

HYQUAL

NATIONAL INSTITUTE OF HYDROLOGY

JAL VIGYAN BHAWAN

ROORKEE - 247667 (U.P.)

INDIA

1995-96

Preface

The data management is the very prime requirement for any type of research. To get predictions very close to the reality the true and exact data are required. A single data involves various information around itself such as: parameter name, place, date of sampling, date of analysis, name of sampler, name of analyst, purpose of sampling, time of sampling, MSL value of sampling point and climatic conditions at the time of sampling. Each and every data requires all such general information to get a real picture of nature. Once we get these general information regarding single data we step further to have various parameters of single sample. Any research activity requires various samples to be analysed or continuous monitoring.

Hence it becomes very tedious job to handle this bulk of data manually with a register. Therefore a computer based software may be very helpful to us for data management. The HYQUAL provides this facility and makes the job of data storage very easy to us. In this report the applicability of HYQUAL has been tested by Mr. Daya Ram Sc. 'B' of Environmental Hydrology Division.

S. M. SETH

Director

Content

| Title | Page no. |
|--|----------|
| List of figures | i |
| 1. INTRODUCTION | 1 |
| 1.1 Recording and Processing of Data | 1 |
| 1.2 Water Sample Information | 1 |
| 1.3 Field and Laboratory Results | 3 |
| 1.4 Water Quality Data Interpretation | 5 |
| 2. WORKING WITH THE HYQUAL | 7 |
| 2.1 The HYQUAL overview | 7 |
| 2.2 Starting and Logging on | 7 |
| 2.3 Setting up sampling point | 7 |
| 2.3.1 Define Region | 8 |
| 2.3.2 Define Cluster | 8 |
| 2.3.3 Determinand Types | 9 |
| 2.3.4 Define Sample Type | 9 |
| 2.3.5 Define Cluster details | 10 |
| 2.3.6 Sample Point Details | 10 |
| 2.3.7 Samplers, Analysts and Laboratories | 11 |
| 2.3.8 General Sample Information and Sample Editing | 11 |
| 2.3.9 The options menu | 12 |
| 2.3.10 Daily flow | 12 |

| | |
|---------------------------|----|
| 3. OUTPUT | 13 |
| 3.1 Tabular output | 13 |
| 3.2 Graphical output | 14 |
| 3.3 Time Series | 14 |
| 3.4 Multi site plots | 16 |
| 3.5 Scatter plots | 16 |
| 3.6 Bar Graphs | 17 |
| 4. DATA IMPORT AND EXPORT | 18 |
| 4.1 Importing data | 18 |
| 4.2 Exporting data | 19 |
| 5. CONCLUSION | 20 |

List of figures

| S.No. | Figure Caption | Page no. |
|-------|--|----------|
| 3.1 | Tabular representation of cluster report | 13 |
| 3.2 | Tree diagram of sample database | 14 |
| 3.3 | A time series graph | 15 |

1.1 Recording and Processing of Data

Analytical data obtained in water quality surveys should be filed in an orderly and systematic manner to ensure rapid access to the data when it is required by any water quality data user. Water quality data may be stored manually on report forms and files or automatically on magnetic tapes through the use of computers. The rapid development of electronic devices for processing of data has provided a capability for handling large volume of data quickly and efficiently.

1.2 Water Sample Information

The recording and storage of data collected in any water quality survey programme must meet a minimum set of requirement. For each sample, the following is a minimum set of items which should be recorded:

- Location of the sampling point
- Date of collection
- Project/study for which the sample had been collected
- Laboratory number
- Field and laboratory results and the methods used.

The sampling location point should be recorded as

geographical coordinates, i.e., latitude and longitude values.

The station number is used to identify a given sampling point in a concise form. Because of this, the station number must be unique, i.e., no two different sampling locations can have the same number, and also no one sampling point can have two different numbers. The simplest station numbering system would be a table in which sequential numbers starting with 1 (one) are assigned to sampling points as they are established.

In addition to the geographic coordinates and the station numbering system, the location of the sampling point should be described in a narrative. The narrative description contain the name of the river, lake, etc., together with landmarks which will help identify the sampling point.

When a sample is collected, the following information should be recorded: day, month, year and time (hours and minutes).

The project for which the sample has been collected may also be identified by a number. Each project number should have a separate file with the associated information: project name, stations to be sampled under the project, frequency of sampling at each station, parameters to be measured, and other information pertinent to the project.

For identification purpose, each sample on receipt at the laboratory as assigned a laboratory number. Each number must have a unique number and the number should be recorded both on the bottle(s) containing the sample and sample card.

1.3 Field and Laboratory Results

The following items should be recorded for field and laboratory results:

- Parameters measured
- Analytical methods used
- Results obtained.

In most systems the parameter which is measured and the analytical method which is used is reflected in the parameter code number. A 3-digit subfield indicating the parameter measured followed by a 2-digit sequential number for the method used is usually sufficient to handle the analytical tests for water quality data. The parameter code number thus have 5 digits.

Each parameter code number represents a specific analytical method for carrying out the measurement. This information is most important to water users of water quality data. A file of parameter code numbers in numerical order should be established.

There are various methods of recording results in a laboratory. The method described below is generally followed in most of the laboratories. The sample entering the laboratory for analysis is accompanied by a laboratory card. The card contains information about the sample location, date of sampling, the sample number, and a list of parameters for which the sample is to be analysed.

The analyst performing a certain test might choose record his results in a laboratory book, or in a table which have the following headings:

| Sample | Parameter Code | Parameter Code |
|--------|----------------|----------------|
| Number | Number 1 | Number 2 |
| | | |
| | | |
| | | |

The results may then be transferred from the book to the laboratory card at the end of the day.

Analytical results recorded on cards, magnetic tapes discs, etc., must be filed in a systematic manner to allow easy rapid access of the results. The data file may be organized in the following manner:

- Station numbers in numerical sequence; and
- Samples under each station in chronological order.

Once the data obtained, they must be filed in such a manner that they are easily retrievable in a format suitable for interpretation, report writing, plotting etc.

1.4 Water Quality Data Interpretation

Once water quality data have been assembled in a data storage system, the next step is to interpret the data with respect to specific questions, environmental problems and water resources management requirements. The most commonly asked questions about at any specific locations about water are:

- What is the water quality at any specific location
- What are the water quality trends in a particular region; is the quality improving or getting worse
- How do certain parameters relate with one another
- What are the total mass loadings of materials
- Is it possible to develop a water quality model based on the available data.

The general water quality conditions may be ascertained by scanning long column of tabulated data, the most common and reliable method is the statistical approach. A statistical analysis of the data provides a deeper understanding of the data and enables the water resources manager to arrive at better decision in water quality management.

There are several well known statistical techniques for examining water quality data, some of which are:

- Mean values
- Standard deviation
- Cumulative frequency distribution
- Mass loadings
- Graphical display
- Least square fitting to straight lines or curves
- Comparison of average values
- Comparison of cumulative distribution
- Q-Q plots
- Frequency analysis
- Power spectrum analysis.

2.

HYQUAL

2.1 The HYQUAL Overview:

HYQUAL is a water quality database for IBM and compatible personal computers running under Windows Version 3.0 or 3.1. It facilitates the storage of water quality data with tabular and graphical reports. It uses a hierarchical structure to store information. It contains following building blocks:

DETERMINAND --> SAMPLE --> SAMPLING POINT --> CLUSTER --> REGION

HYQUAL can be set up to provide standard determinand dictionaries for chemistry, microbiology, hydrology, sewage disposal, water supply and biology. Information once stored, can be accessed and used to produce tables of information or graphical plots.

2.2 Starting and Logging on

HYQUAL is started from within the Windows environment by double clicking HYQUAL icon. Select *System* from the menu by clicking on it or press *Alt-S*, then select *Logon* from sub-menu to display the dialog box for **User Information**. Enter a valid user name and password by clicking in the appropriate text box and then typing the required entry. Then click on appropriate user group category. Finally click on the *OK* button.

2.3 Setting up sampling point

Before we store any information, set up a structure.

To do this go through the following steps:

2.3.1 Define Region

Select the *System* from main menu and then choose *Regions* from sub-menu. The region is used to describe the geographic area from which the water samples are taken. Type region code of up to 9 characters and then region name of up to 29 characters long. The region code indexed alphabetically in HYQUAL. When the region code and name have been entered , select the *Save* command button. The procedure can be repeated until all the required regions have defined. The region name can be changed by double clicking on region code the region code can be changed by *Delete* command. Click on the *Exit* command button when all the entries have correctly entered. The output can be saved to a disk file by selecting *File* from the main menu and *Save as* from sub-menu. Close the output window before continuing.

2.3.2 Define Cluster

Cluster types provides a general description of a sample source, e.g. a lake , river or sewage works. Select *System* menu and choose *Cluster types* from the sub-menu to display Cluster Type Details dialog box. Enter a short code (up to 9 characters) and a full name (up to 29 characters) into the code and name text boxes, then select the *Save* command button. The cluster name can be changed by double clicking on cluster code, the cluster code can be edited by *Delete* command. Click on the

Exit command button when all the entries have correctly entered. The output can be saved to a disk file by selecting *File* from the main menu and *Save as* from sub-menu. Close the output window before continuing.

2.3.3 Determinand Types

Determinands are the physical properties of the sample being measured, e.g., oxygen content, alkalinity, temperature etc. Select *System* menu and choose *Determinands* from the sub-menu to display *Determinands Details* dialog box. Enter a short code (up to 9 characters) and a full name (up to 29 characters) into the code and name text boxes, then select the *Save* command button. The procedure is repeated for each entry. Standard upper and lower limits to sample values and levels can be included if required.

2.3.4 Define Sample Type

This describes the kind of measurements being made at a particular sampling point, e.g. quality, flow of rainfall etc. Select *System* menu and choose *Sample types* from the sub-menu to display *Sample Type Details* dialog box. Enter a short code (up to 9 characters) and a full name (up to 29 characters) into the code and names into respective text boxes, and click once, in the determinand list box, on each of the determinands required for the sample. To unselect a determinand click on it again. When all the determinands are selected click on the *Save* command button.

The procedure repeated for each sample type defined or edited and then click on the *exit* command button.

2.3.5 Define Cluster details

Details about each specific cluster can be entered after all the information about regions, cluster types, determinands and sample types have been defined and stored. Select *System* menu and choose *cluster* to open Cluster Details dialog box. This menu choice allows a specific cluster of sampling points to be described. Choose the Region first then Cluster Type from list boxes levelled Region and Cluster type by clicking on them, then enter a code and a name for the new cluster. Select *save* to complete the operation.

2.3.6 Sample Point Details

Select *System* from the main menu and *cluster* from sub-menu. First select the required cluster by double clicking on its code name in the cluster codes list box and then click on *New SP*. A dialog box named, Sampling Point Details will appear, this dialog allows the user to record a description of a sampling point. Select the *Save* command button. Repeat the process until all the required sample points have been set up then choose *Exit*. To edit the sample point details first select the region and cluster type from the cluster details dialog box.

2.3.7 Samplers, Analysts and Laboratories

The *System* menu offers three choices, *Samplers*, *Analysts* and *Laboratories*, that lead to dialog boxes for entering the relevant information.

2.3.8 General Sample Information and Sample Editing

When a sample point has been defined and saved, information about samples for that point can be entered. Choose *Input* from main menu and *New Data* from sub-menu to open the Select a Sample Point Dialog box. Use the appropriate list boxes to select a region, a cluster type, a cluster and a sampling point in that order then click on the *Sample* command button. A dialog box, named *General Sample Information* will appear having general information options for a sample. The cluster and sampling point codes are shown at the top, below these are text boxes for entry of a sample code, a sampling date and time and an analysis date and time. The sampling and analysis dates default to the last ones entered and the time defaults via the choice *Option* on the *System Menu*. The determinands for the sample are entered by clicking on the *Determinands* button to display a dialog box titled *Determinands Detail* for the Sample. After entering the determinands click on the *Save* button for that dialog box and again on *Save* button for the *General Sample Information* dialog box. Existing sample details can be edited using similar dialog boxes but accessing them through the *Input* choice on the main menu and the *Edit data* choice on the sub-menu.

2.3.9 The options menu

To get Data base options select *Options* from the system. This allows modification of the appearance of some of the default values and flags used by the database. If any value is missing for the sample determinands they will be marked by the word 'Missing' unless it is changed through Option dialog box. Similarly year and default time will be assumed unless altered here.

2.3.10 Daily flow

If daily flow is being measured at a sample point, this data may be stored in the HYQUAL database. From the main menu select *Input* then *New data* from the sub menu and the *Daily flow* button from the *select a sampling point* dialog box. A dialog box called *Daily flow edit* will be displayed. Date and flow values can be entered. Each value must be terminated by using the *Enter* key (Note: use of the "daily flow" check box is reserved for future release of HYQUAL; it is inoperative in version 1.0).

HYQUAL provides two types of data presentation, one is tabular form and other one is graphical presentation of data.

3.1 Tabular output

Tabular output is obtained through the *Output* option on main menu. The first four items on the sub menu *Cluster*, *Sample point*, *Sample* and *daily flow* are able to provide tabular output. Each of these leads to dialog boxes that can be used to regulate the amount of information that will appear in the report. Figure 3.1 shows an example where cluster has been presented in tabular form.

HYQUAL V1.0

Institute of Hydrology

Cluster Report

| Code | Cl-Type | Region | Name |
|---------|---------|--------|---------|
| LAKE Q | LAKE | SOUTH | Lake Q |
| MINE I | MINE | NORTH | Mine I |
| MINE J | MINE | NORTH | Mine J |
| RIVER X | RIVER | EAST | River X |
| RIVER Y | RIVER | EAST | River Y |

 Institute of Hydrology

Fri Jul 16 10:42:49 1993

Figure 3.1 Tabular representation of cluster report.

There is one other tabular report obtainable from the *Database structure* option on the *output* menu. This produces a simple tree showing the relationship between regions, cluster types, cluster and sample points. For example the part of tree is shown in the figure 3.2

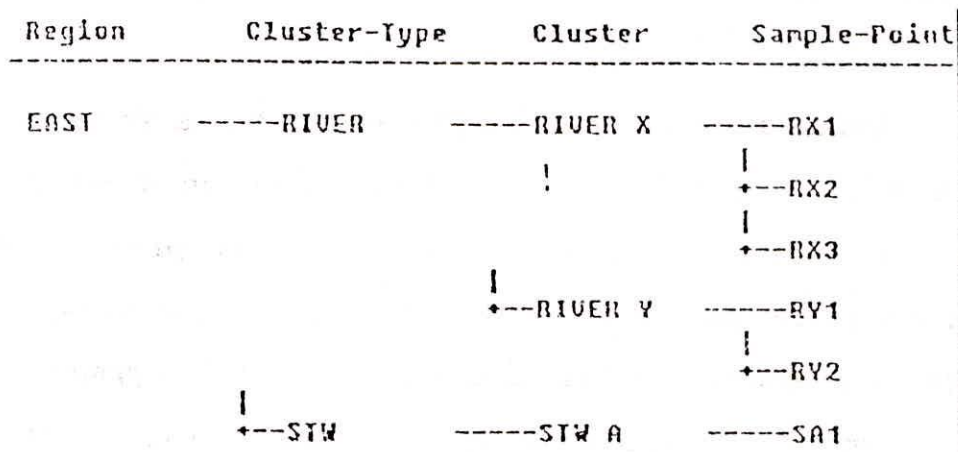


Figure 3.2 Tree diagram of sample database

3.2 Graphical output

Graphs are produced through *Graphics* option of *Output* menu. Four choices offered in graphical output are *Time series*, *Multisites*, *Scatter* and *Bar graphs*. With the HYQUAL Graphs can display a maximum of 500 years of information and a minimum of 20 minutes.

3.3 Time Series

It provides a visual description of the change of levels of determinands with time. Time is plotted on X-axis and

determinands on Y-axis. It presents two Y-axis, Y1 and Y2, the Y1 axis will be plotted on right side of the graph, and the Y2 axis on the left side. Figure 3.3 shows the type of the time series graph produced.

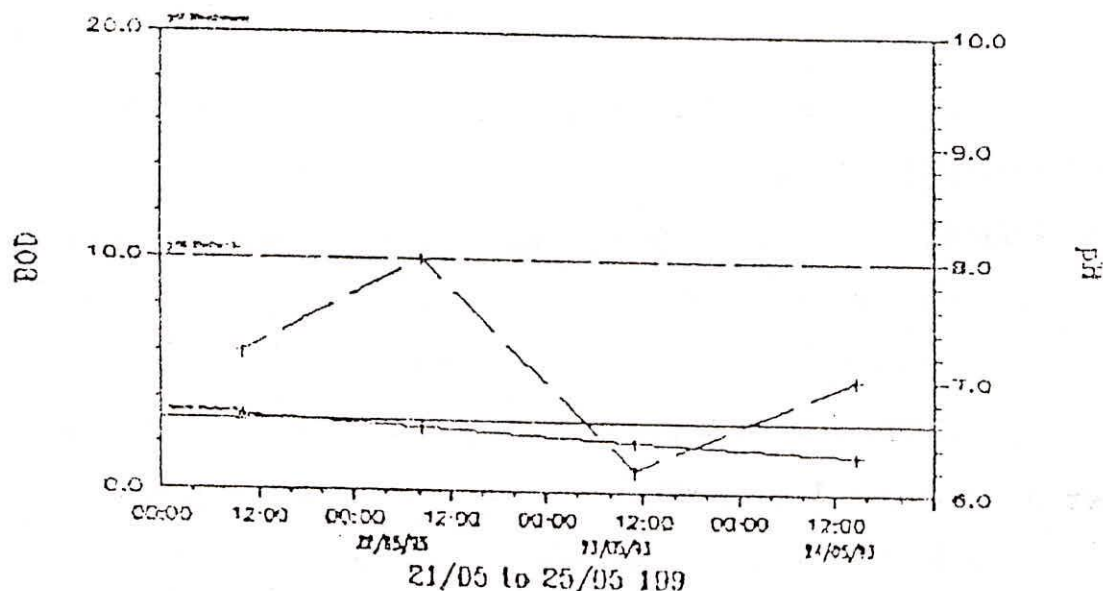


Figure 3.3 A time series graph

The appearance of graphs can be altered by setting a range of parameters accessible through the *Graph define* button. The various options provided in graph definition option are Titles, axis information, option and others. It also has a option to identify missing value with standard missing symbol (an Upturned empty triangle).

The time series graph is extremely flexible, and can

be used to plot any number of determinands in one plot, from the same or different sample points. This will allow the comparison of different determinands, with time, at one point, and also the comparison of same determinand at different sample point with time, as well as the combination of the two.

3.4 Multi site plots

It is possible to plot information from many sample points using the *Multi sites* option the *Graphics* sub-menu. The Graph definition dialog box for multi-site plots is similar to that for time series plots but there are some optional variation described below.

Log-X, **Log-y1**, and **Log-Y2** replaces the safe level and upper and lower limit options in the time series plots.

HLC is an important option for multi sites plots. If HLC is switched on, the statistics (highest, lowest and average) of the determinands for the period requested are calculated, and a bar is drawn for the sample point displaying these three phases. If HLC is switched off only the average of the determinands for the period is calculated, and these are plotted as a histogram for each of the sample points.

3.5 Scatter plots

This option allows the plotting of many determinands values against one dependent determinand from the same sample point. A linear regression is carried out between each

independent and dependent determinands and this can be displayed on the plots. The plots may also be drawn on logarithmic scale.

3.6 Bar Graphs

Bar graphs are intended to provide a time series histogram plotting one determinand at a single sample point, in either two or three dimensions.

3D Bar: The histogram may be plotted as either a 2D or 3D representation using this switch.

4.

DATA IMPORT AND EXPORT

4.1 Importing data

The HYQUAL has a facility to import data from any other software using *Import data* option of *Input* option in main menu.

The ASCII file to be imported has to be formatted in a particular way, this may be explained with example:

TESTIMP.DAT

SAMPLES

//

#C1 ANGLIA abstract cambridge

//

#SP cam0001

//

#FF TRUE TRUE TRUE TS

//

#DT BOC BOD Fe Test Cu S PB

//

DU Date ml/l mg/l mg/l mg/l mg/l mg/l mg/l

//Finally the may be entered as follow:

//

12/03/96 14.3g 2.31 44.4s 7.8m 18.6n 2.3g 3.0n

13/03/96 25.2 12.5 52.6 7.8n 28.6m 2.3g 3.0g

#FE

The explanation of formate is as follow:

// represents comment lines as required by user.

represents command lines

SAMPLES states that

Data file title: SAMPLES

#CL ANGLIA abstract cambridge

Region : ANGLIA

Clust type : ABSTRACT

Clust : CAMBRIDGE

The fourth command line identified by #ff command.

#FF TRUE TRUE TRUE TS

Use default time :TRUE

Data follow by flag :TRUE

Generate sample code :TRUE

Sample prefix :TS

The next command line should contain the #DT command followed by the list of determinands in the import file.

The #DU command line then identifies the Date and Time together with a list of the units of the determinands. The end of the file is then identified with the #FE command.

4.2 Exporting data

The HYQUAL has a option to export determinand values in a standard format. The *Export* choice on the *output* sub-menu is designed to facilitate the production of an ASCII file with determinand values for samples at a valid sample point.

A single data may be identified by a address assigned to it. Before assigning a numerical value to a data, it is necessary to know about its address where it has to go. The HYQUAL defines a unique address to each and every data/measurement. If there is continuous monitoring of data with time, then the HYQUAL may be the good tool to manage the bulk of data. The time is the prime factor in HYQUAL. It has graphical as well as tabular ways of data representation. The HYQUAL also useful to represent data even if few data are missing in the series. In the time series graph data variation graph is traced with time, which based on intrapolation concept. Therefore even if few data are missing from the series, the trend may be find out and the expected value of the data may be obtained. The HYQUAL does not help if inadequacy is regarding sampling date or time. In this it is not possible to place such data at its right place.

DIRECTOR

S M SETH

DIVISIONAL HEAD

K K S BHATIA

PREPARED BY

DAYA RAM