TRAINING COURSE

ON

SOFTWARE FOR GROUNDWATER DATA MANAGEMENT

UNDER

WORLD BANK FUNDED HYDROLOGY PROJECT

LECTURE NOTES
ON

APPLICATION OF HYMOS SOFTWARE

BY

HEMANT CHOWDHARY

ORGANISED BY

NATIONAL INSTITUTE OF HYDROLOGY
ROORKEE - 247 667
INDIA

SYSTEM FOR STORAGE, PROCESSING AND RETRIEVAL OF HYDROLOGICAL DATA (HYMOS)

1.0 GENERAL 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. It 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. It runs on stand alone computers, but can also be used in a network system. Securities have been built in to restrict the access 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.

Here, an overall view of HYMOS is given, the type of data it can handle, the way the data can be stored and processed and the hardware required to run the system successfully.

1.1 HYMOS in a Nutshell

1.1.1 Structure of HYMOS

HYMOS integrates the distinctive phases in the processing of hydrological data. The activities are carried out in specific processing modules; each module consists of a number of programs enclosing particular compilations/computations. The modules are structured according to a logical sequence of activities in data processing. All modules are linked to the HYMOS database, structured and controlled by a database management system. Different databases can be operational under HYMOS. In summary, HYMOS comprises of following systems:

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

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.

The structure of HYMOS is shown schematically in Fig. 1.

HYMOS allows user programs to communicate with the database. For that purpose the FORTRAN subroutine library HYSUB is enclosed with the package. HYSUB comprises an extensive set of subroutines to store and retrieve data and to carry out statistical analysis.

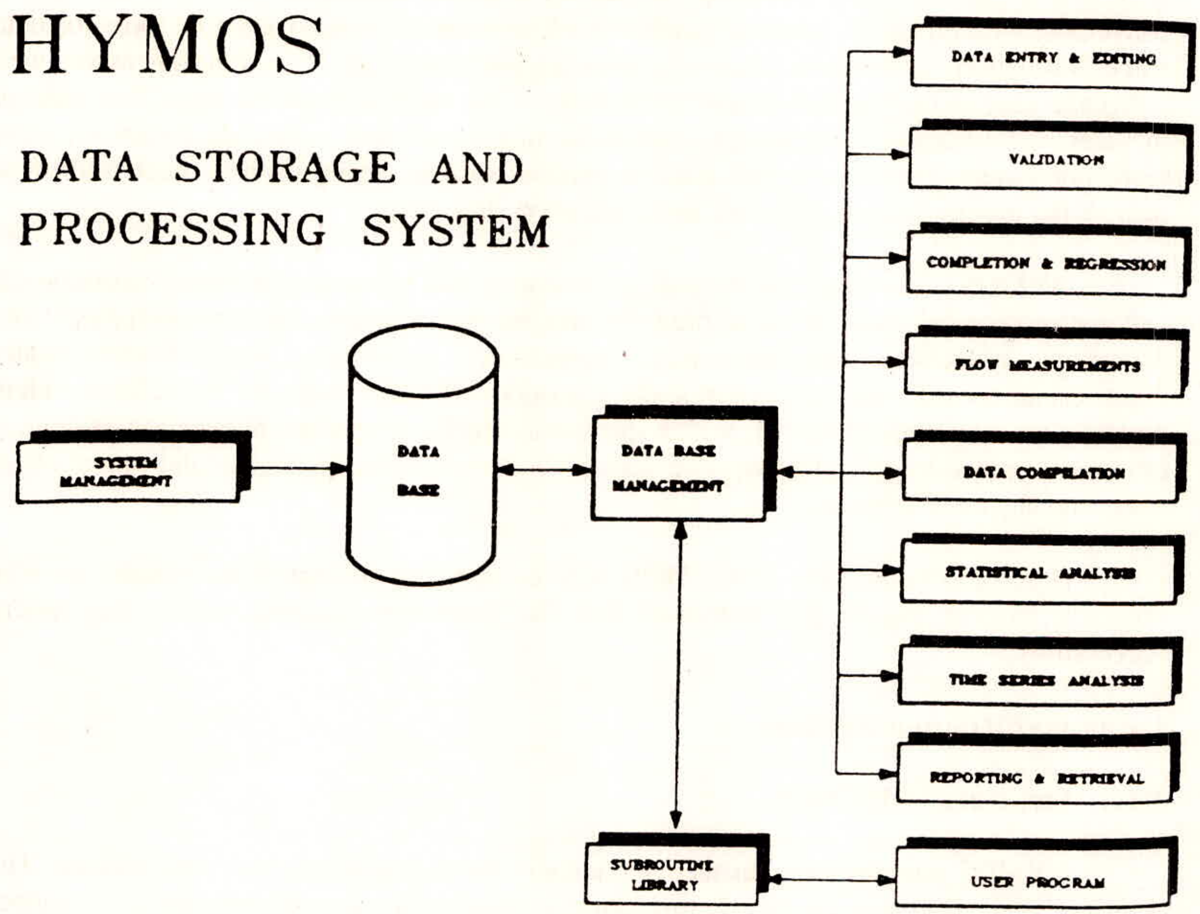

Fig. 1 : Structure of HYMOS

1.1.2 Data types

The types of data, handled by HYMOS, can be categorized in the following groups.

1. Space-oriented data, covering:
   - catchment characteristics
   - station particulars
   - station histories
   - geo-hydrological profiles

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.

3. Relation-oriented data,
   - stage-discharge data divided in: relation or rating curves parameters, valid for a certain time period.

1.1.3 Database management

Database management deals with the creation and structuring of data bases and with the definition of user identifications.

Structure of databases:

A HYMOS created database comprises the hydro- meteorological data of one or more catchments, or sub- catchment(s). As many databases as required can be used. Each database is stored in a separate directory.

In a particular database the data are, except for the catchment characteristics, structured station-wise. Data type and time interval are used for further identification of station data. The overall structure is shown in the following scheme:

		Station particulars
		Station history
	Station a	Geo-hydrological
profiles		
Catchment j ----------------	Station b -----------	Equidistant time
series		
	Station c	Non-equi. time series
		Stage discharge data
		Relation curve
parameter		
		Rating curve
parameter		

Definition of users:

User identifications can be categorized into three parts:

1. group name, 2. user name, and 3. password.

The access to a particular database is reserved to specified groups. Within a group users are defined with different levels of authority, ranging from access to all facilities to data retrieval operations only. Each user must enter a unique password, which can only be changed by the System Manager.

1.1.4 Data storage and retrieval

Under HYMOS the data are stored in dBase look alike files (space-oriented data, non-equidistant time series and relation-oriented data) and in dedicated direct-access files (equidistant time series). Hardly any limits exist to the size of a database as may be seen under section of 'Data Limitations'.

1.1.5 Data processing

The validation, completion and analysis of hydro- meteorological data under HYMOS is logically structured in the following data processing modules:

- → data validation:
- → data completion and regression:
- → flow measurement:
- → data compilation:
- → statistical analysis:
- → time series analysis:

1.1.6 Special features

To facilitate the use of HYMOS special function keys are available for on-line help, display of temporary output files, use of text editors, change of data base and packing of data base files.

The station and series selection is enhanced by selection from displayed tables or catchment maps on the screen.

Finally, keys are available to set the screen colours and hardware configuration, and to switch to DOS during a HYMOS session.

1.1.7 Computer Configuration Requirements

The HYMOS software package is written in the programming languages FORTRAN 77 and C. The HYMOS version 3.0 runs on computer systems of the following specifications:

. IBM XT, AT or compatible personal computer, with:
  . 640 kb core memory,
  . > 20 Mb hard disk,
  . EGA or VGA graphics card,
  . MS-DOS 3.* (or higher) operating system,
. Printer
. HP or Calcomp compatible plotter;
. tape streamer (recommended for back-up purposes):
. protection key HASP II (Supplied by the developer)

For fast and pleasant operation an AT computer with color monitor is preferred but not strictly necessary, nor is a co-processor required. Graphics cards less advanced than EGA or VGA do not support all HYMOS graphics, but they will be acceptable for common data processing activities.

1.1.8 Database limitations

It was mentioned earlier that the equidistant time series are stored in special direct access files. Some limitations exist as to the amount of data, the number of series and the length of the series that can be stored in these files. The following limitations apply to a particular data base.

- ➩ the amount of data is at maximum 1.8 billion (i.e. 7 Gigabyte);
- ➩ the amount of equidistant time series is at maximum 10,000;
- ➩ all series can have a length of at least 54,500 data (e.g. about 150 years of daily data) but, in addition, some 40 million data positions are available for extension of one or a number of series.

The last observation allows one to use for a limited number of series a length far beyond the standard of 54,500 data, e.g. 25 series of 1.6 million data (i.e. 45 years of 15 minute data) can be used additionally. For non-equidistant and non time series data no limits exist, but the size of the disk.

If the above limits are exceeded, more databases have to be defined, e.g. a temporary data base with a few permanent data bases, where the latter are categorized according to location, data type or time interval, depending on the objectives of the Information System.

2.0 DATA STORAGE, RETRIEVAL AND PROCESSING

Data processing covers all activities after transmission up to and inclusive data dissemination. It includes:

2.1 Data Storage and Retrieval System

Data storage and retrieval activities are comprised in two modules:

A. Entry and editing
B. Reporting and retrieval

A. Entry and editing

The entry & editing module includes following options:

- * entry, editing and graphical display of catchment data
- * creation of stations and series, and entry and editing of station particulars and

log books.

* 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.
* full-screen editors for editing and display of equidistant and non-equidistant time series, with on-line graphical display
* entry and editing of current metering data and discharge and sediment transport rating data
* entry and editing of geo-hydrological profiles and filter data
* entry and editing of data files for particular analysis.
* 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.

B. Reporting and retrieval

The reporting and retrieval module comprises:

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

2.2 Data Processing System

The data processing system of HYMOS includes following modules:

A. data validation
B. data completion and regression
C. flow measurement
D. data compilation
E. statistical analysis
F. time series analysis.

Details on the contents of the modules are given below.

A. Validation

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

- data screening by flagging, printing and tabular comparison of time series and computation of basic statistics
- 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 balances

- 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.
- investigation of series homogeneity by means of double mass analysis, presented in graphs and tables.
- statistical tests on data homogeneity and randomness, (median run, turning point, difference sign, Spearman rank correlation, Wilcoxon-Mann-Whitney-U, Student-t, Wilcoxon-W, linear trend and range tests)
- spatial homogeneity tests (near neighbour-technique), where data at a base station are compared with weighted averages of neighbouring stations, selected on distance and orientation.

B. Series completion and regression

The completion and regression module comprises:

- interpolation techniques for filling in missing data based on time and space interpolations
- regression models to establish relationships, or to fill in missing data; the models can be of the following types:
  - polynomial equation
  - power equation
  - logarithmic equation
  - hyperbolic equation
  - exponential equation
  - simple and multiple linear regression equations
  - stepwise regression
- Physically based lumped parameter rainfall-runoff model SAMO (Sacromanto Model), for filling in missing runoff data.

C. Flow measurements

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

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

- processing of flow velocity measurements by profile and moving boat methods, allowing:
  - various methods for measurements 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 database
- computation of stage-discharge relations 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
- validation of rating equations for different periods and new measurements
- 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
- stage-discharge transformation, using:
  - rating curves stored in the data base
  - rating equations of measuring or control structures for critical and sub-critical flow conditions and variable sill level.

D. Data compilation

The data compilation module comprises:

* aggregation and dis-aggregation of time series, where accumulative and instantaneous data are treated differently
* series transformation with various arithmetic transformation options
* minimum, mean and maximum series computation for selected time periods and transfer to the database
* computation of areal rainfall by:
  - (weighted) average of point rainfall data
  - Thiessen method
  - Kriging method
* interpolation and computation of best linear estimates of and uncertainties in areal quantities by point and block kriging method
* computation of potential evapotranspiration, using:
  - Penman method
  - Pan-evaporation method
  - Christiansen method
  - Radiation method
  - Makkink method

- Jensen-Haise method
- Blaney-Criddle method
- Mass Transfer method.

E. Statistical analysis

The module for statistical analysis includes:

* computation of basic statistics and histograms
* 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. Some distributions allow left and/or right censoring of data. 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

* statistical tables, i.e. computation of probabilities and variates for the various probability distributions
* generation of normal and gamma distributed random numbers
* computation of IDF-curves (Intensity-Duration-Frequency curves) for monthly and annual maximum precipitation.
    - computation and plotting of:
    - frequency curves
    - duration curves
    - average duration curves.

F. Time series analysis

The time series analysis module comprises:

* auto- and cross-correlation function computation
* spectral analysis
* run analysis: computation of up- and downcrossings, run lengths and runsum
* range analysis: computation of range of cumulative departures from the mean
* analysis of storage requirements by the sequent peak algorithm.

3.0 TIME SERIES, TIME INTERVALS AND TIME LABLES

For time series a distinction is made between:

- ❶ equidistant time series: the time interval between successive observations is constant
- ❷ non-equidistant time series: the time interval between successive observations is variable
- ❸ special series: series of monthly and annual rainfall extremes for specified durations

The major bulk of the data will usually be equidistant time series. Even if the time interval between successive observations is not constant, the series may sometimes be stored and treated as equidistant time series provided that the exact timing of the observation is not of importance.

3.1 Coding

The coding of series is based on the principle that a monitoring network consists of number of stations. In each station various parameters may be observed at different time intervals. This leads to the following coding of series:

1. for equidistant time series	:	station code + data type + time interval
2. for non-equidistant time series	:	station code + data type
3. for special series	:	station code + fixed data types

Station code

The station code is at a maximum 10 character identification of the station.

Data type

The data type is a unique two character specification for the type of series/parameter. For the equidistant and non-equidistant series one may not use the same data types. The data types AE and AM are reserved for the special series.

Each data type has one unit.

The data type may refer to:

1. accumulative observations
2. instantaneous and average observations
3. constant values during a time interval

Equidistant time series may be of the types 1 & 2, whereas non-equidistant time series either of the types 2 or 3.

Time interval

The time interval consists of two parts:

- time interval unit (1-4) ◇ 1 = year
  2 = month
  3 = day
  4 = hour
- divider (1-99)

The divider specifies in how many parts the time interval unit is split up. In case the divider does not split up the time interval unit into equal parts, then one has to define in addition to the time interval a basic time interval to define the division of parts in time unit.

Basic time interval

The basic time interval consists of two parts:

- basic time interval unit (2-5) 2 = month
  3 = day
  4 = hour
  5 = minute
- replicator (1-99)

The replicator is applied to the basic time interval unit to specify the size of the time interval; e.g. decades in a month comprise 10 days, then the basic time interval unit is day (= 3) and the replicator is 10.

3.2 Time Label

The time of observations is expressed by time labels. The time label definition for the elements of equidistant and non-equidistant time series is different.

The time label of an equidistant time series element reads:

- year, month, day, hour, s(ubdivision) i(ndex)

with:

year	: $\geq$ 1850
month	: 1 - 12
day	: 1 - 31
hour	: 0 - 23
si	: 1 - 99

If $\Delta t(1,j)$ then t = year,0,0,0,k
If $\Delta t(2,j)$ then t = year,month,0,0,k
If $\Delta t(3,j)$ then t = year,month,day,0,k
If $\Delta t(4,j)$ then t = year,month,day,hour,k
and $1 \leq k \leq j$

The subdivision index is 1 if the divider is 1. If the divider is > 1 then the subdivision index specifies the timing within the time or calendar unit; e.g. if the time interval is $\Delta t(2,3)$ then the time label of the second decade in a month is: t = year, month, 0, 0, 2.

Examples of time intervals and time labels:

Time intervals	unit, divider
- Seasonal data (4 season) - Monthly data - Decade data (10days) - Weekly data - 15-Minute data - 10-Minute data - Half-daily data - Pentad data (5 days) - 8-Hourly data - 3-Hourly data	1,4 2,1 or 1,12 2,3 *) 1,52 *) 4,4 4,6 3,2 2,6 *) 3,3 3,8

The examples marked with *) need also a basic time interval:

Data	Time Interval	Basic Time Interval
Decade	2, 3	3, 10
Week	1, 52	3, 7
Pentad	2, 6	3, 5

In case of a decade the basic time interval (3,10) means that the first two intervals in a month will have 10 days, whereas the last one the rest of the days in a particular month.

Time of Observation	Δt	Basic Δt	Time Label
- 4 Seasons,2nd Season of 1991	1,4	0,0	1991,0,0,0,2
- June 1895	2,1	0,0	1895,6,0,0,1
- 3rd Decade August 1956	2,3	3,10	1956,8,0,0,3
- 6-10 May 1972	2,6	3,5	1972,5,0,0,2
- 21 April 1947	3,1	0,0	1947,4,21,0,1
- 1600-2000 hrs. 31 Jan. 1985	3,6	0,0	1985,1,31,0,5
- 1830-1900 hrs. 5 March 1967	4,2	0,0	1967,3,5,18,2
- 0905-0906 hrs. 22 April 1991	4,60	0,0	1991,4,22,9,6
- 14th Week 1980	1,52	3,7	1980,0,0,0,14
- 2400 h 28/2 - 0015 h 29/2 1960	4,4	0,0	1960,2,29,0,1

3.4 Missing Value

For each series a value has to be given to identify missing data. This should be a value less than the physically possible minimum value. Default is -99.

3.5 Data Limits

Likely minimum and maximum values and maximum rate of rise and of fall (variation per time interval) may be given for data screening purposes. HYMOS will only flag data exceeding these limits during screening to be investigated further; it will not eliminate those data!

4.0 MENUS AND FUNCTION KEYS

4.1 Main and Sub-menus

The layout of the menus depend on the accessibility level of the user. The following rules apply:

- Options a user has no access to are not shown
- The system manager has level 0
- Users are ranked 1 to 3
- A user of level 1 has access to all data entry, editing, analysis and retrieval options
- A user with level 3 can only retrieve data from database and is allowed to make use of some of the analysis options
- For user with level 2 a middle position is reserved.

The menus are sequenced conformable to the required succession of steps in processing of hydrometeorological data. If so required, with little effort the sequence of steps can be altered by the SYSTEM MANAGER if dBase IIIplus or higher is available.

The default setting of the functions under the main menu options with their user accessibility is shown in Table 1.

4.2 Function Keys

The function keys on the keyboard (<F1> to <F10>) offer a number of actions that can be performed while working with HYMOS. A brief description of the function keys is given under.

<F1> Help
Help texts for support while working with HYMOS.

<F2> Spooler
The spooler provides a temporary storage of output files generated by the HYMOS.

<F3> Editor
While working with HYMOS user files can be created or adjusted with the aid of an editor. This editor can be evoked by this key.

<F4> Change/select current database
To switch to the database of interest this key is pressed.

<F5> Selection of time series
Series codes of available series and data availability will be displayed, page-wise, with 10 series per page. By pressing <Enter> any series may be chosen.

<F6> Logbook
The logbook contains information about the status of processing of series and about the required action. By pressing this key the available logbook are displayed from which any one can be chosen for display.

<F7> Reorganisation of regular time series data file(s) in the database.

<F8> Set configuration
This option allows one to set the hardware configuration according to ones need.

<F9> Temporarily to DOS
To execute a DOS action while running HYMOS this can be used. Strike the space-bar to return to HYMOS.

<F10> Exit
This key is pressed to quit from HYMOS.

Table 1 : Main-menu and Sub-menu with user accessibility.

Main menu	Sub-menu	User level
Entry & editing	Catchment data	1
	Station/series defination	1
	Equidistant time series	1
	Non-equidistant Time series	1
	Combined data/parameters	1
	Geo-hydrological profiles	1
	Dedicated files	1
	Series availability	3
Validation	Screening	3
	Time series graphs	3
	Relation curves	2
	Double mass analysis	3
	Series homogeneity tests	3
	Spatial homogeneity tests	3
Completion & regression	Interpolation	1
	Regression methods	2
	Simulation	1
Flow measurements	Entry & editing data/parameters	1
	Processing of flow measurements	1
	Fitting of rating curve	2
	Shift adjustment	1
	Validation of rating curve	3
	Extrapolation of rating curve	1
	Stage-discharge transformation	1
Data compilation	(Dis-) Aggregation of series	1
	Series transformation	1
	Min-Mean-Max series	1
	Areal rainfall	1
	Kriging	1
	Evaporation	1
Statistical analysis	Basic statistics	3
	Fitting distribution	3
	Statistical tables	3
	Random number generator	2
	IDF-curves	1
	Frequency and duration curves	3
Time series analysis	Correlogram	3
	Spectrum	3
	Range	3
	Runs	3
	Storage	3
Reporting & retrieval	Reports	3
	Mixed tables	3
	Graphics	3
	Retrieval	3
	Transfer	1
System configuration	User profile	0
	Databases	0
	Data types	0
	Log-files	0

5.0 DATABASES

5.1 General

Following considerations should be kept in mind to create databases:

1. Communication between databases directly is only possible via series transfer.
2. All data required for data validation should be located in the same database; hence data outside the catchment being worked upon, but which are useful for the validation of the data in that specific catchment, should also (temporarily) be loaded in that database.
3. Large databases reduce the operational speed and require large disks.

It may be worthwhile to distinguish between:
- workbases, and
- final databases.

In a workbase temporarily the field data of the current hydrological year are stored to be validated and filled in, which are, after being processed, transferred to the final database, to be available for retrieval by the users.

5.2 Database Creation

To store the data of a catchment a database is created as follows:

1. Select <System configuration> from the main HYMOS options.
2. Select <Databases> from the <System configuration> options.
3. Select <Add>, to add a database.
4. Fill in the following data:
   - <Database drive>:
   - <Database code>:
   - <Database name>:
   - <Group name>:
   - <Number of data files>:

5. Press <PgDn> to confirm the entries.

6.0 CATCHMENT DATA

6.1 General

Before entering station and series data it is recommended to enter catchment and subcatchment data and boundaries for reference and later use in the computational procedures, in particular for estimation of catchment rainfall and to visualise the spatial distribution of parameters.

A full description of the entry of (sub-)catchment characteristics and boundaries is presented in the HYMOS manual. Note that when entering this HYMOS option FOR THE FIRST TIME first the latitude and longitude of the local origin (lower left corner) of the entire catchment stored in the database has to be entered. This local origin applies to all catchments or subcatchments in the database!!

The boundaries of the (sub-)catchment consists of an assembly of line segments. Note that for each sub-catchment the total assembly should make closed contour (i.e. the co-ordinates of the first point in the first line segment should be the same as the co-ordinates of the last point in the last line segment).

6.2 Entry of Catchment Data

As outlined in Section 6.1 the entry of the catchment boundaries goes in three steps:

1. Entry of latitude and longitude of the local origin,
2. Entry of catchment characteristics, and
3. Entry of catchment boundaries.

Entry of latitude and longitude of the local origin

The coordinates are entered as follows:

1. Select <Entry & editing> from the HYMOS main menu.
2. Select <Catchment data> from the <Entry & editing> options.
3. Select <Subcatchment Characteristics> from the <Catchmentdata> options
4. Enter degrees, minutes and seconds latitude North or South, Enter degrees, minutes and seconds longitude East or West.
5. Confirm the entries by <PgDn> .

Entry of catchment characteristics

For entry select:

1. <Subcatchment Characteristics> from the <Catchment data> options
2. <Add subcatchments> and fill in the following table:

. Sub-catchment name:	unique name (max. 10 characters)
. River name:	river draining the sub-catchment (max.25 characters)
. Tributary of:	name of the river (max.25 characters) of which the sub-catchment river is a tributary.
. At km:	chainage along main river at which the tributary enters
. Left or Right bank:	side of river at which the tributary enters (L to R)

. Area (km²):	sub-catchment area (is computed from the boundary)
. Stream order:	stream order (1-9) of sub-catchment river, see also note below
. Length (km):	length of river
. Slope (-):	average slope of the river ($*10^{-3}$)
. Latitude:	}geographical coordinates of the
. Longitude:	}local origin of the sub-catchment.

3. Press <PgDn> to confirm the entries.

Entry of catchment boundaries

The catchment boundaries are a set of x,y coordinates (in km), to describe the boundary of the sub-catchment, given in one or more data blocks. Data on coordinates, etc., are stored in data blocks. Each block gets a unique number: BL01, BL02, BL03, etc. construct a boundary a number of blocks may be assembled, where each block represents a line segment. The assembly of line segmentsfor a boundary should make a closed contour.

The boundaries are used for:

- boundaries for computation of Thiessen-weights,
- boundaries for point-kriging.

To add catchment boundaries following steps are required:

1. Select <Subcatchment data> from the <Catchment data> options.
2. Select <Add/Edit block> .
3. Select <Boundary data>
4. Enter data blocks in one of the following ways:
   - when use is made of existing data blocks in the database, then press key <F5> and select a block name from the displayed data blocks. Press <PgDn> to conform the block sequence.
   - to enter a new data block enter first a block name and press <PgDn> . Then select <File> or <Screen>.
     - <File>: data are to be read from file; enter full path and file name of block data file; the layout of the file is explained under "Data file" below
     - <Screen>: to enter the data via the screen and press <PgDn> to confirm sequence.
- Press <Esc> or <PgUp> to switch to other options.

Data file

Generally the layout of the data file for a particular block is as follows:

First line : Block name, i.e. BLXX
Second line: number of rows, number of columns
Next lines : data matrix (x,y) or (x,y,z) n rows*n columns

Data blocks of boundary data:

BL01 block name
5,2 rows and columns
.5,.6 coordinates in km
.6,2 - do -
2,4 - do -
3,5 - do -
5,5 - do -

BL02
6,2
5,5
5,3
4,2
2,.6
.6,.6
.5,.6

Note that the coordinates of the first point of the first data block is equal to the coordinates of the last point of the second data block. This is necessary to compute the area enclosed by the segments correctly.

The boundaries may be shown in a plot on screen by selection of <Plot> from the <Catchment data> options.

7.0 PROCESSING OF RAINFALL DATA

7.1 Summary of Steps

The following sequence of steps for validation ond completion of daily and monthly rainfall data is required:

1. Entry of series codes
2. Creation of input data files
3. Loading of database

- daily data

4. Listing of availability of time series

5. First screening of data
6. Correction/editing of data
7. Computation of correlation distance
8. Validation by series tabulation
9. Plotting of rainfall series
10. Spatial homogeneity test
11. Filling-in of missing data

- monthly data

12. Aggregation of series
13. Screening of monthly series
14. Double mass analysis

- daily and monthly data

15. Final editing and filling-in of missing data
16. Reporting

- data analysis

17. Computation of catchment rainfall
18. Computation of isohyets
19. Statistical analysis
20. Time series analysis

These actions are described in the following sections.

7.2 Entry of Series Codes

When data of a particular series are transferred to the data-base for the first time, the series have to be initialized by specifying the series code through:

- station code
- data type
- time interval

In addition to the series code following information will be of value (although not strictly necessary):

- station characteristics
- station history
- value for missing data
- physical range of the data

Station code

Use one national standard for station coding.

Data type

For observed rainfall the letter combination PH is appropriate. Any other 2 character combination (except AE & AM and those used for non-equidistant time series) may be uses as well; note however that in some applications grouping of series is required (e.g. area homogeneity check, catchment rainfall computation) therefore it is advised to use for; specific parameter one data type code.

Time interval

Time intervals of one day are coded by: 3,1
Time intervals of one month are coded by: 2,1

Station characteristics

The station characteristics include:

- Station code : a unique set of 1 to 10 characters to identify the station This code forms also the first part of the time series code
- Station name : a free name with no operational consequences.
- River : name of the river the station is located on.
- District : district of station location.
- Country : country of station location.
- Latitude : latitude in degrees, minutes and seconds North or South (N/S). Beside in geographical coordinates data may also be entered in UTM see note below
- Longitude : longitude in degrees, minutes and seconds East or West (E/W). Options for latitude also apply here.
- Altitude : altitude in metres relative to reference level ($-10 \leq$ altitude ≤ 8000).
- Catchment area : area upstream of the station (km^2) ($0 \leq$ area $< 10^8$).
- Agency : name of the agency who operates the station.

NOTES

① In the above list of station data is the entry of the <Station code> obligatory; the entry of the rest is optional, but it may be selection-items for <Reporting>

② The latitude and longitude of a station may be given in one of the following ways:
- geographical coordinates, via the station characteristics menu
- UTM (Universal Transformal Mercator) coordinates, press <F2>

For rainfall stations it is particularly of importance to enter the station latitude, longitude (in view of station search for areal homogeneity check, catchment rainfall computation and plotting of isohyets by kriging) and altitude to adjust for orographical effects if apparent.

Station history

The station history file comprises information on station particulars. A fixed layout for all rainfall stations is recommended. The following type of information is advised:

- year of establishment
- description of location and accessibility
- parameters observed
- type of equipment
- type /name of monitoring network
- details of station inspections

Missing data

A value of -1. is appropriate for missing rainfall data

Data limits

The likely minimum rainfall value is 0.; the likely maximum value as well as the rate of rise/fall depend on the time interval and on the location. Consult yearbooks or climatological summaries for this.

The entry of station and series codes is executed by selecting/entering in sequence:

1. <Entry & editing> from the main menu
2. <Station/series definition> from the <Entry & editing> options
3. Select <Add> and enter the station code
4. Press <Enter> to move to the station characteristics and press <PgDn> after entry of the characteristics
5. Select <Series characteristics>
6. Select <Equidistant> from the displayed options
7. Fill in:
   - Data file number to store the series (only required if more than one data file was specified when creating the catchment)
   - Data type (PH for observed rainfall)
   - Time interval ((3,1) for daily rainfall)
   - Basic time interval (not required for $\Delta t(3,1)$)
   - Time shift (enter the hour the rainfall is daily being read)
   - Missing value (-99 or -1)
   - Minimum and maximum rate of rise and rate of fall (see above under "Data limits")

7.3 Preparation of Data Files

In this section the preparation of the data files is elaborated. Rainfall data are generally available as equidistant time series, e.g. hourly, daily, decade, monthly or annual totals.

Definition

A data block is a set of data of a series of a continuous time period. A data file may consist of one or a number of data blocks.

Headers

For data entry data files of equidistant time series may be supplied:

- without file/block header, or
- with file/block header.

For reasons of file documentation and safety in data transfer the second option is to be preferred.

a. Data files without file/block header

Only if the content and the period of record of the data in the file can clearly be expressed in the file name and the layout fulfills the requirements specified below one should apply this option.

The requirements are:

- the file data should be in fixed format,
- the data have to be time sequential row-wise
- the file contains only one data block

When file and block headers are not supplied, HYMOS assumes that the file fulfills the above requirements. Then the series code, start date, number of rows and columns, data conversion factor and reference level, start position of first data field and field length are entered via the screen when loading the database.

b. Data files with file/block header

A detailed outline of the requirements for data files with file/block headers is presented in the HYMOS manual. A summary is given below.

Number of data blocks

If a file header and block header(s) are used then in one file one or more data blocks may be stored.

File/block headers

Each file starts with a file header and each data block starts with a block header. It implies that there is only one file header and as many block headers as there are data blocks.

Data file layout

Based on the previous statements a data file should have the following layout:

- File header
- Block header data block 1

- Data block 1
- Block header data block 2
- Data block 2

...

...

...

- Block header data block n
- Data block n
- etc.

File header

In the file header specify:

- whether the data are in FIXEd or in FREE format, and
- whether the data are:
. time sequential Row-wise and one series per data block
. time sequential COLumn-wise and one series per data block
. time sequential column-wise with in one data block a number of series PARallel.

Data format

In a file the data may be presented in free or in fixed format:

- free format means the field length is variable and a blank or a comma is used to separate the data. If the data are in free format all data in a block will be read; there is no possibility to skip a part of a data block.

- fixed format means a specified fixed field length for each value in a data block. The formats may differ from one block to another. In this case not necessarily all data will be transferred; A PART OF THE DATA BLOCK MAY BE SKIPPED by specifying per data block the column number of the first data field, the field length and the number in columns.

Block header

If the data block contains one series it is to be specified in the block header:

- the seriescode of the data in the block.
- the start date,
- the number of rows and columns,
- the conversion factor and reference level to be applied to the data in the transfer
- in case of fixed format data: the start position of the first data field and the field length
- in case the data are column-wise time sequential additionally the actual number of rows per column have to be specified.

Basically the same type of information is to be supplied in the headers of data blocks with more than one series (parallel series option).

Data blocks

The layout of the data blocks is specified in the block headers. Note that when the data are provided format free then:

- in case of row-wise time sequential data an incomplete last row has to be ended with /
- in case of column-wise time sequential data all columns have to be filled up to the maximum column length with some value. This value will not be transferred but it is only there to put the other values in the right position.

7.4 Loading of Database

To transfer the data from data files to the database, select subsequently:

1. <Entry & editing> from the main menu
2. <Equidistant time series> from the <Entry & editing> options
3. <Read data from file> from the <Equidistant time series> options

Messages will appear if data file or file or block header does not comply with the requirements. If the message reads "series code not found" then check whether the series has been created (check through <Review> under <Station/series characteristics>) as well as check the spelling of the series code in the data file. Trouble may also be caused by a comment/copyright line under the last data block when an ASCI file is created via a spreadsheet program. This line is not seen if e.g. the Norton Editor is used for file editing. Hence delete the last blank line(s).

Important

When the data of a file have been transferred to the database, then before continuing with the next one, a look at the series through option <Edit time series> from the same <Equidistant time series> menu is worthwhile. Once the data are shown on the screen, press <F3> to display the series graphically. (Note that in one selection not more than 400 data will be displayed, so repeat the action if necessary).

7.5 Listing of Availability of Time Series

Prior to validation make a barchart of the availability of data in the time series. This will help to select stations for mutual comparison only for time periods where the stations have data and to reduce the amount of output.

The barchart is made by selecting/entering in sequence:

1. <Entry and editing>, from the main menu
2. <Series availability>

3. Time interval: <Day> and divider = 1
4. Select:
   - <Series sequential>, or
   - <Series random>

   If at this stage more than one data type has been created then choose the <Series random> option for selection of series by scrolling through the displayed series codes and select the series by pressing <Enter> when the cursor bar is over the required series code. The availability of up to 20 series can be shown on a graph.
5. Start and end date: when the time series is selected, the date of the first and the last value loaded will be displayed as well as the number of months or years to be plotted in one graph; the latter is default set equal to the number of months/years between the start and end date. In order to improve the visibility of the availability select only a limited number of months or years in one plot. The rest follows automatically in the next plots.

7.6 First Screening of Data

The first screening comprises of:
1. Listing of the data, and 2. Flagging of outliers.

To obtain a listing of the data in a convenient form is of most importance for reference and to administer further screening and correction. Rainfall amounts outside the expected physical range should be flagged for further investigation

The screening options are dealt with in the HYMOS manual. The listing of daily rainfall data and flagging of outliers is obtained by selecting/entering in sequence:

1. <Validation> from the main menu
2. <Screening> from the <Validation> options
3. <Screening> from the <Screening> options
4. Output device: <HYMOS>, <Printer> or <File>:
   - if one wants to have first a look at the data before the tables are sent to the printer, select <HYMOS>, and the output is accessible via <F2> from the main screen
   - if one wants to print the data later or from a different computer and printer then select <File>; this option is of particular interest as it does not hamper the progress of data processing at the main computer where HYMOS has been installed
5. Select the time interval: select <Day> and divider= 1
6. Display the available series via <F5> and select by pressing <Enter> when the cursor bar is over the required series;
   - note that a number of series can be selected in one run, but to avoid many blank pages make sure that the selected series all have data for the time period to be entered next (for this inspect the barchart of the availability of data)

- note also that for all selected series the same flagging criterion will be used, see also under 10.

7. Enter the start date and end date of the period for which the listing should be made
8. Select from the layout of tables: <dedicated>
9. Enter 1 for the number of decimals in the output as the accuracy of the rainfall measurement is generally .1 mm
10. Select <absolute> boundaries for flagging of outliers; default the data limits entered at the time the series codes were created will be displayed; at this stage these limits may be changed for the screening if so required.

In the output one will find all the details of the screening. Inspect the results carefully on unrealistic extremes.

HYMOS can also provide a list of all the outliers only. For this select in sequence:

① <Validation> from the main menu
② <Screening> from the <Validation> options
③ <Less/greater than> from the <Screening> options

and enter a minimum and maximum value for the selected series. All values outside this range will be printed.

7.7 Correction/Editing of Data

Corrections to data can be of the following types:

a. Editing some single series data
b. Elimination of a sequence of erroneous data
c. Adjustment by mathematical operations

Editing some single series data is done by selecting/entering in sequence:

1. <Entry & editing> from the main menu
2. <Equidistant time series> from the <Entry & editing> options
3. <Edit time series> from the <Equidistant time series> menu
4. Press <F5> to select the code of the required series
5. Enter the start date series part to be edited; from this date onward 400 data will available for editing
6. Enter the number of decimals in the displays
7. Correct the required data
8. Press <Esc> to quit and answer "Y" to <Update database>
9. Answer "Y" to <Print updated data>; the output is stored in the Spooler.

Elimination of a sequence of erroneous data is done by selecting/entering in sequence:

1. <Entry & editing> from the main menu
2. <Equidistant time series> from the <Entry & editing> options
3. <Entry of constant> from the <Equidistant time series> menu

4. Press <F5> to select the code of the required series
5. Enter the start and the end date of the sequence to be eliminated
6. Select the missing value load to be loaded.

Editing option c. includes <Series transformation>, which is available under <Data compilation>. in particular attention is to be given to the linear equation option of the <Series transformation> options. As an example the adjustment of a part of a series by some multiplication factor is given below. The following selections and entries are required:

1. <Data compilation> from the main menu
2. <Series transformation> from the <Data compilation> options
3. <Transform series> from the <Series transformation> menu
4. Select the time interval of the series
5. <Linear equation> from the series transformation options
6. Select option: <Free>, i.e. a free set of coefficients can be entered
7. Press <F5> to display the available series and select with <Enter> the series to be adjusted
8. Enter the required value for coefficient <c1>, answer X_j ?(Y/N) with "N" and accept the default for <c6>
9. Enter the start and end date of the part of the series to be adjusted

7.8 Computation of Correlation Distance

To determine which stations can be grouped to be considered for intercomparison the correlation between daily rainfall observed at the stations as a function of the distance between the stations should be computed, see e.g. Figure 7.1.

The distance between the stations is obtained from the map showing the station locations.

The correlation between the daily rainfall is computed by the regression option correlations are obtained by selecting/entering in sequence:

1. <Completion & regression> from the main menu
2. <Regression methods> from the <Completion & regression> options
3. <Input>
4. Data from: <HYMOS>
5. Time interval: 3,1
6. Data time sequential Y/N?: enter "Y" if all data from start to end have to be considered; however, if only a period in the year is to be considered the; "N" and after the start date the number of data per year and the number or have to be specified
7. Press <F5> to select the first series and enter the start date and end date or number of data per year and the number of years, depending on the choise under 6; note that no more than 1000 data per series can be selected

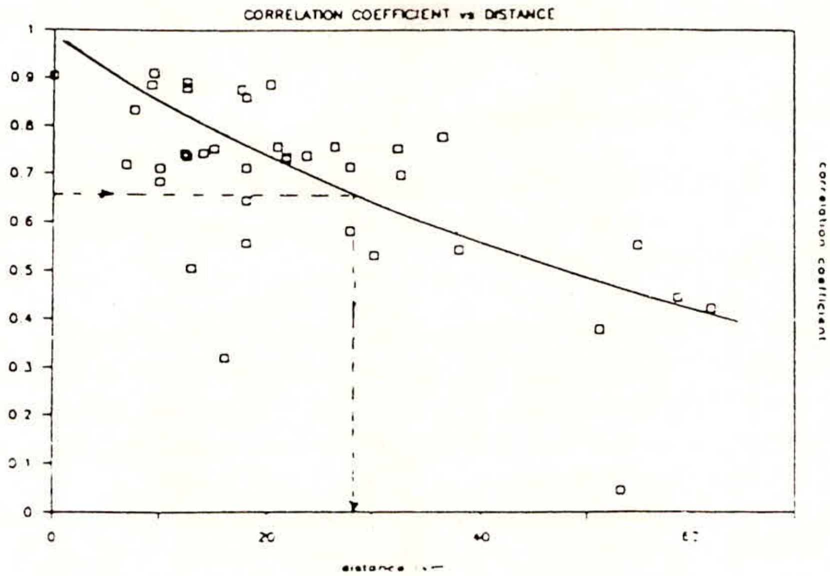

Fig. 7.1 : Correlation-Distance Graph

8. Press <F5> to select other series (max. 10)
9. Press <Esc> to stop the series selection
10. <Function>
11. <Correlation matrix> from the <Function> options
12. <Select>
13. Select the series pair-wise by entering codes 3 (for dependent variable) and 1 (forced variable) for the first two and press <PgDn> to enter the <Lower>and <Upper> limits for data in the correlation analysis; use 1 mm as lower limit to eliminate all "0" from the computation
14. Repeat steps 12 and 13 for:
    - the 1^{st} and the 2^{nd} series
    - the 1^{st} and the 3^{rd} series

      .

      .
    - the 1^{rst} and the 10^{th} series
    - the 2^{nd} and the 3^{rd} series, etc.

Then enter the <Spooler> via <F2> from the main menu to see the results.

Note that the correlation computation has deliberately being executed pair-wise to get as much as possible data in the computation. It is possible to enter all stations at one go by coding the first series with 3 and all others with 1, but then all days will be eliminated where one or more stations have zero rainfall (because the <Lower> limit is set to or have no value at all.

Next plot the result of correlation coefficient versus distance and draw a line of best fit through the data. Mark the distance where the correlation falls below approximately 0.7

or 0.6. This distance is called the correlation distance. It gives an indication of what distance between stations is acceptable to call it neighbour station.

7.9 Validation by Series Tabulation

By series tabulation up to 6 series will be displayed in columns parallelly for date-wise comparison. With this option two types of errors may be detected, provided that neighbour stations are being selected:

1. day-shift in data entry; gauges read in the morning of day n should be entered at day n-1. In particular at stations with non-professional staff one may see the entry being set on day n. By inspection of sequences of shifts from non-rainy to rainy days this type of error is easily traced with the help of a ruler to ease the day-wise comparison
2. infrequent data entries; another common error is that the gauges are not always read every day and the entered totals refer to a couple of days. This policy may not affect decade or monthly totals to great extent but certainly will affect the daily rain fall statistics (extremes, number of rainy days) and also rainfall-runoff modelling (rainfall intensity relative to infiltration capacity determines what part runs off via the surface or percolates to the groundwater). From a group-wise intercomparison, such an error is easily seen.

The day-shift error and the way to detect it is shown in Figure 7.2.

The table is obtained by selecting/entering in sequence:

1. <Validation> from the main menu
2. <Screening> from the <Validation> options
3. <Tabulation> from the <Screening> options

Next after specification of the time interval up to at maximum six stations can be selected.

Once again select only groups of neighbour stations, otherwise this type of analysis is not meaningful and leads to erroneous conclusions.

If a series appears to be shifted one day eliminate the data first by overwriting it with missing values and re-enter the series data to the database by adjustment of the start day with one day.

Date (1979)	Measured Rainfall (mm)		
	Heath No.	G'Fields	Ch. Ercall
1/1			
1/2	19.5		
1/3	2.1	1.9	
1/4	0.4	0.8	
1/5	15.4	14.4	0.8
1/6	0.0	0.0	17.2
1/7	0.0	0.0	0.0
1/8	10.3	10.5	0.0
1/9	10.1	9.5	14.3
1/10	9.2	10.2	8.3
1/11	4.1	5.1	12.3
1/12	20.6	21.1	1.8
1/13	4.8	4.5	18.8
1/14	9.3	9.2	8.0
1/15	0.3	0.4	10.7
1/16	12.7	7.9	0.9
1/17	0.2	0.0	1.5
1/18	27.4	19.1	9.3
1/19	0.0	0.0	17.8
1/20	1.7	2.4	0.0
1/21	46.2	30.9	1.8
1/22	0.0	0.0	12.8
1/23	0.4	0.5	0.0

Shows obvious sequence error

Error

Fig. 7.2 : Detection of Day-Shift Errors

7.10 Plotting of Rainfall Series

The intercomparison of daily rainfall data is further supported by a graphical display of the time series year by year or even for smaller time periods. To improve the discrimination between the series further, use different origins by applying the shift option. At maximum 5 series can be displayed in one plot. As before this plotting is only meaningful if neighbour stations are shown in the same graph. An example is shown in Figure 7.3.

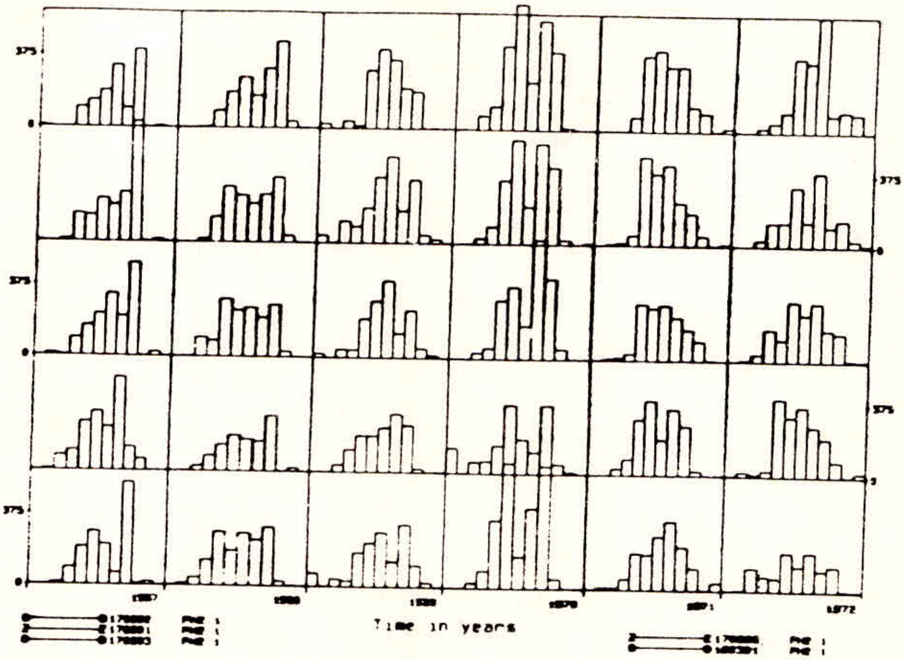

Fig. 7.3 : Plot of Rainfall Data

Graphs are obtained by selecting in sequence:

1. <Validation> from the main menu
2. <Time series graphs> from the <Validation> options
3. <Time series> from the <Time series graphs> options
4. <Equidistant> series
5. Time interval: <Day> and divider = 1
6. Press <F5> to display the available series and select by pressing <Enter> when the cursor bar is over the required series
7. Enter the required time period and specify the number of months in one plot
8. Then the plot options are displayed and select:
   - <Scale>
   - <Shift>
   - <Graph type>

   and move the cursor bar to <Draw>
9. Specify the scale of the vertical axis: for daily rainfall round the maximum value to e.g. the nearest 50 or 100 mm
10. For the shift apply if the scale maximum was x_m:
    - first series : 0
    - second series: x_m
    - third series : $2.x_m$
    - fourth series: $3.x_m$, etc.
11. Select for rainfall series as graph type <bar>
12. Select the device: <Screen>, <Plotter> or <File>; if one writes the plot

to file, then, to have it plotted later on, one has to copy the plot file to the plotter port: if this is e.g. com2 then type: copy <plotfile name> com2

If the rainfall stations are close to each other, the series should show more or less the same pattern; only in periods of typical local convective showers the daily patterns may show less resemblance.

7.11 Spatial Homogeneity Test

Spatial homogeneity of rainfall data is tested by comparing the observations at a station (= test station) with estimates for that station computed from observations at surrounding stations. In the estimates, the weights of the surrounding stations are inversely proportional to some power of the distance to the test station. Surrounding stations are selected based on distance and orientation relative to the test station. If the absolute difference between the observation and the estimate exceeds absolute and/or relative limits then the observation will be flagged (not eliminated!) to stress the need for further investigation.

The test is executed by selecting/entering in sequence:

1. <Validation> from the main menu
2. <Spatial homogeneity test> from the <Validation> options
3. Enter the data type; note that only series of this type will be used in the test
4. Select the time interval of the series
5. <No normalization> under the normalization options unless orographical effects are expected
6. Press <F5> to display the available series and <Enter> to select the required test station (= basic station)
7. Enter the start and end date of the part of the series to be investigated and select the output device
8. Enter the radius of area of influence in km; this is the maximum distance between the test station and its neighbours for being available for selection
9. Enter the power to which the distance is raised in the weight of a particular neighbour station in the computation of the estimate
10. Enter the absolute admissible deviation between observation and estimate and the multiplier to the standard deviation of the observations at neighbour stations
11. Select the series per quadrant around the test station. The selected ones are marked with <*> based on distance; this may be changed by pressing <Enter> when the cursor bar is over a series

In the results only the observations are shown which did not fulfill the criteria. If there is clear evidence that a value is wrong, eliminate or correct the value with the editor.

7.12 Filling-in of Missing Data

Generally, filling-in of scattered missing data is advised to improve the value of the data set. It is doubtful whether filling-in of large missing parts is be executed when preparing for publication of data in yearbooks. However, for analysis purposes also the filling-in of large missing parts may be considered.

Filling-in missing or eliminated daily rainfall totals is best achieved by spatial interpolation. This is exactly the same technique as used to flag data by the spatial homogeneity test discussed above. Note that with this procedure the extremes as well as the number rainy days will be affected by the (weighted) averaging; with respect to the latter if at o of the selected neighbour stations rainfall is observed also the series being filled-in will get a non-zero value!! Note also that normalization has to be applied if orographical effects or some other type of physically based spatial inhomogeneities do occur.

The procedure is executed by selecting/entering in sequence:

1. <Completion & regression> from the main menu
2. <Interpolation> from the <Completion & regression> options
3. Select the data type of the series being filled-in and of the neighbour stations
4. <Spatial interpolation> from the <Interpolation> options

The remainder is the same as for the spatial interpolation test except that now no error criteria have to be defined.

7.13 Aggregation of Series

After the validation, editing and completion of the daily data a control validation takes place for monthly data. The daily data are therefore aggregated to monthly totals.

To aggregate data the following steps are required:

a. Create the series codes of the monthly series, $\Delta t(2,1)$, similar to the creation daily series codes; select/enter:

b. Aggregate the daily values to monthly values; select/enter:

1. <Data compilation> from the main menu
2. <(Dis-)Aggregation of series> from the <Data compilation> options
3. Select the time interval of the basic series: <Day>, divider = 1 Select the time interval of the aggregate (monthly series): <Month>, divider = 1
4. Select the series to be aggregated (= basic series, i.e. the daily rainfall series); note that a number of basic series can be aggregated in one run
5. Next the basic series are displayed and combine them one by one with the displayed series codes of the aggregates: the choice is displayed in the right hand window; press <Esc> to terminate the selection to continue
6. Enter the start and end date of the period for which the daily data have to be aggregated.

Note that if one or more data in a month are missing, no monthly total will be computed and the month will be stored as missing.

7.14 Screening of Monthly Series

Repeat the validation procedure as explained in 7.5 to 7.11 for the monthly data in full, (mind the changes in the time internal definition!!). As monthly data will show a higher correlation the intercomparison will be easier and more discriminative.

7.15 Double Mass Analysis

Instead of looking at individual observations as in the spatial homogeneity test, in the double mass analysis the existance of long term inhomogeneities are investigated. The double mass curve shows the ratio between accumulated totals at the test station and one or a group of maximum 5 neighbour stations (= base stations). If this curve deviates from a straight line, then a change has occurred at the test station provided that the base stations are correct, hence it is important to select reliable stations as base stations; if the group of neighbours is choosen large enough then errors in one are hardly felt in the total sum.

The double mass analysis is used in two steps:

1. first at annual (or a part thereof) basis; hence daily observations are accumulated to annual (or seasonal) values
2. if the first curve shows a distinct break then the analysis is continued at monthly basis for the years around the break to determine in detail when the inhomogeneity started

The double mass analysis for monthly rainfall data is performed by selecting/entering in sequence:

1. <Validation> from the main menu
2. <Double mass analysis> from the <Validation> options
3. Select <Year> as interval of analysis
4. Select the time interval of base and test station: <Month> and divider =
5. Press <F5> to display and <Enter> to select the base station series and enter multiplier to each series to compute the total (default the average of the series be taken)
6. Press <F5> to display and <Enter> to select the test station series
7. Enter for the base stations (X) and the test station (Y):
   - conversion factor,
   - lower and upper limit if censoring of data is required,
   - unit of data in the output
   - name of the series in the output
8. Enter:
   - the start date,
   - the number of elements per year (as <Year> was selected under 3)

or part of the year; note if one would enter here e.g. 10, then from each year 10 month will be included in the "annual" total
the minimum number of elements in an "annual" total; if at the test station or at base stations some days would be missing, the same months will be omitted in totals for all stations, but if the minimum number is not reached that year will be considered as being missing

9. Select the output device and plot-text can be entered if required

In case of a distinct break, the test series data are adjusted by multiplying the inhomogeneous part with the ratio of the slope of the curves before and after the break vise versa. With the third correction option explained in Section 7.7 this adjustment can be executed.

7.16 Final Editing and Filling-in of Missing Data

Based on the validation of monthly data a second editing of daily data may be required. Adjust/fill-in the data with the above explained procedures. Then repeat the aggregation get a consistent set of daily and monthly data.

In case there is only interest in final storage of monthly data, then instead of completing series at daily level the filling-in may be done at monthly level. This option is to be considered when a large number of daily data in a month is missing; in case only a 1 daily data is missing then the spatial interpolation at daily level is to be preferred. ' filling-in of missing data at monthly level may be done by:

1. spatial interpolation, see Section 7.12
2. multiple regression.

The multiple regression option goes in two steps:

1. creation of the multiple regression equation
2. filling-in the missing data using the regression equation

To create the multiple regression equation following steps are to be taken in sequence:

1. <Completion & regression> from the main menu
2. <Regression methods> from the <Completion & regression> options
3. <Input>
4. Data from: <HYMOS>
5. Time interval: 2,1
6. Data time sequential Y/N?: enter "Y" if all data from start to end have to be considered; however, if only a period in the year is to be considered the answer "N" and after the start date the number of data per year and the number of years have to be specified
7. Press <F5> to select the series and enter the start date and end date or number of data per year and the number of years, depending on the choise

under 6;

8. <Function>
9. Select <Multiple> or <Stepwise> from the <Function> options; if the <Stepwise> option is used enter also:
   - maximum number of steps (do not use more than 4)
   - output every time step (Y/N):
   - F-values to enter and to delete variables from the regression equation
10. <Select>
11. Select the dependent variable by entering code 3 and the independent variables (max = 4)
    - in case of multiple regression by code 1 (forced variables)
    - in case of stepwise regression by code O (variables free for selection) and press <PgDn> toenterthe <Lower> and <Upper> limits for data in the correlation analysis
12. Store the created regression equation

The filling-in of missing data goes along the following lines; select/enter in sequence:

1. <Completion & regression> from the main menu
2. <Interpolation> from the <Completion & regression> options
3. Select the data type of the series being filled-in and of the neighbour stations
4. <Series relation> from the <Interpolation> options
5. Select from the displayed relations the required one by pressing <Enter> when the cursor bar is over the relation
6. Enter the period to be filled-in using the selected regression equation; note that the filling-in may be executed for complete years or parts of successive years (to be specified by the <Period in year>

Filling-in by regression may sometimes lead to negative values. This is physically impossible for monthly rainfall. Hence after the filling-in has been completed, check on the existance of negative monthly totals and replace those negative values by zero's. For this use can be made of the <Replace> option of <Equidistant time series> under <Entry and editing> .

7.17 Reporting

After the daily and monthly rainfall data have been validated, corrected and missing values have been filled-in, the series are ready for reporting. This can for hourly, daily, decade and monthly data conveniently be done in the following way; select/enter in sequence:

1. <Reporting & retrieval> of the main menu
2. <Reports> from the <Reporting & retrieval> options
3. <Report> from the <Reports> options
4. Select a station from the displayed available stations
5. Select the time interval <Day> or <Month>
6. Select a series code from the displayed available series

7. Specify the contents of the report:
   - Station characteristics (Y/N):
   - Station history (Y/N):
   - Start date:
   - Number of years:
   - Title: name of series in report
   - Number of decimals in report; if -1 is entered the number of decimals will be variable (data up to 3 digits accurate)

7.18 Computation of Catchment Rainfall

Catchment or areal rainfall can be computed from point rainfall by the following procedures:

1. equal station weights
2. user provided station weights
3. weights according to Thiessen method.and
4. weights according to kriging method.

As an example the steps necessary to execute the computation of areal rainfall using the Thiessen method will be described.

Make sure that:

- the series code for areal rainfall has been created !!
- the (sub) catchment boundary has been entered !!

The following selections/entries are required:

1. <Data compilation> from the main menu
2. <Areal rainfall> from the <Data compilation> options
3. <Thiessen> from the <Areal rainfall> options
4. Select an output device
5. Select the time interval of the series to be considered
6. Enter the data type of the point rainfall series
7. Select the series code of areal rainfall
8. Enter start and end date for areal rainfall computation
9. <Plot of Thiessen polygon (Y/N)>:
10. <Percentage of enlargement for station search>: to allow stations outside the catchment to participate
11. Enter point rainfall series code either from file (if the weights were already computed before) or from the screen
12. Select the (sub-)catchment for which the computation is to be made (boundary data should have been entered)
13. Next the catchment with bi-sectors is shown provided the stations. were entered from the screen (otherwise no plot will be made)
14. Store the station weights for later use, if required; to get consistent areal

rainfall data through the years it is important that as much as possible the same stations are included.

7.19 Computation and Plotting of Isohyets

The plotting of isohyets goes in two steps:

1. Computation of rainfall values on a dense grid over the catchment using the kriging technique; this results in a datamatrix, and
2. Drawing of isohyets by the <Graphics> option under <Reporting and retrieval>, making use of the data matrix developed in step 1.

Kriging

For computation of catchment rainfall and plotting of isohyets use is made of the kriging technique. For this, a rectangular grid is placed over the catchment (= area of interest). Estimates of the rainfall at the grid points are obtained as a weighted average of the rainfall at surrounding stations. The station weights are determined by kriging and depend on the distances between the grid points and the rainfall stations on the one hand and the adopted covariance model, representing the spatial correlation structure, on the other. The estimates for the grid points are expressed in a statistical way: the kriged value, which is the best linear estimate (unbiased and minimum variance), and the uncertainty in this value indicated by the standard deviation of the estimate. The technique can also be applied to other quantities with a spatial character. Well known is the application to derive groundwater level contours.

Covariance models

The models for the generalized covariance are of order k=O and have following forms:

1. $K(h) = C.\delta + A_1 . \mid h \mid$ (7.1)
2. $K(h) = C.\delta$ (7.2)
3. $K(h) = A_1 . \mid h \mid$ (7.3)

Where:

$K(h)$ = generalized covariance of order k=O
h = distance
δ = Dirac function
C, A_1 = parameters, with $C \geq 0$ and $A_1 \leq O$

The parameter values C, A, are either determined by HYMOS or provided by the user. The parameters may by determined for the actual values or for their natural logarithms.

Parameter estimates

For the estimation of the model parameters a system of linear equations, called the kriging system, is solved. At least 5 rainfall stations should be available. The rainfall stations in the analyses are selected based on their location and on the series type.

Orographical effects

The model assumes spatial homogeneity. In order to incorporate orographical effects,kriging technique may be applied to normalized values, i.e. normalized e.g. by long term averages. In that case for all rainfall stations as well as for the grid points long term average values have to be provided. The values may be based on e.g. topographical information, prevaling directions of inflow of moist air.

Catchment rainfall

The average rainfall in the catchment is determined as the average of the kriged rain values at the grid points. The boundary of the catchment is slected from a scrolling menu. Only the grid points with the catchment boundaries are considered. The number of rainfall stations to be included for computing the rainfall estimate for a grid point can be limited to the nearest stations specifying the number of data points included in kriging (> 5).

Isohyets

Based on the rainfall values at the gridpoints a linear interpolation will be carried out to arrive at the isohyets.

Network optimization

Beside the isohyets also isolines of the standard deviation of the rainfall estimates can obtained to investigate the acceptability of the rainfall network density. By adding or deleting stations its effect on the size of the standard deviation can be evaluated, until fulfills the requirements.

Station weights

As an optional output the weight of each station in the areal rainfall computation can obtained and stored in the data base.

Application

To compute rainfall values on a dense grid over the catchment following steps are required :

1. Select <Data compilation> from the main HYMOS options
2. Select <Kriging> from the <Data compilation> options
3. Select <Point-kriging>
4. Select <Isolines>
5. Choose output device: <HYMOS>, <Printer> or <File>
6. Enter the time interval unit and divider and the data type of the series considered in the computation, e.g. <Time interval unit>: Month and <divider> 1 monthly data and for rainfall data <Data type> = PH
7. Enter the date (year, month, day) for which the computation should be made
8. Select a sub-catchment from the displayed list, necessary for the selection of station
9. Answer the next questions <Orographical adjustment required?> and <Take n log. of the obs. values?> with N
10. Select the model $K(h) = a_1 * |h|$ It from the available generalized

covariance functions

11. Enter the grid size in x and y-direction in km; enter 10
12. Enter the number of data points included in kriging; the maximum available da points are default
13. Answer Y to <Select an area of interest>, i.e. the catchment boundary selecte at step 8 will be used to bound the computation of rainfall values for the grid points.
14. To store the gridpoint data with the rainfall estimates and their standard deviation enter a file name; the extension of this file name is MAT. Subsequently, if at 13 the answer was Y. the catchment boundary is stored under the same file but with extension INT. So if the file name is e.g. TEST, at this step two data matrices have been created: TEST.INT: which stores the catchment boundary TEST.MAT: storing the grid co-ordinates, rainfall estimates and standard deviation
15. After leaving the option select <F2> to inspect the output file (if <HYMOS> was selected at step 5).

To draw the isohyets following steps are required:

1. Select <Reporting & retrieval> from the main HYMOS optios
2. Select <Graphics> option under <Reporting and retrieval>
3. Select <Create> from the <Graphics> options
4. Use an old file and select MEKPREC.PCT from the available list. MEKPREC.PCT is an earlier made picture file for drawing of isohyets. The contents of the file will now be adjusted to the requirements.
5. Select EDIT from the available options.
6. Adjust the MEKPREC; consult Annex 8 for the details. Note that the matrix files TEST.INT and TEST.MAT of step 14 above have to be included.
7. After finishing the EDITing of all tasks press <Esc> .
8. Select SAVE and enter a new picture file name, e.g. TEST.PCT, to avoid overwriting the MEKPREC.PCT file.
9. Select DRAW and select <Screen>, <Plotter> or <File> for plotting/storage of the .PCT file

7.20 Statistical Analysis

As an example of statistical analysis the fitting of theoretical distribution functions to annual extreme monthly rainfall data will be executed.

To fit distribution functions following steps have to be taken:

1. Select <Statistical analysis> from the main HYMOS options.
2. Select <Fitting distributions> .
3. Select an output device: <HYMOS>, <Printer> or <File> .
4. Read the data from <Data base> .
5. Select the required series by pressing <F5> .
6. Select type: <Annual maximum> .

7. Select period: <Full years>.
8. Enter start and end date.
9. Next one may edit the data or introduce thresholds on the series.
10. Select the required distribution function and when applicable select the number of parameters and the parameter estimation procedure.
11. Next the series statistics with the model parameters will be shown.
12. Now select <Extremes for distinct return periods> .
13. Select subsequently <Plot of distribution function>
14. Select <Scale> and <Conf.lim.> and then <Draw>

7.21 Time Series Analysis

As examples of time series analysis the autocorrelogram and spectral density function will be computed and plotted.

Autocorrelogram

The autocorrelogram of monthly rainfall series is obtained as follows.

1. Select <Time series analysis> from the main HYMOS options.
2. Select <Correlogram>.
3. Select <Autocorrelation analysis> .
4. Select an output device: <HYMOS>, <Printer> or <File> .
5. Select <Month> and <1> as time interval.
6. Press <F5> and select the required series from the displayed table.
7. Enter the start and end date of the part of the series to be analysed.
8. Enter the maximum number of lags in the correlogram, e.g. 36.
9. Select the required drawing options and plot device.

Spectral density function

The spectrum of monthly rainfall series is obtained as follows.

1. Select <Time series analysis> from the main HYMOS options
2. Select <Spectrum> .
3. Select an output device: <HYMOS>, <Printer> or <File> .
4. Select <Month> and <1> as time interval.
5. Press <F5> and select the required series from the displayed table.
6. Enter the start and end date of the part of the series to be analysed.
7. Enter the truncation lag of the covariance function, e.g. 36.
8. Enter the number of frequency points: e.g. 72
9. Select the required drawing options and plot device.

8.0 CLIMATOLOGICAL DATA

8.1 Types of Data

Beside rainfall the following types of climatological data are likely to be processed:

1. Pan evaporation
2. Radiation
3. Sunshine duration
4. Minimum, mean and maximum temperature
5. Wind speed
6. Relative humidity
7. Vapour pressure

8.2 Processing of Climatological Data

The processing of climatological data generally goes along the same lines as the first steps for processing of rainfall. A problem usually is that the density of the climatological network is much less than of rainfall. However, on the other hand the spatial variability of the most of the climatological parameters is also much less, so station intercomparison can be considered. Furthermore, use can be made of the physical relationship of the parameters: evaporation increases with windspeed and vapour pressure deficit (low humidity). The set of parameters 2 to 7 allows computation of potential evaporation, which can be compared with the pan evaporation measurements.

Generally, following steps are relevant:

1. Entry of series codes; it is assumed that at the climatological stations also rainfall is one of the parameters being measured, so the station characteristics would have already been entered
2. Creation of input data files
3. Transfer of data to the database
4. Listing of availability of time series
5. Screening of data, by listing the data and flagging of outliers
6. Plotting of time series, with the same parameter at neighbour stations in one graph
7. Computation of evaporation from climatological parameters; consult the methods shown in the HYMOS manual
8. Double mass analysis on:
   - pan evaporation data of neighbour stations
   - pan evaporation and computed evaporation at the same station
9. Correction of data and where possible filling-in of missing data either by time series interpolation or regression
10. Aggregation of daily series to monthly series for pan evaporation and mean daily temperature
11. Create series of daily minimum and maximum temperatures per month, using

the <Min-Mean-Max series> option under <Data compilation> .

12. Reporting

9.0 WATER LEVEL DATA

9.1 Required Basic Station Information

The information required to execute the processing of stage data includes:

1. The station location and date of establishment
2. Site description
3. Full details of the gauge setting, gauge reach and gauge zero
4. Hydraulic conditions upstream and downstream of the station: existance of reservoirs, intakes, outlets, possible backwater effects
5. Benchmark particulars
6. Type of equipment installed
7. Whether or not discharge is being measured
8. Site accessibility
9. History of malfunctioning, break downs, repairs, maintenance and relevelling

The frequency with which water levels are being read depends on the variability of the water level and on the objective of the measurements. Basically, the readings taken at the smallest interval are the data to be entered for further processing. Where computerized data processing has started, it appeared to have been common practice, that only daily averages are being entered. This is, however, incorrect if water level data are used to compute discharges; due to the non-linear relationship between stage and discharge, averaging of water levels before computation of the discharge will lead to underestimation of the actual discharge.

9.2 Types of Errors in Water Level Records

To design effective procedures for validation of stage records one need to know what kind of errors can be expected. Typical errors related to manual or automatic recording of stages include:

1. Entry or transcription errors; easily detected as unrealistic sharp peaks in time series plots.
2. Overtopping of gauges; result shown as a flat top of the hydrograph.
3. Shifts in the setting of one or more of the staff gauges; this causes shifts in the stage-relation with nearby stations.
4. Readings are not taken at the required intervals; a stepped hydrograph results.
5. "Upside down readings", i.e. the above water metre mark is put in front of the correct centimetres reading; easily detectible by comparison with hydrographs of nearby stations: reverse behaviour of the water table at the two sites and at the metre change there is a distinct break.

6. Blockage of intake pipe of stilling well, or a staff gauge disconnected with the river; the recorded levels will show a constant value.
7. Recorder pen is malfunctioning; the hydrograph will show an unrealistic behaviour: compare the data with the staff gauge readings.
8. Clock of the recorder is malfunctioning; resettings of the pen apparent from interruptions in the chart record; compare with staff gauge readings.

From the above it follows that stages plotted versus time and versus stages of nearby stations will be most effective to detect erroneous data.

9.3 Steps in Data Processing

The following steps are required in processing water level data:

1. Entry of series codes. The data type "HH" may be used.
2. Creation of input data files.
3. Loading of data.
4. Listing of the availability of water levels.
5. First screening of data. The data limits are the minimum and maximum level of the gauge, whereas flashy floods may give realistic entries for the maximum rate of rise and of fall. Mark the data flagged and correct where necessary.
6. Plotting of water level time series. Combine the hydrographs of stations along the same river in one plot and do not use the shift option here; due to that do not shc more than three hydrographs in one plot. Inspect the hydrographs carefully 4 smooth behaviour and on resemblance with neighbour stations. Correct where necessary.
7. Fitting of stage-relation curves; this activity is explained in more detail later.
8. Adjustment of data.
9. Filling-in missing data:
   - for minor gaps use linear interpolation,
   - for larger gaps use stage-relation equations.
10. Aggregate water levels to daily averages.
11. Reporting of daily average water levels.

9.4 Fitting of Stage-Relation Curves

In natural streams between two sites along the same river a one to one relationship between stages may be established provided that the inflow between the two sites is comparative small and that the proper time shift correction is applied to adjust for time the flow requires to travel from a to b. If the time shift correction is not properly applied the stages will show a looped relation.

Following selections/entries are required:

1. <Validation> from the main menu

2. <Relation curves> from the <Validation> options
3. <Compute relation> from the <Relation curves> options
4. Specify the time interval of the series to be investigated
5. Select series code for the dependent variable Y
6. Select series code for the independent variable X; then the type of relation is shown in the right window of the screen. If this is the relation of choice then continue with select <Continue>.
7. Enter the start and end date of series Y.; note that at maximum 500 data can be considered in one run.
8. Enter the start date of series X; usually, on physical grounds a time shift can be specified (accounting for travel time of the flood wave: approx. $c = 1.5 (B_{river} / B_{storage})*u$)
9. <Optimize time shift (Y/N)>, answer with "N" if for the shift has already been accounted for. If the answer is "Y", then the number of intervals for time shift adjustment have to be entered as well as the lower and upper boundaries for each interval; HYMOS computes per interval the "optimum" time shift from the maximum of the cross-correlation function of Y and X which may be overruled by the entry on the required time shift.
10. Next a number of options are displayed:
    - <Display X-Y data>: this will produce a graph of Y versus X and next to it plot options are displayed; special attention is asked for the option <Reduce> which allows to reduce the number of X-Y values to be plotted in order to save plotting paper and plotter pen
    - <Compute relation curves>: this option leads to fitting the relation by polynomials. Polynomials will be computed for at maximum three intervals of X. If more than one interval is used, apply an overlap in the definition of the boundaries to force an intersection between the polynomials of the intervals.
    - <Compare with curve>: this option allows to make a comparison between the presently selected X-Y data and previously made relation curves. When selecting this option the available relation curves will be displayed, available for selection <Change>: to change period, time shift or series
11. After relation curve computation the following options are displayed:
    - <Graphics>: to plot the data and the curve.
    - <Error analysis>: to give details on the goodness of fit of the curve.
    - <Store results>: to store the parameters of the relation curve in the database for purpose of comparison with other periods or to fill in missing data.
    - <Change>: to change period, time shift or series.

10.0 STAGE-DISCHARGE DATA

10.1 Required Steps

Assuming that combined stage-discharge measurements are available, the following steps have to be taken to compute a stage-discharge relation:

1. Collection of station particulars and cross-sectional data
2. Preparation of data files
3. Transfer of data to the database
4. Fitting or validation of rating equation
5. Extrapolation of rating equations

10.2 Station Particulars

Details of the flow measurement equipment, instrument calibrations, measurement procedures and measuring teams have to be collected.

Of high importance is the collection of

- a layout of the river stretch and existance of structures in the neighbourhood
- a detailed layout of the cross-section from year to year for breaking up the rating equations in intervals, and
- the type of roughness in the cross-section for extrapolation purposes.
- information on possible backwater or unsteady flow adjustment

Backwater

When stations are located upstream but in the reach of influence of structures like dams or near confluences than there may not be a single relation between stage and discharge. This is usually clearly observed from a plot of the discharge measurements: for the same water level there is a range of discharges. To get a first order estimate of the possible extent of backwater the following equation can be used for a rectangular cross-section:

$$\Delta h_x = \Delta h_0 \exp(-3\ S\ L_x\ /\ (h_n(1 - F_r^2))) \qquad (10.1)$$

with ;

Δh_X = backwater effect at $x = L_x$
Δh_0 = initial set up of water level at $x = 0$
S = river bottom slope
L_x = distance
h_n = normal or equilibrium flow depth, $h_n = (q/(K_M.S^{1/2})^{3/5}$
q = discharge per unit width
K_M = K-Manning = (1/n)-Manning
F_r^2 = square of Froude number = $(h_c/h_n)^3$
h_c = critical flow depth, $h_c = (q^2/g)^{1/3}$
g = acceleration of gravity

For low Froude number from eq.(10.1) it follows that the length for a reduction of the initial set up by 90 per cent is:

$$L_{90\%} = 0.77\ h_n/S \qquad (10.2)$$

Equation (10.2) shows the extent of the backwater to be proportional to normal flow depth and inversely proportional to the river slope.

Unsteady flow

To investigate the existance of possible unsteady flow effects, shown as a looping in rating equation, the Jones equation can be investigated:

$$Q_u = Q_s\ (1 + (c\ S)^{-1}\ dh/dt\ )^{1/2} \qquad (10.3)$$

Q. = Q. . (l + (c . s)-l . dh/dt) hi (9

where:

Q_u = discharge under unsteady flow conditions
Q_S = uniform discharge
c = celerity of flood wave, c = 1.5 $u^*(B_{river}/B_{storage})$
dh/dt = rate of change of stage

From eq.(10.3) it observed that flashy floods in flat rivers may show a looped rating curve E.g. for a water level change of 0.20 m in one hour in a river with a slope of 10^{-4} and velocity of 2 m/s the right hand side term in eq.(10.3) is 1.09, i.e. the actual discharge deviates by 9 per cent from the uniform discharge. If the river slope is 10^{-3} under the same conditions, the deviation is less than 1 per cent, hence negligible.

10.3 Preparation of Data Files

Measured stage-discharge data can be entered via screen or read from data files. If in the computation of the rating curve corrections are required for unsteady flow or backwater do not forget to enter the gradient respectively the fall; in the standard case just a dummy at this place will do.

10.4 Transfer of Data to the Database

Two options are available to load stage-discharge data to the database:

1. Transfer of data from file, or
2. Entering data via the screen

The entry via the screen is executed by selecting/entering in sequence: (Note that to transfer from file goes along the same line up to step 5, where one select <File>)

1. <Flow measurements> from the main menu
2. <Entry/Editing data/parameters> from the <flow measurement> options
3. Select <Add> and <Current metering data>

4. Following options are displayed for selection:
   - <Simple rating>: at least stage and discharge data are required.
   - <Unsteady flow correction>: at least stage, discharge and rate of change of the water level (gradient) are required.
   - <Backwater corr. constant fall>: at least stage, discharge and fall are required
   - <Backwater corr. normal fall>: requirements as for constant fall.
5. Select: data from <Screen>
6. Enter date, number of observation, gauge zero, stage, discharge, gradient or fall, and optionally: width, wetted perimeter and cross-sectional area

10.5 Fitting and Validation of Rating Equation

Rating curves are advised to be presented by power type equations, which have the following general form:

$$Q = c\ (h + a)^b \tag{10.4}$$

where:

Q = discharge
h = water level
a,b,c = coefficients

A rating curve may be built up of at maximum three equations, each of them valid for a certain water level range. Corrections can be included for unsteady flow and backwater effects.

The following steps are recommended to establish a rating curve:

1. Before fitting a stage-discharge relation analyse the hydraulic conditions at the station and in its vicinity.
2. Determine which one of the following situations does exist:
   - simple rating equation
   - looped relation due to unsteady flow
   - complex relation due to backwater
3. Thereafter inspect the cross-section and determine the levels where distinct changes in the cross-sectional profile occur.
4. When available, check the flow measurements on the application of the standards of ISO or WMO (points in vertical and number of verticals)
5. Make a listing and plot of the stage-discharge data and compare the maximum water level range with the stage range of the discharge measurements; extrapolation is required if the full water level range is not covered by the stage-discharge measurements.
6. Adjust the data set by flagging unreliable data with "0"; then these measurements will temporarily be eliminated from the data set.
7. Fit the rating equation to the measurements; make sure that the number of

measurements per unit water level does not vary too much, so that in the fit (based on least squares principle) applies for each water level equally well. If more intervals are used apply an overlap in the definition of the intervals to force the equations for each interval to intersect within the overlap! In case of backwater with normal fall the fitting goes in two steps. First the backwater free data have to fitted by a rating curve; for this the backwater free data have to be flagged with "2" in the data <Adjust> option!! Next the backwater correction is computed.

8. Inspect the error analysis and the stage-discharge plot. If more than one interval are used, check whether the intersection of the equations is within the interval overlap; if this is not the case try to improve with an adapted definition of the intervals.
9. Store the rating curve when it matches satisfactory.

The fitting of a simple rating equation is executed by selecting/entering in sequence:

1. <Flow measurements> from the main menu.
2. <Fitting of rating curve> from the <Flow measurement> options.
3. Select station code.
4. Enter start and end date of the data to be analysed and select the output device.
5. Then following options are displayed:
   - <Standard procedure>
   - <Unsteady flow correction>
   - <Backwater with constant fall>
   - <Backwater with normal fall>

   Select <Standard procedure> .
6. Next the following options are shown:
   - <Print>: to list the retrieved stage-discharge measurements.
   - <Plot Q/h data>: to plot the stage-discharge data with linear and double logarithmic scales.
   - <Adjust>: to eliminate data temporarily from the dataset by changing data flag "1" to "0".
   - <Continue>: to proceed with the fitting.

   When the data set is used for the first time make use <Print> and <Plot Q/h data> .
7. Select <Power> type of rating equation and do not fix parameter "a" .
8. Enter the required water level intervals for which an equation has to be established and enter the boundaries for each interval (with overlaps!!). Also keep in mind that the lower boundary of the first interval is less and the upper boundary of the last interval is larger than the range of the water level series, otherwise missing values for the non-covered range will be entered. Press <PgDn> to execute the computation.
9. Next the following options are displayed:
   - <Plot curves>: to obtain a plot of data and rating curve.

- <Error analysis>: to obtain a comparison between observed and computed discharges; the output is sent to the selected output device, and a summary is presented on screen.
- <Store parameters>: to store the parameters of the rating equation; the validity period can be adapted at this stage.
- <Change condition> to:
    - <Adjust data>: to change data flags.
    - <Adjust intervals>: to change the water level interval.
    - <Return to main> to:
        - . select another <Station>
        - . select another <Period>
        - . change the type of <Computation>
        - . <Exit> .

The validation of rating curves goes along the same lines. Note that for each interval in the validation the number of stage-discharge data used in the fitting of the rating curve and the corresponding standard error has to be entered!!

11.0 DISCHARGES

11.1 Required Steps

The following steps are recommended for the processing of discharge data:

1. Create discharge series codes if a series is not yet initialised.
2. Stage-discharge conversion using rating curves; make sure that the conversion takes place using the water level series with the smallest sampling/time interval.
3. List the availability of discharge series.
4. Screen/list discharge time series.
5. Plot discharge time series, with series along the same river in one plot.
6. First adjustment/editing of data.
7. Check the water balance for sequential stations along the same river or for stations around confluences, see <Time series graphs> option: <Balance> and adjust the data if required; lack of balance may occur if intake or release of water takes place in between the stations.
8. Execute double mass analysis of discharge series. A clear break in the curve may be observed if the flow at one of the sites is affected from some period onward due to intake/release of water. This information may be used to check the start of and changes in the water intakes/releases.
9. Aggregate discharge series to monthly and annual values for those sites where natural flows can be expected. If series are affected by intakes/releases then virgin flow series have to be created by correcting the series for the water withdrawal or release; then execute the aggregation.
10. For the natural created virgin flow series, establish a relation between annual or seasonal runoff and catchment rainfall (u/s of the station) with the same

time interval and use this relation to create synthetic discharge series. It is of high importance that the point-rainfall series used to estimate the catchment rainfall have thoroughly been checked on consistancy.

11. Execute double mass analysis on the historical and the synthetic discharges. Assuming that the rainfall series are correct the double mass curve between the historical and synthetic discharges should result in a straight line. Deviation from a straight line may may be caused by the application of erroneous rating curves.
12. Correct erroneous data by re-evaluation of rating curves for the suspected part and execution of stage-discharge transformation.
13. Fill-in missing data by water balance considerations for the smallest interval available or rainfall-runoff simulation; aggregate the data and if required fill-in the remaining part by rainfall-runoff regression as from monthly intervals onward.
14. Report on daily and monthly flows.

11.2 Stage-Discharge Transformation

Water level series may be transformed into discharge series by use of the stage-discharge relation of one of the following form :

1. Simple rating curve
2. Rating curve with unsteady flow correction
3. Rating curve with constant fall backwater correction
4. Rating curve with normal fall backwater correction
5. User structure equation
6. Measuring structure relation.

11.3 Filling-in Missing Data

To fill-in missing discharge data one may use:

1. water balances, if a closed balance exists between two or more stations
2. rainfall-runoff simulation; HYMOS includes an adapted version of the well known Sacramento Streamflow Model
3. rainfall-runoff regression, usually carried out for monthly or seasonal flows.

Water balances

A linear relation, fulfilling the mass conservation condition, is created via <Series transformation> under <Data compilations and applied to a period of missing data. Then this relation is stored and subsequently picked up at <Interpolation> under <Completion & regression> as a <Series relation>.

Rainfall-runoff simulation

The Sacramento model is available under <Simulation> of the <Completion &

regression> options. Its use is beyond the scope of the workshop. Reference is made to the HYMOS manual for details.

Rainfall-runoff regression

With the help of multiple and stepwise regression, monthly or seasonal relations between discharge and rainfall in the same and previous months/seasons can be made:

$$Q_t = F(P_t, P_{t-1}, \ldots) \qquad (11.1)$$

where:

Q_t = discharge in month/season t
P_t = precipitation in month/season t
P_{t-1} = precipitation in month/season t-l, etc.

This type of relations is made by the <Regression methods> under <Completion & regression> as follows:

1. Select output device and data source (if from database select HYMOS)
2. If the data are read from the database: specify the time interval unit and divider (for monthly data 2 1)
3. Next specify whether continuous records with data from start to end are considered or parts of years (e.g. a season); in the former case the data are considered 'time sequential'.
4. Press <F5> to obtain a list of available series and select the series one by one.
5. Enter the start and end date of the first series and the start dates of the next series. The precipitation series refer to the same series but are only shifted in time; this is achieved by entering different starting dates!!
6. Subsequently, the names of the selected series may be adjusted, which is particularly useful in this case; to remember the time shifts include them in the name.
7. Select stepwise regression to investigate the significance of the rainfall at previous months/seasons in the relation.
8. Mark the dependent variable = discharge with code 3 and the independent ones available for selection with code 0.
9. Set limits to the data if so required.
10. Execute the regression and check the result in the Spooler (<F2>).

* * * * * * * *