3. Data Handling

The main topics covered in this Chapter are the types and frequency of measurement recorded at the field stations and the data handling procedures both in the field station and at project headquarters.

Table 2: Yarsha *Khola* Watershed: List of Hydrological Stations and Catchment Characteristics

Site No	River (site)	Elevation (masl)	Catchment size (sq.km.)	Alt. Range (masl)	Land cover (%)	
1	Yarsha *Khola* (hydro main)	1000	53.38	1000-3050	*Khet*: 13.9 *Bari*: 37.4 Forest: 31.5	Grass: 5.7 Shrub: 5.0 Other: 6.5
2	Gopi *Khola* (Gopi)	1040	17.37	1040-2495	*Khet*: 13.3 *Bari*: 40.8 Forest: 31.4	Grass: 3.9 Shrub: 5.0 Other: 5.7
3	Gopi *Khola* (Gairimudi)	1440	4.90	1440-2407	*Khet*: 3.7 *Bari*: 58.0 Forest: 22.7	Grass: 3.9 Shrub: 6.1 Other: 5.7
4	Yarsha *Khola* (Forest Site)	1970	1.11	1970-2594	*Khet*: 0 *Bari*: 16.2 Forest: 82.0	Grass: 0.9 Shrub: 0.9 Other: 0
5	Khahare *Khola* (Thulachaur)	2280	0.32	2280-2731	*Khet*: 0 *Bari*: 15.6 Forest: 40.6	Grass: 3.1 Shrub: 25.0 Other: 15.6
7	Khahare *Khola* (Bagar)	1740	2.08	1740-2731	*Khet*: 0 *Bari*: 50.5 Forest: 14.4	Grass: 10.1 Shrub: 20.2 Other: 4.8

Note: *Khet* – irrigated rice land
Bari = rainfed agricultural land

Table 3: Yarsha *Khola* Watershed: List of Meteorological Stations

Site No	Site name	Elevation (masl)
1	Hydro main	1,020
3	Gairimudi	1,530
4	Forest site	1,960
5	Thulachaur	2,300
6	Jyamire	1,960
7	Bagar	1,680
8	Nimkot	1,420
9	Namdu	1,380
10	Thuloban	2,640
11	Mrige	1,650
12	Pokhari	2,150

Table 4: Yarsha *Khola* Watershed: List of Erosion Plots

Site No	Site name	Elevation (masl)	Plot characteristics
5	Thulachaur	2,300	Grass, shrub; dark brown soil; Slope (deg): 19.1
6	Jyamire	1,960	*Bari*; brown soil (Chimteilo Pusro); Slope (deg): 17
9a	Namdu	1,380	*Bari*; red soil; Slope (deg): 17.5:
9b	Namdu	1,380	Grass, fallow, degraded; red soil; Slope: (deg): 17.5 (see also Fig. 6)

Note: *Bari* = rainfed agricultural land

Figure 11: The forest subcatchment (monitored at site No 4) in the Yarsha Khola Watershed, Nepal: 82% forest, 16.2% bari, 0.9% grass, 0.9% shrub (photo: November 1996)

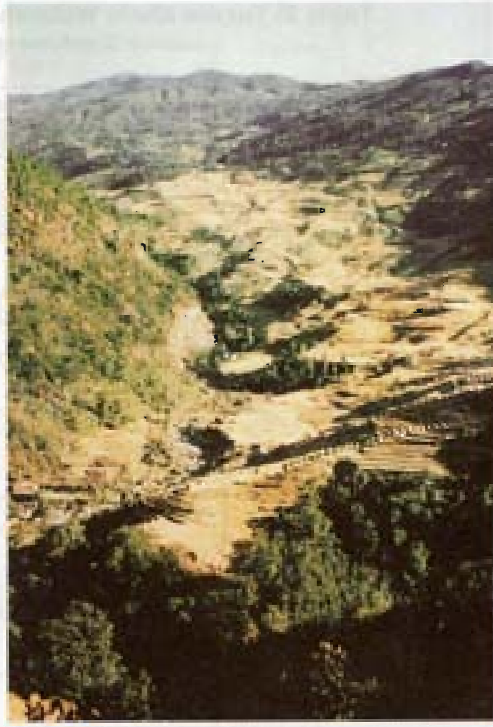


Figure 12: The subcatchment of the Gopi Khola (monitored at site No 2) in the Yarsha Khola Watershed, Nepal: 13.3% khet (lower part of the watershed), 40.8% bari, 31.4% forest, 3.9% grass, 5% shrub, 5.7% others (photo: November 1996)

Sub-topics describe the actual types of measurement recorded at the different field research stations and sites, the vital role played by the field teams maintaining the stations and recording the results, and the means through which quality data collection, recording, management, and storage are ensured in both field and office.

A well-organized and transparent data handling system is essential for good results. The data handling has to be defined and coordinated right from the beginning of project activities in order to achieve comparable data sets and results in the different watersheds of the PARDYP project. Figure 13 provides a flow chart for the data handling. In the following sections, the most important elements of this flow chart are discussed. The analysis is covered in Section 4.

3.1 The Measurement Programme

3.1.1 Overview

Common procedures and methodologies are crucial in the PARDYP project. This key Section specifies the basic, compulsory measurement programme of the hydrology and meteorology components in the PARDYP watersheds (Table 5). On top of these compulsory measurement

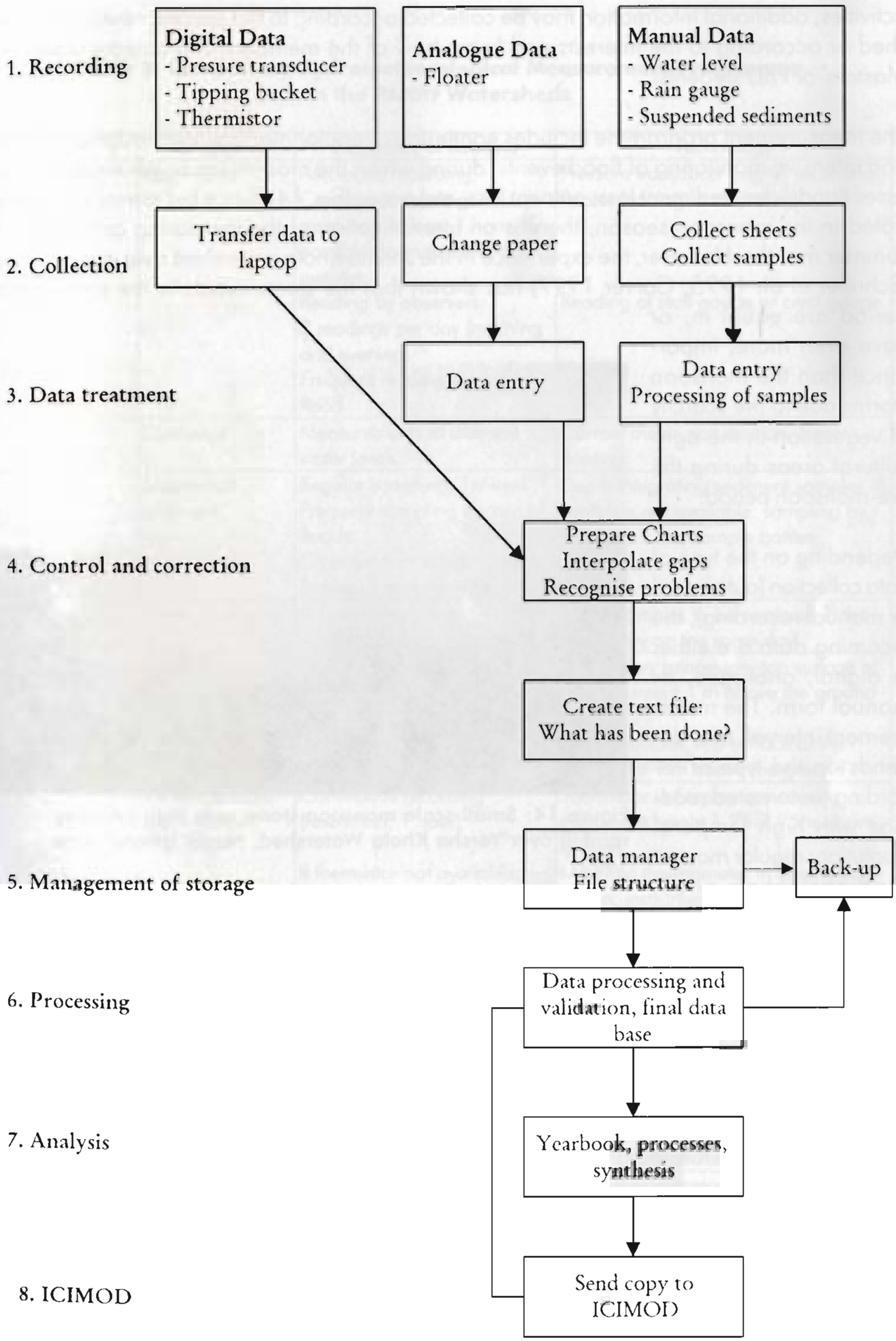


Figure 13: Handling of the hydrological and meteorological data in the PARDYP watersheds

activities, additional information may be collected according to the specific needs of a watershed or according to the interests and know-how of the members of a country team (e.g., Masters' or PhD students).

The measurement programme includes continuous, regular monitoring throughout the year and intensive monitoring of flood events, during which the most dramatic geomorphic processes (landslides, sediment loss, nutrient loss, etc) occur (Fig. 14). Since big storms are concentrated in the monsoon season, there is an intensification of the measuring activities in the summer months. However, the experience in the Jhikhu *Khola* watershed over previous years (Schreier et al. 1995, Carver 1997) has shown that the storm events in the pre-monsoon period are equal in, or have even more, importance than the monsoon storms due to the scarcity of vegetation in the agricultural areas during the pre-monsoon period.

Figure 14: Small-scale monsoon storm with high intensity rainfall over Yarsha Khola Watershed, Nepal (photo: June 1997)

Depending on the type of data collection (automated or manual recording), the incoming data are either in digital, analogue, or manual form. The measurement interval, too, depends on the type of recording (automated readings with high temporal resolution; regular manual readings taken once, twice, or three times a day). The time of the manual measurements has to be defined according to the timing of the readings in the national network; furthermore, it has to be the same for all the stations in the watershed.

The management of the measurement network in the PARDYP watersheds is the responsibility of field technicians (downloading of data, discharge measurements) as well as of local readers (manual data recording, assistance in discharge measurements, supervision of stations). The jobs of the field technicians and the readers are very important and they have a lot of responsibility. The fact that the main monsoon storms occur at night makes their task even more difficult!

The following sections provide details on the measurement programme listed in Table 5.

3.1.2 Water Level

For automated water level recording, pressure transducers (Fig. 15) as well as floaters (Fig. 16) are in operation in the PARDYP watersheds. The installation of pressure transducers is essential on

Table 5: Compulsory Hydrometeorological Measurement Programme in the PARDYP Watersheds

	Parameters	Frequency of Reading	Technology
Hydrological stations	Water level	Automatic recording: Continuous reading: resolution at least 1 hour, if possible down to 15 or 5 minutes Reading by observers: 2 readings per day (morning and evening) Frequent reading in case of flood	Pressure transducer (digital recording), floater (analogue recording) Reading of staff gauge or crest gauge
	Discharge	Measurements at different water levels.	Current meter, salt dilution, tracer, floating
	Suspended sediment	Regular sampling, 1x/week Frequent sampling in case of floods Cross-section sampling (only in special programmes)	Depth integrating sediment samples (if samplers not available: sampling by hand). 0.5 litre sample bottles
	Water chemistry	4-6 samples a year	Water samples to be analysed in the laboratory on the same day!
Meteorological stations	Rainfall (each station)	Daily rainfall (reading in the morning) Rainfall intensity with high temporal resolution	8" ordinary raingauge, top surface of the instrument 1 m above the ground Tipping bucket, top surface of the instrument 1 m above the ground
	Air temperature (each station)	Continuous recording (resolution: 1 hour) If thermistor not available: max/min in 24 hours	Thermistor in Stephensen screen, instrument height 1.25-1.75m above ground Max/min thermometer in Stephensen screen, instrument height 1.25-1.75m above ground
	Soil temperature (at least main stn.)	Continuous recording (resolution 1 hour)	Thermistor, 20 cm below surface If 3 sensors at the main met. station: 5,20, 50 cm below surface
	Air humidity (at least main stn.)	Readings 3x per day (morning, 2pm, evening)	Hair hygrometer in Stephensen Screen, instrument height 1.25-1.75m above ground Dry/wet bulb thermometer in Stephensen Screen, instrument height 1.25-1.75m above ground
	Other parameters (main met station)	Daily or continuous readings, depending on the equipment	Manual or automated instruments
Erosion plots	Erosion and runoff	Reading and sampling after each event	Measurement of water level and sampling from each drum

sites in which fixing a vertical floater pipe is impossible. Crest gauges, which mark the maximum water-level reached during an event, are only useful if there is no automatic water-level recording device or for purposes of comparison. In the monsoon season, the temporal resolution of the automatic water level recording is higher than in the dry season. For example, in the Yarsha *Khola* watershed, the recording interval in the monsoon season is 5' (in statistical mode in the order of average wave actions) and in the dry season 15' (momentary values). In the case of floaters, weekly charts are used in the monsoon season, monthly charts in the dry season. The hydrological monitoring sites are equipped with a station form on which each manipulation carried out at the station, or each observation (functioning of the equipment, comparison of staff gauge reading and reading of the instrument), have to be recorded with date and time.

Figure 15: Pressure transducer installed at the main hydrological station in Yarsha Khola Watershed, Nepal (photo: May 1997)

3.1.3 Discharge

There are three methods for sophisticated discharge measurements: current meter, salt dilution, tracer. These methods are briefly presented. (For more information see WMO 1980; LHG 1982; Fischer 1982; Chow et al. 1988; LHG 1994; and Spreafico and Gees 1997.)

Figure 16: Floater House at the hydrological sub-station in the Xi Zhuang Watershed, China (photo: February 1998)

- For current meter measurements (propeller device), a more or less fixed, homogeneous cross-section is required with a smooth, regular flow.
- The salt dilution method is essentially used in mountain rivers with rough cross-sections, big boulders, and turbulent flows. A conductivity meter is needed for the measurement. Turbulent river flows are a condition for this measurement technique in order to guarantee a good mixing of the

salt. Before starting with 'serious' measurements, experimental measurements are necessary in order to find the sites suitable for salt injection and for conductivity measurements which ensure an optimum mixing length. For river flows above $2m^3$/sec, the application of the dilution method becomes unrealistic due to the amount of salt required (5-10kg).

The tracer measurements (e.g., using fluorescent tracers) are again appropriate in turbulent mountain rivers, and their application has no upper limit in terms of flow. In preparing for the measurements, similar steps to those used with the salt dilution method (identification of injection point, sampling point, optimum mixing length) have to be carried out. The tracer is injected with the help of a Mariott Bottle (Fig. 17). The disadvantage of this techniques is that the calculation of discharge cannot be carried out directly in the field. The water samples (16 per measurement in Yarsha *Khola*) have to be taken to the laboratory for analysis, and the spectrofluorometer required for the analysis is extremely expensive.

Figure 17: Preparation for a tracer measurement with a Mariott Bottle at the main hydrological station, Yarsha Khola Watershed, Nepal (photo: monsoon season 1997)

This short introduction to the discharge measurement techniques illustrates that, for each hydrological station and for each season, the appropriate method has to be used. In Yarsha *Khola* watershed, all three measurement techniques are used. The current meter at Site 1 in low and medium flow, salt dilution at Sites 2,3,4,5, and 7, and the tracer (uranin) technique at Sites 1,2,3, and 4 during high flows (location of sites: Fig. 10). For comparison purposes, it may be interesting to apply different methods simultaneously. Current meter and salt dilution methods are applicable in all the five PARDYP watersheds. Measurement with fluorescent tracers, however, is only feasible in two watersheds of Nepal at present, as a spectrofluorometer for the analysis of the tracer samples is only available at the laboratory of the Department of Hydrology and Meteorology in Kathmandu.

Based on discharge measurements, rating curves are established – these describe the relationship between water level and discharge (see Section 4.4.2 and Fig. 18). In order to get a reliable rating curve for a specific station, discharge measurements for as many water levels as possible have to be carried out. For low flow conditions, this is not difficult as there are many occasions for measurement, and the measurements are easy. Much more difficult is the situation at high or flood flow conditions, as high water levels do not occur very often. Each discharge measurement technique requires a certain time during which the flow should remain constant. High or peak flows, however, are usually of short duration only. The attempt has to be made, therefore, to measure as many high flows as possible, for which the water level does not

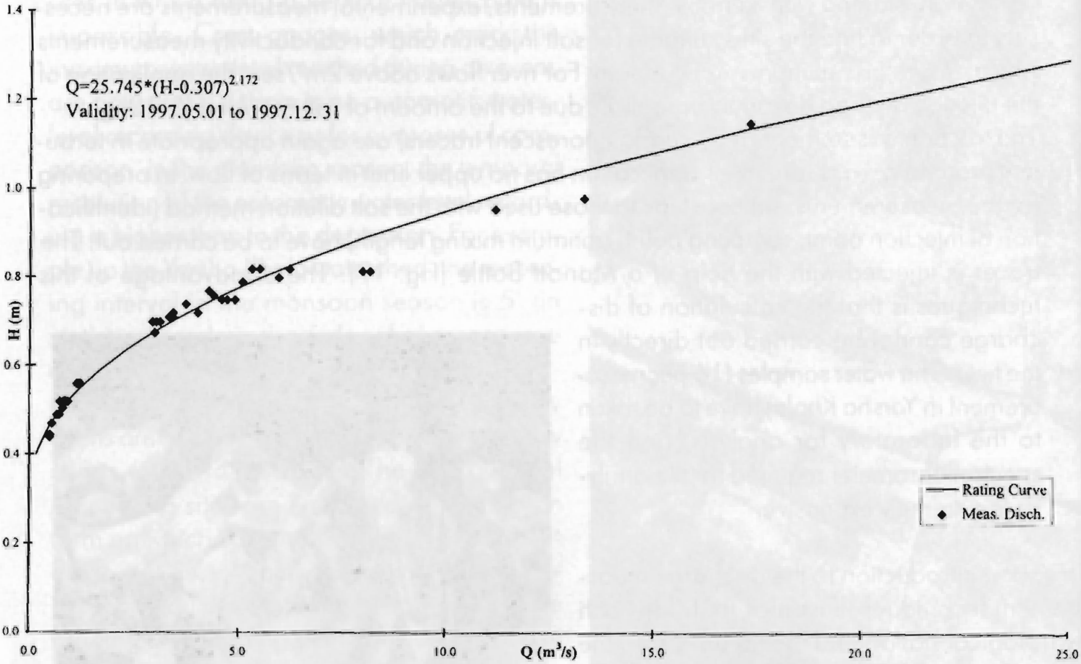


Figure 18: Water level/discharge rating curve of the Yarsha Khola at the main station, 1997

change too fast. However, it is almost impossible to measure peak flows accurately, although a rough estimate can be made, e.g., by using floats, by making one current meter measurement in the middle of the river, etc.

Obtaining good quality discharge measurements is a challenge. To realise such measurements, a small team of three to four persons is required, ideally consisting of a hydrologist, a field technician, the reader, and the assistant reader of the particular measuring site. In view of these personnel requirements and the expensive equipment needed, discharge measurements cannot usually be carried out simultaneously at each station. The team has, therefore, to work out a measurement programme for the monsoon season in which each station receives sufficient attention. Ideally, the team has to spend approximately two weeks at a stretch at a specific station. Within this period, the chances of encountering many different flow situations, including flood flow conditions, are at least moderate.

3.1.4 Sampling of Suspended Sediments

Only the sampling of suspended sediments is included in the compulsory measurement programme of the PARDYP watersheds. The investigation of bedload transport is optional (Table 6). Ideally, the sampling is carried out with depth-integrating sediment samplers, either from the bank of the river using a measuring rod or from a bridge using a handline. If sediment samplers are not available, the sampling can also be carried out manually from the river bank, slowly moving the bottle down and up over a vertical profile (Fig. 19). The sampling is carried out with 0.5 litre sample bottles; each sample has to be labelled properly with date, time, water level,

Table 6: Level of Disciplinary Analyses - Compulsory and Specialised Analyses

	METEO	HYDRO
Compulsory	• Spatial rainfall patterns • Temporal rainfall patterns • Analysis of high rainfall intensities • Comparison of rainfall of a specific season/year with long-term data series • Spatial and temporal temperature patterns • Calculation of evaporation • Analysis of other meteorological parameters	• Rating curve: water level/discharge • Rating curve: discharge/suspended sediments • Annual hydrographs of the 'main values' • Duration curve of discharge • Balances of runoff, discharge and sediment transport • Discharge comparison of different stations
Specialised	• Frequency and probability of wet and dry spells • Calculation of potential evapotranspiration • Analysis of soil temperature and soil humidity	• Concentration times • Space correlation, point correlation of different stations • Map: areas of high water potential and high water demand • Groundwater investigations • Assessment of sediment potential from different areas • Bedload transport

event number, station, and reader's name. For the establishment of a discharge/sediment rating curve, the water level has to be recorded parallel to each sediment sample.

A considerable portion of the sediments and nutrients is transported during events. It is important to know when the peak concentration of sediments is reached in a flood and when the maximum amount of nutrients is washed out. Therefore, sediment sampling periods concentrate on flood events during which at least 10 to 12 samples should be taken on the rising limb, the peak, and the falling limb of the flood hydrograph. In addition, flood sampling is crucial for the establishment of a reliable discharge/sediment rating curve, since the information for the upper part of the curve is rare. For the flood sampling, the reader has to be instructed to start sampling once the water level rises considerably above the base flow of the respective season. Not each sampled event, however, will then be included in the analysis (see 3.2.2).

Figure 19: The reader taking a sample of suspended sediment at the main hydrological station in Yarsha Khola Watershed, Nepal (photo: monsoon season 1997)

Ideally two to four events in the monsoon season, one to three events in the pre-monsoon season, and one to three events in the post-monsoon season should be measured successfully. This temporal differentiation is important as the sediment output during pre-monsoon, monsoon and post-monsoon storms is considerably different (Fig. 20).

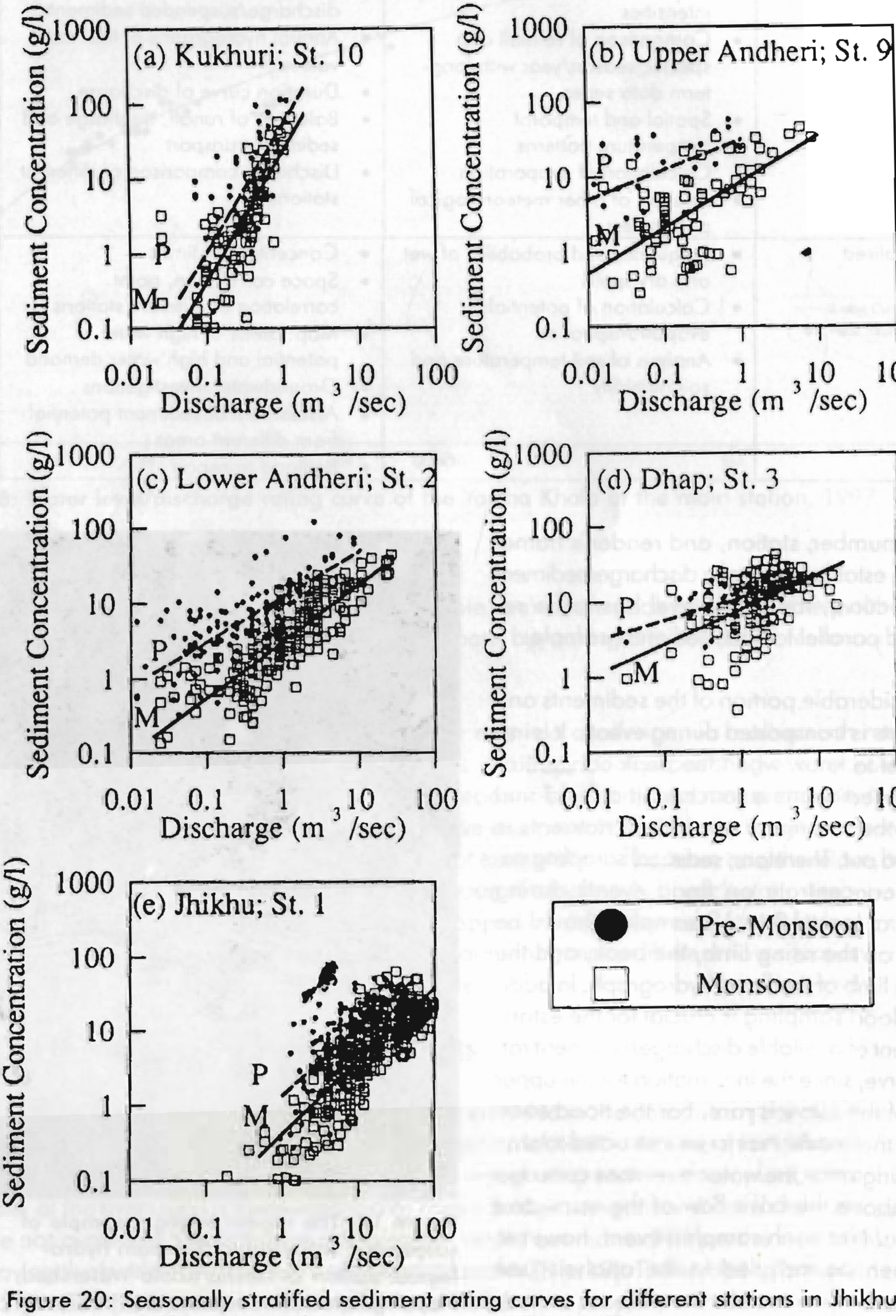


Figure 20: Seasonally stratified sediment rating curves for different stations in Jhikhu *Khola* watershed over several years (source: CARVER, 1997)

In special campaigns and in the different seasons, cross-section investigations should be carried out by sampling the whole transect of the river in order to procure information about the distribution of the sediments across the river.

3.1.5 Water Chemistry

The analysis of water chemistry provides information on the loss of nutrients as well as on water quality, both crucial elements for the monitoring of nutrient loss and balance. As the laboratory analysis is time consuming, the number of samples per station has to be reduced to about four to six per year: one sample during minimum flow, one during pre-monsoon period, and three to four samples during monsoon season.

3.1.6 Meteorological Parameters

For the project, rainfall (amount and intensity) and air temperature are the most important parameters to be monitored at all the measuring sites. For temperature and rainfall intensity, automatic recording devices (tipping buckets for rainfall, and thermistors or thermographs for temperature) are required in order to procure a high temporal resolution of the information. Ideally, devices for the measurement of air humidity and soil temperature are installed at the meteorological sub-stations as well.

Other meteorological parameters (soil humidity, evaporation, wind speed, wind direction, and radiation) are measured at the main meteorological station. As these data should be more or less representative for the watershed, the site selection for the main meteorological station is crucial (Fig. 21).

3.1.7 Erosion Plots

Surface runoff and soil loss are recorded at the erosion plots. In most of the watersheds, four drums are installed to catch the water and sediments with an overflow device from one drum to the next (Fig. 22). As the volume of runoff from an erosion plot is calculated based on the water height in the drums, the latter have first to be calibrated by systematically adding a known volume of water and measuring the depth of water.

Figure 21: The fully automated meteorological station in Yarsha Khola Watershed, Nepal (photo: October 1997)

The recording and the sampling, carried out by the reader, have to be taken after each event. The data recording begins with the measurement of the runoff height in the collection drums. If the depth of the runoff water in the drums equals or exceeds five cm, sediment sampling is carried out for each event. For sampling, water in the drum is first agitated to mix the fine and coarse sediments. A composite sample of a half a litre is taken from each drum. During heavy storms and when

Figure 22: The gutter and the collection drums of an erosion plot in Namdu, Yarsha Khola Watershed, Nepal (photo: May 1998)

there are large sediment depositions in the drums, two samples are taken, one from the upper part of the drum (suspended sediment) and one from the lower part (deposited sediments). Each sample has to be tagged with a reference number indicating the site, drum and sample number, date, and reader's name. Further details on the collection and analysis of data from erosion plots are specified by the Department of Soil Conservation and Watershed Management (1998).

3.1.8 The Readers

The work of the readers is very important.

- The readers are the key persons for the collection of raw data on hydrology, meteorology, and erosion.
- The readers are the watch-persons of the stations and their installations.
- The readers have always to be on the spot for the regular readings, for intensive monitoring during flood events, and to assist during discharge measurements. They have to carry out these duties during day and night.
- The readers are expected to be punctual and precise in carrying out the readings (Fig. 23).

In view of such responsibilities, the readers need incentives. The most important, of course, is a decent salary which motivates them to do a good job. Furthermore, basic equipment such as a raincoat, torchlight, umbrella, watch, and rucksack (for the hydro and erosion plot readers) have to be provided. Finally, the readers need forms on which they can enter their readings and observations. Examples of such forms are given in Annex 1.

In order to fulfill their duties, the readers need regular training. This training has to include, for example, how to carry out accurate readings, methods of sediment sampling, and the level of assistance expected during discharge measurements. In addition to this technical

Figure 23: The reader of the meteo substation in Gairimudi, Yarsha Khola Watershed, Nepal (photo: April 1998)

training, the readers need to be informed about the project, the reasons behind the collection of information, and the analyses that are carried out with 'their' data. They need to have an opportunity to exchange ideas, express their views, and discuss their experiences. If such training is provided regularly, the readers will realise that they are part of a whole network and that their contribution is crucial. If they understand the project goals and feel they are a part of the project team, then the readers will be motivated to collect good quality data.

The experience in the Nepal watersheds has shown that such one-day training sessions should be carried out once or twice a year (e.g., before and after the monsoon). All the readers are invited to come to the field station. In the first part of the morning session, the project objectives are highlighted and the state of the project is discussed; readers then ask questions and express ideas or complaints. In the second part of the morning session, results based on the data collected by the readers are illustrated and discussed. The readers are asked to comment critically on these results and to contribute from their long experience information about the physical processes in the area. The afternoon session takes place in the field, directly at the measuring sites (Fig. 24): the measuring and reading techniques are demonstrated and repeated, difficulties or key issues arising from the measurements are emphasised, the measurement programme is highlighted, and the uncertainties of the readers are discussed.

Figure 24: Training of the readers in Jhikhu Khola Watershed, Nepal (photo: May, 1988)

Such training courses have proved to be extremely useful and successful and are an important contribution to the achievement of common methodologies and high quality data.

3.2 *Data Handling in the Field Station*

3.2.1 Data Collection

It has to be ensured that all the hydrometeorological data recorded at the different monitoring sites reach the project field station safely, completely, and on time. Particularly important is the downloading of digital data, either by changing the storage module or by downloading the data to a portable computer. Equally important is the change of the charts of recording instruments (e.g., floaters).

Each digital instrument (e.g., tipping bucket, thermistor) has a different storage capacity and, accordingly, the necessary interval for downloading varies for the different recording devices. In addition, the floaters are equipped with either monthly or weekly charts, depending on the season. Therefore, for each watershed, a strict programme for data collection has to be defined which ensures timely downloading of the data or changing of charts. Ideally, this job is carried out in special campaigns during which all the data from automatic recording devices are collected--even when the storage capacity of certain instruments is greater.

The readers are responsible for organizing the flow of the manually-collected data as well as of the sediment samples from the measuring site to the field office. The sample bottles have to be promptly delivered to the field station in order to ensure timely processing of the samples in the laboratory and the rapid turn-around availability of sufficient empty bottles, particularly during the monsoon season. The labelling of the sample bottles reaching the field laboratory has to be checked thoroughly. The record forms for each specific month should be brought to the field office at the beginning of the subsequent month when the reader receives his salary. The forms have to be checked for clarity, accuracy, and completeness.

3.2.2 Data Treatment

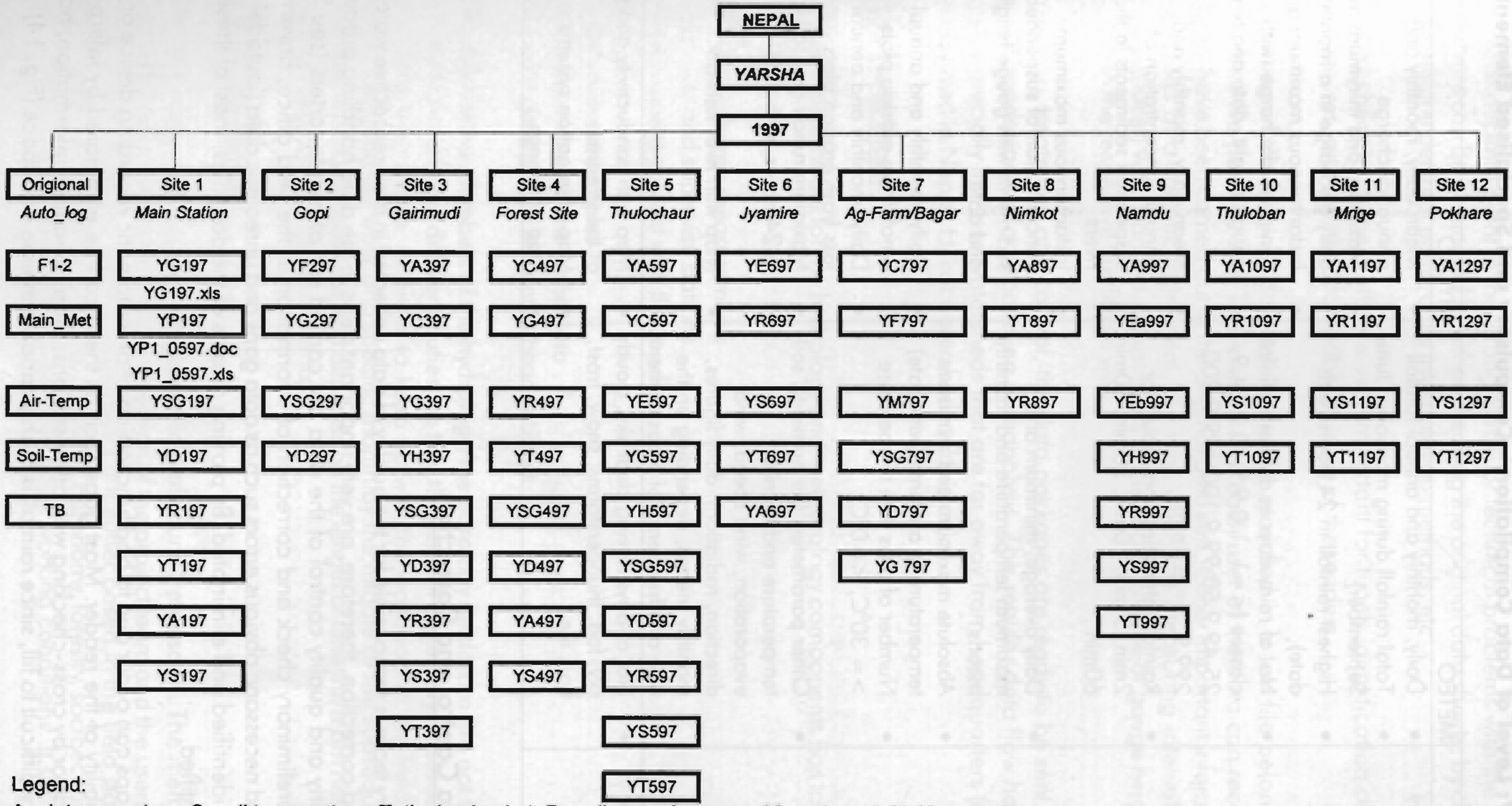
Once the data, forms, charts, and samples have reached the field office, they have to be processed. The necessary steps depend on the type of information, as follows.

- In most cases the digital data are already in the computer as a result of the downloading process. The data files have to be prepared for data storage (Fig. 25), with appropriate file name, file head, and data format.
- The information recorded on the charts has to be transformed into data files. If facilities for digitising charts are available at the project headquarters, then this transformation is not carried out at the field office. If digitising facilities are not available, the information required (see Table 7) has to be extracted manually from the charts by reading the values at regular intervals, and then it has to be entered into the computer. This step can be carried out in the field office.
- Entering the manual data into the computer is the most important, most time consuming, and most challenging work in the field office; it needs a lot of concentration and is tiring. In order to avoid errors in data entry, the data files have to be cross-checked by other members of the team.

The initial processing of the sediment samples (from both the hydrological stations and the erosion plots) is carried out in the field laboratory. The samples are filtered, dried and weighed, and the results are entered into the laboratory book. Chemical analysis (phosphorus content compulsory, calcium and carbon content desirable) of a selection of sediment samples (e.g., samples during particular storm events, single samples during base flow) has to be carried out in a professional laboratory. Ideally, hydrological and erosion plot samples from the same storm events should be selected for chemical analysis.

Based on Table 5, water chemistry is analysed only in special campaigns. Treatment of the water samples has to be carried out on the day of sampling. For these campaigns, an analysing kit has to be available in the field laboratory. The analysis of water chemistry might include nitrate, ammonia, phosphate, calcium, magnesium, conductivity, pH, and hardness; at present, water chemistry is in the test phase in the watersheds of Nepal.

Figure 25: The file structure developed for Yarsha Khola Watershed, Nepal



Legend:
A=air temperature, S=soil temperature, T=tipping bucket, R=ordinary rain gauge, M=main met, H=Hygrometer,
C=crest gauge, F=floater, G=staff gauge, P=pressure transducer, SG=sediment sampling,
E=Erosion plot

Table 7: Level of Data Compilation - Compulsory and Specialised Elements

	METEO	HYDRO
Compulsory	• Daily, monthly and annual rainfall • Total rainfall during monsoon (June-September) • Highest rainfall in 24 hours (with date) • No. of rainy days in different rainfall classes (<=1, 1.0-9.9,10.0-24.9, 25-49.9,50-99.9,100-199.9,200-299.9,>=300mm) • Rainfall intensities (mm/hour) for 2min, 5 min, 10min, 20min, 30min, 60min • Daily average, minimum and maximum temperature and monthly means • Absolute maximum and minimum temperature in a month (with date) • Number of days with temperature >= 30°C, <=0°C • Other parameters (air humidity, soil temperature and humidity, evaporation, wind speed, wind direction, radiation): daily figures, monthly means, depending on the type and frequency of measurements	• Mean daily, monthly and annual discharge • Maximum and minimum mean daily discharge in a month • Instantaneous maximum and minimum discharge (with gauge height, date and time) • Mean daily, monthly and annual concentration of suspended sediments in the rivers • Instantaneous maximum concentration of suspended sediments with gauge height and date • Daily, monthly and annual runoff from erosion plots • Daily, monthly and annual soil loss from erosion plots • Maximum runoff and soil loss in 24 hours, with date • Map with river network
Specialised	• No. of days having clear sky, cloudy sky, fog, thunderstorms, snow, frost, hail, ice	• Map with conductivity and pH of the streams • Map of irrigation systems • Map of springs

3.2.3 Data Control and Correction

Each country team is responsible for the quality of data collected in their respective watersheds. Control and correction, therefore, are very important steps in the data handling process. The main scrutiny and quality control of the data are carried out at headquarters (see Section 3.3.2). A preliminary check and correction of information in the field office, however, are possible and necessary: obvious errors such as data gaps and unrealistic data ('outside' values) have to be identified and eliminated. By plotting graphs of the data files, most of these errors can be identified.

- Data gaps can occur due to bad functioning of an automatic recording device or due to unreliability of the reader. Most gaps such as these can be eliminated by interpolation methods or by cross-checking with other stations. Gaps in rainfall information, however, are very difficult to fill, since rainfall is highly variable in time and space (Fig. 14).

- Unrealistic data can occur due to mistakes in data entry, mistakes by the reader (e.g., a comma or decimal point in the wrong place) or technical problems with the instrument. Such errors can be eliminated by either checking the original data sheets, by changing the location of the decimal point, or by interpolation methods, etc.

This first data check in the field office is very important for timely identification of problems, instrument failures, and the need for further training of the readers.

Each correction in the original data set has to be recorded in a text file belonging to the respective data file (see Fig. 25). This ensures that each user of the data can reconstruct the changes that have been carried out in the original data sets. Other important information, such as observations at a specific instrument, the time when the measuring interval (e.g., of a pressure transducer) was changed, adjustments of the floater (in time or gauge height), have to be recorded in the respective text file as well.

After collection, treatment, and control, the data packages are ready to be sent to the data manager at headquarters for processing and analysis. This regular data flow has to be well organized, especially in those watersheds that are far away from headquarters (e.g., Yarsha *Khola*, watershed in Nepal, Hilkot watershed in Pakistan).

3.2.4 The Field Team

The field team responsible for the hydrology and meteorology components has to fulfill many important tasks.

- Carry out discharge measurements
- Collect data and supervise stations
- Be in permanent contact with the readers
- Enter data
- Process sediment samples
- Make the first corrections to the data
- Prepare the data sets to be sent to headquarters

It is important that the members of the hydrology/meteorology field team are not too specialised, but well trained in the different duties, so that they can, if necessary, replace each other and also rotate their duties. In order to keep up with the workload in the field office, it is desirable to complete the work of a specific month by the middle of the subsequent month. However, this is not always possible and depends on the length of time the team has to spend in the field, e.g., for discharge measurements.

3.2.5 The Field Station

The field station (Fig. 26) is a busy place, particularly during monsoon. The members of the hydrology/meteorology team are only one section of the project team and the users of the field station. A number of other project collaborators are either based in the field station or visit frequently during specific campaigns (e.g., geology, agronomy, forestry, socioeconomics). The field station has to be a decent living place with good working conditions. It should include sufficient accommodation facilities, a living room, a good kitchen, an office (with sufficient working space, a laptop computer for downloading, at least one desktop computer, and sufficient storage facilities for files and stationery), a laboratory (with facilities for filtering, weighing, and drying samples), and decent toilet and washing facilities. It is important that a good cook is part of the permanent field staff, so that the technicians can concentrate on their technical duties.

Figure 26: The field station in Yarsha *Khola* Watershed, Nepal (photo: November 1997)

3.3 *Data Handling at the Headquarters*

3.3.1 Data Management and Storage

Transparent data management is essential for several reasons.

- In each watershed a huge amount of hydrological and meteorological data is recorde and collected.
- The information is to be used by many institutions and individuals for analysis.
- The hydrometeorological data sets will be used for synthesis with other components of th PARDYP project.
- Data from one watershed will be compared with data from other watersheds.

The Data Manager

Each country team has to identify one person to be in overall charge of data management an storage. This person should have considerable experience in computer technology, preferably background in physical science, and a good overview of the project. All the data should b stored centrally on the data manager's computer. The data manager:

- has the overview of all the data,
- is in permanent contact with the team at the field office and keeps a record of the data th have arrived from the field,
- is in contact with those individuals or institutions responsible for data processing,
- has to make sure that no data are given out before they are officially cleared and th several versions of the same file are not circulating at any one time, and
- is the 'custodian' of the data: he decides to which users the data are provided and h keeps in touch with them.

The Software

For the time being, it has been decided to use Excel as the common software for data management and storage. Excel is easily available and used widely, and thus ensures access to the information for a variety of users. However, it may create problems for data processing (see Section 3.3.2) since Excel is not a specific hydrological/meteorological software package. At present, there seems to be no alternative to storing the data in Excel, but the processing may be carried out with other, more specialised software packages. In future, it is envisaged that the procedure for data storage and processing for all the watersheds will be standardised by introducing a specialised hydromet programme, but the costs and the training needs have to be evaluated first.

File Structure

Figure 25 provides an example of the data file structure developed for Yarsha *Khola* watershed, Nepal. The files from the erosion plots are not included. This structure is based on the following principles.

- The files are structured according to measuring sites and measuring parameters. Experience has shown that it is easier to handle the file structure if separate files are created for each parameter at one particular measuring site.
- The files in the directory 'original' include the raw data from automatic recording devices; the data files in Excel (extension: xls) include the finalised data (after data processing, see Section 3.3.2); in the text files (extension: doc) interpolated data, station specifications, control measurements, problems observed, and so on are recorded.
- Monthly files are basically created for the parameters measured with automatic recording devices and annual files for manual readings.
- The creation of file names must receive particular attention. File names should include as much information as possible in order to achieve maximum transparency of the data base. Part of the annual file, YG197.xls (Y: Yarsha *Khola*; G: staff gauge; 1: site #1; 97: year 1997), is given in Table 8.

Data Format

The data format provides guidelines for the internal file structure. As for the file structure, a clear and uniform data format is important for the user. Table 8 provides an example of the file for manual water level readings at the main station in Yarsha *Khola*, Nepal. This data structure is based on the following principles.

- The data file includes a text block with the relevant information on the specific station and measuring infrastructure and a data block.
- The data are structured into columns: date, readings in different columns (8 a.m. reading, 4 p.m. reading), and remarks.
- The identical structure of each file of a specific data set and station is a condition for the application of macros in the analysis of the data.

This structure may have to be modified according to the needs of the data processing procedures in any specialised software package that might be acquired for the project in years to come.

3.3.2 Data Processing

Before the analysis can commence, the data have to undergo a strict process of scrutiny, checking, and correction. Data checking in the field office (see Section 3.2.3) is only the first step. Data processing is a very responsible job. It has to be carried out by a specialist who is well aware of the methods of analysis and of the statistical procedures involved. In Nepal, data processing is delegated to the Department of Hydrology and Meteorology of His Majesty's Government of Nepal. The following text specifies the data processing procedure and is extracted from a hand out by Mr. Sunil Kansakar (Department of Hydrology and Meteorology) provided to trainees at the PARDYP Training Workshop in November 1997.

In practice, there is always a difference between the measured and the true values. This difference is known as error. The errors are classified into three groups:

- incompleteness of data,
- errors in observation, and
- administrative errors.

Incompleteness of Data

Incompleteness of data is caused by failure of instruments or by discontinuity in observers' recordings. The missing data can be estimated by:

Table 8: Data Format for the Manual Water Level Reading

Staff Gauge

Instrument:	Ordinary Staff Guage
Site #:	1 (Yarsha *Khola* Watershed)
Station:	Main Hydro Station
Year:	1997
Interval:	24 Hour
File Name:	YG197

File Name: YG197

Day	8:00	16:00	Remark
1-Jun	42	40	
2-Jun	41	40	
3-Jun	40	44	
4-Jun	42	38	
5-Jun	40	41	
6-Jun	43	41	
7-Jun	42	41	
8-Jun	44	42	
9-Jun	43	42	
10-Jun	45	43	
11-Jun	40	39	
12-Jun	40	38	
13-Jun	39	38	
14-Jun	40	38	
15-Jun	40	38	
16-Jun	39	38	
17-Jun	39	40	
18-Jun	44	42	
19-Jun	38	38	
20-Jun	41	39	
21-Jun	38	36	
22-Jun	45	39	
23-Jun	48	40	
24-Jun	50	43	
25-Jun	48	52	
26-Jun	50	52	
27-Jun	48	44	
28-Jun	50	46	
29-Jun	46	42	
30-Jun	52	64	
1-Jul	56	50	
2-Jul	50	50	
3-Jul	52	50	
4-Jul	56	52	
5-Jul	80	68	
6-Jul	64	58	
7-Jul	95	68	
8-Jul	78	74	

- interpolation (linear in the case of slowly changing processes, polynomial in the case of rapidly changing processes),
- comparison with other observations at the same or from nearby stations (double mass curve, regression), and
- rainfall-runoff models to fill in the missing runoff data.

Errors in Observation

- Random errors (accidental errors) occur by chance, both above and below the true value. They do not greatly affect the mean value and hence the effect can be reduced by increasing the number of observations or by improving the accuracy of individual measurements.
- Systematic errors cannot be eliminated by increasing the number of readings or observations. They can be caused by the instrument and also by incorrect rating curves. Systematic errors can be detected by calibration. Systematic errors affect mean results as well as the extremes.
- Gross errors occur occasionally, resulting from random instrument failure, from human error during measurement, from transfer of data or from misinterpretation of data. If the error is small, it can hardly be separated from random errors. If it is large, the extreme values will be incorrect and, in severe cases, also the mean. Gross errors depend largely on the measurement and registration devices, the maintenance and the method of processing (manually or with computer). These errors may be detected in different ways: visual checking (plot), comparison of values with a given physical range, comparison of the consecutive data differences with a given range (delta check), and comparison of data with those from other stations (plotting hydrographs, double mass analysis).

The total error in a measurement consists of random error, systematic error, and gross error. If the systematic and gross errors are removed, there will still be a residual error caused by random errors. Random errors cannot be corrected but must be kept to the minimum.

3.4 Administrative Errors

These errors are usually made by the observer and also at the office where the data are being processed. These errors can easily be detected and corrected. Frequent visits to the station(s) by technical personnel can minimise the errors caused by observers. Errors made at the office can easily be detected through the double entry system in the computer and by comparing the files.

During data processing, the information is not only checked, but is also prepared for the first step in the analysis procedure, and the publication in the Yearbooks (see Section 4.3). In reality, the data control and the first level of analysis (Section 4.2) cannot be separated. However, differentiation is made here in this discussion paper in order to keep the sequence of the considerations clear and transparent. Ideally, the data processing is carried out with special software packages. In Nepal, the HYMOS programme is used at present by DHM; as already mentioned, this package may be introduced in the other watersheds in future. Such software packages usually include facilities to carry out different tests for data control, to identify outside values, to produce graphs, to calculate statistical values or to digitise floater charts.

After completion of the data processing, the information has to go back to the data manager. As long as the data are stored in Excel-files, all these files have to be updated with the changes necessary for data processing. If the data management and storage are standardised in HYMOS (or another package), then the data manager needs to get a copy of the actualised version of the data base.

The data are now ready for analysis and for publication. Analysis is the topic of Section 4.