DATA STORAGE & RETRIEVAL SYSTEM

MRS. DEEPA CHALISGAONKAR*

INTRODUCTION

Hydrology is an applied natural science and therefore hydrologists have to deal with vast amount of data. Moreover, with the increase in population and industrial activities, water is gradually becoming a scarce commodity rather than a freely available natural resource. Due to this, the analysis of water resources systems is becoming more and more complicated and detailed and this demands, inter alia, more frequent observation of larger number of variables (1). It has been estimated that the annual volume of primary information received from a single gauging station is in the neighbourhood of 150,000 characters. Other observations such as water quality might produce between 300 to 600,000 characters. The summarization of these data in a water year book is a time consuming and voluminous job. Moreover, several types of statistical information is required for the analysis of hydrological systems, i.e. flood frequency determination, design flood estimation, low flow analysis and reservoir operations (8). The proper storage of these data in conventional way is next to impossible because of sheer volume of labour involved.

CONVENTIONAL WAYS OF DATA STORAGE

In the conventional way, data are mostly kept in registers or files and whenever demanded, they are manually copied and supplied. Usually, no data inventory is available and many times data are lost with the passage of time due to several reasons like physical damage to the storage media. With the introduction of digital computers, data handling is done more rapidly, easily and economically (1).

USE OF COMPUTER MEDIA FOR DATA STORAGE

Up to 1980, punch cards were mainly used for storing the data in computer compatible form but this type of data storage had a number of limitations like the cards may be damaged, they are inconvenient to carry, visual checking of the data is difficult and time consuming and proper sequence of the cards is to be maintained (1).

To avoid the problems associated with the cards, from 1982 onwards some organisations started storing their data on computer compatible magnetic storage media. In the simplest way, data are organized in sequential files which are stored on magnetic media. Whenever needed, the data are picked up from the media and are used. To access a specific data record from a sequential file, it is required to check each record in sequence until the desired one is encountered. Clearly, sequential access holds a distinct disadvantage for

*Scientist 'C', National Institute of Hydrology, Roorkee-247 667 (UP)

a user interested only in specific records. The contents of a record on a magnetic tape cannot be edited, nor can the records be added or deleted in between.

To overcome the drawbacks of magnetic tapes, magnetic disks are used as the storage devices. The magnetic disks are the direct access medium on which data are recorded in concentric circles and a read/write head is used to directly locate the place where the read/write operation is to be performed (1).

Now a days Computer based Database Management Techniques are being applied for efficient handling of Hydrological Data. It consists of creating a database, updating it frequently as and when new data are available and retrieving the data from the database when desired. This updating is done at some pre-defined frequency usually monthly or fortnightly. The data file is kept either on-line storage or off-line storage. In the on-line storage mode, the data files are available for operations at all the times. However, this mode of storage is slightly expensive. In the off-line mode the data files are stored on magnetic tapes. This mode of operation is cheaper than on-line and is mostly used to store the data (7).

A brief description of developments in India for storing various types of data using computerized system is given below:

Meteorological Data

India Meteorological Department (IMD) has been storing climatalogical data on punch cards for last four decades. The daily rainfall data were being punched in 31 cards format until 1970. These are being punched in 24 cards format since 1971 onwards. In 31 cards format, each card contains a catchment number, subdivision number, latitude and longitude of the station alongwith station number.

IMD has assigned a unique 3 digit catchment number to different catchments in India. In the 31 cards format, the catchment number, sub-division number and station number were recorded in each card. For each year, 31 cards were required and each card contained data for a specific data of each month.

For storage of data in 24 cards format, 2 records are needed for each month. The field in each record are catchment number, latitude, longitude, station number, year, month and 16/15 rainfall values.

A different schemes is used for storage of hourly rainfall data of the self recording raingauges. This format includes element code, index number of raingauge, state, year, month, date, card number (either 1 or 2) and hourly rainfall values. The second card also has field for amount and duration of maximum one hour precipitation during 24 hour period. These formats are shown in fig.1.

The daily rainfall data collected by state government agencies are published by respective state government. The data from IMD is also available on magnetic tapes in coded from (7).

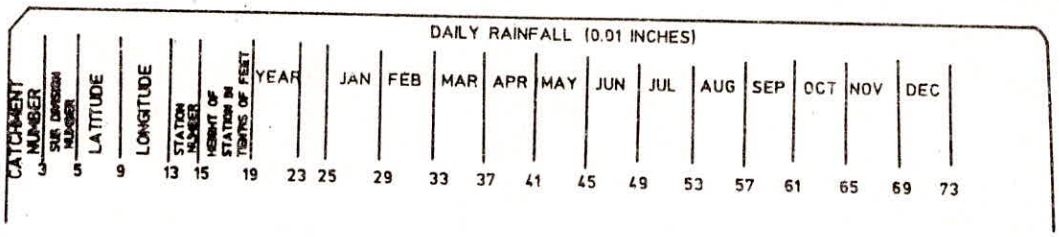

FIG. 1A: DAILY RAINFALL - 31 CARD FORMAT

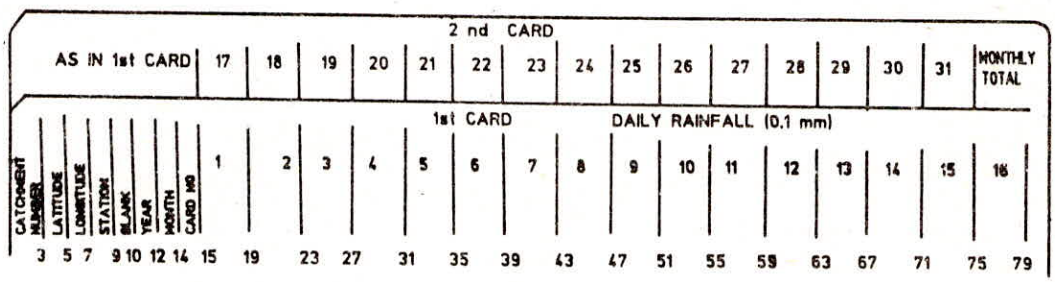

FIG. 1B DAILY RAINFALL - 24 CARD FORMAT

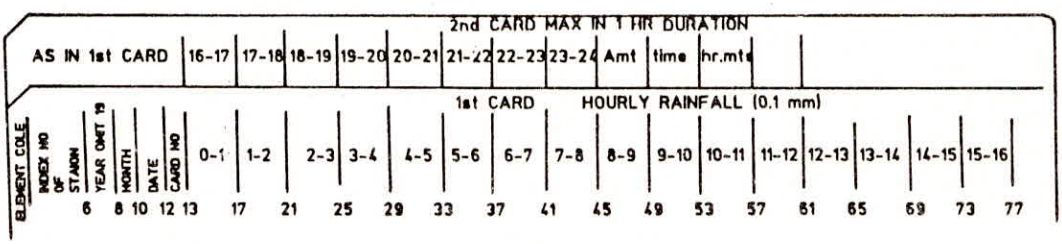

FIG. 1C HOURLY RAINFALL FORMAT

In 1991, a Data Acquisition, Storage and Retrieval System has been developed by National Institute of Hydrology under an Indo-Dutch Collaborative Project WAMATRA-II. This system collects the meteorological parameters like air temperature and rainfall along with other parameters without human intervention. The collected data is automatically stored on semiconductor memory of the data logger at the field which can be retrieved either through radio communication or through a battery operated Lap-top computer. The user of the data can print and/or process the data on the personal computer. He can also plot the data to visualize the variation of data with respect to time. However, this is not a general purpose system as it operates on specific hardware developed for this purpose.

Surface Water Data

Presently, the Central Water Commission (CWC), River Valley authorities and State Irrigation Departments collect and compile basic surface water data of the key-stations through out the country. On the basis of a resolution passed by CBIP in 1948, the CWC started publishing the surface water data of the important river basins in India in the form of water year books beginning from the year 1951. After some time the State government agencies also started publishing the water year books in respect of important river basins in the respective states(5).

With the increase in pace of developmental activities, it became essential to use the computers for storage and retrieval of surface water data. The River Data Directorate and Statistics Directorate of CWC took the responsibility for establishment of a data bank for systematic collection, processing, storage and retrieval of surface water data. For this, the data from the water year books formed the basic input of computerized hydrological data bank of statistics Directorate of CWC.

The data collected since inception of various river sites upto the year 1987 have already been computerized. The basic surface water data such as Gauge and Discharge, Sedimentation and water quality for about 250 river sites in various basins have been computerized using the computer hardware facilities consisting of Micro-32 computer system and NEC-S/1000 Super Computer System in addition to several personal computers provided by National Informatics Center (NIC). The Hydrological data bank is regularly updated with the addition of data for the latest available year. Several retrieval programs have been developed by them to generate outputs for different river sites/basins in respect of the followings, viz.,

i. Daily data on gauge and discharge
ii. Ten daily data on stage and discharge
iii. Monthly summaries along with respective averages for gauge and discharge.
iv. Daily, Ten daily and monthly outputs in terms of coarse, medium and fine particles for sedimentations.

Ground Water Data

At present, Central Ground Water Board (CGWB) and State Ground Water Departments are the main organisations involved in ground water data collection and storage.

Previously, hydrological surveys were being carried out since long in the country along with the geological surveys and investigations. The data collection of hydrogeological parameters on regional basis commenced in 1969 with establishment of hydrograph network stations all over the country (2).

With the growing demand of water, complexities in the hydrogeological frame work and data requirements, the data base expanded and variety of data was being collected and stored on regional and local basis. Keeping this in view, from 1984, CGWB started the work for computerized storage and retrieval of ground water data. An elaborate coding mechanism has been developed by CGWB. For this purpose codes have been assigned to each state, district and river basin in which the station lies. The information stored also includes the latitude of a well and a well number. Three character codes have been chosen for the geology of the well site and water and water quality parameters. Further, the lithology of the geological units have also been codified. The data has been stored on IBM-compatible personal computers using dBASE-III plus package, in the format shown in fig. 2 and Fig. 3 for easy storage and retrieval (7). The PC is linked to NEC-S/1000 and multi user possibilities of the system is being explored (3).

In 1988, National Institute of Hydrology has also developed a Ground Water Data Storage and Retrieval System using dBASE-III plus for U.P. Ground Water Investigation Organization (GWIO). As GWIO maintains data district wise, the data using the developed package is also stored district wise. Thus, to specify a particular well location, the name of the well station name, block name and district name have to be specified. Since the spellings of the district, block and station name may differ sometimes and also for achieving saving in the computer storage space, codes have been assigned to district name, block name and station name.

Addition of new data to the existing data can be achieved using simple commands. A menu driven interface has also been developed for ease of operation of the system (7).

The Data Acquisition, Storage and Retrieval System of NIH, described in meteorological data section, also stores some ground water parameters like soil temperature and soil suctions along with other parameters.

Water Quality Data

In India, the water quality monitoring, storage and retrieval are being attempted by mainly following agencies:

i. Central Water Commission
ii. Central Board for Prevention and Control of Water Pollution
iii. State Pollution Control Boards
iv. State Irrigation Departments
v. Central and State Ground Water Boards
vi. Ganga Project Directorate

The Central Water Commission has a large network quality monitoring stations. It monitors about 270 sites in India and water samples and sediment samples are collected at 10 days frequency also. The data is then recorded on prescribed format and then is

CENTRAL GROUND WATER
MASTER FILE
FINDEX

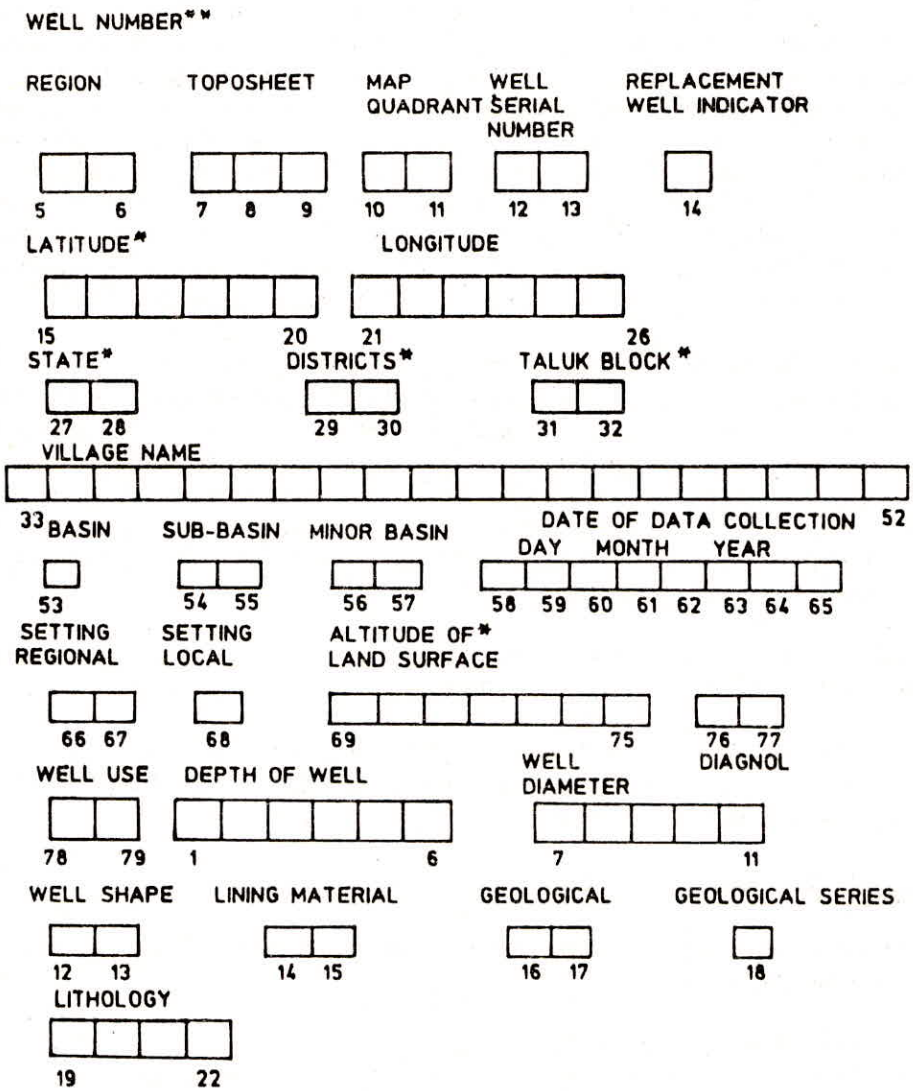

FIG. 2 FORMAT FOR GROUND WATER DATA STORAGE BY CGWB

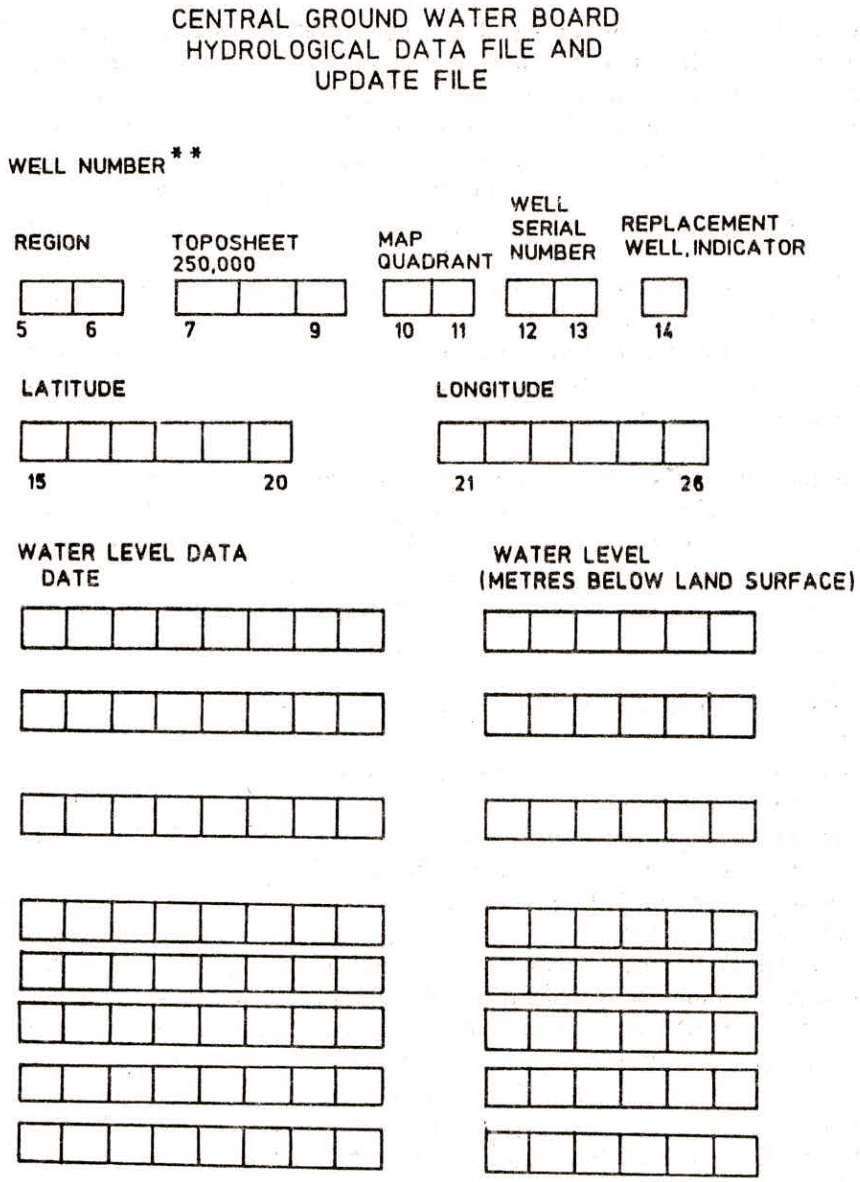

FIG. 3 FORMAT FOR GROUND WATER DATA STORAGE ADOPTED BY CGWB

transferred to Delhi for feeding to computer system in a standard format. They have also developed the program for retrieval of fortnightly data in respect of parameters such as pH, conductivity, TDS etc.

Central Pollution Control Board has adopted a three tier grid system of monitoring water quality of Indian rivers at Global level under Global Environmental Monitoring System (GEMS), at National level under Monitoring Indian National Aquatic Resources (MINARS) and at state level under Minimum Action Plan. The monitoring under these systems was started in 1979. The complete data collected and coordinated by CBPCWP, Delhi are fed to the IBM compatible PC in a standard format and a data storage and retrieval system is being used to store and retrieve data.

The Central Ground Water Board has identified about 8000 locations, where water quality samples are collected and data is analyzed. The data is being recorded on standard formats concerning the activities of Irrigation Departments, their main interest is to locate at the water quality from irrigation utility. The data at present is stored in the State in standard sheets and use of data storage and retrieval system is not there.

In the universities, sporadic studies are taken up for a specific research project. However, after the Ganga Action Plan has been taken up, 14 universities have been identified on the banks of Ganga and each has been given the responsibility of a stretch of the Ganga and the samples are regularly collected and detail analysis for about 30 parameters is carried out. The data are being stored on a data storage and retrieval system at Delhi.

National Resources Data Management System

It has been pointed out at various forums that the planning by Government of India can be made more effective if it is decentralized. Accordingly now a days more emphasis is being placed on district level planning. As an aid in this planning, a district wise data base is being created under the project NRDMS of department of science and technology. The main purpose of NRDMS is to make an assessment of natural resources, to identify the information required for various category of users, to evolve methodologies for computer based storage, processing and retrieval of information and to use it for the purpose of planning and development. The NRDMS project is working as a multi agency team. The approach adopted so far has been to assign various districts to various agencies. Water resosurces form an important part of the natural resources for which the data are being compiled and collected. The data being collected includes for each block/village, rainfall, no. of dug/tube wells, depth of water table, water quality in the wells, no. of tanks, rivers and streams in the block village, canals and minors and their discharge. The other data which are used in hydrological analysis includes the cropping pattern and irrigated area, and temperature (7).

ROLE OF NICNET IN DATA STORAGE AND RETRIEVAL

The existing system of information collection, compilation and inference not only took considerable amount of time but also did not have mechanism of correcting various pieces of information (data) due to non-standardization of collection and compilation procedures, which played an instrumental role in enormous delays. Hence a satellite based government

information network NICNET has been evolved by National Informatics Center to ensure a systematic procedure for information exchange among various departments, ministries of the state and central governments. To this end, NIC has established an inter city computer network to cover the entire country. They have provided seven terminals to CWC at Delhi which are connected to NEC-S/1000 system of NIC, thus providing an access to other places covered under NICNET. Similar Ministry of Water Resources, New Delhi and Central Soil and Material Research Station, New Delhi have been connected to NEC-S/1000 for access to NICNET to speedy retrieval of data through satellite. With the expansion of NICNET to the District level where the real data are generated, the capture of the rainfall, ground water and water use data could be carried out for the respective districts for onward transmission to the concerned state computer centres for the updation of the state hydrological data bank, which on further consolidation could be transmitted for the updation of National Hydrological Data Bank. The regional centres of CWC and CGWB covered through NICNET are shown in fig.4 and fig.5(3).

Central Ground Water Board,New Delhi
(CGWB Regional Network under the coverage of NICNET)

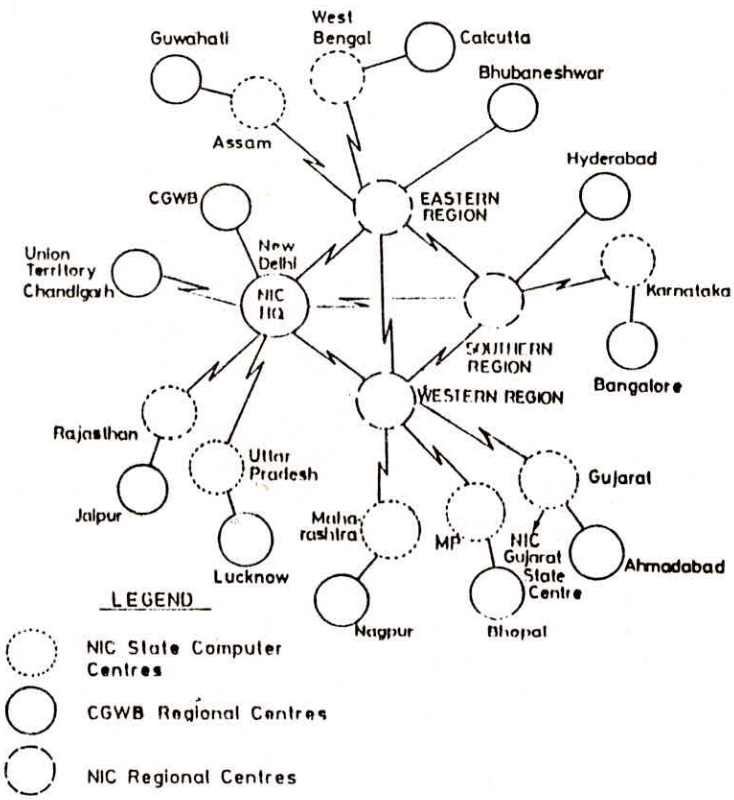

FIG. 4

Central Water Commission New Delhi
(Regional Centres under the coverage of NICNET)

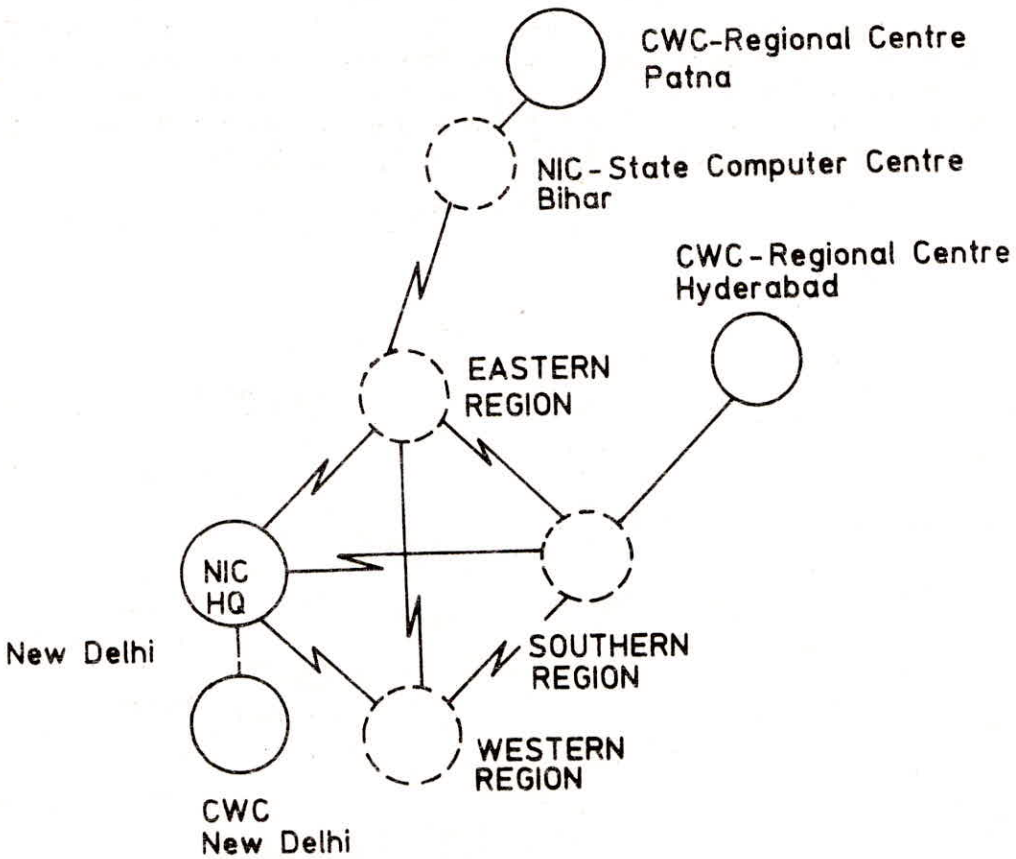

FIG. 5

REMARKS

The suggestions for the future developments in the area of data storage and retrieval may be :

1. Standardization of data formats
2. The data storage and retrieval system should be independent of the Computer Systems.
3. The system should include all type of hydrometeorological time series and spatial data required for hydrological studies.
4. The retrieval system should be capable of providing the data usually needed for most of the hydrological studies.

REFERENCES

Chalisgaonkar D et.al., ''Data storage and retrieval system on personal computer', Report No. UM-30, NIH, 1988-89.

Dutt D.K., 'Status of data base management system in Central Ground Water Board', Seminar on National Hydrological Data Bank, 19-20 May, 1988, Institution of Engineers (India), Bangalore.

Mehta N.K., 'NICNET in computerization of hydrological data', Seminar on National Hydrological Data Bank, 19-20 May, 1988, Institution of Engineers (India), Bangalore.

Parashar R.K. and Gupta S.K., 'Hydrological data processing, storing and retrieving system in the river data directorate of Central Water Commission', National Seminar on Use of Computers in Hydrology and Water Resources, 16-18 Dec., 1991, IIT, Delhi.

Rajgopalan K.S. and Mathew F.T., 'Need for Improved Access to Water Data', Seminar on National Hydrological Data Bank, 19-20 May, 1988, Institution of Engineers (India), Bangalore.

Sastri N.S. et.al., 'Use of computers of hydrological data processing, storage and retrieval systems', Seminar on Use of Computers in Hydrology and Water Resources, 16-18 Dec., 1991, IIT, Delhi.

Satish Chandra, 'Current status of hydrological data observation, recording, processing, storage and retrieval', Seminar on National Hydrological Data Bank, 19-20 May, 1988, Institution of Engineers (India), Bangalore.

UNESCO-WMO, 'Hydrologic Information System' Paris, Geneva, 1972.

Wisar E.H., 'HISARS Hydrological Information Storage and Retrieval System'.

WMO, 'Case studies of National Hydrological Data Banks', Operational Hydrology Report No.17, World Meteorological Organisation, Geneva, 1981.