A DATA STORAGE AND RETRIEVAL SYSTEM FOR GROUND WATER DATA

S.K.Jain D.Chalisgaonkar

National Institute of Hydrology,
Roorkee - 247667.

1. INTRODUCTION

Hydrology is an applied science and like other applied sciences, vast amount of data is handled by the investigators working in the area of hydrology. This data becomes the basic input for various hydrological designs and management activities like design of reservoirs and spillways, ground water exploitation, embankments for flood protection, management decisions for reservoir operation, flood control and crop production. It is well known that the quality of the design depends upon the quality of the input and that better and informed management decisions may be taken if accurate and extensive data is timely available.

Traditionally, the data measured by different field organizations is recorded in registers. In most of the cases, no inventory of the data is maintained and thus it is very difficult for the user agencies to find out whether the required data is available or not and if available then what is the location where it is available. Once the location is known, the data has to be manually copied from the registers. The process of obtaining the data is, thus very tedious, cumbersome, error prone, and time consuming. Computerized data bases are developed to overcome these drawbacks.

A data base is defined as a collection of inter-related data stored together with controlled redundancy to serve one or more applications. The data are stored in such a fashion that they are independent of the programmes which make use of them and a common and controlled approach is used for adding new data items, modifying and retrieving the required data from the database. Thus, a complete inventory of the data may be maintained. Moreover, whenever the data is needed by a user agency, it can be quickly supplied through magnetic tapes or print outs. Some consistency checks can also be applied to the data to detect the previous errors such as negative discharges.

2.0 DATABASE MODELS

The three major models used for organizing a data base are hierarchical, network and relational. A particular DSR system is generally designed to support only one of these models. The examples of each kind of model are shown in fig.1, fig.2 and fig.3 representing data regarding ground water level (GWL), a water quality parameter (WQ) and withdrawal (WD) observed at three different observations wells lying in one river basin. The The first model shows a hierarchical representation of the data. It is ordered by station name, then the observed values of GWL, WQ and WD. The hierarchical data structure, also called a tree structure, has a "one-to-many" relationship among the data records. This means that a higher level data record, called a parent (for example river basin) can be related to one or more of the next lower-level records (for example stations) each called a child. Thus, on the hierarchy in fig.1, the parent STATION1 has three children GWL1, WQ1 and WD1. Each parent can have one or more children, but each child can have only one parent.

A different way of structuring the data is possible in the network model using multiple links specifying relationships between data items. The network database structure allows "many-to-many" relationships among parent and child records as shown in figure 2. If in the Fig.1, the GWL at STATION1 is sâme as GWL at STATION3, this can be represented using network model as shown fig.2.

The third data base model is the relational model, and is based on presenting data in the form of tables or relations as shown in fig.3. Each horizontal row is a record and data items are shown in the column. The relational data base structure shows relationships among records by linking tables together as needed.

Each of these models has its strengths and weaknesses and needs to be matched to the data base requirements of an organization. The network model allows the addition of lateral connections to the tree. The major advantages of the hierarchical and network models are economy and speed. In general, these models in comparison to the relational model, have less data redundancy and allow faster access to the information. However, these data base models are complex to update since all affected links must be reconfigured.

The major advantage of the relational model is that any combination of data in the data base can be easily retrieved. Links between data records can be established by user's commands as the need arises. This great flexibility allows the relational data base to be easily configured to answer new and unanticipated questions. In contrast, the hierarchical and network models require that the links be built into the design of the data base. Programms must follow the established paths to find information. These paths are designed to quickly retrieve the desired information but retrieval may be slow for information that requires access not directly supported by the built-in links.

3.0 CHARACTERISTICS OF AN IDLE1 DATA STORAGE AND RRTRIEVAL SYSTEM

The immediate objective of Data Storage and Retrieval System is to make application development easier, cheaper, faster and more flexible. Some of the features of a Data Storage and Retrieval system are as follows:

a. Performance:- A DSR system designed for interactive use must give quick response.

b. Minimum Cost:- To keep the cost down, the data should be represented in such a way so that the total storage requirement is minimum.

c. Minimal Redundancy:- The DSR system should eliminate redundant data where it is economical to do and should control the inconsistencies that are caused by redundant data values.

d. Search Capability:- The DSR system should achieve fast and flexible search capability.

e. Integrity:- The storage, updating and retrieval procedures must be such as to avoid harm to the data. Range checks and other controls should detect data inaccuracies where possible.

f. Privacy and Security:- Data must be kept secure and private. The data must be protected from a person who may falsely update them. Unauthorized access to the data must be prevented.

In practice, no software package gives all the characteristics that an ideal Data Storage and Retrieval System should have and the designer has to select a compromise among different qualities.

4.0 STRUCTURE ADOPTED

The information regarding the structure of a record i.e. characteristics of each field in each record is shown in table -

Structure definition:-

Field name	Field type	Field size	Decimal
DISTRICT	Character	2	
BLOCK	Character	2	
STATION	Character	2	
WELLNO	Character	6	
LATITUDE	Character	6	
LONGITUDE	Character	6	
LEVEL	Numeric	7	3
JAN	Numeric	2	0
FEB	Numeric	6	2
MAR	Numeric	6	2
APR	Numeric	6	2
MAY	Numeric	6	2
JUN	Numeric	6	2
JUL	Numeric	6	2
AUG	Numeric	6	2
SEP	Numeric	6	2
OCT	Numeric	6	2
NOV	Numeric	6	2
DEC	Numeric	6	2

5.0 STORAGE AND RETRIEVAL OF GROUND WATER DATA

This data storage and retrieval system was developed for the computerization of the ground water data of Ground Water

Investigation Organization(GWIO), U.P.

In the present practice, the GWIO maintains their data district wise. A district is further subdivided into blocks and blocks in to stations. Thus to specify a particular station name, district name and block name and station name have to be specified. To be consistent with the practice being followed by them, this data storage and retrieval system has been developed using the relational data base model. Since the spelling of the district, block and station names differ sometimes and also for achieving saving in the computer storage space, two digit codes have been assigned to them.

The package dBASE III-PLUS has been used for the development of this system. The addition of new data to the existing data base and retrieval/modification of the existing data can be achieved using this menu driven interactive system.

6.0 CONCLUSION

An interactive menu driven data storage and retrieval system for ground water data has been developed on IBM compatible personal computers using dBASE III-PLUS. The same system can be used for the storage and retrieval of other hydrologic data with slight changes in the data structure.

The improvements which should be made to the system are:

i) Data security aspect should be incorporated to avoid the mishandling of the data

ii) Data validation checks should be provided.

iii) Ideally a hydrological data base must also contain information about river basin code. This was not incorporated as all the station lie in the same basin. Sub-basin information can, however, be built up.

iv) In many studies, it is also required. to have ground water data canal commandwise. This way be incorporated as an additional field.

v) Attempts are also being made to link a graphics package with the DSR.

7.0 REFERENCES

(i) Townsend,C., 'Mastering dBASE-III PLUS - A Structured Approach',1987

(ii) Ullman,J.D., 'Principles of Data base Systems', 1984

(iii) UNESCO, Hydrological Information System, 1972

(iv) Wisar,E.H., HISARS, Hydrological Information Storage and Retrieval System Reference Manul

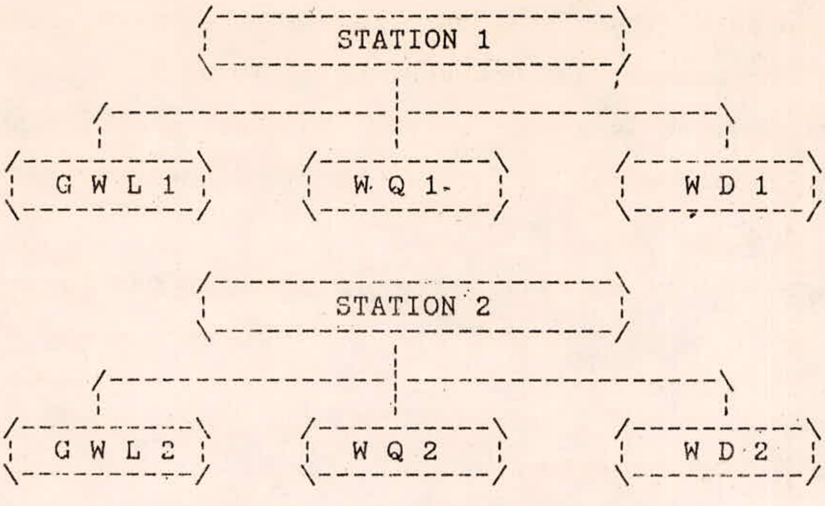

FIGURE 1. A HIERARCHICAL MODEL

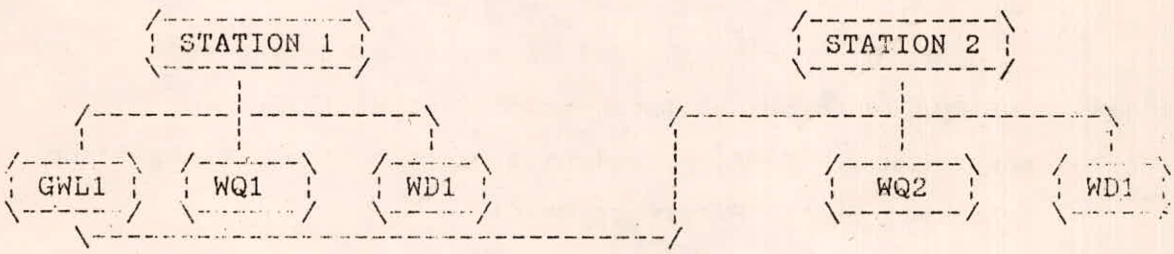

FIG. 2 A NETWORK MODEL

STATION1	GWL1	WQ1	WD1
STATION2	GWL2	WQ2	WD2
STATION3	GWL3	WQ3	WD3

FIG 3 A RELATIONAL MODEL