HYDROLOGY PROJECT

Technical Assistance

World Bank & Government of The Netherlands funded

Training module # SWDP - 16

How to carry out primary validation for climatic 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 for climatic data
Target group	:	Assistant Hydrologists, Hydrologists, Data Processing Centre Managers
Duration	:	One session of 60 minutes
Objectives	:	After the training the participants will be able to: • carry out primary validation of climatic data
Key concepts	:	• tabular and graphical scrutiny • comparison of data from two equipment at same station • validation against various data limits • consistency between related climatic variables • re-computation of reported values • hourly rainfall and related climatic variables
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 Principles*** • Overhead – highlighted text • Overhead highlighted text and bullets	5 min	
2	***An example with lessons*** • Overhead – highlighted text • Flip chart – example • Overhead highlighted text	10 min	
3	***Primary validation of temperature*** • Overhead – headers 3.1 to 3.4 • Overhead – typical measurement errors	10 in	
4	***Primary validation of humidity***	5 min	
5	***Primary validation of wind speed***	5 min	
6	***Primary validation of atmospheric pressure***	5 min	
7	***Primary validation of sunshine***	5 min	
8	***Primary validation of pan evaporation*** • Overhead – headers 8.1 to 8.4 • Overhead – bulleted points typical measurement errors • Cartoon – Evaporation pan errors due to animals	10 min	

Note **It may be considered appropriate to skip very quickly through humidity, wind speed, atmospheric pressure and sunshine and to use the examples of temperature and pan evaporation to illustrate the general principles**

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 for climatic 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.

Improvement in computing facilities now enables such validation to be carried out whereas in the past, the volume of data and the time required to carry out comprehensive manual validation was prohibitive.

Primary validation of climatic data will be carried out at the Sub-divisional level using primary module and is concerned with data comparisons at a single station:

- for a single data series, between individual observations and pre-set physical limits
- for a single series between sequential observations to detect unacceptable rates of change and deviations from acceptable hydrological behaviour most readily identified graphically
- between two measurements of a variable at a single station, e.g. dry bulb thermometer and a thermograph recorder.

Comparisons between stations for climatic data are not considered in primary validation

2. An example with lessons

Data validation in its simplest form involves inspection of a data set and being able to say what is realistic and what is unrealistic and correcting the unrealistic if it is possible.

The question can then be asked - what is realistic? To the uninitiated it is just a set of numbers. How do we know that one set of numbers is better or worse than the ones which have been provided? To illustrate some general principles a simple example is used of a climatic number set. We are given dry bulb temperatures at a station at hourly intervals;

0700	25.0^{o}C
0800	28.0^{o}C
0900	40.0^{o}C

We could reasonably assume that there was a mistake in one or more of the readings, probably the last. We don't expect the temperature to rise by 12^{o}C in one hour. If the next reading at 1000 was 32^{o}C we would be even more confident that the figure of 40^{o}C entered for 0900 was incorrect and that the most likely cause was that the observer had incorrectly read 40^{o}C for 30^{o}C (or had read it correctly but written it down wrong). From this simple example we can draw several lessons.

(a) **The numbers in a data set are representative of a variable and the sequence of the variable represents a physical process which has physical behaviour and physical limits.** In the example above the numbers represent a physical process of atmospheric warming under the influence of solar radiation. A rise in temperature of 12^{o}C in one hour is quite unrealistic. **The lesson is that if we want to know what is realistic we must first understand the physical process, its behaviour and limits.**

(b) **The method of measurement or observation influences our view of why the data are suspect.** In the example above, we can make a reasonable guess of the source of the error because we know that the measurement was taken manually by an observer. Observers are not infallible; they make mistakes - even the best observers. **The lesson here is that 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 physical process and field practice.

(c) **Graphical inspection permits detection of errors which are more difficult to identify by numerical techniques.** Take for example a variation on the above set of temperatures. Say, the temperature listed in the data set at 0900 was 35°C would we still consider that the value was a mistake? Perhaps yes, but we would not be so certain; we would consider it suspicious but not impossible. If such a value were plotted on a graph for the day as available in SWDES and HYMOS, an error both for 40°C and for 35°C becomes a near certainty

Primary validation will therefore be considered, variable by variable, looking in summary at the physical process, behaviour and limits, and then its field measurement and typical errors, for the following variables:

- Temperature
- Humidity
- Wind speed
- Pan evaporation
- Atmospheric pressure
- Sunshine

3. Primary validation of temperature

3.1 Temperature variation and controls

Temperature is a measure of the ability of a body (in this case the atmosphere) to communicate heat to other bodies and to receive heat from them (IMD definition). Temperature varies primarily with the magnitude of solar radiation and observes cyclic diurnal and seasonal patterns. It is influenced at particular times by prevailing air masses and by the incursion of air masses from other source areas with different insolation properties and by prevailing cloudiness which limits incoming radiation. These factors limit the maximum and minimum temperatures which are expected at a given location for given season and time of day. They also limit the rates of change expected from hour to hour and from day to day.

With respect to location, temperature varies with latitude (which controls solar radiation), altitude and proximity to the ocean. Generally temperature is a spatially conservative element which has strong correlation (at least on an averaged basis) with neighbouring stations within the same air mass. There is normally a regular decrease in mean temperature with altitude at a rate of approximately 0.6°C per 100 metres for moist air and 0.9°C for dry air. Stations in close proximity to the sea have their temperature moderated by its influence so that maxima tend not to be so high, minima not so low, thus giving a reduced diurnal range.

Specific site conditions also affect the temperature measured. Stations in topographic hollows may experience temperature inversions in calm conditions, reversing the normal lapse rate of temperature with altitude. Stations sited in urban areas have generally higher temperatures than adjacent rural areas. The nearby prevailing ground cover, whether bare or vegetated, influences the measured temperature - including the proximity of trees which shade the site or of buildings which alternately shade and reflect heat to the station.

Validation based on location and site conditions are best considered in the comparison between stations and is discussed under secondary validation.

3.2 Temperature measurement

Temperature is periodically observed (once or twice daily) using a set of four thermometers, located in a thermometer (or Stevenson) screen, which from its construction and installation provides a standard condition of ventilation and shade. The four thermometers are:

Dry bulb thermometer -	measuring ambient air temperature
Wet bulb thermometer -	which provides a basis for calculating relative humidity
Maximum thermometer -	to indicate the highest temperature reached since the last setting
Minimum thermometer -	to indicate the lowest temperature reached since the last setting.

Graduations are etched on the glass stem of the thermometer. In the case of the dry bulb, wet bulb and maximum thermometers, observations are of the position of the end of the mercury column but in the case of the minimum thermometer, the reading is taken of the position of the end of the dumb-bell shaped index furthest from the bulb. Each thermometer has a calibration card which shows the difference between the true temperature and that registered by the thermometer. Corrections for a given temperature are applied to each observation. When the maximum and minimum thermometers have been read they are reset (twice per day for minimum; once per day for maximum) using a standard procedure.

Temperature is also measured continuously using a thermograph in which changes in temperature are recorded through the use of a bi-metallic strip. The temperature is registered on a chart on a clock-driven revolving drum and the measurement (chart) period may be either one day or one week. The observer extracts temperatures at a selected interval from the chart. The manually observed reading on the dry bulb thermometer is measured and recorded at the beginning and end of the chart period and if these differ from the chart value, a correction is applied to the chart readings at the selected interval.

3.3 Typical measurement errors

- Observer error in reading the thermometer, often error of 1^oC (difficult to detect) but sometimes 5^oC or 10^oC. Such errors are made more common in thermometers with faint graduation etchings.
- Observer error in registering the thermometer reading
- Observer reading meniscus level in minimum thermometer rather than index
- Thermometer fault - breaks in the mercury thread of the dry, wet or maximum thermometer
- Thermometer fault - failure of constriction of the maximum thermometer
- Thermometer fault - break in the spirit column of minimum thermometer or spirit lodged at the top or bubble in the bulb.
- Thermograph out of calibration and no correction made.

Thermometer faults will result in individual or persistent systematic errors in temperature.

3.4 Error detection

Many of the above faults will have been identified by the field supervisor or at data entry but others may be identified by setting up appropriate maximum minimum and warning limits for the station in question. These may be altered seasonally. For example, summer maximum temperature can be expected not to exceed 50^{o}C nor winter maximum temperature to exceed 33^{o}C.

Other checks carried out by SWDES include:

- Dry bulb temperature should be greater than or (rarely) equal to the wet bulb temperature.
- Maximum temperature should be greater than minimum temperature
- Maximum temperature measured using the maximum thermometer should be greater than or equal to the maximum temperature recorded by the dry bulb during the interval, including the time of maximum observation. The value of the maximum will be set to the observed maximum on the dry bulb if this is greater.
- Minimum temperature should be less than or equal to the minimum temperature recorded by the minimum thermometer during the interval, including the time of observation of the minimum thermometer. The value of the minimum will be set to the observed minimum on the dry bulb if this is lower.
- Thermograph readings at time of putting on and taking off should agree with the manually observed readings.

4. Primary validation of humidity

4.1 Humidity variations and control

The standard means of assessing the relative humidity or moisture content of the air is by means of the joint measurement of dry bulb and wet bulb temperature. From these two measurements, the dew point temperature, actual and saturated vapour pressures may also be calculated. The relative humidity (%) is the ratio of the actual vapour pressure to saturated vapour pressure corresponding to the dry bulb temperature. Whilst the actual vapour pressure may vary little during the day (except with the incursion of a new air mass), the relative humidity has a regular diurnal pattern with a minimum normally coinciding with the highest temperature (when the saturation vapour pressure is at its highest). It also shows a regular seasonal variation.

Generally relative humidity is a spatially conservative element which has strong correlation (at least on an average basis) with neighbouring stations within the same air mass. Stations in close proximity to the sea have higher relative humidities than those inland and a smaller daily range.

4.2 Humidity measurement

Wet and dry bulb thermometers used for temperature assessment also used for calculating various measures of humidity. The wet bulb is covered with a clean muslin sleeve, tied round the bulb by a cotton wick which is then led to a water container, by which the wick and muslin are kept constantly moist.

The observer calculates the relative humidity from the wet bulb depression using a set of tables.

Relative humidity may also be measured continuously by means of hygrograph in which the sensor is human hair whose length varies with relative humidity. The humidity is registered on a chart on a clock-driven revolving drum and the measurement (chart) period may be either one day or one week. The observer extracts humidity at a selected interval from the chart. A manually computed reading from dry and wet bulb thermometers is recorded at the beginning and end of the chart period and if these differ from the chart value, a correction is applied to the chart readings at the selected interval.

4.3 Typical measurement errors

Measurement errors using dry and wet bulb thermometers in the assessment of humidity are the same as those for temperature. In addition an error will occur if the muslin and wick of the wet bulb are not adequately saturated. Similarly there will be an error if the muslin becomes dirty or covered by grease. These defects will tend to give too high a reading of wet bulb temperature and consequently too high a reading of relative humidity. Errors in the hygrograph may also result from poor calibration or the failure to correct for manually observed values at the beginning and end of the chart period.

4.4 Error detection

Errors may be detected by setting up upper and lower warning limits appropriate to the station and season. The maximum is set at 100%. If the wet bulb is greater than dry bulb and the resulting calculated relative humidity is greater than 100%, then the observation will be rejected by SWDES. Graphic inspection of the daily series can be used to identify any anomalous values.

The observer calculated values of relative humidity may be compared with those calculated by the computer. However in view of the fact that the calculation can be done very simply in the office there seems little point in continuing the field calculation except in those cases where it is used to calibrate the hygrograph.

Accompanying field notes should be inspected for observations by the supervisor of errors in the thermometers or of a dry or dirty muslin. Hygrograph records can be inspected for departures of starting and finishing values measured by manual methods.

5. Primary validation of wind speed

5.1 Windspeed variations and controls

Wind speed is of particular importance in hydrology as it controls the advective component of evaporation. Wind speed exhibits wide variation not only from place to place but also shows strong diurnal variation at the same place. Wind flow and speeds are controlled by local pressure anomalies which in turn are controlled by the temperature. It may also be influenced by local topographic features which may funnel the wind and increase it above the areal average; conversely some stations will have wind speeds reduced by shelter. Extraordinary wind speeds may be experienced in some parts of the country through the incursion of tropical cyclones.

5.2 Measurement of wind speed and direction

Wind speed is measured using an anemometer, usually a cup counter anemometer. The rate of rotation of the anemometer is translated by a gear arrangement to read accumulated wind total (km) on a counter. By observing the counter reading at the beginning and end of a period, the wind run over the period can be determined and the average speed over the interval can be determined by dividing by the time interval. Standard Indian practice is to measure the wind speed over a three minute period as representing an effectively instantaneous wind speed at the time of observation. Daily wind run or average wind speed is also calculated from counter readings on successive days at the principal observation times.

Wind direction is commonly measured and may be used in the calculation of evapo-transpiration with respect to finding the fetch of the wind. It is observed using a wind vane and reported as 16 points of the compass either as a numerical figure or an alpha character

5.3 Typical measurement errors

Errors in windspeed might arise as the result of observer errors of the counter total, or arithmetic errors in the calculation of wind run or average wind speed. Instrumental errors might arise from poor maintenence or damage to the spindle which might thus result in reduced revolutions for given wind speed.

5.4 Error detection

Because of extreme variability in wind speed in space and time, it is difficult to set up convincing rules to detect suspect values. Nevertheless simple checks are as follows:

- Wind speeds should be zero where the direction is reported as ‘0’ (calm)
- Wind speeds cannot exceed 5 km/hr when the wind speed is reported as variable.
- Wind speeds in excess of 200 km per hour should be considered suspect and will result in a warning flag.

6. Primary validation of atmospheric pressure

6.1 Atmospheric pressure variations and controls

Atmospheric pressure is a measure of the weight of the air column vertically above a unit area. The principal variation is with altitude but a correction is always made to reduce the observed measurement of pressure to a standard sea level pressure so that spatial variations can be more readily investigated.

Pressure changes within a relatively narrow range and rates of change are comparatively slow. Lowest pressures and the most rapid rates of change are experienced in tropical cyclones.

Variations in atmospheric pressure are of great importance in weather forecasting but its influence on evapotranspiration is very limited and can often be assessed with acceptable accuracy with the use of a mean value of atmospheric pressure. It is of importance for pressure correction where non-vented pressure transducers are used for the measurement of water level.

6.2 Measurement of Atmospheric pressure

Atmospheric pressure is usually measured using a mercury barometer where the weight of the mercury column represents the atmospheric pressure. Commonly the Kew pattern barometer is used in India. It is read using a Vernier scale. Corrections are made for index error and for temperature (reducing to a standard temperature of 0^oC using a set of tables. It is also reduced to mean sea level pressure.

A barograph is also used for the continuous measurement of pressure. It consists of an aneroid sensor which expands and contracts with changes in pressure. These are registered on a clock-driven drum chart. Values of pressure may be extracted at hourly or other intervals from the chart and it is calibrated and set up to correspond with the reading using the more accurate mercury barometer.

6.3 Typical measurement errors

Observer errors may result from incorrect observation incorrect registration or in the application of corrections for temperature or reduction to sea level. Observation problems can result from the use of the Vernier scale.

Instrumental errors result from the entry of air into the space above the mercury and mechanical defects in the Vernier head.

6.4 Error detection

Primary validation is mainly through the setting up of upper and lower warning and maximum and minimum limits. Values outside the maximum and minimum limits are rejected; values outside the warning limits are flagged.

7. Primary validation of sunshine duration

7.1 Sunshine variations and controls

Sunshine duration is a very important contributor to the evapotranspiration equation and is widely used in the absence of direct measurements of radiation. The potential maximum sunshine duration varies regularly with latitude and with season. Actual sunshine also varies with ambient weather conditions and is generally lower during the monsoon than during the dry season. In urban areas the amount of bright sunshine may be reduced by atmospheric pollution and in coastal areas it may be reduced by sea mists.

7.2 Measurement of sunshine

The only instrument in common use in India for sunshine measurement is the Campbell Stokes sunshine recorder. This consists of a glass sphere mounted on a section of a spherical bowl. The sphere focuses the sun's rays on a card graduated in hours, held in the grooves of the bowl which burns the card linearly through the day when the sun is shining. The card is changed daily after sunset. Hence the sunshine recorder uses the movement of the sun instead of a clock to form the time basis of the record. Different grooves in the bowl must be used in winter summer and the equinoxes, taking different card types. The total length of the burn in each hour gives an hourly sunshine duration.

7.3 Typical measurement errors

The instrument is very simple in principle and the use of the sun rather than a clock as a time base avoids timing errors. Potential errors may arise from the use of the wrong chart which may result in the burn reaching the edge of the chart, beyond which it is not registered. Possible errors may result from extraction of information from the chart by the observer.

7.4 Error detection

SWDES may be used to detect and if necessary reject suspect values. Thus:

- Values of hourly sunshine greater than 1.0 or less than 0.0 are not permitted
- Sunshine records before 0500 and after 1900 are rejected and hence daily totals greater than 14.0 hours are rejected.

Daily warning limits may be set seasonally within SWDES based on the maximum possible sunshine for the location and time of year.

8. Primary validation of pan evaporation

8.1 Pan evaporation variations and controls

Evaporation is the process by which water changes from the liquid to the vapour state. Pan evaporation provides an estimate of open water evaporation. It is a continuous process in which the rate of evaporation depends on a wide range of climatic factors:

- amount of incoming solar radiation (represented by sunshine hours)
- temperature of the air and the evaporating surface

- saturation deficit - the amount of water that can be taken up by the air before it becomes saturated (represented in measurement by the wet bulb depression)
- wind speed

Evaporation again maintains a regular seasonal pattern with highest totals before the onset of the monsoon, during which evaporation is suppressed by decreasing saturation deficit.

8.2 Measurement of pan evaporation

The standard measurement in India is made using the US Class A pan evaporimeter. It is a circular pan 1.22 m in diameter and 0.255 m deep. It rests on a white painted wooden stand and is maintained level. The pan is covered by a wire mesh to avoid loss of water due to birds and animals. The inner base of the pan is painted white. A stilling well is situated in the pan within which there is a pointer gauge. Measurement must take account not only of evaporation losses but also gains due to rainfall; the raingauge nearby is used to assess the depth of rain falling in the pan.

On days without rain at daily (or twice-daily) reading, water is poured into the pan using a graduated brass cylinder (cup) to bring the level up precisely to the top of the pointer gauge. The number of cups (and part cups) is recorded and represents a depth of evaporation.

On days with rain since the last observation the rainfall may exceed evaporation and water must be removed from the pan to bring it to the hook level. The adjacent raingauge is used to assess the rainfall inflow.

On days with forecast heavy rainfall a measured amount of water may be removed from the pan in advance of the rainfall occurrence (to avoid pan overflow)

8.3 Typical measurement errors

- Observer errors - the observer over- or underfills the pan - such values will be compensated for the following day
- Instrument errors
  - ❖ Leakage: this is the most serious problem and it occurs usually at the joint between the base and the side wall. Small leaks are often difficult to detect in the field but may have a significant systematic effect on measured evaporation totals.
  - ❖ Animals may gain access to the pan, especially if the wire mesh is damaged
  - ❖ Algae and dirt in the water will reduce the measured rate of evaporation
  - ❖ Errors arise in periods of high rainfall when the depth caught by the raingauge is different in depth from the depth caught in the pan as a result of splash or wind eddies round the gauges.

8.4 Error detection

Warning and maximum limits may again xxxbe allocated to screen spurious values arising from observer error, leakage, animal interference or dirty water.

Where leakage has been identified and is recorded in the field record book, the records for a period preceding the discovery will be inspected and flagged as suspect and for review under secondary validation