SR-25

आपो हि ष्ठा मयोभुवः

UNDP PROJECT

DEVELOPING CAPABILITIES FOR HYDROLOGICAL RESEARCH

IND/90/03

HYDROLOGICAL DATA PROCESSING AND ANALYSIS

Y R Satyaji Rao

NATIONAL INSTITUTE OF HYDROLOGY
JALVIGYAN BHAWAN
ROORKEE-247 667 (INDIA)

CONTENTS

PREFACE

Advances in scientific and Engineering Hydrology dependend on good, reliable and continuous recording, processing and analysis of Hydrological variables. The successful and efficient execution of modern hydrological studies depend on a vast and diverse amount of information in the form of data base system. In India, with the improvement of modern needs and the use of hydrologic events data process, storage and analysis are of utmost importance. It is also very important aspect in the Hydrologic modeling studies.

The report entitled 'HYDROLOGICAL DATA PROCESSING AND ANALYSIS' prepared by sh. Y R Satyaji Rao, Scientist'B' Deltaic Regional centre of NIH, Kakinada under the guidance of Dr. P V Seetapathi, Scientist'F' in connection with the ongoing UNDP project.The report presents the concepts of the data processing, processing methods of Hydrological variables, and status of Data processing and Analysis in India and Eleswhere.

Satish Chandra
DIRECTOR

LIST OF FIGURES

LIST OF TABLES

1.INTRODUCTION

Hydrologic data hold the key to the orderly and efficient development and control of water resources. Advances in scientific Hydrology and Engineering Hydrology are dependent on good, reliable and continuous recording ,processing and analysis of Hydrologic variables. The present Hydrologic data collection is designed in a way that each of the Hydrologic components is dealt with separately by each individual organisations. There is no single body within India that is responsible for all hydrological services. Though data collected by any one organisation can be utilised by others, unfortunately in the present system the same data is collected by different agencies independently. Even the collected data by these agencies suffer from the following two major drawbacks. i). The data collected by any single organisation independentely is observed to be inadequate to define the entire water system. ii). The data which is not really required or data pertaining to unrelated elements are frequently observed and collected.

Once data is collected, the next important and necessary step is their proper storage. In the conventional ways, the data are mostely kept in manuscript form in registers or files. Usually no processing is done before the data is stored in the registers/files. With the advancement of Technology, computer memories, magnetic tapes,microfilms etc., these conventional storage media are being replaced. In the modern procedures, the data are stored only after a particular level of processing. Most countries have well established and expanded hydrological networks, producing correspondingly increasing quantity of data. This results in manual processing techniques becoming totally obsolete , except for limited individual studies.

The term 'Data Processing' covers data preparation, data entry and transfer to the data base, data validation, data correction, filling -in- of missing data, data compilation and analysis, data retrieval and data dessimination/publication of yearbooks. Basically, the aim of data processing is to manuplate the raw data and to put it in a proper form and extract the required information from it. The processing activities are divided into primary and secondary processes. The primary processing consists of the data validation i.e., removing the random and systamatic errors if any from the raw data and putting it in a proper format. The secondary processing includes deriving some information from the data such as fixing of rating curves using gauge-discharge data, Flood forecasting using precipitation data etc.

In recent years sophisticated instruments and high speed computers are being developed and they can now be employed to measure delicate hydrologic phenomenon and to solve complicated mathamatical equations involved in the application of hydrologic theories. It has been noted that by many authors that hydrologic education has become extremely important for training, manpower and research and planning for developing and controling water resources.

1.1 Scope of the Report:

The present report attempts to bring out the status of Hydrologic data processing and analysis in India and elsewhere. The concepts of data processing, available computer systems, machine methods, software packages, Hydrological data banks etc. have been discussed in this report. Importance of Data processing in the field of Hydrology in developing countries like India is also highlighted.

The report has been divided into ten chapters. The basic concepts of data processing are discussed in second chapter. The third chapter deals with various hydrological variable processing and analysis like climatological data, Rainfall data, Surface water data, Ground water data, Snow and ice, Sediment and water quality. The processing of Remotely sensed data have been discussed in chapter four. Data processing, with machine methods especially for climatological data have been discussed in detail in chapter five.Chapter six explains about the various Indian organisations involved in data collection and processing. The seventh chapter deals with Hydrological data banks and processing elsewhere. A typical Hydrological data processing software HYMOS has been discussed in chapter eight, and finally conclusions, purpose, and scope of the study are in chapters nine and ten respectively.

2.0 CONCEPTS OF DATA PROCESSING

Quantative scientific data may be classified into two kinds i.e i. Experimental data and Historical data. The experimental data are measured through experiments and usually can be obtained repeatedly by experiments. The historical data on the other hand are collected from natural phenomena that can be derived only once and then will not occur again. Most hydrologic data are historical data. The hydrologic data are the only source of information upon which quatitative hydrologic investigations are genarally based. Further more natural hydrologic phenomenon are highly erratic and commonly stochastic in nature,and therefore the statistical interpretation, probability analysis and data processing are essential.

2.1 Data Processing Cycle:

Data processing is the manipulation of data into more useful form and it includes not only numerical calculations but also operations such as classification of data and the transmission of data from one place to another. In general, these operations are performed by some type of machine or computer, although some of them could also be carried out manually.

The component of data collection includes capturing of data and the transmission of data. Data processing covers all the activities after transmission up to and including data dissemination. Basically data processing procedure consists of the following steps.

Origination---->Input--->Processing---->Output------>Distribution

Storage

2.1.1 Origination:

This step refers to the process of collecting the original data and is called the source document.

2.1.2 Input:

In this the initial data, or input data, are prepared in some convenient form for processing. The form will depend on the processing machine.

2.1.3 Processing:

In this step the input data are manipulated and usuall combined with other information to produce data in a more usefu form.

2.1.4 Storage:

This step is crucial in many data processing procedures. Data processing results are frequently placed in storage to be used as input data for further processing or use at a later date.

2.1.5 Output:

In this the results of the preceding processing steps are collected. The particular form of the output data depends on the use of the data.

2.1.6 Distribution:

This step refers to the distribution of the output data. Recording of the output data is offen called report documents.

2.2 Data Processing Operations:

Data processing procedure normally consists of a number of basic processing operations performed in some order. They are as follows :

2.2.1 Recording:

Recording refers to the transfer of data into some form or document. It occures not only during the origination step and during the distribution step, but throughout the processing cycle.

2.2.2 Duplicating:

This operation consists in reproducing the data into many forms or documents. Duplicating may be done while the data are being recorded manually, or it may be done afterwards, by some machine. On the other hand, one may record a sales transaction by punching the data onto a card, and then duplicate the card using a duplicating machine.

2.2.3 Verifying:

Since recording is usually a manual operation, it is important that recorded data be carefully checked for any errors. This operation is called verifying.

2.2.4 Classifying:

This operation separates data into various categories Classifying can usually be done in more than one way.

2.2.5 Sorting:

Arranging data in a specific order is called sorting.

2.2.6 Merging:

This operation takes two or more sets of data, all sets having been sorted by the same key, and puts them together to form a single sorted set of data. When one deck is empty, the cards in the other deck are put at the end of combined deck.

2.2.7 Calculating:

This is performing numerical calculations with the data.

2.3 Data Base Management System

Throughout the world an immense amount and variety of meteorological/Hydrological data have been accumulated over the years. However, no country need to process all the world's data. At most a selection, largely consisting of observations produced within it's own frontiers, is required and aftedr a time some condensation even of thhis may be possible. For the smallest countries 'hand and eye' methods of processing may still suffice. Satisfactory solultions to many water resources problems require readily accessible, reliable observational data on the elements of the hydrologic cycle and related factors. The efficient data base management system can be help in improveing the following aspects.

(a) Establishment and supervision of hydrological networks
(b) Collection, processing and publication of basic data
(c) preparation of reports on water resources
(d) Preparation and dissemination of hydrological forecasts
(e) Analysis and design studies
(f) Research and development
(g) Training

3.0 HYDROLOGICAL DATA PROCESSING AND ANALYSIS

The following procedures have to fallow all Hydrologic data variables in it's collection and processing.

Any Hydrologic data is stored on computer capatible devices for computer processing, it becomes necessary to carry out preliminary checks, manual scrutity etc. The reports received from manually observed stations by telephone or other communication channels are checked by a repeat back system. Im proper registering of data includes entering data against wrong time and date, alteration of figures etc. The official at receiving station could check the reasonableness of report by judging the report, based on past experience and statistics of the station and region to which the station belongs.

A distinction should be made between data which is given and information which is derived from data by logical processes (Gilliland, 1972). The information output may be the ultimate requirement or it may be used as data input to a higher level system.

Method 1

```
-----------------------------------------------------------------
Data            Data           Data            Data         Information
Aquisition --> Storage --> Relocation --> Processing--> Display
                                            information
-----------------------------------------------------------------
```

Method 2

```
-----------------------------------------------------------------
Data            Data          Information    Information   Information
Aqusition--> processing--> Storage   -->   Relocation --> Display
             Information
-----------------------------------------------------------------
```

In general the first method of processing data is preferred because in this system the original data is stored and can therefore be recovered. In the second system only processed information is stored and the original data are not recoverable. However, sôme time the quantity of data is so large that some processing is necessary to reduce the quantity stored. This processing is called data compaction.

The following errors may occur in data collection

(a) Error built into instruments

(b) Errors involved in reading instruments and transmitting or recording data

(c) Error due to improper instruments exposure or to the lack of representativeness of the instruments site to the area for which it is to be used as an index.

(d) Error occuring during the processing of data.

Most of the errors described above could be further sub classified as systamatic errors and randam errors.

3.1 CLIMATOLOOGICAL DATA:

A large number of climatological data generally collected are:

(1) Daily evaporation from Class-A Pan or from Piche's Evaporimeter.

(2) Wet and dry bulb temperatures at selected hours.

(3) Soil temperature at selected hours at various depths from ground level.

(4) Daily wind run from anemometers and various direction

(5) Short wave radiation

(6) Sunshine duration from sunshine records

(7) Air pressure

(8) Relative humidity

3.1.1 Data Validation :

It is noted that validation of climatological data by method of comparison may not be possible, because of the scarcity of the climatological station network. Thus the basic validation techniques applied are range checks, rate of change check and, of particular importance, consistency checks between related parameters observed at the same site i.e. using plotting techniques. For all climatological data, station & parameter codes should be tested for validity . where relevant, sensor calibration values and ranges should be output with suspect values.

3.1.2 Data Processing :

First step in processing of climatological data is to derive average values usually on a daily basis, for temperature, relative humidity, vapor pressure and for some indirect evaporation techniques and the slope of the vapor pressure curve. An important derived factor is the degree day index, the accumulated departure of temperature from a standard reference temperature. The psychrometer constant for the station should be stored in the station description file. Average daily wind speed must also be calculated, or wind run obstructed from integrated totals. These are several climatological parameters which need to be transformed to standard condition for storage and/or application.

Where direct measurement techniques are used, the computer may be used to verify evaporation estimates by checking the water levels (or lysimeter weights), and the water additions/subtractions recorded. Estimate of evaporation (and evapotranspiation) should be made over time intervals of sufficient length to minimize heat flux errors.

3.2 RAINFALL DATA :

It is common experience that the precipitation data in it's raw form would contain many gaps and inconsistant values. As such preliminary processing of precipitation data is essential before it is put to further use in analysis. The methodology for executing the various steps involved in the processing system are briefly described.

3.2.1 Estimation of missing data:

While retrieving data for climatological purpose or inputing data in real time, one often comes across missing data situations. Since blank in a data set is read as zero by computer, necessary software for identifying the blanks and marking them appropriately need to be developed.

Data for the period of missing rainfall data could be filled using estimation technique. The length of period upto which the data could be filled is dependent on individual Judgement. Rainfall for the missing period is estimated either by using the normal ratio method or the distance power method.

Normal ratio method:

In this method the rainfall at a station is estimated as a function of the normal monthly or annual rainfall of the station under question and those of the neighbouring stations for the period of missing data at the station under question.

Distance power method:

In this method, the rainfall at a station is estimated as a weighted average of observed rainfall at the neighbouring stations. The weights are taken as the reciprocal of the distance or some power of distance of the estimator stations.

3.2.2 Internal consistency check:

The internal consistency or self consistency checks are applied by using statistical information based on historical data of the station and current data in case of short duration rainfall.

3.2.3 Spatial consistency check:

Spatial consistency checks for precipitation data are carried out by relating the observations from surrounding stations. This is achieved by interpolating the rainfall at the station under question with rainfall data of neighbouring stations.

3.2.4 Adjustment of data:

To obtain homogeneity among and within measurments of precipitation, adjustment of data becomes necessary. It has two principal objectives. First is to make the record homogeneous with a given environment and the second is to eliminate or reduce extraneous influences by correcting for change in gauge location or exposure. Adjustment for these errors is made by 'Double Mass Analysis.

3.2.4.1 Double Mass Analysis:

It is a graphical method for identifying and adjusting inconsistencies in a station data by comparing with the trend of reference stations data. In a double mass curve, the accumulated seasonal or annual precipitation values of reference stations is taken as abscissa and those of the station under test as ordinate.

A change in the regime of the raingauge such as change in exposure, change in location, etc. is revealed by a change in the slope of the straight line fit. The other records are adjusted by multiplying the precipitation values by the ratio of the slope of the later period to the slope of the earlier period.The Fig.no 1 shows the method of Double Mass Analysis.

3.2.5 Data Computation:

Analysis of rainfall data includes the determination of

(i) For daily series: Daily maximum and minimum values total per month and year the date of occurence of the extremes as well as the number of wet or dry days.

(ii) For monthly and annual series 'basic' statistics extremes and fractiles.

Apart from analysis of point rainfall the aerial rainfall is also investigated and computed by catchment or sub catchment wise. Numerous methods of computing arial rainfall from point rainfall have been proposed. The most commonly used methods are

(i) Arithmatic mean method
(ii) Thiessen polygon method
(iii) Isohyetal method
(iv) Quadrant characteristics method

For studies of short duration events the regular time series must be produced from recording raingauge data. It is obvious that the time interval selected for producing this time series should be compatible with the duration of interest.

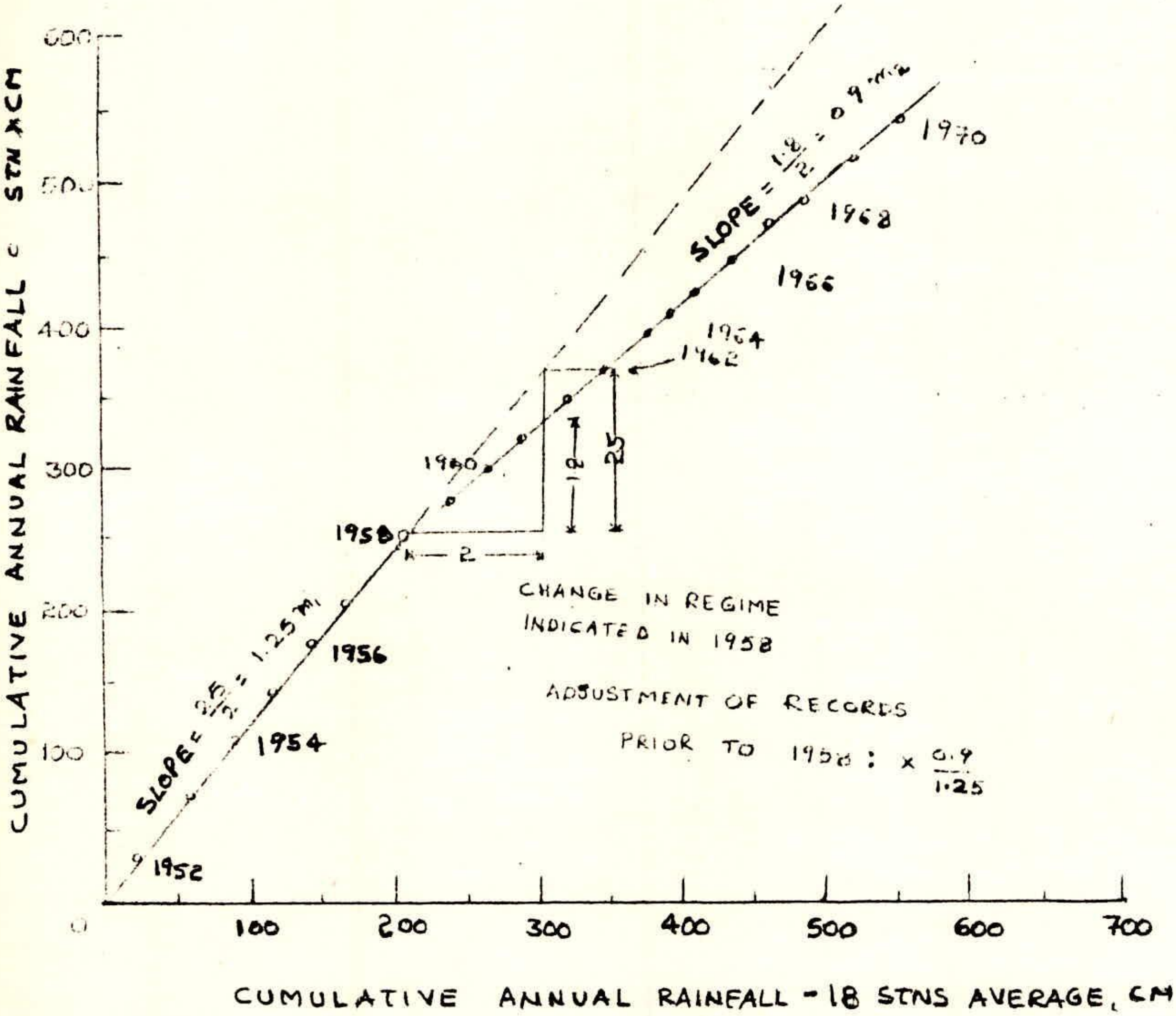


Fig. 1

DOUBLE MASS ANALYSIS

3.3 SURFACE WATER DATA :

The components of surface water data collection are given in Fig. no 2 . Although there are parallel needs for data on both quantity and quality in all components of surface water monitering. The primary differentiation of data is made between flowing and non flowing bodies of water. The reason for distinction lies in the processing requirments. For non flowing bodies of water it is generally the level (for some simple level volume relationship) which is of direct intrest. For flowing water there is the additional need to compute flow rates, the form in which the data will most generally be used.

3.3.1 Processing of river flow:

There are four components of the processing required to produce river flows which are ;

i) River gauging data
ii) Development of rating curves
iii) Flow computation
iv) Routine post computation tasks

i) River gauging data :

The river gauging data directly influence the river flow computation. It is, therefore, recommended that flow gauging data are submitted for computer varification. The Fig.no 3 shows the some of Time series plots for checking river flow data.

ii) Development of rating curves :

Rating curves define the relationship between river level (stage) and flow. This relationship is determined by performing many river gaugings over wide range of flows and using the level/flow points to define a continuous rating curve. while gauging structure have standared theoritical ratings, it is good

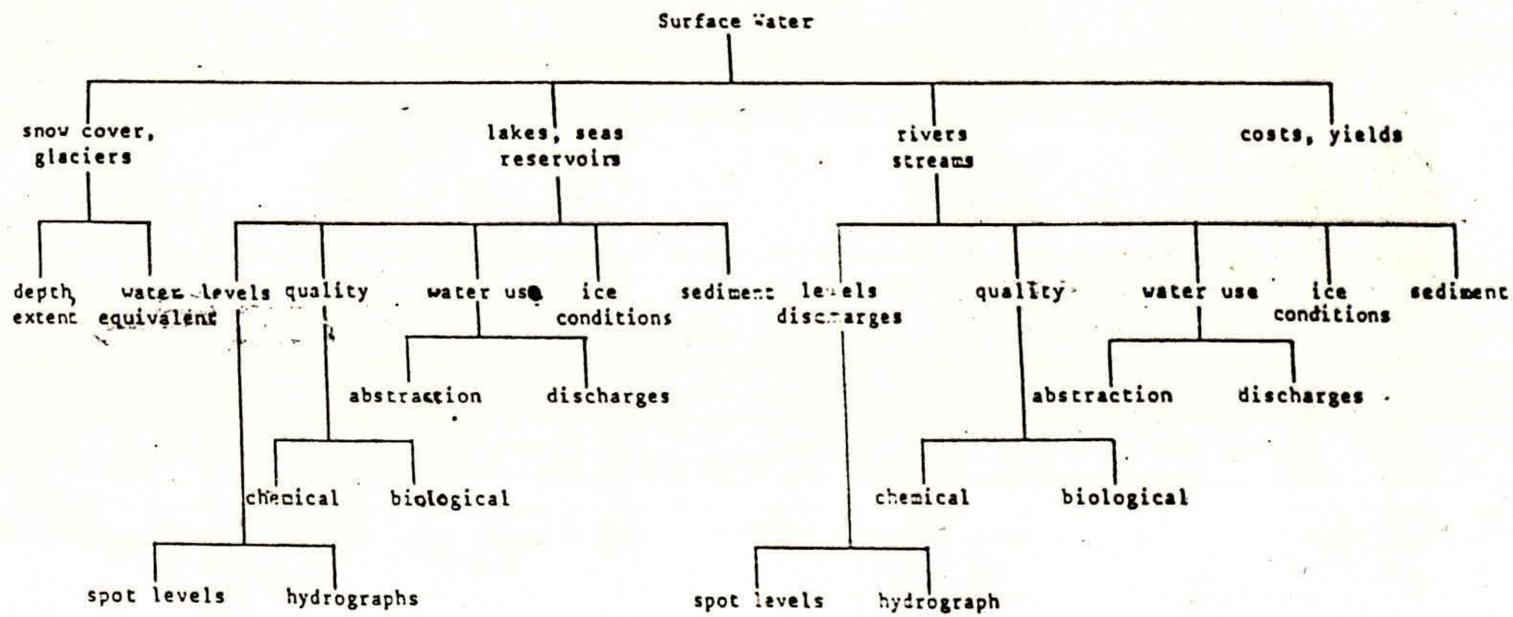


Figure 2 - The components of surface water data

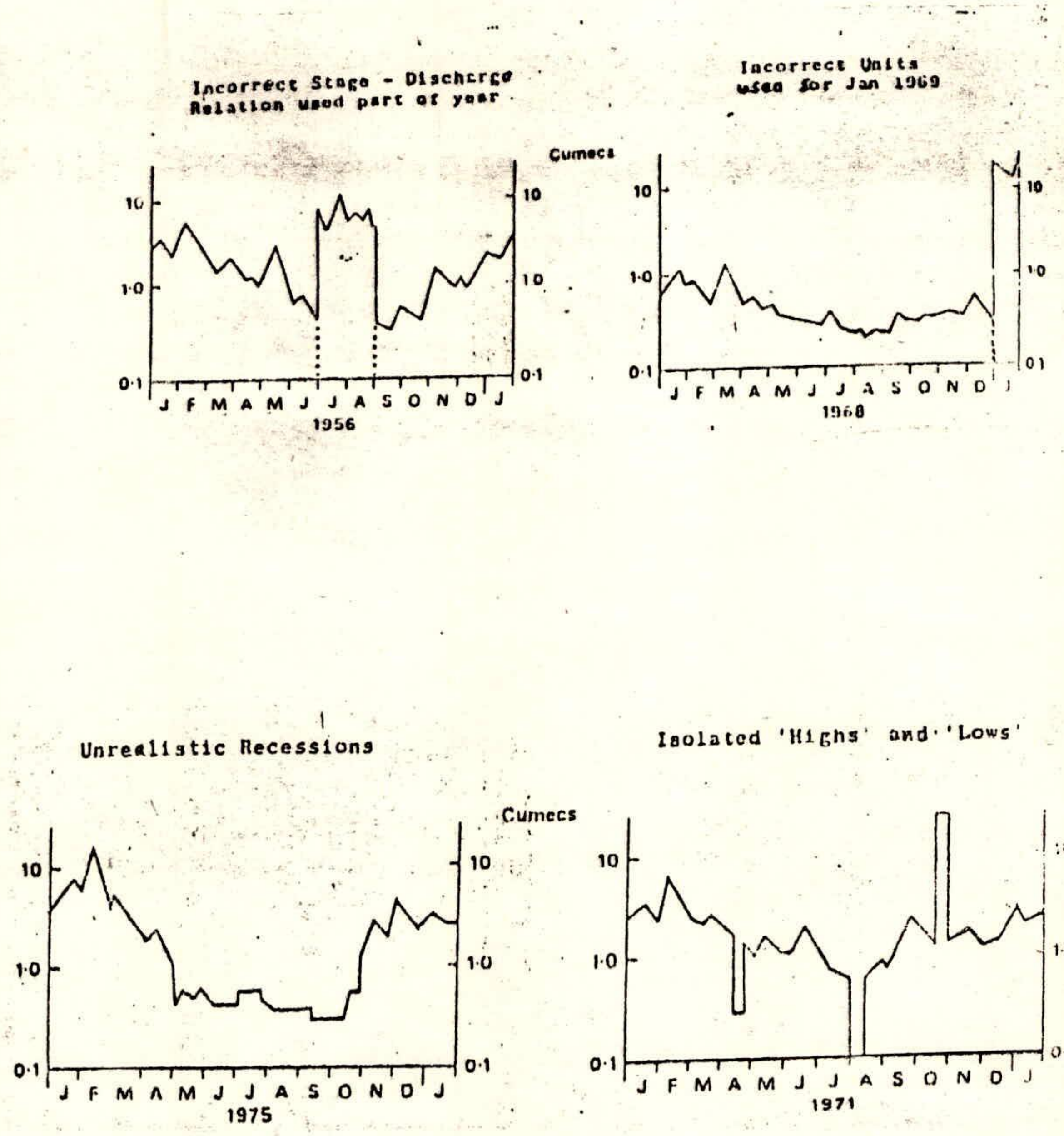


FIG. NO 3 TIME SERIES PLOTS FOR CHECKING RIVER FLOW DATA.

practice to rate structures in the feild.

The functional form of the rating curves has one of the following origins:

a) A theoretical (or modified) equation for a gauging structure.
b) A function fitted by the computer to the gauge points i.e the automation of the manual curve fitting processes.
c) A function fitted to the points of the table prepared i.e a smoothing of the manually fitted curve.

iii) Flow Computation :

In order to evaluate flows the following sets must be held by the computer:

__ A quality controlled set of water levels i.e those which have been corrected for datum, gauge height, and timing errors and having subsequently been validated. If slope methods of flow computation are used, two sets of water levels will be required.

__ Rating curve(s) corresponding to the time and range of the water levels, Where rating curves are related to frequently changing artificial controls, e.g gates and sluices, a time series of control settings may be input to guide computer selection of the relavent rating curve.

__ Any shift corrections which need to be applied to the water levels or not. This requires the amount and duration of the shift to be specified.

Having verified that all the necessary data sets are available, flow computation may proceed in the following steps:

1. Apply shift corrections to the water level . At very low flows it is possible for negative shifts to produce a modified water level below zero datum. Such an eventuality should be detected and reported.

2. Check that the rating curve being used is still appilicable to the time of the water level being processed. If the time validity of the rating curve has passed, identify and locate the new curve. Failure to locate the appropriate curve should be reported.

3. Check that the water level being processed lies within the valid water level range of the rating curves. If the level lies outside the valid range, the processing system must know whether extrapolation of the rating curve is allowed, and if so, to what extent. If extrapolation is not allowed or if the water level lies outside the allowable extrapolation range, as out of range error should be reported.

4. Apply the rating curve to the water level and obtain the corresponding flow value.

5. Return to stop 1 until all water levels have been processed.

6. Aggregate the flow time series to obtain the standard time unit (normally one day) average flows.

iv) Routine post-computation tasks:

Some attempt is normally made to fill data gaps by cross-correlation with nearby gauging stations, particularly those in the same river system. In the absence of reliable cross correlation relationships, raianfall-runoff models including the use of conceptual catchment models may be used. All estimated data should be suitably flagged.

Many river systems are significantly affected by man's activities and these effects tend to change (increase) with time. For hydrological and water resources studies it is frequently necessary to try and obtain a stationary time series. This process is known as naturallizing river flows, and requires extensive background information on all forms of direct and indirect abstractions, discharges and impoundments in the catchment. These water use effects may be aggregated into a single time series of net modifications to river flow. When these corrections are applied to the measured river flow a naturalized time series is obtained.

A further routine need is to investigate the frequency distribution of flows, and update existing probability based relationships.

A significant task for all data-processing, but particuarly relevant for flow systems, is to perform the necessary house keeping operations on data sets.

3.4 GROUNDWATER DATA:

Groundwater is not an isolated resource because, except in special cases, precipitation and surface water of an area are part of the boundary conditons for the groundwater body, they are part of the input output of a complex system. Groundwater basin geology, basin precipitation, surface streams are either inherent properties or boundary conditions of the groundwater body. As such, these data may have equal or greater rank than data derived from water level measurements.

Groundwater data collected and analysed fall into three broad categories 1) data needed to describe the environment, 2) data needed for analysis to formulate concepts about the flow system and 3) data needed to define the mass transport phenomenon. Environmental data includes relatively unchanging parameters, such as those defining the geometry of the aquifer system, its storage and transmission properties, the physical and chemical nature of water prior to stress. Data needed to define the flow system include changeable parameters, such as those defining hydraulic boundary condition, hydraulic head distribution, water input and output of the system and response of the system to physical and chemical changes. Data needed to define mass transport are much more difficult to obtain. Data for environmental and hydrologic parameter described above must be more accurately known, and dispersion, absorption and diffusion coefficients, iron exchange capacity are involved.

Data compression and compaction :

Data compression is concerned with reducing data volume without rejecting data, whereas data compaction involves discarding unwanted data. Data compaction is a function of the objective. Data compression is not a function of the objective but rather a technique for improving the efficiency of the data aquisition, storage, and relocation stage.

It has been proposed by Gilliland to record the time at which a given water level change occurs rather than to record the water level at given time interval. This method can lead to reduction in the size of the record of the order of several hundred with no loss of accuracy and insignificant record loss.

3.5 SNOW AND ICE:

Data relating to snow and ice conditions are an important element of hydrological data-processing systems particularly for snow fed basins. Data are collected on the extent, depth and water equivalent of snow and ice by a wide variety of technique, including multi-sampling along snow courses. Satellite imagery is increasingly used for estimation of snow extent,and even depth/water content in some countries. However,the constraints of complex digital image processing place automated processing outside the normal range of most hydrological Electonic Data Processing systems.

Even if the water equivalent of falling snow caught in raingauges may be validated along with rainfall data, other snow and ice parameters are more difficult to handle.

Data on the extent of snow cover may only be validated by a time consuming manual synthesis of field obervations, aerial survey data and satellite imagery. Techniques to perform automated interpretation of satellite imagery for snow extent are being developed. But there are still problems of differentiating between snow and cloud cover, and of insufficient image resolution.

Data on snow depth and water equivalent again demand much manual validation and verification, integrating from snow courses, snow gauges and conventional precipitation gauges. The large spatial variation in snow cover makes inter station comparisons difficullt. However, there are techniques to estimate the statistical reliability of snow course observations. Under conditions of melting snow, degree day factors are widely used for correlation purposes and, where snowmelt represents a significant proportion of river flow, established relationships between runoff and snow water equivalents may be used. Air and water temperature relationships are valuable not only for the computaton of degree day factors, but are also used for assisting in the validation of ice cover and thickness data, and in the forecasting of ice formation and break-up data.

3.6 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.

Many of the cross-section plotting techniques described for river gauging are directly relevant to suspended sediment data validation. Indeed, the two sets of data may be validated together if both sediment and velocity observations were made.

As with water quality data mass balance calculations may be performed if sufficient quantities of data exist. A useful check where catchments are reservoired is to test that river sediment loads within a resonable distance downstream of the reservoir are less than those upstream. If a sediment rating curve exists for the section sampled, the departure of the sampled value from the curve may be estimated for its statistical significance and/or plotted for manual scrutiny.

3.7 WATER QUALITY DATA:

The usefullness of supply of water depends on it's chemical quality. The number of observation sites should vary with the variability of water quality and should increse with an increase in concentration. In humid regions where concentrations of dissolved mater are low, fewer observations are needed than in dry climates, where concentrations, particularly of critical ions, such as sodium may be high.

Validation of procedures for water quality data are obsolute check of analysis and physical/chemical checks of determined relationships. If some parameter values have been determined in the laboratory and all the relevant associated data are available to the computer, they may be re-computed for verification purposes.

The following setpes are involved in primary processing:

(i) verification of laboratory derived values

(ii) conversion of measurement units and adjustment of values to some standard reference scale.

(iii) Computation of water quality indices

(iv) Mass balance calculations.

The operations necessarily comprise the conversion of measurements units, chemical units and correction of values to match a reference standard. The standadisation of units is import in obtaining consistency of values stored in the data base.

Water quality indices are generally based on emperical relationships which attempts to classify pertinent characteristics of water quality for a specified purpose. Mass balance calculations are performed to monitor pollution control and as a further check on the reliability of water quality data.

Ex:

1. **Total Hardness;**

Total Hardness results from the presence of cations, of which calcium and magnesium are the most abundent in groundwater. Total hardness is custorily expressed as equivalent of calcium carbonate(caco3) thus

$$TH = ca * caco3/ca + Mg*caco3/Mg$$

where total hardness calcium and magnesium are expressed in milligram per litre (ppm) and ratios in equivalent weights, the above equation is reduced to

$$TH = 2.497\ ca + 4.115\ Mg$$

2. Typing of water samples using piper's Trilinear Diagram:

Trilinear diagrams were used for analysing the water types. It consists of two triangles, one on the left and on the right side and central diamond shaped projection space.

Here, cations (in epm) expressed as percentages of total cations are plotted as a single point on the left triangle, while anions similarly expressed as percentages of total anions appear as a point in the right triangle. Those two points are then projected upwards parallel to the sides of the triangles to give a point in the diamond. The water quality types can be identified quickly by the lacations of the point in the different zones of the diamond. The Piper's Trilinear Diagram as shown in Fig no. 4

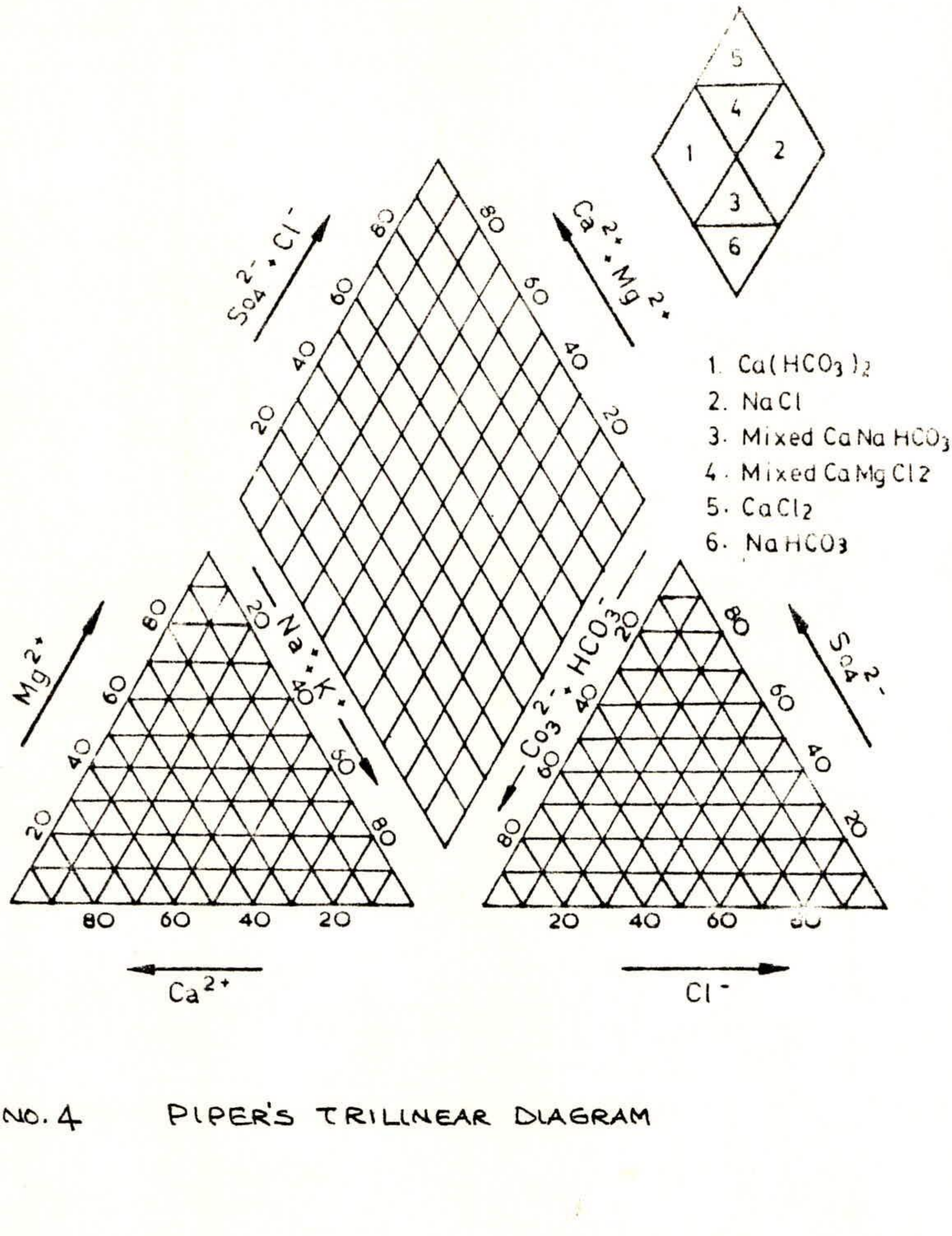


FIG NO. 4 PIPER'S TRILINEAR DIAGRAM

4.0 PROCESSING OF REMOTE SENSING DATA

Remote sensing is the art and science of obtaining information about an object, area of phenomenan through the analysis of data aquired by a device that is not in contact with the object, area of phenomenan under investigations. The term remote sensing is commonly restricted to methods that employ electromagnetic energy as the means of detecting and measuring target characteristics. This definition of remote sensing exceeds electrical, magnetic and gravity surveys that measure force fields rather than electromagnetic radiation. Magnetic and radioactivity surveys are frequently made from aircraft but are considered air borne geophysical surveys rather than remote sensing .

4.1 Remotely Sensed Data:

The two basic processes involved in electromagnetic remote sensing of earth resources are data acquisition and data analysis. The elements of the data acquisition process are, energy sources, propagation of energy through the atmosphere, energy interaction with earth's surface features, air borne or space borne sensors, resulting in the generation of sensor data in pictorial (image) and numerical form.

The image oriented system is one of the oldest and very popular. Common examples of image oriented data products are black and white pancromatic, black and white infrared and colour infrared photographs. Data can also be obtained at various resolutions or scales depending on the purpose. The photographs obtained by this technique are interpreted visually using photo interpretation techniques.

Numerical oriented system is relatively new but it is very fast. Multispectral scanners are often used as sensors to record the data. The data obtained through scanner are stored on magnetic tapes in digital form. The multispectral scanner are capable of collecting spectral data in narrow bands of various region of electromagnetic spectrum and the data can be analysed rapidly with aid of computer system. This information is then presented generally in the form of maps, tables and written discussion report.

4.2 Digital Image Processing:

Digital image processing is an extremly broad subject and it often involves procedures which can be mathematically complex. The computer is programmed to insert these data into an equation or series of equations and then store the results of the computation for each pixel.

4.2.1 Pre-processing:

The data received from the satellite is distorted due to the atmosphere effects, Geometric errors due to the satellite platform mis-orientation, scanning mirror non-linerity, earth rotation and curvature effect, radiometric anomalies of different detectors etc. Now some of pre-processing procedures will be discussed.

4.2.2 Geometric corrections:

Earth rotation effect, because the earth rotates West to East when satellite moves from North to South to location of the starting points of successive lines are displaced in a west-east direction. The earth rotation corrections is applied by adding a calculated integer number of 'Zero' fill pixels before and after each line of imaged data. The Fig.no 5 shows the Geometric distortion of the Landsat images and Nonsystamatic Distortions.

4.2.3 Mirror scan non-linearity:

The MSS mirror's scan rate is not constant. It varies non-linearily due to the imperfect nature of the eletromechanical driving mechanism. The mirror velocity is not constant because of the speeding up at the beginning and slowing down at the end of the scan. It is more nearly represented by a cosine function. Because of the varying velocity the pixels are varying degree on the ground. To correct this, using the cosine curve the input location for each output pixel is competed and it's gray level is estimated.

4.2.4 Geocoded Data:

A data registered to a particular map projection is reference to as geocoded. By registerting different satellite data to one common projection, satellite independent data can be generated.

4.3 Image Enhancement :

Enhancement is the special processing of data to improve it's visual quality.

4.3.1 Density Slicing:

This method involves the humping together of grey levels in a range into a single gray level. This method works best with single band data. If several features, each have different values of means and standard deviations then several gray level slices may be produced. The new sets of gray level slices may be assigned different colours on a display.

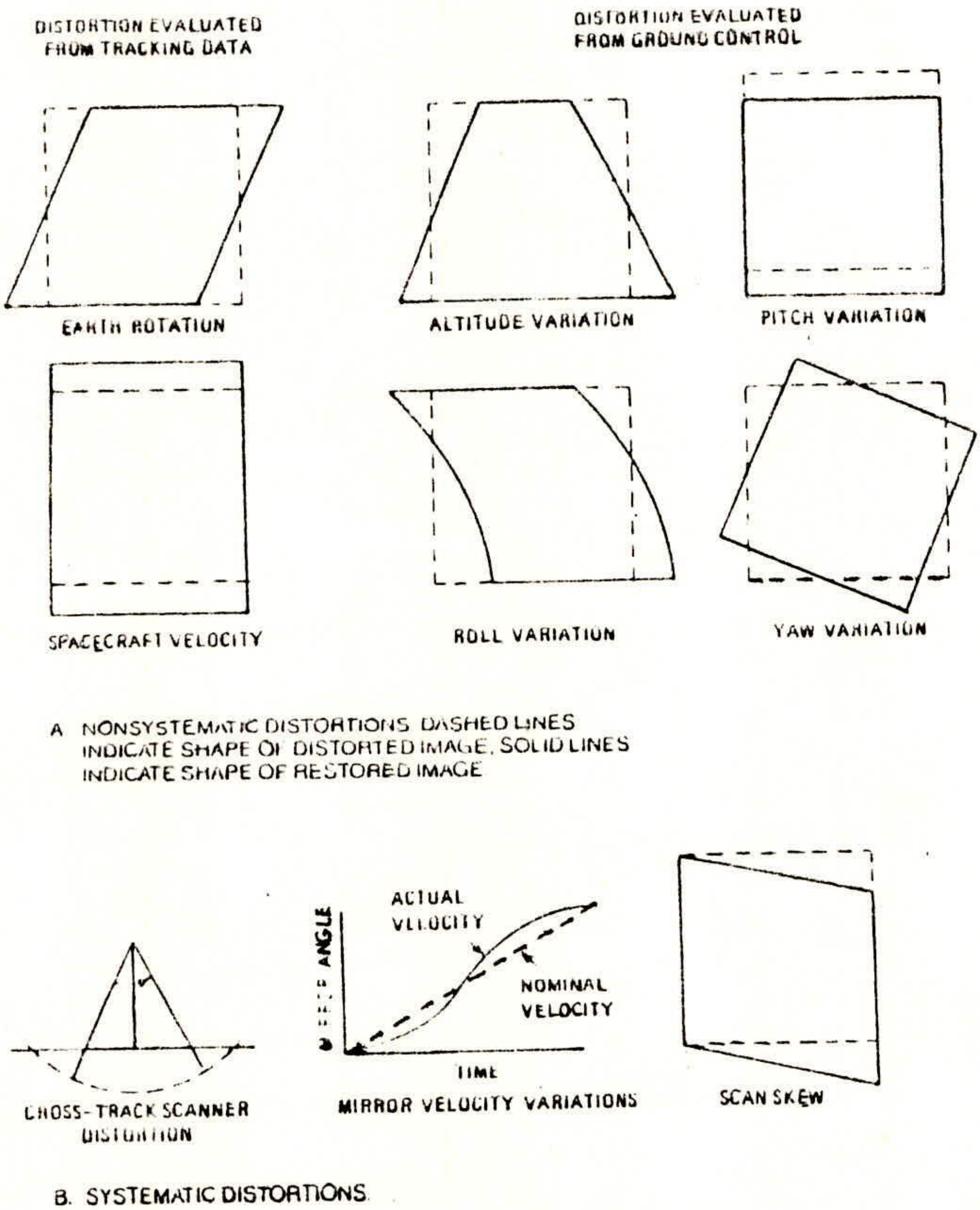


FIG NO. 5 GEOMETRIC DISTORTION OF THE LANDSAT IMAGE

4.3.2 Contrast stretching:

The range of gray level found in a given scene is usually much smaller than the available range of the film or display. So by artificially increasing the range, the contrast in the image can be increased. The Fig.no 6 shows the original image with no contrast enhancement and with Linear contrast stretch.

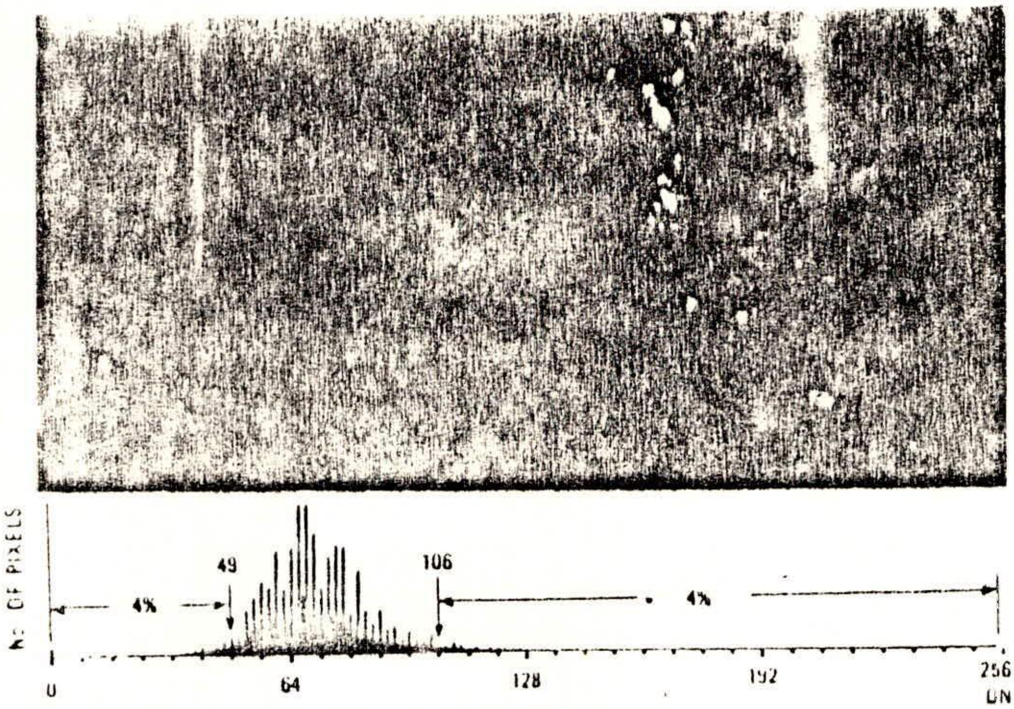
Commonly used non-linear stretches
- i) Piecewise Linear
- ii) Gaussian
- iii) Logarithome
- iv) Probability distribution function

4.3.3 Spatial filtering:

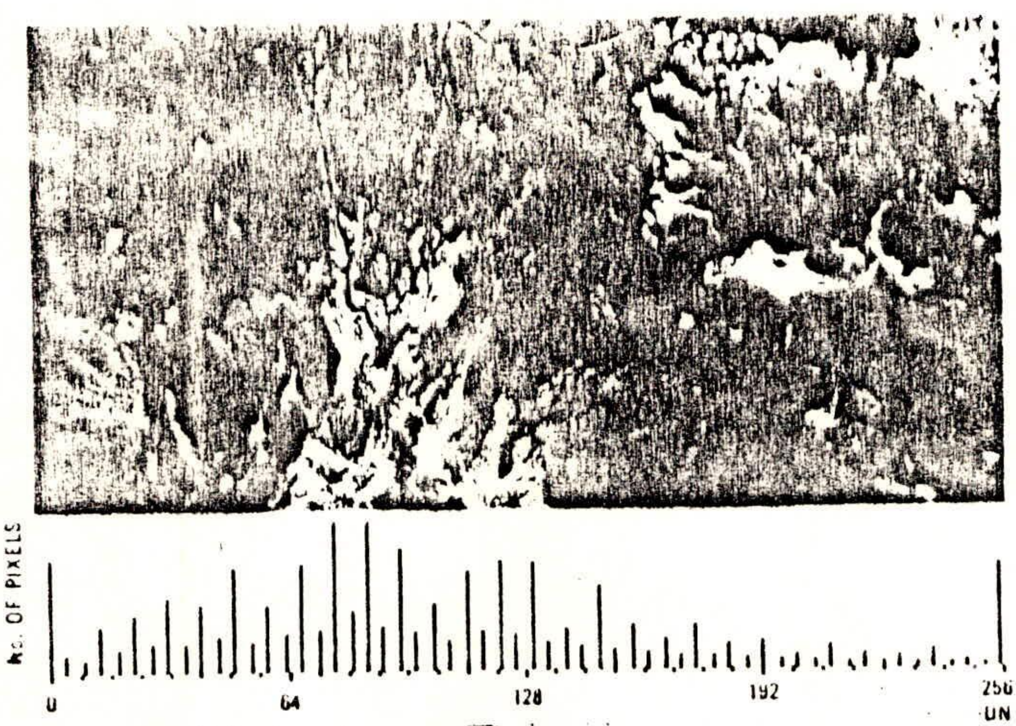
The above enhancement techniques handle each pixle independently of its neighbours. The relation between neighbouring pixles can also be used in enhancement. The mathamatical technique for separating an image into its various spatial frequency components is called Forier analysis. After an image is separated into its components spatial frequencies it is possible to emphasize certain groups of frequencies to produce an enhanced image. Both high and low frequencies can be emphasized.

4.4 GIS Based Approach :

All the hydrological data and other study elements as lineaments, geomorphological maps prepared visually, along with existing information from the georeferenced database are integrated and analysed in a GIS based approach. The published maps of soil, landuse/landcover, geology, alluvium thickness etc are digitized and georeferenced according to UTM co-ordinates to facilitate "map overlay" based analysis.Target areas of interest



A. ORIGINAL IMAGE WITH NO CONTRAST ENHANCEMENT.



B. LINEAR CONTRAST STRETCH WITH LOWER AND UPPER FOUR PERCENT OF PIXELS SATURATED TO BLACK AND WHITE RESPECTIVELY.

FIG NO. 6

are isolated based on the concepts of convergence of evidences". The GIS model functions as a simple additive scheme where areas of interests are highlighted and those that are not significant to work are masked out. For example the final ground water target map is obtained by combining weighted GIS input files i.e. vegetation + Geomorphology + lineaments + Drainage texture + Bed rock mask. The fig no.7 shows the conceptual component of a Geographic Information System data types and various methods of representation.

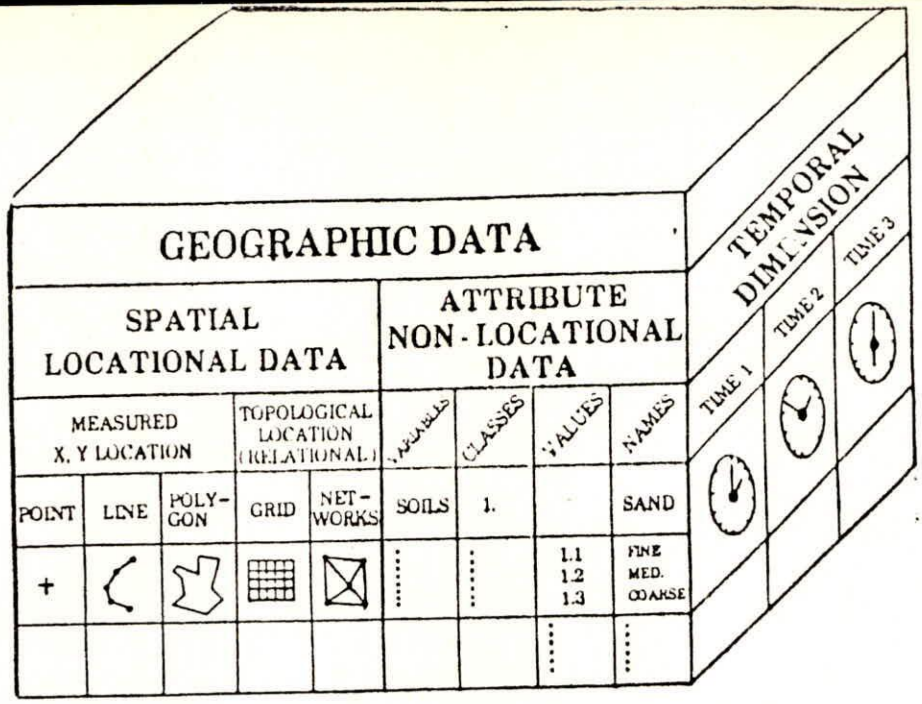
5.0 DATA PROCESSING BY MACHINE METHODS

5.1 Introduction:

Data processing may be defined as any systematic procedure through which basic information is transposed into more accessible or more directly usable forms. The data have first to be recorded at source, dispatched to a collecting centre and finally distributed to the users. Nowadays data production and transmission are becoming so bound up with data processing, both on the organisational and on the equipment sides that it is sometimes difficult to differentiate between them at all clearly. Throughout the world an immense amount and variety of meteorological data have been accumulated over the years. Moreover their rate of production is rising steadily because of ever-growing demands both practical and scientific and because of invention of new recording and observing devices (instrumental digitizers, radiation sondes, space satellites etc.).

5.2 Data Processing Media :

Many services have already adopted punch-cards for data processing. The punch-card is the internationally accepted medium



Three conceptual components of a geographic information system (GIS).

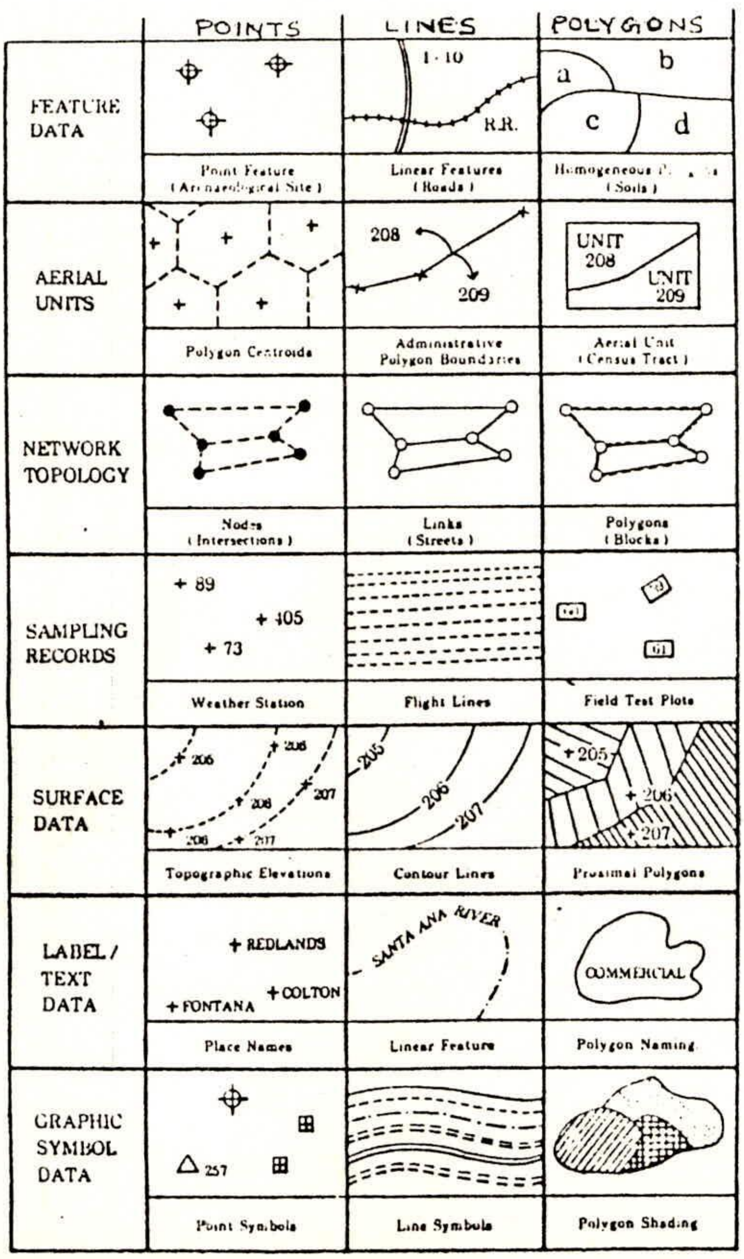


FIG.NO.7 Geographic data types and various methods of representation.

of exchange for meteorological data in machinable form and the use of a standard 80 column card is recommended by the WMO. Punchcards have a long history of use with mechanical and electro mechanical machines and more recently with electronic computers. But they are not the only possibility; paper tape, familier in telecommunication work, and magnetic tape are now widespread, especially with computers and for applications for which compactness and continuity of record are a prerequisite. A more practical alternative and one that is increasingly being employed is to transpose instrumental readings directly on the required processing medium by means of digitising equipment. Tape is usually preferred to cards for this because the euipment is more compact and more reliable for long periods of continual use.

Calculation may only be a small part of climatological data processing and often only the simplest statistics are needed. For example a single card sorting machine can rearrange and classify a large volume of data and prepare frequency distribution in a fraction of the time and at much less cost than by purely manual means.

Climatological data need normally to be held in safekeeping for long periods not merely for posterity but for use, preferrably on a medium adopted to operating automatic data-processing equipment. The ideal medium, if one existed, would need to be (a) durable and non-erasable to ensure permanence of the data; (b) flexible, to allow easy arranging, editing, and modifying of the data; (c) versatile, to allow ready use for all purposes and by all means, from visual browsing and manual manipulation, through automatic processing by the entire range of data-processing machines; and (d) compact, to allow accumulation and storage over long periods in manageable proportions and at reasonable costs. No ideal medium exists, however; nor can one be expected, since the requirements are to be large extent mutually incompatible.

Punch-cards, punched paper tape and magnetic tape are used extensively in data-processing both inside and outside meteorology; and punch-cards, magnetic tape and digitally coded microfilm are used extensively for automatic data-storage and retrieval. Manuscript records, autographic records, charts and printed paper documents are time-honoured media for the storage and manual retrieval of data: recently equipment has been developed for automatic retrieval from autographic records and printed documents. To ensure better permanence and more compact, cheaper storage and reproduction the practice of microfilming paper records has grown in importance.

The main characteristics of the above media are set out in Table no 1. From this table, it can be seen that each medium has its advantages and weaknesses.

5.3 Data Processing Machines :

Data processing equipment may be broadly classified either by function (recorders,sorters,tabulaters,calculators,etc.)or by mode of operation (manual,mecanical,electonic). In data processing machines the mental load has some times been reduced while the physical effort has been incresed,but in general the trend has been to diminish both with a consequent increase per operater in the number of functions and tasks that can be carried

TABLE 1

COMPARISON OF DIGITAL DATA-PROCESSING AND STORAGE MEDIA

Medium	Storage density Characters per cc	Ability to arrange, edit or modify content	Archival life and safety of data	Special storage requirements
Punch-cards (Hollerith code)	25	Good	Poor	Compression
Paper tape	125 250	Poor	Poor to fair	Spools
Magnetic tape	6,000 60,000	Excellent	Fair	1. Spools 2. Magnetic shielding 3. Dust-free 4. Temperature and humidity control
Microfilm (digitally coded)	6,000 30,000	None (except by splicing or by reel)	Good	1. Spools 2. Dust-free 3. Temperature and humidity control
Microfilm (plain language)	30,000 2,000,000	None (except by splicing or by reel card)	Good	1. Spools (for rolls) 2. Dust-free
Printed sheets or rolls (plain language)	1,000 2,000	Poor to fair	Fair to good	Little or none

Medium	Speed (characters/sec.) Reading	Writing or punching	Advantages	Disadvantages
Punch-cards (Hollerith code)	20- 2,666	20- 333	1. Flexible 2. Input-output for all classes of machines	1. Wear and age rapidly 2. Easily lost 3. Slow 4. Bulky and costly storage
Paper tape	20- 2,000	10- 1,000 (300 maximum for punching)	1. Data not easily lost 2. Cheap readers and punches	1. Inflexible 2. Slow 3. Wear and age rapidly if poor quality paper 4. Bulky storage

TABLE 1 (continued)

Medium	Speed (characters/sec.) Reading	Writing or punching	Advantages	Disadvantages
Magnetic tape	10,000-300,000	10,000-360,000	1. Flexible 2. High speed 3. Compact storage	1. Data signals age and can be erased 2. Costly writing and reading equipment
Microfilm (digitally coded)	133-100,000	10-30,000	1. High speed 2. Compact storage 3. Long, safe life 4. Cheap reproduction	1. Inflexible 2. Costly reading equipment
Microfilm (plain language)	200-2,000	10,000-200,000	1. Human readable 2. Ultra-compact storage 3. Long, safe life 4. Cheap reproduction	1. Inflexible 2. Read automatically only by slow, costly equipment
Printed sheets or rolls (plain language)	200-2,000	10-60,000	1. Human readable 2. With good paper quality, long safe life	1. Bulky storage 2. Read automatically only by slow, costly machines

out on one machine. Bearing in mind that most problems can be broken down into smaller elements and that processing units are repeatable, there is no theoretical limit to the performance of any category by equipment.

5.3.1 Punch-card Machines

Key-punch and verififying machines:

The transcription from the observational records to cards is done on keyboard operated machines, one column at a time proceeding from left to right on the card. The power for operating the simplest type of key-punch is supplied manually by the operator, who also feeds the cards in one at a time. Electrically-driven punches are also available with automatic card feed; these are easier and faster to operate, having a lighter touch. They can also be "programmed" to control such automatic facilities as skipping adjacent columns (or fields) not required to be punched, duplicating standard data when the "auto" key is depressed.

Verifiers are similar to key punches in construction and operation but are designed to test for the presence or absence of holes rather than to cut them. A previously-punched card is placed in the verifier and the operator proceeds to go through the motions of repunching form the original coded documents. When a key is struck, plungers are depressed which should match the punched holes and penetrate them, thereby completing an electrical circuit and allowing the carriage to advance by one column. If the depressed plungers do not match the punched holes, the keyboard locks and a warning light may be shown. The operator then has to decide what is wrong and whether the card has to be repunched. Proof that cards have passed through a verifier is usually provided by a single notch cut in the edge of the card or by a small ink dot recorded at the bottom of each tested column.

Reproducing machines :

These machines are primarily for punching data from one set of cards to another, usually with automatic checking as the copies are made. All or part of each card can be reproduced, relocating the data fields if need be. Some reproducers are also capable of comparing the punchings on separate sets of cards even when their field locations do not correspond.

Sorting and collating machines :

Sorters are used for rearranging cards in any designed sequence and for segregating or grouping particular categories of data. Each card is fed beneath a sensing device, set to operate on one selected column at a time, and is automatically directed into one of thirteen separate pockets, corresponding to each of the twelve possible punching positions with an additional "reject" pocket.

5.3.2 Special and peripheral equipment:

A growing variety of equipment has been developed to operate n association with automatic data-processing (ADP) systems or to provide linkages between otherwise incompatible systems.Essentially they are converters of some kind and often have to be specially made or modified to suit particular applications.

Analogue-to-Digital converters:

Digitizers accept the analgoue output from cbserving instruments in the form of voltages, pressure, temperatures, shaft positions, rate counts, measurements of weight or length, etc. This output is converted to digital form for recording on cards or tape (punched or magnetic), in print or directly to line. The sampling rate may be as high as several hundred readings per second, and a range of different instruments may be tied through one or more converters to a single recording unit. Automatic weather stations producing their observations, directly in a machineable medium and/ or broadcasting them by radio, are examples of this.

Digital- to- analogue converters:

Numerical data may be required in graphical or line form e.g frequency curves, time or space cross-sections, the contour charts in numerical weather prediction of the isopleths required for climatological atlases. There are several automatic linedrawing or graph- plotting machines commercially available which accept digital input form cards or tape, or directlyl from computers.

5.3.3 Electronic Data-Processing(EDP)systems :

The design and construction of digital computers are still a vigorously expanding technological field with an impressive variety of models to its credit. Improved components (e.g. transistors which have completely ousted the thermionic valve in computer manufacture) and logical designs have resulted in a large reduction in size, weight, power consumption and heat dissipation, and a large increase in speed, reliability and ease of operation. This in turn has made it feasible to build large multipurpose EDP systems round a single computer.

All computers are now assembled from electronic "packages" of high intrinsic reliability, on a building-block principle. Thus on the few occasions when a fault in the computer does develop, it is merely necessary to trace the affected package which can be replaced in a matter of seconds so that normal working can be quickly resumed while the faulty component is repaired at leisure. Another advantage of this method of construction is that a minimum system may be acquired initially and extra capacity (memory and ancillaries) can be added as and when required with very little interruption of day-to-day working.

In climatology a card or tape-oriented computer equipped with two readers, a punch and a printer would be capable of carrying out all the collating and sequential processing required, a task normally done by a much larger array of conventional card machines. Sorting, in the sense of grouping items or of a limited rearrangement of orderly data, can be done with very little internal storage but a large random sort (e.g.of marine data) would require a magnetic drum at least, and it may be more economic to punch and sort on cards for this class of data. Where there are international exchange obligations for card data an extra card punch would be necessary.

5.3.4 Parallel-processing and time-sharing systems :

The fastest modern computers require the speed of magnetic tape and/or the combined speed of a large number of the slower peripheral units to keep them busy. Also external transfers must be allowed to proceed in parallel with each other and with the internal processing of data already in store. This facility of Parallel processing need not be in respect of a single programme:the available computer time may be shared between several under the control of a master or director programme.

In some time-sharing systems the programmes are dealt with in strict rotation, and time of the central control and arithmetic units is shared more or less equally: in others there is a priority setting so that a particularly important job may be made to take precedence over another. In either case the aim is to utilize every piece of equipment and every microsecond of computer time as efficiently as possible and with the minimum of inconvenience to the programmer and the operator. The master programme supervises the switching arrangements between the peripheries and programmes being worked on and confines each to its own portion of the store, all levels of which have to be decided in advance by the programmer. In the most recent systems, however, this is also taken care of by the master control as is the optimum allocation of available equipment. If more peripheral equipment or memory is added later there is no need to rewrite the programmes in order to utilize it. The point of time sharing is that several concurrent progrmme can be completed in much less time than if run one after the other, and with a well-balanced schedule or work the delay to any one may be quite insignificant.

6.0 INDIAN ORGANISATIONS IN DATA COLLECTION AND PROCESSING

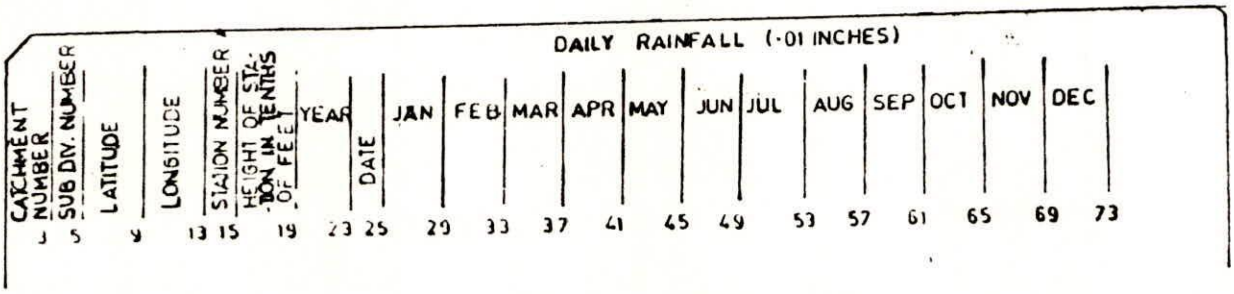
The following organisation are involved in the data collection and processing.

1 India Meteorological Department (IMD)
2 Central Ground Water Board (CGWB)
3 Central Water Commission (CWC)
4 National Informatic Centre (NIC)
5 National Remote Sensing Agency (NRSA)
6 Central Water and Power Research Station (CWPRS)
7 Andhra Pradesh Engineering Research Laboratory (APERL)
8 Institute of Hydraulics and Hydrology (IHH)
9 National Institute of Hydrology (NIH)

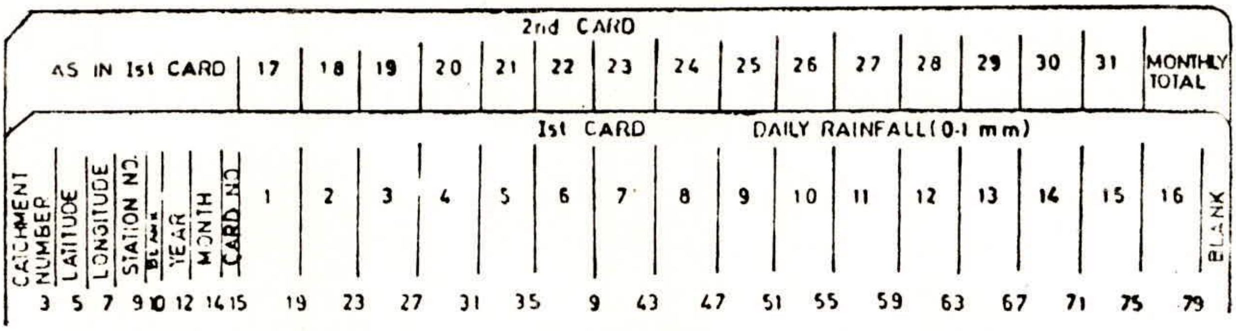
Among these organisations CWPRS, APERL, IHH and NIH are collecting various Hydrological data from other agencies and processing the data for their individual studies.

India Meteorological Department:

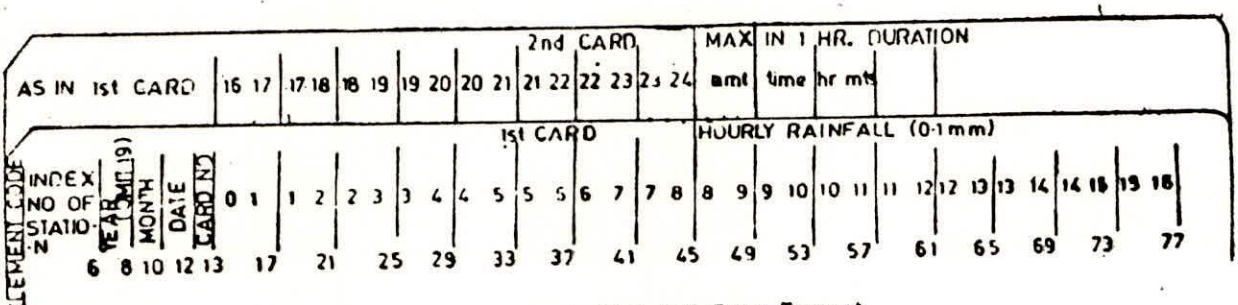
IMD has been storing climatological data for more than four decades. The formats for recording daily and hourly rainfall records are as shown in Fig no 8. In earlier times processing was generally done manually. The need for use of faster methods has been felt with increase in the number of data record for processing and this led into the use of computers. Currently IMD is using two computer systems namely ROBOTRON EC - 1040 mainframe computer system and VAX 11/730 digital computer system for the, Storage and processing of the collected data. Types of storage media used are magnetic tapes, disks and floppy diskettes. Different types of formats are used for the storage of various types of data as devised by IMD/WMO, with record length varying from 80 to 125 bytes. After keying in and varification of data on data entry machines, they are transferred into magnetic tapes of checks viz multipunch check, duplicate check, validity check,



31 Card Daily Rainfall data Format



24 Card Daily Rainfall data Format



: Hourly Rainfall Data Format

Figure 2 The daily and hourly Rainfall Formats absured value check and climatological check etc are applied to make the data as free of errors as possible before final archival on magnetic tapes.

Central Ground Water Board :

The data collection of hydrological parameters on regional basis commenced in 1969 with establishment of hydrograph network sations all over the country. The data collection was limited to 410 hydrograph network stations in the country. The chemical data was being taken once a year.

The data processing including data acquisition and application related to the CGWB activities can be enumerated as follows. The data base system establishes a linkage between data aquisition and application programme. It includes.,

(a) Data-collection i.e. hydrograph network measurments and water sample collection and chemical analysis.

(b) Data entry-through computers terminal.

(c) Storage and retrieval of the data using software package-d base III plus.

(d) Report presentation.

(e) Analysis presentation-trend calculation,rise and fall maps, time series analysis, and chemical results presentation.

The PC/XT has been liked with NEC 1000 and multiuser possibilities of the system is being explored. Provision for plotter and digitizer is also being made. The DBMS in dbase III plus software consists of 18 programms interlinked to store data in main and datafiles. The programme is menu driven with following choice.

(a) Data entry
(b) Data modify
(c) Data delete
(d) Data listing

The programm is further linked to software GRAPHER to produce well hydrograph and to calculate the general trend of water levels.

Central Water Commission :

The Central Water Commission is engaged in computarised storage and retrieval of hydrological data for quite some time. Mainly river data directorate and statistics directorate in the headquarters of CWC are primarily responsible for establishing a data bank for systamatic collection, processing, storage and analysis of hydrologic data. For rapid and accurate processing and retrieval of large volume of data, the electronic data processing facilities of the NIC, Govt. of India are utilized. Data on computer is initially processed mechanically with two fold objectives.

Storing of hydrologic data on magnetic tapes and processing of data on computers for various hydrological computations. One of these software packages processes the gauge and discharge data of the various rivers flowing in the country and produces the reports usually known as water year books.

The daily gauge and discharge data for all the months for particular site on a river is printed in a tabular formate. Computed discharge is substituted in case of missing data. The maximum and the minimum water level/discharge and it's dates of occurrence are presented monthwise. The water levels and discharge is further carried out in a stratified spectrum such as monsoon season, non-monsoon season and annual.

National Informatic Centre :

NIC has launched it's district information system (DISNIC) under which slected information from the districts will be transmitted to the state computers. DISNIC has already evolved a data base with the parameters of the water use data at the district level. Water use related statistics and water level profiles of the various sources are maintained. Remotely sensed data obtained from satellite and aircraft surveys are used to study and monitor natural resources, natural phenomena and the dynamics of human activities with them. Rapid, synoptic and accurate studies on crop, soil, water potential, forests, environment and pollution etc. are now feasible with satellite remote sensing and digital image analysis. The DISNIC software division has undertaken the irrigation system study and analysis in DISNIC irrigation sector at the district as well as state level. In this the following sectors have been studied in detail. i.e. (1) Minor irrigation, Medium irrigation, Major irrigation, Flood and Drought relief schemes, Anti-sea erosion, land Navigation etc.

National Remote Sensing Agency :

It is an Autonomus Organisation under Department of Space, Govt. of India. In India various Hydrologic and water resources data required for various activities in water resources are not being collected and maintained at one place or available with one organisation, since various Central and State Govt. agencies are involved in it. Under this circumstances remote sensing technology can be gainfully used for water resources (Hydrology) data collection. Remote Sensing as a tool in hydrologic data collection, qualitative observations and quantative measurements were made in to two modes i.e. spatial (by Integration of point data) and temporal. Remote sensing provides the capability to expand data collecton in the temoporal mode tremendously.

Modes of Hydrologic Data Collection by Remote Sensing:

Spatial	Temporal	Spectral
Delineate	Continuous (real-time)	Detect Identify
Measure (area)	Intermitent (fast-reaction)	Relate or Infer
Quantity (parameters)	Periodic (repetative) one time	Interpret

Using the Remote Sensing technique the following studies are being carryied out by NRSA.

(i) Surface water mapping and monitoring
(ii) Identification of potential zone of groundwater
(iii) Flood mappig and Damage assessment
(iv) Snow cover area mapping and melt runoff forecasting.
(v) Drought assessment and monitoring.
(vi) Monitoring of vegitation status.
(vii) Mapping and monitoring of watersheds.
(viii) Water quality studies.
(ix) Irrigation studies.

Central Water and Power Research Station

It is primarily a sophisticated user of hydrometeorological data and carries out problem oriented collection, processing and analysis of water data , with the basic data genarally obtaining from other organisations such as IMD,CWC, state governament departments etc.,It collects rainfall data, stream flow data, evaporation and evapo-transpiration data , drainage basin charactaristics such as elevation , stream lengths, soil types, vegetation etc.,and details about water resources projects/

hydraulic structures in the region. These data are stored in a CYBER 180/840. A system with a large network of terminals which can also work as individual PC's when required.

Pre-processing/quality control is being done mainly to rainfall and stream flow data which are being used as inputs to the parametric hydrological models. Those data that are actively used in any specific analysis are kept in a data base system on hard disk. The rest of the data are compiled and stored in computer compatible form such as magnetic disks , floppy disketies and data cartridges.

Andhra Pradesh Engineering Research Laboratory :

It is a research Institute at Hyderabad under Andhrapradesh government. They are doing studies in various fields in Engineering and Technology. Development of data base management system in respect of hydrological data was taken up as research study in this institute. The APERL have used dbase III and developed the system called Hydrometeorological data creation and retrieval system (HYDCARS).

APERL has taken up another CBIP project namely Development of software for pariliminary processing of hydrological data using micro-computers and the same is in progress the following are the features of this software.

--Conversion of daily rainfall into, weekly, monthly and seasonal.
--to find maximum one hourly and six hourly rainfalls.
--estimating the missing rainfall data by normal ratio method.
--estimation of weighted rainfall using the weights of the polygons.
--estimation of absolute, maximum and minimum rainfall.
--conversion of hourly gauge data into different periods.
--estimation of peak flows in the reservoir for different return periods.

--stage discharge relations.
--to check the internal consistancy of data.
--computatioin of sediment rating curve for rising and falling floods.
---geometric elements of open channel systems.
---reservoir routing.

Institute of Hydraulics and Hydrology :

It is a research institute under Govt. of Tamilnadu public works Department. The data such as rainfall and discharge are collected and stored in the magnetic tapes in the data bank. There are about 34 river basins and 388 raingauge stations maintained by various agencies namely meterological, PWD and the Railways. The rainfall and inflow data of any stations are avilable from 1930.

The punching machine, DCM microsystem is also available.In olden days the data were stored in the punching cards and from 1981 onwards these data were stored in the magnetic tapes. They are using VAX/UMS version V 4.5 computer main frame system available at IRS, Anna University, Madras. In selected area, by surveying the sedimentation data, Area-elevation capacity data were brought to the data bank and which are also stored in the magnetic tape for use in computers. The processing, analysis and results of Vaigai reservoir is available.

National Institute of Hydrology :

National Institute of Hydrology is basically a research institute under ministry of Water Resources. It collects Hydrologic data from various agencies and related state government departments. But so far there is no centralised storage and retrieval system for these collected data. A large number of programmes are developed in Fortran IV language and are implemented on the VAX-11/780 computer system available with NIH for reading rainfall data from IMD tapes, to check the accuracy of the collected data, to fill the gaps in data series and to convert the collected data into desired form and format. The computer programme TAPE.FOR is to read daily rainfall data from IMD tapes and identify the missing data periods. A program DAILY reads the daily rainfall data (after indentifying the card format and cheeking for leap year) and writes it into a 12 X 31 matric. The computer programme GAPE. FOR is used for estimating the missing station rainfall data using normal ratio method and programme DISPOW.FOR is meant for estimation of rainfall using the distance power method.

The programm DOUBLE.FOR is to check the consistancy of particular record using double mass curve analysis.

Programme TENDAY.FOR uses the output file from the TAPE.FOR to compute the ten daily rainfall totals and average rainfall uring the ten day period.

Programme MAX.FOR is meant for the computation of the maximum 1 day, 2 days, 3 days, 4 days and 5 days rainfall from the daily rainfall data processed for missing rainfall amounts.

The programme CATCH.FOR computes the weighted average catchment rainfall using the theissan weights supplied by user. ISO.FOR computes the variation of depth area over the catchment using ISO-hyetal method.

The software HYMOS has been instaled on the PC which is a typical Hydrological data processing system in NIH. It covers almost all Hydrological data. The processing system covers data entry, Editing, validation, completion and regression, flow measurments, Data compilation, statistical analysis, time series analysis, reporting and retreval.

7.0 HYDROLOGICAL DATA BANKS AND PROCESSING ELSEWHERE

A complete National Hydrological data bank would contain data and statistics covering all hydrologic variables for all parts of a country and would permit any potential user to easily extract and use these data without extensive manipulation. The main elements of a data bank are the computer software and hardware required to handle these hydrologic data. It normally also includes access to the data series by users and data exchange. Depending on the way in which data are collected and processed, the data banks may be centralized, distributed or limited. A fully centralized hydrological data bank is physically located at one site and comprises data covering all hydrological variables for all parts of the country. A distributed national hydrological data bank consists of several data banks covering all hydrologic parameters with same type of co-operating arrangement between the different banks or agencies operating them. A limited hydrological data banks are centralized, which covers only a few of the many hydrological variables. Computers have been used for hydrological data collecton and processing for some years now. The WMO survey (1973-1976) of National Hydrological data banks indicated that many countries had either started to use or intended to use computers. i.e. around 64 countries employed computers. Among them in few countries the status of data banks and data processing will be discussed in this chapter.

Australia

Data Bank:

Australia has a co-ordinated national Hydrological data bank consisting of various separate data banks covering hydrological data for a region or state within Australia, responsibility for hydrological data collection is vested in the states and Territories with one authority in each state having prime responsibility. This has necessiated the development of procedures to enable data to be transferred between data bases and users in a universally acceptable format. In western Australia a state computer-based hydrological data bank has been established to ensure that the records are safely stored in a well-organised manner and are readily available. All hydrological data of long-term value are now held in this data bank with the exception of groundwater data for the densely monitored capital city metropolitan area which is held in an associated data bank. There is excellent cooperation between the various data collection agencies ensuring minimal duplication of data collection effort and only minor problems with the centralised data bank.

Computer facilities:

The data bank is held on a control data corporation (CDC) CYBER 1972, operating under the NOS/BE operating system. The machine has 131000(decimal) 60-bit words of central memory, two central processor units and 13 peripheral processors. The agency has ready access to a card reader and line printer terminal. A digital equipment corporation mini-computer (POP

11/34) acting as a concentrator for eight visual display units provides the main access to the computer. The eight VDU'S are used to control the flow of jobs and correct the data and in addition are used for system development and hydrological studies. The mini-computers also provides the paper-tape reading facility which forms an essential part of the data processing system. A graphic display unit (TEKTRONIX 4014) is used for the data varification and analysis.

Data Processing :

Processing continuous data:

At present the digitization process has been geared to the rapid handling of data charts. The quality of the data is maintained by a programme designed to check the data thoroughly. The analogue-to-digital conversion process is checked by occasional computer plotting of the data onto overlays, which are then placed over the original chart. Rating curves are developed from current metering for each gauging station and checked at least annually for stability. These rating curves are stored as a set of X and Y co-ordinates in a disk resident data base for ready access by other programmes. Daily flow volumes and peak flow rates are calculated from the level data and the rating curve routinely.

Processing non-continuous data:

The results of water-quality analysis and ground water levels are to be non-continuous data set. These data are to be entered into the machine by a key-to-disk operation directly from the field. The water quality data passes through the same series of checking and updating programms. The checking of the data for correctness and internal consistency is thorough. Calculations carried out as part of the laboratory analysis are checked again by the computer on data entry. Where parameters are correlated, such as electrical conductivety and chloride ion concentration general relationships are used to ensure that the data values are within an acceptable tolerence.

BELGIUM

Data Bank :

The Belgium National Hydrological data bank consists of a group of five data banks, i.e. central data bank managed by the Hydrology section of the Royal Meteorological Institute and four specialised data banks set up by various public administrations as shown in Table no 2.

Computer facilities:

The Royal Meteorological Insitute, the Royal obervatory and the Institute for space Aeronomy uses a UNIVAC 1100/40 computer with four disk units (60 000 000 characters) and four magnetic tape units. The input-output facilities include various punched cards and magnetic tapes, high-speed printer, and graph plotter. Other data banks uses SIEMENS 7738, SIEMENS 7000 and mini-computers with disk unit (10 000 000 characters) and a magnetic tape unit.

Data Processing:

Primary processing comprises data collection, decoding, identification and composition of the data in conventional units validation of both the identification and the data themselves addressing and storing in appropriate files. In secondary

TABLE 2

Central hydrological data bank of the Royal Meteorological Institute of Belgium (IRM)

Source and volume of stored data

Source	Type of data	Points of measurements
Ministry of National Education		
- Royal Meteorological Institute		
- Climatology Section	- Daily rainfall	Approx. 350 rainfall stations (1)
- Hydrology Section	- Daily rainfall and rainfall at 10-min intervals	37 stations (2)
	- Air and soil temperature; air humidity; run-of-the-wind at 2 m; evaporation and temperature of evaporation pan	19 stations (3)
	- Areal averages of precipitation over hydrographic basins	62 basins
	- Potential evapotranspiration from hydrographic basins	43 basins
- Radiometry Section	- Global solar radiation.	7 stations
	- Duration of sunshine	14 stations
Ministry of Economic Affairs		
- Mines Administration	- Groundwater levels	171 wells
Ministry of Agriculture		
- Agricultural Hydraulics Service	- Daily stages and discharges of rivers	175 stations
	- Hourly stages and discharges of rivers	86 stations
	- Comments on the stream-channel conditions	261 stations
Ministry of Public Health and Environment		
- Administrative Unit for Water	- Groundwater levels	(Being developed)
Ministry of Public Works		
- Hydrological Studies Service	- Daily stages and discharges of rivers	
	- Comments on the stream-channel conditions	75 stations
- Maritime Services for Antwerp	- Monthly discharges of rivers	5 stations
Brussels Inter-borough Company for Water	- Daily discharges of rivers and tappings	6 stations
	- Groundwater levels	50 wells

processing the following manual checks, quality control by machine methods to be taken into account.

---Intercomparison of recorded values and direct readings
---plotting hydrographs for manual comparison with the original charts
---Regional comparisons based on the ratio of reduced values of stages
---For stream gauging, plotting a graph of the points at which measurements are made and comparison with the original cross-section
---Plotting the graph of the annual regime of specific discharges and regional comparisons.
---Regional comparisons of monthly and annual streamflow deficits

Some of specific tasks in secondary data processing include.

---Calculation of the average areal precipitation over river basins
---Calculation of daily mean values of evapotranspiraton from river basins
---Calculation of mean velocity and discharge based on stream gauging
---Analytical fitting of stage-discharge relations conversion of stages to discharges
---Drawing ground water level maps
---Evaluation of the indices of water quality

And also preparation of regular time series containing monthly tables of hourly values with means and extremes, annual tables of daily values with means and extremes, and miscellaneous graphs showing variations with time.

CANADA

Data Bank:

The Inland waters Directorate, Environment Canada, maintains two hydrological data banks, one operated by the Water Survey of Canada (WSC) contains data on water levels, discharges and sediment and the other, operated by the water quality Branch, contains data on water quality.

The Water Survey of Canada (WSC) has been collecting and publishing streamflow and water level data, on a regular basis since 1908 and sediment data since 1961. The present hydrometric network consists of approximately 2450 gauging stations. Daily streamflow and water level data have been published annually or biannually on a regular basis since about 1908 in a varity of formats in over 230 publications. Daily discharge and daily waterlevel data for rivers and lakes are now published annually. The monthly and annual mean discharges and annual extremes and total discharges published for the period of record for all stations where streamflow data have been collected. These summaries are published every three years in eight volumes as for the surface water data publications. Likewise the waterlevel data and sediment data are also published periodically.

Computer facility:

NAQUADAT is run at a commercial time sharing vendor, where, there are IBM 360/85, 370/168 and 3033 computers. Normal access is via a high-speed reader printer terminal in place Vincent Massey, Hull. Secondary access is possible via portable telephone coupled terminals. The programms are written in COBOL, MARK IV, and FORTAN. Limited statistics are generated in the reporting programmes, eleborate stastical analysis is achived via intergrated stastical packages such a SPSS or BMDP.

Data Processing:

Data are forwarded annually to Ottawa by the regions either on punched cards or magnetic tape processed on a CDC CYBER 74 computer and print-outs returned to the regions for verification of data. These data are then processed at Ottawa to give various types of publication using computer programmes. Publications are printed in accordance with a pre-arranged contract to ensure minimum delay. Data are available to users in publications or computer compatable form about three months after being submitted at Ottawa. The computer science centre at the federal department of engery, Mines and Resources in Ottawa has been used for data processing since 1966 when a CDC 3100 computer was in operation.

COLOMBIA

Data Bank:

The hydrometeorological network in the country comprises the following approximate numbers of stations shown in Table .

	Hydrological	Rainfall	Climatological
Stations operated by HIMAT	599	921	447
Stations operated by other bodies	584	400	327
Stations suspended	265	642	61
Total	1448	1963	835

At present the regional offices send to the central office coded data sheets, on which the curves traced by the pen have been digitized. These data are then recorded on technical carriers, processed and stored. So far the work undertaken in building up the information system may be summarised as follows.

i. Introducing a coding system for the stations based on dividing the country into hydrographic zones.

ii. Storing on magnetic technical carriers the original and derived information which is at present being transferred to the final files.

iii. Studying other data banks for this purpose, certain centres in the united states and the meteorological and Hydrological services in England, were visited Estimating the potential input of the system.

iv. Enquiring into the actual needs for information in order to determine priorities for parameters and for particular analysis of these parameters, according to frequency (hourly data, daily data, etc).

Computer facility:

HIMAT has acquired a TEXAS 990-10 (64k) computer which is equiped with two diskette units, a printer and a plotter and which will be linked to the large, well equiped IBM 370-145 (1000 k) belonging to another government agency. At present, the IBM computer is used for the hydrological data bank.

Data processing:

In general terms, data processing and quality control for the hydrological parameters consists of the following.

Stage:

Staff gauge reading are compared graphically with the recorder charts, resulting in adjustments to the scales of instruments or in the methods of observation. The pairs of values of time/stage for significant points of slope variation are coded and subsequently checked. Once the data have been recorded the NIVCAU program for basic processing is run, which detects inconsistencies and an annual series of mean daily values of stage is obtained.

Calibration tables:

Using the conventional methods for interpolation and extrapolation the relationships stage discharge are defined or re/defined.

Discharges :

The aforementioned NIVCAU program enables mean daily discharges and extreme values to be obtained from the original values of stage using the appropriate calibration table. The resulting series are compared by another computer program which effects a comparitive balance for various stations.

Applying the relationship for the concentration of sediment in the vertical a computer program calculates the mean daily concentration of sediment. The same program calculates the daily transport, multiplying the concentrations by the corresponding mean daily discharge.

Air temperature and humidity data are processed by computer to obtain daily values of humidity vapour pressure and dew point. A new program has been completed for processing enemograms, enabling frequencies to be calculated and stored. A program is being written which will process other meteorological parameters such as pressure, precipitation and sunshine.

MOROCCO

Data Bank :

The Moroccan national hydrological data bank is of the centralized type. The files were built up and are managed by the Water Resources Division of the Direction of Hyraudlics (Ministry of Equipment and National Development) in close collaboration with the Computer Service of the Direction of Hydraulics. Morocco has approximately 30 synoptic stations, 850 rainfall recorders. The stations are essentially equipped with direct-reading or graphic recording instruments. With a few exceptions, the networks are not automated. The Water Resource Division (WRD) operates a network of 226 permanent hydrometric stations, including 142 principal stations, 65 secondary stations, 9 stations on water springs, 8 stations on irrigation canals and 2 stations on natural lakes. The secondary stations are not equipped to measure water levels during flood events. Of the 226 existing stations, 87 are equipped with water-level recorders. There are also about 90 SSB-type radio stations for the provision of flood forecasts and warnings. In addition, WRD makes regular measurments of grounwater levels at about 3000 points.

Computer facility:

The computerized handling of data is under the responsibility of the Computer Service of the Water Engineering Department. The collection, checking and preparation of hydrological data is performed by WRD.

The computer service is equipped with an IBM-370/125 computer (256k real memory), two 3,344 disk drives, two magnetic tape drives, one card reader, one mini-disk reader, one printer, one IBM-3776 terminal and one BENSON 2222 graphic system. Data are put at present on mini-disks through twelve IBM 3740 stations. The data files are at present on magnetic tape and will soon be on disks. There are duplicates of all the files and one copy is kept in for fireproof cabinet.

Processing of Data:

Daily rainfall totals are checked before putting on technical media, but it is not yet possible to make sure of the day to which the 0700-h rainfall readings should be assigned. A method involving the use of data from recording raingauges and lists of stations classified into catchment areas and sub-areas is being worked out.

Stream-gauging and water-level measurements are checked against the original values after punching. Water levels are compared with the level recordings at random for the preparaton of the hydrological year-book and systematically for hydrological studies. The graphical and computer analysis of stream gaugings are compared, permitting the detection of coding, punching or anlaysis errors.

When the rating curves have been fed into the computer and stages checked, the computer separates the periods which can be processed with the rating curves from the periods which must be treated using logarithmic interpolation between gaugings (during depletion).

The water-quality analysis outputs are checked against the original bulletins and corrected as necessary. A number of checks are also carried out on the computer (ion balance, theoretical and measured conductivity, total of dissolved ions against the dry residue). These checks indicate the quality of the analysis. A number of secondary processings, such as the determination of hardness, quality, salinity, and suitability for drinking,as requested by users, are also carried out on the computer.

NORWAY

Data bank :

The hydrological data bank at the Norwegian Water Resources and Electricity Board (NVE), located in Oslo, is a centralized bank for surface-water and groundwater quality. It covers hydrological data from all parts of the country. A standardized retrieval system permits any user to extract and treat these data with out extensive manipulation. Close to 1300 hydrological observing stations are being operated of which about 800 provide data for computing discharge. About 500 statons are equipped with automatic water-level recorders. Most of these stations are constructed to operate under extreme weather conditions.

Computer facility:

Data are processed on a CONTROL DATA CYBER 171, 132K 64 bits word computer, with disk, magnetic tape, card reader and printer equipment. The installation is based on a typical time-sharing operating system and has facilities for up to 64 terminal connections. The ownership is shared between all departments within the institution. The daily administration of the machine as well as primary punching of incoming data is taken care of by a separate electronic data-processing department. The bulk of programs operating the data bank will be run in remote batch mode. Retrieval and analysis will be mostly interactive.

Data processing:

Data are generaly grouped as monthly records and punched twice to avoid punching errors. When the data are read into a work file the raw data are listed and plotted as water stages or converted to daily discharge.

The data are corrected for ice damning in the winter season. This correction is today made manually and utilizes data from adjacent ice-free gauging stations and meteorological data as well as discharge measurements during the ice period. The daily mean discharges are estimated during the ice season and afterwards reconverted into equivalent water stages according to the rating curve valid under ice-free conditions.

When the data from one year are corrected, the extremes and monthly means are compared with values from previous years. If these statistics and plots or listings of daily values seems reasonable, the series are accepted and may be merged into the historical archives of annual records.

The other types of data are processed by similar procedures. During the processing several controls are made depending on the data type. These controls, however, are simple and fail to discover many errors.

UNITED STATES OF AMERICA

Data Bank:

The U.S. Geological Survey investigates the occurrence, quantity, quality, distribution, and movement of the surface and underground waters that comprise the water resources of the United States. The data collected during field activities are entered into the national Water Data Storage and Retrieval system (WATSTORE).

Streamflow data are gathered at about 9000 sites. The National Stream Quality Accounting Network (NASQAN) maintains 525 stations at which uniform sampling techniques and analytical methods are used to ensure comparability and reliability of data. Many districts use small mobile laboratories built into trucks or vans to perform required field analysis such as dissolved oxygen and pH and to prepare samples for shipment to laborataries.

Digital recorders are used at many field locations to record values for parameters such as river stages, conductivity, water temperature, turbidity, wind direction, fluorides, and chlorides.

Data are recorded on 16-channel paper tape, which is removed from the recorder and transmitted over telephone lines to the receiver at Reston. Va. The data are recorded on magnetic tape for use on the central computer.

Computer facility:

The IBM 360/65 computer system was acquired by the survey in late 1965. This system, for the first time, offered the capacity of remote access to the computer. The WATSTORE system is now accessed by over 100 computer terminals.

Data processing:

Each data collector is responsible for verifying the accuracy of data to ensure that what was intended to be entered into the file was entered without error.

Data recorded on magnetic tape from digital recorder paper tapes are processed by programs from an auxiliary file. For example, gauge heights or dial readings from this magnetic tape are processed through rating tables obtained from an auxiliary file to produce output values,such as discharge, temperature, dissovled oxygen, and conductivity. In addition, shift and datum corrections can be applied to these values to obtain an updated output value. These output values are then stored in the main file as mean daily values or maximum and minimum values.

Edit checks are made on data being entered into many of the files, but these checks only identify impossible values, and are no substitute for accurate data entry techniques and visual examination of data. Data that have been accepted and integrated into the files are then available for application programs.

Secondary processing products such as tables, graphs, plots, and maps are output from these application programms.

8.0 A TYPICAL HYDROLOGIC DATA PROCESSING SOFTWARE "HYMOS"

8.1 Genaral overview:

HYMOS is a database management and processing system for hydrometeorological quality and quantity data, designed for use on personal computers (pc's). It arranges a convenient structuring of data in a database and provides an extensive set of tools for data entry, validation, completion, analysis, retrieval and reporting.

HYMOS is comprehensive, well tuned and easy to use via full screen menus with on-line help to guide the user. The package includes many tabular and graphical options facilitating efficient reporting. HYMOS runs on stand-alone computers, but can also be used in a network system. Securities have been built in to restrict the acess for certain activities to qualified staff only.

HYMOS is developed to streamline the storage and processing of (geo-) hydrological and meteorological data. It is tailored for use by hydrological and meteorological data processing branches, water resources management authorities, water boards, water engineering consultants and hydrological advisers.

HYMOS data are to a large extent typically time-oriented. Together with a space oriented Geographical Information System, it covers all data storage and processing requirements for planning, design and operation of water management systems.

HYMOS comprises following systems:

A. a database management system, to create database, to structure the database and to define user idententifications;

B. a data storage and retrieval system, covering data entry, editing, reporting in tabular and graphical form as well as the transfer and retrieval of data;

C. a data processing system, including validation, series completion by interpolation, simulation and regression techniques elaboration of flow measurements, data compilation, statistical analysis, and time series analysis.

8.2 Data types :

The types of data, handled by HYMOS, can be categorized in the folloing groups:

8.2.1 Space-oriented data, covering :

--catchment characteristics
--station partiiculars
--station histories
--geo-hydrological profiles.

Catchment characteristics :

The catchment data comprise typically time-independent characteristics like area, river length, slope, course, (sub-) rivers) and topography catchment boundaries and features can be displayed for graphical series selection.

Station particulars :

The station particulars stored in HYMOS include:

--data on the location of the station, river basin, district, geographical latitude and longitude (UTM may be used for entry), altitude, catchment area and data collecting agency

--extremes of climatological parameters, water level and discharge.

Station history :

Station history data consists of free text covering e.g. station maintenance information, status of data processing, etc.

Since most commerically available text editors can be linked to HYMOS for the entry of station history information the user can apply the text editor he/she is familiar with.

Geo-hydrological profiles :

Groundwater-well data, including filter dimensions and a full description of the geo-hydrological profile (dimensions, soil type, porosity and permeability) can be stored under HYMOS.

8.2.2 Time-oriented data, including:

- --equidistant time series, i.e. series with regular time intervals
- --non-equidistant time series, i.e. series with irregular time intervals.

Time series types :

Time series represent various types of data (e.g. rainfall, temperature, water levels, chlorinity, sediment concentration, etc.). A two-letter combination is used in HYMOS for identification. A maximum of 675 different data types can be defined.

A data type refers either to equidistant or to non-equidistant series and has two attributes: a unit, or an observation code, to distinguish among instantaneous, accumulative, averaged and interval constant data.

Equidistant time series :

Various time intervals for a particular data type can be applied. The intervals can vary from one minute to one year (e.g. quarterly, hourly, daily, monthly and annual rainfall). The equidistant time series are identified by a combination of station code, data and interval code.

A special data type and interval is reserved for monthly maximum rainfall amounts at intervals of 5-10-15-30 min, and 1-3-6-12-24-48 hrs for purpose of Intensity-Frequency-Duration anaylsis.

Non-equidistant time series :

Non-equidistant time series are identified by: station code plus data type. In deviation from the equidistant time series each series element in the data base is stored with a data and time label.

8.2.3 Relation-oriented data, divided in:

--stage-discharge data
--relation or rating curves para-meters, valid for a certain time period.

Stage-discharge data :

Flow velocity measurements can be stored temporarily under HYMOS for validation and further eloboration in speical files. The condensed results of the flow measurements: i.e. sets of water level, discharge, velocity and cross-sectional data are stored in the data base. In case backwater or unsteady flow effects, the stage-discharge relation, additionally water levels at a second location or gradients are stored as well.

Relation and rating parameters :

Sets of coefficients, ranges of applicability and validity period for relation curves, stage-discharge relations (with or without unsteady flow and backwater effects) and sediment rating curves can be stored in the database to be used in computations.

8.3 Data storage and retrieval system :

Data storage and retrieval activities are comprised in two modules: The "Entry and editng" module and the module "Reporting and retrieval".

8.3.1 Entry and editng

The entry & editing module includes following options:

1. entry and editng of catchment data
2. creation of stations and series, and entry and editing of station particulars and geohydrological profile data
3. loading of equidistant and non-equidistant time series data from user files or manually; direct links with telemetering systems can be established, while also fast data transfer from EPROMs to the data base is possible
4. full-screen editors for editing and display of equidistant and non-equidistant time series, with on-line graphical display
5. entry and editing of current metering data and discharge and sediment transport rating data
6. entry and editing of data files for particular analysis (statistical analysis, regression, etc.).

8.3.2 Reporting and retrieval :

The reporting and retrieval module comprises:

1. preparation of reports on station and series characteristics and time series
2. preparation of mixed tables of database quantities
3. plotting database quantities in a user specified layout
4. retrieval data stored in the database
5. transfer of data from one database to another.

8.4 Data processing system :

The data processing system of HYMOS includes following modules:

1) data validation
2) data completion and regression

3) flow measurement

4) data compilation

5) statistical analysis

6) time series analysis.

Details on the contents of the modules are given below:

8.4.1 Validation

For data validation use can be made of the following procedures:

i) data screening by flagging, printing and tabular comparison of time series and computation of basic statistics

ii) graphical evaluation of time series (lines or bars), including:
   - -plotting of time series (max.5 per graph)
   - -residual time series graphs
   - -residual mass curves
   - -moving averages
   - -water balance
   - -graphical presentation of series availability in the database, i.e. bars for equidistant time series and number of data per month for non-equidistant time series

iii) relation curve analysis, to establish relation equations and to investigate shifts in the relationships; it includes:
   - --graphical display of relations
   - --fitting of relation curves by polynomials, divided in 1,2 or 3 intervals to account for effects of irregular or compound cross-sections
   - --storage of relation curve parameters in the data base
   - --comparison of relation curves of different time periods

iv) investigation of series homogeneity by means of double mass analysis.

v) statistical tests on data homogeneity and randomness.

vi) spatial homogeneity tests (near neighbour-technique), where data at a base station are compared with weighted averags of neighbouring stations, selected on distance and orientation.

8.4.2 Data completion and regression

The completion and regression module comprises:

i) interpolation techniques for filling in missing data based on time and space interpolations

ii) regression models to establish relationship, or to fill in missing data; the models may vary seasonally (max 12 periods), and can be of the following types:

- -polynomial equation
- -power equation
- -logarithmic equation
- -hyperbolic equation
- -exponential equation
- -simple and multiple linear regression equations
- -stepwise regression

iii) Physically based lumped parameter rainfall-runoff model SAMO, for filling in missing runoff data. SAMO is derived from the Sacramento streamflow Simulation Model. The model is a system of parallel and serial reservoirs. In the simulation of the runoff process a distinction is made between the land phase and the channel phase. The landphase is approached by an explicit moisture accounting lumped parameter model. The catchment area is divided into one or more segments, discharging to a channel. Within every segment areal homogeneity with respect to rainfall and the basin characteristics is assumed. The propagation and attenuation of flood waves in the channel can be simulated by hydrological routing methods.

8.4.3 Flow measurements :

The module comprises a number of techniques for validation and elaboration of flow measurements and rating curves, including:

i) entry and editing of flow velocity measurements, stage discharge data and rating curve parameters with cross-sectional parameters

ii) processing of flow velocity measurements by profile and moving boat methods, allowing:

- -various methods for measurments in the vertical
- -wet-and airline corrections
- -mean-and midsection method to compute the discharge
- -graphical and computational validation of measurements
- -transfer of condensed results to the data base

iii) Computation of stage-discharge relation given as parabolic and power type equations, with:

- -coefficients for up to 3 water level ranges per relation
- -corrections for backwater effects
- -corrections for unsteady flow
- -detailed error analysis
- -transfer of coefficients with validity period to the data base

iv) validation of rating equations for different periods and new measurements

v) extrapolation of rating curves:
   - computation of cross-section capacities and parameters
   - graphs of cross-sectional parameters versus stage
   - computation of synthetic stage-discharge data beyond the measured range

vi) State-discharge transformation, using:
   - rating curves stored in the data base
   - rating equation of measuring or control structure for critical and sub-critical flow conditions and variable sill level.

8.4.4 Data compilation

The data compilaton module comprises:

i) Aggregation and dis-aggregation of time series, where accumulative and instantaneous data are treated differently

ii) series transformation with various arithmetic transformation options

iii) minimum, mean and maximum series computation for selected time periods and transfer to the database

iv) Computation of areal rainfall by:
   - (weighted) average of point rainfall data
   - Thiessen method
   - kriging method

v) interpolation and computation of best linear estimates and uncertainties in areal quantities by point and block kriging method

vi) computation of potential evapotranspirataion, using:
   - Penman method
   - Pan-evaporation method
   - Christiansen method
   - Radiation method
   - Makkink method
   - Jensen-Haise method
   - Blaney-Criddle method
   - Mass Transfer method.

8.4.5 Statistical analysis

The module for statistical analysis includes:

i) computation of basic statistics and histograms

ii) fitting of distribution functions of the following type:
   - Normal distribution (with Box-Cox transformation to normality)
   - Log-normal distribution

-Exponential distribution
-Pearson-3 and general Pearson distribution
-Log-Pearson distribution
-Raleigh distribution
-Extreme Type 1 (Gumbel), 2 and 3 distribution
-Goodrich distribution
-Pareto distribution for peaks over threshold

The parameters are estimated using the method of moments and a mixed moment-maximum likelihood method. Once the distribution is fitted extremes for various exceedance probabilities can be computed, binomial, Kolmogorov-Smirnov and Chi-squares goodness of fit tests can be applied and a graphical display can be made of the fit of the distributions with confidence limits.

iii) statistical tables, i.e. computation of probabilities and variates for the various probability distributions

iv) generation of normal and gamma distributed random numbers

v) computation of IDF-curves (Intensity-Duration-Frequency curves) from monthly maximum precipitation in short intervals

vi) Computation and plotting of:
-frequency curves
-duration curves
-average duration curves.

8.4.6 Time series analysis

The time series analysis module comprises:

i) auto-and cross-correlation function computation

ii) spectral analysis

iii) run analysis:computation of up-and downcrossings, run lengths and runsum

iv) range analysis:computaton of range of cummulative departures from the mean

v) analysis of storage requirements by the sequent peak algorithm.

9.0 CONCLUSIONS AND RECOMMENDATIONS

(i) There is no centralized National Hydrological Data Bank in India.

(ii) Hydrological observing networks and their operation, field instrumentation, and data-transmission systems are maintained by different individual organisations like Central Govt. organisations and State Govt. agencies.

(iii) Number of organisations dealing with hydrological data are attempting to develop some computer based systems for data storage and retrieval.

(iv) Data processing procedures are adopting only for individual studies not at the time of recording the data. Most of the data are unprocessed and stored in the form of Manuscript or registers especially in State Govt. agencies.

(v) The development in the field of Hydrology is suffering from lack of Trained manpower for research, planning, developing computer softwares, and control on the available water resources.

(vi) An information system should be established with a network of data banks and data bases by integrating and strengthening the existing Central and State level agencies. It will help in improving the quality of data through better processing procedures.

(vii) The practice of preparing Hydrological data year books is yet to be followed by all the State Governments invariably.

ACKNOWLEDGEMENT

The author express his sincere thanks to Dr. S M Seth, Scientist 'F' and Project Coordinator, UNDP assisted project (RC Component) for his valuable suggestions in improving the report.

REFERENCES

1. Chow, V.T. (1964) 'Hand Book of Hydrology' McGraw Hill Book Co, INC, New York, NY; USA

2. Data processing in Meteorology WMO No.180. TP 90

3. Development of a process oriented calibration scheme for the HBV Hydrologic model. Nordic Hydrology 22, 1991, 15-36.

4. Delt Hydraulics (1991), Hydrological data processing and analysis software 'HYMOS' Lecture notes, WAMATRA II.

5. Elizabeth M.Shaw, (1983), Hydrology in practice. Van Nostrand Reinhold (UK)

6. Fortin, et al, (1991), Processing of Remotely sensed data to derive useful input data for the HYDROTEL (Hydrology model). Institute of National de la Recherche, Scientific, Queboc ontario, canada.

7. Gilliland J.A. 1972, principles of Ground water data acquisition. Water Resources Research, Vol.8, No 1, pp 182-87.

8. ISHERWOOD, W.L (1970), use of Digital computers in processing, publishing and Analysing Hydrologic data. "Nordic Hydrology", 1.

9. Proceedings of the second international symposium in Hydrology (1972), Decisions with inadequate Hydrologic data. Fort collins, Colarodo, USA.

10. Richard H, Mcluen (1985) Hydrologic analysis, and design. Deptt of Civil Engineering, University of Maryland.

11. Raghunadh H.M. (1985), Hydrology, (Planning, Design, Analysis).

12. Toebes, C and Ouryvaev.V, (1970) Representative and experimental basins, studies and reports in Hydrology No.4. UNESCO.

13. WMO (1971), Machine processing of Hydrometeorological data, Tech note No 115, WMO No 275.

14. Workshop on Data storage and retriveal system, 1987, (March 9-13) course material, NIH, Roorkee.

15. Workshop on processing and analysis of precipitation data, (1987-88) (course manual), NIH, Roorkee.

16. WMO 1974, Guide to Hydrologic practices, volume I, II and III, WMO No 168

17. WMO 1981, case studies of National Hydrological data banks (Planning Development and Organisation); WMO No 576.

DIRECTOR	SATISH CHANDRA
HEAD AND COORDINATOR	P V SEETHAPATHI
STUDY GROUP	Y R SATYAJI RAO