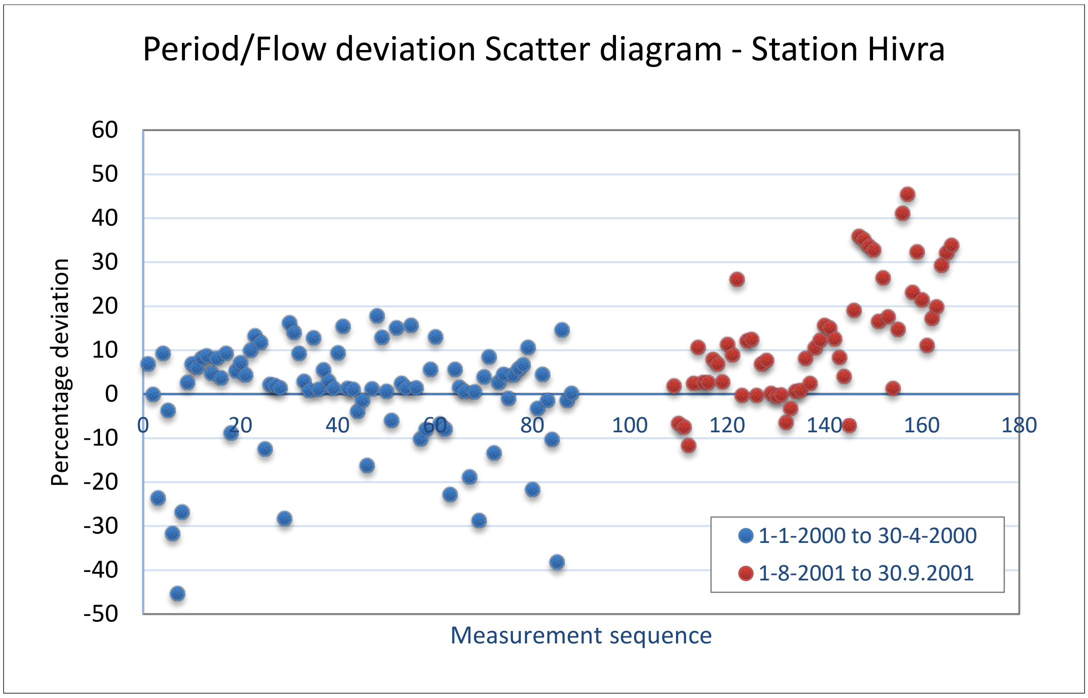


Figure 13.4: Period-Flow Deviation Scatter Diagram for Hivra Rating Curve Data and New Records

13.2.4 Stage/flow deviation diagram

A similar scattergram plot shows the percentage deviation with the stage (Figure 13.5) and it serves as a means of illustrating whether the relationship is biased over certain ranges of the stage. Most recent records can also be placed within this context.

The example shown in Figure 13.5 brings out that there is some difference in deviation at different stages; particularly at the lower stages the differences are substantial. This plot therefore confirms the necessity for revising the rating curve.

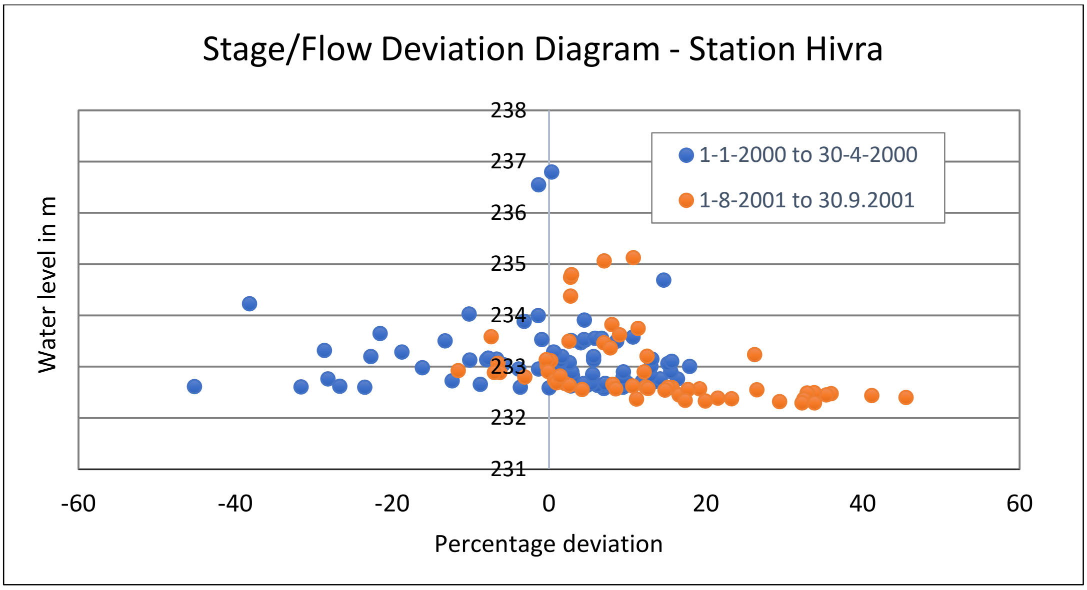


Figure 13.5: Stage-Flow Deviation Scatter Diagram for Hivra Rating Curve Data and New Records

13.2.5 Cumulative deviation plot of records

A plot of the cumulative deviation of flow records from the rating curve gives another indication of bias and whether that bias has changed with time. Figure 13.6 shows such a plot for the station Hivra. From the upward trend of the line for the new measurements, it is concluded that the new measurements produce consistently higher flow values for the same stages, as compared to the earlier measurements.

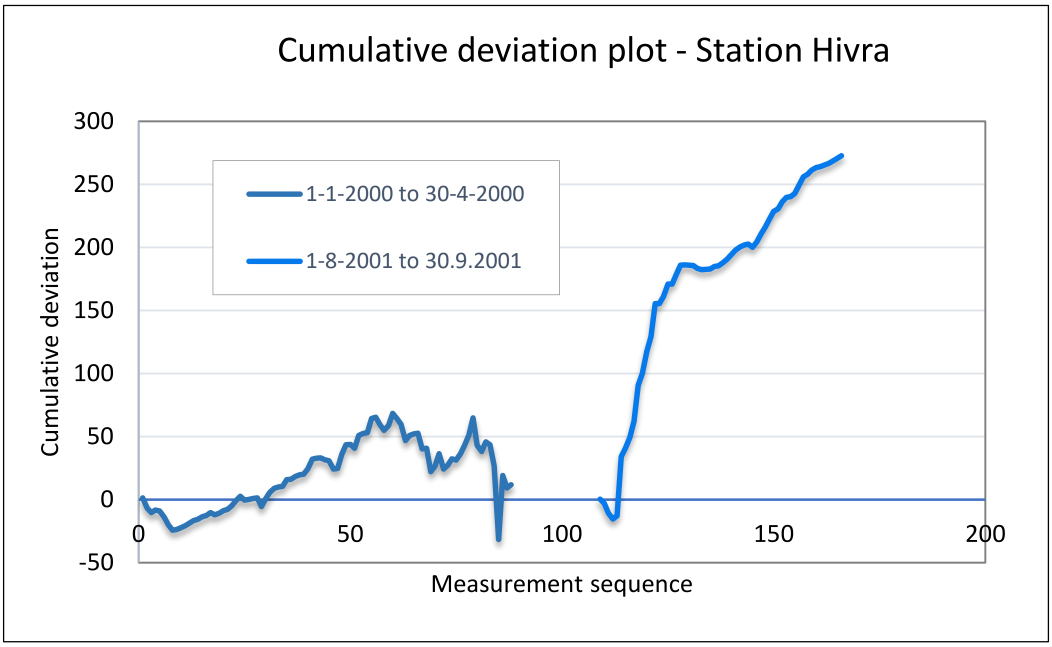


Figure 13.6: Cumulative Deviation Plot

13.2.6 Stage discharge plots with gauging distinguished by season

It is sometimes helpful to separate records between seasons to demonstrate the effects of varying weed growth or other seasonal factors on the stage-discharge relationship. The effects of weed growth may be expected to be at a maximum in low flows before the onset of the monsoon. Monsoon high flows wash out the weed which grows progressively from the end of the rains. The discharge for a given level may thus differ from one month to another. This shows up more clearly in rivers where winter low flows are not much affected by weed growth as compared to the summer low flows, and thus show much smaller spread. Where an auxiliary gauge is available, a backwater rating curve (normal fall method) may be used. Otherwise, a simple rating curve may be used for the periods when the weeds are absent, and Stout's shift method used during periods of variability.

13.3 Numerical Validation Tests

13.3.1 Use of Student's 't' test for verification records

A test such as the Student's "t" test may be used to decide whether verification records can be accepted as part of the homogeneous sample of observations making up the existing stage-discharge curve. Such a test will indicate whether or not the stage-discharge relation requires re-calculation or the section requires recalibration.

In this test, the 't' statistic is calculated as the ratio of the mean deviation and the standard error of the difference of the means as:

$$t = \frac{\overline{d_1}}{s} \qquad \text{Equation 13.1}$$

Where,

$\overline{d_1}$ is the mean deviation of the new records from the existing curve (%)

and s is the standard error of the difference in the means expressed as $s = a\sqrt{\frac{N+N_1}{NN_1}}$

N = Number of records used to derive the existing rating

N_1 = Number of new records

"a" is given by the following expression:

$$a = \sqrt{\frac{\sum d^2 + \sum\left(d_1 - \overline{d_1}\right)^2}{N+N_1-2}} \qquad \text{Equation 13.2}$$

$\sum d^2$ = Sum of the squares of the percent differences for the old records from the existing rating curve.

If the computed value of 't' = $\frac{\overline{d}}{s}$ is greater than the critical value of 't' for (N + N_1 -2) degrees of freedom at 95% probability level, a new rating needs to be developed or a request is to be made to the field staff for additional verification records.

The critical values of Students '*t*' statistic at the 95% confidence level can be obtained from the standard tables available for the Student's '*t*' distribution. It should be noted that the rating changes are more frequent and more noticeable in the low flow range. The review and validation are therefore done for each range and, unless there is evidence to the contrary, unaffected ranges should retain the old rating but with the range limits adjusted for the new intersection. If the computed value of the 't' statistic exceeds the critical value for the chosen range of confidence, then the null hypothesis is rejected, indicating that the new observations are different from those predicted by the rating curve.

As an example, the validation of the new records at the Hivra station is shown in Table 13-1. The results of the 't'-test is seen to not support the earlier observation of significant deviation.

Table 13-1: Results of Validation using Students 't' Test to Check Records

Segment	N		N+N1-2	a	S	t (computed)
Segment 1	32.00		51	8.1468	2.287924	0.596
Segment 2	88.00		127	275.7471	52.14019	0.342

Segment	Df	T (critical value of t)	t	T value result
Segment 1	51	2.007584	0.596	Accept
Segment 2	127	1.97882	0.342	Accept

13.3.2 Test for absence of bias in signs

A well-balanced rating curve must ensure that the number of positive and negative deviations of the observed values from the rating curve is evenly distributed. That is, the difference in number between the two should not be more than that, which can be explained by chance fluctuations. The test is employed to see if the curve has been established in a balanced manner so that the two sets of discharge values, observed and estimated (from the curve), may be reasonably supposed to represent the same population.

This test is performed by counting observed points falling on either side of the curve. If Q_i is the observed value and Q_c the estimated value, then the expression, Q_i - Q_c, should have an equal chance of being positive or negative. In other words, the probability of Q_i - Q_c being positive or negative is ½. Hence, assuming the successive signs to be independent of each other, the sequence of the differences may be considered as distributed according to the binomial distribution (p+q) N. Here, N is the number of observations, and p and q, the probabilities of occurrence of positive and negative values are ½ each. The expected number of positive signs is N_p. Its standard deviation is ($\sqrt{N_{pq}}$). The "t" statistic is then found by dividing the difference between the actual number of positive signs N_1 and the expected number of positive signs N_p by its standard deviation ($\sqrt{N_{pq}}$):

$$t = \frac{|N_1 - N_p| - 0.5}{\sqrt{N_{pq}}} \qquad \text{Equation 13.3}$$

The resulting value is compared with the critical value of "t" statistic for a 5% significance level for the degrees of freedom equal to the total number of stage-discharge data. If the value of the critical "t" statistic is more than that obtained for the observed data then it can be considered that the data does not show any bias for the sign of the deviations between observed and computed discharges.

Table 13-2: Test for Absence from Bias in Signs

Segment	Df	T (t critical)	t (computed)	T value result
Segment 1	32	2.036933343	0.177	Accept
Segment 2	88	1.987289865	2.878	Reject

13.3.3 Test for absence from bias in values

This test is designed to see if a particular stage-discharge curve, on an average, yields significant underestimates or overestimates as compared to the actual observations on which it is based (compare the graphical test using the period/ flow deviation and stage /flow deviation scattergrams). The percentage differences are first worked out as:

$$P = 100\,(Q_i - Q_c) / Q_c \qquad \text{Equation 13.4}$$

If there are N observations and $P_1, P_2, P_3, \ldots, P_N$ is the percentage differences and P_{av} is the average of these differences, the standard error of P_{av} is given by:

$$S_e = \sqrt{\frac{\sum (P_i - P_{av})^2}{N(N-1)}} \qquad \text{Equation 13.5}$$

The average per cent P_{av} is tested against its standard error to see if it is significantly different from zero. The "t" statistic, in this case, is computed as:

$$t = (P_{av} - 0) / S_e \qquad \text{Equation 13.6}$$

If the critical value of "t" statistic for 5% significance level and N degrees of freedom is greater than the value computed above, then it may be considered that there is no statistical bias in the observed magnitudes compared to that obtained by using the rating curve.

The percentage differences have been taken as they are comparatively independent of the discharge volume and are approximately normally distributed about zero mean value for an unbiased curve.

Table 13-3: Test for Absence from Bias in Values

Segment	Df	T (t critical)	t (computed)	T value result
Segment 1	32	2.036933343	0.466	Accept
Segment 2	88	1.987289865	1.150	Accept

13.3.4 The goodness of fit test

Due to changes in the flow regime, long runs of positive and/ or negative deviations may be obtained at various stages. This may also be due to inappropriate fitting of the rating curve. This test is carried out for long runs of positive and negative deviations of the observed values from the stage-discharge curve. The test is designed to ensure a balanced fit of the deviations over the different stages.

The test is based on the number of changes of sign in the series of deviations (observed value minus expected or computed value). First, the signs of deviations in discharge measurements (+/—) in the ascending order of stage are recorded. Then, starting from the second sign of the series, "0" or "1" is placed under the sign respectively, based on whether the sign agrees with or does not agree with the sign immediately preceding it. For example, if there are N numbers in the original series, then the ($N - 1$) numbers in the derived series would be 11000100010001, as shown below.

+	-	+	+	+	+	-	-	-	-	+	+	+	+	-
	1	1	0	0	0	1	0	0	0	1	0	0	0	1

If the difference of the observed values from the predictions by the rating curve happen to arise from random fluctuations, the probability of a change in sign could be taken to be ½. However, this assumes that the estimated value is the median rather than the mean. If N is fairly large, a practical criterion may be obtained by assuming that the successive signs are independent (i.e., by assuming that they arise only from random fluctuations). Therefore, the number of "1" s (or "0" s) in the derived sequence of ($N - 1$) members may be judged as a binomial variable with parameters ($N - 1$) and ½.

From the series derived above, the actual number of changes in the sign is noted. The expected number of changes of the sign is computed by multiplying total possible numbers (i.e., $N - 1$) with the probability of change of sign (i.e., ½). The statistical significance of the departure of the actual number of change of signs from the expected number is known by finding the "t" statistic as follows:

$$t = \frac{|N' - (N-1)p| - 0.5}{\sqrt{(N-1)pq}}$$ Equation 13.7

where N' denotes the actual number of changes of sign. If the critical value of "t" statistic for $(N - 1)$ degrees of freedom is more than the value computed above, then it can be considered to have an adequate goodness of fit. Otherwise, the results will indicate that there is a significant bias in the fitted curve, with long runs of positive or negative deviations.

Using the data from the previous Example of Hivra flow records, the results of the Goodness of fit test are provided in Table 13-4.

Table 13-4: Goodness of Fit Test

Segment	N'	(N-1)p	Abs(N'-(N-1)p)-0.5	(N-1)pq	t (computed)
Segment 1	15	15.5	0.0	7.75	0.000
Segment 2	33	43.5	10.0	21.75	2.144

Segment	Df	T (t critical)	t (computed)	t value result
Segment 1	31	2.039	0.000	Accept
Segment 2	87	1.987	2.144	Reject

14 EXTRAPOLATION OF THE RATING CURVE

14.1 General

Extreme flows are crucial for design and planning. Hence, their best possible estimates should be made. Taking discharge measurements for high instantaneous flow is particularly difficult as they occur infrequently and are of short duration. Such peak flows mostly arrive at the time when it is highly probable that the gauging team is not present on the site. Also, the conditions prevailing during floods are usually not safe for flow gauging. Under such high flood condition, the gauging site becomes inaccessible, the gauging facilities remain no longer serviceable and the river may have spread from a confined channel to the flood plain.

Extrapolation of rating curves is frequently required because the range of levels over which gauging is carried out for developing the rating curve does not cover the full range of all observed levels. The rating curve may fall short at both the lower and the upper end. Extrapolation is not simply about extending the rating from existing records to the extreme levels, although in some cases this may be acceptable. There are instances where a different control may apply, the channel geometry may change, and flow over the floodplain may occur. Also, the channel form and vegetation roughness coefficients may change.

The applicable methods of extrapolation depend on the physical condition of the channel, whether in the bank or overbank and whether there are fixed or shifting controls. Consideration must also be given to the phenomenon of the kinematic effect of open channel flow when there may be a reduction of the mean velocity in the main channel during inundation of the flood plain. Methods given below are suitable for rivers with defined banks and fixed controls, as well as for the channels with a spill.

As per the current Hydrological Information System, extrapolation of stage-discharge relationships will be carried out at the State Data Processing Centre.

14.2 High Flow Extrapolation

Specific hydraulic conditions during a flood event should be kept in mind when extrapolating rating curves. The Q(H) relationship may abruptly change above some threshold stage, due to the following factors:

- change in the shape of the flow section when the water stage rises and overflows into the flood plain

- change in the downstream control such as backwater effect from a hydraulic structure or supercritical flow producing a hydraulic jump near the gauging station
- change in the roughness because of water flowing through vegetation in the flood plain
- change in the hydraulic behaviour of a flow structure in cases of overtopping of a bridge or a culvert
- secondary circulations in the river section resulting in energy losses
- by-passing flow occurring upstream of the gauging station not accounted for in the rating extrapolation
- temporal change occurring in the hydraulic conditions during the flood, caused by a change in the cross-section geometry due to erosion or sediment deposition or vegetation and wood debris blocked by a bridge.

In the absence of peak discharge measurements available, estimates corresponding to high values of the stage may be made using one or more of the following techniques:

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

14.2.1 The double log plot method

Simple extrapolation of the logarithmic stage-discharge relationship may be applied if the hydraulic characteristics of the channel do not change much beyond the measured range. In this case the relationship can simply be extended beyond the measured range by projecting the last segment of the straight-line relationship in the log-log domain. Such an extrapolation is illustrated by the dashed straight line in Figure 14.2 for the cross-section profile shown in Figure 14.1.

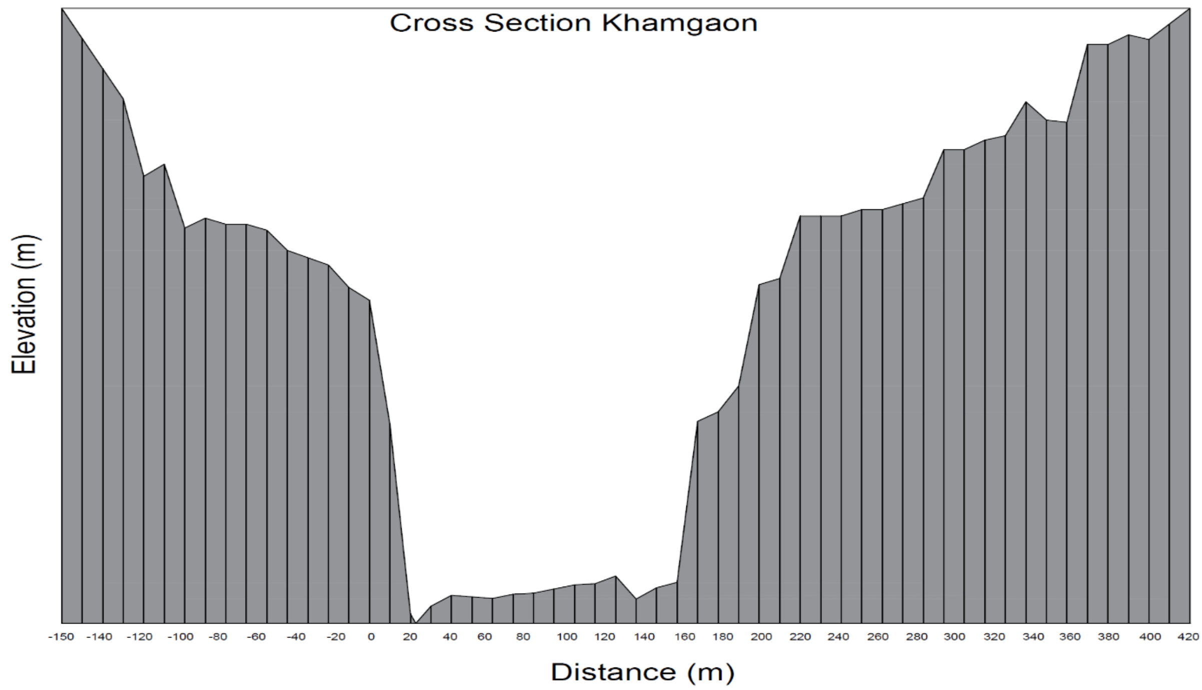


Figure 14.1: Cross-Section of the River at Khamgaon

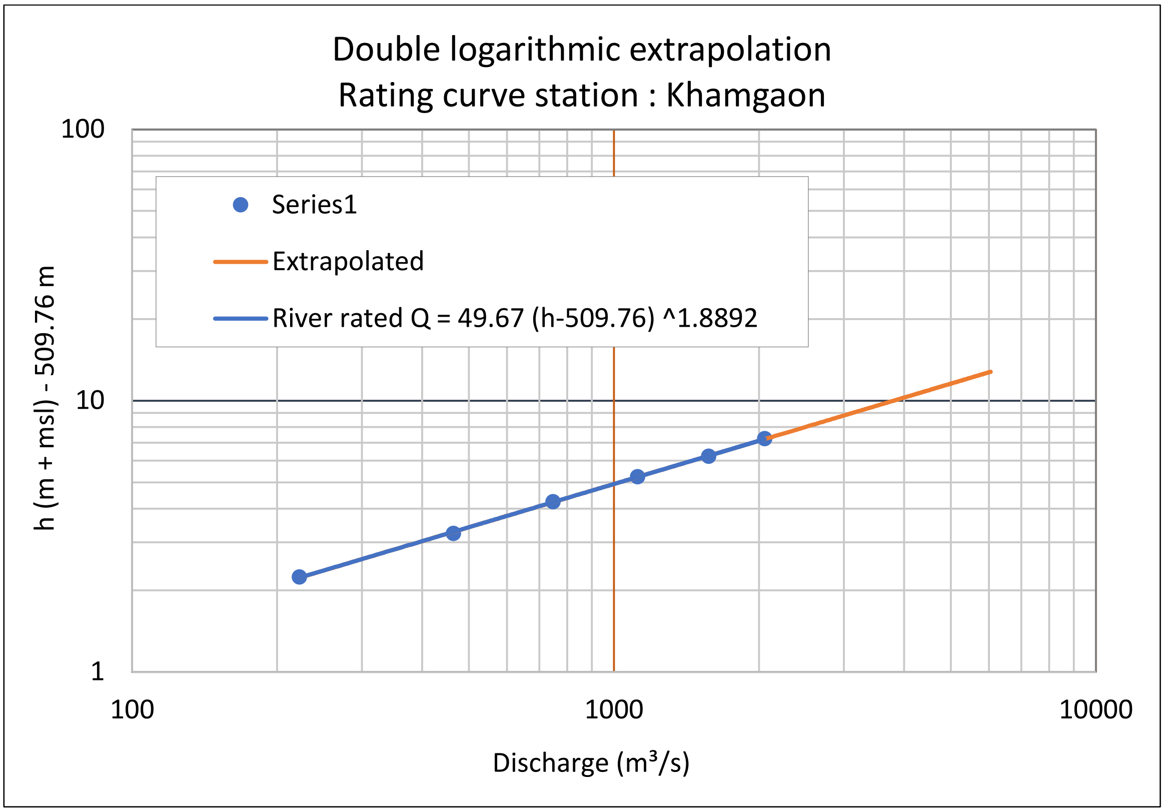


Figure 14.2: Example of Double Logarithmic Extrapolation of Rating Curve

In the example presented in Figure 14.2, a rating curve was established for the river flows up to the level of the flood plain. This curve had to be extended to cover the highest observed water level, which was about 4 m above the flood plain level. The double logarithmic technique was applied for this extrapolation. Double-logarithmic extrapolation implies that the same power type equation is applicable for the higher stages as well. The correctness of the use of this technique for the cross-section shown in Figure 14.1 is doubtful, since there is a presence of the flood plain. One of the basic conditions for the application of the double logarithmic method, namely no change in the hydraulic characteristics at the higher stages, is not fulfilled. This method will likely lead to an underestimation of the discharge since the contribution of the floodplain flows to the total river flow is not taken into consideration.

14.2.2 Stage-area / stage-velocity method

Where extrapolation is needed either well beyond the measured range, or there are known changes in the hydraulic characteristics of the control section, then a combination of the stage-area and stage-velocity curve may be used. Stage-area and stage-mean velocity curves are extended separately. For stable channels, the stage-area relationship is fixed and is determined by the survey up to the highest required stage. The stage-velocity curve is based on the current meter records within the measured range. Since the rate of increase in velocity at higher stages diminishes rapidly, this curve can be extended without much error for in-bank flows. Discharge for a given (extended) stage is then obtained by the product of area and mean velocity read using extrapolated stage-area and stage-mean velocity curves (Figure 14.3). This method may be used for extrapolation at both the upper and lower end of the rating.

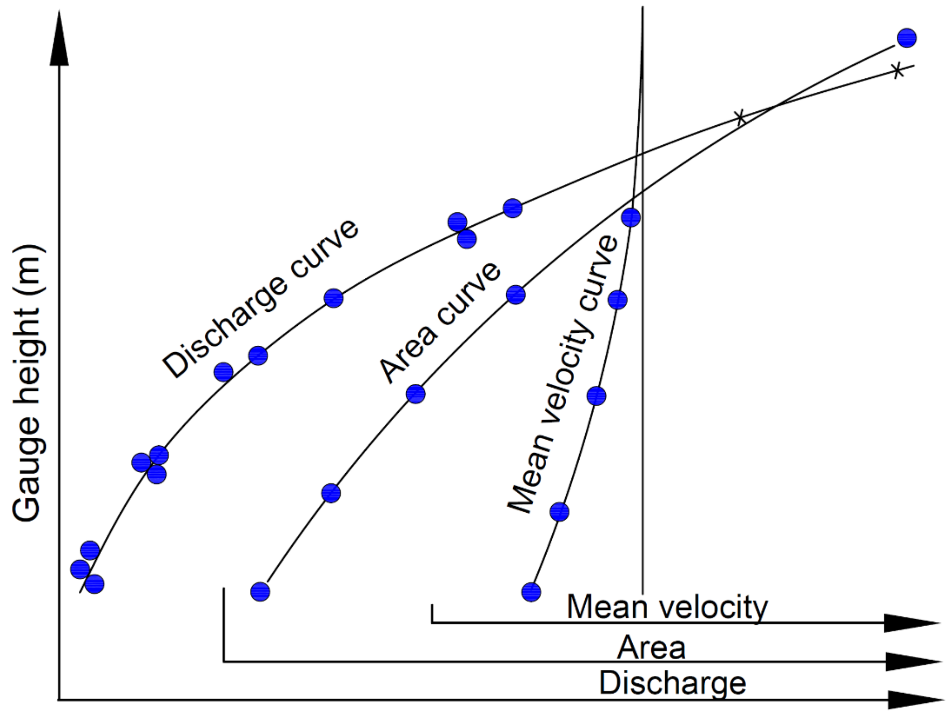


Figure 14.3: Extrapolation Based on Stage-Area/ Stage-Velocity Technique (Adopted from Herschy, 2009)

The mean velocity curve can also be extrapolated by the use of a logarithmic plot of mean velocity against the hydraulic radius (Figure 14.4). The hydraulic radius can be found for all stages from the cross-section survey. The logarithmic plot of mean velocity and hydraulic radius generally shows a linear relationship and thus can be extended linearly beyond the extent of measurements. Mean velocity in the extrapolated range can be obtained from this curve. Extrapolated discharge as before is obtained as the product of mean velocity thus estimated and the corresponding area from the stage-area curve.

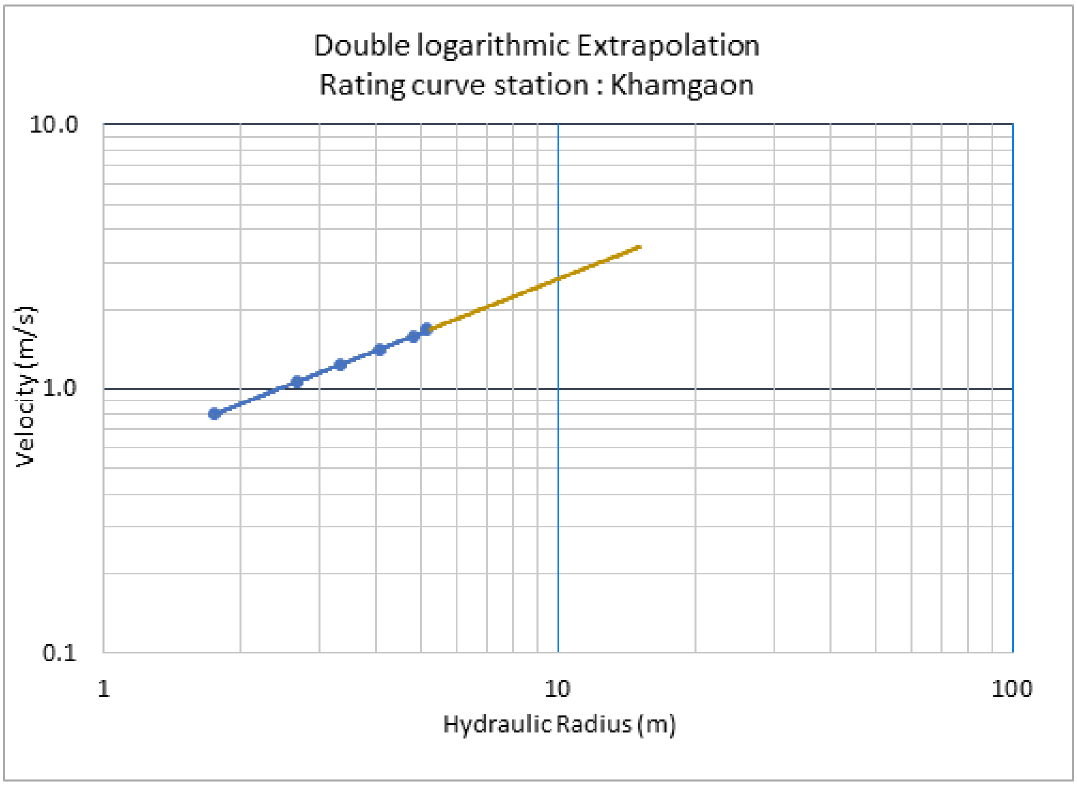


Figure 14.4: Example of Double Logarithmic Extrapolation of Mean Velocity Against Hydraulic Radius

14.2.3 The Manning's equation method

A slight variation of the stage-area-velocity method is the use of Manning's equation for steady flow. In terms of mean velocity, the Manning's equation may be written as:

$$v = K_m R^{2/3} S^{1/2} \quad \text{Equation 14.1}$$

Since for higher stages the value of $K_m S^{1/2}$ becomes nearly constant, the equation can be rewritten:

$$v = K^* R^{2/3} \quad \text{Equation 14.2}$$

or $$K^* = v/R^{2/3} \quad \text{Equation 14.3}$$

The relationship of the stage (h) to K^* is plotted from discharge measurements. This curve often approaches a constant value of K^* at higher stages (Figure 14.5). The value of K* may then be used in conjunction with extrapolated relationships between h and A and, h and $R^{2/3}$ based on the surveys. The values of discharge for the extrapolated stage are then obtained by applying the Manning's equation with K^* and extrapolated values of A and $R^{2/3}$.

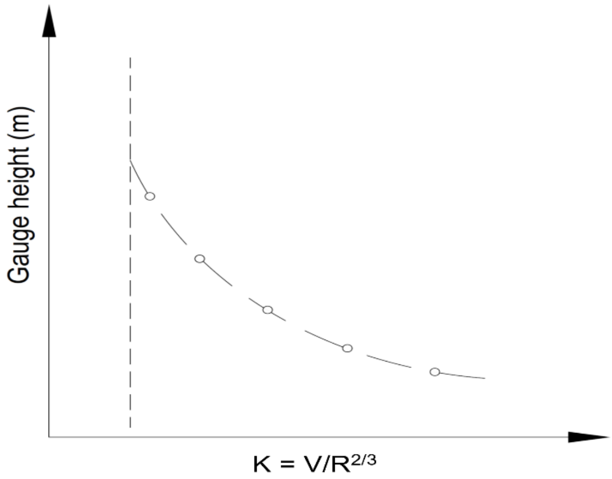


Figure 14.5: K* Versus Gauge Height (Adopted from Herschy, 2009)

Above the bank full stage, the discharge that passes through the floodplain must be determined separately by assuming an appropriate K_m value, as is done using the conveyance slope method described later. This method was applied to the Khamgaon river cross-section data shown in Figure 14.1 and the observed discharges. The steps and results are shown in Figure 14.6 to Figure 14.8.

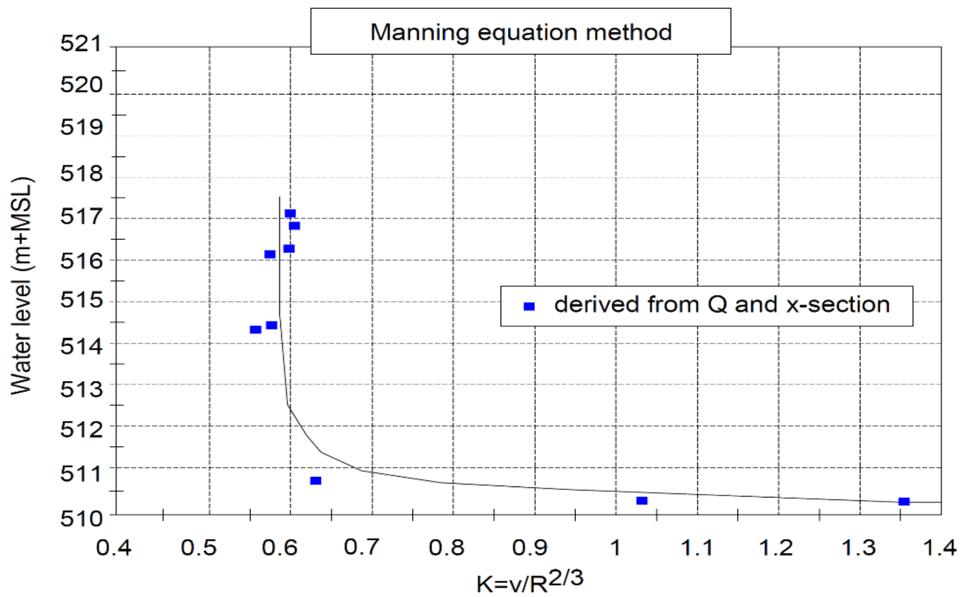


Figure 14.6: K* Versus Gauge Height for the Khamgaon Example

It can be seen from Figure 14.6 that K^* indeed tends to an approximately constant value for the higher stages, which was subsequently applied for the extrapolation. Together with the cross-section area shown in Figure 14.7 and the hydraulic radius of the river, the flow through the main section was computed. For the flood plain, Manning's equation was applied separately. The result is shown in Figure 14.8. In this figure, the result of the double logarithmic extrapolation technique has also been shown for reference. It is observed that the flow through the main river is approximately the same by the two methods. However, the total flow with the Manning's equation is larger, since in this method the flow through the floodplain is duly accounted for.

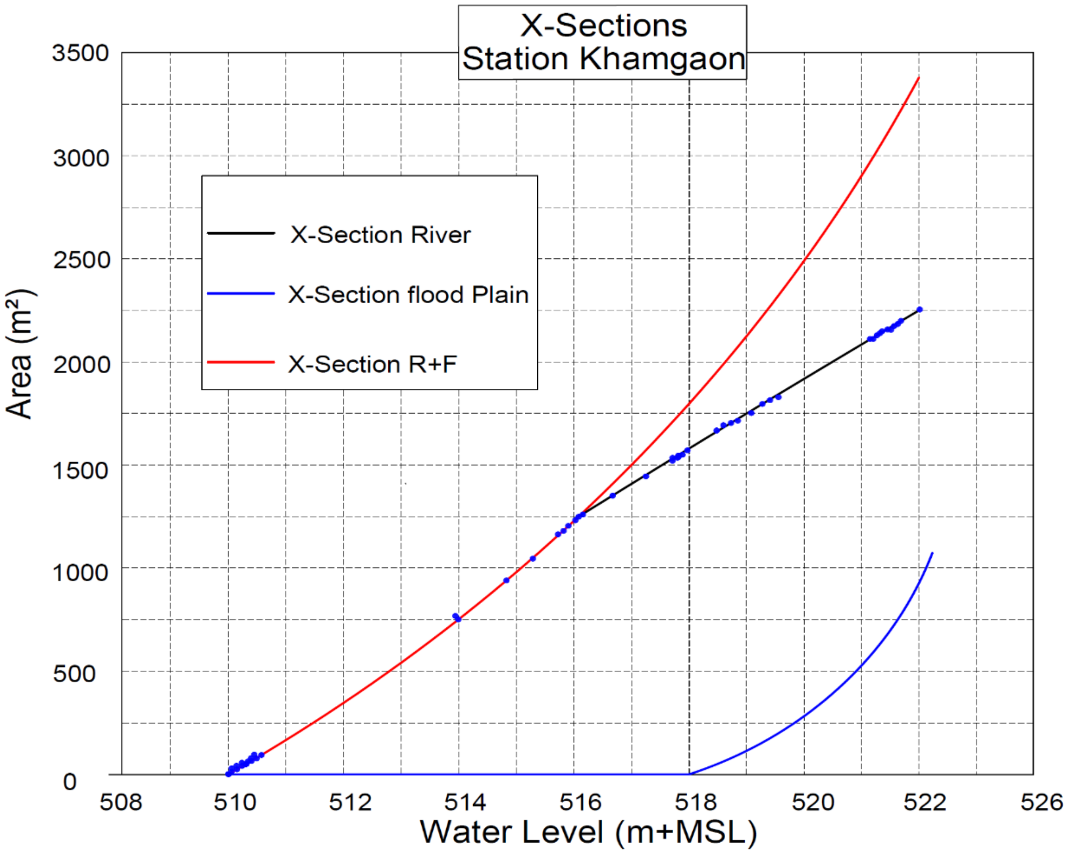


Figure 14.7: Cross-Sectional Areas of River and Flood Plain in Khamgaon Example

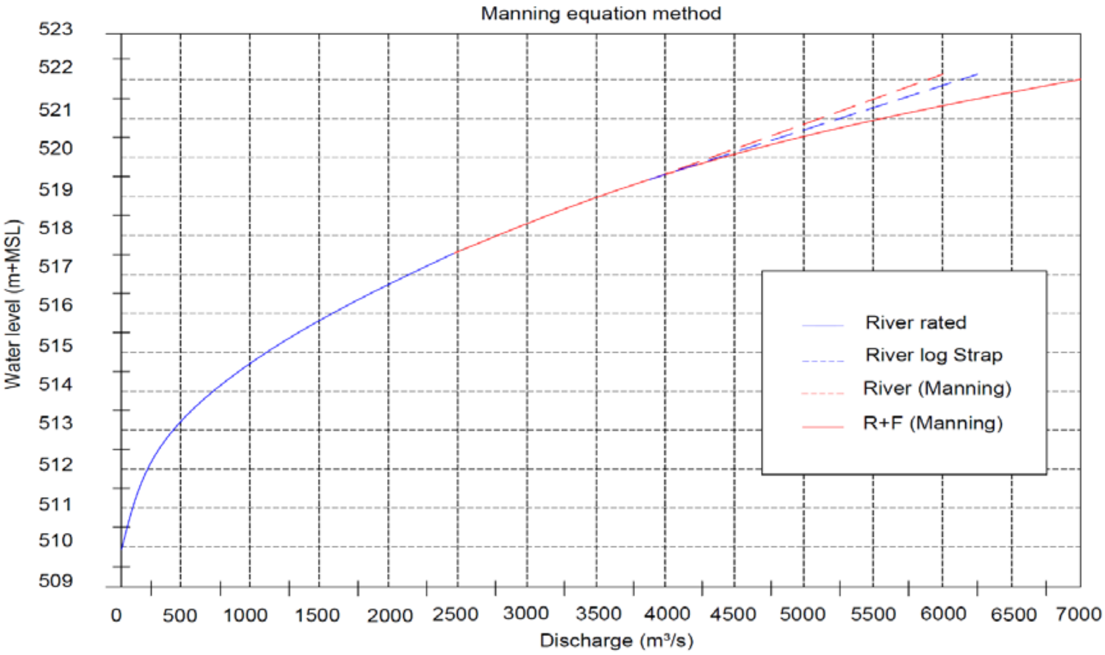


Figure 14.8: Extrapolation Based on the Manning's Equation Method Compared with Double-Logarithmic Extrapolation

14.2.4 The conveyance slope method

In the conveyance slope method, the conveyance and the energy slope are extrapolated separately. It has greater versatility than the methods described above and can be applied on sections with the overbank flow. It is therefore recommended for use. It is also based on Manning's equation:

$$\boldsymbol{Q = K_m R^{2/3} S^{1/2} A} \qquad \text{Equation 14.4}$$

or:

$$Q = K\,S^{½} \qquad \text{Equation 14.5}$$

where the conveyance is:

$$K = K_m\,A\,R^{2/3} \qquad \text{Equation 14.6}$$

For the assessment of K for the given stage, A and R are obtained from the field survey of the discharge measurement section, and values of n are estimated in the field. Values of K are then plotted against the stage up to the maximum required level (usually on natural graph paper), as shown in Figure 14.9. The extrapolation of conveyance has been demonstrated in Figure 14.10.

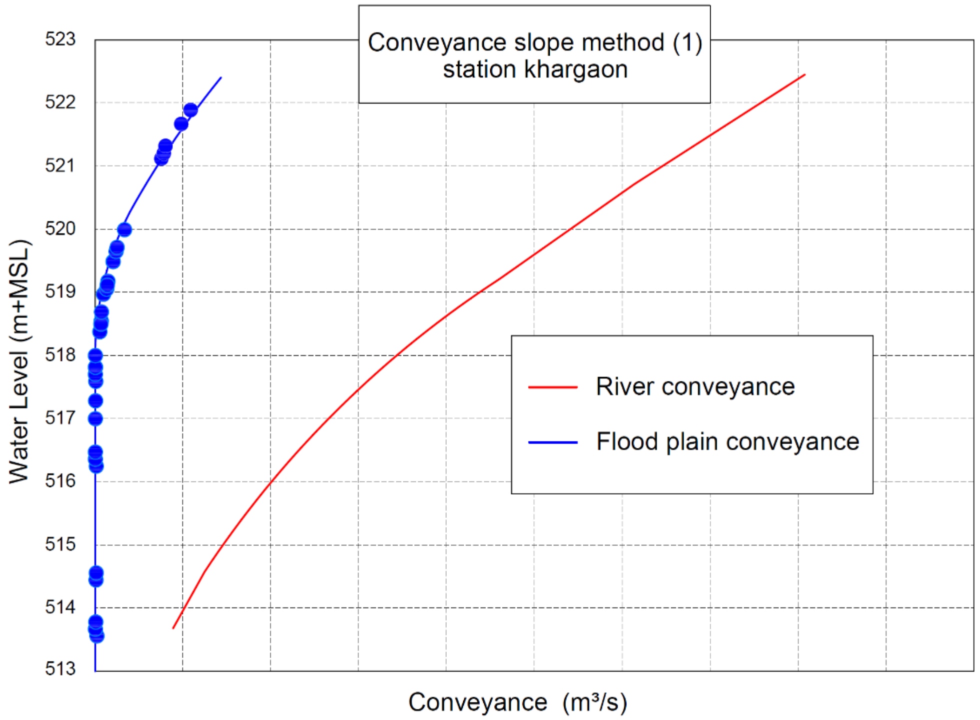


Figure 14.9: Conveyance as f(h)

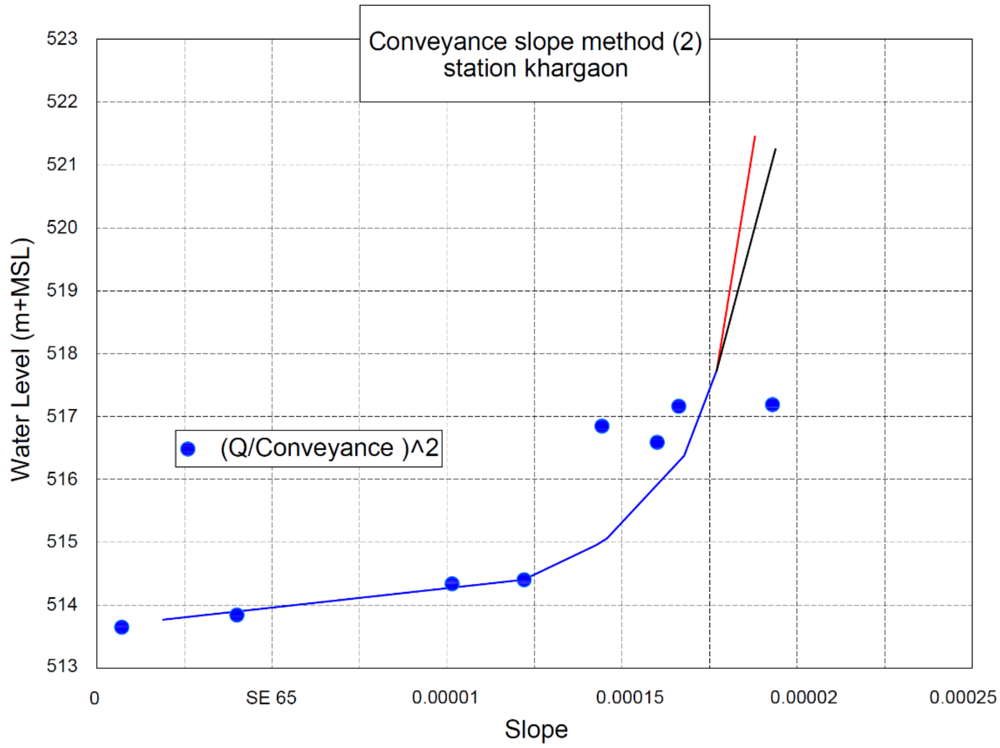


Figure 14.10: Slope Extrapolation

Values of *S*, which is the energy gradient, are usually not available. But for measured discharges, $S^{1/2}$ can be computed by dividing the measured discharge by its corresponding K value. The value of S is then calculated and plotted against stage on natural graph paper and extrapolated to the required gauge height, with the knowledge that S tends to become constant at higher stages in most of the cases. It has been shown for the Khamgaon case in Figure 14.10.

The discharge for a given gauge height is obtained by multiplying the corresponding value of K from the K curve by the corresponding value of $S^{1/2}$ from the S curve. It should be noted that in this method, errors in estimating K_m have a minor effect, because the resulting percentage error in computing K is compensated by a similar percentage error in the opposite direction in computing $S^{1/2}$.

The whole procedure can be accomplished in various stages as given below:

Computation of cross-section data

The cross-section should first be obtained as:

- distance (x) from an initial point
- depth (y) and
- depth correction (y_c)

The depth correction y_c may be introduced to quickly evaluate the effects of changes in the cross-section on geometric and hydraulic parameters.

The actual depth y_a is computed from:

$$y_a = y + y_c$$

A plot of the cross-section and the levels at fixed interval can be made, and the following quantities computed (see Table 14-1):

- surface width, (B)
- wetted perimeter, (P)
- cross-sectional area, (A)
- hydraulic radius, (R): R = A/P
- factor {area × (hydraulic radius)$^{2/3}$}, ($AR^{2/3}$)

Computation of cross-section data & Computation of cross-section parameters

Station: Khamgaon

Date: 1977 / 1/ 1

Coordinates of cross section profile:

Table 14-1: Example on Stage-Discharge Extrapolation

Distance from the Origin (m)	Elevation (m)
-160.00	521.54
-150.00	521.30
-140.00	520.90
-130.00	520.68
390.00	521.03
400.00	520.93
410.00	521.30
420.00	522.03

Section 1

Left bound -160.00 m, right bound 00 m from the initial point

Water boundaries included

K-Manning: 30.0, $S^{1/2}$K - Manning: 0.4135

Stage (m)	Width (m)	Wetted Perimiter (m)	Area (m^2)	Hydraulic Radius (m)	A $R^{2/3}$	Flow (Q) m^3/s
512.00	0.00	0.00	0.00	0.00	0.00	0.00
513.00	0.00	0.00	0.00	0.00	0.00	0.00
514.00	0.00	0.00	0.00	0.00	0.00	0.00
515.00	0.00	0.00	0.00	0.00	0.00	0.00
516.00	0.00	0.00	0.00	0.00	0.00	0.00
517.00	7.14	8.74	0.72	0.08	0.13	0.06
518.00	45.57	47.23	25.86	0.55	17.30	7.16
519.00	106.14	107.98	113.08	1.05	116.61	48.22
520.00	125.14	127.20	230.06	1.81	341.53	141.23
521.00	142.50	144.76	360.64	2.49	662.77	274.07

Section 2

Left bound.00 m, right bound 230.00 m from the origin (initial point)

Water boundaries included

K-Manning: 40.0, $S^{1/2}$K - Manning: 0.5514

Stage (m)	Width (m)	Wetted Perimiter (m)	Area (m^2)	Hydraulic Radius (m)	A $R^{2/3}$	Flow (Q) m^3/s
512.00	158.04	158.48	277.67	1.75	403.55	222.50

Stage (m)	Width (m)	Wetted Perimiter (m)	Area (m^2)	Hydraulic Radius (m)	A R $^{2/3}$	Flow (Q) m^3/s
513.00	163.95	164.72	438.66	2.66	842.81	464.69
514.00	181.98	182.98	608.46	3.33	1355.54	747.39
515.00	195.83	197.00	798.55	4.05	2030.17	1119.36
516.00	205.89	207.25	999.41	4.82	2852.65	1572.84
517.00	226.34	227.86	1218.69	5.35	3727.16	2055.02
518.00	230.00	232.73	1447.88	6.22	4897.61	2700.35
519.00	230.00	234.73	1677.88	7.15	6226.16	3432.87
520.00	230.00	236.73	1907.88	8.06	7669.16	4228.48
521.00	230.00	238.73	2137.88	8.96	9219.35	5083.20

Section 3
Left bound 230.00 m, right bound 420.00 m from initial point Water boundaries included
K-Manning: 30.0, $S^{1/2}$K - Manning: 0.4135

Stage (m)	Width (m)	Wetted Perimiter (m)	Area (m^2)	Hydraulic Radius (m)	A R $^{2/3}$	Flow (Q) m^3/s
512.00	0.00	0.00	0.00	0.00	0.00	0.00
513.00	0.00	0.00	0.00	0.00	0.00	0.00
514.00	0.00	0.00	0.00	0.00	0.00	0.00
515.00	0.00	0.00	0.00	0.00	0.00	0.00
516.00	0.00	0.00	0.00	0.00	0.00	0.00
517.00	0.00	0.00	0.00	0.00	0.00	0.00
518.00	62.39	63.27	27.71	0.44	15.98	6.61
519.00	100.21	101.23	99.98	0.99	99.16	41.00
520.00	134.52	135.69	219.43	1.62	302.31	125.01
521.00	167.39	168.80	359.94	2.13	596.30	246.58

Stage	Discharge/section			Total Discharge
512.00	.00	222.50	.00	222.50
513.00	.00	464.69	.00	464.69
514.00	.00	747.39	.00	747.39
515.00	.00	1119.36	.00	1119.36
516.00	.00	1572.84	.00	1572.84
517.00	.06	2055.02	.00	2055.07
518.00	7.16	2700.35	6.61	2714.12
519.00	48.22	3432.87	41.00	3522.09
520.00	141.23	4228.48	125.01	4494.72
521.00	274.07	5083.20	246.58	5603.85

These parameters may be determined for the whole cross-section or parts of cross-section, i.e., for the main river and flood plain separately.

It must be noted that when the cross-section is divided, the wetted perimeter for each part may be determined in two ways (refer Figure 14.11):

- the water boundary not considered:
  - for flood plain : Pfloodplain = ABC
  - for the main river : Priver = CEFG
- the water boundary is treated as a wall:
  - for the flood plain : Pfloodplain = ABCD
  - for the river : Priver = DCEFG

The latter option appears to be more realistic to account for the lateral transport of momentum between the river and the flood plain. In general, it reduces the discharge carrying capacity of the main channel. Both options are included to maintain consistency with hydraulic computations, where generally the first approach is used.

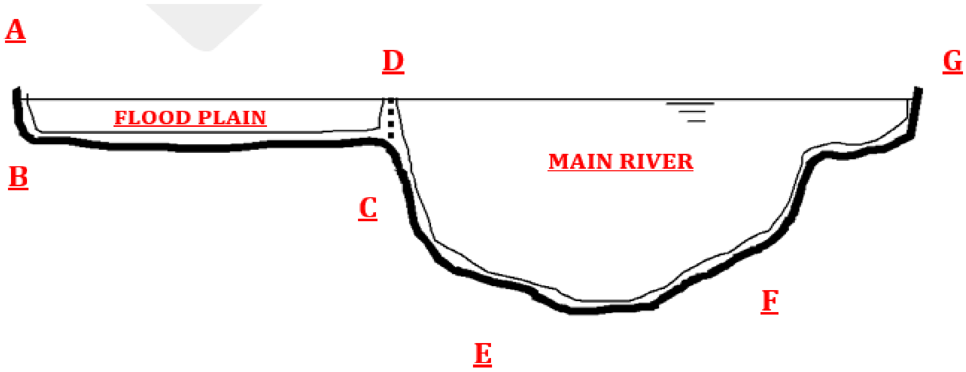


Figure 14.11: River flow in the Main Channel and Flood Plain During High Stages

Computation of hydraulic quantities in the measured range

Next, the geometric and hydraulic quantities are obtained in one of the following ways:

- from the stage-discharge database, provided that the cross-sectional parameters are also given
- from a combination of the cross-section profile and the rating curve

The following parameters are obtained for various depths:

- surface width (B)
- wetted perimeter (P)

- cross-sectional area (A)
- hydraulic radius (R): R = A/P
- factor {area x (hydraulic radius)$^{2/3}$}, (AR$^{2/3}$)
- discharge (Q)
- average velocity (v): v = Q/A
- Conveyance (K): K = (1/n) (AR$^{2/3}$); where n is an estimated value
- Slope (S): S = (v/K)^2

Estimation of discharge in the extrapolated range

The estimated values of slope (S) in the measured range are plotted against the stages. Extrapolation is made when this curve is asymptotic to the bed slope at higher stages. The conveyance curve is also plotted to make use of the estimated values of K for the full range of stages. For any stage in the extrapolated range, the value of K and S are read from the two curves and the product of these two quantities and the area of cross-section (A) yield the estimated value of discharge (Q).

After synthetic stage-discharge data is obtained for the extrapolated range, these data are incorporated in the set of stage-discharge data. Subsequently, new attempts can be made to fit the rating equation to the measured and estimated stage-discharge data.

A comparison of various methods of stage-discharge extrapolation is provided in Table 14-2 for further reference.

Table 14-2: Methods of Rating Curve Extrapolation

Sl. No.	Methods	Description	Limitations
1	Logarithmic plotting	If the control shape does not change significantly and the channel roughness remains fairly constant	Suited to channel control conditions for medium and high flow. Should not be used to extrapolate more than 1.5 times the highest measured discharge
2	Manning's or Chezy's equation	Special care is needed if the shape of the cross-section changes appreciably, because friction slope may also change significantly	Friction slope of bank full discharge and flow within banks may be significantly different
3	Velocity Area method	Dependant on the stage-velocity relationship, which can be accurate only in the	Accuracy is questionable in the range above the highest measurements

Sl. No.	Methods	Description	Limitations
		range where discharge measurements are available	

14.3 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 cannot be plotted on such paper. Also, the values are all in the low range, so there is no benefit from using the log paper. A visually-guided curve is drawn between the lowest point of the known rating to the known point of zero flow, obtained by observation or by the survey of the low point of the control. There is no assurance that the extrapolation is precise, but improvement can only be made from further low flow discharge measurements. However, low flows persist for a sufficient period for the measurement to be carried out and the record to be updated, and there is little physical difficulty in obtaining such measurements.

Often, the power type equation is used for the lowest segment with stage-discharge observations assumed to hold good below the measured range as well. The shift parameter 'a' is either determined based on the measurements by the system or is introduced based on cross-sectional information.

15 BASIN MODELLING CONCEPTS

15.1 Introduction

There is abundant river infrastructure in river basins in India, where irrigation has been a tradition for centuries. India has well over 5000 dams today, with plans to build more. Proper design of future infrastructure must therefore address the issues of joint operation of the existing infrastructure with the planned infrastructure, so as to assess the best configuration of the system that will maximize the potential benefits while minimizing the investment and maintenance costs. This assessment of the best configuration and the best way to operate the existing (or future) basin configuration is part of the development of a river basin plan. To this end, computer modelling has become an integral part of river basin planning in most developed countries, and India is also making steps in this direction. This section will provide some insight into the required input data for river basin modelling, as well as highlight the current state of the art related to the limitations and capabilities of the existing tools.

15.2 Input Data Requirements for River Basin Planning Models

In most general terms, river basin planning is conducted by using lengthy time series of runoff estimates that are matched in the model with the present of future water demands. There are four types of input data:

- Runoff time series data
- Water demand time series data
- Physical information (storage and canal capacities, network connectivity, etc.)
- Operational / management objectives

The usual approach is to develop estimates of runoff based on historical flow records that provide input into the existing storage reservoirs. Another popular approach is to use rainfall-runoff models, which are typically calibrated on a few years of data, and then run on a longer series of rainfall inputs to generate flow estimates. The advantage of this approach is that it can help generate flow estimates for the years when there were no flow records. However, this approach should involve the validation (also known as the "verification") phase, where the calibrated model parameters are used to verify the ability of the model to predict runoff in historic years for which the data are available but were not used in the initial model calibration. It is often the case that rainfall-runoff models show a rather poor performance in this phase, as shown previously in Figure 3.11, since the difference between the historical flow records and simulated flows can be significant. If is for this reason that the use of rainfall-runoff models should be restricted to ungauged basins, and used only as the last resort option. In river basins with available records, a preferable way to generate runoff estimates is to conduct naturalization of flows, which is further detailed below.

15.3 Development of Natural Flow Estimates

The Natural flows are river flows that would have been observed at selected locations in a river basin assuming there had been no human intervention by operation of large storage reservoirs or withdrawals. The most common approach to estimate natural flows in gauged basins is the Project Depletion Method, which is essentially aimed at "undoing" the impacts of human intervention in a systematic way, reach by reach, in a downstream progression.

The following section explains the calculation procedure on a small example shown in Figure 15.2 that has most of the elements found in complex river basins. There are two river reaches with a reservoir R_1 at their confluence. In this example natural flow is calculated at the reservoir site. There is one diversion (D_1) and one return flow (RT_1) into the reservoir, one diversion channel out of the reservoir (D_2), and regulated outflow from the reservoir into natural channel reach C_3. The general approach to calculate natural flows at any location is to estimate local runoff which originates between the given location and the closest upstream locations at which natural flows had already been evaluated. Denote the natural flow at reservoir as Q_{R1} and the local runoff between natural flows Q_1, Q_2 and the reservoir as LR. The natural flow at the reservoir site can then be calculated as:

$$Q_{R1} = Q_1 + Q_2 + LR \qquad \text{Equation 15.1}$$

Consequently, the principal component of estimating natural flows is determination of the local runoff LR. Assuming Qr1 and Qr2 are the recorded flows at locations 1 and 2, LR for the reservoir in Figure 15.2 can be calculated using the following equation assuming average flow over time step t:

$$LR = Q_{C3} + Q_{D2} - Q_{RT1} - Q_{D1} + \Delta V/t - Q_{r1} - Q_{r2} \qquad \text{Equation 15.2}$$

where:

Q_{C3} the recorded flow in channel C_3

Q_{D1} flow in diversion channel D_1

Q_{D2} flow in diversion channel D_2

Q_{RT1} flow in return flow channel RT_1

$\Delta V/t$ reservoir storage change over time step t

It should be noted that the use of short (daily) time steps in the calculation would usually require that the upstream flows be routed to their downstream ends by using a hydrologic routing scheme such as the SSARR routing explained earlier in this document. For short time steps, the routing is also required for calculating the final natural flows at a given location by summing up the local runoff with the routed estimates of natural flows at the immediate upstream locations.

Reservoir storage change is further evaluated using the starting and ending storage (Vs and Ve) for a time step, along with adjustments for net evaporation (evaporation minus precipitation) for a given time interval t (seconds). Note that the sign for net evaporation is reversed since the idea is to remove the effect of net evaporation (i.e. put the evaporation loss back in the river):

$$\frac{\Delta V}{t} = \frac{Ve-Vs}{t} + \frac{(E-P)[A(Ve)+A(Vs)]}{2t} \qquad \text{Equation 15.3}$$

where:

V_e volume at the end of time step t (m^3)

V_s volume at the start of time step t (m^3)

P total precipitation over time step t (m)

E total evaporation from the reservoir surface over time step t (m)

$A(V_e)$ surface area (m^2) corresponding to the ending volume Ve

$A(V_s)$ surface area (m^2) corresponding to the starting volume Vs

To summarize, local runoff LR can in general be assessed by conducting a water balance calculation for a sub-catchment which is delineated by the downstream point for which LR is evaluated and the upstream control points where recorded flow series are available. The general expression is:

$$LR = \sum_{i=1}^{m} Qi - \sum_{j=1}^{n} Qj + \sum_{k=1}^{l} \frac{\Delta Vk}{t} \qquad \text{Equation 15.4}$$

where:

Q_i average outflows (i=1, m) from a sub catchment within time step t

Q_j average inflows (i=1, m) into a sub catchment within time step t

while the storage change term $\Delta V/t$ is summed up over all storage reservoirs in the sub-catchment area under consideration. Inflows and outflows into a sub-catchment include all diversions and return flows into it, as well as diversions out of it.

Normally, natural flows should be calculated at all on-stream reservoir locations, especially when reservoirs have sizeable live storage.

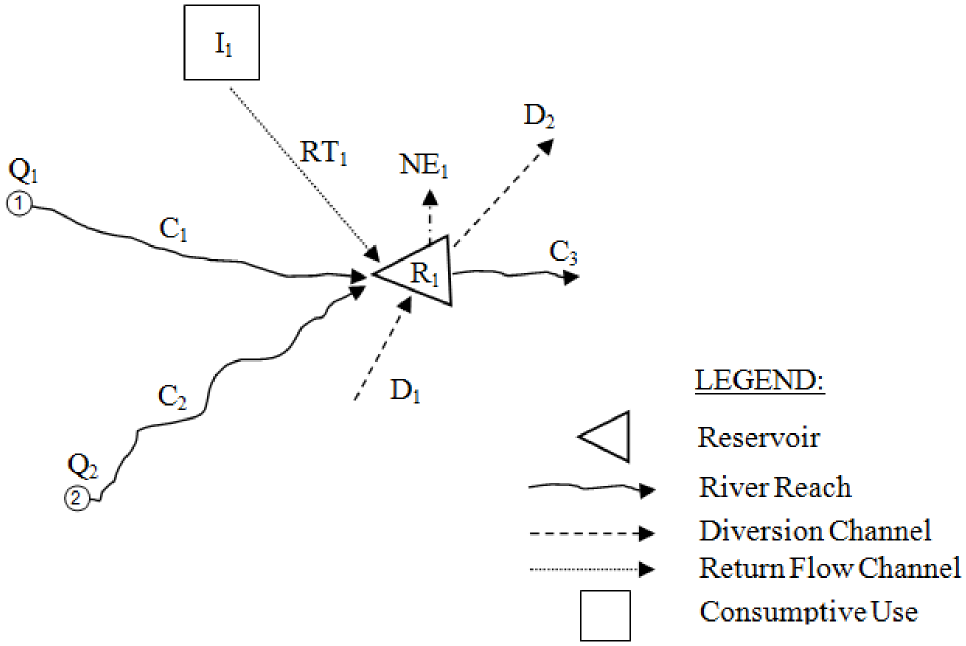


Figure 15.1: Sample Schematic for Calculation of Natural Flows

Equation (15.1) suggests that natural flows be first determined at upstream locations (e.g. locations 1 and 2 in the example in Figure 15.2). The calculation then proceeds in the above manner for all requested locations in the river basin in a downstream progression. Routing should be used for daily calculation. Otherwise, weekly or 10-daily calculations can be conducted without routing, assuming steady state conditions.

15.4 Development of Modelling Schematic

The term "Model" implies the representation reality. Since real world systems can be large and complex, the first task for river basin modellers is to decide on the scale and the level of detail that will be included in the model. That requires the development of a modelling schematic. A modelling schematic is usually built on top of a map, by breaking down the river system into a set of interconnected river reaches, diversion canals, irrigation blocks, return flow channels, reservoirs and junction nodes. An example modelling schematic is shown in Figure 15.2 .

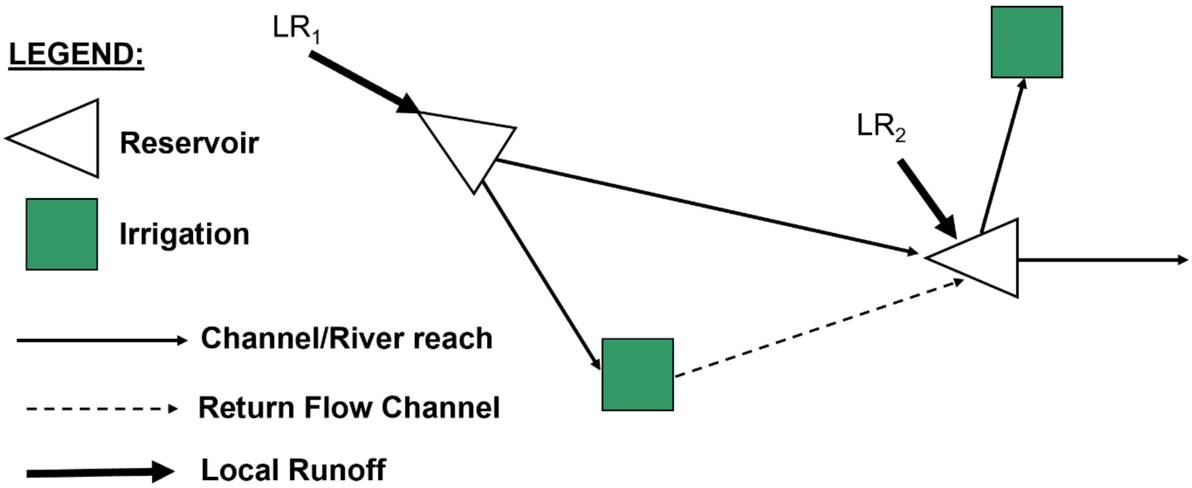


Figure 15.2: Example of a River Basin Modelling Schematic

It can be observed from the above schematic that the total inflow into a reservoir generally consist of the local runoff and the outflow from upstream reservoir. A return flow channel shown above typically represents a portion of consumptive use that is returned to the stream, indicating the point of return as the downstream node of the return channel.

Selection of the proper modelling scale involves the appropriate breakdown of the entire river basin into a manageable number of components. Small diversions along a river reach can be combined into a single equivalent water use component, and similar approach can be used for smaller tributaries, which can be summed up along one reach and represented as an equivalent tributary at the end of the reach, as for example the local runoff (LR_2) inflow into the downstream reservoir. In general, the following data should be collected and inspected as part of the model setup:

Spatial data (Physical structures definition and location)

- The list of all hydro-meteorological stations with their data length
- List of all structures to be included in the model – Dams / Barrages / Diversion canals / Hydro Power Plants / command areas / municipal and industrial water users. These are used in the development of modelling schematic.

Physical structures data

- Elevation-Area-Capacity/tables for reservoirs
- Outflow vs elevation curves / tables for all reservoir outlet structures
- Pump / canal maximum flow capacities

Hydropower Plants

- Installed capacity

- Net head – flow – efficiency curves for each turbine
- Tail water elevation (fixed level, or rating curve or pool level od downstream reservoir)
- Head losses (if any) as a function of flow

Hydro-meteorological historical timeseries data

- Daily Rainfall (include the stations at reservoirs if available)
- Daily/Monthly Evaporation data (pen evaporation or previous estimates of potential ET from other studies)
- Daily Observed Discharge from all gauged sites (include dam releases through all existing outlets)
- Daily observed water levels for all reservoirs
- All water demands (Irrigation/domestic/industrial/inter-basin transfers or any other demands in basin) timeseries (typically weekly or 10-daily time steps).
- Estimates of infiltration losses to seepage from reservoirs and irrigation canals if available

Operational data

- Reservoir Guide Curves, maximum, normal and minimum operating level throughout the year
- Water management guidelines, operating priorities and / or deficit sharing policies (if any)

Some of the above data related to historical reservoir levels and outflows is used in the process of developing natural flows that is later used as model input. While the choice on the number of modelling components defines the spatial resolution of the model, it is also important to properly define the temporal resolution, i.e. the adequate length of the modelling time step.

The proper selection of the time step is important due to the following assumptions made in most river basin models:

a) Theoretically, the model should be able to inspect many different operating scenarios, which are based on the assumption that any user can be supplied from any reservoir located upstream of it; and,
b) All model releases are driven by downstream demands, which can be off-stream (irrigation) or on stream (environmental flow targets). This helps implement a water conservation policy, since the model should have no unnecessary spills from any of the reservoirs.

Numerous studies in the past have involved modelling large basins by using a daily calculation time step. This is not feasible since the above assumptions a) and b) are not sustainable in river basins where the travel time between the most upstream reservoir and the most downstream user significantly exceeds the calculation time step length. Consider for example the basin in Figure 15.3

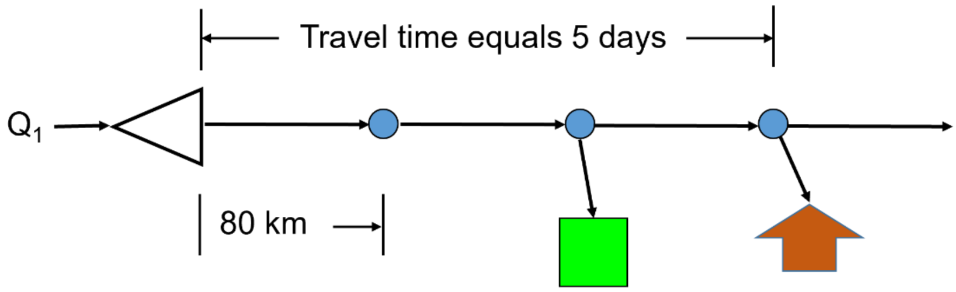


Figure 15.3: Sample Modelling Schematic with Five Days of Travel Time

Assuming the average streamflow velocity of 1 m/s, any water released from storage will only reach up to 86.4 km per day. Hence, the assumption that the reservoir release would be available to the most downstream water user within the length of the calculation time step is incorrect. Most modellers claim that the introduction of river routing in the model will address this issue, but this assumption is also incorrect. River routing will modify the steady state flow by introducing the channel storage change, such that the total inflow at the upstream end of the channel no longer equals the total outflow at its downstream end, but it will not make water releases from the reservoir travel any faster than in the steady state solution. For typical river basins in India, the total travel time through the basin exceeds 5 days, which justifies the selection of a 10 daily time step, since the calculation time step should be at least twice the length of the total travel time. For very large basins, steady state modelling is limited to monthly time steps. This has disadvantages as it negatively affects the accuracy of modelling. Monthly time steps are unrealistic when it comes to modelling reservoir spills, which appear to be significantly reduced in monthly simulations compared to weekly simulation of the same systems. Local in-house studies conducted internally by Alberta Environment and Parks, Canada have shown the average difference in system spills of 28%, meaning that the same simulation with weekly time steps has 28% higher spills that the results of identical simulation with monthly time steps. What was spilled in a weekly run was allocated to water users in a monthly simulation, thus making the results of the monthly simulation much more optimistic. This should come as no surprise when the shape of the monthly hydrograph is compared with weekly or daily, as in Figure 15.4.

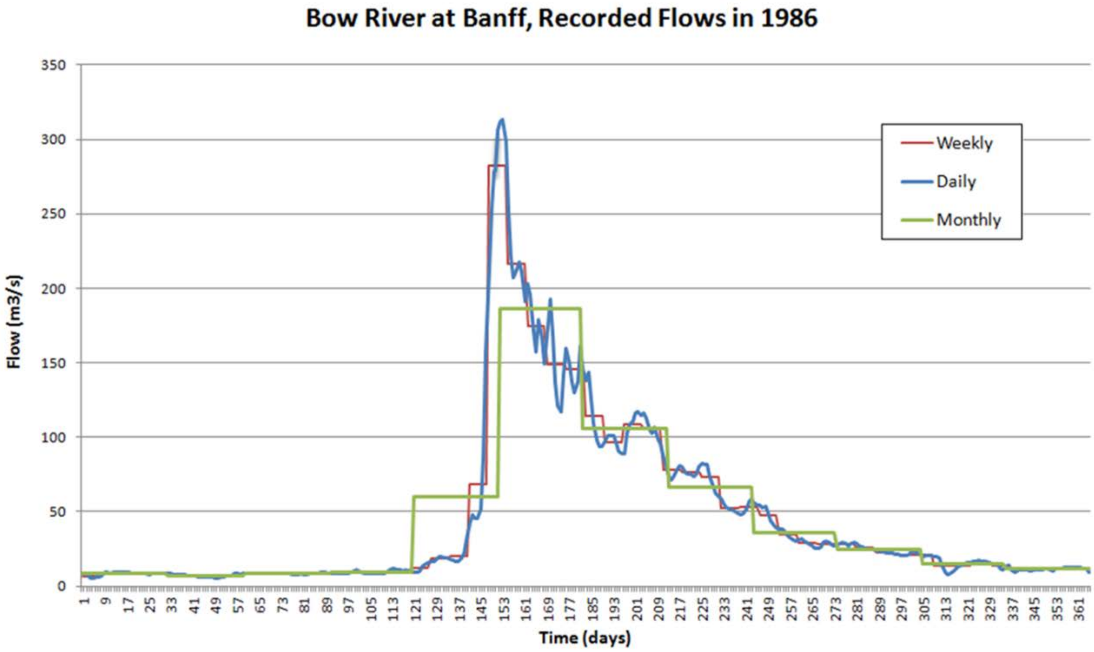


Figure 15.4: Comparison of daily, Weekly and Monthly Hydrographs

It is easy to see the that hydrograph peaks are lost in the monthly hydrograph which has a constant flow for the entire month, thus making it easier to handle than the weekly or daily inflows, where reservoir spills are inevitable during high peak flows.

15.5 River Basin Model Constraints

Constraints are functions that limit flows in the network. They set certain mathematical conditions that are expected to be satisfied by the model solution. They are typically divided into physical and operational constraints.

15.5.1 Physical constraints

15.5.1.1 Mass balance at the nodes:

This constraint implies that the total sum of all inflows, outflows (and storage change in case of reservoirs) has to equal zero for any node.

15.5.1.2 Return flows from irrigation blocks:

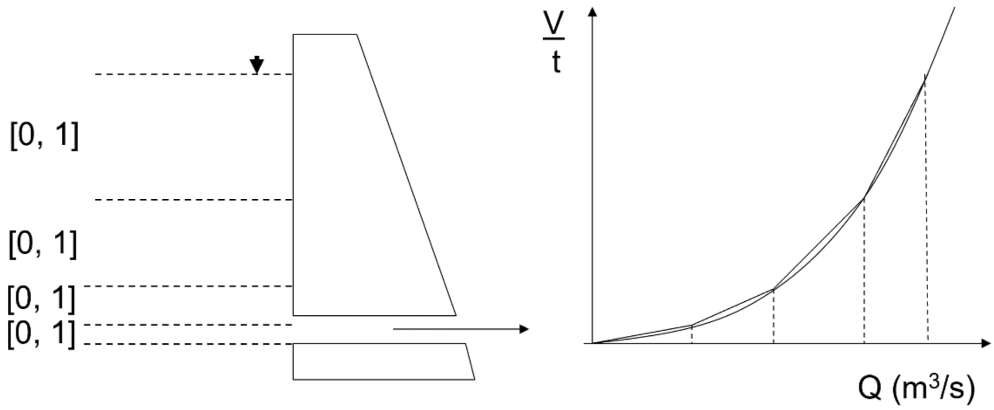
Return flows are dependent on the level of consumptive use, and the model should be able to set the return flows as a fraction of consumptive use dynamically during the simulation run. This fraction should be flexible, in the sense that it can vary from one time interval to another, if so desired.

15.5.1.3 Net evaporation at reservoirs:

Reservoirs incur losses that are a function of both meteorological input (evaporation and precipitation) as well as the water surface areas at the end of each simulated time step, which depend on the model results. A relationship between surface area and stored volume needs to be taken into account to allow the calculation of net evaporation as a function of the average storage over each simulated time step. This constraint is particularly challenging when models are asked to solve more than one time step simultaneously. The relationship between storage and water surface area can be linearized using the proper number of linear segments.

15.5.1.4 Reservoir outflow constraints:

Reservoir outflows are a function of the average storage over a simulated time step, based on the outflow vs elevation curve given for a particular outlet structure. This is a non-linear relationship which can be linearized using piece-wise segmentation of the outlet curve. One challenging issue encountered by the model developers was the tendency of the model to fill the upper storage zones, which have a higher outflow capacity, while leaving the lower storage zones empty (see Figure 15.5).



$$S = \frac{Vs - Ve}{Q1 - Q2} \cdot \frac{1}{t} \qquad Qmax(o) = \frac{1}{S} \cdot \frac{1}{2}\left(\frac{Vs}{t} + \frac{Ve}{t}\right) \qquad Qt(o) \leq \sum_{i=1}^{n} \frac{1}{Si} \cdot \frac{1}{2}\left(\frac{Vs(i)}{t} + \frac{Ve(i)}{t}\right)$$

Figure 15.5: Piece-wise Linearization of the Reservoir Outlet Curve

To ensure the proper sequence of filling reservoir zones from the bottom up, it was necessary to introduce binary variables associated with each storage zone that corresponds to the segment of the reservoir outflow curve. When binary variable related to a particular zone equals 0, there is no storage in that zone. When it equals 1, the zone will have water stored in it. The model constraint is that the values of the binary variables of any zone are greater or equal to the values of the zone immediately above it. If this condition is satisfied, the storage zones will be filled properly. Based on this and other similar constraints, most river basin models that use linear programming solution technique (RiverWare, Oasis, WEB.BM), use the so-called Mixed Integer Programming solvers, which can include a mix of decimal and integer variables (binary variables are a special type of integer with only two possible values, 0 or 1).

15.5.1.5 Weir flow_constraint:

Diversion from the river may sometimes be a function of the incoming flow in the river at the point of diversion. Typically, the lateral weir function will allow portion of the diverted flow from the river based on the hydraulic head and the opening of the weir cross section. Modelling should take into account the maximum flows that can be diverted from the river as a function of the incoming river flow. This function can also be linearized using piece-wise segmentation. This constraint could represent both the physical flow limits as a function of flow in another channel, or in some instances it could represent operational constraints imposed by the basin managers.

15.5.1.6 Channel routing constraints:

A function that converts inflow into a channel to its outflow based on the channel storage changes. This function is necessary when modelling time steps that are shorter than the total travel time in the basin, and it should be used in such instances in combination with multiple time step solution mode, as explained din the following section.

15.5.1.7 Hydropower_constraints:

This constraint calculates hydro power as a function of the net head, flow through the turbines, and efficiency factor. The Net Head is the difference between the head water elevation and tail water elevation, where the options for either the head or the tail water elevation should be:

- constant level;
- rating curve of the river reach (upstream or downstream); and,
- average storage level for upstream (Head Water Level) or downstream (Tail Water Level) storage.

15.5.2 Operational constraints

The following operational constraints are commonly required in river basin modelling:

15.5.2.1 Biological minimum flows:

These constraints are typically modelled as a soft target for various river reaches. As such, they may not always be met in extremely dry years. This operational constraint can also be used to model minimum flows at border crossings, minimum flows required to maintain navigation routes or required reservoir releases for sediment flushing.

15.5.2.2 Maximum instantaneous diversion flow:

This limit incorporates the opening and closing policy for irrigation canals, as well as flow maintenance policy that the management wishes to implement.

15.5.2.3 Maximum diversion volume for an irrigation season:

most water users in many jurisdictions around the world have an annual maximum limit that they are not allowed to exceed in any year when diverting water from the river. This constraint is part of their water license.

15.5.2.4 Apportionment agreements:

This constraint represents the minimum annual flow volume that has to be passed to a downstream state. This volume can be a percentage of natural flows and as such it can differ from year to year.

15.5.2.5 Equal Deficit Sharing among selected components:

This component ensures equal deficits throughout the irrigation season for one or more irrigation blocks.

15.6 Solution Modes

There are three possible solution modes for river basin management models:

15.6.1.1 Single Time Step (STO) solution mode.

When in this solution model, the model makes a decision on reservoir releases in a single time step without and considerations of the demands and inflow forecasts in subsequent time steps. The underlying assumption is that the time steps are long enough compared to the travel time in the basin such that each water user can be supplied by any of the upstream reservoirs. The STO model setup requires the use of reservoir rule curves so as to prevent the reservoirs from premature emptying.

15.6.1.2 Multiple Time Step (MTO) solution mode.

This mode should be executed for a selected number of time steps, which should typically include a number of time steps solved simultaneously. Figure 15.6 shows a scheme that solves three time steps simultaneously, where the ending storage from one time step is provided as the starting storage for the same reservoir in the next time step that are connected using the carry over storage arcs.

The MTO solution concept is necessary when modelling large river basins with daily time steps, regardless of whether the model derives solutions on the basis of using optimization or simulation algorithms. It shows that a release made from the reservoir on the first day will have to travel through the system over the next two days before it can reach the downstream diversion on the third day. This process should also include the channel storage changes properly, but taking into account additional tributaries and other diversions along the way. A numerical example is provided at the end of this section that demonstrates these concepts on a real-world problem. There have not been many models which have the capability of MTO, however, it is observed that River Ware, Oasis and WEB.BM can provide MTO solutions, with the WEB.BM being the one that takes hydrologic channel routing and the related channel storage change constraints into account. MTO solution mode does not require reservoir rule curve to be used as input. In fact, it generates optimal reservoir operating rule for each reservoir in each simulated year. It is also capable of fining the best level of demand hedging if the equal deficit constraint is used. In other words, the model can simultaneously optimize the use of storage and the management of water demands.

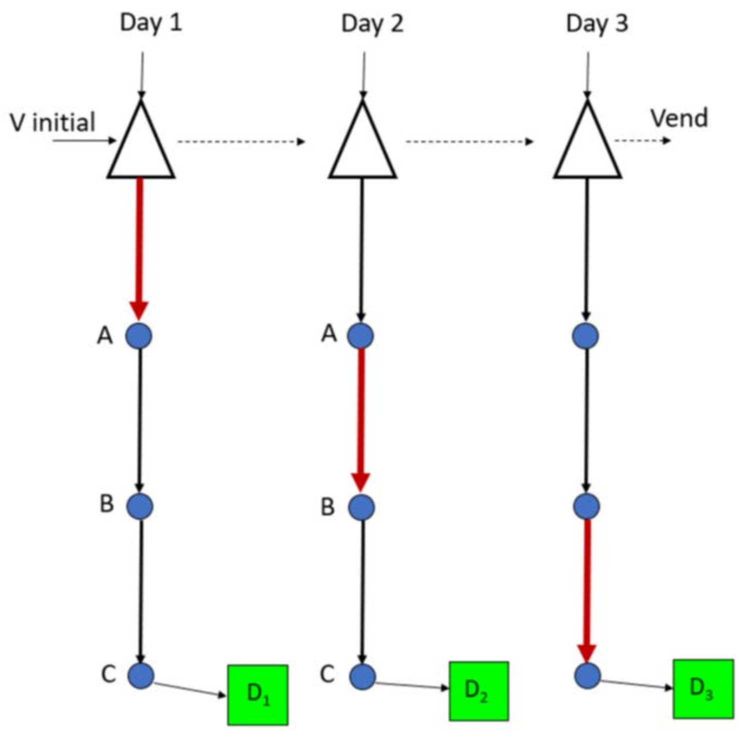


Figure 15.6: MTO Solution Concept

15.6.1.3 Combined STO/MTO solution mode:

This solution model is designed for real time operation, or for planning scenarios that would mimic the real time operation. In it, the solution is derived using the MTO solution mode for only a few time steps for which the inflow forecasts are available, and only the solution for the first day is adopted as final. For example, if the travel time in the basin is 5 days, and there is a 5-day runoff forecast obtained from one of the available forecasting tools, the model will be called every day to solve a small MTO problem over a 5-day time horizon. Only the solution for the first of the five days will be adopted to decision making in each day. This scheme is presented in Figure 15.7.

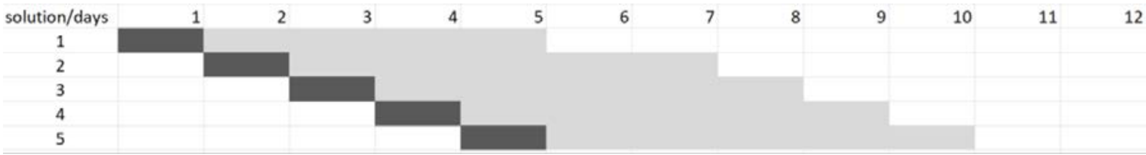


Figure 15.7: Combined STO/MTO Solution Concept

The accepted solutions for real time operation on a daily basis are shown with black colour in Figure 15.7.

In general, the use of models should proceed as follows:

1. Develop all required input data and the appropriate modelling schematic
2. Conduct model verification scenario, which should reproduce the historic water levels by enforcing the same reservoir outflows as those that are on the historical record and by using the local runoff estimates developed previously in step 1.
3. Conduct MTO scenario with perfect foreknowledge of runoff for the entire year and with the use of equal deficit constraint. Conduct analyses of model solutions to determine the best reservoir operating rules and deficit sharing policies.
4. Run an STO/MTO scenario with shorter time steps and with the new operating rules based on the analyses of the results in step 3.

The ultimate use of river basin models is linked to automated reservoir operation in real time. This requires the use of a runoff forecasting model in the monsoon season as an accompanying tool in addition to the use of the river basin management model, and it also requires calibrated hydrological channel routing for all river reaches included in the model.

15.7 Optimization and Simulation Based Models

Model vendors seem to promote the models they developed, and the same can be concluded regarding the model users: those who are familiar with certain models are prepared to favour using them rather than switching to other models. A river basin management model mimics decision making, i.e. it has a built-in algorithm that determines water allocation and reservoir releases. In general, there are two types of solution algorithms that determine reservoir releases:

1. Simulation models that rely on the use of various “what if” rules; and,
2. Optimization models

Since reservoir operation is an interesting research topic, the literature is full of various papers, each presenting their own solutions. Numerous heuristic solvers are being tested (although on relatively small problems with simplified constraints), and there is a lot of academic work on multi-objective and stochastic optimization procedures, both of which are unlikely to engender serious interest by the river basin management authorities. For example, multi-objective optimization tends to generate a large family of solutions that are considered to be "pareto optimal", thus leaving the task of selecting the "best one". To properly evaluate various models requires comparison of the solutions to a common set of test problems.

There are a number of studies that provide solid evidence related to the use of optimization models for managing multi-purpose multi-reservoir river basins. Optimization models require specification of the weight factor Pi allocated to each type of water use. The relative difference between the weight factors for different components represents the priority of allocation. There are two ways to specify the objective function within an optimization model, and with the same inflows and the same weight factors, either specification would result in the same flow allocation:

$$Max \sum_{i=1}^{n} Q_i P_i \leftrightarrow Min \sum_{i=1}^{n} ((I_i - Q_i)P_i \qquad \text{Equation 15.5}$$

The second formulation that uses the functional minimization form is known as "goal programming" since the goals is to minimize the deficits between the targets that are given as the upper bound on flow in each model component.

A number of optimization models use a simplified Network Flow Algorithm (NFA) that cannot handle constraints that include dynamic flow limits, such as for example the limits on reservoir outflows. Among else, these include the Modsim, Aquatool, Realm and the WEAP model. All of these models can handle only the STO solution mode.

The only known models so far that can handle MTO solution mode within the optimization framework are the RiverWare, Oasis and the WEB.BM models. The numerical example at the end of this Section demonstrates the use of MTO solution technique over a 10-day period on a river basin in Canada, where the proper time of travel vs flow tables were available to enable proper routing.

15.8 The Need for River Basin Management Models

River basin management models hold out a promise of automated computerized operation of river infrastructure. This will require combining several modules into a Decision Support System, as well as meticulous study to first create a river basin plan, since real time operation should rely on following:

1. River basin plan that includes optimized reservoir zoning and deficit sharing policies based on analysing the results of long-term basin operation under known historic (on stochastic) inflows and the current or future level of water demands.
2. A short to medium term runoff forecasting module that will provide runoff estimates over a realistic planning horizon with acceptable reliability
3. SCADA or other real time data monitoring system that provides real time information from the storage sites (reservoir levels) and from hydrometric and meteorologic stations.
4. The use of reservoir optimization module that can apply multiple time step optimization (MTO) solution method by using the recommendations from the river basin plan developed in step 1 and the results of runoff forecasts developed in step 2.

This kind of DSS may significantly improve the water use efficiency and increase the benefits of the current and future river infrastructure in a similar way the driverless cars and modern navigation systems will revolutionize the transportation industry over the next few decades.

Example 15-1

Details regarding this numerical example including all necessary input data and instructions how to build the project are given in the final section of the WEB.BM's User Manual. For brevity, the problem is briefly explained here and the results are discussed to demonstrate the proper working of the model. The modeling schematic is shown in Figure 15.8. The simulation is based on a daily time step, it starts on April 30th 2013 and runs continuously in MTO mode over 11 days. The basin is regulated by the Dickson Dam which forms Gleniffer Lake at its upstream end, and it contains one tributary inflow at node 5 as shown in Figure 15.8. Although the system appears to be very simple, the requested solution in this example may not be so easy to find using mere simulation models without resorting to numerous repetitive iterative runs, which are internally automated within the LP modeling framework. The operating priorities in this modeling example are listed below:

1. Maintain minimum flow of 16 m^3/s in Channels RD 20, RD 11 and RD 60 at all times;
2. Provide water demands to RD 100 and node 8 (as given with the remaining input data); and,
3. Keep the maximum flow through the City of Red Deer at 950 m^3/s (Channel RD 11 and entry into the channel RD 40) at all times as much as possible during floods.

The example simulation run relies on the following input data:

1. Storage capacity table of Glennifer Lake created by the Dickson Dam;
2. Outflow vs elevation table for combined Dickson Dam bottom outlet and spillway;

3. Reconstructed historic 2013 daily inflows into Glennifer Lake;
4. Historic flows at the Little Red Deer River tributary (inflow at node 5 in the Schematic);
5. Travel time vs flow tables for river reaches RD 30, RD 40 and RD 50 (travel time for reach 20 is ignored due to its short length); and,
6. Key reservoir elevations, including the maximum operational water level (949.5 m), normal water level (944.0 m), and the top of dead storage (926.0 m).

For simplicity, net evaporation has been ignored in this example. The input data for this example is provided in the WEB.BM User's Manual. The problem is presented graphically in the schematic in Figure 10 for only three consecutive time steps with the same reservoir connected in time via carry-over storage arcs. The storage release movement is shown as a thick line in the schematic, depicting the movement of the release from one river reach to another, assuming that the length of channels was chosen such that they correspond roughly to one day of travel time for average flow conditions, while the calculation time step is also conducted on a daily basis.

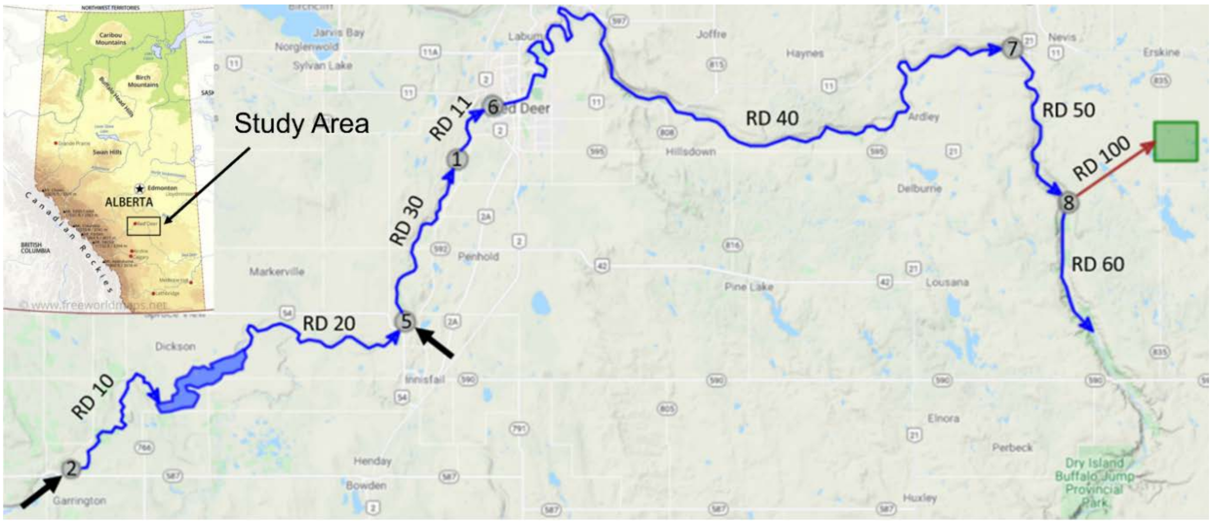


Figure 15.8: Red Deer River Modelling Schematic

The use of optimization in this framework should determine the amount and the timing of the release from the reservoir such that the water demand of 100 m^3/s is met at the downstream location on May 6, along with the minimum flow requirement of 16 m^3/s in the most downstream river reach RD 60 of the system which should be met for all simulated days. If a simple time lagging method was employed, the model would release the 100 m^3/s on May 4 such that this amount could be later consumed on May 6. However, this does not take into account the requirement to re-fill channel storage so as to enable the flow of 116 m^3/s to be available at the end of Reach 50 on May 6th. With the starting reservoir elevation of 942.5 m, WEB.BM derived the solution that is displayed in Table 15.1 (reservoir levels are given in m while all other components are given in m^3/s):

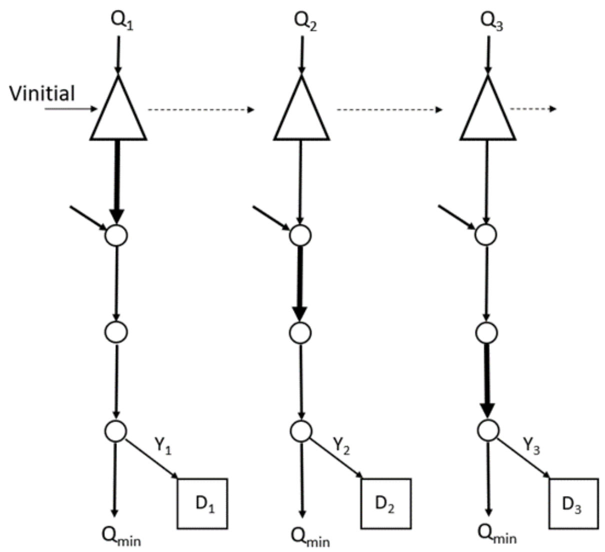


Figure 15.9: Example of Channel Routing with MTO Solution Framework

Table 15-1: WEB.BM Model Solution – Reservoir Units m, all others m³/s

Compone	No	30-	01-	02-	03-	04-	05-	06-	07-	08-	09-	10-
Reservoir	2	942.6	942.5	943.0	943.5	941.8	941.8	941.9	942.0	942.2	942.3	942.4
River	10	32.16	24.40	92.65	83.31	53.69	26.15	28.74	29.39	41.45	36.74	34.85
River	20	16.00	34.16	16.00	16.00	300.1	23.51	16.00	16.00	16.00	16.00	16.00
River	30	26.10	41.86	22.31	22.31	306.5	30.09	22.31	21.97	21.81	21.42	20.89
River	11	28.95	31.86	31.99	26.82	135.8	161.9	67.36	40.10	29.81	25.49	23.29
River	40	28.95	31.86	31.99	26.82	135.8	161.9	67.36	40.10	29.81	25.49	23.29
River	50	31.35	30.80	31.45	30.29	65.96	131.5	118.3	72.52	48.71	36.27	29.58
Diversion	10	15.00	15.00	15.00	15.00	15.00	15.00	100.0	15.00	15.00	15.00	15.00
River		16.45	16.23	16.17	16.00	26.46	68.71	16.00	85.05	57.11	37.86	25.78

It should be noted that the above table shows the flows at the upstream end of each river reach. Hence, the flow in channel 40 is shown as 118.3 m^3/s on May 6, implying that 2.3 m^3/s will end up in channels storage on May 6, while the amount at the downstream end of this channel will be 116 m^3/s, which is distributed between water demand D_3 (100 m^3/s) and river reach 60 (16 m^3/s). Note that there is no additional spill above 16 m^3/s in Channel 60 on May 6, which means that the model has determined the right amount of storage release in days 4 and 5 to ensure that the available flow at the top of channel 60 is equal to the sum of the required diversion and the minimum flow in Channel 60. The initial storage release of 300.111 m^3/s is increased by the amount of local runoff into node 5 to give the total inflow into Channel 30 of 306.571 m^3/s, and after two days of routing transformations this flow will provide 116 m^3/s at entry of Channel 60. The routing process through a sequence of channels is shown in Figure 11. It should be noted that the actual entry into channel 50 is a sum of two flows, diversion flows at channel 100 and the river flow at channel 60, which are summed up in Figure 11 for the purpose of analyzing the routing sequence. If it is desirable to do this in the model, a channel with no routing can be inserted before node 8 to show the result of routed flows from channel 60.

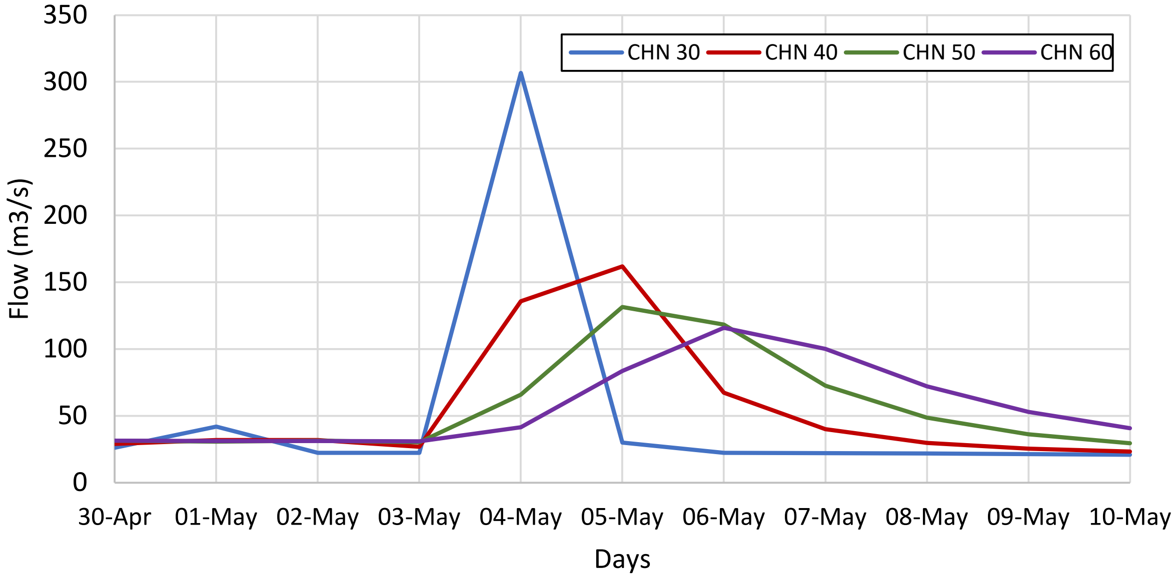


Figure 15.10: Inflows at Nodes 5, 6, 7, and 8 (Entry of Channels 30, 40, 50 and 60)

Based on the model setup, the incoming flow in channel 60 would have the diversion flows at channel 100 subtracted first. An additional numerical example which demonstrates optimal management of storage releases during a severe historic flood in June 2013 is provided in the WEB.BM User Manual accessible by signing in to the web site where the model resides: www.riverbasinmanagement.com.

The only requirement to use the WEB.BM software is an email account and a user defined password. The terms of use are available at www.riverbasinmanagement.com which contains links to the instructional videos and the User's Manual. To help verify the claims made in this paper, the backup copies of the two projects that contain numerical example in this paper and two more examples in the manual can be downloaded from the data link provided in the Data Availability section. The projects that can be downloaded and restored include:

a) Bargi Reservoir Test.zip, which includes an example of simultaneous optimization of reservoir operation and demand hedging, as previously detailed in Ilich at al (2020)
   Downloadable link:
   https://data.world/nilich/webbm-repository/workspace/file?filename=Bargi+Reservoir+Test.zip
   and,

b) Red Deer River two hydrologic routing tests.zip which include two numerical examples with daily time steps, one related to a sudden increase in reservoir release during a low flow period explained in this paper, and the other one showing alternative management of the 2013 flood as explained in the User's Manual Downloadable link:
https://data.world/nilich/webbm-repository/workspace/file?filename=Red+Deer+River+two+hydrologic+routing+tests.zip

Glossary of Terms

Alluvium - Sediments deposited by erosional processes, usually by streams.

Annual Flood - The maximum instantaneous peak discharge observed in a year. The maximum discharge peak during any hydrologic year (1st June to 31st May).

Annual Flood Series - A list of the maximum flood peak discharges occurring in each year for the period of record.

Annual Runoff/ Water Yield / Annual Yield - The total natural discharge of a stream for a year, usually expressed in millimetres or centimetres depth or hectare metre or millions of cubic metres.

Areal Rainfall - The average rainfall over an area, usually derived from or discussed in contrast with point rainfall.

Artificial Control - A weir or other man-made structure which serves as the control for a stream-gaging station.

Attenuation - The process where the flood crest is reduced as it progresses downstream.

Backflow - The backing up of water through a conduit or channel in the direction opposite to normal flow.

Backwater Curve - The longitudinal profile of the surface of a liquid in a non-uniform flow in an open channel, when the water surface is not parallel to the invert owing to the depth of water having been increased by the interposition of an obstruction such as a dam or weir. The term is sometimes used in a generic sense to denote all water surface profiles; or for profiles where the water is flowing at depths greater than the critical.

Backwater Effect - The effect which a dam or other obstruction or construction has in raising the surface of the water upstream from it.

Backwater Flooding - Upstream flooding caused by downstream conditions such as channel restriction and/ or high flow in a downstream confluence stream.

Bank - The margins of a channel. Banks are called right or left as viewed facing in the direction of the flow.

Bank Storage - Water absorbed and stored in the void in the soil cover in the bed and banks of a stream, lake, or reservoir, and returned in whole or in part as the level of water body surface falls.

Bank full Stage/Elevation - The stage below which all discharge is confined to the main channel and above which part of the flow occurs in overbank areas of the floodplain. An established river stage at a certain point along a river which is intended to represent the maximum safe water level which will not overflow the river banks or cause any significant damage within the reach of the river.

Barrage - Any artificial obstruction placed in water to increase water level or divert it. Usually, the idea is to control peak flow for later release.

Base Flow - Streamflow which results from precipitation that infiltrates into the soil and eventually moves through the soil to the stream channel. This is also referred to as ground water flow, or dry-weather flow. The sustained or fair-weather discharge that persists after storm runoff and associated quick return flow are depleted. It is usually derived from groundwater discharge or gradual snow or ice melt over extended periods of time, but need not be continuous flow. It can be based on annual or seasonal periods depending upon when major floods usually occur. It may also be defined as the stream discharge derived from groundwater sources.

Basin - An area having a common outlet for its surface runoff. A basin will comprise multiple catchments.

Basin Boundary - The topographic dividing line around the perimeter of a basin, beyond which overland flow (i.e.; runoff) drains away into another basin.

Basin Lag - The time it takes from the centroid of rainfall for the hydrograph to peak. The measure of the time between the centre of mass of precipitation to the centre of mass of runoff (on the hydrograph); basin lag is a function of not only basin characteristics, but also of storm intensity and movement. Some hydrologic texts define lag from the centre of mass of rainfall to the hydrograph peak.

Bed Load - Sand, silt, gravel, or soil and rock detritus carried by a stream on or immediately above its bed. The particles of this material have a density or grain size such as to preclude movement far above or for a long distance out of contact with the stream bed under natural conditions of flow.

Benchmark (BM) - A permanent point whose known elevation is tied to a national network. These points are created to serve as a point of reference. Benchmarks have generally been established by the Survey of India, but may have been established by other agencies.

Braided Stream - Characterized by successive division and rejoining of streamflow with accompanying islands. A braided stream is composed of anabranches.

Calibration - The process of using historical data to estimate parameters in a hydrologic model that may include forecast, routings, and unit hydrographs.

Catchment Area - An area having a common outlet for its surface runoff.

Channel (watercourse) - An open conduit either naturally or artificially created which periodically, or continuously contains moving water, or forms a connecting link between two bodies of water. River, creek, run, branch, anabranch, and tributary are some of the terms used to describe natural channels. Natural channels may be single or braided. Canal and floodway are some of the terms used to describe artificial channels.

Channel Inflow - Water, which at any instant, is flowing into the channel system form surface flow, subsurface flow, base flow, and rainfall that has directly fallen onto the channel.

Channel Routing - The process of determining progressively timing and shape of the flood wave at successive points along a river. The outlet of each sub-catchment is located far upstream of the outlet of the main catchment. The outflow from a sub-catchment will have to pass through the channels before reaching the catchment outlet. The hydrograph entering a channel from a sub-catchment will get modified by the temporary storage of channel. So, it is necessary to estimate the outflow hydrograph of the channel to find the flow at the main catchment outlet by a process is known as channel routing.

Complex Rating - Discharge rating that relates discharge to stage plus some other independent variable such as rate of change in stage or fall in a reach between two gauge stations.

Control - Closest section or reach of a channel downstream from a gage, usually a natural constriction or artificial weir, where the channel is shallower, narrower, or rougher than it is elsewhere and where the water-surface slope is significantly steeper.

Conveyance Loss - The loss of water from a conduit due to leakage, seepage, evaporation, or evapo-transpiration.

Correlation - A statistical index that measures linear variation/ relationship between variables.

Crest - The highest stage or level of a flood wave as it passes a point.

Critical Depth - The depth of water flowing in an open channel or conduit, partially filled, corresponding to one of the recognized critical velocities.

Critical Flow - A condition of flow where the mean velocity is at one of the critical values; ordinarily at the Reynolds' critical velocities which define the point at which the flow changes from streamline or nonturbulent to turbulent flow.

Cross Section - The shape of a channel, stream, or valley determined by a line approximately perpendicular to the main path of water flow, along which measurements of distance and elevations are determined.

Cross-sectional area - Area perpendicular to the direction of flow.

Cubic Metres per Second - A volumetric unit of water flow. Abbreviated m^3/s.

Current Meter - Device used to measure the water velocity or current in a river. A current meter of the vertical axis type (most commonly used in India) has a series of conical cups fastened to a flat framework through which a pin extends. The pin sets in the framework of the meter, and the cups are rotated around it in a horizontal plane by the flowing water, registering the number of revolutions by acoustical or electrical devices, from which the velocity of the water may be computed. In case of current meter with horizontal axis, a helical screw or impeller rotates in correspondence with the stream velocity.

Daily Flood Peak - The maximum mean daily discharge occurring in a stream during a given flood event.

Dam - Any artificial barrier which impounds water.

Datum - A reference "zero" elevation for a stream or river gage. This "zero" is generally referenced to the mean sea level.

Degradation - The geologic process by means of which various parts of the surface of the earth are worn down and carried away and their general level lowered, by the action of wind and water.

Degrees of Freedom - The number of independent pieces of information, or parameters, required to form a statistical estimate.

Delta - An alluvial deposit, often in the shape of the Greek letter "delta", which is formed where a stream drops its debris load on entering a body of quieter water.

Depletion Curve - That part of the hydrograph extending from the point of termination of the Recession Curve to the subsequent rise or alternation of inflow due to additional water becoming available for stream flow.

Depression Storage - The volume of water contained in natural depressions in the land surface, such as puddles.

Depth of Runoff - The total runoff from a drainage basin, divided by its area. For convenience in comparing runoff with precipitation, the term is usually expressed in centimetres or millimetres of depth during a given period of time over the drainage area expressed in square kilometres.

Detention Basins - Detention basins are normally dry, but are designed to detain surface water temporarily during, and immediately after a runoff event. Their primary function is to attenuate the storm flows by releasing flows at a lower flow rate. There are no gates or valves allowed on the outlet so that water can never be stored for long durations.

Detention Storage - The volume of water, other than depression storage, existing on the land surface as flowing water which has not yet reached the channel.

Direct Runoff - Water that enters the stream channel during a storm. It mainly consists of rainfall on the stream surface, surface runoff, and quick return flow. The runoff entering stream channels promptly after rainfall or snow melt. Superposed on base runoff, it forms the bulk of the hydrograph of a flood.

Discharge - The rate at which water passes a given point. Discharge is expressed in a volume per time with units of cubic metres per second. Discharge is often used interchangeably with streamflow.

Discharge Curve - A curve that expresses the relation between the discharge of a stream or open conduit at a given location and the stage or elevation of the liquid surface at or near that location. This is also known as Rating Curve and Discharge Rating Curve.

Discharge Table - A table showing the relation between the gage height and the discharge of a stream or conduit at a given gaging station. Also known as a Rating Table.

Diversion - The taking of water from a stream or other body of water into a canal, pipe, or other conduit.

Divide or Drainage Divide - The boundary line, along a topographic ridge or along a subsurface formation, separating two adjacent drainage basins. The high ground that forms the boundary of a watershed. A divide is also called a ridge.

Drainage area - The area of a watershed draining into a stream which passes through a specified outlet point on it (also known as Watershed and Catchment Area). The area may be of different sizes for surface runoff, subsurface runoff or flow, and base flow. Generally, the surface runoff area is used as the drainage area.

Drainage Basin - A part of the surface of the earth that is occupied by a drainage system, which consists of a surface stream or a body of impounded surface water together with all tributary surface streams and bodies of impounded surface water.

Drawdown - The lowering of the surface elevation of a body of water, resulting from the withdrawal of water therefrom.

Dredging - The scooping, or suction of underwater material from a harbor, or waterway. Dredging is one form of channel modification. It is often too expensive to be practical because the dredged material must be disposed of somewhere and the stream will usually fill back up with sediment in a few years. Dredging is usually undertaken only on large rivers to maintain a navigation channel.

Drought - A period of abnormally dry weather sufficiently prolonged from the lack of precipitation to cause a serious hydrologic imbalance.

Dry Weather Flow - Streamflow which results from precipitation that infiltrates into the soil and eventually moves through the soil to the stream channel. This is also referred to as base flow, or ground water flow.

Effective Precipitation (Rainfall) - That part of the precipitation that produces runoff. Precipitation that is "effective" in correlating with runoff.

Effluent Stream - Any watercourse in which all, or a portion of the water volume came from the Phreatic zone, or zone of saturation by way of groundwater flow, or baseflow.

Embankment - Fill material, usually earth or rock, placed with sloping sides and usually with length greater than height. Restrains water from overflowing into the countryside.

Energy Dissipator - A structure which slows fast-moving spillway flows in order to prevent erosion of the stream channel.

Energy grade line - A graphical representation of the kinetic head of water flowing in a pipe, conduit, or channel. The line is plotted above the hydraulic grade line at a distance equal to the velocity head. Abbreviated EGL.

Ephemeral stream – A stream that flows in response to runoff producing precipitation events and thus discontinuing its flow during dry seasons. Such flow is usually of short duration.

Erosion - Wearing away of the lands by running water, glaciers, winds, and waves, can be subdivided into three processes: Corrasion, Corrosion, and Transportation. Weathering, although sometimes included here, is a distant process which does not imply removal of any material.

Evaporation - Process by which liquid water is converted into water vapor.

Evapotranspiration - Combination of evaporation from free water surfaces and transpiration of water from plant surfaces to the atmosphere.

Exceedance Probability - The probability that a random event will exceed a specified magnitude in a given time period, usually one year.

Excess Rain - Effective rainfall in excess of infiltration capacity.

Fall - Difference between the water-surface elevations of two locations on a stream, usually base and auxiliary gage sites for a slope station.

Flash Flood - A flood which follows within a few hours (usually less than 6 hours) of cloudburst or heavy / excessive rainfall, dam failure, or the sudden release of water impounded by ice.

Flood - A relatively high flow as determined by either gage height or discharge quantity. An event during which a stream overflows its normal banks.

Flood Crest - The maximum height of a flood wave as it passes a location.

Flood Frequency Curve - A graph showing the number of times per year on the average, plotted as abscissa, that floods of magnitude, indicated by the ordinate, are equalled or exceeded. A similar graph but with exceedance probability of floods plotted as abscissa.

Flood Profile - A graph of elevation of the water surface of a river in flood, plotted as ordinate, against distance, measured in the downstream direction, plotted as abscissa. A flood profile may be drawn to show elevation at a given time, crests during a particular flood, or to show stages of concordant flows.

Flood Routing - Process of determining progressively the timing, shape, and amplitude of a flood wave as it moves downstream to successive points along the river or through a reservoir.

Flood Stage - A gage height at which a watercourse overtops its banks and begins to cause damage to any portion of the defined reach. Flood stage is usually higher than or equal to bank full stage.

Flood Wave - The rise and fall in streamflow during and after a storm. A rise in streamflow to a crest and its subsequent recession caused by precipitation, snow melt, dam failure, or reservoir releases.

Flood way - A part of the flood plain, otherwise leveed, reserved for emergency diversion of water during floods. A part of the flood plain which, to facilitate the passage of floodwater, is kept clear of encumbrances. The channel of a river or stream and those parts of the flood plains adjoining the channel, which are reasonably required to carry and discharge the floodwater orfloodflow of any river or stream.

Floodplain - A strip of relatively flat and normally dry land alongside a stream, river, or lake that is covered by water during a flood. The low-lying areas adjacent to a stream that are occasionally, are predicted to be, or have been covered by water when the stream overflows its banks.

Floodwall - A long, narrow concrete, or masonry embankment usually built to protect land from flooding. If built of earth the structure is usually referred to as an embankment or a levee. Floodwalls and levees confine streamflow within a specified area to prevent flooding. Ring bunds confine streamflow out of an area.

Flow – the rate of water discharges from a source expressed as a volume per unit time. Synonymous with discharge.

Flow Duration - The percentage of time during which specified flow rates are exceeded.

Flow Duration Curve - A cumulative frequency curve that shows the percent of time during which specified units of items (e.g., discharge, head, power, etc.) were equalled or exceeded in a given period. It is the integral of the frequency diagram.

Fluvial - referring to processes occurring in a river.

Freeboard - The vertical distance between the normal maximum level of the water surface in a channel, reservoir, tank, canal, etc., and the top of the sides of a levee, dam, etc., which is provided so that waves and other movements of the liquid will not overtop the confining structure.

Frequency - The number of occasions that the same numerical measure of a particular quantity has occurred between definite time periods. Often stated in terms such as return interval, recurrence interval, or percent chance.

Frequency Analysis - An analysis of the frequency at which a given event occurs or repeats over a particular time period or in a given sample.

Frequency Curve - A curve that expresses the relation between the frequency distribution plot, with the magnitude of the variables as abscissas and the number of occurrences of each magnitude in a given period as ordinates. The theoretical frequency curve is a derivative of the probability curve.

Frequency Distribution - A generalized cumulative density function of known shape and range of values.

Friction Head - The decrease in total head caused by friction.

Friction Slope - The friction head loss per unit length along an open channel or a conduit.

Gage Zero - The elevation of zero stage.

Gaging Station - A particular site on a stream, lake, reservoir or other body of water where systematic observations of stage and/or flow are made.

Gate - A device in which a leaf or member is moved across the waterway from an external position to control or stop flow. There are many different kinds of gates, automatic or manually operated.

Gauge - A device for indicating the elevation of a water surface.

Gauge Datum - The arbitrary zero datum elevation which all stage measurements are made from.

Gauge Height - The water-surface elevation referred to some arbitrary datum of the gauge. Gauge height is often used interchangeably with the more general term stage, although gage height is more appropriate when used with a reading on a gauge.

Gauge Height of Zero Flow - Gauge reading corresponding to zero or extremely small discharge at a gaging station; the gauge height of zero flow is often used interchangeably with the "point of zero flow," which is more appropriately used for a physical location in the streambed near the gauge.

Gauging - The operation, including both field and office work, of measuring the discharge of a stream of water in a waterway.

Ground Water - Water within the earth that supplies wells and springs; water in the zone of saturation where all openings in rocks and soil are filled, the upper surface of which forms the water table. Also known as Phreatic water.

Ground Water Flow - Streamflow which results from precipitation that infiltrates into the soil and eventually moves through the soil to the stream channel. This is also referred to as baseflow, or dry-weather flow.

Ground Water Runoff - That part of the runoff which has passed into the ground, has become ground water, and has been discharged into a stream channel as spring, or seepage water.

Gumbel Distribution - Gumbel distribution is a member of family of Extreme Value distributions. It is a two-parameter distribution and is widely used in hydrology for flood frequency analysis.

Head Loss - The decrease in total head caused by friction, entrance and exit losses.

Headwaters - Streams at the source of a river.

Historical Series - A systematic record or series of all events, including both measured and non-measured events, in a given period of years, with the date of each event being known.

Hydraulic Grade Line - A line or an elevation representing the hydraulic head in a closed conduit or open channel. In an open channel, the hydraulic grade line is the water surface. Abbreviated HGL.

Hydraulic Head - The height of the free surface of a body of water above a given point beneath the surface. The height of the water level at the headworks, or an upstream point, of a waterway, and the water surface at a given point downstream. The height of a hydraulic grade line above the centre line of a pressure pipe, at a given point.

Hydraulic Mean Depth/ Hydraulic Radius - The right cross-sectional area of a stream of water divided by the length of that part of its periphery in contact with its containing conduit; the ratio of area to wetted perimeter.

Hydrograph - A chronological depiction of water level or discharge, which includes base flow, or one which corresponds to a net rain storm of duration longer than one unit period.

Hydrograph Separation - The process where the storm hydrograph is separated into baseflow components and surface runoff components.

Hydrographic Survey - An instrumental survey to measure and determine characteristics of streams and other bodies of water within an area inclusive of areas under water, including such things as location, areal extent, and depth of water in lakes or the ocean; the width, depth, and course of streams; position and elevation of high-water marks; etc.

Hydrologic Budget - An accounting of the inflow to, outflow from, and storage in, a hydrologic unit, such as a drainage basin, aquifer, soil zone, lake, reservoir, or irrigation project.

Hydrologic Cycle - The natural pathway water follows as it changes between liquid, solid, and gaseous states.

Hydrologic Model - A conceptual or physically-based procedure for numerically simulating a process or processes which occur in a watershed.

Hydrologic Unit - A geographical area representing part or all of a surface drainage basin or distinct hydrologic feature such as a reservoir, lake, etc.

Hydrology - The applied science concerned with the waters of the earth, their occurrences, distribution, and circulation through the unending hydrologic cycle of: precipitation, consequent runoff, infiltration, and storage; eventual evaporation; and so forth. It is concerned with the physical and chemical reaction of water with the rest of the earth, and its relation to the life of the earth.

Hydrostatic Head - A measure of pressure at a given point in a liquid in terms of the vertical height of a column of the same liquid which would produce the same pressure.

Hyetograph - A graphical representation of rainfall intensity with respect to time.

Impermeable/ Impervious - Material that does not permit fluids to pass through it or does not let water to infiltrate.

Infiltration - Movement of water through the soil surface into the soil.

Infiltration Capacity - The maximum rate at which water can enter the soil at a particular point under a given set of conditions.

Infiltration Index - An average rate of infiltration, in centimetres or millimetres per hour, equal to the average rate of rainfall such as that the volume of rainfall at greater rates equals the total direct runoff.

Infiltration Rate - The rate at which infiltration takes place expressed in depth of water per unit time, usually in millimetres per hour.

Influent Stream - Any watercourse in which all, or a portion of the surface water flows back into the ground namely the, vadose zone, or zone of aeration.

Initial Detention - The volume of water on the ground, either in depressions or in transit, at the time active runoff begins.

Initial Loss - Rainfall preceding the beginning of surface runoff. It includes interception, surface wetting, evaporation and infiltration unless otherwise specified.

Inter-basin Transfer - The physical transfer of water from one basin/ catchment to another.

Interception - The process by which precipitation is caught and held by foliage, twigs, and branches of trees, shrubs, and other vegetation, and lost by evaporation, never reaching the surface of the ground. Interception equals the precipitation on the vegetation minus streamflow and through fall.

Interception Storage - Water caught by plants at the onset of a rainstorm. This must be full before rainfall reaches the ground.

Interflow - Water that infiltrates into the soil profile and moves laterally until it returns to the surface or stream. The lateral motion of water through the upper layers until it enters a stream channel. This usually takes longer to reach stream channels than runoff. This also called subsurface storm flow.

Intermediate Zone - The subsurface water zone below the root zone and above the capillary fringe.

Intermittent Stream - A stream that flows periodically.

Isohyet - A line that connects points of equal rainfall.

Lag (Time) - The time it takes a flood wave to move downstream.

Laminar Flow - Streamline flow in which successive flow particles follow similar path lines and head loss varies with velocity to the first power.

Land Use - A land classification, such as row crops or pasture, that indicates a type of land use. Roads may also be classified as a separate land use.

Length - The distance in the direction of flow between two specific points along a river, stream, or channel.

Log – Normal Distribution - If the logarithms of a variable are normally distributed, then the variable is said to be log normally distributed. Details are available in any statistical text.

Log Paper/ Log-Log Paper - A graph paper that has logarithmic scales on both horizontal and vertical axes. The scales may be any number of cycles, but usually in combinations such as 1x1, 2x2, 3x3, 3x5, 4x7, etc.

Log-Normal Paper - Graph paper used in estimating frequencies of floods, etc. The paper has a logarithmic scale for the flood (or other event) amounts and a cumulative distribution scale (also called frequency or percent chance scale) for the probability plotting positions.

Loss - The portion of precipitation lost as runoff from the surface of the land due to evaporation and/or deep percolation.

Manning's n - A coefficient of roughness, used in a formula for estimating the capacity of a channel or a pipe to convey water. Generally, n values are determined by inspection in the field.

MAP (Mean Areal Precipitation) - The average rainfall over a given area, generally expressed as an average depth over the area in centimetres or millimetres.

Mass Curve - A graph of the cumulative values of a hydrologic quantity (such as runoff or rainfall), generally as ordinate, plotted against time or date.

Mean - The average of a series of numbers. It can be arithmetic or geometric, depending on the equation used to compute the mean.

Mean Daily Flow - The average or mean discharge of a stream for one day. Usually given in cubic feet per second.

Mean Depth - The average depth of water in a stream channel or conduit. It is equal to the cross-sectional area divided by the surface width.

Meander - The winding of a stream channel.

Meander Belt - The area between lines drawn tangential to the extreme limits of fully developed meanders.

Median - The value in an array of numbers that has as many lower values as it has higher values.

Movable Bed - A stream bed made up of materials readily transportable by the stream flow.

Moveable Bed Streams - Streams where steep slopes and lack of vegetation result in a lot of erosion. During a flood, a channel may be eroded more deeply, or it may become filled with sediment and move to a different location.

Muskingum - A flood routing technique that applies to channel or reach routing.

Natural Control - A stream gaging control which is natural to the stream channel, in contrast to an artificial control structure by man.

Net Rainfall - The portion of rainfall which reaches a stream channel or the outlet point as direct surface flow.

Normal - A mean or average value established from a series of hydrological or meteorological observations.

Normal Distribution - The Normal distribution is one of the most important distributions. It is also the distribution most commonly used. This is used to fit empirical distributions with skewness coefficient close to zero. Details are available in any statistical text.

Normal Year - A year during which the precipitation or stream flow approximates the average for a long period of record.

Orifice - An opening with closed perimeter, usually sharp edged, and of regular form in a plate, wall, or partition through which water may flow, generally used for the purpose of measurement or control of water. The end of a small tube, such as a Pitot tube, piezometer, bubbler type water level recorder etc.

Outflow Channel - A natural stream channel which transports reservoir releases. Outlet- An opening through which water can be freely discharged from a reservoir.

Outliers - Extreme event represented by data points which depart from the trend of the rest of the data.

Overland Flow - The flow of rainwater or snowmelt over the land surface toward stream channels. After it enters a watercourse, it becomes runoff.

Peak Discharge - Highest rate of discharge of a volume of water passing a given location during a given period of time (during the year, or a flood event, etc.).

Pearson Type III Distribution - Pearson type III is a three-parameter distribution, also known as Gamma distribution with three parameters. Details are available in any statistical text.

Percolation - The movement of groundwater in streamline flow in any direction through small interconnected and saturated interstices of rock or earth.

Percolation - The movement of water, under hydrostatic pressure, through the interstices of a rock or soil, except the movement through large openings.

Percolation Loss - Water that percolates downward through the soil beyond the reach of plant roots.

Perennial Stream - A stream that flows all year round. Compare intermittent stream.

Permanent Control - A stream gaging control which is substantially unchanging and is not appreciably affected by scour, fill, or backwater. Natural or artificial control, the location and dimensions of which remain unchanged for very long periods.

Phreatic surface - The free surface of ground water at atmospheric pressure.

Phreatic Zone - The locus of points below the water table where soil pores are filled with water. This is also called the zone of saturation.

Pluvial - Anything that is brought about directly by precipitation.

Point Discharge - Instantaneous rate of discharge, in contrast to the mean rate for an interval of time.

Point Precipitation - Precipitation at a particular site, in contrast to the mean precipitation over an area.

Population - The entire (usually infinite) number of data points from which a sample is taken or collected. The total number of past, present, and future floods at a location on a river is the population of floods for that location even if the floods are not measured or recorded.

Precipitation - Precipitation is the water received in a liquid or solid state, out of the atmosphere, generally onto a land or water surface. It is the common process by which atmospheric water becomes surface, or subsurface water. The term "precipitation" is also commonly used to designate the quantity of water that is precipitated. Precipitation includes rainfall, snow, hail, and sleet, and is therefore a more general term than rainfall.

Predictor Variable - The independent variable or variables in a regression equation or the variable used to predict the criterion variable or dependent variable.

Probability - The likelihood that a certain event will occur.

Probability Paper - Any graph paper prepared especially for plotting magnitudes of events versus their frequencies or probabilities.

Quick Return Flow - The diminishing discharge directly associated with a specific storm that occurs after surface runoff has reached its maximum. It includes base flow, prompt subsurface discharge (commonly called interflow), and delayed surface runoff. This flow reappears rapidly in comparison to base flow and is generally much in excess of normal base flow. It is common in humid climates and in watersheds with soils of high infiltration capacities and moderate to steep slopes. Abbreviated QRF.

Random Error - Errors that occur in any kind of measured data from time to time because of a variety of unrelated causes.

Rating Curve - A graph showing the relationship between the stage, usually plotted vertically (Y-axis) and the discharge, usually plotted horizontally (X-axis).

Rating Table - A table of stage values and the corresponding discharge for a river gaging site.

Reach - The distance between two specific points outlining that portion of the stream, or river for which the forecast applies. This generally applies to the distance above and below the forecast point for which the forecast is valid.

Recession Curve - The part of the descending limb on a hydrograph that extends from the point of inflection to the time when direct runoff has ceased.

Recurrence Interval - The average number of years within which a given event will be equalled or exceeded. A 50-year frequency flood has an average recurrence interval of 50 years, and so on. It is the inverse of percent chance. It is often referred to as return interval.

Regional Analysis - An analysis of parameters on gaged watersheds in a region that is used to estimate the same parameters for ungauged watersheds in the same region. It is often used in making flood frequency or other types of hydrologic analyses.

Regression - A method of developing a relationship between a criterion variable and one or more predictor variables, with the objective of predicting the criterion variable for given values of the predictor variable.

Reservoir-A Pond, lake, tank, basin, or other space, either natural in its origin or created in whole or in part by the building of engineering structures. A reservoir stores, regulates, and controls water. Usually, a man-made facility for the storage, regulation and controlled release of water.

Reservoir routing - Flood routing through a reservoir. The hydrograph of a flood entering a reservoir will change in shape as it emerges out from the reservoir. This is because of the volume of water stored in reservoir temporarily. The peak of the hydrograph will be reduced (attenuated), time to peak will be delayed (translated) and base of the hydrograph will be increased. The extent up to which an inflow hydrograph will be modified in the reservoir is computed by a process is known as reservoir routing.

Reservoir Surface Area - The surface area of a reservoir when filled to the normal pool or water level.

Residual - The difference between the value predicted with the regression equation and the criterion variable.

Return period – The average elapsed time between occurrences of an event with a specified magnitude or greater. For example, a 100-year discharge measured on a given river is equalled or exceeded, on average, once every 100 years. This does not mean that the 100-year discharge occurs once every 100 years, but that the average time between events of that magnitude or greater is 100 years. Stated another way, there is a 1% chance of a discharge equal to or greater than the 100-year flood event occurring in any given year. This does not rule out the possibility of two major floods occurring at a place on consequent years, even though the probability for the same will be low.

Riparian – Pertaining to the banks of a river, stream, or other typically, flowing body of water as well as to plant and animal communities along such bodies of water. This term is also commonly used for other bodies of water, e.g., ponds, lakes.

River Basin – The total drainage area of a river and its tributaries.

River Gauge - A device for measuring the river stage. The Gauge may be of manual or automatic type.

River Gauge Datum - The arbitrary zero datum elevation which all stage measurements are made from. This refers to the level above the mean sea level, so that water levels at upstream or downstream sites along a river can be compared.

Routing - The methods of predicting the attenuation of a flood wave as it moves down the course of a river.

Runoff - That part of precipitation that flows toward the streams on the surface of the ground or within the ground. Runoff is composed of baseflow and surface runoff.

Sediment – Soil particles, sand, and minerals dislodged from the land and deposited into aquatic systems as a result of erosion.

Seepage - Infiltration which reaches the water table.

Semi-log-paper/ Semi-logarithmic Graph Paper – A graph paper with an arithmetic scale along one axis and a logarithmic scale along the other. Either scale is used for the independent variable as the data require.

Sheet Flow - Flow that occurs overland in places where there are no defined channels, the flood water spreads out over a large area at a uniform depth. This also referred to as overland flow.

Shift Adjustment - Adjustment, usually varying with time and stage, applied to gage heights to compensate for a change in the rating shape or position.

Shifting-control Method – Systematic use of shift adjustments as a substitute for revised ratings.

Simple Rating – Discharge rating that relates discharge to stage only.

Skew - Skew is a shape parameter and the third moment about the mean, which measures the symmetry of a distribution.

Slope Rating - Complex rating that relates discharge to the observed gauge height at one gauge (base gauge) and to the fall in water-surface elevation between the base gauge and an auxiliary gauge at another site.

Soil Moisture - Water contained in the upper regions near the earth's surface.

Specific Yield - The ratio of the water which will drain freely from the material to the total volume of the aquifer formation. This value will always be less than the porosity.

Spillway - A structure over or through which excess or flood flows are discharged. If the flow is controlled by gates, it is a controlled spillway, if the elevation of the spillway crest is the only control, it is an uncontrolled spillway.

Spillway Crest - The elevation of the highest point of a spillway.

Spring - An issue of water from the earth; a natural fountain; a source of a reservoir of water.

Stable channel - Channel whose discharge rating remains unchanged for relatively long periods of time, generally between major floods.

Staff Gauge - A staff gauge that is placed on the slope of a stream bank and graduated so that the scale reads directly in vertical depth. A vertical staff graduated in appropriate units is placed so that a portion of the gage is in the water at all times. Observers read the river stage off the staff gage.

Stage - The level of the water surface above a given datum at a given location. Observed gauge height reading.

Standard Deviation - A measure of dispersion of data. Data grouped closely about their mean have a small standard deviation; data grouped less closely have a larger standard deviation. Abbreviated s.

Storage - Water artificially impounded in surface or underground reservoirs for future use. Water naturally detained in a drainage basin, such as ground water, channel storage, and depression storage. Also used to depict the capacity of a reservoir below the elevation of the crest of the auxiliary spillway. Usually expressed as million cubic metres or thousand hectare-metre of storage.

Storm Hydrograph - A hydrograph representing the total flow or discharge past a point.

Stormwater Discharge - Precipitation that does not infiltrate into the ground or evaporate due to impervious land surfaces but instead flows onto adjacent land or water areas and is routed into drainage.

Stream – A general term used for a body of flowing water, a natural water course containing water at least part of the year.

Stream Gauge - A site along a stream where the stage (water level) is read either by eye or measured with recording equipment.

Stream Segment - Refers to the surface waters of an approved planning area exhibiting common hydrological, natural, physical, biological, or chemical processes. Segments will normally exhibit common reactions to external stresses such as discharge or pollutants.

Streamflow - Water flowing in the stream channel. It is often used interchangeably with discharge.

Streamline - A vector drawn tangentially to the flow of water or other moving fluid.

Subsurface Runoff - Water that infiltrates the soil and reappears as seepage or spring flow and forms part of the flood hydrograph for that storm. Difficult to determine in practice and seldom worked with separately.

Subsurface Storm Flow - The lateral motion of water through the upper layers until it enters a stream channel. This usually takes longer to reach stream channels than runoff. This also called interflow.

Surcharge Capacity - The volume of a reservoir between the maximum water surface elevation for which the dam is designed and the crest of an uncontrolled spillway, or the normal full-pool elevation of the reservoir with the crest gates in the normal closed position.

Surface Impoundment - An indented area in the land surface, such as a pit, pond, or lagoon.

Surface Runoff - The runoff that travels overland to the stream channel. Rain that falls on the stream channel is often lumped with this quantity. Total rainfall, minus interception, evaporation, infiltration, and surface storage, that moves across the ground surface to a stream or depression.

Surface Storage - Natural or human-made roughness of a land surface, that stores some or all of the surface runoff of a storm. Natural depressions, contour furrows, and terraces are usually considered as producing surface storage, but stock ponds, reservoirs, stream channel storage, etc., are generally excluded.

Surface Water - Water that flows in streams and rivers and in natural lakes, in wetlands, and in reservoirs constructed by humans.

Systematic Errors - Errors that may occur because of defects in the instruments, in their exposure, or in the observational procedure. A gradual change in the surroundings of a station may be a source of systematic error.

Thiessen Polygon Method - A method of using a rain gage network for estimating average depth of rainfall over a watershed.

Time of Concentration - The time it takes runoff to travel from the hydraulically most remote point of the watershed to the outlet. Time of concentration varies from storm event to storm event, but is often used as a constant. Time of concentration consists of three hydraulic components: sheet flow, shallow concentrated flow, and channel flow. Abbreviated Tc.

Transmission Loss - A reduction in volume of flow in a stream, canal, or other waterway due to infiltration or seepage into the channel bed and banks.

Travel Time - The time required for a flood wave to travel from one location to a subsequent location downstream. The average time for water to flow through a reach or other stream or valley length.

Undercurrent - A current below the upper currents or surface of a fluid body.

Unit Hydrograph (or Unit graph) - The discharge hydrograph generated due to one centimetre of excess runoff distributed uniformly over the entire basin for a given time period.

Unit Hydrograph Duration - The time over which one inch of surface runoff is distributed for unit hydrograph theory.

Unit Hydrograph Theory - Unit Hydrograph Theory states that surface runoff hydrographs for storm events of the same duration will have the same shape, and the ordinates of the hydrograph will be proportional to the ordinates of the unit hydrograph. For example, the hydrograph for 0.5 centimetre of storm runoff will be half that of that from the unit hydrograph. Also, the response will not vary with time. This allows break up of storm rainfall into smaller duration rainfalls, the impacts of which are calculated independently and superimposed with applicable lag time.

Unstable Channel - Channel whose discharge rating is changed frequently by minor rises or, in alluvial channels, continually during all flow conditions.

Water Table - The level below the earth surface at which the ground becomes saturated with water. The water table is set where hydrostatic pressure equals atmospheric pressure.

Water Year/ Hydrologic Year - The time period from June 1 through May 31.

Watercourse - Any surface flow such as a river, stream, or tributary.

Watershed - Land area from which water drains toward a common watercourse in a natural basin. A watershed is always defined with respect to a specified outlet point on the watercourse.

Weibull Plotting Position - Values used to plot a frequency curve.

Weir - A structure built across a stream or channel for the purpose of measuring flow (measuring or gaging weir). A structure built across a river that allows water to flow from the main river channel into a side canal as required.

Wetland - An area that is regularly wet or flooded and has a water table that stands at or above the land surface for at least part of the year.

Zone of Saturation - The locus of points below the water table where soil pores are filled with water. This is also called the phreatic zone.

References

Boiten, W. (2005). Hydrometry, IHE Delft Lecture Note Series, A.A. Balkema Publisher, The Netherlands.

Buishand, T.A. (1984), Tests for Detecting a Shift in the Mean of Hydrological Time Series, Journal of Hydrology, 73, 51-69.

CWC (2001). Manual on Estimation of Design Flood, Central Water Commission, New Delhi.

Fulford, J.M. and Davies, W. J. (2005). Radar Water-Level Measurement for Open Channels, Water District Management and Governance, Proceedings of the Third International Conference on Irrgation and Drainage, USCID, San Diego, California.

Herschy, R. W. (2009). Streamflow Measurement, Taylor & Francis, New York, USA.

IS 15119 (Part 2): 2002. Indian Standard Measurement of Liquid Flow in Open Channels, Part 2, Determination of the Stage-discharge Relation, Bureau of Indian Standards, New Delhi.

Kahya, E. and Kalayci, S. (2004), Trend Analysis of Stream Flow in Turkey, Journal of Hydrology, 289, 128-144.

Kiely, G., Albertson, J.D. and Parlange, M.B. (1998), Recent Trends in Diurnal Variation of Precipitation at Valentina on the West Coast of Ireland, Journal of Hydrology, 207(3-4), 270-279.

Levine, D.M., Stephan, D.F., Krehbiel, T.C. and Berenson, M.L. (2010), Statistics for Managers Using Microsoft Excel, 5e, PHI Learning Private Limited, New Delhi.

Mueller, D. S. and Wagner, C. R. (2006). Application of the Loop Method for Correcting Acoustic Doppler Current Profiler Discharge Measurements Biased by Sediment Transport, Scientific Investigations Report 2006–5079, U.S. Department of the Interior, U.S. Geological Survey.

Rennie, C.D. and Rainville, F. (2006). Case Study of Precision of GPS Differential Correction Strategies: Influence on aDcp Velocity and Discharge Estimates, ASCE Journal of Hydraulic Engineering, Vol. 132, No. 3, 225-234.

Shahin, M., Vanoorschot, H.J.L. and Delange, S.J. (1993), Statistical Analysis in Water Resources Engineering, A.A. Balkema Publishers, The Netherlands.

Snedecor, G. W., and Cochran, W. G. (1994). Statistical methods (8e), Calcutta, India: Oxford & IBH Publishing Co.

Tilrem, O.A. (1979). Ministry of Water, Energy and Minerals Tanzania and Norwegian Agency for International Development, Manual on Procedures in Operational Hydrology, Volume 4, Stage- discharge Relations at Stream Gauging Stations.

USDI (2017). Field Performance Evaluation of Water Level Sensors, Research and Development Office, Science and Technology Program, (Final Report) ST-2017-

7833-01, Hydraulic Laboratory Report HL-2017-07, U.S. Department of the Interior, Bureau of Reclamation.

Yu, Y.S., Zou, S. and Whittemore, D. (1993), Non-parametric Trend Analysis of Water Quality Data of Rivers in Kansas, Journal of Hydrology, 150, 61-80.

Yue, S., Pilon, P. and Cavadias, G. (2002) Power of the Mann-Kendall and Spearman's Rho Tests for Detecting Monotonic Trends in Hydrologic Series, Journal of Hydrology, 259 (1-4), 254-271.

NATIONAL HYDROLOGY PROJECT

Department of Water Resources, RD & GR

Ministry of Jal Shakti

2nd Floor, Rear wing, MDSS (MTNL) Building, 9, CGO Complex, Lodhi Road, New Delhi 110003

Email: nhp-mowr@gov.in I Website: https://nhp.mowr.gov.in