HYDROLOGY PROJECT
Technical Assistance

World Bank & Government of The Netherlands funded

Training module # SWDP - 22

How to carry out primary validation of water level 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

Page

1. Module context

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

2. Module profile

Title	:	How to carry out primary validation of water level data
Target group	:	Assistant Hydrologists, Hydrologists, Data Processing Centre Managers):
Duration	:	One session of 60 minutes
Objectives	:	After training, the participants will be able to • carry out primary validation of water level data
Key concepts	:	• scrutiny of tabular and graphical data • comparison of water level data from two equipment at same station • validation against data limits • multiple graphs of water levels from adjacent stations • graph of water level u/s and d/s of the structures
Training methods	:	Lecture, software
Training tools required	:	Board, OHS, Computer
Handouts	:	As provided in this module
Further reading and references	:	

3. Session plan

No	Activities	Time	Tools
1	***General*** • Important points - (a) • Important points - (b)	5 min	 OHS 1 OHS 2
2	***Instruments and observational methods*** • General • Staff gauges • Automatic water level recorders • Digital water level recorders	10 min	 OHS 3 OHS 4 OHS 5 OHS 6
3	***Scrutiny of tabular and graphical data - single record*** • General • Example 3.1: Cross sectional profile • Example 3.1: Absolute limits • Example 3.1: Upper and lower warning limits • Example 3.1: Rates of rise and fall limits • Fig. 3.4: Graphical inspection of hydrograph • Fig. 3.5: Graphical inspection of hydrograph • Validation of regulated rivers	15 min	 OHS 7 OHS 8 OHS 9 OHS 10 OHS 12 OHS 13 OHS 14 OHS 15
4	***Comparison of daily time series for manual and autographic or digital data*** • General & validation procedure	5 min	 OHS 16
5	***Multiple graphs of water level at adjacent stations*** • General	5 min	 OHS 17
6	***Exercise*** • Setting min. & max. limits for a station from the cross-sectional profile available in SWDES • Inspection of water level hydrographs against the data limits • Inspection of water level hydrographs for possible anomalies	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 carry out primary validation of water level data

1. General

- **Data validation is the means by which data are checked to ensure that the final figure stored in the HIS is the best possible representation of the true value of the variable at the measurement site at a given time or in a given interval of time.** Validation recognises that values observed or measured in the field are subject to errors which may be random, systematic or spurious.

- **Stage or water level is the elevation of water surface above an established datum; it is the basic measure representing the state of a water body. Records of stage are used with a stage-discharge relation in computing the record of stream discharge. The reliability of the discharge record is dependent on the reliability of the stage record and on the stage discharge relation.** Stage is also used to characterise the state of a water body for management purposes like the filling of reservoirs, navigation depths, flood inundation etc. Stage is usually expressed in metres or in hundreds or thousandths of a metre.

- **Primary validation of water level data will be carried out at the Sub-divisional level using Primary module of dedicated hydrological data processing software and is concerned with data comparisons at a single station:**

  - for a single data series, between individual observations and pre-set physical limits
  - between two measurements of a variable at a single station, e.g. staff gauges and levels recorded by an automatic or digital water level recorder

- **Data entry checks will already have been carried out to ensure that there have been no transcription errors from the field sheets to the database. Some doubtful values may already have been flagged.**

- **Where a doubtful or incorrect value is identified, this should be marked with a 'flag' to indicate that it is 'suspect'. In some instances it may be possible to replace this value with a corrected value, in which case the replacement value is flagged 'corrected'. Otherwise 'suspect' values remain flagged for secondary validation.** Missing values may be interpolated as stage or as discharge depending on the nature and duration of missing or faulty records and the availability of records on which to base interpolation.

2. Instruments and observational methods

The method of measurement or observation influences our view of why data are suspect. To understand the source of errors we must understand the method of measurement or observation in the field and the typical errors of given instruments and techniques.

Data validation must never be considered a purely statistical or mathematical exercise. Staff involved in it must understand the field practice.

Three basic instruments are in use at river gauging stations for measurement of water level.

- Staff gauges
- Autographic water level recorders
- Digital water level recorders

2.1 Staff gauges

2.1.1 Instrument and procedure

The staff gauge is the primary means of measurement at a gauging station, the zero of which is the datum for the station. It is a manually read gauge and other recording gauges which may exist at a station are calibrated and checked against the staff gauge level. Staff gauges are located directly in the river. An additional staff gauge may be situated within the stilling well but this must not be used to calibrate recording instruments as it may be affected by blockage of the intake pipe. Where the staff gauge is the only means of measurement at a station, observations are generally made once a day in the lean season and at multiple times a day during a flood period - even at hourly intervals during flood season on flashy rivers.

2.1.2 Typical measurement errors

Staff gauges like other manual measurements, are dependent on the observer's ability and reliability and it must not be assumed that these are flawless. Checking on the performance of the observer is mainly the task of the field supervisor, but the data processor must also be aware of typical errors made by observers.

A common problem to note is the misplacement of decimal point for readings in the range .01 to .10. For example a sequence of level readings on the falling limb of a hydrograph:

4.12, 4.10, 4.9, 4.6, 4.3, 4.1, 3.99 - should clearly be interpreted as:
4.12, 4.10, 4.09, 4.06, 4.03, 4.01, 3.99.

Experience suggests that where the record is maintained by a single observer who is left unsupervised for extended periods of time, that it may contain some 'estimated' readings, fabricated without visiting the station. This may show up as sequences which are hydrologically inconsistent. Typical indicators of such 'estimates' are:

- Abrupt falls or a sudden change in slope of a recession curve.
- Long periods of uniform level followed by a distinct fall.
- Uniform mathematical sequences of observations, for example, where the level falls regularly by 0.05 or 0.10 between readings for extended periods. Natural hydrographs have slightly irregular differences between successive readings and the differences decline as the recession progresses.

In addition, precise water level measurement may be difficult in high flows, due to poor access to the gauge and wave action and, in flood flows the correspondence between staff gauges and recording gauges may not be as good as in low flows. Quality of gauge observations is of course, also affected if the gauge is damaged, bent or washed away. The station record book should be inspected for evidence of such problems.

2.2 Automatic water level recorders

2.2.1 Instrument and procedure

The vast majority of water level recorders in use in India use a float and pulley arrangement in a stilling well to record the water level as a continuous pen trace on a chart. The chart is changed daily or weekly and the recorder level is set to the current level on the staff gauge, which is also written on the chart at the time of putting on and taking off.

2.2.2 Typical measurement errors

Automatic water level recorders are subject to errors resulting from malfunction of the instrument or the stilling well in which it is located. Many of these errors can be identified by reference to the chart trace or to the level figures which have been extracted from it.

The following are typical malfunctions noted on charts and possible sources of the problems.

(a) Chart trace goes up when the river goes down
- Float and counterweight reversed on float pulley

(b) Chart trace goes up when the river goes down
- Tangling of float and counterweight wires
- Backlash or friction in the gearing
- Blockage of the intake pipe by silt or float resting on silt

(c) Flood hydrograph truncated
- Well top of insufficient height for flood flows and float sticks on floorboards of gauging hut or recorder box.
- Insufficient damping of waves causing float tape to jump or slip on pulley.

(d) Hydrograph appears OK but the staff gauge and chart level disagree. There are many possible sources including operator setting problems, float system, recorder mechanism or the operation of the stilling well, in addition to those noted above. The following may be considered.

<u>Operator Problems</u>
- Chart originally set at the wrong level

<u>Float system problems</u>
- Submergence of the float and counterweight line (in floods)
- Inadequate float diameter and badly matched float and counterweight
- Kinks in float suspension cables
- Build up of silt on the float pulley affecting the fit of the float tape perforations in the sprockets

<u>Recorder problems</u>
- Improper setting of the chart on the recorder drum
- Distortion and/or movement of the chart paper (humidity)
- Distortion or misalignment of the chart drum
- Faulty operation of the pen or pen carriage

<u>Stilling well problems</u>
- Blockage of intake pipe by silt.
- Lag of water level in the stilling well behind that in the river due to insufficient diameter of the intake pipe in relation to well diameter.

(e) Chart time and clock time disagree
- Chart clock in error and should be adjusted

In particular it should be noted that the partial blockage of the stilling well or intake pipe will result in a serious systematic error in level measurement.

2.3 Digital water level recorders

2.3.1 Instrument and procedure

Like the chart recorder many DWLRs are still based on a float operated sensor operating in a stilling well. One significant improvement is that the mechanical linkage from the pulley system to the chart is replaced by the shaft encoder which eliminates mechanical linkage errors and the imprecision of a pen trace. The signal from the shaft encoder is logged as level at a selected time interval on a digital logger and the information is downloaded from the logger at regular intervals and returned for processing. The level is set and checked with reference to the staff gauge.

Alternative sensors for the measurement of water level do not require to be placed in still water, notably the pressure transducer. Loggers based on such sensors have the advantage that they do not need to be placed in a stilling well and thus can avoid the associated problems.

2.3.2 Typical measurement errors

Measurement except for the sensors noted above is still subject to the errors caused by the float system and by the operation of the stilling well. Many of the same or equivalent checks are therefore necessary to ensure the continuity and accuracy of records. In particular the comparison and noting of staff gauge and logger water levels (and clock time and logger time) at take off and resetting, in the Field Record Book are essential for the interpretation of the record in the office.

Procedures in the office for checking the reliability of the record since the previous data download will depend on the associated logger software but should include a graphical inspection of the hydrograph for indications of malfunction (e.g. flat, stepped or truncated trace). Comparisons as for the chart recorder should be made with the observer's readings and bad or missing data replaced by manual observations.

3. Scrutiny of tabular and graphical data - single record

3.1 General description

The first step in validation is the inspection of individual records from a single recorder or manual measurement for violations of preset physical limits or for the occurrence of sequences of data which represent unacceptable hydrological behaviour.

Screening of some unacceptable values will already have been carried out at the data entry stage to eliminate incorrectly entered values.

Numerical tests of physical limits may be considered in three categories:

(a) Absolute maximum and minimum limits
(b) Upper and lower warning limits
(c) Acceptable rates of rise and fall

3.2 Absolute maximum and minimum limits

Checking against maximum and minimum limits is carried out automatically and values violating the limits are flagged and listed. The values of the absolute maximum and minimum levels at a particular station are set by the data processor such that values outside these pre-set limits are clearly incorrect. These values are normally set for the full year and do not vary with month or season.

The cross-section plot of the river gauging line in conjunction with the cross section of the control section at higher flow depths provides an appropriate basis for setting these minimum and maximum limits. With respect to minimum values for stage records, since it is normal to set the zero of the gauge at the level at which flow is zero then, for many stations a zero gauge level may be set as the absolute minimum. However, for some natural channels and controls negative stage values may be acceptable if the channel is subject to scour such that flow continues below the gauge zero. Such conditions must be confirmed by inspection of the accompanying Field Record Book.

Similarly, the absolute maximum is set at a value after considering the topography of the flood plains around the control section and also the previously observed maximum at the station. If long term data on water levels is already available (say for 15 –20 years) then the maximum attained in the past can be taken as an appropriate maximum limit.

3.3 Upper and lower warning limits

Validation of stage data against an absolute maximum limit does not discriminate those unusually high or low values which are less than the maximum limit but which may be incorrect. Less extreme upper and lower warning limits are therefore set such that values outside the warning range are flagged for subsequent scrutiny. The underlying objective while setting the upper and lower warning levels must be that such limits are violated 1–2 times every year by a flood event. This would ensure that on an average the one or two highest flood or deepest troughs are scrutinised more closely for their correctness. These limits may also be worked out using suitable statistics but care must be taken of the time interval and the length of data series under consideration. Statistics like 50% ile value of the collection of peak over the lowest maximum annual values used to set the upper warning level for the case of hourly data series of say 15-20 years. Of course, such statistics will also be subjected to the nature or shape of hydrograph which the station under consideration experiences. And therefore the appropriateness of such limits have to be verified before adopting them.

3.4 Rates of rise and fall limits

The method of comparing each data value with immediately preceding and following observations is of particular relevance to parameters exhibiting significant serial correlation such as water level data. The limit is set numerically as the maximum acceptable positive or negative change between successive observations. It should be noted however that what is an acceptable change in level during a rising flood hydrograph during the monsoon may be unacceptable during the dry season. Violations of rise and fall limits are therefore more readily identified from graphical plots of the hydrograph whilst numerical tests provide a relatively coarse screen.

It is a good practice and an essential requirement for an organised data processing activity that the listing of the entered data is obtained as a result of the above checking against various data limits. Such a listing will provide an appropriate medium for recording remarks/comments of the data processing personnel while validating the data. Any glaring deviations of the entered data from the expected one will also be apparent from such a listing.

Example 3.1:

Hourly water level data of Pargaon river gauging site in Pargaon catchment is considered during the period of 1995-1997. It is required to set various data limits for the purpose of validation of water level data at this station.

The cross-section of the gauging section is given in Fig. 3.1 from which it is apparent that the lowest bed level is about 507 m and the top of the bank is about 520 m. If the zero of the gauge is set at 507 m then the minimum data limit can be easily set at 507 m since it is not expected for water to reach this level even after moderate scouring of the bed. This gauging section happens to be at the bridge location it is therefore not appropriate to consider the cross section details of the same for setting up the maximum limit since the control section may be somewhere downstream of the bridge section and may have different levels at the flood plain levels. Such flood plain levels and the topography around the control section would be governing the levels at the gauging cross section. As a first estimate the top of the cross section at the bridge section can be taken as the maximum limit which is 520 m. An alternative is to use the highest water level ever recorded at the site which is about 519.5 m in the year 1997 (as per the available data). Fig. 3.2 (a) to Fig. 3.2 (c) shows the application of these two data limits for four years during 1994-97 respectively.

Considering that the objective of setting the upper warning level is to flag 1-2 flood events each year for closer scrutiny, a limit of 513 m as the upper warning levels will be effective. Such limits can also be arrived at using suitable statistics on the data derived for peak over a certain threshold limit.

For limits on maximum rate of rise and rate of fall it is best to use the historical data and obtain the derivative of the hourly water level series. After having seen the rate of change of water levels for a few years the maximum limits on rate of rise and fall can be easily obtained. For this case, limits of 2 and -1 m/hr is set for the maximum rate of rise and rate of fall respectively (Fig. 3.3). It may be seen from this plot that a few observations which are clear errors results in abnormal values of rate of rise or falls beyond the set limits even for small lower stages.

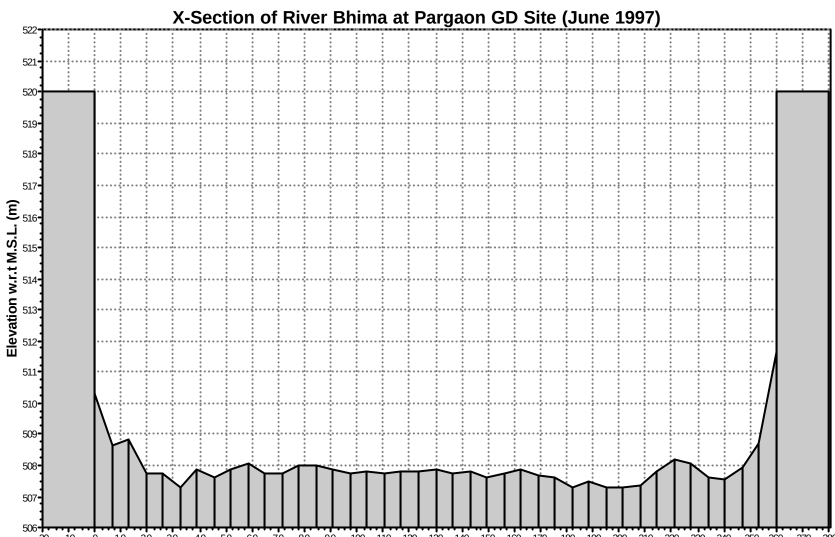

Fig. 3.1: Cross sectional profile of river gauging station at PARGAON.

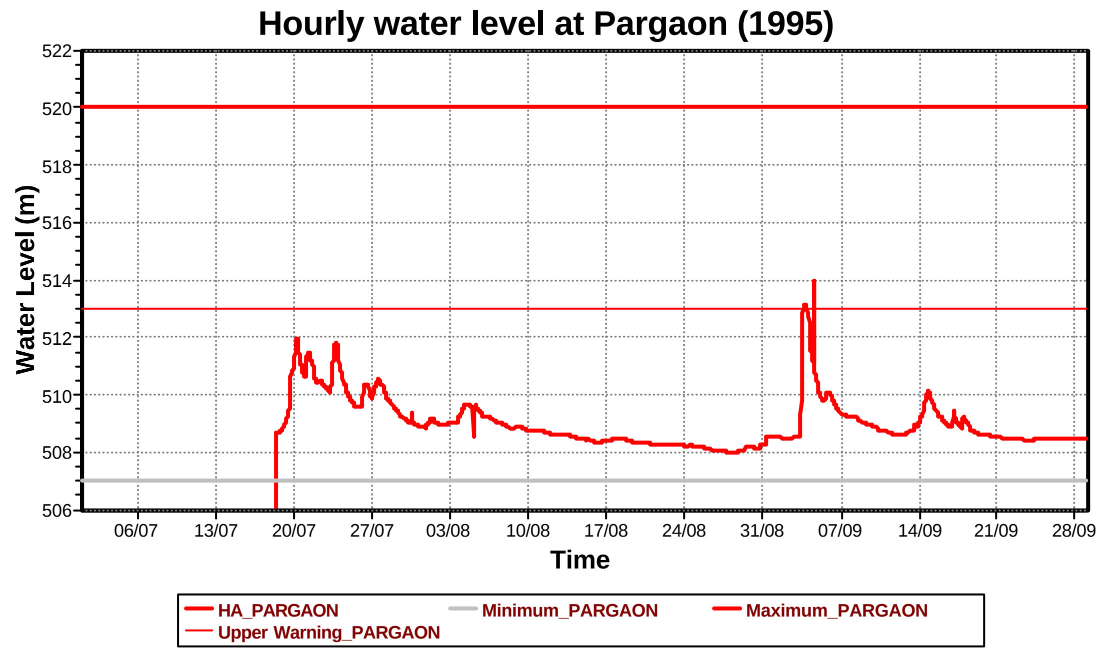

Fig. 3.2(a): Hourly water level data with data limits at PARGAON for 1995

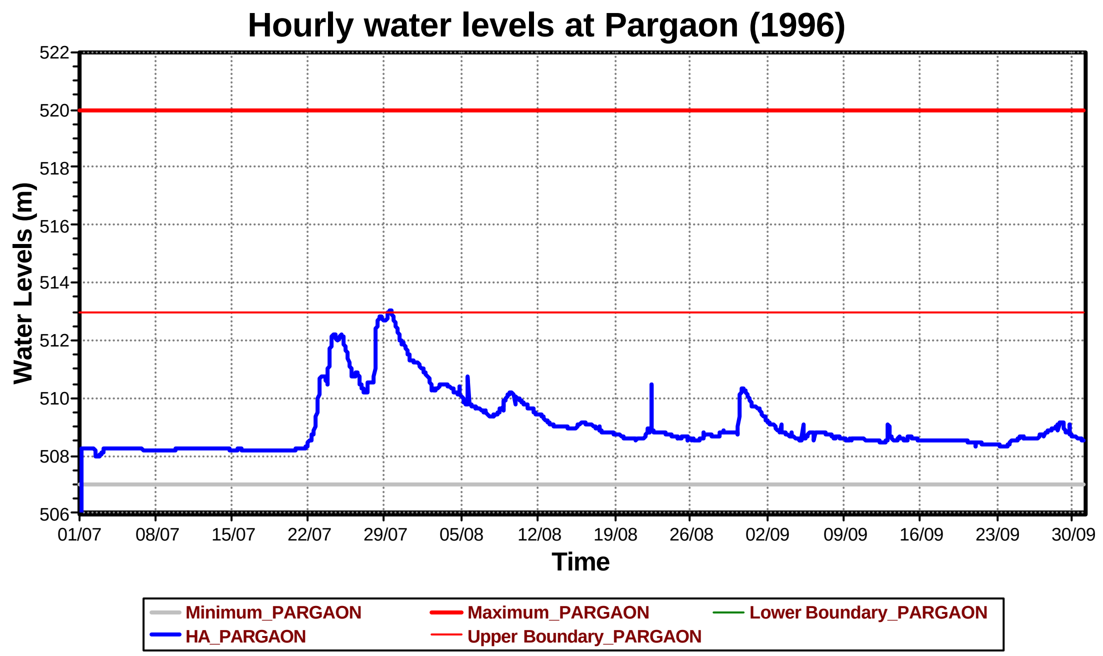

Fig. 3.2(b): Hourly water level data with data limits at PARGAON for 1996

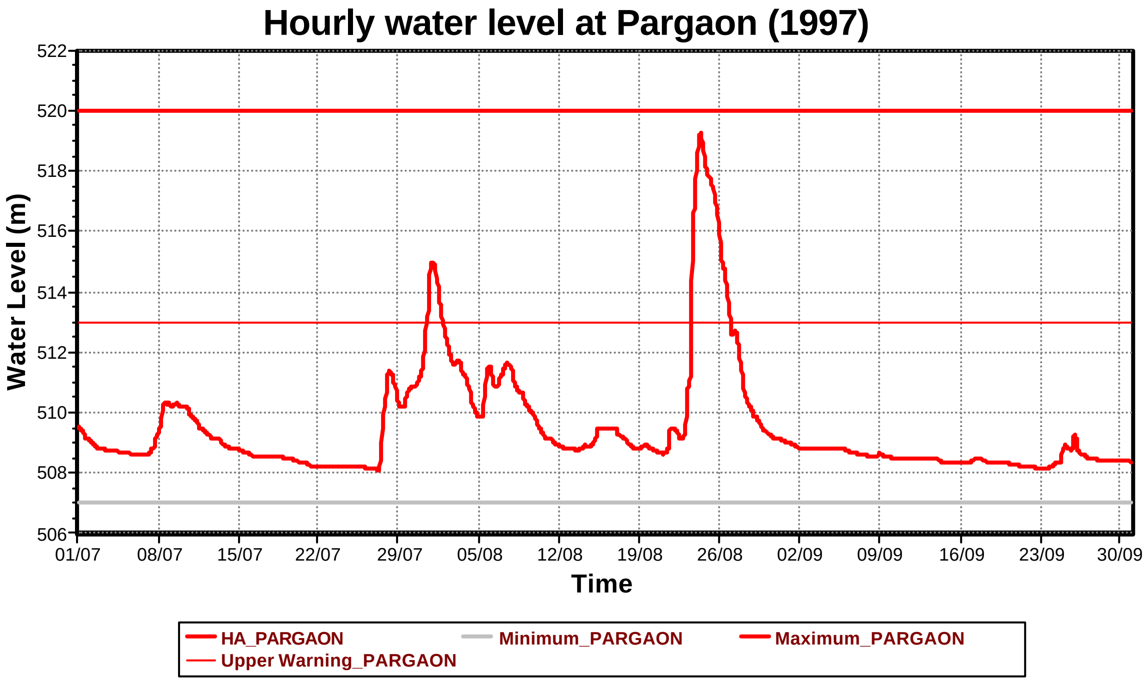

Fig. 3.2(c): Hourly water level data with data limits at PARGAON for 1997

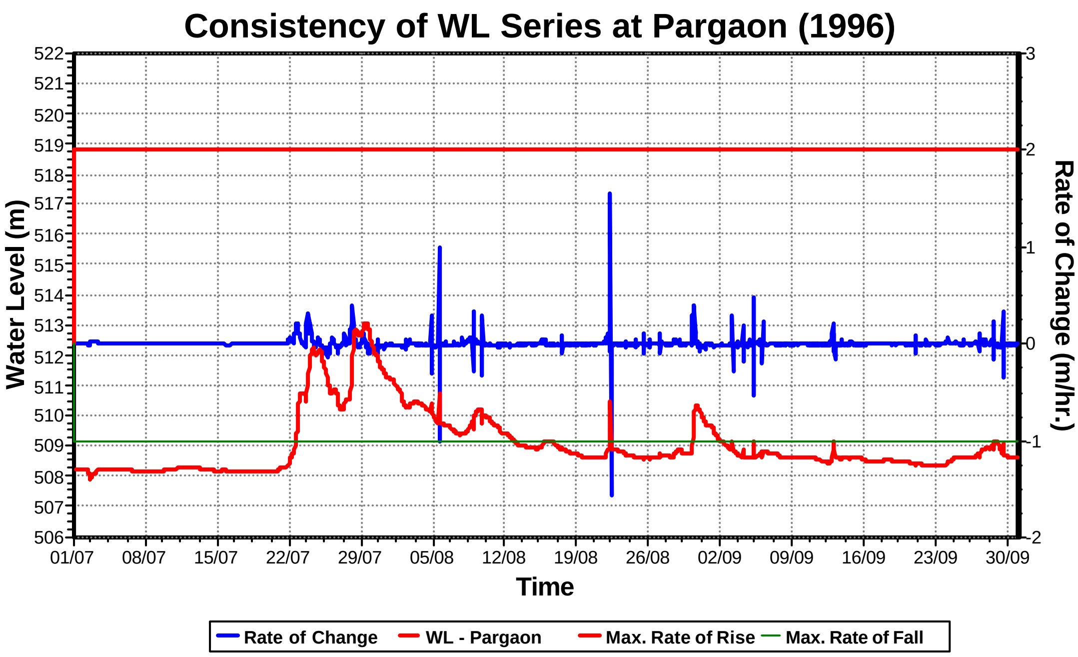

Fig. 3.3: Rate of change in hourly water level data with data limits at PARGAON for 1996

3.5 Graphical inspection of hydrographs

Visual checking of time series data is often a more efficient technique for detecting data anomalies than numerical checking and must be applied to every data set with an inspection of the stage hydrograph on screen. Screen display will also show the maximum and minimum limits and the upper and lower warning levels. Potential problems identified using numerical tests will be inspected and accepted as correct, flagged as spurious or doubtful and corrected where possible. An attempt must be made to interpret identified anomalies in terms of the performance of observer, instruments or station and where this has been possible to communicate this information to field staff for field inspection and correction. A few examples cases are discussed below:

Case 1

Case 1 represents a false recording of a recession curve (Fig. 3.4) caused by:

(a) An obstruction causing the float to remain hung
(b) Blockage of the intake pipe
(c) Siltation of the stilling well

This also shows where the obstruction was cleared. It may be possible to interpolate a true recession.

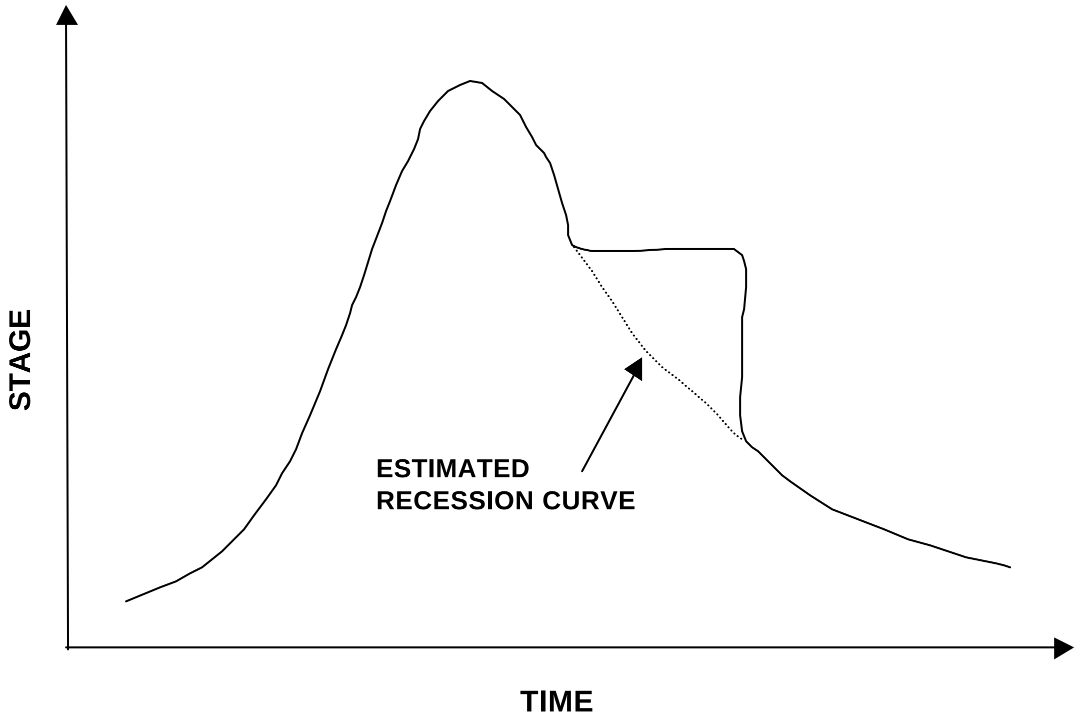

Fig. 3.4: False recording of recession curve

Case 2

Case 2 involves steps in the stage recording because of temporary hanging of the float tape or counterweight as the result of mechanical linkages in the recorder (Fig. 3.5). Such deviations can be easily identified graphically and true values can be interpolated.

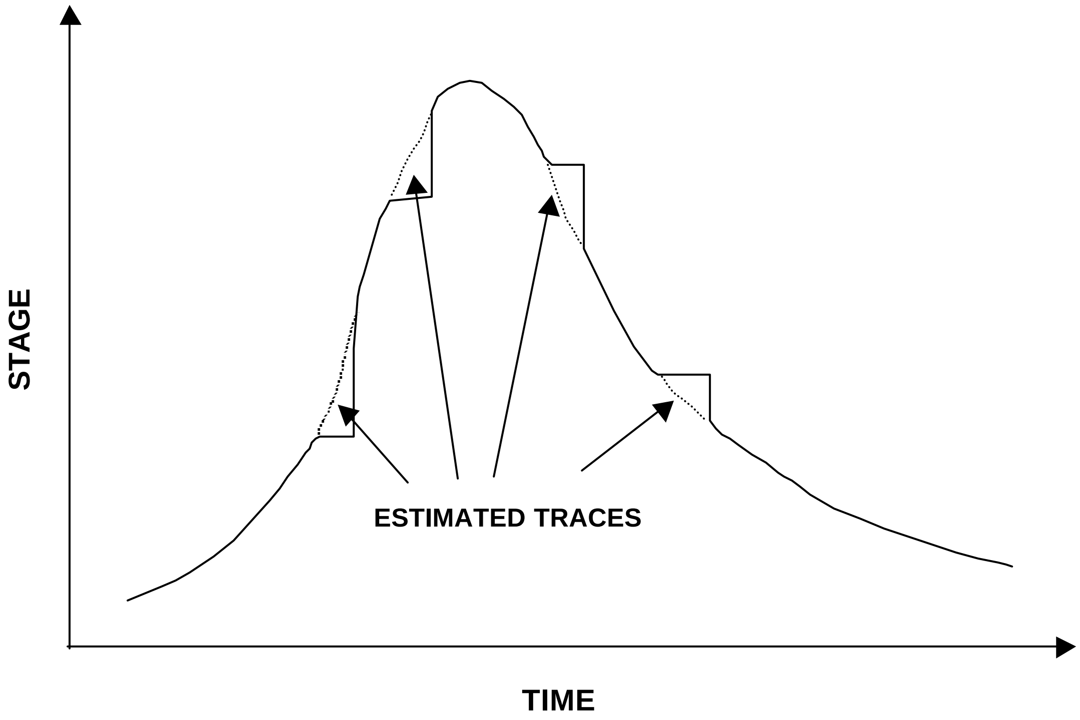

Fig. 3.5: Temporary hanging of the float

Case 3

Spurious peaks and troughs (spikes) in the hydrograph. These may be generated by observer error or occasionally by electronic malfunction of transducers or shaft encoders. It should be noted however that in some instances real variations of a similar nature may be generated:

(a) In rivers with small flow, switching on and off of pumps immediately upstream from the observation station will generate rapid changes

(b) The building of a bund or obstruction upstream from a station and its subsequent failure or release will generate first a negative spike followed by a positive one. In this case the levels as observed are correct.

3.6 Validation of regulated rivers

Rivers which are unaffected by river regulation and abstraction have a flow pattern which is determined by the rainfall and the transformation and storage of that rainfall on that basin on its way to the outlet. The hydrograph at such stations follows a natural pattern on which errors and inconsistencies can readily be identified.

Such natural rivers are not common in India; they are influenced artificially to a greater or lesser extent. One example is listed above (Case 3) **More generally on regulated rivers the natural pattern is disrupted by reservoir releases which may have abrupt onset and termination, combined with multiple abstractions and return flows.** The influences are most clearly seen in low to medium flows where in some rivers the hydrograph appears entirely artificial; high flows may still observe a natural pattern. In such rivers validation becomes more difficult and the application of objective rules may result in the listing of many queries where the observations are in fact correct. **It is therefore recommended that the emphasis of validation on regulated rivers should be graphical screening** by which data entry and observation errors may still be readily recognised.

The officer performing validation should be aware of the principal artificial influences within the basin, the location of those influences, their magnitude, their frequency and seasonal timing, to provide a better basis for identifying values or sequences of values which are suspect.

4. Comparison of daily time series for manual and autographic or digital data

4.1 General description:

At stations where water level is measured at short durations using an autographic or a digital recorder, a staff gauge reading is always also available. Thus, at such observation stations water level data at daily time interval is available from at least two independent sources. Discrepancies between reading may arise either from the staff gauge readings, the recorder readings or from both. The typical errors in field measurement have been described above and these should be considered in interpreting discrepancies. In addition errors arising from the tabulation of levels at hourly intervals from chart records or from data entry are possible.

4.2 Validation procedure

Two or more series representing the same level at a site are plotted on a single graph, where the two lines should correspond. A residual series may also be plotted showing the difference between the two methods of measurement. The following in particular should be noted:

- If there is a systematic but constant difference between staff gauge and recorder, it is probable that the recorder has been set up at the wrong level. Check chart annotations and the field record book. Check for steps in the hydrograph at the time of chart changing. The recorder record should be adjusted by the constant difference from the staff gauge.
- If the comparison is generally good but there are occasional discrepancies, it is probably the result of error in the staff gauge observations by the observer or incorrect extraction from the chart.
- If there is a change from good correspondence to poor correspondence in flood conditions, a failure associated with the stilling well or the recorder should be suspected and the staff gauge record is more likely to be correct and may therefore be used to correct or interpolate missing records in the recorder record.
- A gradual increase in the error may result simply from the recorder clock running fast or slow. This can be easily observed from the graphical plot and the recorder record should be adjusted.

5. Multiple graphs of water level at adjacent stations

Comparison of records between stations will normally be carried out as part of secondary validation at Divisional level and will not only be done with respect to discharge but also for stages. However an initial inspection may be done at Sub-divisional level where records for a few neighbouring stations are available. Such stations, especially if on the same river will show a similarity in their stage plot and inspection may help in screening out outliers.