HYDROLOGY PROJECT
World Bank & Government of The Netherlands funded

Training module # SWDP - 51

How to back-up SW data at various DPCs

New Delhi, February 2002

CSMRS Building, 4th Floor, Olof Palme Marg, Hauz Khas,
New Delhi – 11 00 16 India
Tel: 68 61 681 / 84 Fax: (+ 91 11) 68 61 685
E-Mail: hydrologyproject@vsnl.com

DHV Consultants BV & DELFT HYDRAULICS

with
HALCROW, TAHAL, CES, ORG & JPS

Table of contents

Page

1. Module context

While designing a training course, the relationship between this module and the others, would be maintained by keeping them close together in the syllabus and place them in a logical sequence. The actual selection of the topics and the depth of training would, of course, depend on the training needs of the participants, i.e. their knowledge level and skills performance upon the start of the course.

2. Module profile

Title	:	How to back-up SW data at various DPCs
Target group	:	HIS function(s): ……
Duration	:	x session of y min
Objectives	:	After the training the participants will be able to*:*
Key concepts	:	•
Training methods	:	Lecture, exercises
Training tools required	:	Board, flipchart
Handouts	:	As provided in this module
Further reading and references	:	

3. Session plan

No	Activities	Time	Tools
1	***Preparations***		
2	***Introduction***:	min	OHS x
	Exercise	min	
	Wrap up	min	

4. Overhead/flipchart master

5. Handout

Add copy of the main text in chapter 7, for all participants

6. Additional handout

These handouts are distributed during delivery and contain test questions, answers to questions, special worksheets, optional information, and other matters you would not like to be seen in the regular handouts.

It is a good practice to pre-punch these additional handouts, so the participants can easily insert them in the main handout folder.

7. Main text

Contents

How to back-up SW data at various DPCs

1 Backup of hymos databases

1.1 Backup and its advantages

Any information stored on the computer or on a disk is vulnerable to damage or loss for a variety of reasons like theft, fire, wear and tear, computer viruses, power failure, magnetic fields etc. One of the major threats is the user himself – a single wrong command can destroy months of effort. Therefore, backup of data becomes critical and has to be performed regularly. A backup routine has not been provided as in integral part of either the SWDES or the HYMOS system because of the number of options available in the WINDOWS environment and the extremely easy availability of general purpose software in the public domain.

Various levels of backup usually performed are:

Local backup: The first level of backup can be kept on the same computer, preferably in a different partition.

On-line backup: The second level of backup can be kept on the hard disk of another computer, if connected on a network.

Off-line backup: This is the most important backup and serves as the permanent archive of data. This could be on floppies, CD's, magnetic tapes or other devices, depending on the hardware configuration of the computer and the facilities available in the data centre.

Off-line incremental backup: This option is useful for very large volumes of periodic data.

In this procedure only that information which has changed after the last backup is backed up and thus saves lot of space.

Off-line backup can be taken on floppies, CDs., cartridge tape or DAT and JAZZ drives etc.. Typically, the capacity of floppies is 1.44 MB, of CDs is 650 MB and of DAT cartridges is between 2 and 24GB. Backup of a large file(s) (say upto about 6-8 MB) can be spanned on multiple floppies, if required. However, since floppies normally are more prone to corruption and become unreadable frequently, backup of information on a set of multiple floppies is difficult to work with and thus is not preferred. It becomes quick and easy though to backup smaller file(s) on one or two floppies.

CDs and tapes on the other hand have substantial storage space and are preferred methods for backing up large amount of information. The availability of such specialised backup tools and the higher technical skill required for operation and maintenance may limit their application in some cases.

Backup on CD's, cartridge tapes etc. will be done using the software made available with the backup device. All procedure for CD-ROM drives, cartridge tapes, ZIP drives, DAT and JAZZ drives are different and equipment specific, and thus would not be conducive to a system based back-up protocol.

It is self defeating to take backups on the same piece of physical removable media cycle after cycle. This leads to the following problems:

- Corruption of data because of media failure : Loss of previous backup also
- Loss of backup in case undetected virus in current backup
- All corruption, errors in previous data made in the current cycle transmitted to the backup

To safeguard against these problems, it is suggested that the Grandfather..father..son technique is used, whereby backups are taken cyclically on three different sets of media. This ensures that at any point of time, at least two different secure backups are always available. It also provides for redundancy in the backup system and for checking on data sanctity and validity.

Restoring Data: The backup system followed will always have a corresponding restore utility. In case a restore is required it must be ensured that it is done such that no existing valid data is over-written.

Procedure to be followed for backing up surface water data as available in SWDES and HYMOS databases at various Data Processing Centres is discussed in the following sections. However, similar policy can be followed for any other important data being worked with at these offices which obviously also needs backing up at suitable time interval.

1.2 Employing compaction and compression before backing up of data

The databases grow each time data is added to them. However, it is important to note that when data is deleted, they do not become smaller automatically. This is done as a standard practice in most ***RDBMS*** and other databases to provide faster response times. Thus, it becomes important to 'compact' the database periodically as the database keeps deleting its own temporary and work data. Compaction reduces the size of the database and will therefore help in creating smaller backups. It is therefore recommended that a Microsoft Access database be compacted frequently, especially prior to a backup process.

Another tool commonly available is data compression. Compression reduces file sizes for the purpose of storage by "packing" the data in a smaller space. Since a compression process reduces the size of storage required for backup it is always useful when the amount of data to backed up is larger and at the same time the receiving media is having relatively lesser space. Such is case when the backup of few moderately sized (say 6 to 8 MB) files is attempted to taken on the floppies. There are several utilities available for compressing the files such as WINZIP, Microsoft backup and other compression utilities like LHA, RAR etc. Most of these utilities also support simultaneous process of compression and backup functions (with backup spanning on multiple floppies).

While taking backups on floppies it is very important to ensure that all backup on floppies are tested by copying the files back into a temporary directory after ejecting and reinserting the floppy diskettes!!!

1.3 Backup policy at SDDPCs

SDDPCs will routinely work with a maximum of three SWDES workareas (in many cases there will be only one workarea) only. Each SWDES workarea has only one physical file (with ".MDB" as the extension) that contains all the data; i.e.; each work-area file is complete in itself. Only where the DWLR data would also be dealt, there will be an additional file of the same name as the workarea itself but with ".MDD" as the extension.

Thus the regular work at the SDDPCs will be on few SWDES workarea files. Since these files would contain data for about 4 to 5 years, they will be about 8 –10 MB in size. Compaction and compression can bring them to the size of one or two floppies.

Off-line back up at SDDPCs: Since the data files will be of comparatively smaller size, it might be a simpler and quicker option to take an off-line back up the data on floppies. Such off-line back up is recommended to be taken every 10 days. It is appropriate to fix the dates of such back ups as 1st, 11th and 21st of every month or the next working day in case it happens to be a holiday. In case the staff at any SDDPC feels well acquainted with taking back up on CDs then it can be the preferred medium as there will be less chances of corruption etc.

On-line back up at SDDPCs: As a second level of safety, it is appropriate to keep the copies of content of datafiles on the second computer available at the SDDPCs. This can be easily done by using the same set of back up taken on floppies or CD, every 10 days, to restore on the second computer. This would incidentally also check if the floppies are working properly.

1.4 Backup policy at DDPCs

At DDPCs, both SWDES and HYMOS databases will be in use. SWDES databases are as received from SDDPCs and thereafter consolidated into unified databases for each SDDPC. HYMOS database(s) will be operational of each of the river basin(s) or part thereof within the jurisdiction of the DDPCs. Normally, a DDPC will have one HYMOS database but in some cases it may be 2-3 as well. There will also be availability of “HYMOS transfer databases” as data from DDPCs HYMOS database will be transferred regularly to the SDPCs/RDPCs.

Thus four types of databases available at DDPCs, on a regular basis, are: (a) SWDES fragmented databases as received from SDDPCs for incremental data for each month, (b) unified SWDES databases for each SDDPCs in which all incremental SWDES databases are consolidated, (c) HYMOS databases and (d) incremental processed “HYMOS transfer databases” as sent to SDPC/RDPC every month. Databases of the first and the last types are not important to be backed up as the information is already available in the unified SWDES databases or the HYMOS databases.

It is desired that all the four types of databases be organised in well-defined folders. For example, all the SWDES fragmented databases received from various SDDPCs every month can be organised in a folder called “FRAG_SWDESDB”. Under such folder the databases can be organised under separate directories for each SDDPC separately. Similarly, all the SWDES unified databases for each SDDPC would available in the SWDES program directory. All HYMOS databases will similarly be available under the “HYMOS Databases” directory with one more sub-directory level to identify different databases. All the HYMOS transfer databases, exported from DDPC and to be sent to SDPC/RDPC, can be stored in a prescribed directory as “TRAN_HYMOSDB”. Databases of the first and the last types are not important to be backed up as the information is already available in the unified SWDES databases or the HYMOS databases respectively.

Off-line back up at DDPCs: It would be very convenient to back up the contents of the SWDES and HYMOS databases from the two directories on the CDs as the amount of data would be difficult to handle by using floppies. Regular off-line backups of these databases at the DDPCs is recommended to be taken regularly at a ten daily interval as in the case of SDDPCs.

On-line back up at DDPCs: In case the computers available at DDPC are interconnected with the help of network or other link, it would be useful and easy to copy the contents of the

four folders on to the other computer. If SWDES and HYMOS work is carried on 2-3 computers available at the DDPC then such copying from one computer to another computer must be done for all computers so that every computer keeps a back up of useful data of other computers.

1.5 Backup policy at SDPCs/RDPCs

At SDPCs/RDPCs there will be availability of full-fledged network and also 4-6 computer nodes working with comparatively larger amounts of data.

At SDPCs/RDPCs also, both SWDES and HYMOS databases will be in use. SWDES databases are as received from DDPCs and thereafter consolidated into respective databases for each SDDPC. HYMOS database(s) will be operational of each of the river basin(s) or part thereof within the State. There will also be availability of "HYMOS transfer databases" as data from DDPCs will be received regularly at the SDPCs/RDPCs.

Four types of databases available at SDPCs/RDPCs, on a regular basis, are: (a) SWDES fragmented databases as received from DDPCs for incremental data for each month, (b) unified SWDES databases for each SDDPCs in which all incremental SWDES databases are consolidated, (c) HYMOS databases and (d) incremental processed as "HYMOS transfer databases" as received from DDPCs every month. Databases of the first and the last types are not important to be backed up as the information is already available in the consolidated SWDES databases or the HYMOS databases.

On-line back up at SDPCs/RDPCs: As all the computers at SDPC/RDPC will be connected with the network it would be useful and easy to take a weekly backup of important data on all the computers on the server computer.

Off-line back up at SDPCs/RDPCs: As the data at SDPC/RDPC will be very important, it is essential to keep a off-line back up on CDs or DAT drives as well. As a regular activity after the on-line back up is taken it is appropriate to ensure the off-line backups as well. The frequency