TRAINING COURSE

ON

SOFTWARE FOR GROUNDWATER DATA MANAGEMENT

UNDER

WORLD BANK FUNDED HYDROLOGY PROJECT

LECTURE NOTES
ON

HYDROLOGICAL DATA AND ITS CHARACTERISTICS

BY

A K KESHARI

R D SINGH

ORGANISED BY

NATIONAL INSTITUTE OF HYDROLOGY
ROORKEE - 247 667
INDIA

HYDROLOGICAL DATA AND ITS CHARACTERISTICS

1.0 INTRODUCTION

Computer simulation techniques in hydrology encompass a wide range from relatively simple models to the highly complex models. Each model requires data in a computer-compatible form sufficient to conduct the analysis to the required level of accuracy. As a rule, the more complex is the model, the requirement for detail and accuracy in the data used by the model is greater. In this lecture, general requirement of data for hydrological studies alongwith the important hydrological variables and their characteristics, from view point of computerised storage are discussed. The formats for different types of hydrologic data are also discussed.

Maps provide an excellent means of summarizing the large quantities of hydrological information which are now made available by modern technology in an easily understood manner. They also provide an efficient and effective way of communicating information and ideas for the purposes of education, design, planning or convincing political, social and economic decision makers. This lecture describes the different types of maps and atlas used for the hydrological studies.

Before taking up the data processing tasks, different sources of errors in hydrological data should be known. This would be helpful for updating and validating the observed hydrological records. The lecture discusses the possible sources of errors in hydrological data.

2.0 GENERAL DATA REQUIREMENT FOR HYDROLOGICAL STUDIES

The data required for most of the hydrological studies may be broadly classified in the following groups:

(i) Hydrometeorological data
- Precipitation
- Evapotranspiration
- Radiation
- Temperature, Humidity and Vapour Pressure
- Wind speed and Wind direction
- Cloud cover
- Sunshine

(ii) Surface water data

(a) Water quantity data
- Water levels in the streams
- Discharge
- Levels of lakes and reservoirs

(b) Sediment data
- Bed sediments
- Suspended sediments

(iii) Ground water data
- levels of wells
- Log data
- Spring flows
- Well/Aquifer test data
- Pumping and Recharge data

(iv) Water quality data (Biological and Chemical)
- Water quality of rivers and streams
- Water quality of lakes and reservoirs
- Water quality of spring flows
- Water quality of groundwater reservoirs
- Water quality of waste disposals

(v) Process based data
- Infiltration
- Soil moisture
- Seepage
- Baseflow

(vi) Water use data
- Effluent to rivers, lakes and reservoirs (quantity as well as quality)
- Abstractions from rivers, lakes, reservoirs and ground water regimes

(vii) Reservoir Data
- storage-elevation relationship
- Area-elevation relationship
- Elevation-discharge relationship

(viii) Physical parameters

(a) Land surface physical parameters
- Area
- Elevation (topography)
- Slope
- Overland length and slope
- Vegetation
- Soil

(b) Natural drainage channel network
- Geomorphology
- Channel cross sections
- Roughness characteristics

(c) Aquifer systems characteristics
- Thickness
- Porosity
- Storage coefficient
- Transmissivity
- Dispersivity

3.0 DATA FORMATS

Except for remote sensing applications, all hydrological data is collected from individual sampling points distributed unevenly in space. This is very important as it defines a single basic format necessary for the spatial representation of data. Since it is impossible to infer the location of a sampling point from knowledge of the position of other sampling points, all data must carry with it a location marker; either geographical coordinates, or a unique sampling point number which enables the coordinates to be found in a separate reference file.

The format of data with regard to time has several different but classifiable structures. In fact, most natural resource data needs two storage formats for each sampling point. The first format is required for fixed (or stationary, or non-time series) data which does not vary, or varies only slowly, with time. The second format is for the time series: the series of data values needed to represent physical parameters having significant temporal variation. Raingauge location and elevation are items of fixed data, whilst the daily rainfall values recorded at the gauge represent a time series. Time series data may be categorized as continuous, regular, or irregular in time. Examples of these three series are water level tracings on charts, daily rainfall totals, and flood gaugings. Fig. 1 shows the format of time series records.

The above distinctions are made because the different categories need handling in different ways. Regular time series require only the start date, the interval between observations, and the number of observations stored in each physical computer record, i.e. whether one record contains ten days, one month or one year of daily data. For irregular time series, each observation must be accompanied by its relevant time reference. Continuous series need to be converted into digital form for storage, and this digitization process will create a regular or irregular series. When a digitizer is used to abstract chart data, an irregular time series format is produced, even though the points may have been abstracted at regular time intervals. This is because the digitizer generates pairs of time value (X,Y) coordinates. It is possible to convert this to a regular series for storage purposes.

Repeated pattern
of parameter values

Station Number	Parameter type code	Time of 1st value	No. of cycles	

(a) Fixed interval

Observation pairs

Station Number	Parameter type code	Time of 1st value	No. of value pairs	First Observation		Second Observation	
				Time value	Parameter value	Time value	Parameter value

(b) Irregular interval

Values for each parameter in the set

Station Number	Parameter set code	Time of observation		Parameter 1 value	Parameter 2 value	

(c) Multiple Parameter Series with Fixed Parameter sets

Pairs of values for each parameter

Station Number		Time of Observat ion	No. of Values	Parameter 1		Parameter 2		
				Code	Value	Code	Value	

(d) Multiple Parameter Series with Variable Parameter sets

Fig. 1 : Format of Time Series Records

A special type of time series common in water data is the multiple parameter time series. These are series where several observations are made at the same site. Examples are climate stations and water quality sampling points. Much of this data normally comes from the field on a single data sheet, and is input in this format. Whilst the data could also be stored in this way, there are advantages in an alternative format which separates all the variables into several single parameter time series files.

A variation of the multiple parameter series is encountered when data are sampled at various depths at the same location, e.g. water quality and certain sediment samples. These types of data require space, time and depth references. In an exactly analogous way to time, the regular or irregular spacing of sampling depths will determine the type of depth reference required.

Typical spatial data structures are shown in Fig.2. These include gridded or vectoral data structures, and it must be recognized that in future, the collection of some types of water data by satellite remote sensing will require such geocoded data base systems. However, very specialized software such as Geographical Information System (GIS) is required to handle data of this type, a major problem for remote sensing data being compression of the vast quantities of data to an amount which may be included in the usual scale of most data bases. A typical representation of well/aquifer test data sheet is shown in Fig. 3. Fig. 4 shows a stage discharge data format adopted by CWC.

AREA IDENTIFIER	DATA CATEGORY	--------------------	CATEGORY N

↓

DATA CATEGORY 1		
ITEM 1	--------------------------	ITEM N

↓ EACH ENTRY GIVES THE PROPORTION OF THE SUB-AREA WITH THE GIVEN ATTRIBUTE

ITEM 1			
VALUE 1	VALUE 2	VALUE 3	VALUE 4
24	0	0	76

OR

ALTERNATE PRESENTATION OF SUB-AREA GIVEN OVER TO PARTICULAR ATTRIBUTE VALUE

ITEM 1			
VALUE CODE	AMOUNT	VALUE CODE	AMOUNT
1	24	4	76

Fig. 2 : Table representation of area and sub-area data

Sheet No.:_______

Circle No.:_______ Test Data for Pumping Well Observation Well Well No.:_______

Well Location:____

Well Owner: ______________________ Distance to pumping well:________________

SWL Depth:_________ Reference point elevation: _________ Test Date:________________

Data File:________________

Circle One: STEP DRAWDOWN CONSTANT RATE RECOVERY

Time of Day	Time	Depth to Water	Drawdown	Discharge	Comments

Fig. 3 : Well/Aquifer test data sheet

1-5	6-12	13-14	15-16	17	18-22	23-29	30-34	35-41	42-46	47-53	54-58	59-65	66-70	71-76
Blank	Site Code	Year	Month	Card No.	Gauge Day-1	Discharge Day-1	Gauge Day-2	Discharge Day-2	Gauge Day-3	Discharge Day-3	Gauge Day-4	Discharge Day-4	Gauge Day-5	Discharge Day-5

Note: In Card No. 7; Col. 30-34 contain zero of R.L., and
Col. 42-46 contain temporary R.L.

Fig. 4 : Stage Discharge Data Format adopted by CWC

4.0 IMPORTANT HYDROLOGICAL VARIABLES AND THEIR CHARACTERISTICS

There are three distinct phases of the water cycle which provide an initial sub-division of hydrological variables; atmospheric, surface and sub-surface. The atmospheric hydrometeorological components of the water cycle of most interest to the hydrological sector are precipitation, evapotranspiration, temperature and radiation. Apart from some monitoring of soil moisture and temperature in the unsaturated zone, the sub-surface phase is restricted almost entirely to groundwater.

4.1 Hydrometeorology

The quantitative estimation of the atmospheric phases of the hydrological cycle requires observation of precipitation and of other climatological elements, such as air temperature, air humidity, wind, pan evaporation, etc. (Fig. 5). For hydrological purposes, the latter climatological elements are used either directly or indirectly to assess evaporation and evapotranspiration. Other uses include a range of agrometeorological services to farmers.

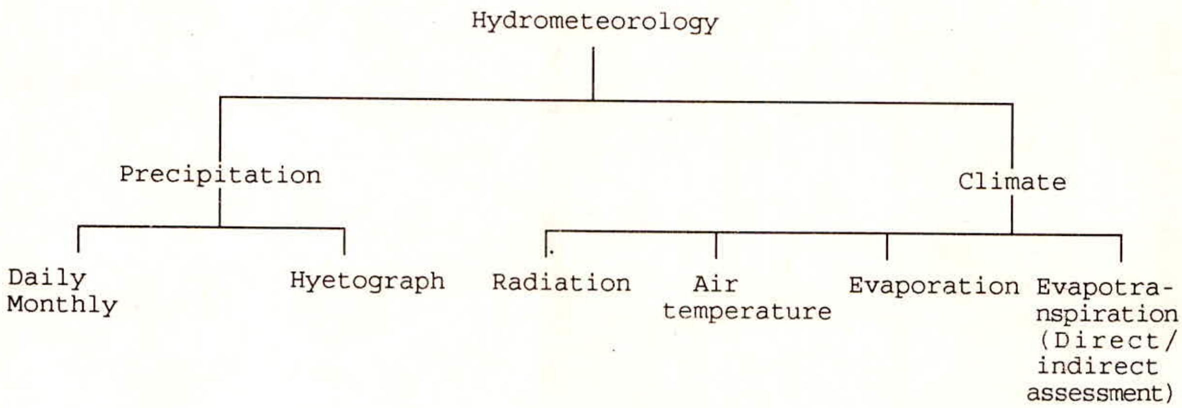

Fig. 5 : The Components of Hydrometeorological Data

4.1.1 Precipitation

It is imperative to have reliable precipitation estimates as they represent the upper bound on available water resources. A particular feature of precipitation is its extremely wide variation in time and space, and for this reason, it will always be a significant component of any hydrological data collection and analysis system. Precipitation includes all forms of water that falls from the atmosphere to the earth's surface. The meteorologists are interested in a variety of forms of water, but the hydrologists are interested in distinguishing only between liquid precipitation and frozen precipitation. Rainfall is a liquid precipitation, and is the most important form of precipitation for the most parts of India. Rainfall runs off to the stream soon after it reaches the ground, and is the cause of most floods.

Data from conventional precipitation networks is most commonly in the form of daily read rainfall. This observation technique yields a simple regular interval time series. In remote areas, particularly where a reliable on-site observer cannot be found, monthly read gauges may be used, again producing a regular time series. For the same reasons, and usually in areas where precipitation is very sporadic, *accumulating* gauges may be used. These are visited infrequently, usually after precipitation was expected, or was known to have occurred. Such data are processed as an irregular time series. An alternative approach is to apportion the accumulated total precipitation according to the totals recorded at adjacent daily stations, producing a regular time series. Data from these manually read totalizing gauges comprises by far the greatest input into any precipitation processing system.

For precipitation in the form of rainfall, automatic gauges are widely used, normally recording data in chart form. There are two main considerations in abstracting digital data from this type of chart. First, whether the abstraction should be performed manually or automatically, and second, whether the abstraction should be made at regular or irregular time intervals.

Regardless of the abstraction technique, digitized rainfall charts will ultimately be stored as irregular series. It is easy to see that storing rainfall totals for, say, every 15 minutes during extended periods without rainfall would be extremely wasteful of storage space. For hydrological network design and data processing, the time interval between stored data points should be the maximum possible whilst allowing interpolation of values at intermediate times to an acceptable level of accuracy.

4.1.2 Other climatological data for hydrological purposes

Co-operation between the Meteorological and Hydrological Services should also extend to climatological data as both will operate climatological stations. Climatological data is derived from stations measuring a limited sub-set of variables, or from stations specifically built to measure the complete range of variables required for hydrometeorological (or agrometeorological) purposes. The highest order of climatological stations are normally operated and manned by staff from the Meteorological Service, their purpose being to provide synoptic weather information. Data returns from climatological stations appear as multiple parameter time series, for which a standard storage format may be used. Rainfall data will be copied over into the single parameter rainfall system files.

The difficulties in processing climatological data arise mainly from the wide range of parameters which can be measured at a single station. There is also a wide variation in the type of instrument used to sense any given parameter, each of which may require some conversion or correction when being processed. The processing system must therefore be aware of the observation method and instrument used at different stations in order that the relevant adjustments can be made. Where evaporation values are derived by indirect methods, look-up tables need to be stored to compute theoretical incoming radiation and daylength, or alternatively they may be derived from the basic physical equation. If it is required to compute theoretical evapotranspiration from several surface types, data on roughness coefficients and albedoes are required for each surface.

Evaporation is an essential component of hydrological cycle which plays a major role in water balance studies and assessment of water availability from lakes, reservoirs, and tanks. Anticipated evporation is a decisive element in the design of water impounding structures. Further, the knowledge of evaporation from water bodies plays a vital role in drought alleviation schemes. A rought estimate indicates that about 33% of the total storage from the reservoirs, tanks and lakes, is lost through this process in our country. Nearly two-thirds of the precipitation that reaches the land surface of the earth is returned to the atmosphere by the evapotranspiration. In arid regions, evaporation may consume a large protion of water stored in reservoirs. Variability of evaporation from year to year is whether less or more than the variations in streamflow or precipitation, is an important characteristic for hydrologists.

In those areas where evaporation potential substantially exceeds annual precipitation, construction of a reservoir means a large loss of water through evaporation. Where evaporation potential is high, runoff tends to be low, since runoff is essentially the residual after evapotranspiration requirements are substracted from precipitation. The transpiration pattern is modified by the vegetal characteristics of the different regions of the country. In a desert region with little or no vegetation, transpiration is necessarily low, although evaporation potential is high. In no case evapotranspiration can exceed precipitation except when water from another basin is available to augment the local supply. Actual evapotranspiration drops below its potential level as the soil dries out.

The total quantity of transpiration by plants over a long period of time is limited primarily by the availability of water. In areas of abundant rainfall, well distributed through the year, all plants will transpire at about the same rates, and the differences in total will result from the differences in the length of the growing seasons for the various species. Where water supply is limited and seasonal, depth of roots becomes very important. Here, shallow rooted grasses wilt and die when the surface soil becomes dry, while deep-rooted trees and plants will continue to withdraw water from lower soil layers. The deeper-rooted vegetation will transpire a greater amount of water in the course of a year. The rate of transpiration is not materially reduced by decrease in soil moisture until the wilting point of the soil is reached. On a natural catchment with many vegetal species, it is reasonable to assume that evapotranspiration rates do vary with soil moisture since shallow rooted species will cease to transpire before deeper rooted species.

4.2 Surface Water

The components of surface water data collection are given in Fig. 6. Although there are parallel needs for data on both quantity and quality in all components of surface water monitoring, water quality is shown as a separate sub-division. This reflects the fact that water quality is generally developed as a major processing sub-system. In consequences, the discussion of surface water systems may be similarly divided.

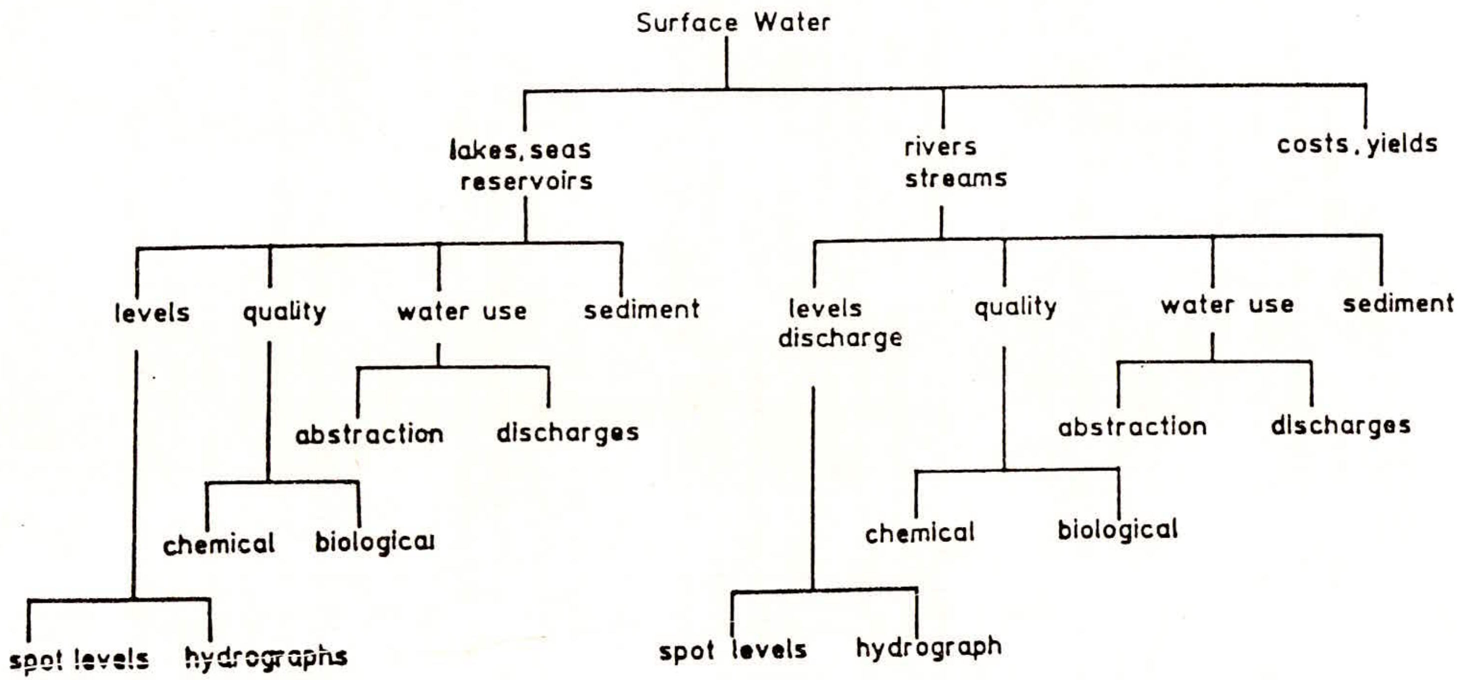

Fig. 6 - The components of surface water data

4.2.1 Water quantity

The primary breakdown of data is made between flowing and non-flowing bodies of water. The reason for this distinction lies in the processing requirements. For non-flowing bodies of water, it is generally the level (or some simple level-volume relationship) which is of direct interest. For flowing waters, there is the additional need to compute flow rates, the form in which the data will most generally be used. This level-discharge conversion is only straightforward if some constant cross-section hydraulic control is used for which the level-discharge relationship is known. Such sections are frequently impractical because of the nature and size of watercourses to be gauged. In these cases, current metering, slope-area estimation and dilution gauging may be used for single measurements of discharge, and a series of measurements for a range of levels enable the construction of the level-discharge relationship. All such methods rely on either the selection of a stable bed profile or, more usually, the periodic redifinition of the profile. This redefinition modifies the level-discharge relationship.

Additional complexities arise in applying the calibration equations because of other variable hydraulic controls, e.g. different sluice settings or gate openings, or seasonal differences in the channel roughness conditions. The processing system must store the various calibration equations in table, or preferably in equation format, know when to apply them, and flag each set of processed flows with an identification code for the particular calibration curve used. Where level data are taken from chart recorders, exactly the same considerations apply as for the processing of rainfall charts. One important difference is that there is generally much more persistence in flow data, most gauges being positioned where there are perennial streams.

A particular problem of flow processing is deciding what to do with the stage measurements once flows have been derived. Many applications require only flows; levels may be re-computed subsequently from the relevant calibration relationship. However, this level data, by whatever means it was collected, does represent the basic data source, and as a general rule such data should be preserved. Autographic charts may be stored on microfilm or microfiche. A common solution is to process raw level data into a validated regular time series and store it off-line on a magnetic tape. This allows complete reprocessing if say, it was found that the wrong calibration curve had been used to compute flows. Another solution is that the original series of field observations of levels should form this fundamental series which are archived together with the stage discharge relations. Discharge series are computed whenever discharge data are required. This avoids the need to maintain in effect, two archives, allows for easy inspection of original field observations, and permits ready alteration of stage discharge relationships should that subsequently prove necessary.

4.2.2 Sediment

Two types of sediment sampling are used in most hydrological studies; bed sediments and suspended sediments. The suspended sediments can be grab samples, continuous point samples or depth integrated samples. Analysis may include particle size analysis and total concentration. In addition to information on the exact type and location of the sample, the discharge and perhaps the velocity at the site or point of sampling must be measured/calculated and recorded.

The problem with identifying any kind of water sample is accurately specifying the location of the sample. In the case of a river, the usual way is to measure from one bank and give the depth of sampling if relevant. This information, together with the water level, is sufficient to specify the exact location of the sample rather than its relative location. The water level can also be used to estimate water discharge if a rating curve is available for the site. In the case of lake or sea, the x-y coordinates or grid reference of the site has to be given as well as depth and water level. In the case of tidal waters, it is useful to know the state of the tide at the time of sampling.

The storage of sediment data can be treated in exactly the same way as other water quality data, or a special file can be kept. In a small system, it may be better to collect all sediment data together and treat it separately. If the standard water quality format described below is used then the records for each sample are divided into two parts, a sample identifier and environmental information and details of the analysis which may then also include chemical analysis. This is true for bed samples and suspended samples. If an individual sample forms part of a sequence of observations, this should be clear from the coding system used.

When sediment discharge has been computed, it is best to store the results in a file or group of files containing at least the following information:

- Site details
- Sediment discharge - by size fraction
- Water discharge
- Cross-sectional area
- Wetted perimeter
- Water temperature
- Date and time of sampling
- Water level-gauge reading

These data enable sediment rating curves to be prepared and periodically updated.

4.3 Groundwater

Groundwater data can be dealt with under three headings (Fig. 7): springs, wells and yields/costs. Well water level is still most often obtained by manual dipping techniques and usually at irregular intervals, producing an irregular time series format. If the wells are pumped, estimates of quantities abstracted may come from meters, from duration of pumping, or from quantity of power consumed. These estimates require knowledge of the pump specification and the pumping head, and may be performed by the processing system. In addition to defining the surface drainage system of which the spring or well is a part, it is necessary to code the aquifer system(s). In the case of wells there is, in addition to the basic

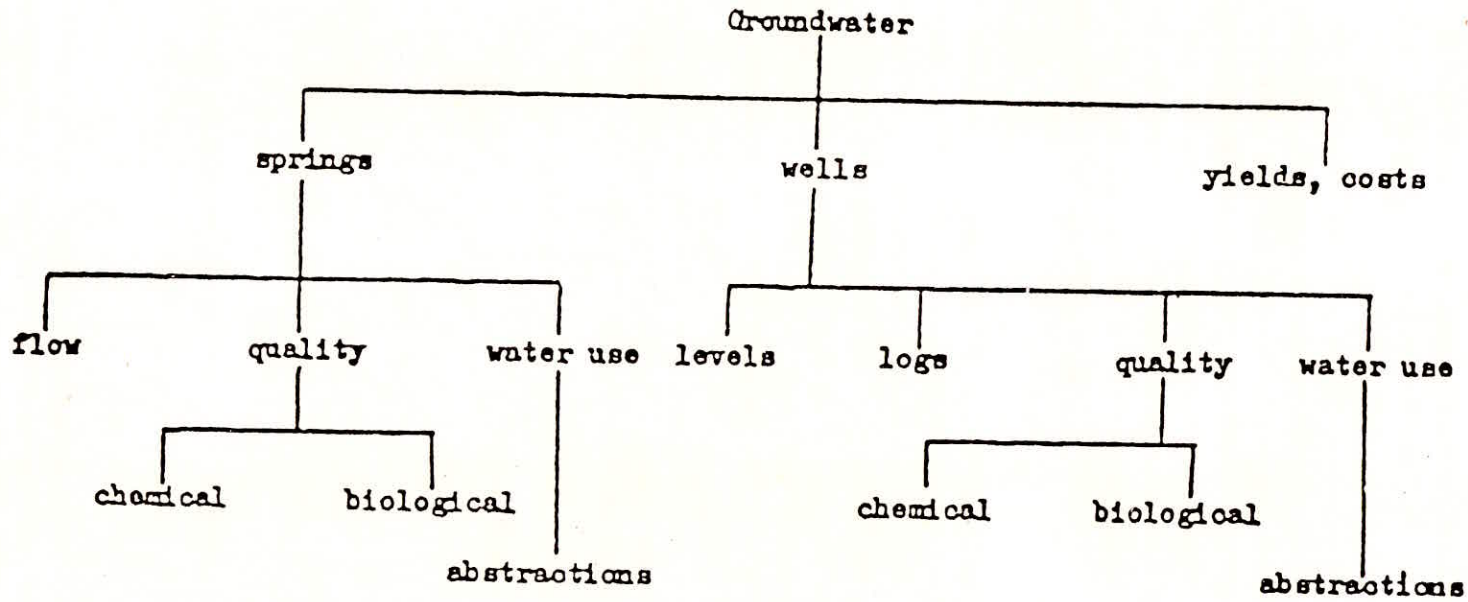

Figure 7 - The components of groundwater data

site description data, the full geological and hydrogeological logging of the hole.

It is probably better to save separate collections of data on wells, the details of discharge and water quality, the site location including licensing arrangements, and the details of the logging of the hole. Water level, discharge and water quality are time series data sets which can be stored in the standard formats. Site location and logging are single sets of observations which do not change with time. More information on the specification of fixed hydrological station data is given in section 5.0.

The natural resource data base system (WMO-634) is well suited to storing geological and hydrological bore-hole data. Each horizon can be coded and the parameters of each horizon recorded. To devise the coding system for such a data bank is a large job and existing systems should be first reviewed to assess their suitability. However, a simple start can be made by noting, against each bore-hole, the location of reports on that hole in conventional form. This system would start by being an aid to data retrieval in a conventional library sense. For many kinds of data, the coding and abstraction effort required to store the full, detailed information is simply not worthwhile for the number of occasions on which it would be used.

4.4 Water Quality

Water quality has been divided into two groups, chemical and biological. A feature of both these groups is the enormous range of tests and analyses that may be performed on any sample, both in the field and, more particularly, in the laboratory. The processing problems which this presents are accentuated by the different sets of analyses performed on samples from different stations and from the same station at different times. Water quality analyses yield a multiple parameter time series but the range of parameters investigated requires that each data value be associated with a parameter code and perhaps a code indicating the analysis method. This necessitates the preparation of dictionaries for chemical and biological terms and analyses and the formulation of codes for each. Thus, the final multiple parameter format used will be the station identifier, the time and possibly depth identifiers, the number of analysis results in the record, and the corresponding number of pairs of parameter/analysis codes and values. This format is shown in Fig. 1(d).

Some water quality sampling points are located at flow gauging stations in order that mass balance calculations may be performed. Thus, some common link between quantity and quality files is required.

Other problems of water quality data handling are the identification of samples by location, type, purpose, etc. In general, a sample is taken at a point and so its unique identification is given by its x, y and z coordinates. However, some samples are depth integrated and some are taken within the framework of a river cross-section. In almost all cases, the x, y and z coordinates given are relative rather than absolute. It is almost universal to specify the depth of the sample, this implies that the water level must also be recorded if an absolute value of the vertical coordinate is to be known. In the case of cross-sections, the distance from a datum on one bank is given. For open water, the x - y coordinates are given, either as a map reference or in some arbitrary framework, the origin of which must be specified somewhere.

In a comprehensive data storage system, it is very desirable, except where the same site can be used each time a sample is taken, that site locations are given some standard grid reference. This is particularly relevant for lake and sea data. When the same site is used repeatedly, as at a river cross-section, then relative coordinates may be recorded.

It is impractical to give every sample collected an individual number to identify it. The sample's full identity should be given by its location, the time the sample was taken and a local identification number which would be retained through the laboratory analysis. It is not necessary to save the local sample number within the computer storage system after data validation. It is impractical to give every sample a different number because most sample collectors would not bother to check if a given number had already been used before using it again, especially when there were long gaps between sampling.

To identify a sample, the following items of data may be used:

(a)	Site number	⇔	Unique site number
(b)	Sample number	⇔	Only to label the data for collection, analysis and collation
(c)	Type of sample	⇔	Point, depth integrated, bed
(d)	Date and time		
(e)	Set number	⇔	Identifies all samples in a set
(f)	Water level	⇔	Water discharge
(g)	Water temperature	⇔	Environmental parameter may be coded in water quality format
(h)	Depth		
(i)	Distance of x coordinate	⇔	Distance from bank marker or grid coordinates of sample set
(j)	y coordinate	⇔	For lake or sea samples

These data enable records to be sorted and stored, placing all data from one site, of a single kind, at the same time and of the same set all together. This facilitates rapid retrieval in the most frequently required form.

While the analyses may be identified for, and presented to, the computer as a series of parameter codes and associated data values, it is not essential to store the data in the computer in that form. The data may be stored as a simple two-way table in which the rows represent the samples and each column is used for a particular analysis or water quality parameter. Data compression techniques may be used to conserve computer space if many columns are frequently empty. The advantage of this format is that the analysis codes need not be stored with each data value. Using this method of coding data, it is possible to permit the use of a wide variety of forms of data which can be collected by different agencies to suit their own needs, in addition to those of any central data base facility.

4.5 Water Use

Water use data may be divided between abstractions and effluents. There are of course several water uses such as transportation an amenity which fall into neither category, but data on such water uses are rarely kept. Abstractions and discharges have a direct effect upon both the quantity and quality of a water resource. It is difficult to generalize on the format of these data as they vary considerably in type. Some large abstractions and effluents may be measured directly, most will be estimated, and the quantities involved may be constant or widely fluctuating with time.

In basins where there is a high degree of water use, it becomes difficult to assess the true *natural* hydrological regime, and several processing systems have recognized this fact by archiving gauged flows, abstractions (and associated net water use coefficients), effluents and reservoir volumes, and subsequently attempting to assess the *natural* flows for hydrological studies, particularly those which seek to assess the effects of various resource development options through time.

4.5.1 Effluents

It is necessary to be aware of all significant effluents to the surface water system and their associated water quality. Depending on their size, effluents may be measured continuously or, more usually, on a daily or monthly time basis. Quantities may be stored in the same way as river discharges and water quality using the standard formats referred to elsewhere. Both quantity and quality data of effluents are of prime importance for hydrological analysis. The growing concern of environmental consequences emphasizes the measurement and storing of water quality data. In the site description, any licensing arrangements will be included and these may vary from month to month and/or in relation to the natural flows or levels in the water bodies receiving the discharges.

The location of an effluent site must be specified in relation to the drainage network in the same way as a flow-measuring station. This may be done in a station description file.

4.5.2 Abstractions

Knowledge of abstractions is required for the same reasons as effluents. Once again the actual abstractions as well as the licensed abstractions are required. Actual abstractions are presented in the same format as other discharge measurements. If water quality is involved, the standard water quality formats are used.

The licenced and private groundwater draft are required for groundwater balance studies, and groundwater and contaminant transport modelling. Therefore, all abstractions from groundwater reservoirs should be recorded as a time series data. The private groundwater drafts may not be recorded with sufficient accuracy, but its direct/indirect assessment must be carried out. In addition to the drafts, the pumping location and its history also should be recorded. The groundwater draft may be for municipal supply, irrigation demand, or industrial supply. The storing of draft data separately for each head helps in various types of analysis.

Table 1 summarizes the general data collection and storage formats of the hydrological variables discussed above. Codes given in the Table 1 reflect the most usual data observation techniques.

Table 1 : Classification of hydrological data by format

Data	Type
Daily Precipitation	SPFITS
Hyetograph	CTS
Climate (temperature, evaporation,	MPFITS
humidity, etc)	SPFITS,
Lake level	CTS
Lake quality, chemical	MPVITS
Lake quality, biochemical	MPVITS
Sediment	SPVITS
River levels	SPFITS,
River discharges	CTS
River quality, chemical	SPFITS,
River quality, biochemical	CTS
Spring discharge	MPVITS
Spring quality, chemical	MPVITS
Spring quality, biochemical	SPFITS,
Well logs	CTS
Well levels	MPVITS
Abstractions	MPVITS
Well quality, chemical	NTS
Well quality, biochemical	SPVITS,
Yields - cost	CTS
Licensing	SPVATS,
	CTS
	MPVITS
	MPVITS
	NTS
	NTS

Code for data types:

SPFITS	- Single parameter fixed interval time series
SPVITS	- Single parameter variable interval time series
MPFITS	- Multiple parameter fixed interval time series
MPVITS	- Multiple parameter variable interval time series
CTS	- Continuous time series
NTS	- Non time series

5.0 STATION DESCRIPTION OF DATA

Previous reference has been made to the two categories of data which are used to characterize a station: data which is fixed in time and data which varies in time, e.g. water levels and discharge. Generally the fixed data is input when the station is first incorporated into the data base, and the sets of observations made at the station are added periodically to the station time series file.

It has also been noted that this general view may be too simplistic for some types of hydrological station. For instance, whilst the prime objective of a flow gauging station is to record a time series of water levels, there is for many stations a second time series formed by the set of level-discharge relationship. In recognition of the time variation of some station characteristics, data structures of the type shown in Fig. 8 have been utilized in several water data bases. The station history file is useful for all types of stations, whilst the level-discharge calibration file is required particularly for flow gauging stations. This structure allows the current station description and calibration files to be of fixed size and format, a useful feature when designing retrieval programs. The station description file contains data on the current instrumentation and time series data, location and format, and if this needs revision it is first copied to the station history file then replaced in the station description file by the new data.

For the purposes of discussion, all the data except the specific station time series, e.g. daily rainfall values, water levels and water quality values will be classified as station description data. Whether conceived as a single file or broken down as in Fig. 8, the variables needed to describe the location, purpose, equipment, administration and operation of hydrological stations can be readily identified. Table 2 summarizes these data, but also gives examples of station data specific to some of the hydrological variables identified in the previous sections.

Site description data files are essentially as found in any manual system, although more emphasis is placed on the location and format of other relevant data. For example, computerized hydrological station files need to contain explicit references to the disposition elsewhere in the computer of the associated time series files, or the dictionary files which convert codes for station, instrument and analysis types, watershed names, data reliability etc.

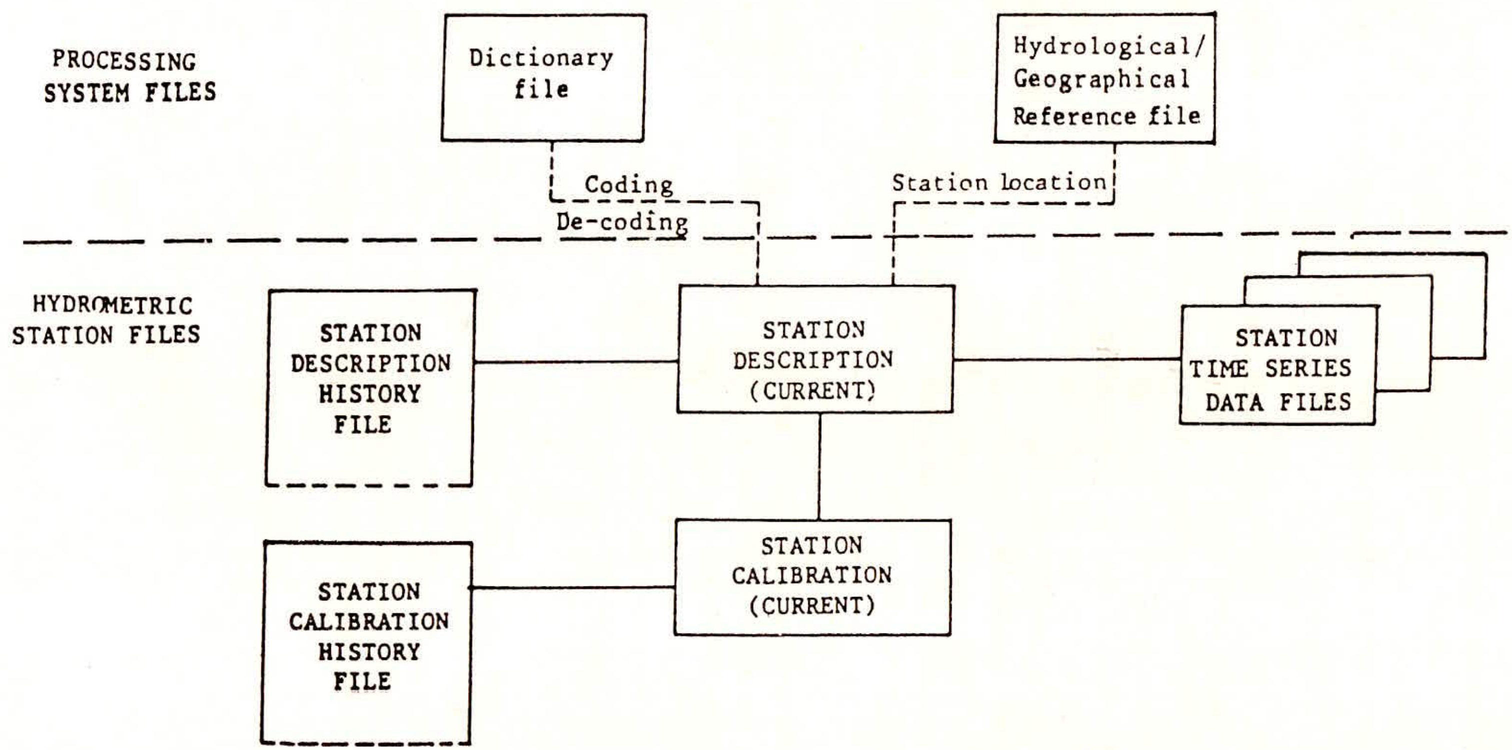

Figure 8 - Relationship of hydrometric station data files

Table 2 : Typical hydrological station description data

1. Type of station	-	river level, bore-hole etc.
2. Station number		
3. Grid coordinates	-	latitude, longitude or Universal Transverse Mercator projection (UTM)
4. Description of site location	-	how to get there, details of where to find surveys and full original details
5. Hydrological location	-	hydrological reference. Drainage area and/or aquifer(s)
6. Details of each instrument at station	-	type
	-	manufacturer
	-	serial number
	-	pointer to rating/calibration curves
	-	limits of calibration
	-	time and measurement scales of data recorder
	-	date installed
	-	date last serviced
	-	frequency of servicing
7. Details of licensing	-	name of licensee
	-	address
	-	consent conditions
	-	amount to discharge (+) or abstract (-) conditions may depend on time of year or flow condition (level or discharge)
	-	water quality - range of values allowed for each coded parameter, for each discharge condition
8. Structure of the time	-	for each kind of input record the position, series input data formats format and units of each data field.
9. Structure of the time series storage data formats	-	format (Figures 1 and 6)
	-	number of data values per record are multiple parameter data values identified by position or by parameter code plus data value?
10. Physical organization of time series files	-	which disks/tapes etc.
	-	for what period is data available (including missing segments)
11. Details of environment	-	data include exposure of site, altitude, etc.
12. Details of cross-sections	-	for river sections only, includes datum of location of verticals and level gauge. This may contain only the latest section or it may contain a selection of previous sections. If frequency of survey is high it may be necessary to create a separate time series file for this category.
13. Bore-holes	-	a basic minimum set of well data includes the depth, top of well datum, diameter schedules and which aquifer(s) is tapped (coded).
	-	reference to detailed manual or computer records to facilitate rapid access.
	-	coded geological succession and analytical parameters
	-	pump test data in standardized form
	-	constructional details

NOTE :The drainage area or aquifer codes must enable the discharge to be logically related to the basin or aquifer. In this case of the aquifers, x-y coordinates and definition of the pumped aquifers are needed. In the case of surface discharge, the site should at least be topologically ordered.

6.0 MAPS AND ATLAS

Many types of maps and atlas are needed for water resources development, day-to-day operations and hydrological design. A recent survey undertaken by UNESCO disclosed that major gaps exist in the map archives and mapping programmes of most countries, particularly in the domain of surface and sub-surface waters.

If the potential of mapping is to be realized, then certain rules must be followed:

(i) The water balance, energy balance and the demands of other physical controls be satisfied in the mapping process.

(ii) The nature of the phenomena being mapped and its response to other environmental factors, as well as the nature of the data and the limitations of measurement must be understood and considered from the earliest stages of mapping.

(iii) Standardization should be the rule to as great an extent as is possible; of data, of units, of isolines, scales and other factors which will make the product comparable and eliminate the risk of confusion in the mind of the user.

(iv) The basic rules of cartography should apply since these facilitate preparation and production. They also aid perception by reinforcing or diminishing detail so as to convey quickly a clear, concise meaning while safeguarding integrity.

6.1 Consistency

The extrapolation in space of point measurements has received much attention because of the need for consistency posed by the water balance. For example, in the mapping of precipitation, the water balance is being increasingly used as it provides a form of quality control for precipitation mapping. The spatial variability of the precipitation depends on vegetation, topography as well as meteorological processes of varied dimensions. It is not sufficient for mapping to have accurate measurements; the analyst must also be able to meaningfully interpret, extrapolate and interpolate a highly heterogeneous set of data, according to the aforementioned factors. By the use of grid maps, the conundrum of excessive detail can often be overcome and space averaged values obtained which are very convenient for large scale mapping of the water balance. However, a good knowledge to physical relationships remain the most valuable asset in this process.

Time consistency is usually possible for precipitation or runoff alone, but frequently the two sets of information have not been collected on an integrated basis. Groundwater data may pose particularly great difficulties. Although well levels may be recorded continuously, other basic data needed for scientific evaluation of yields, flows and storage changes tend to be obtained on an irregular basis which may span periods of years and have little regard for the season. Particularly, this is most often common in case of measurement and recording of water quality parameters. Water quality parameters are most oftenly measured at irregular time intervals, and also a different set of parameters is observed. Data selection is such that the networks are different at each season, e.g. spring and autumn. As a consequence, even the estimates of seasonal changes based on observations taken over a period of years may not be completely valid because of the lack of consistency in the measuring networks. This relative lack of 'synoptic' information makes a 'model' approach necessary in groundwater mapping in most instances, and poses a challenge for mapping in the future.

6.2 Data

Most data networks and the data themselves have imperfections because of space and time variabilities of elements as well as instrumental and procedural inadequacies, to say nothing of economic forces. The analyst is, therefore, required to resort to statistical procedures and conceptual models to overcome these data deficiencies. He must evaluate the completeness of records, inherent errors, the variability in space and time, and other characteristics of the element to be mapped, including external physical restraints. The number and quality of data as well as their statistical nature impose limitations as to what can be usefully mapped, the map scale, the precision of boundaries, the placing of isoliness and the significance of estimates obtained from the completed map. It is, therefore, most important that all aspects of the data, the supply, the quality and the statistical character be thoroughly understood before mapping proceeds. Furthermore, the resulting map or supporting text should clearly identify the nature of the data base, its quality and the period of records so that the user will not assign greater accuracy to his interpretations than is warranted.

The need for estimation arises when records are broken or sparse, and in the preparation of probability-type maps. Many estimation techniques are available, such as regression, the use of ratios and frequency-analysis transformations. The increasing availability of the electronic computer has made possible the use of fairly complex estimation procedures. As a consequence, it is now reasonable for such procedures to include expressions for land use, soils, vegetative cover and other physiographic detail on a rational basis such as through the use of modelling. Conventional techniques such as graphical regression remain, nevertheless, very practical tools for most mapping purposes.

6.3 Characteristics of Maps

At the same time that data are being accumulated, made homogeneous and checked for quality and statistical character, decisions must be made concerning the map, its size, scale, projection and base-map detail. The purpose of the map and the availability and nature of data are key factors in making this decision. On the other hand, the expert recommendations of national and international organizations should also be considered.

Small-scale maps are normally used in general planning and education. Large-scale maps are used in project analyses, basin studies etc. Maps which display details of networks or graphs of regimes at a number of representative stations may, of course, require larger scales than those used only to present the areal variations of an element.

No flat map is entirely true and the various projections employed make trade-offs in the accuracy of area, distance and direction, distorting one to the benefit of the others. For hydrological purposes, the errors in maps covering small areas are of little importance, but for continental or global purposes, these distortions do become important. For operational and water balance purposes, equal-area projections are desirable, but for hemispheric and global maps, equal-area projections such as the sinusoidal projection of Goode's projection should be used whenever possible. Where regional maps are used, a cylindrical projection is usually preferable near the equator, while conical-equal-area projection is best suited for regions in middle latitudes. Polar projections are commonly used at high latitudes. Larger-scale maps, such as for urban areas, are commonly drawn on universal transverse mercator projections.

The base-map information should complement, but not detract from the main theme of the map. At times, the base-map information is very pertinent to the interpolation of the hydrological fields, for example, when the river systems or topography are shown. Well-known land features, principal cities and administrative boundaries are included to aid in determining geographical locations. Hydrological networks are frequently shown as well, since they provide a good index of the reliance that can be placed on the analysis of the data that a map displays.

The nature of the data upon which the map is based should be identified in the legend or the adjoining text since it provides excellent guidance to the confidence which the user may place on the analysis. If a large number of data are used and they are homogeneous, it usually suffices to identify that fact and the period of observation. If they are not homogeneous, it may be necessary to use numbers or symbols on the map to indicate the quality of the data for each location. A description of estimation techniques and their validity should also be given.

Among the more commonly mapped characteristics of hydrological elements are the following :

(1) Networks of instruments
(2) Means or medians
(3) Departures from the mean
(4) Variability
(5) Total amounts for a specific event or duration
(6) Extremes, including the mean extreme and variability
(7) Number of days or months with specified conditions
(8) The time of beginning, duration and ending of phenomena
(9) Intensity with a specified frequency of occurrence
(10) Ratios for different durations or frequencies
(11) Combinations of several elements or components
(12) As graphs, variability in time intensity-duration frequency, annual regimes and mean directional array.

By and large, maps depicting these factors are very similar from the viewpoint of construction, isolines and point, line or areal signals being the main means of data depiction.

Isoline maps are by far the most common. From the scientific standpoint, it is important that the isolines do not convey more than is permitted by the combination of data and supplementary knowledge. They must be placed with due consideration of the errors of measurement, the surrounding data field and physical relations. The isoline should not just satisfy values plotted on the working map, for it can usually be fitted with confidence in any one of a large number of positions depending on the sample error. It is usually necessary, for practical reasons, to prepare a map with a specific isoline interval which is less than that allowed by the data. A note of caution should be given on the map in these instances, recognizing that for prognostic purposes the division between the isolines should be at lease twice the standard deviation of the considered variable.

Uniform isoline intervals should normally be used, however, emphasis of particular features may require a departure from this practice. Geometrically-increased spacing is often necessary over mountainous areas and often zones displaying great differences in precipitation or evaporation amounts. The isolines are sometimes reinforced by flat colours, hatching or other forms of infill. Isolines are used extensively in hydroclimatic mapping, but other hydrological maps make considerable use of point, line and areal symbols as well as grid-values and computer graphics.

Typical point-value maps are those showing depth measurements, point observations of chemical or biological concentration, or temperature. Bar graphs, such as those of seasonal runoff variation or of the frequency distribution of measured elements, also fall within this category.

Line symbol differences are achieved by variations in linear form, width and shading, and the use of bands and colour. The variations in linewidth are used to show differences in volumes or other characteristics; bands of different shading show such flood or runoff volumes for different probability levels.

Areal symbols are used to depict areal extent, such as flooded or contributing areas. The areas may be identified simply by boundaries, e.g. the boundary of a zone with a specified range of ratios of the 50-year-to-mean annual flood. Alternatively, hatchings, patterns or colours or alphanumeric codes may be used to distinguish between areas having specified characteristics.

When more than one element is to be presented there are several options which may be taken up. Different line symbols or colours may be used to distinguish two or more types of isolines.

Another approach is to use several maps, each to the same scale and projection, placed adjacent to each other on the map sheet. Transparent overlays can be very effective for displaying several elements; however, they are costly to produce, and therefore, are most commonly used in the preparation stage.

While there are as yet too many subjective factors in hydrological mapping to turn the task over completely to automation, there is great merit in using computers for many purposes. Computers are very useful in the quality control and evaluation of data, in preparing estimates, in the plotting of data and the drawing of isolines, and in using shading as an areal symbol. The difficulties in interpreting topographical effects have been a major deterrent to their more complete application. But this problem disappears if the network is sufficiently intensive or if maps of *departures from normal* are being prepared. Maps for many different probability levels can be prepared rapidly from a single suitable data set and programme once the network problem is resolved.

The electronic computer has made grid mapping very attractive for it allows the integrated use of the water balance equation and other relationships which express landform, soils, land use, etc. for the rational synthesis of data on an orderly basis. Estimates of many of these values can be obtained, with acceptable accuracy, from conventional maps through the use of a grid overlay. Others can be synthesized as needed by using the water balance or regression procedures. Isolines of space-average values can be fitted to the grid-square averages by automated processes if desired.

The spacing or *mesh* used in the grid is of paramount importance. A wide spacing between grid points yields space averages for a very large area. Care must be taken to ensure the mesh is sufficiently fine to show the desired detail; yet compromise may be necessary if the number of data and computations are to be kept manageable. The grid should also consider the data network density; in practice the interval usually should not exceed one-half the average distance between data points. *Graphic* maps, where different degrees of shading can be produced to indicate ranges of values, are effective in showing different percentiles and other characteristics of hydrological data, and can be produced very quickly.

6.4 Hydrological and Hydrometeorological Maps

6.4.1 Maps of precipitation and evaporation

Apart from differences introduced by the physical nature of the elements, the mapping of precipitation and of evaporation are achieved in a virtually identical manner. The most common maps are those portraying annual or seasonal accumulations and their variability. It is possible to construct a great variety of precipitation maps which depict amounts for specified durations and probabilities. Isohyetal maps for different duration and probabilities, space and time characteristic maps of monthly rainfall as hyetographs, and severe storms for different duration are most common types of map. These are of great utility in engineering design. Maps of drought (and moisture excess) are also worthy of attention because of their additional importance in land-use planning.

Isoline maps are most frequently employed, and it is most important that the isolines are consistent with topography. Smoothing is desirable to the level permitted by the sampling error and as dictated by purpose and scale. Continental-scale maps are smoothed to agree with large-scale features of topography, whereas catchment area maps usually attempt to reflect most of the detail indicated on the appropriate topographic map. The fitting of isolines to grid-square values results in a smoothed map, the degree of smoothing being dependent on the size of the grid squares. Good physical reasons ought to exist to explain all wriggles and bends; isolines drawn smoothly across pronounced ridges and valleys should be treated as suspect. Isolines fitted to highly variable fields on the basis of conventional measuring network are invariably grossly over-simplified.

Much more detail is required for most purposes and, while this presents few problems over level terrain, the problem is formidable in mountainous areas. One solution to mapping in mountainous areas is the use of spatial-mean maps. Where there is a good basic network, it is possible to prepare detailed maps using elevation-dependency curves and topographic contours for interpolation. The curves must be used with caution, the relationship holding only within fairly confined regions of similar slope, aspect and climate.

There is a large variety of design and operational maps such as those of percentiles, depth-duration-frequency, and ratios which can also be shown as isolines with or without detail. Percentile maps can be readily obtained from frequency tabulations prepared by electronic computers, or by statistical procedures using the mean, the standard deviation and a frequency factor.

Estimates of short-duration extremes of precipitation are obtained by the statistical analysis of precipitation records and by the analysis and maximization of major storms which are transposable to the area of interest. Depth-duration maps are commonly used to depict the results. Their preparation requires measurements for the selected duration, but their quality can often be substantially improved by supplementing these values with estimates obtained using ratio techniques. Ratios obtained from maps can be used to estimate values for a variety of durations from 24-hour precipitation values which are the most common precipitation measurement. Where there is sufficient length of record, return-period values and quartiles may be estimated statistically using equations which express the estimate as a function of the sample mean and a product of the frequency factor and the standard deviation.

Maps of rainstorms are basic to the evaluation of floods and to the preparation of depth-area-duration curves, and generalized maps of depth-area-duration such as are used in the rational approach to design floods. The quality of the results is highly dependent on detail, particularly in the area of intense precipitation. Supporting data obtained by field surveys, radar or other sources are therefore usually a requisite to the preparation of good detailed maps.

There are three basic types of evaporation maps : those for water surfaces, potential evaporation maps and actual evaporation maps. As previously noted, their format differs little from that for precipitation maps. However, because of the limitations and paucity of evaporation measurements, estimation techniques must usually be employed.

6.4.2 Surface water maps

The principal use of surface water maps is in the evaluation of water resources and the design and operation of resource projects. In addition, the maps are frequently used to convince the public or officials, who are not hydrologists, of the value or feasibility of a project, as well as for scientific and educational purposes.

Measurements of streamflow and stage provide the basis for most analyses, but there are many other valuable data sources such as field surveys and aerial photography. Many maps are derived from other information, such as those for water surplus which are computed from climatological measurements and assumed soil characteristics. By a suitable combination of stream symbols and colours, streamflow properties such as the magnitude of the average discharge, the permanence of flow, the seasonal distribution of streamflow, the dominant source of water, duration of ice cover in cold climates and salinity can be shown in broad classes without using more map space than ordinarily required for the stream pattern.

Because of the linear nature of streamflow, point and linear symbols are the simplest methods of depiction. Point symbols retain the integrity of the initial measurements. This is not present in isolines of runoff which are arbitrarily fitted to satisfy climate, topography and superficial geology. Maps which identify the seasonal flow and its extremes at the gauging station are fundamental and of utility, not only in assessing conditions at intermediate locations but also in determining areas of data deficiency. The use of histograms which depict the volume of flow by months, of graphs showing probabilities of occurrence, or of trends in annual runoff, are ways of amplifying the characteristics of point data.

The changes in volume along a watercourse can be shown effectively on small-scale maps by flow lines whose widths are proportional to the volume. Contiguous bands or superimposed bands can be used to show individual discharge classes, including extremes; the use of different areal symbols or colours allows other characteristics of streamflow to be shown. A linear scale relating width to volume can be used where the regime is stable, but non-linear scales may be required when there are large variations.

Runoff is an integrated value representative of a contributing area and, for purposes such as understanding the water balance or the development of small water projects, there are advantages in having it expressed in an areal manner. With a large number of small catchments, it may suffice to identify the contribution of each numerically, or by the use of areal symbols. Another alternative is to fit isolines to the area in a manner which is consistent with the areal aspect of runoff, physiography and climatology.

Isoline maps usually express runoff as a depth (mm) or a yield ($l/s/km^2$). When there are many small, adjacent catchments, the runoff value is often assigned to the geometric centre of gravity of each, and isolines are fitted to the plotted data. This practice provides generalized regional patterns, but does not properly apportion surface runoff within each catchment area. Generally the measuring network does not provide the detail needed for the simple placement of isolines, and it is necessary to use rational procedures such as physical relationships linking runoff with slopes, soils, geology and climate. The use of the water balance and a grid-mapping approach is another possibility. In most approaches, the apportionment of runoff across a catchment is still largely an art which is successful in the macro-scale, but which does not necessarily stand up under detailed scrutiny. Climatic data are particularly suited as an aid in placing isolines of runoff in regions with humid climates.

Variability may be indicated by histograms and graphs constructed for each gauging station which show probability by percentiles. It may also be indicated by collocated maps which show the minimum and maximum runoff as a percentage of the long-term mean, by maps of the coefficient of variation as calculated for representative streams and by maps which show the ratio of high or low percentile values to the mean annual (monthly) flow. The areal extent of flooding is often depicted as isolines on a mosaic prepared from aerial photographs. In more formal mapping the flooded areas are identified by different types of areal symbols, shading or colour tints, according to their probability of occurrence.

Other *flood maps* in addition to those showing the areal extent are required for engineering design. The mean annual flood can be related to drainage area within homogeneous zones, thereby providing a useful tool in estimating water supplies. Similarly, zones can be defined within which the frequency characteristics of the annual floods, the project design floods and the probable maximum floods are similar.

There are many operational maps relating to time which are used as working diagrams and are seldom published. Isochronal maps are used in flood routing techniques to estimate the time of travel of storm-produced surface water to the basin outlet. They may also be used to assess meteorological processes in the construction of storm hyetographs. Maps of the time of travel of flood peaks along a river system are needed for river forecasting. The distances travelled per unit time by a flood of a specified probability are defined on the maps using line symbols.

The water budget approach based on climatic data and empirical estimates of water use by vegetation has resulted in the preparation of a large number of maps of water surplus and deficiencies for various probability levels. These maps have value in the planning of irrigation systems and in evaluating regional variations in water use by crops. In format, these maps are generally similar to the isoline maps of runoff. Water budget diagrams offset on these maps enhance their utility.

6.4.3 Maps of water in the zone of aeration

Water in the zone of aeration, popularly known as soil moisture is very important for agricultural purposes. The soil moisture plays a critical role in partitioning the precipitation into various hydrological components. Relatively few maps of water in the zone of aeration have been published. As a result of agricultural requirements, it has been possible to prepare maps of soil characteristics, thereby providing indices of water-holding capacity and infiltration rate. However, soil moisture measurements themselves are generally not suitable for the preparation of maps for large areas. Soil moisture accounting procedures have been used in the preparation of isoline maps. While these maps have value, they must generally be considered as speculative, in recognition of the errors in the measurement of precipitation and in estimating evaporation and assumptions made concerning soils.

The major complication in mapping is the present inability to extrapolate from a point observation over a large area. The establishment of a suitable measuring network is presently impracticable. A logical alternative is to employ estimation procedures which allow for differences in soils, geology, topography, vegetation and climate across the area to be mapped. In the absence of such a capability, mapping has proceeded on a local basis and has been largely confined to depiction of means, extremes and frequencies for the points of observation. A description of the soil profiles, land use, topography & climate of the area should be included on these maps.

6.4.4 Groundwater maps

It is virtually impossible to consider any component of the water balance out of the context of its physical environment. This is particularly true of groundwater. Maps depicting environments are not considered in this section, yet they are fundamental components of, or inputs into, the mapping of groundwater which can be extremely complex.

A number of very simple groundwater maps exist, such as those showing the locations and performance of domestic wells, and the changes in well levels at network stations. These require simple mapping skills, being composed mainly of point symbols and numbers. However, for scientific understanding of the quantities and processes involved, a geological approach to mapping is preferable. In many instances, the problem is three-dimensional, thereby complicating the mapping task. Since it is not possible to measure the systems adequately, their description is often highly conceptual.

Hydrogeological maps identify features such as the distribution, location, dynamics and chemical quality of groundwater, as well as characteristics of the aquifers and impermeable rock structures. To achieve this, it is necessary to depict thicknesses, contours of buried valleys, karst areas, and boundaries between fresh and saline water. Cross-sectional diagrams are a necessary adjunct to many maps because of the complexity of the systems which must be described. Availability and quality are among the first requirements for water development. Geological information can be interpreted in terms of the more limited information on underground water that is available.

A three-dimensional approach is obtained from the assembly of information from wells and from maps of surface and bedrock topography, geological maps, isopach maps, surface soil maps and groundwater flow maps (isolines of the piezometric surface, the flow being at right angles to the surface contours). The isolines of water level superimposed with the pumping and observation well locations, the flow field maps showing the velocities at nodal points by vectors, the spatial and temporal hydraulic head variation, and the aquifer response to the imposed stresses are most common groundwater maps. The use of colour is usually necessary for achieving clarity in view of the number of factors that must often be shown.

6.4.5 Water quality maps

Water quality maps are very important in understanding, and reflect the quality status of water bodies. In case of rivers and streams, a line sketch can be prepared showing different stretches of rivers and streams of a marked quality difference. The whole river/stream can be divided into different quality stretches in terms of BOD (Biological Oxygen Demand), DO (Dissolved Oxygen), or subjective categorization of usable/ portable water or in terms of degree of pollution. Different stretches can be coloured also to make more convinceable. The variations of concentration of different chemical species along the channel network can be also drawn. In case of lakes and reservoirs, the vertical distribution of concentration is more important to assess the degree of vertical mixing. In groundwaters, the contours of concentration of various chemical species are drawn in two dimensions. The concentration contours depict the different pollution zones of groundwater regimes. These zones can be coloured also for more clarity.

The water quality maps prepared at different times provide the information about the deterioration rate of groundwater quality. The superimposition of the maps of hydraulic head distribution and contaminant concentration distribution depict how the pollutants are migrating from one point to other in the aquifer domain, and determine the potential usability of groundwater reservoirs for various uses. The capture zone curves, the contaminated zones, plumes, and spatial and temporal distribution of a chemical constituent in aquifer domain are common type of water quality maps.

6.5 Thematic Maps

Thematic map is a map designed to illustrate a particular theme. The theme may be topographic pattern, landuse pattern or soil pattern. The topographic maps are prepared by the Survey of India. The landuse and soil maps are prepared using remote sensing data on a basin scale. However, these maps are also prepared for the entire country or state or district wise (on the basis of political boundary) for education puposes. Hydrological studies require these maps as a basin scale. The soil maps are also prepared by the Geological Survey of India. A reference map for the Indian Remote Sensing Satellite is available with National Remote Sensing Agency (NRSA) at the scale of 1:6000,000 to help the user to procure the desired data identified by the corresponding scene on the map from the NRSA.

Survey of India has followed certain norms to prepare toposheets on 1:50,000 and 1:250,000 scales covering the entire India. Now, Survey of India is preparing toposheets on the scale of 1:25,000 also. The toposheets on the scale of 1:250,000 represent the area covered under 1° longitude and 1° latitude. On the other hand, the area covered on the scale of 1:50,000 comes under 15′ longitude and 15′ latitude. The toposheets map are the key maps for preparing the base maps for most of the hydrological studies.

Landuse and land cover are dynamic features over space and time. Mapping of these features is an essential part for many planning and management activities concerned with the hydrological aspects of various regions. Ongoing advancement in remote sensing satellites, data is being available with very high resolution. Land use maps prepared using remote sensing data at different times enable to assess the rate of change in the land use pattern. Land use maps classify the study area in different categories such as forests, barren land, pastures, culturable area, fallows, agricultural area etc. Soil maps classify the region in different groups such as red soils, black soils, mixed soils etc. These maps are very helpful in carrying out various hydrological studies.

6.6 Atlas

Different types of Atlases are prepared by various organisations in the country. Some of them are listed below:

- National Atlas of India Vol. 1 & 2
  National Atlas & Thematic Mapping Organisation
  Deptt. of Science and Technology
  Calcutta (W.B.)

- Irrigation Atlas of India
  National Atlas & Thematic Mapping Organisation
  Deptt. of Science and Technology
  Calcutta (W.B.)

- River Basin Atlas of India
  Central Board for Prevention & Control of Water Pollution
  New Delhi

- Flood Atlas of India
  Central Water Commission, New Delhi

- Severe Rainstorms of India
  Indian Institute of Tropical Meteorology, Pune

- Probable Maximum Precipitation Atlas
  Indian Institute of Tropical Meteorology, Pune

- Atlas of Agricultural Resources of India
  National Atlas & Thematic Mapping Organisation
  Deptt. of Science and Technology
  Calcutta (W.B.)

- Climatological Atlas of India
  Indian Meteorological Deptt.

- Agroclimatic Atlas of India
  Indian Meteorological Deptt.

- Forest Atlas of India
  National Atlas & Thematic Mapping Organisation
  Deptt. of Science and Technology
  Calcutta (W.B.)

7.0 SOURCES OF ERRORS IN HYDROLOGICAL DATA

The sources of errors in the hydrological data may be classified as: (1) Instrumental errors, (2) Manual errors, (3) Generic errors, (4) Errors due to inadequate hydrological network, (5) Errors due to external influences, and (6) Sampling errors.

Instrumental errors: Instrumental errors in the hydrological data arise due to structural defects in instruments and imperfect adjustment of the instrument. After correcting the structural defects and adjusting the instrument components properly, some small errors may remain unadjusted, and these are called residual errors. The amount of error depends upon how these factors affect the measurement. Breakage in measuring jar or collector, tilt in gauge, zero reading error in point gauge, and tilt in measuring levels are some typical examples which introduce instrumental errors. Sometimes instrument fail at the time of observation, and observer could not correct because of insufficient knowledge. Therefore, there is need of educating the observer so that he can rectify the minor failure problems.

Manual errors: Manual errors may be either observational or manipulating. These are manual mistakes. Observational errors arise due to carelessness in reading the observations and carelessness in recording the observed readings. These are basically personal mistakes. The amount of error depends solely upon the observer, and the observer's sincerety in taking the measurement, his education, perceiveness and realization of the importance of hydrological data will determine the reliability of data.

Sometimes instruments may not be in proper order, the recorded observations without noticing the instrument faults will introduce errors in the measurement. Non-recording gauges do not involve manual mistakes and carelessness while recording the measurement.

Manipulating errors arise due to carelessness in transferring the data from one source to other source, feeding and storing the data in computers, rounding off the data, reporting the data in different units and converting the data from one unit to other unit. The precision of the data depends upon how much care has been taken during handling the data.

Generic errors: It is often observed that data are missing for some periods in a length of record. The data gap is mainly caused when instruments are not in order. Rectification or replacement of an instrument takes some time, and thus, during this period, data could not be recorded. The missing data are filled by using some techniques. Missing data are generated using historical data of nearby or climatologically similar stations. Truely speaking, the generated data cannot replace the observed ones because each technique has its own limitation. Thus, filling the gaps by the generated values introduce errors in the hydrological data, but this is binding to carry out hydrological studies. Such errors can be referred as generic errors. Moreover, any error in the parent measurements will propagate in the generated ones.

Errors due to inadequate hydrological network: If hydrological network in a particular drainage basin is not adequate, the measurements made at existing stations will not reflect the true characteristics of the drainage basin. Analysis based on this inadequate hydrologic data will not commensurate with the realism. There is a fundamental conflict of interest between the data gatherer and the water resource system designer and builder. The data expert wishes to improve the accuracy of his product which may require more years of observations with more accurate equipment, and more observation points and more sampling frequency. The designer and builder wish to get on with their work. Thus, special attention is needed and methodologies involve decision making process in handling inadequate hydrologic data.

Errors due to external influences: Many external factors, which may be natural, man-made, or accidental, affect the measurement, and thus, may introduce error. For example, heavy wind, and tilting of raingauge due to different settlement, subsidence or shocks will influence the rainfall measurement. Turbulent flow characteristics will affect the discharge measurements made by the conventional techniques.

Sampling errors: Hydrologic data are variant in space and time. The amount of variability is the deciding factor to judge how many times samples should be taken to capture the requisite information. Thus, sampling frequency of various types of time series of hydrologic data must be properly decided in acordance with the characteristics of hydrologic variables and the data users. Errors may be introduced during sampling either due to instrument, sampling mechanism, disturbances at the time of sampling, operator mistakes, and carelessness in handling and labelling. Fig. 9 shows a typical representation of steps and sources of error in groundwater sampling.

Step In situ condition	Sources of Error
Establishing a Sampling Point ↓	Improper well construction/placement; inappropriate materials selection
Field Measurements ↓	Instrument malfunction; Operator error
Sample Collection ↓	Sampling mechanism bias; Operator error
Sample Delivery/Transfer ↓	Sampling mechanism bias; sample exposure, degassing, oxygenation, field conditions
Field Blanks, Standards ↓	Operator error; matrix interferences
Field Determinations ↓	Instrument malfunction; Operator error; Field conditions
Preservation/Storage ↓	Matrix interferences; handling/labeling errors
Transportation	Delay; Sample loss

Fig. 9 : Steps and Sources of Error in groundwater Sampling

7.1 Errors in Rainfall Measurement

Take the example of measuring the rainfall. In the measurement, there may be instrumental errors in gauges, or in their recording or measuring arrangements. Errors may be introduced due to following factors : (1) some rain water may get lost due to slash from the collector; (2) some water from an initial rain may get lost in moistening the gauge funnel and other inside surfaces; (3) blowing winds may tilt the rains from vertical, thus bringing losses catch in the vertical gauge, (4) dent in the collector rim may change its receiving area, etc., (5) vertical upward air currents may impart upward acceleration to precipitation, thus bringing lower catch in the gauge etc. Some of the factors may increase the catch, and some of them may decrease the catch. However, in general, it can be stated that almost all the errors that are introduced in rain catch measurements have a tendency to yield measurements which are too low. Thus, the observed rain catch needs to be increased for the likely errors introduced in its measurement.

Of all the possible errors, the most serious error is introduced due to winds, which may result in a vertical acceleration of air forced upward over the gauge. Such upward air current will impart an upward acceleration to the precipitation about to enter the gauge, thus resulting in a deficient catch. The deficiency is greater for small rain drops than for large drops, and is thus greater for lighter rains than for heavy rains. It is not possible to precisely and reliably evaluate the usual errors, because of problems involved in determining the true precipitation reaching the ground. The errors introduced by wind effects increase with wind velocity; and since wind velocity is more at high altitudes, the gauge installed at higher altitude will definitely catch lesser rain catch. Higher the gauge, greater will be the wind error and hence more deficient will be the rain catch.

7.2 Errors in Discharge Measurements

Various direct and indirect methods are employed to measure the discharge flowing in the river. Direct methods which are most commonly adopted are weir station, control station and velocity area station. Discharge is also measured indirectly from power-plant records and some other techniques such as dilution and nuclear techniques. In case of weirs, the head of water over the weir-crest is measured, and the discharge is calculated by using the appropriate formula derived for that particular kind of weir. Any error in the measurement of the head introduces error in the measured discharge. For example, error may be manual due to observer's carelessness in registering or reading, or instrumental due to zero reading error in the point gauge.

In velocity area station popularly known as rating curve method, due to the presence of a control across a river, the water on the upstream side gets affected. Natural undulations in the river bed may affect the water surface, and hence may act as a control. The gauge radings are required to be corrected before plotting them against discharges to obtain the rating curve (Fig. 10). However, manual mistakes and carelessness are not encountered in the non-recording gauges, because measurements have not been read by an observer. After installing the gauge, its datum must be checked annually, and the gauge must be reset as soon as a change in its datum is detected. Unless this precaution is taken, long periods of records are likely to be doubtful, and it is not possible to correct them afterwards. Two corrections are employed while using the rating curve method to estimate the discharge. These corrections are needed because of the effects of changing stage and shifting control. These are illustrated below :

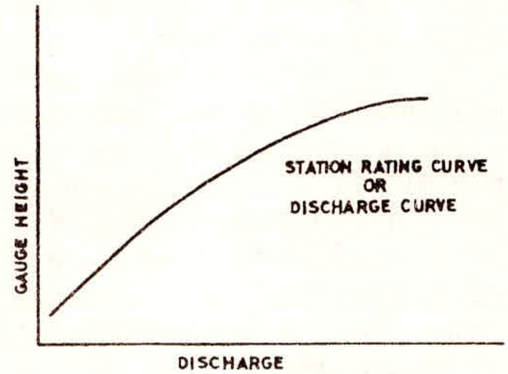

Fig. 10 : Typical Representation of Rating Curve

(1) Effect of changing stage on the stage-discharge curve:

The discharge measured for a particular given stage is different for a changing stage than for a constant stage. The stage of a river may be constant, rising or falling. During the rising stage of the river, the measured discharge is more than that for a constant stage (Fig. 11). Similarly, during a falling stage of the river, the measured discharge is less than that for constant stage.

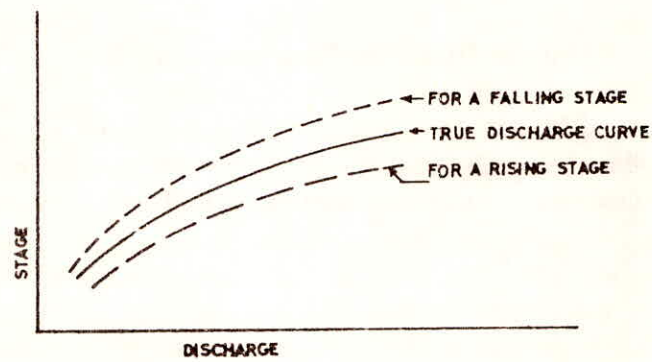

Fig. 11 : Rating Curve for Different Stages

While plotting a stage-discharge curve, the values of discharges must be taken for a constant stage. But this is not the practical case, and the values of discharges are generally measured either during a rising stage or during a falling stage. Hence to obtain a true discharge curve, the measured discharges must be corrected for the changing stage. The amount of deviation from the true discharge curve does not depend upon the total rise or fall during a particular set of measurement, but depends only on the rate at which this change occurs.

(2) Effect of shifting control on stage-discharge curve:

A control does not always behave as a permanent control, but changes its position many times. Due to this shift or a change in the control, the observed discharge measurements are not the true values, and will not fall on a true discharge curve. If the provision for a changing stage is also made, the observed points should fall on the right side of the true discharge curve for a rising stage and the point should fall on the left side of the true discharge curve for a falling stage.

If a control is found to be changing, the next step is to find out the speed with which the control is changing. If the control is changing very slowly or only at the time of floods, then the best method is to draw a new rating curve as soon as sufficient change in the control is detected. Then this new curve is applied for a period till sufficient deviation due to the change in control is detected. On an average, a control changes sufficiently once a year. In other words, the rating curve has to be changed almost once a year. On many other rivers, the control changes quite rapidly and constantly. In such cases, Stout method is generally used in order to correct the observed measurement.

8.0 KINDS OF ERRORS

Errors of measurement are of three kinds : (i) mistakes (ii) systematic errors and (iii) accidental errors. Mistakes are errors that arise from inattention, inexperience, carelessness and poor judgement or confusion in the mind of the observer. If a mistake is undetected, it produces a serious effect on the final result.

A systematic error is an error that under the same conditions will always be of the same size and sign. Systematic errors follow some definite mathematical or physical law and a correction can be determined and applied. Such errors are of constant character and are regarded as positive or negative according as they make the result too great or too small. There effect is, therefore, cumulative. By cumulative errors, we mean that these occur in the same direction and tend to add up or accumulate. While the compensating errors are those which are liable to occur in either direction, and hence tend to compensate. If undetected, systematic errors are very serious.

Accidental errors are those which remain after eliminating the mistakes and systematic errors, and are caused by a combination of reasons beyond the ability of the observer to control. They tend sometimes in one direction and sometimes in the other. Accidental error represents the limit of precision in the determination of a value. Accidental errors obey the laws of chance and, therefore, must be handled according to the mathematical laws of probability.

9.0 CONCLUDING REMARKS

The hydrologic data is a key point for all kinds of hydrological studies. The efficient planning, development and management of water resources projects depend upon the adequate and reliable data. Hydrological data are variant in space and time. Various hydrological data must be recorded with appropriate sampling frequency. Modern use of computers in simulation studies necessitate the data in an efficiently designed format. Keeping the importance of hydrologic data in mind, it is necessary to acquire sufficient knowledge about the general and specific requirement of data for hydrological studies, and their characteristics in terms of its variability in space and time, and storing in and retrieving from the computer-compatible formats. Time-series and non-time-series data must be recorded and stored in accordance with the data characteristics and data users.

Maps and Atlas provide an excellent means of summarizing the large hydrological information in an easily understandable manner. They provide an efficient and effective way of communicating information and ideas for the purposes of education, design, planning or convincing political, social and economic decision makers. Maps have also become more purposeful, there being increasing trends to serve operational and developmental needs in addition to filling the traditional educational and scientific requirements. Their preparation by computer has resulted in a large variety of maps of varied duration and probability which are quickly available for consideration in complex planning problems. Furthermore, the improved quality and availability of data and maps will ensure more widespread use and better understanding of hydrological information in the resolution of important social, scientific and economic problems.

The required level of accuracy in the various hydrological studies rely on the reliability and adequacy of data. The raw data are first processed before using the input to a mathematical model. Different sources of errors in hydrological data and characteristics of errors should be known while taking up the data processing job. This helps in updating and validating the observed hydrological records, and enable to obtain realistic and meaningful solutions of hydrological problems.

Problem 1:

A rectangular notch is installed to measure the discharge. If the error involved in the measurement of head over the crest of the notch is 2%, estimate the error made in the discharge measurement.

Problem 2:

If the discharge is measured using a triangular notch instead of a rectangular notch in Problem 1, what will be the percentage error involved in the discharge measurement.

Problem 3:

A sharp-crested weir of height 0.80 m and length 2.0 m was fitted with a point gauge for recording the head of flow. After some use, the point gauge was found to have a zero error; it was reading heading 2 cm too small. Determine the percentage error in the estimated discharges corresponding to an observed head of 50 cm.

BIBLIOGRAPHY

WMO-634, 1985. *Guidelines for computerized data processing in operational hydrology and land and water management.* WMO No.- 634, Joint FAO/WMO Publication, Geneva.

Rodda, J.C., 1976. *Facets of Hydrology.* Edited by J.C. Rodda, John Wiley & Sons, New York.

Linsely, R.K., Franzini, J.B., Freyberg, D.L. and Tchobanoglous, G., 1992. *Water Resources Engineering.* McGraw Hill Inc., New York.

***