HYDROLOGY PROJECT

World Bank & Government of The Netherlands funded

Training module # SWDP - 07

How to make data entry for rainfall data

New Delhi, November 1999

CSMRS Building, 4th Floor, Olof Palme Marg, Hauz Khas,
New Delhi – 11 00 16 India
Tel: 68 61 681 / 84 Fax: (+ 91 11) 68 61 685
E-Mail: dhvdelft@del2.vsnl.net.in

DHV Consultants BV & DELFT HYDRAULICS

with
HALCROW, TAHAL, CES, ORG & JPS

Table of contents

1. Module context

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

2. Module profile

Title	:	How to make data entry for rainfall data
Target group	:	Data Processing Centre Assistant, Assistant Hydrologists, Hydrologists, Data Processing Centre Manager
Duration	:	One session of 60 minutes
Objectives	:	After the training the participants will be able to*:* • Enter rainfall data at various intervals • Carry out the prescribed data entry checks
Key concepts	:	• Daily rainfall data • Rainfall data at synoptic hours • Tabulated hourly rainfall data • Digital rainfall data from data loggers
Training methods	:	Lecture, exercises, software
Training tools required	:	OHS, computers
Handouts	:	As provided in this module
Further reading and references	:	

3. Session plan

No	Activities	Time	Tools
1	***General principles*** • Important points	2 min	OHS 1
2	***SWDES and rainfall data entry*** • Rainfall data entry and SWDES	2 min	OHS 2
3	***Manual inspection of field records*** • Manual inspection of field records	2 min	OHS 3
4	***Entry of daily rainfall data*** • SWDES - Features of daily rainfall form, data entry checks, graphs	10 min	OHS 4, 5, 6, 7
5	***Entry of rainfall data at twice daily interval*** • SWDES - Features of twice daily rainfall form, data entry checks, graphs	4 min	OHS 8, 9
6	***Entry of hourly data*** • SWDES - Features of hourly rainfall form, data entry checks, graphs	5 min	OHS 10, 11, 12
7	***Exercise*** • Explore features of daily rainfall form and enter one year rainfall data and perform appropriate data entry checks • Explore features of twice daily and hourly rainfall data entry forms	15 min 20 min	

4. Overhead/flipchart master

5. Handout

Add copy of Main text in chapter 8, for all participants.

6. Additional handout

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

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

7. Main text

Contents

How to make data entry for rainfall data

1. General principles

- **There is a requirement to make all rainfall data available on computer for validation processing and reporting - the first step is therefore data entry.**

- **Data entry will be carried out at Sub-divisional offices as near as possible to the observation station to ensure interaction between data processing and observation personnel.**

- **Primary module of the dedicated hydrological data processing software (SWDES) is available for all types of surface water data entry including rainfall. This front-end module is specifically customised to suit the layout of data entry forms used by the field observers for recording the data and to carry out graphical and tabular data validation.**

- **Initial emphasis is on the entry of current rainfall data, but SWDES also provides a very convenient and efficient means of entering historical data, from original data sheets where available and otherwise from published tabulations.**

- **Prior to entry to computer two manual activities are essential:**

  - ❖ Registration of receipt - on the day of receipt (See Module 5)
  - ❖ Manual inspection of rainfall data sheets and charts

- **On completion of data entry and primary validation in the Primary module, data will be exported (transferred) to the Secondary module of dedicated hydrological data processing software at the Divisional office for further validation and processing.**

2. SWDES and rainfall data entry

SWDES has been developed as specialist data entry software based on Microsoft Access to customise data entry for the individual needs of states and agencies, mimicking the forms and data sheets that are used by agencies for particular variables.

SWDES is primarily designed for the entry of time series data but it also incorporates space-oriented data sufficient to locate and catalogue the stations under the control of a particular state or agency. Stations and series can then be accessed from typical Windows Menus and Toolbars by clicking on appropriate buttons. This feature, of course, is common to all variables.

For all equidistant and cyclic time series data, SWDES provides entry screens automatically with date and time labels against which the variable values are entered. This simplifies data entry and avoids the potential errors of date/time entry.

SWDES provides data entry checking capability, rejecting clearly spurious values and flagging suspect ones for inspection. For example it will reject entry of an alpha character in a numeric field or duplicate decimal point and will highlight for inspection values above a pre-set limit. In addition options for plotting time series graphs, at time of entry are available in most cases.

SWDES provides a suitable format for each of the following rainfall instruments and frequencies:

- Standard non-recording raingauge - read daily (at selected standard time)
- Standard non-recording raingauge - read twice daily (at selected standard times)
- Autographic recording (siphon) raingauge – values tabulated at hourly intervals from chart
- Digital recording (tipping bucket) raingauge - digitally logged values at fixed intervals
- Digital recording (tipping bucket) raingauge - digitally logged values of each tip.

Data from standard non-recording raingauges and autographic gauges will be entered from keyboard; digital data will be by file transfer.

3. Manual inspection of field records

Prior to data entry to computer an initial inspection of field records is required. This is done in conjunction with notes received from the observation station on equipment problems and faults, missing records or exceptional rainfall. **Rainfall sheets and charts will be inspected for the following:**

- Is the station name and code and month and year recorded?
- Do the number of record days correspond with the number of days in the month?
- Are there some missing values or periods for which rainfall has been accumulated during absence of the observer?
- Have monthly totals of rainfall and rain days been entered?
- Have the autographic rainfall hourly totals been extracted?
- Is the record written clearly and with no ambiguity in digits or decimal points?

Any queries arising from such inspection will be communicated to the observer to confirm ambiguous data before data entry. Any unresolved problems will be noted and the information sent forward with the digital data to Divisional office to assist in secondary validation. Any equipment failure or observer problem will be communicated to the supervising field officer for rectification.

4. Entry of daily rainfall data

Using SWDES the station and daily series is selected and the screen for entry (or editing) of daily rainfall is displayed as given in Fig. 4.1. Simultaneously, the station and series codes and corresponding Sub-division and local river/basin are also displayed. A window showing year and month, from which the month of entry may be selected. Also displayed are windows showing the upper warning limit which is used to flag suspect values; these values can be altered, for example depending on season. Negative values are rejected as well as values over a maximum limit

There are four columns for the date, daily data, and cumulative amount within the month and remarks respectively. The dates are filled automatically according to the month and year. The data corresponding to each day is to be entered by the user. The cursor goes in vertically downward direction by default after each data entry. When data are missing, the corresponding cell is left blank (not zero) and a remark entered against that day. Where the observer has missed readings over a period of days and an accumulated total is subsequently measured (e.g. rainfall, pan evaporation or wind run) the cells corresponding to the missed days will be left blank (not zero) and a remark will be inserted against the date of the accumulation to specify the period over which the accumulation has occurred (e.g. Accumulated from 23 to 27 Sept.).

Note: There are occasions when the climate observer is legitimately absent from his station, for example on account of sickness. The observer must be encouraged to leave such spaces “Missing” or “Accumulated” rather than guess the missing values. The completion procedures, based on adjoining information, are better able to estimate such missing values.

As the entries are made the cumulative amount within the month is computed and filled automatically. If any remarks are to be entered the same can be done by going to the specific date by using mouse/tab/cursor. At the bottom of the form, for each month, the number of rainy days, total and maximum rainfall for the month as available in the manuscript has to be entered. A rainy day is defined as that day on which the rainfall is more than 0.0 mm. The computed values of these quantities as per the entries by the user is automatically filled.

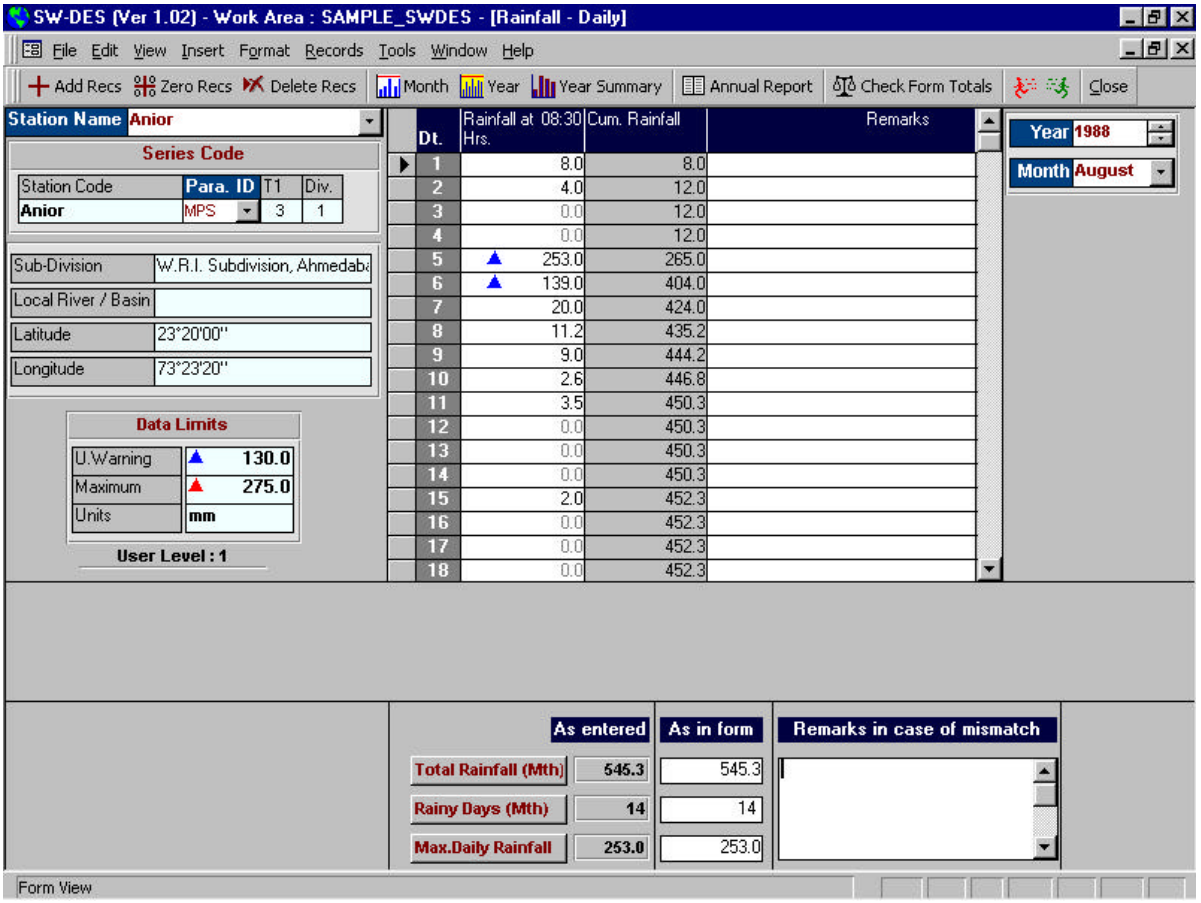

Fig. 4.1: Entry screen format for entering daily rainfall data

During the process of making entries the user can draw the graph for the data being entered. There are three options in which the data being entered could to plotted: (a) to plot the daily data of the month (see Fig. 2) and (b) to plot the daily data of the entire year (see Fig. 3) and (c) to plot the monthly totals (summary, see Fig. 4) for the whole year.

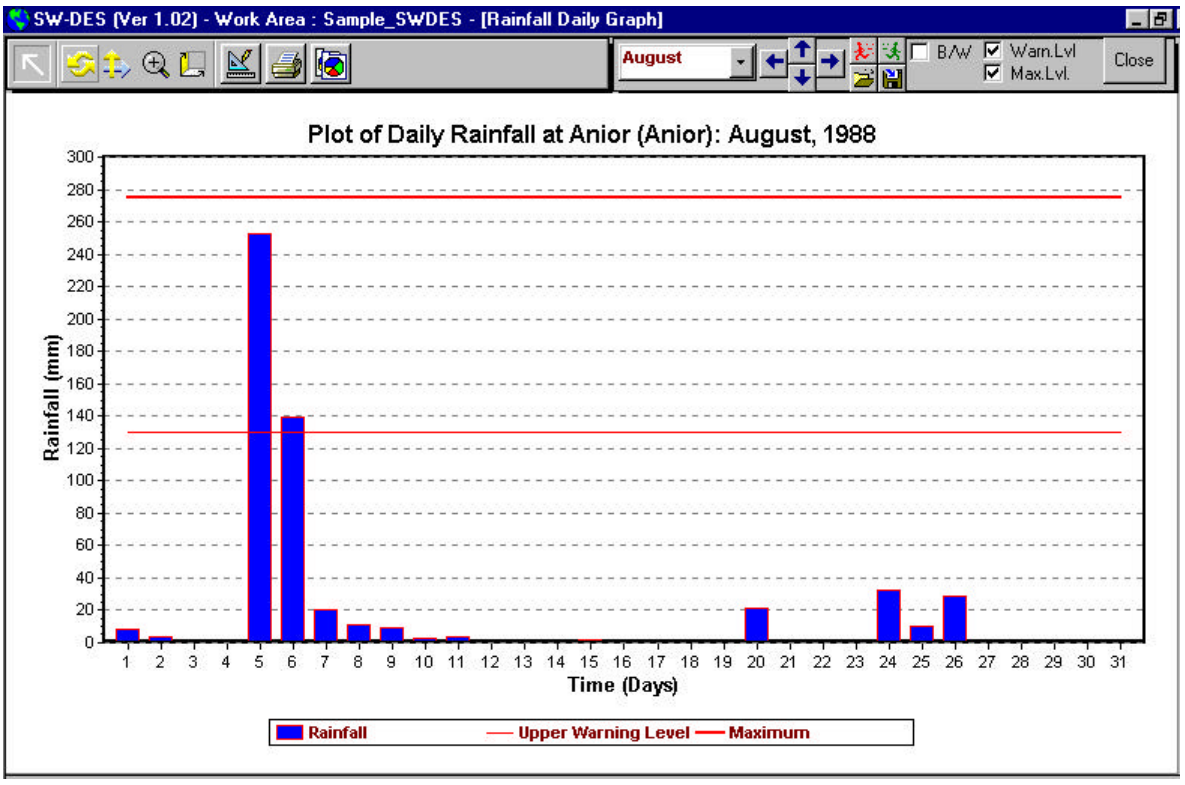

Fig. 4.2: Plot of daily rainfall for a month

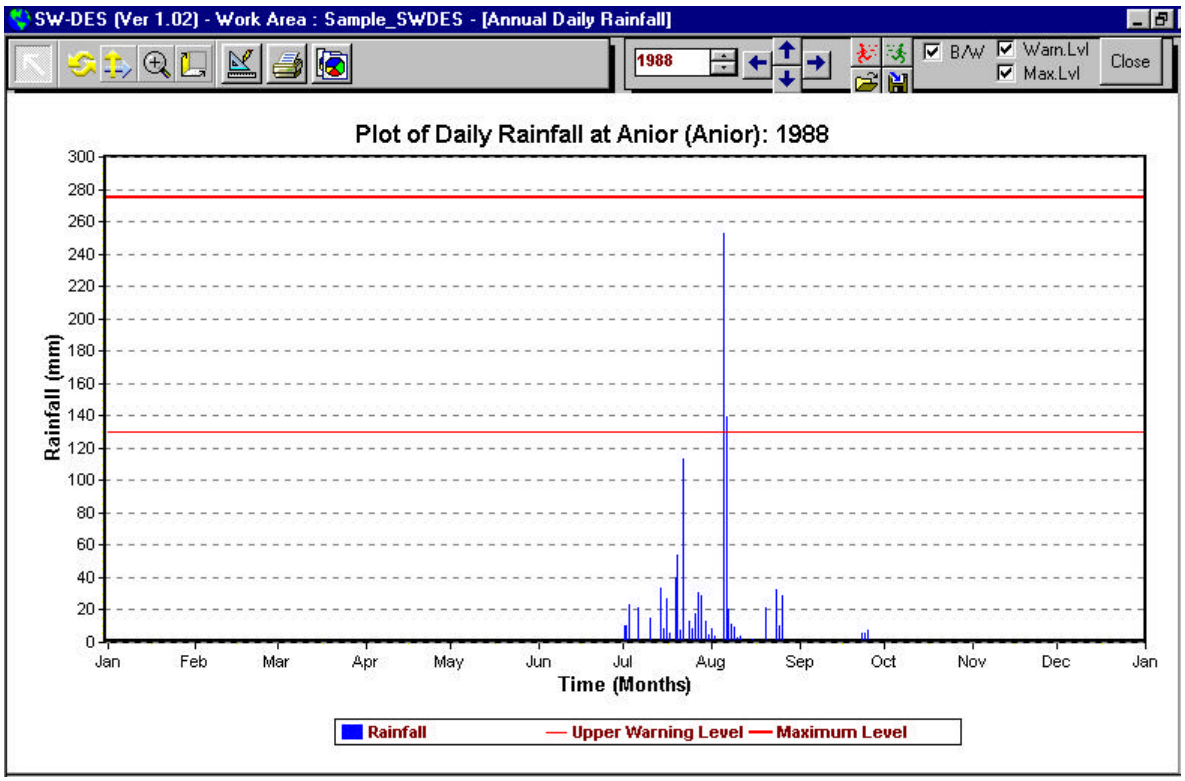

Fig. 4.3: Plot of daily rainfall for a year

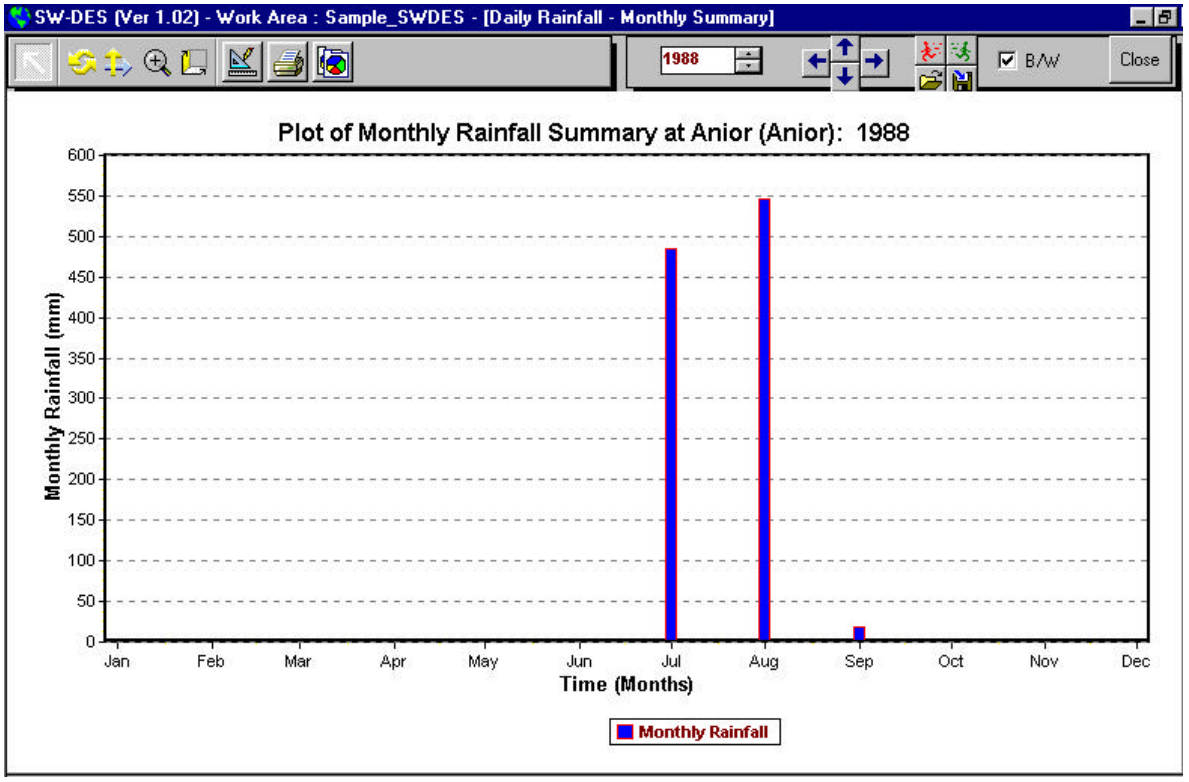

Fig. 4.4: Plot of monthly rainfall summary (monthly total) for a year

Two types of data entry checks are performed for this case of daily rainfall data.

(a) **The entered daily data can be compared against upper warning level and maximum value**. This allows the user to quickly know which data value has violated the prescribed limits. Upon such prompting the user can once again refer back to the manuscript to see if there was some mistake in entering the data. If such values which violated the maximum data limits are found to be actually reported in the manuscript then the user can put suitable remarks to indicate so.

(b) **Checks are carried out to see if there is a proper match between the entered and computed values of number of rainy days in the month, maximum and total rainfall in the month**. In case of any mismatch the user is prompted by colour highlighting of the mismatch, so that he can check back the entries. If cumulated values of daily series are also available in the manuscript then it becomes faster to pin down the mistake.

Scrutinising and checking the daily rainfall data month by month in this manner will leave little scope of any data being wrongly entered.

5. Entry of rainfall data at twice daily interval

The layout of the form for entering twice daily rainfall observed at standard times (at morning and evening synoptic hours) is as given in Fig. 5.1:

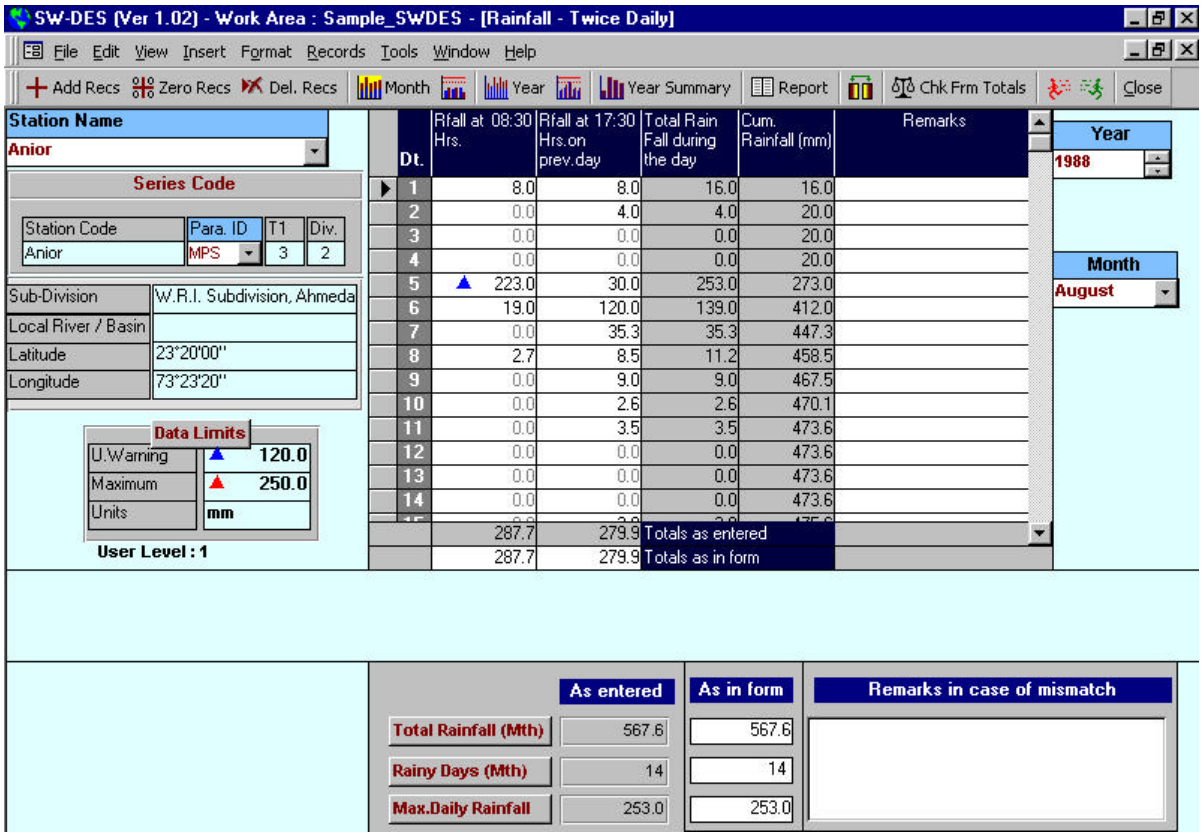

Fig. 5.1: Entry screen format for entering twice daily rainfall values

As for the daily rainfall form this form also displays the information regarding station code, station name, sub-division and local river/basin. Two data limits: (a) Upper warning level and (b) Maximum value are also displayed on the screen. The year and month of the data displayed on the screen at any point of time are also displayed for reference purpose and to set to any desired month and year.

There are six columns in all: one for the date, two for each observation of the day and one each for total daily rainfall, cumulative amount within the month and remarks respectively. The dates are filled automatically according to the month and year. Data corresponding to each observation in the day is to be entered by the user. After entering any data the cursor, by default, goes horizontally to next record of the day and then goes to the first observation of the next day. As the entries are made the daily total and the cumulative amount within the month are computed and filled automatically. Remarks may be entered against any day.

At the bottom of the two data columns the respective column totals are to be entered as available in the manuscript. The corresponding computed values derived from the entries made by the user are computed and filled automatically. Apart from this, for each month, the number of rainy days, total and maximum rainfall for the month as available in the manuscript have to be entered. The computed values of these quantities derived from the entries by the user are automatically filled.

The graphical facility for this case is similar to that provided for the daily rainfall data with rainfall for both the observations in the day shown differently (see Fig. 5.2).

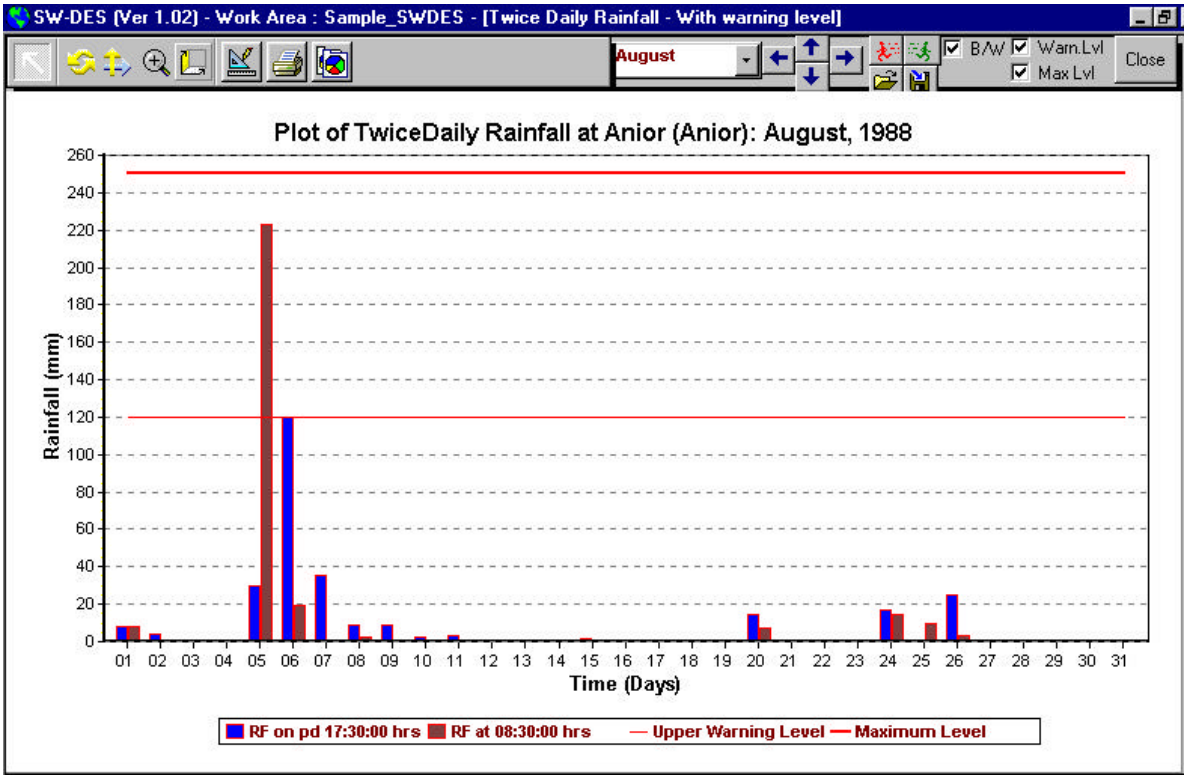

Fig. 5.2: Plot of twice daily rainfall for a month

Two types of data entry checks are performed for this case of twice daily rainfall data.

(a) The entered daily data can be compared against upper warning level and/or maximum value.

(b) The second type of check is carried out to see if there is a proper match between the entered and computed values of totals for each column, number of rainy days in the month, maximum and total rainfall in the month. In case of any mismatch the user is prompted by coloured highlighting of the mismatch so as to check back the entries. If cumulated values of daily series are also available in the manuscript then it becomes faster to pin down the mistake.

Making checks on columnar totals, number of rainy days, monthly maximum and totals, provides adequate data entry checks and enables faster tracking down of errors. Any mismatch remaining after thorough checking with the manuscript must be due to incorrect field computations by the observer and must be communicated to him through the field supervisor.

6. Entry of hourly data

Hourly rainfall data are obtained either from the chart records of the autographic type recording raingauge or by the digital data obtained from a tipping bucket raingauge (TBR). The data originating from data loggers, in digital form, can directly be imported into the database using the import option. A special form for entering hourly rainfall data abstracted from autographic charts is available. However, the option of entering the digital data using this form can also be made use of, if required. Moreover, after digital data have been imported they can be inspected graphically and validated for certain limit checks using this form.

As for the daily rainfall form this form also displays the information regarding station code, station name, sub-division and local river/basin. Two data limits: (a) Upper warning level and (b) Maximum value are also displayed on the screen. The year and month of the data displayed on the screen at any point of time are also displayed for reference purpose.

The layout of the hourly rainfall form is as given in Fig. 6.1:

Hourly rainfall data are entered in the form of a matrix in which the columns are the hourly rainfall values for a day and the rows represents different days of the month. Time-label entries for the dates and hours are filled automatically. The rainfall value is entered to the time following the hour in which the rainfall occurred, e.g., rainfall falling and recorded from 1130 to 1230 is recorded against 1230. All the hourly values are entered by the user by navigating horizontally across the days. At the end of each day's entry the cursor moves to the column for entering the daily total as available in the manuscript. At the end of the last entry for the last day of the month, the cursor moves to the cells for entering the columnar totals for each hourly observation for the month as available in the manuscript. Finally, the monthly total as available in the manuscript has to be entered. The computed totals for each day, each hour across the month and for the month are filled automatically in the respective cells. Similarly, the maximum hourly rainfall recorded in the month and number of rainy days in the month are entered as available in the manuscript. The corresponding totals are computed and filled automatically by the system. Remarks, if available in the manuscript, can be entered on the daily basis.

There are three options for the graphical display of data: (a) to plot the data of any day from 0100 to 2400 hrs. in the form of hourly bar chart form (only for cases when the tabulation is not for standard hours (i.e. 0830 hrs to 0830 hrs next day), (b) to plot the data from 0830 hrs on any day to 0830 hrs of the next day in hourly bar chart form (see Fig. 6.2) and (c) to plot the tabulated rainfall data back in the form of a continuous trace so as to replicate that obtained by the autographic chart recorder (see Fig. 6.3). The least count of the simulated trace, of course, remain 1 hour which is the interval of rainfall tabulation).

The latter option is very useful in comparing the entered data with the analogue chart records. The user can navigate through different days in the month to see the corresponding plots without leaving the graph.

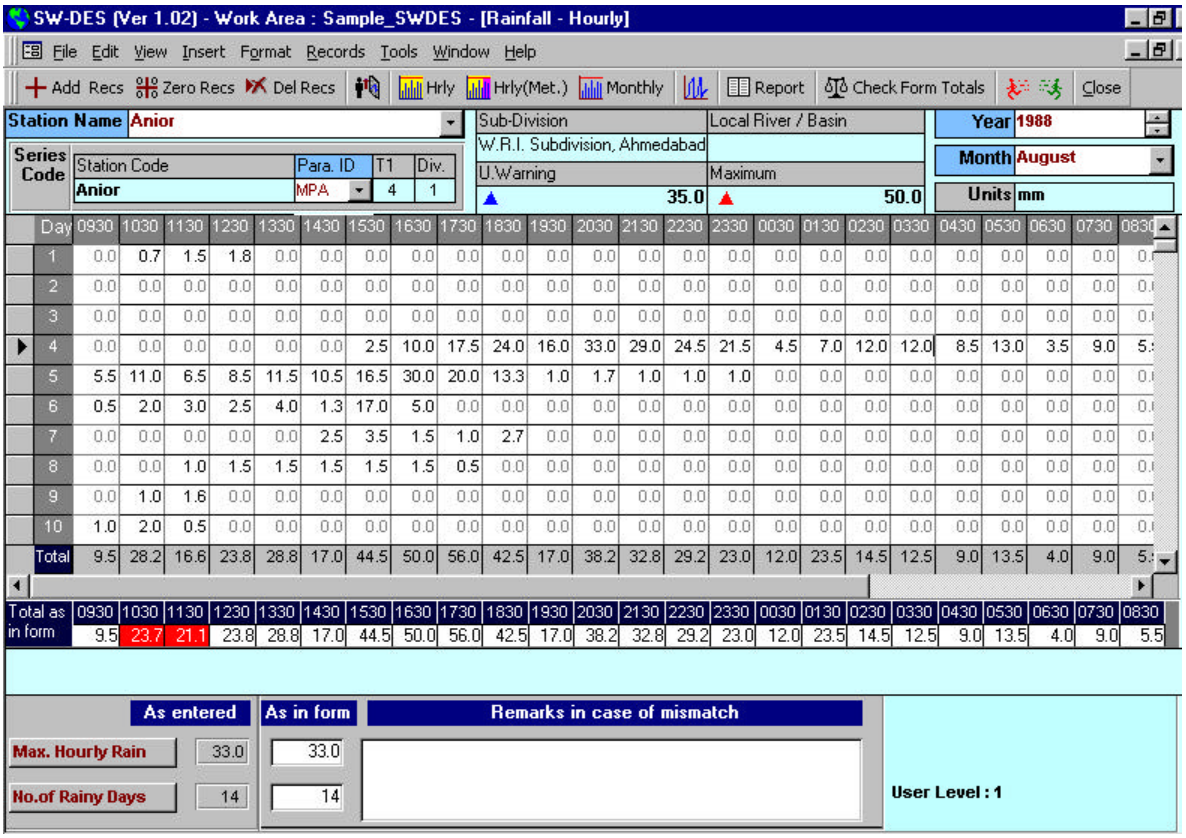

Fig. 6.1. Entry screen format for entering hourly rainfall data

Four types of data entry checks are performed for hourly rainfall data.

(a) The entered daily data can be compared against upper warning level and/or maximum values.
(b) The entered and computed values for daily total for each day are compared.
(c) A check is made to see if the entered and computed values of total for each hour across the month are in agreement.
(d) The monthly total rainfall, the number of rainy days in the month and the maximum hourly rainfall in the month are compared between the observer calculated values and the computed values. In case of any mismatch the user is prompted by colour highlighting so as to check back the entries.

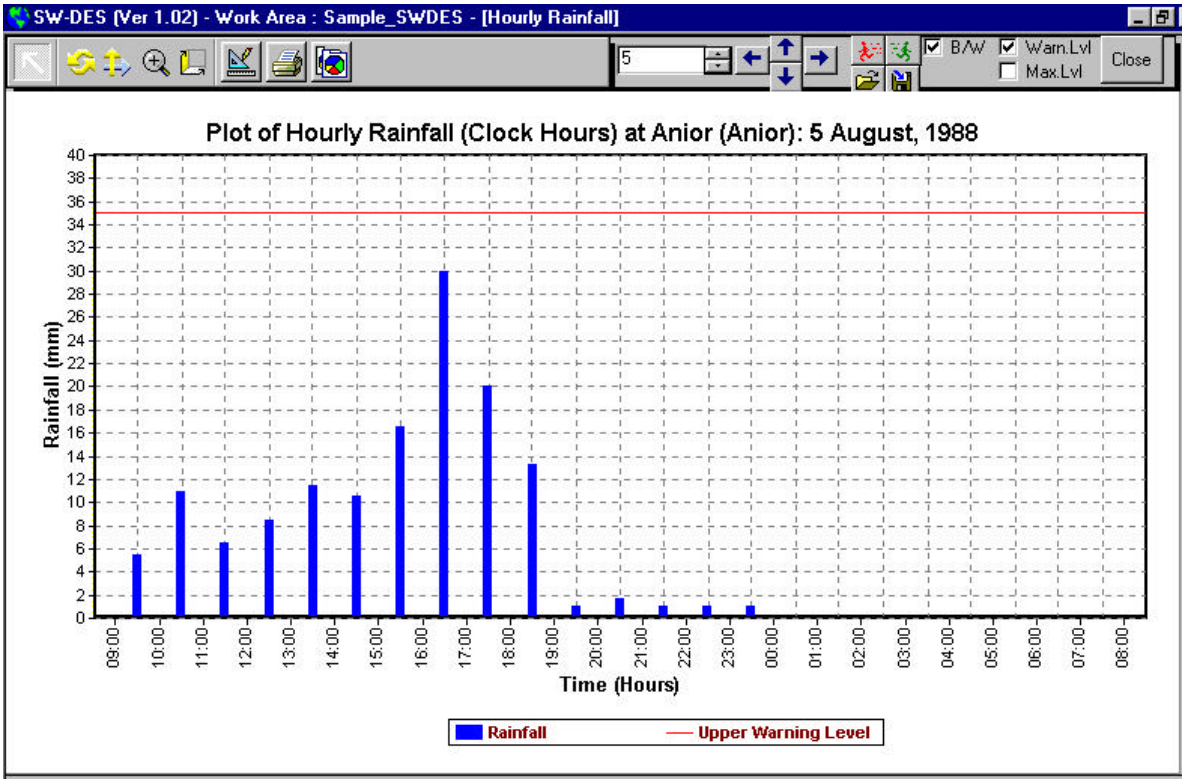

Fig. 6.2. Plot of hourly rainfall for a day

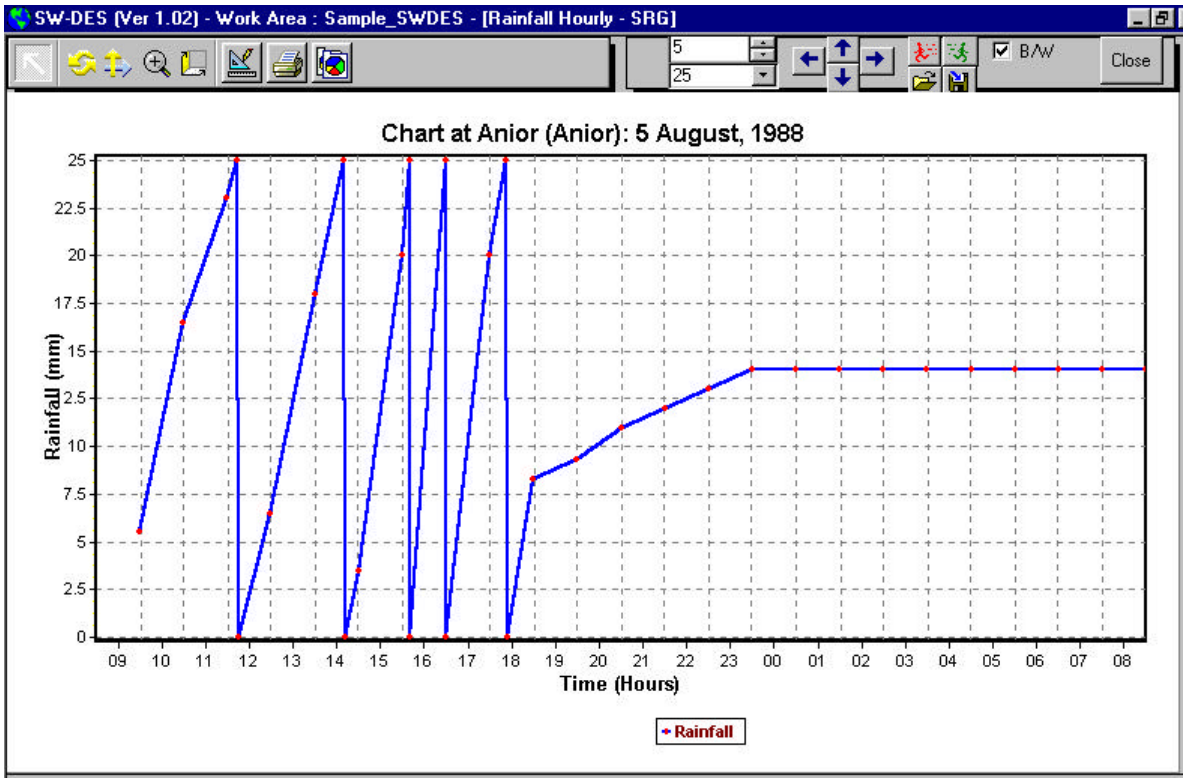

Fig. 6.3. Plot of hourly rainfall simulated as observed by the analogous chart recorder

7. Import/entry of digital data

A digital rainfall record is obtained from raingauge having a mechanism to record rainfall electronically into the data logger. Tipping bucket raingauges which register the occurrence of each tip of a bucket of known rainfall depth by means of a reed switch, provide such a facility.

There are two options by which such data can be stored.

(a) The timings for each tip of the bucket are stored.
(b) The other option is to store the number of tips or the amount of rainfall in a pre-set time interval. The pre-set interval can be 15 min., 30 min., 1 hour or any other desired interval.

In both cases the amount of rainfall corresponding to each tip of the bucket is also recorded. Since the former option gives the timings for each tipping to the nearest second it is possible to derive the rainfall values over any desired interval. Information on the station/instrument identification, year, month and day is also suitably stored.

Data stored on the data loggers can be downloaded in simple ASCII format with the help of the software and hardware as prescribed by the manufacturer. Loggers hold information on the instrument/station identification, amount of rainfall corresponding to each tip of bucket and timings for each tipping or amount of rainfall in the prescribed time interval. This information may be easily imported into the data entry system by having a conversion/import program which will read the output of the logger and import it. Such conversion/import programs have to be made for each type of data logger in the HIS. Since the format of the data obtained from the data loggers can be standardised to some extent there is possibility of developing appropriate templates which can then be employed for a range of data loggers. This feature will be developed according to the equipment acquired by the various states and agencies.

Rainfall observations stored in the data logger in the form of timings for each tip result in a data series of non-equidistant nature. By contrast, data observed as rainfall in the pre-set time interval forms an equidistant data series. Where the rainfall is recorded as a pre-set time interval and the interval is one hour, the same form may be used as for data recorded on autographic charts (As section 6 above). Where a different interval is required (15 min., 30 min. etc.) the general option for equidistant data series may be used. Data stored as tip timings must be stored using the layout for non-equidistant series. It is to note that provision of both these general options is yet to be made in the system.