CASE STUDY

CS (AR) 174

PREDICTION OF WATER QUALITY PARAMETERS USING STATISTICAL APPROACH IN UPPER NARMADA BASIN

NATIONAL INSTITUTE OF HYDROLOGY
JALVIGYAN BHAWAN
ROORKEE - 247 667 (INDIA)

1994-95

PREFACE

Rivers and streams support a multitude of human uses, including water supply for municipal, industrial, and agricultural uses, recreational opportunities, and a means of waste disposal. The quality of stream water is directly affected by the amount of waste discharged into the stream, and this quality governs the extent to which the water can be used for other purposes, such as recreation or water supply.

The advent of basin-wide planning for water pollution control has conceived the need for determination of the status of river basin with respect to pollutant loads. This need has been culminated in the widespread use of models of river basins for planning purposes. The use of river basin models has a requisite a large data base on which verification of the model can be made and modelling constants determined. The availability of such a data base can also permit a statistical evaluation of the data which would provide further insights in to the status of the river basin.

Since the data from river basins usually involve a number of variables, interpretation through standard statistical techniques is not satisfactory and usually requires a multivariate approach. Two problems which arise in the use of any statistical analysis are missing data values and satisfaction of the various assumptions concerning the data which allow the use of a specific statistical technique. In the present study multiple regression models for some of the water quality determinands have been developed which may be useful for filling the missing data values. Besides the regression modelling, quality-quantity relationships for various water quality parameters are also developed. To have an idea about the probability of violation of prescribed limits of water quality standards, probability distributions are also plotted for various water quality parameters. About 12-years (1980-1992) data collected from different agencies e.g. Central Water Commission, New Delhi, Central pollution Control Board, New Delhi, M.P. Pollution Control Board, Bhopal, is used in this study.

The study entitled "Prediction of water quality parameters using statistical approach" has been carried out by shri Aditya Tyagi, scientist 'B' and Dr. K.K.S. Bhatia, scientist 'F' under the 1994-95 work programme of the Environmental Hydrology Division, NIH.

(S.M. Seth)
DIRECTOR

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

As water moves through the drainage basin it come across through different hydrological processes e.g. storage, and transfer pathways, its quality is modified, and the quality of water leaving the drainage basin will be considerably different from that of the moisture entering the catchment as precipitation. The various processes controlling stream water quality may be in a delicate balance and that a slight modification to the catchment, such as a change in landuse, could generate significant changes in water quality.

The present study is an attempt to provide a hydrological overview for better understanding of the quality dimension of the hydrological processes operating within the catchment ecosystem. About 12-years water quality data observed at the Manot site of the upper Narmada River basin is used to carry out the statistical analysis. The analysis has been carried out in three phases. In the first phase quality-quantity relationships have been developed relating the concentration of quality parameters to the magnitude of stream flow for various water quality determinands. In the second phase probability distributions have been plotted which gives the information about the probability of violation of allowable limits of various water quality parameters. In the third phase multiple regression models have been developed for various water quality parameters which may be used for filling the missing data values.

1.0 INTRODUCTION

Water is essential for human survival; the terrestrial eco-system can not function without it. In addition to drinking and personal hygiene, water is needed for agricultural production, industrial and manufacturing processes, hydroelectric power generation, waste assimilation, recreation, navigation, enhancement of fish and wildlife, and a variety of other purposes. When a resource is used for so many diverse purposes it is important that it be developed and used rationally and efficiently.

While the total amount of water available in the world is constant, its distribution over different regions is uneven and this contributes to tremendous variation in terms of local water availability. In addition to the geographical variability of water its availability in all regions varies with time, depending on hydro-meteorological conditions. Much emphasis has been placed in recent decades on assessing and analyzing water availability to satisfy human needs and also simultaneously to reduce the harmful effects of droughts and floods. This is understandable since with rising global population and increase in our standard of living, the total water requirements have continued to grow.

It is important, however, to realize that availability of water does not only mean the quantity of water available for different purposes but also with its suitable quality. It is quite possible that adequate quantities of water may be available for a specific purpose, but the quality of water may be quite inappropriate for that use because it is unacceptable. Furthermore, a certain quality of water that can be used for irrigation may not be suitable for industrial purposes and vice versa. Quality and quantity are closely interrelated, and both must be considered simultaneously for all water resources management practices. Quantity of water without any reference to its quality can easily be a meaningless term for overall planning and management purposes.

During the last two decades it has become increasingly evident that pollution is impairing different uses of water in many parts of the world, and in certain cases even destroying the possible utilization of important sources of supply With accelerated and uncontrolled development, more and more waste products are being discharged to water courses. In many cases not enough is known with any degree of reliability about the impacts of these pollutants on human health. Extensive monitoring and research programs have to be initiated before safe levels of specific pollutants can be determined, either individually or synergistically with other pollutants.

On the global basis there is no doubt that much better data exist on hydrometeorological variables associated with water quantity as compared to information available on different water quality elements. While reasonably accurate estimates of world water balance are known, information on its quality leaves much to be desired. Thus with very few exceptions, even approximate continental or global assessments of different water quality parameters are not available. Nor is much known about the magnitude and type of municipal, industrial, and agricultural wastes entering water bodies which could constitute a growing hazard to human health and pose serious adverse environmental problems. Since it is comparatively easier to control point sources of pollution, much of the emphasis, in recent years has been placed on this aspect. But this clearly is not enough. For example, even for a major industrialized country like the United States, which has made a determined attempt to control water pollution, its National Commission on Water Quality reported in 1976 that 92 percent of suspended solids, 37 percent of biochemical oxygen demand, and 98 percent of coliform bacteria will still remain uncontrolled in natural surface water, even when all discharges from point sources have been eliminated, which by itself is not an easy task. This is largely due to the agricultural activities. Currently there is no general measurement of volumes of synthetic organic compounds and heavy metals reaching water courses throughout the world, some of which eventually entering the oceans.

Water plays an important part as a medium through which toxic chemicals and heavy metals are dispersed to ecosystems other than the ones intended, both through drainage water and by evaporation and subsequent precipitation . The problem of acid rain, which is a great concern to the several countries having large costal areas, is also primarily due to the evaporation-precipitation processes. Such dispersal mechanisms, including the dynamics of water transport mean that toxic material can be detected in locations far away from the points of application. Thus significant quantities of pesticides, including DDT and its derivatives, have been found in animals in Antarctica, like penguins, and their eggs, skua, and fish, even though such pesticides were neither necessary nor need in that area. Water quality management will also have other direct impacts on human health. Currently much of the population of developing countries does not have access to clean water. According to a survey carried out in 1976 by the World Health Organization of the United Nations, in which 67 developing countries participated, only 20% of the rural population and 75% of the urban population is provided with clean water in most developing countries, whereas, the combined figure for rural and urban sectors is only 35%. The United Nations Conference on Human Settlements, held at Vancouver, Canada in June 1976, passed a major resolution using provision of clean water to every one by 1990. This target was subsequently endorsed by the United Nations Water Conference, held at Mardel Plata, in March 1977. The importance of clean water was further demonstrated by the United Nations when the decade 1980-1990 was proclamed as the International Water Supply and Sanitation Decade.

Improvement in the quality of water available for human consumption will contribute to the reduction of health hazards due to diseases like Cholera, typhoid, infections hepatitis, and bacillary dysentery. It would further reduce human contracts with vectors of water borne diseases like schistosomiosis, trypanosomiasis and guinea worm. Current water and health situation on a global basis has been estimated as given in table 1 (Biswas, 1974).

Table - 1

Sl.No.	Diseases	Number of cases every year
1.	Gastro-enteritis	400 million
2.	Schistosomiasis	200 million
3.	Filariasis	200 million
4.	Malaria	160 million
5.	Onchocerciasis	20-40 million

These statistics clearly indicate the need for better water Quality Management. While this need is easy to demonstrate beyond any reasonable doubt, the real question then becomes how to carry out such management processes effectively.

The projections given by different agencies and scientists are acarving but undoubtedly the situation has improved in last two decades. However, to make a dent in this problem, it is significant that newer techniques of water quality forecasting, projections and management are resorted to. This can be achieved by using the modern tools like water quality modelling, computer simulation and conventional practices like public participation and awareness.

The inputs of pollutant constituents to river usually come from natural processes and wastewater discharges of municipal, industrial or agricultural activities in the form of both the point and nonpoint sources runoff. The physical and biochemical characteristics of the point waste sources are much more understood and measurable, than those of the nonpoint sources. The complex, often random, and time variable nature of both point and nonpoint sources is often disregarded because of insufficient knowledge and modelling complexity, however, consequently this resulted in reduction of model output reliability.

Water resources planning is not a new process. In traditional approach to water pollution control that still persists in local planning, the main objective is to clean up raw waste water and comply with the effluent standards imposed by a responsible regulation agency. Due to high cost involved in pollution control and often nonmonetary benefits associated with improved water quality, the planning process is more a series of conflicts between diversified groups rather than a smooth continuous process. The priorities in such a process may frequently be related to the visibility of the pollution input and its consequences. Non visible and often unrecorded nonpoint pollution has thus an obvious handicap when pollution control priorities are considered. Nevertheless, it has been realized that clean up limited to point sources only may not lead to the desired water quality goals. Thus, identification, control, and management of pollution sources particularly non point pollution must be included in the area of water pollution abatement programs.

A knowledge of the various processes involved in the movement of water through the catchment ecosystem and in the control of water quality is essential in evaluating the impact of pollutants on river quality and recognizing potential hazards. It is the general field of non-point or diffuse source pollution that is of great importance in this context because, whereas point sources generally discharge directly into channels and their effects may be readily discerned in terms of increased stream loadings, non-point pollutants may reach the channel by a variety of routes and this movement may not be fully understood. Thus the development of a comprehensive understanding of the constituents of river water specifically storm water runoff, and more importantly, how these constituents vary both temporally and spatially, is an important step in addressing the problem of nonpoint source pollution in a river basin.

1.1 HYDROLOGICAL PROCESSES AND THE QUALITY DIMENSION

Fig. 1 presents a simplified view of the hydrological processes which operate in a catchment ecosystem. Inputs of

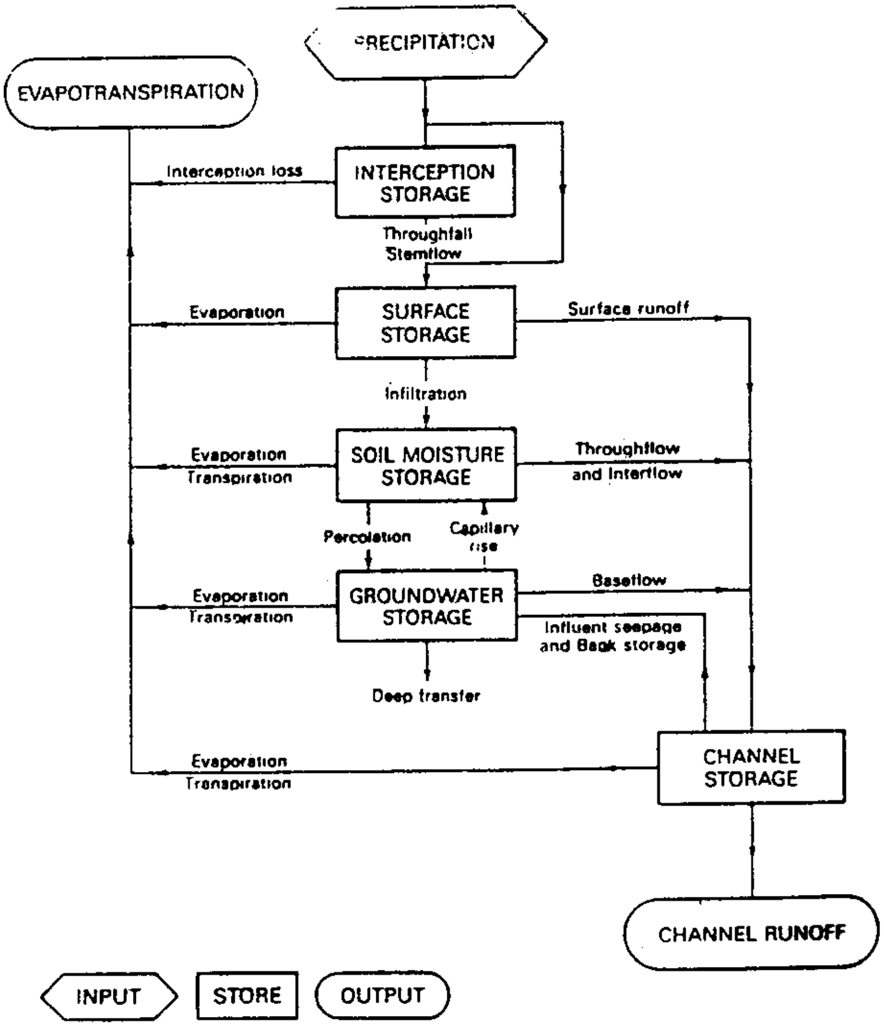


FIG.1 SIMPLIFIED VIEW OF HYDROLOGICAL PROCESSES

precipitation are distributed through a number of stores by a series of transfer processes and are output as channel flow, evapotranspiration, and deep leakage. Thus, in simple terms, water entering a catchment is intercepted by vegetation and subsequently evaporated, or reaches the ground surface as through-fall and stem-flow. In the absence of vegetation, the precipitation reaches the ground water directly, although interception by surface litter may still occur. At the soil surface,water infiltrates into the soil or is retained in surface storage. the surface water may move downslope as surface runoff or be slowly evaporated. Moisture held within the soil is subjected to surface evaporation and plant transpiration, to downward percolation to the water table, and to downslope movement as through-flow and inter-flow. Water held within the groundwater body is similarly subject to losses by plant transpiration and evaporation when the water table is close to the surface, to upward movement into the unsaturated soil by capillary action, to deep leakage to adjoining ground water bodies, and to slow surface outflow into springs, seeps, and river channels as baseflow. The river channel receives variable contributions of surface runoff, through-flow and inter-flow, and base flow which are routed downstream and together contribute to the time variant outflow hydrograph of the drainage basin.

As water moves through this system of stores and transfer pathways, its quality is modified, and the quality of the water leaving the drainage basin will be considerably different from that of the moisture entering the catchment as precipitation. Fig. 2 attempts to portray a number of the more important mechanisms that modify the water quality, and this representation can be viewed as significant in two respects. First it indicates that water quality will change as moisture moves through the catchment system and that water characterizing different stages of the

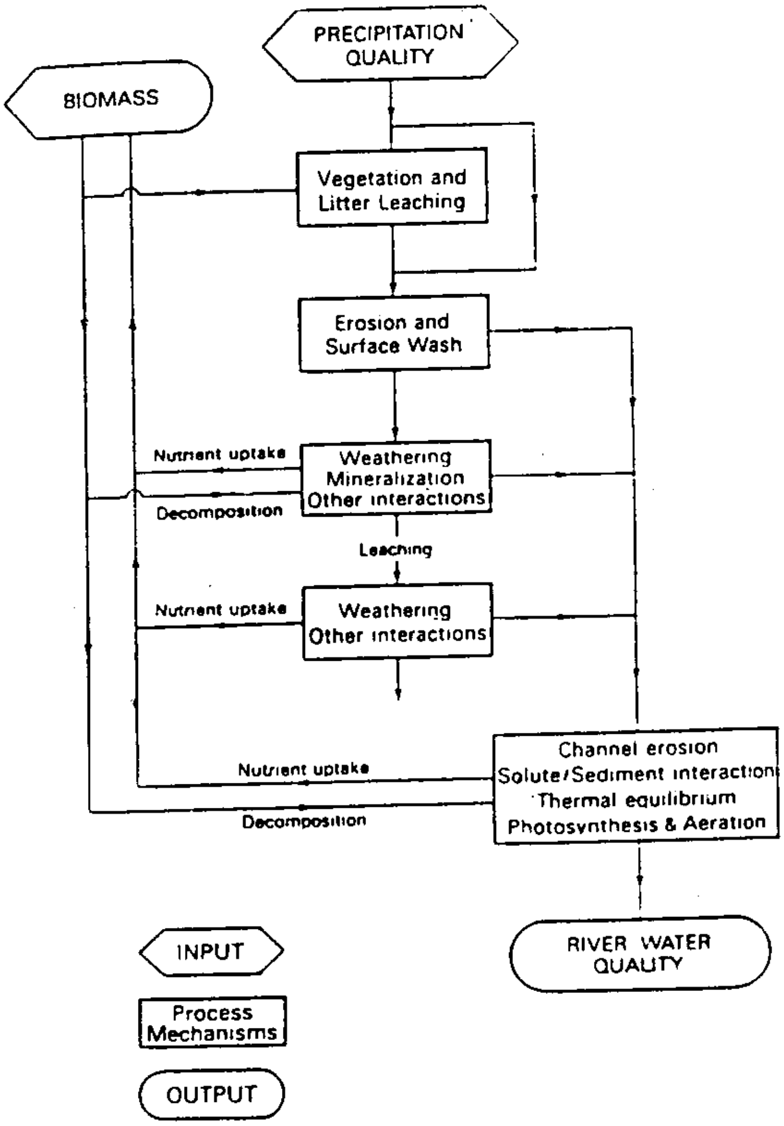


FIG.2 IMPORTANT HYDROLOGICAL MECHANISMS RESPONSIBLE FOR MODIFIYING THE WATER QUALITY

hydrological cycle will exhibit different properties. Secondly, it demonstrates that the quality of water flowing in a river or stream will have evolved through a complex series of interactions with the soil, rock, and biota of the catchment ecosystem, and that it will be necessary to consider these interactions if spatial and temporal variations in stream quality are to be fully understood. Furthermore, it must be appreciated that the various processes controlling stream water quality may be in a delicate balance and that a slight modification to the catchment, such as a change in land use, could generate significant changes in water quality.

2.0 POLLUTION SOURCES

Sources of pollution can be divided basically into two groups : natural and cultural (those caused by man). The sources can be further classified as either point or diffuse (non-point) sources of pollution.

Point sources enter the pollution transport routes at descrete, identifiable locations and usually can be measured directly. Major point sources include effluent from farm buildings or solid waste disposal sites.

Pollution from diffuse sources can be related to weathering of minerals, erosion of virgin lands and forest including residues of natural vegetation, or artificial or semi artificial sources. The last can be related directly to human activities such as fertilizer application or use of agricultural chemicals controlling weeds or insects, erosion of soil materials from agricultural, farming areas and animals feedlots, construction sites, transportation, cumulation of dust and litter on impervious urban surfaces, strip mining, and others.

Since pollution has been defined as the addition of substances to waters above their natural quality, pollution from non point sources is caused mostly by man and his activities and should be distinguished from water quality contributions, sometimes called "background pollution", caused by the contact of water with rocks, undisturbed soils and geological formations, natural erosion and elicitation of chemical and biochemical components from forest litter,migration of salt water into estuaries, or other natural sources.

Besides the sewage and industrial waste water discharges and erosion from agricultural lands or urban areas, many other activities can cause pollution. The curring down of a forest results in a loss of protective cover and exposes soils, thus causing increased surface runoff and erosion; stream channelization and removal of trees and protective grass blanks may have significant ecological effects, sometimes complete destruction of the original biological life in the stream.

2.1 RURAL NONPOINT SOURCES

These sources are mostly related to agricultural activities. Agricultural pollutants have their origin in fertilizer use and pesticide applications, and generally, the primary causes are agricultural methods of disturbing soils by tillage or logging. Several other factors also affect pollution loading, e.g. soil type, climate, management practices, and topography etc. Land uses that produce the most pollution per unit area are animal feed lot operations and farming on steep slopes. Forested lands and pastures, on the other hand, produce the least amount pollution, that is approaching background levels. The impact of pollution on receiving waters depends on the distance of the source from the nearest concentrated flow stream and on the processes taking place during the overland flow phase of the pollutants transport.

2.2 URBAN NONPOINT SOURCES

Urbanization and related hydrologic modifications may cause increased pollution loadings that are significantly above the original or background levels. The sources of urban non point pollutants varies considerably, ranging from urban bird and pet populations, street litter accumulation, tire wear of vehicles, abrasion of road surfaces by traffic, street salting practices, and construction activities. Urban non point pollution may contain many dangerous contaminants such as lead, zinc, asbestos, PCBs (Poly chlorinated biphenyls) oil, and greease.

Evaluation of the effects of nonpoint pollution on surface water quality may differ from the traditional approach designed mainly for point sources. The difference can be summarised in the following table (Table 2).

Table 2: Difference between Point and Nonpoint Sources

S.No	Point Sources	Non Point Sources
1.	Fairly steady flow and quality Variability ranges less than one order of magnitude	Highly dynamic in random intermittent intervals. Variability ranges often more than several orders of magnitude.
2.	The most sever impact during a low flow summer period	The most severe impact during or following a storm event
3.	Enter receiving water at identifiable points.	Point of entry often can not be identified or defined.
4.	Primary parameters of interest: BOD dissolved, oxygen, nutrients suspended solids.	Sediment, nutrients, toxic substances, pH (acidity) dissolved oxygen

A typical non point pollution load to surface waters is a response of the drainage area to a storm event. It has a limited duration, lasting from a fraction of an hour to 2 days. However, some inputs from soil and groundwater zones last longer. The magnitude of the nonpoint pollution load depends on many uncontrollable factors, primarily the rain volume, intensity, and quality, duration of the preceding dry period, and others.

3. SPECIFIC WATER QUALITY PROBLEMS

Water quality problems associated with nonpoint pollution can be divided into several categories:

i) Problems with sediment
ii) Dissolved oxygen and biodegradable organics
iii) Nutrient contribution and eutrophication
iv) Toxic chemicals and metals
v) pH and acidity

Other problems such as bacteriological and radiological nonpoint pollution impact on receiving waters are of lesser importance or are local. Most of the radiological nonpoint pollution originating from atmospheric testing of nuclear weapons is presently below dangerous levels due to the ban of the atmospheric testing by major nations with nuclear capability.

3.1 SEDIMENT PROBLEM

Sediment from nonpoint sources is the most widespread pollution of surface water. Sediment, especially its finer fractions, cuts down light penetration and, thus, greatly reduces algal production. The turbidity caused by the sediment also has deleterious effects on benthic biota and fish, and impairs most of the major beneficial uses of water.

Large amounts of sediment settle out annually in slow moving sections of surface water, particularly in reservoirs, harbors, estuaries, and impounded streams.

In stream, sediment either moves in suspension or is shifted on the bottom. The suspended portion is called wash load while bed load is the portion that moves at or near the bottom in an erratic movement along the stream bed. The bed load portion of the sediment transport, with its instability, eliminates suitable habitats for aquatic life.

The concentrations of suspended sediment in streams are highly variable and are influenced by several factors including rainfall duration and intensity, soil condition, topography, geology, vegetation cover, and disturbing activities taking place in the watershed. Sediment concentration may vary from a few milligrams to more than ten thousand milligrams per liter. Most freely flowing rivers demonstrate significant day to day variations in the concentration of suspended solids. The suspended load may vary with different reaches of a stream particularly when the bed is resuspended or deposited due to changes in the hydraulic regime.

The question of how much suspended sediment is deleterious to surface waters cannot be precisely delineated, and standards are only available for turbidity. However, detrimental effects of sediment loads on the ecosystems can be characterized, and the importance of each effect outlined.

Suspended sediment alters aquatic environments, primarily by inhibiting light, changing heat radiation, blanketing the stream bottom and retaining organic materials and other substances that create unfavorable conditions for benthic organisms.

Sedimentation and blanketing of the stream bottom concern both organisms living in the bottom of the receiving water as well as higher organisms that depend upon bottom organisms for food. It has been long recognized that sediment deposition is primarily harmful to fish eggs buried in the bottom. Heavy or irritating concentrations of solids might interfere sufficiently with the gill movement of fish to affect circulation. However, some studies indicate rather significant resistance of adult fish to turbidities. Regardless of the uncertainties as to the effect of sediment on fish populations, the truth remains that indirect damage to fish through destruction of the food supply, eggs, or alevins, or changes in the habitat probably occur long before any adult fish can be directly harmed.

Increasing turbidity reduces the light penetration which in turn limits the primary productivity by limiting the column of water in which light intensity is sufficient (about 1% of incident light) for the rate of photosynthesis to exceed the respiration rate., Since photosynthetic organisms form the base of the food web, any significant change in their population would have a widespread effect on the organisms depending upon them for food.

The surface of particulate matter may act as a subtracts for bacteria and other microorganisms thus changing their habitat. Sediment also traps nutrients, phosphorus in particular, and other toxic materials, which may be buried with the sediment in places of sediment deposition. Nutrient release and possible release of toxic materials may follow when resuspension of bottom materials occurs during periods of high flow or local scour.

Soil loss, which is the primary source of sediment, contains one or more percent of organic matter, nutrients in amounts of about a fraction of a percent, and often significant amounts of pesticides, toxic metals, and PCBs.

Deposition of organics, which have lower specific density than sand, silt, and clay particles, occur at lower velocities. Most of the organics stay as part of the washload.

As will be seen in the next section, one of the most serious water quality problems is deposits containing large amounts of organics. Their oxygen demand deprive water of oxygen. Furthermore, processes taking place in the sediment layers may also, under certain conditions, release large amounts of nutrients back into the water.

3.2 DISSOLVED OXYGEN PROBLEM

Dissolved oxygen concentration of surface water is the primary parameter on which suitability of water for fish and wildlife is determined. The D.O. concentration standard is also the major parameter used in the waste assimilative capacity determination. The dissolved oxygen problem caused by nonpoint discharge such as runoff from manure applications, animal feedlots, and urban storm or combined sewers, can be devastating to the receiving water. The problem is not limited to runoff discharge containing high amounts of organics - for example, feedlot runoff. Large amounts of relatively clean storm water may resuspend bottom deposits and increase their oxygen demand by several orders of magnitude.

In streams, the dissolved oxygen concentration is a response to various oxygen sinks and sources. The sinks of oxygen that is the biochemical and biological processes that use oxygen, include:

i. Deoxygenation of biodegradable organics whereby bacteria and fungi utilize oxygen as an electron accepter in the bio-oxidation process.

ii. Benthal Oxygen demand, where oxygen is utilized by the upper layers of the bottom sediment deposits.

iii. Nitrification, in which oxygen is utilized during the oxidation of ammonia and organic nitrogen to nitrates.

iv Respiration by algae and aquatic vascular plants which use oxygen during night hours to sustain their living process.

MAJOR OXYGEN SOURCES ARE:

i. Atmospheric reaction, where oxygen is transported from the air into the water by turbulence at the water surface.

ii. Photosynthesis, where chlorophyll containing organisms (algae and aquatic plants) convert CO_2 (or alkalinity of water) to organic matter with a consequent production of oxygen.

3.3 NUTRIENT PROBLEMS AND EUTROPHICATION

Although many elements and chemical compounds are essential for plant and algal growth, only nitrogen and phosphorus are considered the limiting nutrients controlling their growths. On the average, about 50% of phosphorus and an even greater proportion of nitrogen originates from nonpoint sources.

The studies undertaken by the National Eutrophication Survey in the early 1970s clearly indicated that nationwide there is a distinct correlation between general land use and nutrient concentrations in streams. Streams draining agricultural watersheds had on the average considerably higher nutrient concentrations than those draining forested watersheds. The nutrient concentrations were proportional to the percent of land in agricultural and urban land use. It should be pointed out that these conclusions were based on surveys was found between the nutrient concentrations and the geology or drainage density of the watersheds.

The nutrient problem is especially important for lakes, reservoirs, and estuaries. In these water bodies, the classical dissolved oxygen concept may not work in evaluating the waste assimilative capacity and the production of organic matter by phytoplankton and large plants - macrophytes / may greatly exceed the BOD contribution from point sources, runoff, and tributaries. Oxygen levels are also affected by photosynthesis, and respiration and BOD concentrations are affected by the planktonic organisms and their residues.

The eutrophication process takes place in a surface water body in which orgnaic mass production nourished by nutrients exceeds its loss by respiration, decay, and outflow. In today's context, "Eutrophication refers to the natural and artificial addition of nutrients to bodies of water and to the effect of these added nutrients on water quality".

Eutrophication is the process by which a water body progresses from its origin to its extinction according to the level of nutrient and organic matter accumulation. The rate of eutrophication depends on many environmental factors, many of them uncontrollable. It is a dynamic process with highly variable rates that differ from year to year, season to season, and even hour to hour.

3.4 TOXIC METAL PROBLEMS

Almost any element in the earth's crust can potentially be found at least in trace quantities. in surface water bodies. Most of these elements originate from natural sources. However, increased industrialization and urbanization have raised the levels of the trace metals in surface waters, especially the levels of toxic sometimes improperly called heavy, metals.

The periodic table includes over 90 elements from hydrogen to trans-uranians, and all but 20 can be classified as metals. As many as 59 of these elements can be considered. heavy metals.

However, only 17 of the heavy metals are considered both very toxic and relatively accessible. Of these 17 toxic metals (table 2-6) nine are being mobilized into the environment by man at rates greatly exceeding those of natural geological processes.

Most likely each of these nine metals occurs naturally in water to some extent, and all organisms are naturally exposed to some level. In addition, some metals, including Cd, Cu, and Zn have known biochemical roles and others may be required as trace elements by the cell.

Since any metal may have a positive biological role, it follows that the total absence of a metal may be determental, and some concentrations may be optimal, and that some concentrations will be toxic. The toxicity levels will depend on the type of the metal, its biological role, and on the type of the organism and its ability to regulate its body concentrations of the metal. Table 2-7 shows a general ranking of toxicity of the nine toxic metals to aquatic biota.

PLURAG, in studying the impact of pollution of the Crest Lakes, determined that the followings elements should be considered potentially hazardous and requiring further attention.

i. Mercury, lead
ii. Arsenic, cadmium, selenium
iii. Copper, zinc, chromium, vanadium

3.5 ORGANIC CHEMICAL PROBLEMS

The past few decades, especially the period of the 1970s have been an era of controversy and concern over benefits and environmental cost of the use of organic chemicals, namely pesticides and polychlorinated biphenyls (PCBs).

The monitoring of DDT and dieldrin in fish species in the Crest Lakes added greater evidence to the pesticide problem. High salmon mortalities were observed in Lake Michigan, where DDT in fish tissues averaged two to five times higher than those from Lake Superior and 60 times higher than those from the Pacific west coast.

Of most concern to those interested in water quality is the potential concentration increase of toxic materials that can occur via the food web. The process called biomagnification is of particular significance to the pesticide and PCB problem since large quantities of these toxicants can be found in higher trophic level organisms (end of the food web), while lower trophic organisms and water itself exhibit lower, often undetectable concentrations.

There are tremendous amounts of organic chemicals that can be potentially dangerous to the aquatic environment. By their very purpose, all pesticides are toxic and can cause damage. The amount and nature of pesticides reaching surface waters from agricultural lands are primarily functions of the persistence of the compound used, intensity and length of time pesticides have been applied, and transport mechanisms from the area of application to receiving waters.

Many organic chemicals including chlorinated hydrocarbons, are potentially carcinogenic. Some of the potentially dangerous chemicals such as persistent pesticides and PCBs have become a part of the environment and can travel by air, by water, and with the sediment. Many of them, for example DDT and most PCBs are highly insoluble, hydrophobic (water repellent) substances and travel most with air aerosols and aquatic particulate matter.

3.6 p^{H} AND ACIDITY

The pH, which expresses the molar concentration of hydrogen ion as its negative logarithm pH = -log (H) is one of the primary indicators used for evaluation of surface water quality and suitability for various beneficial uses. Most aquatic biota are sensitive to pH variations. Fish kills and reduction and change of other aquatic species result when the pH is altered out side their tolerance limits. Most of the aquatic species prefer a pH near neutral but can withstand a pH in the range of about 6 to 8.5.

The toxicity of other toxic components can also be altered if the pH is changed. The solubility of many metals as well as other compounds (Ammonia) is affected by the pH, resulting in increased toxicity in the lower pH range.

Change of pH and acidity of surface waters resulting from non point inputs can occur mainly from two sources of acidity: (1) acid mine drainage water, and (2) acid precipitation.

Both sources have similar origins. Acid mine drainage is a result of mine water being in contact with sulfur bearing minerals, while acidity of precipitation is caused, primarily by atmospheric sulfur. Oxidation of these sulfuric compounds in surface or atmospheric water produces sulfuric acid, which then dissociates to H and SO ions.

Acid rain, which is defined as rain with a pH less than 5.6 is a result of sulfuric acid nitrate emissions from urban, industrial and electric utility burning operations that use sulfur and nitrogen containing fuel (primarily coal or, to a lesser degree, low quality oil).

Many watershed and surface water systems have a natural ability to neutralize the excess acidity. During the overland flow, rain and snow melt water dissolves calcium and magnesium containing rocks (limestone and dolomite) or soils, and it is enriched by mineral and organic salts such as phosphates and humates. These constituents often provide enough buffering capacity to maintain the pH of surface waters within acceptable ranges. At this time, rain acidity does not seem to have a great adverse effect on large water bodies such as the Gread Lakes that have elevated hardness and salinity contents.

The ability of surface waters to neutralize acidic inputs depends primarily on the carbonate (CO) and bicarbonate (HCO) content that is expressed as alkalinity.

4.0 MODELLING APPROACHES

To find a solution to a particular water quality problem, planners need a model of the system that would reflect the major inter relationships between the inputs and outputs or cause effect relationship of the water quality system. The role of physical models in water quality planning is limited mostly to hydraulic and waste treatability studies but mathematical and sometimes socio economical models have found wide application and have been used extensively, especially for non point source pollution problems. The level of models may vary, and in water quality planning, the models can be categorized basically into four levels (Novotny and Chesters, 1981).

i. Simple unit loading complete mixed systems.These systems relate the pollution generation to units of land areas undue a given use or to leading per capita and industrial production on units. These pollution loadings are then divided by the river flow to obtain the pollution concentration.

ii. Simple statistical Regression Models : The pollution loadings and concentration is receiving waters sometimes can be correlated to various causative factors which include population, geology, land use distribution within the watershed, meteorological factors, and flow, among others.

iii. Deterministic Process Oriented Models : In many cases, it is possible to describe the process participating in pollution generation and transport mathematically using analytical or empirical or semi empirical formulae or sub models. These sub units can then be combined into a larger pollution transport model. Steady state models assume that there is no change of the input parameters over the period of simulation, while more complex dynamic models enable simulation of the system undue dynamic input conditions.

iv. Stochastic Process Models : In many cases both system inputs and system parameters are statistical quantities with certain probabilities of occurrence. In some instances, statistical stochastic models can be designed that would provide the transformation of statistical input variables and system parameters into a range of occurrence of the outputs.

v. Socio - Economic Water Quality Models may apparently represent an ideal tool for planners. These models would relate the water quality response to various socio economical inputs. However, it must be understood that while the state of art of pollution generation and transport models has been brought to an adequate modelling level and accuracy, socio economic water quality models still require large amounts of background information for development and implementation. These models are more simplistic in detail and description of sub processes and more complex in scope.

5.0 LITERATURE REVIEW

5.1 TEMPORAL VARIATION

In the search for general patterns of water quality response, many research workers have shown that the magnitude of both suspended sediment concentrations (or turbidity) and solute concentrations may be closely related to the level of water discharge. The former relationship is nearly very case positive, whilst the latter is generally inverse. From a number of studies, the degree of scatter apparent in the relationships indicate that many other factors besides the level of stream flow influence these two quality parameters, but discharge is a useful surrogate for the major processes involved. With suspended sediment, high concentrations are linked to the incidence of both slope and channel erosion, which in turn are affectively restricted to storm events and high discharges. The inverse relationship between solute concentrations and stream flow may be explained in terms of the dilution of solute rich base flow by storm runoff with shorter residence times and therefore lower solute concentrations.

Several studies of suspended sediment response have shown that the simple relationship between concentration (c) and discharge (Q) which generally takes the form $C = K Q^{b}$ may be seriously complicated by seasonal, hysteric and exhaustion effects (e.g. Walling 1974). However, inspite of these limitations, it is broadly applicable to rivers in many areas of the world and reported values of the exponent b commonly lies in the range of 1.0 - 2.0. Similar relationships of the form $C = K Q^{-b}$ have also been widely used to account for variations in total solute concentrations and attempts to develop this simple function have included the use of more complex concentration/discharge functions (e.g. Hall, 1970; 1971) the use of values of individual flow components in mass balance relationships (e.g. Walling, 1974) and the development of mixing models which attempts to incorporate basin wide processes by considering the mixing of precipitation input with catchment storage (e.g. Johnson et. al. 1969).

The fact that the quality of surface stream names with its quantity of flow has been known for a long time. The formulation of thus relationship has been tried several times with varying degree of success. It is important to know a workable method for predicting the water quality at different stages of flow for purposes of quality improvement by dilution with stored water, catchment of the better waters for storage, estimation of yields of various qualities and determination of the economics affects of changing the quality of water.

W.H. Durum (1953) worked with data for saline River in Kansas and found that a simple hyperbolic relationship gave a satisfactory approximation for estimating chloride concentration from a given flow. Arno T. Lenz and Claire N.Sawyer (1944) reversed the usual process and use the alkalinity viruses the stream flow curves for estimating the flow from a measured alkalinity.

J.C. Ward II (1958) used monthly weighted averages of composited sample data for the Arkansas River near Tulsa, Okla, and the Red River near Gainesville Tex. From this study, Ward concluded that the best estimate of the concentration of the dissolved mineral could be obtained from a parabolic relationship between the logarithms of the concentration and the stream discharge, and that a good estimate could be made from a simple logarithmic relationship between the concentration and the stream flow.

All these attempts have only provided partial explanation of the highly complex patterns of variation exhibited by stream solute concentrations. Furthermore, investigation of the behaviour of individual constituents of the solute load in demonstrate marked contrast in response.

5.2 SPATIAL VARIATION

Whereas the various processes of water storage and transfer and water quality modification are responsible for the complex pattern of temporal variation shown by most measures of water quality, spatial variations in water quality are conditioned by the overall climatic framework within which the hydrological system functions and by the characteristics of the drainage basin involved. For convenience, climatic controls may be viewed as operating at the continental or world scale, while catchment characteristics become more significant at the local scale, although this division is clearly somewhat arbitrary.

Gibbs (1970) attempted to illustrate the major factors influencing the solute composition of river water at the global level. The isolated three major controls on chemical composition of lake and river waters. The were related to the source of the solute load and were defined as, first, atmospheric precipitation dominance, secondly, soil and rock dominance, and thirdly, evaporation chemical precipitation dominance. They varied in importance according to the magnitude of annual precipitation or runoff.

The influence of climate on global patterns of suspended sediment concentration was clearly demonstrated by the work of

Langbein and Schumn (1958) who described a general inverse relationship between mean annual precipitation and mean sediment concentrations for catchments in USA.

Because of the numerous mechanisms involved in regulating streamflow quality, catchment characteristics will exert an extremely important influence on the precise nature of that quality, and small differences in these characteristics may produce very significant effects. Many studies have demonstrated the importance of rock type in controlling the solute content of streamflow both in terms of total concentration and the level of individual ions (Douglas, 1968) Webb and Walling, 1974). Similarly there are numerous examples in the literature where the effects of individual catchment characteristics, such as relief, vegetation cover, and soil type, on sediment loads have been deduced (e.g. Schumm, 1954; Striffler, 1965; Wischmeeier and Smith, 1965).

Led better and Gloyna (1964) used hyperbolic relationship to describe the correlation of the inorganic quality of water with the quantity of streamflow. He further made the improvement in the relationships by using a logarithmic formulation with a variable exponent. The exponent was found to depend on the streamflow at some time and on the recent history of flow that is represented by an antecedent flow index.

6.0 STUDY BASIN AND DATA DESCRIPTION

6.1 STUDY BASIN

The Narmada River Basin is selected as the study basin. The Narmada river is one of the major interstate rivers of the country, flowing westward through the states of Madhya Pradesh, Maharashtra, and Gujarat and draining into the Gulf of Khambat of Arabian sea. In terms of catchment area, it is seventh largest among the fourteen major river basins in our country. The basin is broadly divided into Upper, Middle and Lower sections.

The important urban centers in the basin are Mandla, Jabalpur, Hoshangabad, Khandwa, and Khargone in Madhya Pradesh; and Bharuch in Gujarat. There are 40 large and 70 medium scale industrial units operating besides large number of small scale units in the basin.

In this study, the Manot site at 203 km. of Narmada, Catchment area of 14556 sq.km. (CPCB Report : ADSORBS/25/93-94) in upper Narmada basin is selected.

6.2 DATA DESCRIPTION

About 12 years data (1980-1992) have been collected from various sources. The major sources are CWC, M.P.State Pollution Control Board, and Central Pollution Control Board etc.

The data consisted of streams flow and nineteen water quality parameters e.g. pH, conductivity, total dissolved solids, Fe^{++}, Al^{+++}, NH_4^+, NO_3^-, Po_4^{--}, Oxygen demand in 4 hrs. Turbidity, Na^+,

K^+, Ca^{++} , Mg^{++}, CO_3^- ,HCO_3^-,Cl^-, SO_4^{--} and Hardness etc. It has been observed that data contains a missing values and for some of the constituents e.g. Fe^{++} ,Al^{++} , NH_4^+, Po_4^{--} etc. the is very scanty for some statistical analysis. The data for some of the water quality parameters e.g. stream flow (Q), pH, total dissolved solid, conductivity, Ca^{++}, K^+, HCO_3^-, NH_4^+, Fe^{++}, Mg^{++}, Na^+, NO_3^- etc have been shown graphically from Figs. 3 to 19.

By visual interpretation of the graphical representation of the data it is clear that the pH has an increasing trend i.e. the acidity is increasing with time and 90% of the data shows that pH is more than 7 which indicates that the Narmada Water at Manot is alkaline and its akalinity is going on increasing which may affect beneficial uses of water adversely. The scatter plots of NO_3^-, HCO_3^-, Na^+, K^+, Ca^{++}, and Mg^{++} show that their individual ion concentrations are very low from 2700 to 3600 days (period of 1987-1990). However, the stream flow does not have any such abrupt change in the quantity of stream flow. Because the data during this period has different trend from the mass data and it has been considered as outliers and consequently is not included in further analysis.

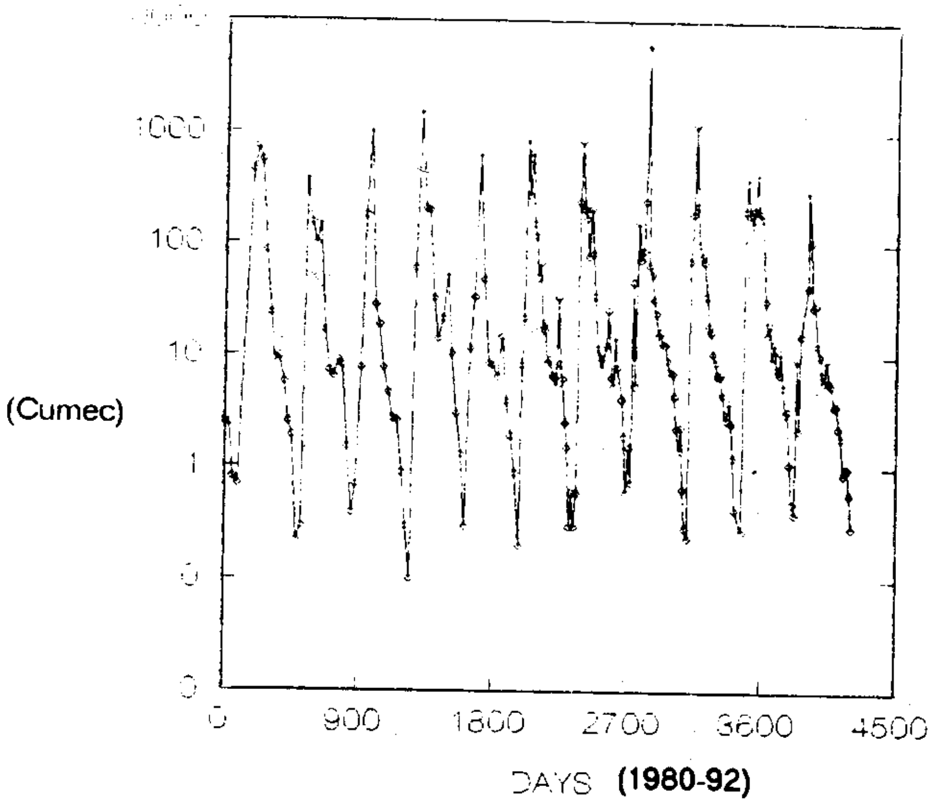


Fig.3 Flow data of river Narmada at Manot

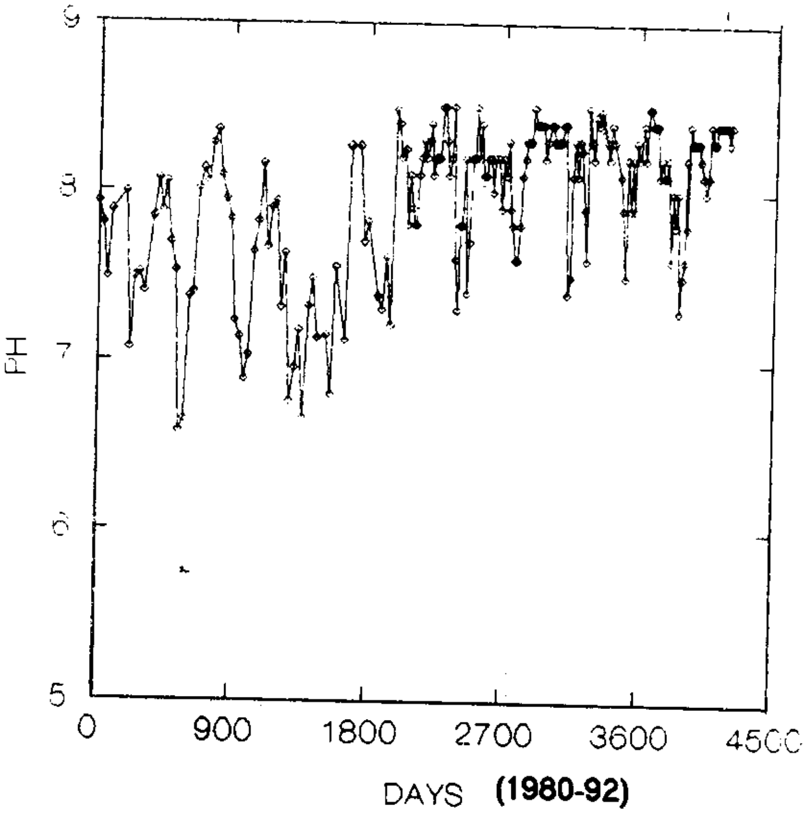


Fig. 4 pH data of river Narmada at Manot

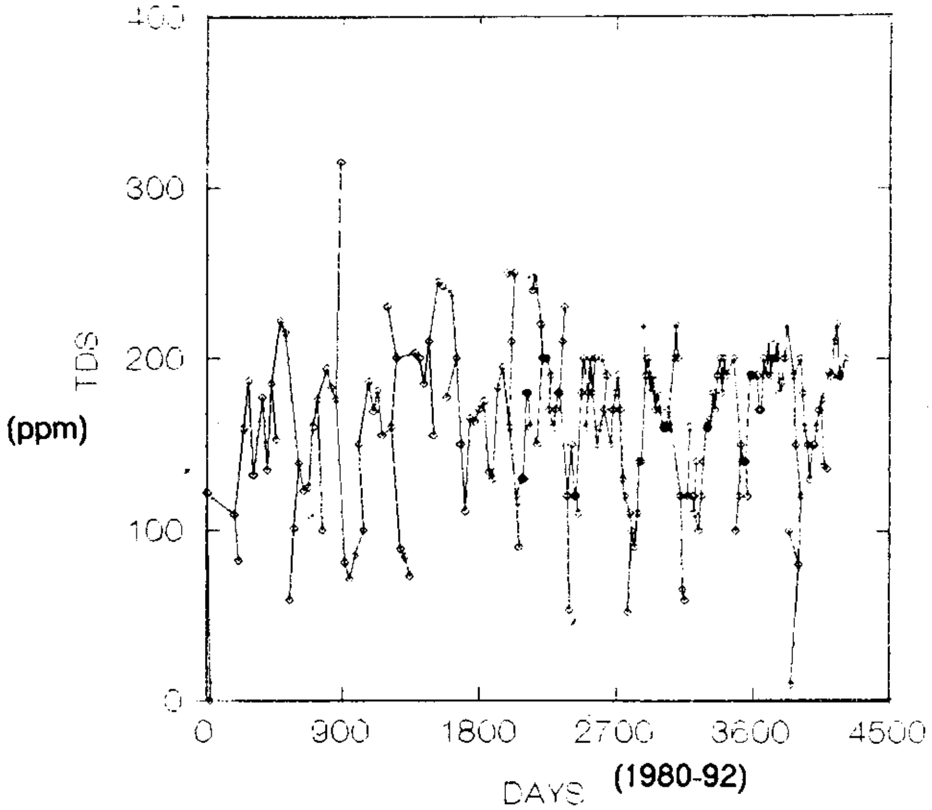


Fig. 5 Total dissolved solids of Narmada water at Manot

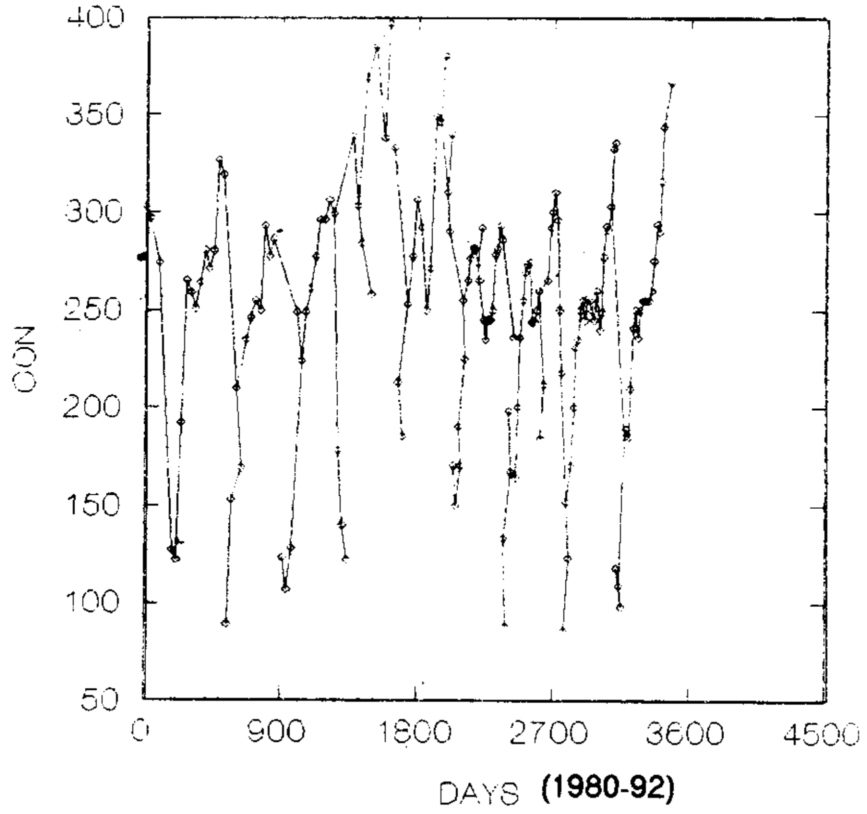


Fig. 6 Conductivity data of Narmada water at Manot

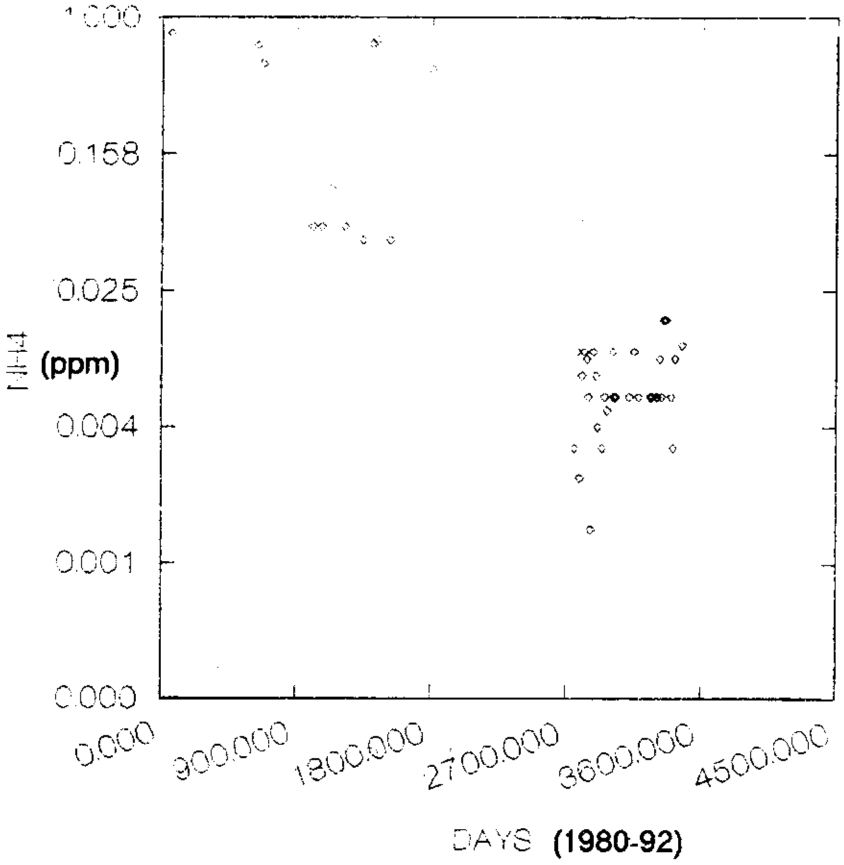


Fig.7 NH$_4^+$ Conc. of Narmada water at Manot

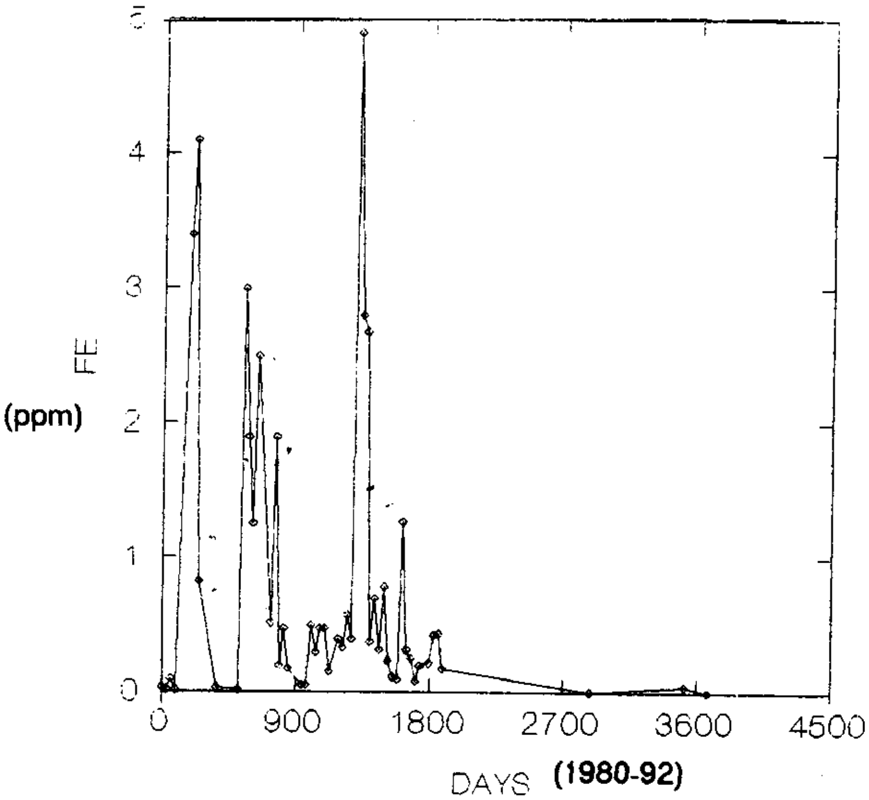


Fig.8Fe^{++} concentration of narmada water at Manot

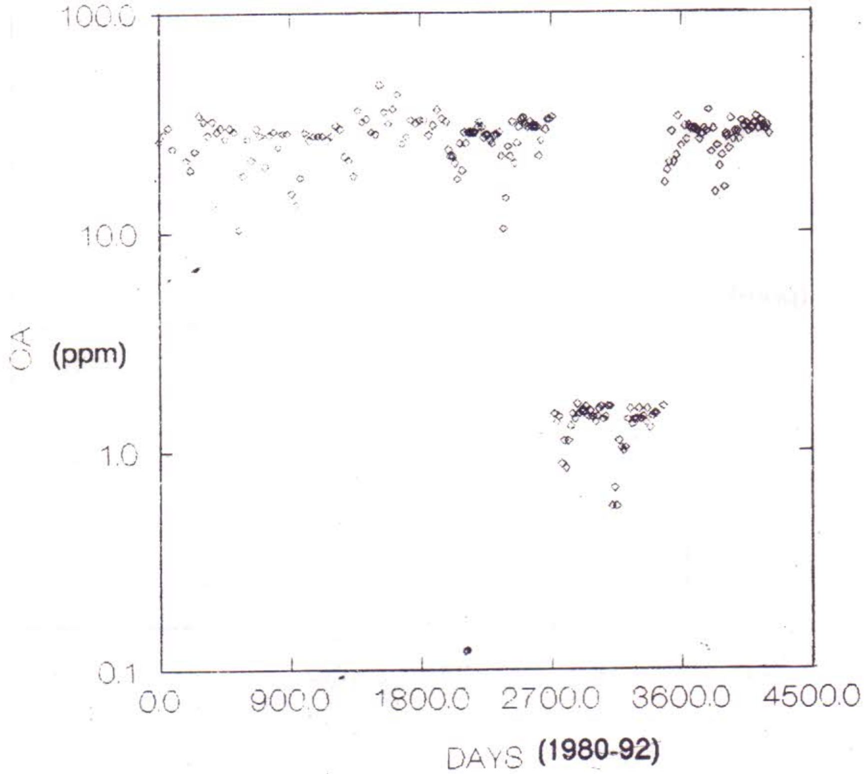


Fig.9 Calcium Ion Conc. in Narmada water at Manot

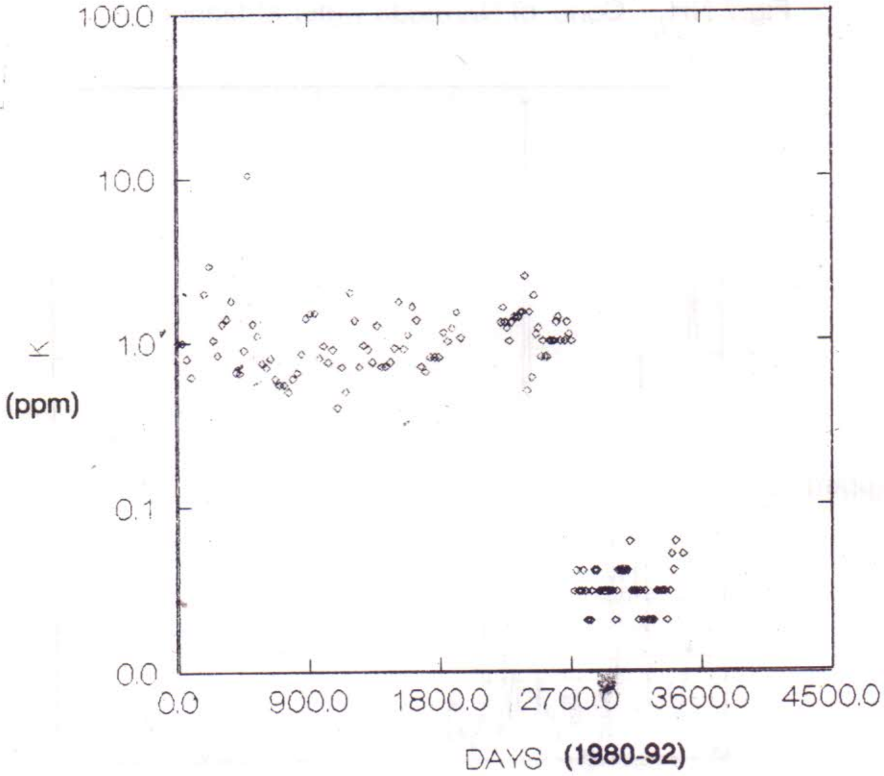


Fig.10 Potacium Ion Conc. in Narmada water at Manot

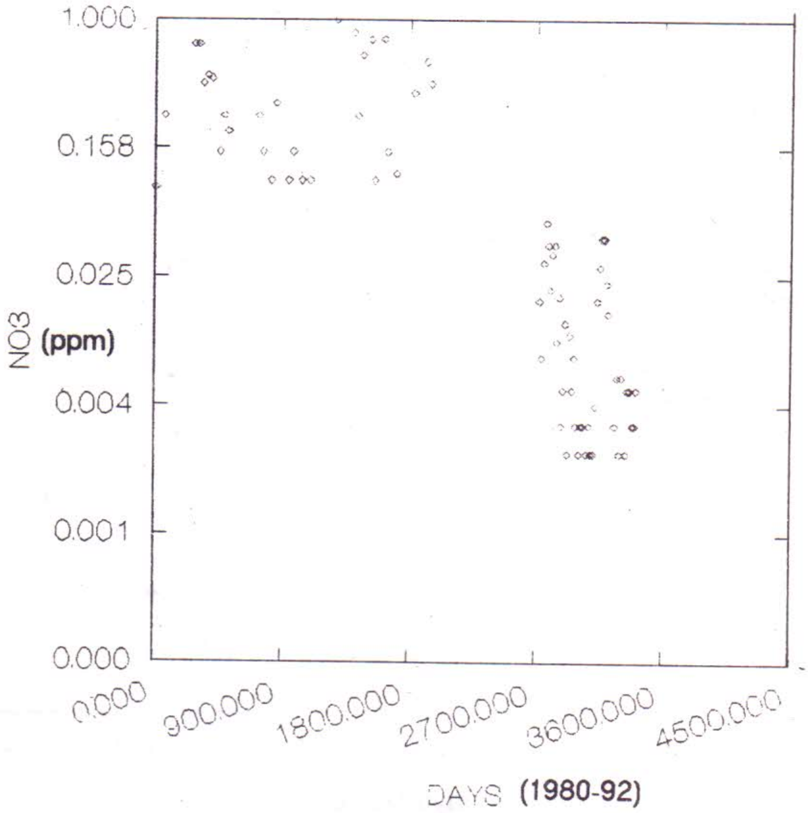


Fig.11 NO_3^- Conc. in Narmada water at Manot

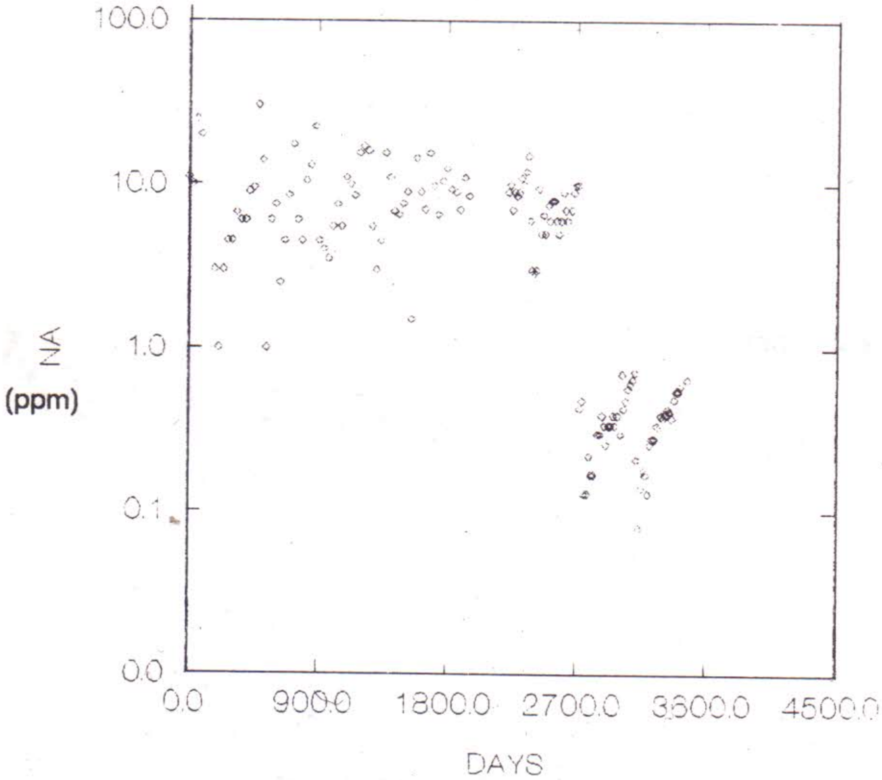


Fig.12 Sodium Ion Conc. in Narmada water at Manot

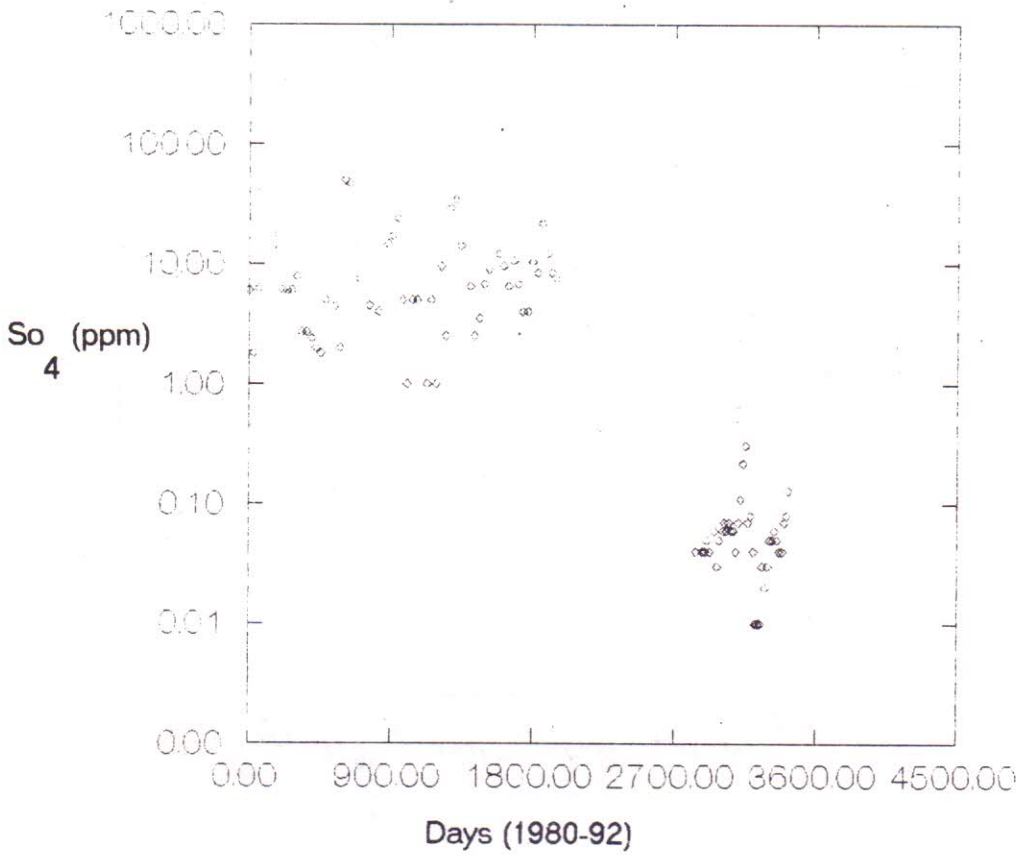


Fig.13 Sulphate Ion Conc. in Narmada water at Manot

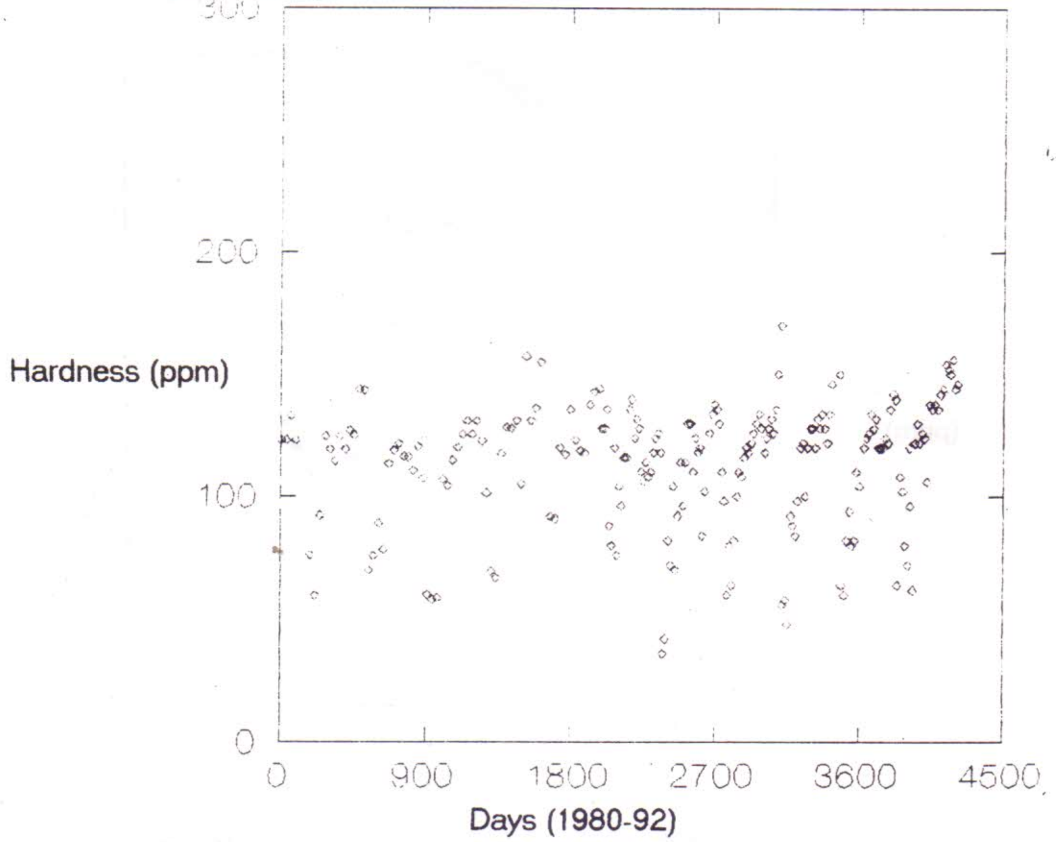


Fig.14 Hardness in Narmada water at Manot

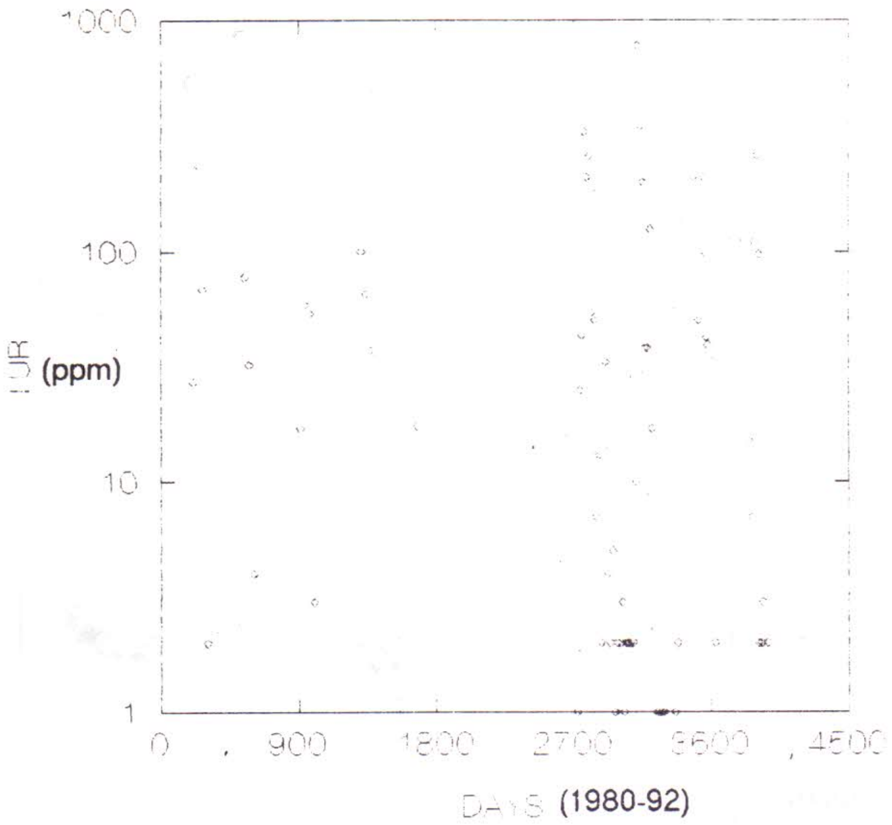


Fig.15Turbidity of Narmada water at Manot

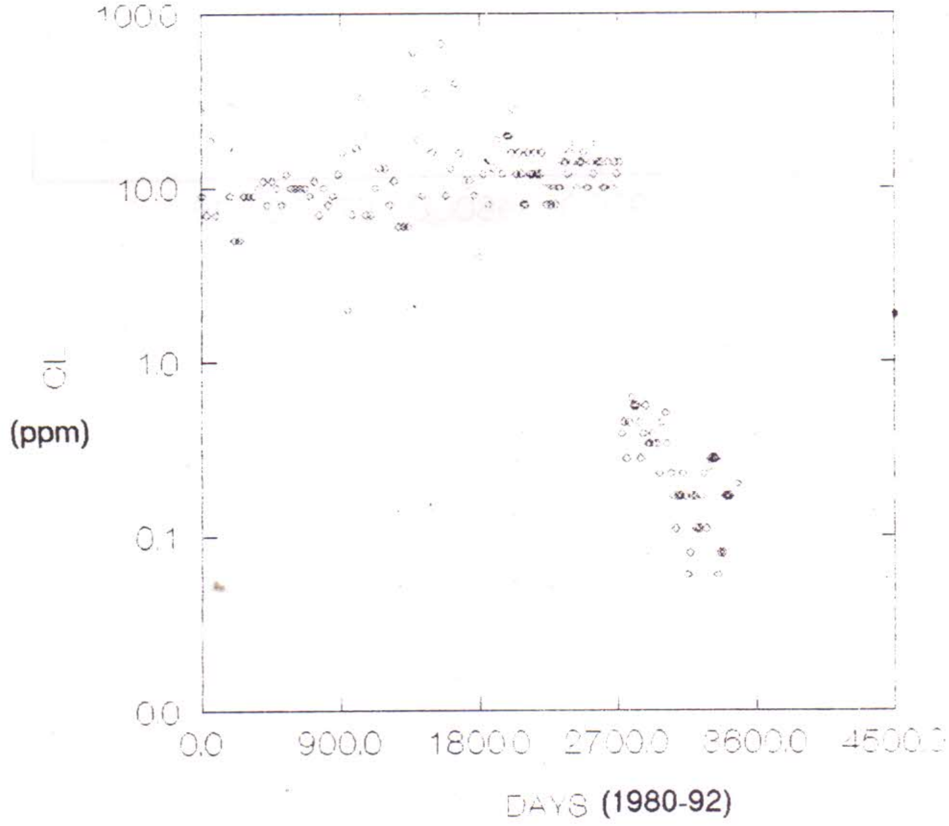


Fig.16 Chloride Conc. in Narmada water at Manot

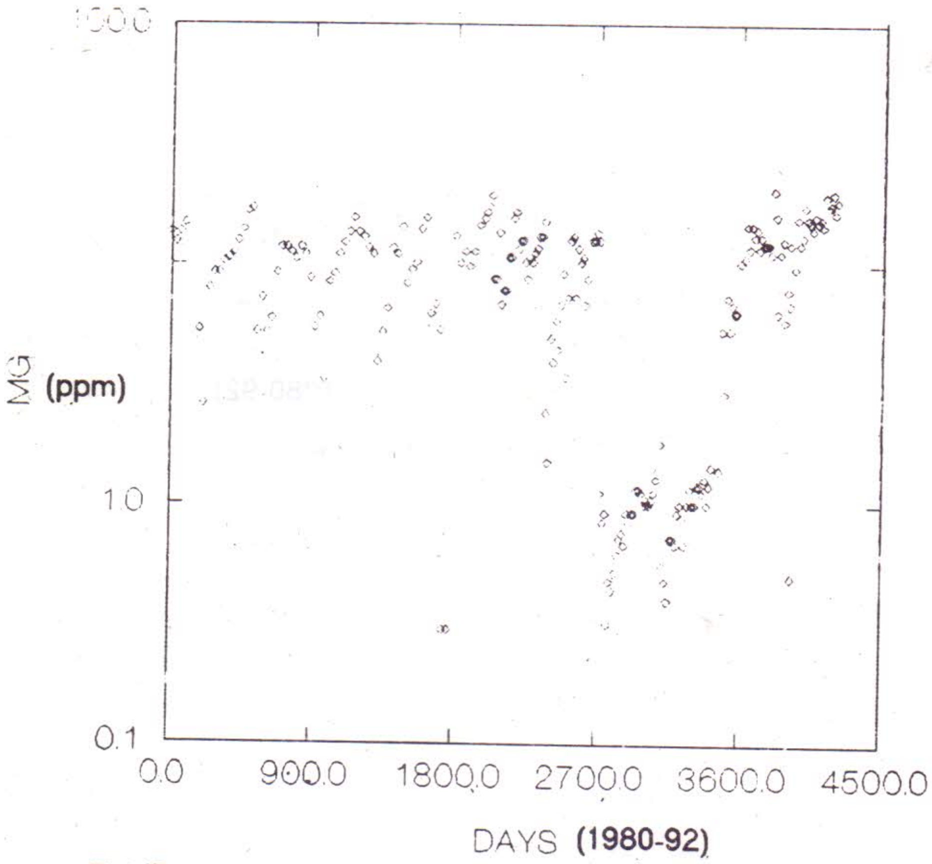


Fig.17Conc. of Mg Ion in Narmada water at Manot

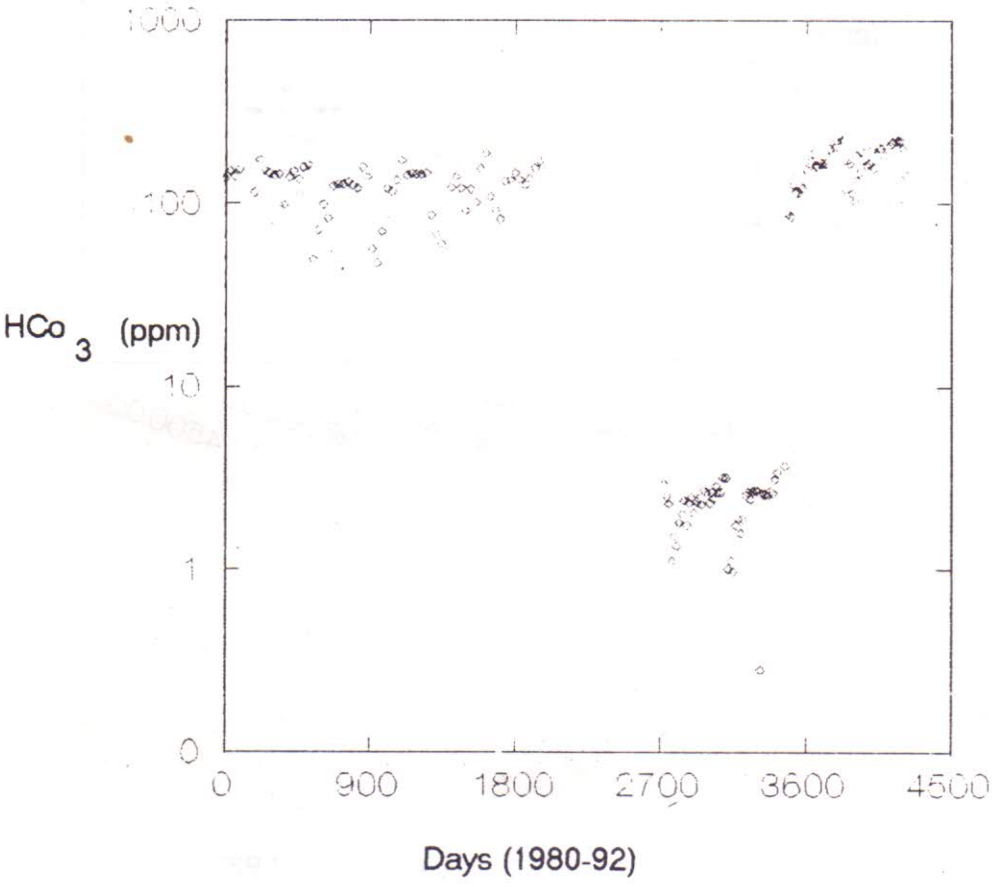


Fig.18 Bicarbonate Ion Conc. in Narmada water at Manot

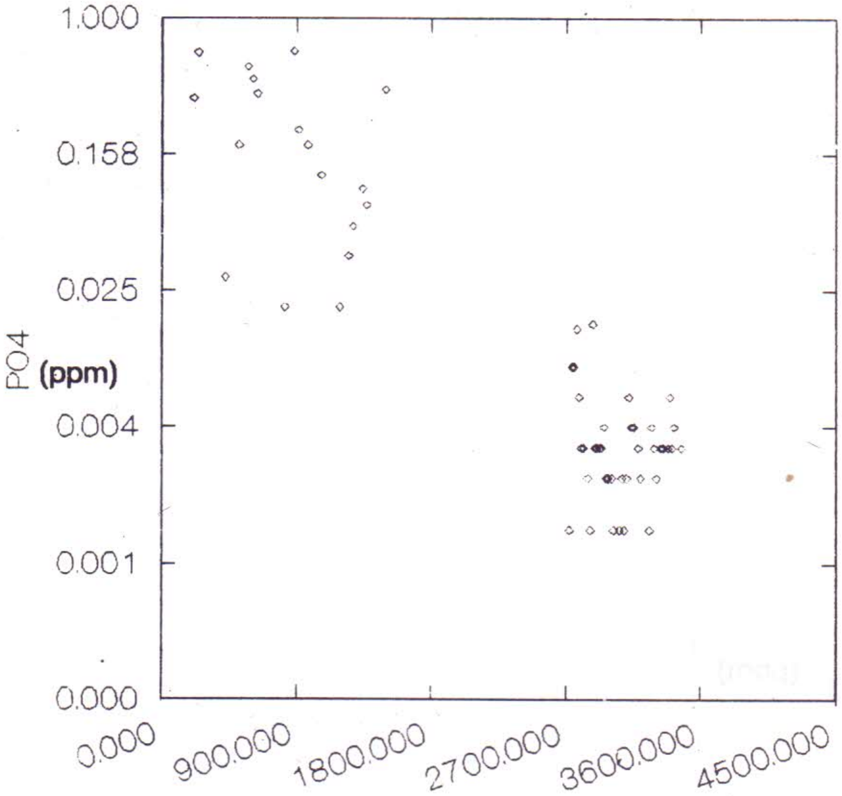


Fig.19 Po_4^{--} Conc. in Narmada water at Manot

7.0 General Model Building

The basic model that we consider is

$$Y = X\beta + e \qquad (1)$$

$$\mathrm{var}(e) = \sigma^2 I_n$$

in which, **X** is data matrix, **Y** is dependent variable vector. $\mathbf{I}_n$ is an identity matrix, β is unknown vector, and **e** is an error vector having zero mean and constant standard deviation (σ).

7.1 Parameter Estimation

The method of ordinary least squares is the most widely used method because of the simple concept and no assumption is necessary on the probability distribution of data. This method will be used for estimating parameters for phosphorus and is briefly described here. Let y_i (i = 1,......,n) be the response variables with mean μ_i. Consider a following linear model expressing the relationship between means μ_i and the explanatory variables x_{ij} (value of jth independent variable for observation i).

$$\mu_i = \sum_{j=0}^{k} \beta_j x_{ij} \qquad (2)$$

where, $x_{i0} = 1$, β_j are the parameters to be estimated, and k+1 are the number of parameters in the model.

The method of least squares minimizes the sum of squares between the observed and model computed-values of dependent variable and in the process estimates the parameters that will provide the least sum of squares of error. The function to be minimized with respect to the model parameters (β_j) is

$$\eta = \sum_{i=1}^{n} (y_i - \sum_{j=0}^{j=k} \beta_j x_{ij})^2 \tag{3}$$

where, y_i are the observed values of the dependent variable and η is the sum of squares of errors. The above function is minimized by partially differentiating the above equation with parameters and equating to zero. The equations so derived are solved for the parameters. The general solution of these equations in the matrix notation are as follows (Draper and Smith,1981).

$$\hat{\beta} = (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{y} \tag{4}$$

where, ^{T} represents the transpose of the matrix, and $\hat{\beta}$ is the vector of estimated parameters, **y** is an n dimensional vector of dependent variables, and **x** is the nx(k+1) coefficient matrix of independent variables.

The vector of model-computed dependent variables $\hat{y}$ used for estimating the standard deviation and other statistical tests in matrix notation is defined as $\mathbf{x}\hat{\beta}$. The point estimates of the parameters obtained from Equation (4) and the regression process itself need to be examined before these could be used in the model application. In order to facilitate statistical diagnostics on parameters and regression, the errors $(y_i - \hat{y}_i)$ are assumed independent, normally distributed with zero mean and constant variance Specifically, the following question must be answered

(1) Is the linear model suitable in explaining the variation in dependent variables (y_i) at various locations and will be model still be useful for another independent data set?,

(2) Is the assumption of normality and constant variance in error structure valid?,

(3) What are the uncertainties in the estimated parameters and are the parameters significant in the statistical sense?, and

(4) Is the overall regression significant in a statistical sense?

7.2 Statistical Tests

Various statistical tests are used to address the above concerns, some of these are briefly described here.

7.2.1 Model Performance

The squared multiple R, represents the proportion of the variation in the dependent variable accounted for, or explained by, the linear model. The higher R square values indicate better performance of model and suitability of model in computing the dependent variable (Draper and Smith, 1981). The other related test to examine the suitability of the model for a new data set (from the same population) of dependent variable is the adjusted squared multiple R (Wilkinson, 1990). The adjusted squared multiple R will be smaller than squared multiple R values as the coefficients would be optimized for first data set.

The model performance is also judged by visual examination of the linear plot of observed and model computed dependent variables. The assumptions of normality and variance can be examined by plotting residuals (errors) and y_i, for all observations. A randomly distributed plot is indicative of constant variance and normal distribution.

7.2.2 Uncertainties in Parameters (Confidence Interval and Hypothesis testing)

A 95 % confidence interval indicates a range which will include the true value with 0.95 probability. The narrower the interval the better is the estimation. For correlated parameters, joint confidence region is a better representation of confidence in estimated parameters than the individual confidence interval as it considers parameters jointly responsible in calculating the dependent variables (Draper and Smith, 1981). Inclusion of zero in confidence interval/region signifies that zero is a possible value of the parameters. The other test frequently used for examining if estimated parameters are significantly different than zero(or any other value) is the hypothesis testing using t-statistics (Draper and Smith, 1981). If the observed t-value is higher than the critical t-value at a certain level of significance, it indicates the estimated parameter to be significantly different than zero.

7.2.3 Analysis of Variance and F-statistics

The analysis of variance (ANOVA) and the value of F-ratio is used for assessing the significance of regression. When the F-ratio is statistically significant, it implies that a significantly large amount of the variation in the data about the mean has been taken up by the regression equation (Draper and Smith, 1981). The above statistical investigations greatly assist in examining the model performance and quality of the estimated parameters.

7.3 Detection of Influential Cases

The technique of case analysis, in which the data are analyzed in detail with attention given to the role of each case in determining values of estimators and test statistics, is used in the detection of influential cases. The primary concerns of case analysis are two interrelated questions. First, how well the model used resembles the data actually observed. The basic statistic here will be the residual. If the fitted model does not give a set of residuals that seems reasonable, then some aspect of the model will be called into doubt. The second question of interest is the effect of each case on estimation and other aspects of aggregate analysis. In some data sets, for example, it may be that the observed aggregate statistics depend on one case in such a way that, if that case were detected, the outcome of the aggregate analysis would change. Such cases are termed as 'influential cases' and could be detected using two case statistics called leverage values and distance measures.

7.3.1 Leverage Value

The vector of residuals $\hat{e}$ is defined by

$$\begin{aligned} \hat{e} &= Y - \hat{Y} \\ &= Y - X(X^TX)^{-1}X^TY \\ &= [\,I - X(X^TX)^{-1}X^T]\,Y \end{aligned} \tag{5}$$

The matrix defined by $X(X^TX)^{-1}X^T$ is very important in the study of case analysis, so we shall give it a name V, defined by

$$V = X(X^TX)^{-1}X^T \tag{6}$$

Using this definition, the fitted values are given by

$$\hat{Y} = VY \tag{7}$$

and the residuals are given by

$$\hat{\mathbf{e}} = (\mathbf{1} - \mathbf{V})\mathbf{Y} \tag{8}$$

By assumption, the errors are uncorrelated random variables with zero means, and common variance σ^2. Using (5) the moments of $\hat{e}$ are given (Weisberg, 1980) as

$$\begin{aligned} \mathbf{E}(\hat{\mathbf{e}}) &= \mathbf{0} \\ \mathbf{var}(\hat{\mathbf{e}}) &= \sigma^2(\mathbf{1} - \mathbf{V}) \end{aligned} \tag{9}$$

Like the errors, each of the residuals has zero mean, but they have different variances and they are not uncorrelated.
The elements of $\mathbf{V}$, the v_{ij}'s, are given by the equation

$$\mathbf{v}_{ij} = \mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_j \tag{10}$$

and, for the diagonal elements,

$$\mathbf{v}_{ii} = \mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})\mathbf{x}_i \tag{11}$$

where $\mathbf{x}_i^T$ and $\mathbf{x}_j^T$ are, respectively, the ith row and the jth row of the data matrix $\mathbf{X}$. From (9), the variance of ith residual is

$$\mathbf{var}(\hat{\mathbf{e}}_i) = \sigma^2(\mathbf{1} - \mathbf{v}_{ii}) \tag{12}$$

and the covariance between the ith and jth residual is

$$\mathrm{cov}(\hat{e}_i, \hat{e}_j) = -\sigma^2 v_{ij} \tag{13}$$

Also, the correlation between the ith and jth residual is

$$\mathrm{corr}(\hat{e}_i, \hat{e}_j) = \frac{-v_{ij}}{(1-v_{ii})^{\frac{1}{2}}(1-v_{jj})^{\frac{1}{2}}} \tag{14}$$

It is clear that each v_{ii} must fall in the range between 0 and 1. The notion of v_{ii} giving a measure of how far the ith case is from the center of the data is central to case analysis. As can be seen from (12), cases with large values of v_{ii} will have small values for $\mathrm{var}(\hat{e}_i)$; as v_{ii} gets closer to one, this variance will approach zero, and as this happens, regardless of the value of y_i observed for the ith case, it is nearly certain to get a residual for the ith case near zero; that is $\hat{y}_i \equiv y_i$. Such a case can be very important in estimating parameters. Cases with large v_{ii} have been called high leverage cases (Weissberg, 1980). In multiple regression, v_{ii} measures the distance from the point x_i to the center of the data, and cases with unusual values for the independent variables will tend to have large values of v_{ii}.

7.3.2 Studentized Residual

As discussed above, $\mathrm{var}(\hat{e}_i)$ will be small whenever v_{ii} is large, so cases with x_i near $\bar{x}$ will be fit poorly, and cases with x_i far from $\bar{x}$ will fit well. This is particularly undesirable because violations of a model may be most likely to occur under unusual conditions. The detection of those violations by simply examining the residuals is not possible. However, an improved set of residuals can be obtained by scaling so that cases with large v_{ii} get larger scaled residuals, and cases with smaller v_{ii} get smaller scaled residuals. Then, all the residuals in the analysis can be compared directly. One good way of doing scaling is to divide each of the residuals by an estimate of its standard deviation. Such scaled residuals are called Studentized residuals (Weisberg, 1980), and defined by

$$r_i = \frac{\hat{e}_i}{\sigma\sqrt{1 - v_{ii}}}, \qquad i=1,2,\ldots\ldots,n \qquad (15)$$

The most important advantage of the Studentized residuals is that $\mathrm{var}(r_i) = 1$ for all i, independent of both σ^2 and the v_{ij}'s, as long as the model is correct; when the model is incorrect, then $\mathrm{var}(r_i)$ will generally not be constant over all i (Weisberg, 1980). As a guideline, the cases having their absolute value of r_i more than 3 (Draper and Smith, 1981) may be suspected to be outliers and hence their influence must be studied.

7.3.3 Cook's Distance

The residuals, or the Studentized residuals, are the most commonly used case statistics to measure the success or failure of fitting a model at each case. Now to declare a particular case to be an outlier, its impact up on the values of the model estimates should be checked. If this is found to be significant then that particular case may be called an outlier and must be removed from the data set ; however, these points could be the most important observations in the data set.

In the context of case analysis, rather than comparing estimation techniques, it is important to study the change in the estimate of β when a case is to be deleted from the data. Viewing the estimate $\hat{\beta}$ from the full sample as a fixed point, let $\hat{\beta}_{-i}$ be the least squares estimate of β obtained from the regression using all the cases except the ith. An empirical version of the influence function is obtained by taking the difference between the full data estimate and the estimate using (n-1) cases (excluding the ith), $\hat{\beta}_{-i} - \beta$. A method of measuring the distance between these points is needed to judge whether the ith case has sufficient influence on the estimation of parameters and deletion of it would result in a substantially different conclusion. Cook's distance (D_i) given below, is used to obtain a confidence region for $\hat{\beta}$

$$D_i = \frac{(\hat{\beta}_{-i} - \hat{\beta})^T (X^T X)(\hat{\beta}_{-i} - \hat{\beta})}{(k+1)\sigma^2} \tag{16}$$

where D_i is the Cook's distance, and (k+1) is the total number of parameters. D_i is compared to F(k+1,N-k-1,1-α) for selected α; a large D_i denotes an influential ith observation.

8.0 ANALYSIS AND DISCUSSION

Analysis of the water quality data is carried out in three different ways.

i. Quality - Quantity relationship
ii. Frequency distribution
iii. Regression based Modelling

The above approaches discussed in detail as follows:

8.1 QUALITY - QUANTITY RELATIONSHIP

The scatter diagrams of stream discharge versus water quality constituents concentration, when plotted an logarithmic graph paper, indicated that an equation of the form:

$$C = K\ Q^{b} \qquad (1a)$$

in which Q is the stream discharge, c denotes the water quality constituent concentration, and k and b are regression parameters, could be used for estimating quality for a given quantity.

The scatter diagrams showing the quality - quantity relationship for various parameters has been shown in Figs. 20 to 29. The quality-quantity relationships as described by eq. (1a) are developed for various water quality determinands. The coefficients K and b are presented in tabular form (Table 3).

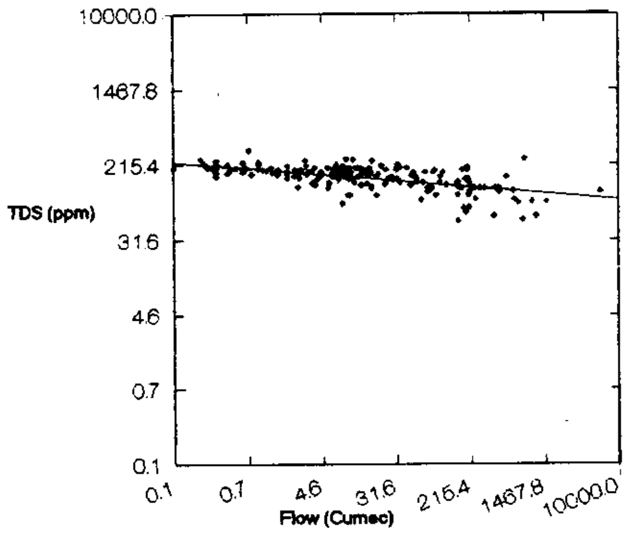


FIG.20 Plot of Flow versus TDS

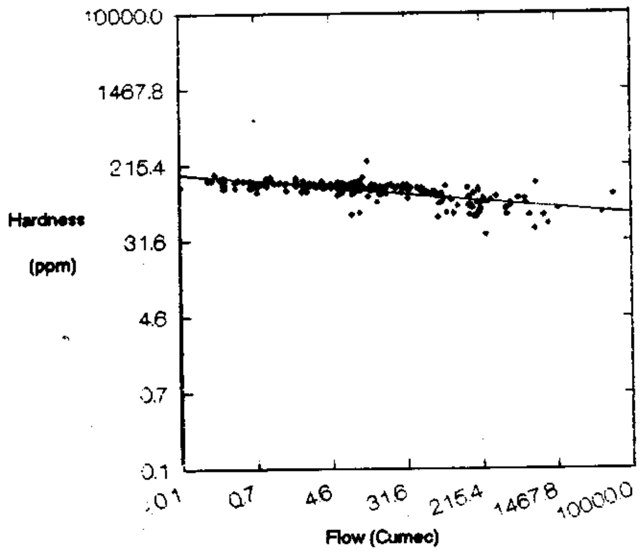


FIG.21 Plot of Flow versus Hardness

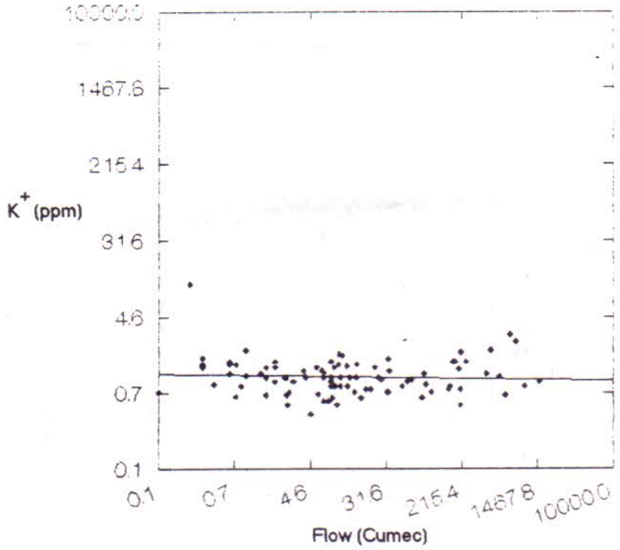


FIG.22 Plot of Flov Versus K^+ Ion

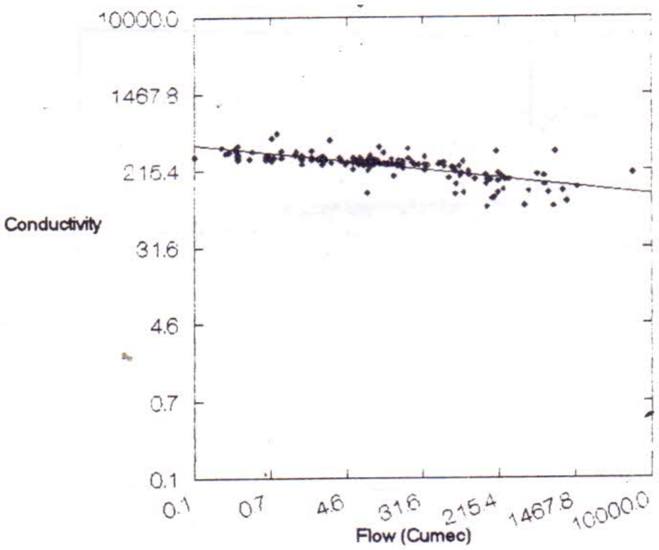
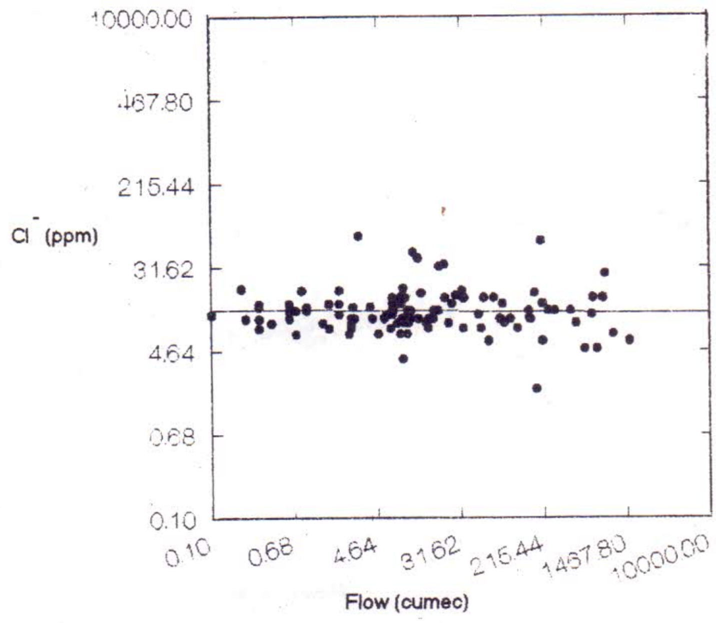
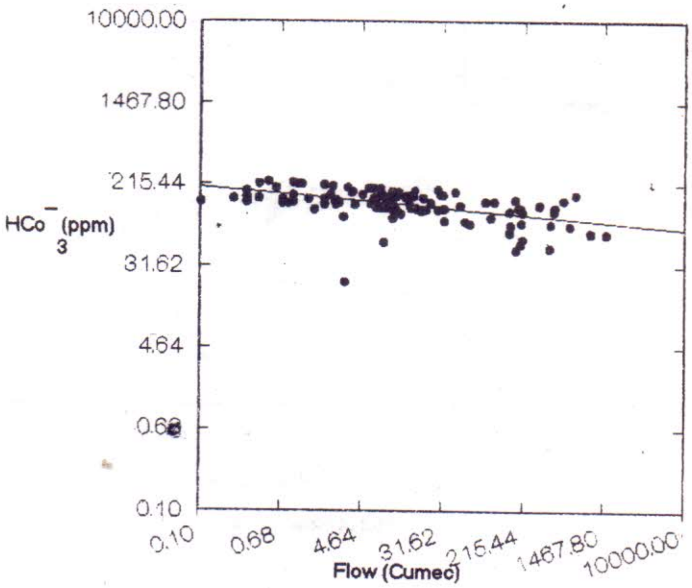


FIG.23 Plot of Conductivity versus Flow



F I G 24 :Plot of Flow Vs Cl-ion



F I G.25 :Plot of Flow Vs $HCo_3^=$

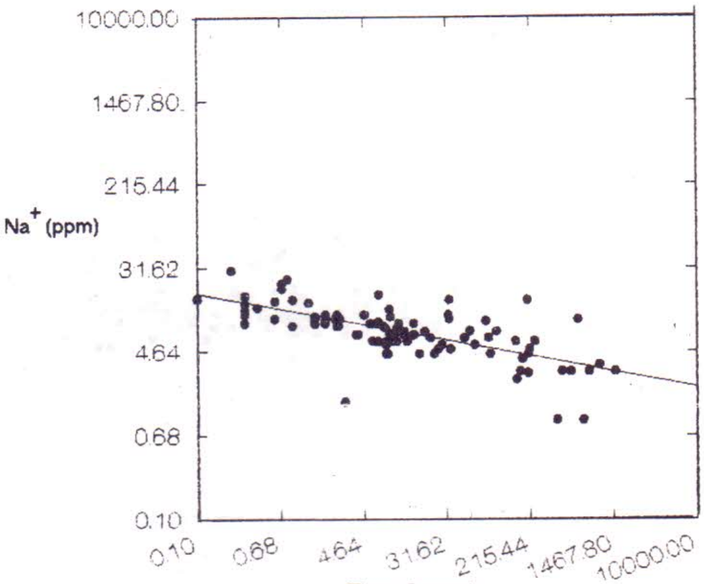


FIG.26 :Plot of Flow Vs Na-ion

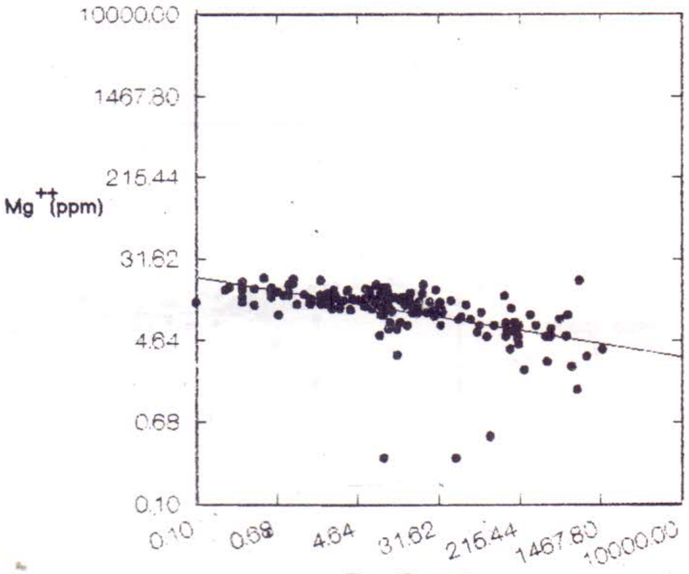


FIG 27 :Plot of Flow Vs Mg-ion

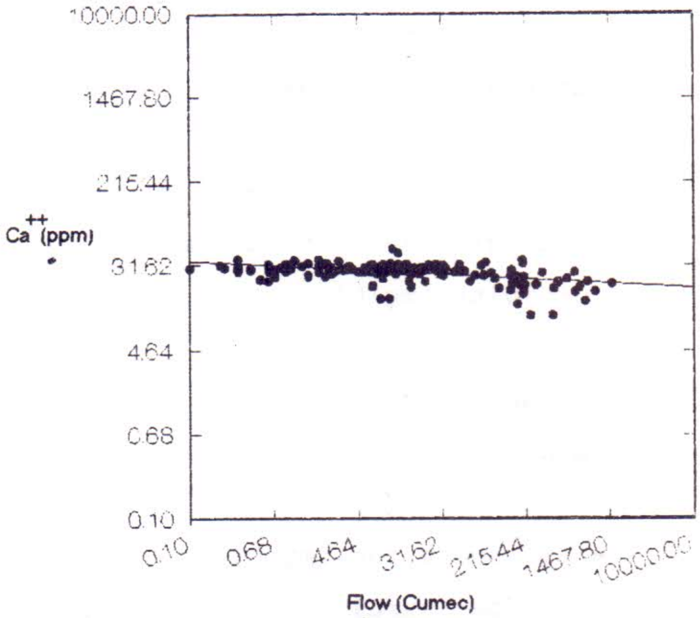


FIG.28 :Plot of Flow Vs Ca-ion

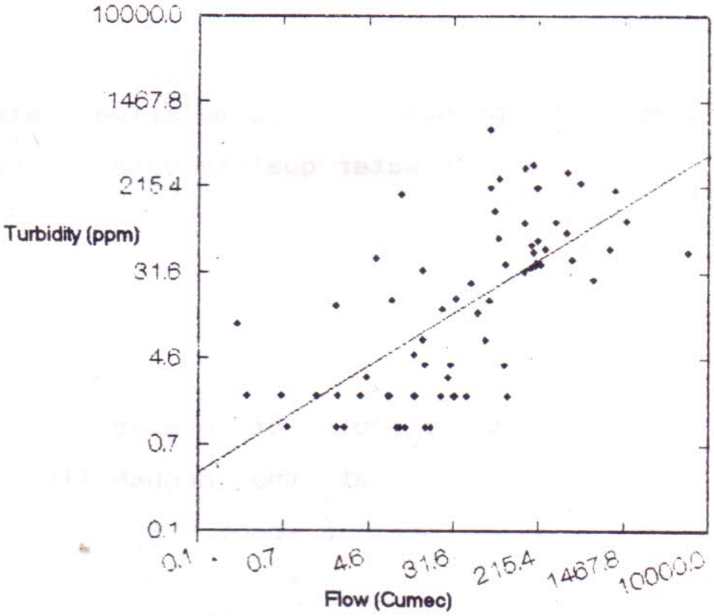


FIG.29 Plot of Flow versus Turbidity

Table 3: Coefficients of Quality-Quantity relationships for various water quality determinands

S. No.	Water quality parameter	K	b
1.	Conductivity (micro-mho/cm)	$10^{5.705}$	-0.104
2.	Total dissolved solids (mg/L)	$10^{5.252}$	-0.074
3.	Turbidity (mg/L)	$10^{0.446}$	0.600
4.	Sodium (mg/L)	$10^{2.457}$	-0.186
5.	Pottacium (mg/L)	$10^{0.061}$	-0.015
6.	Calcium (mg/L)	$10^{3.426}$	-0.052
7.	Magnesium (mg/L)	$10^{2.644}$	-0.161
8.	Bicarbonate ion (mg/L)	$10^{5.102}$	-0.085
9.	Chloride (mg/L)	$10^{2.473}$	-0.008
10.	Sulphate (mg/L)	$10^{4.478}$	1.539
11.	Hardness (mg/L)	$10^{4.419}$	-0.084

8.2 FREQUENCY DISTRIBUTION

The cumulative probability (Cumulative relative frequency) distribution for various water quality determinands are shown in Figs. 30 to 43. On x-axis the concentration of the water quality parameter and on y-axis the respective non-exceedence probability have been plotted.

It is clear from the plots of probability distribution of stream flow (Fig. 30) that the probability of stream flow remaining less than 10 cumec is about 50% and there are very rare chance for exceeding the stream flow beyond 1000 cumec. From Fig. 31, conductivity of stream water at Manot remains always less than 1000 micro-mho/cm which is classified under class-I water (USA classification) and considered good to excellent- suitable for

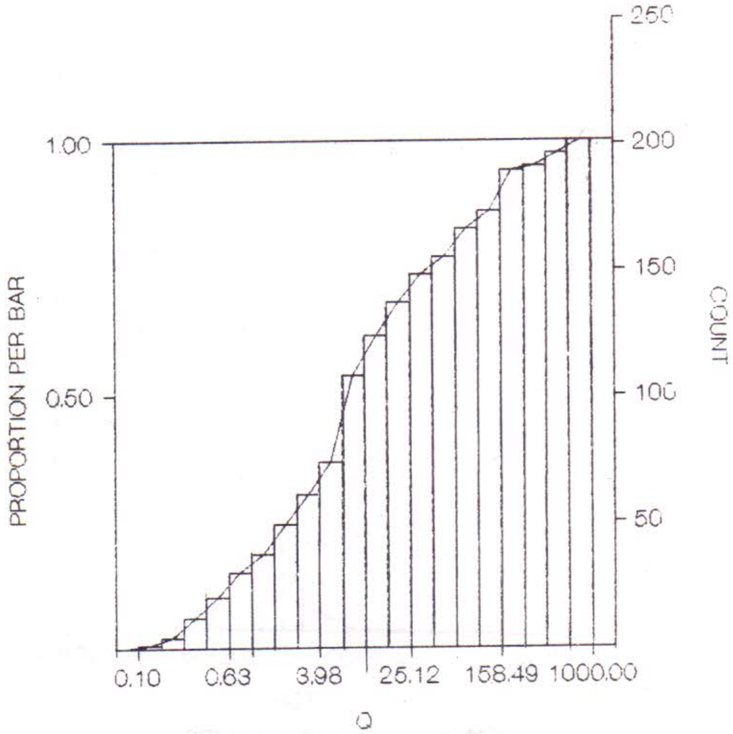


FIG-30 :Frequency distribution of stream flow

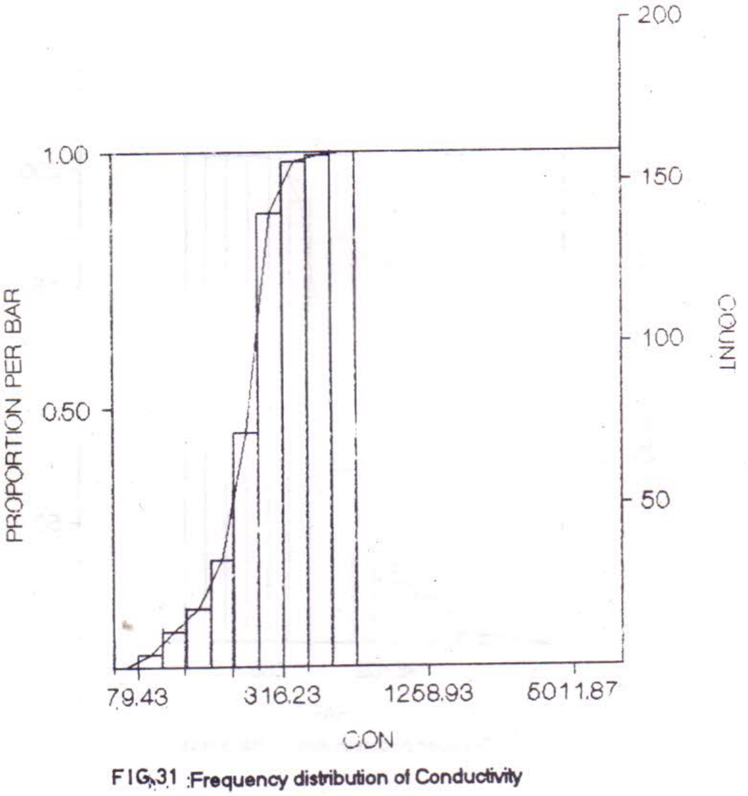


FIG.31 :Frequency distribution of Conductivity

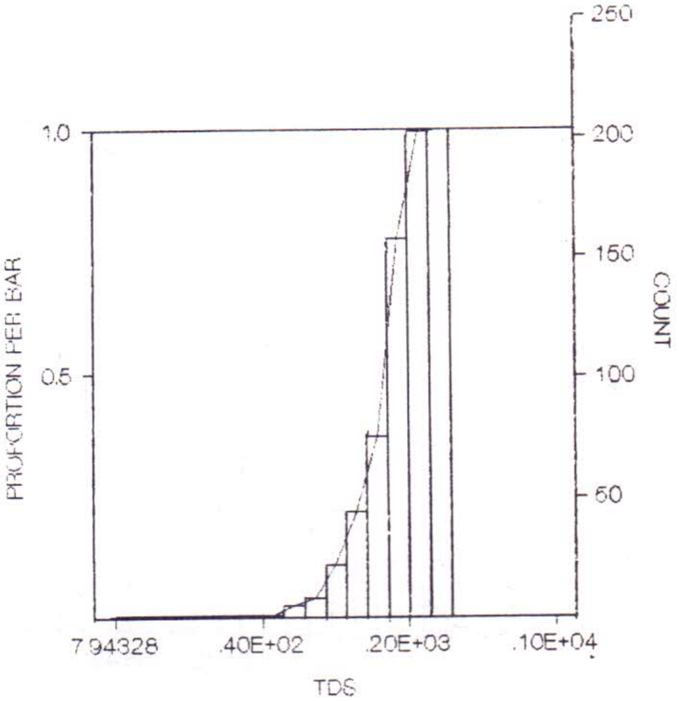


FIG.32 :Frequency distribution of TDS

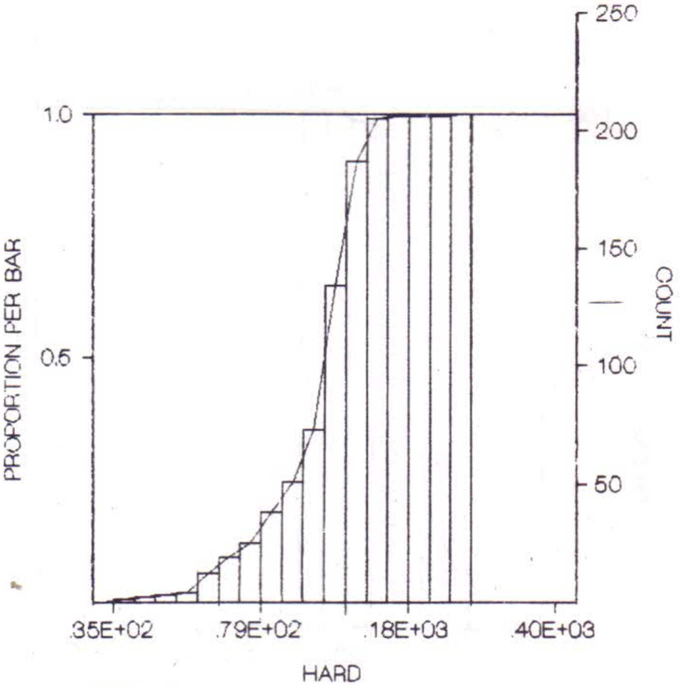


FIG.33 .Frequency distribution of Hardness

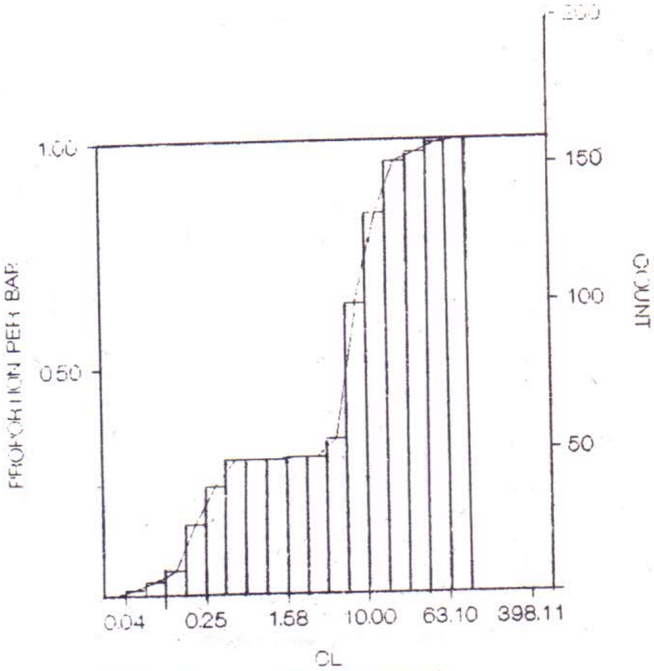


FIG.34 : **Frequency distribution of Chloride**

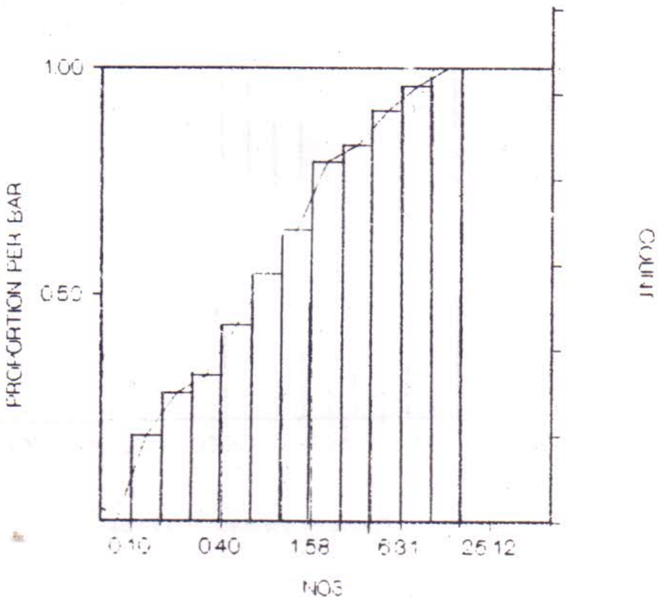


FIG.35 :**Frequency distribution Nitrate**

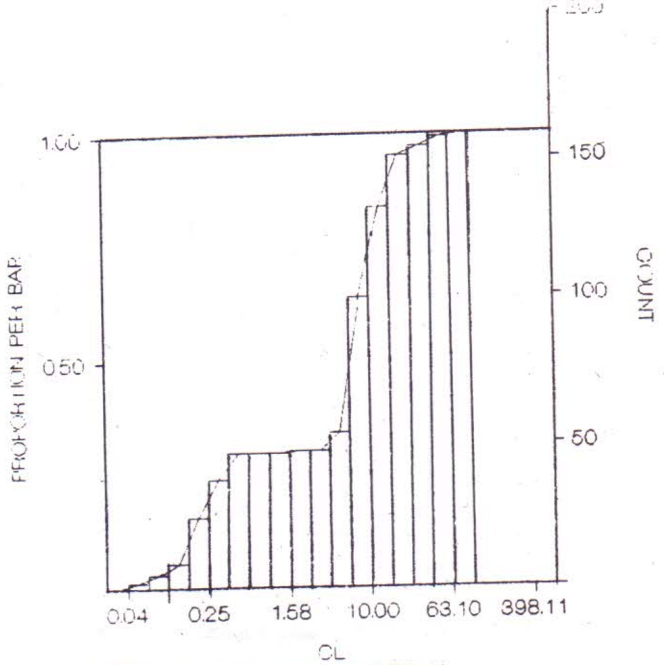


FIG.34 : Frequency distribution of Chloride

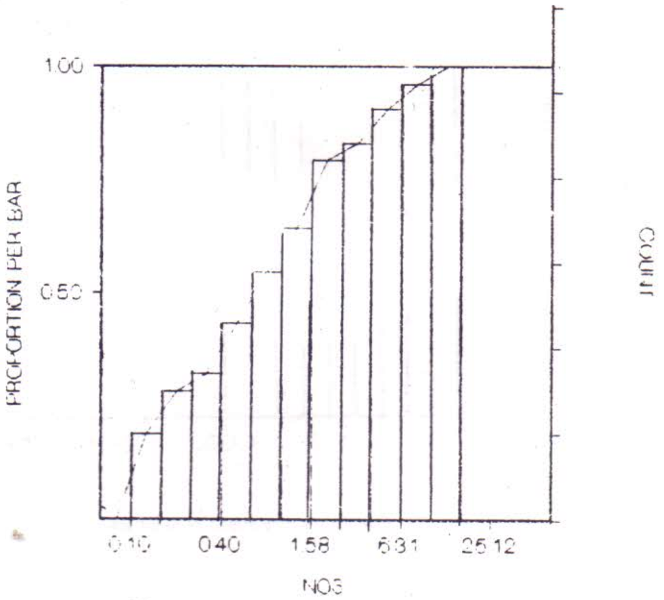


FIG.35 :Frequency distribution Nitrate

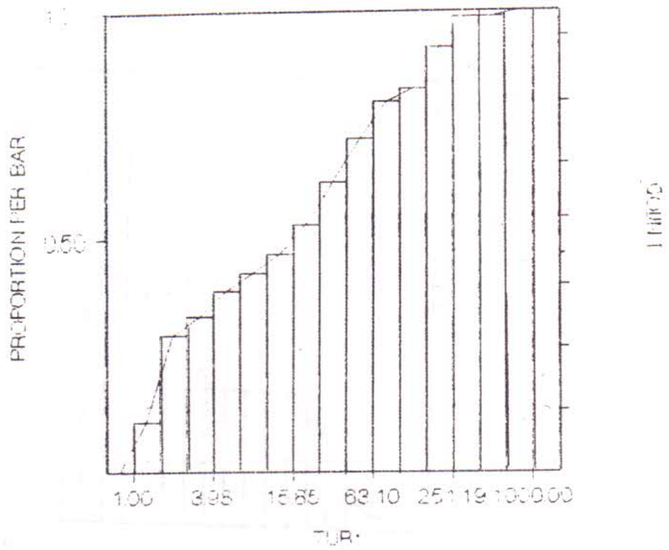


FIG. 36 :Frequency distribution of Turbidity

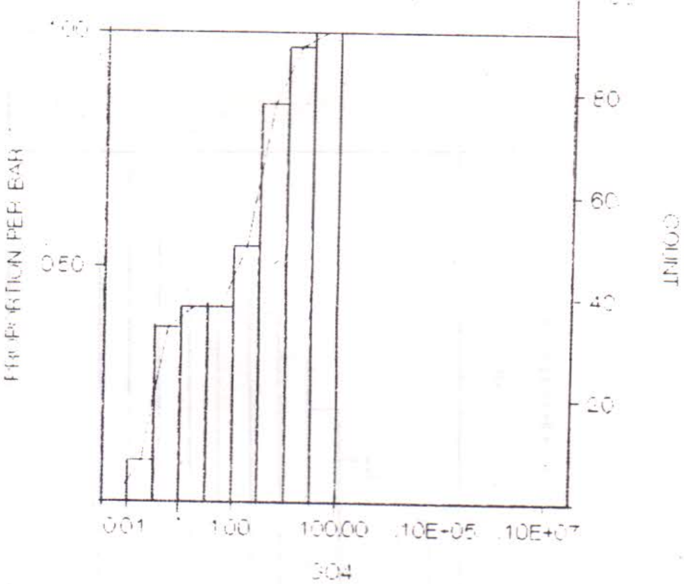


FIG. 37 :Frequency distribution of Sulphate

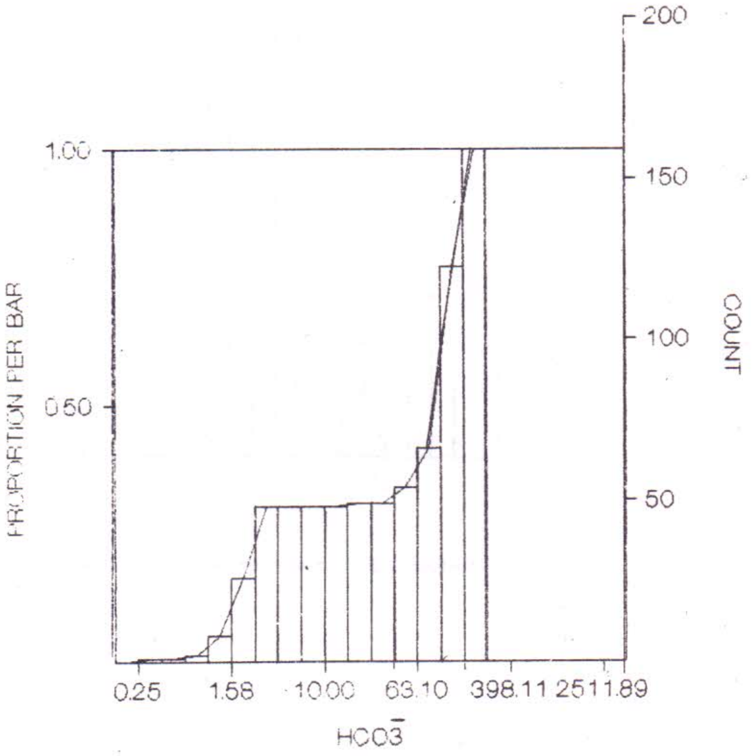


FIG.38 :Frequency distribution of Bicarbonate ion

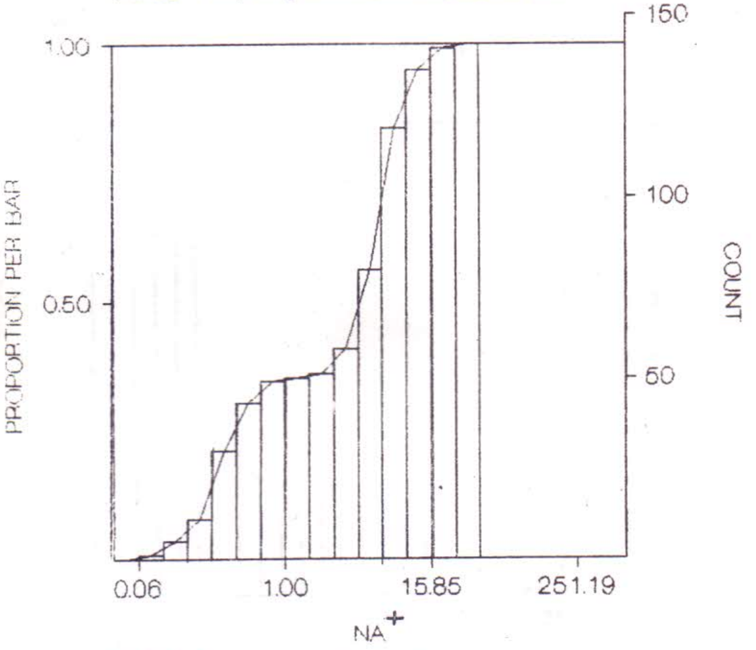


FIG.39 :Frequency distribution of Sodium ion

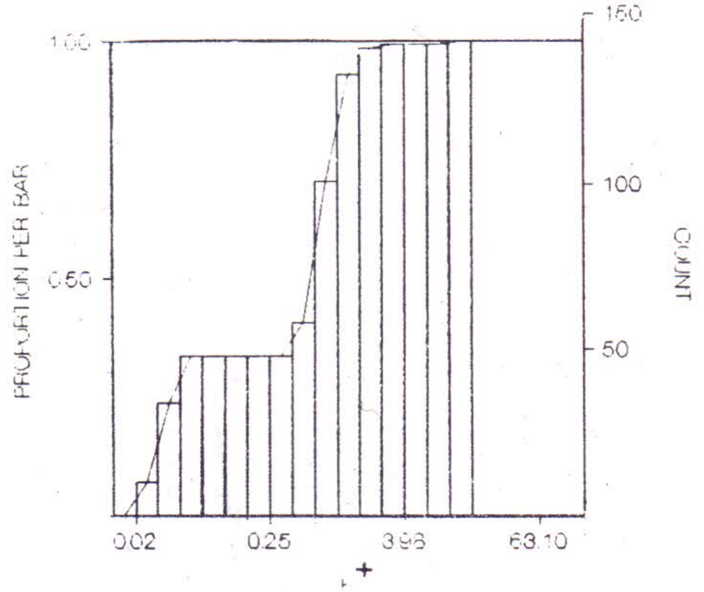


FIG. 40 :Frequency distribution of Pottacium ion

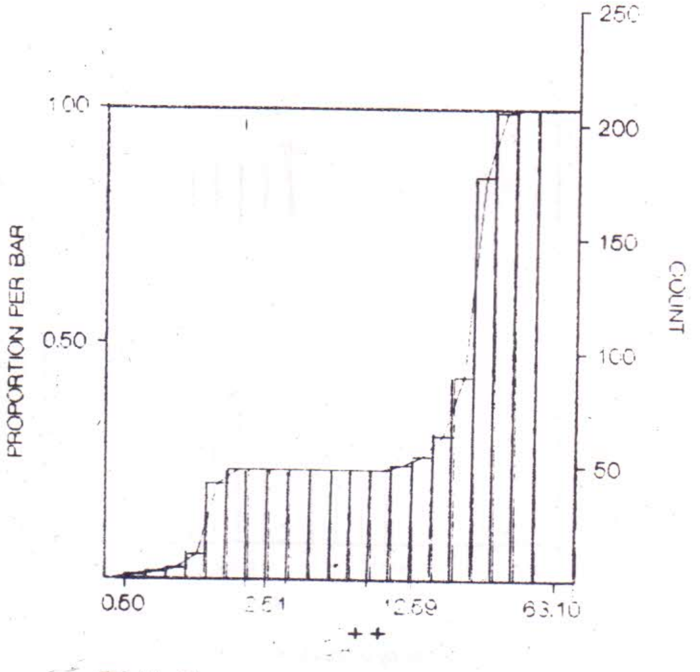


FIG 41 :Frequency distribution of Calcium ion

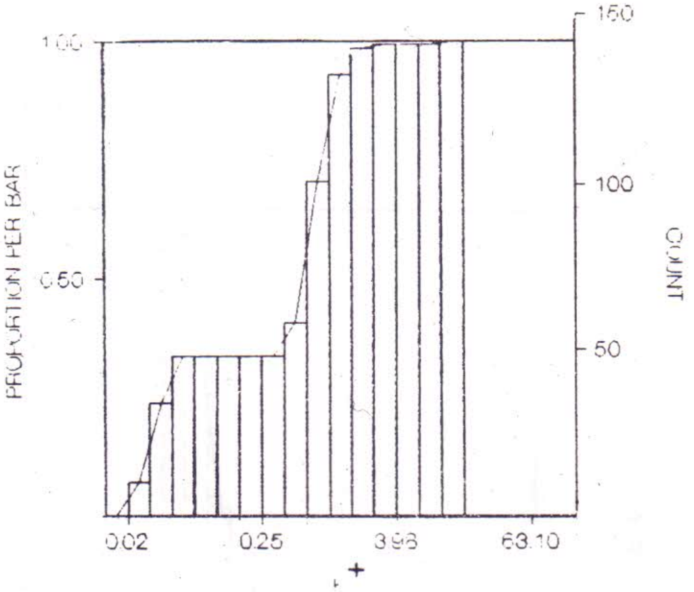


FIG. 40 :Frequency distribution of Pottacium ion

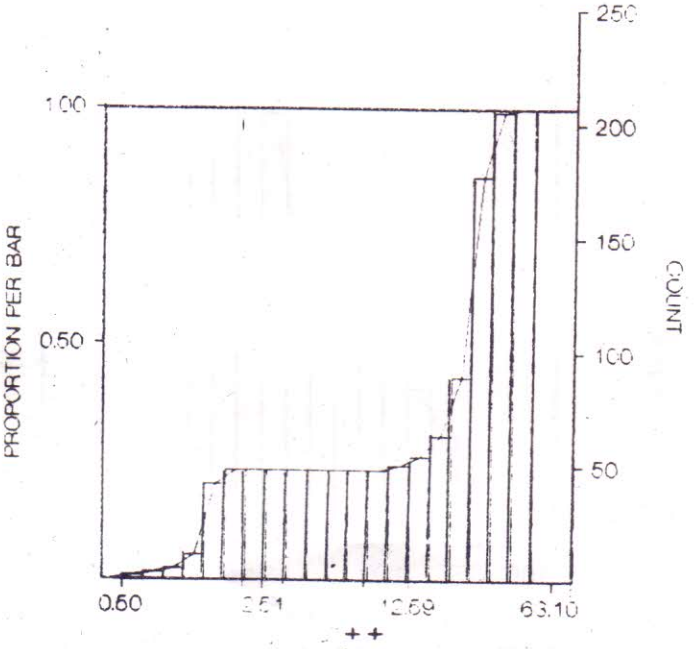


FIG 41 :Frequency distribution of Calcium ion

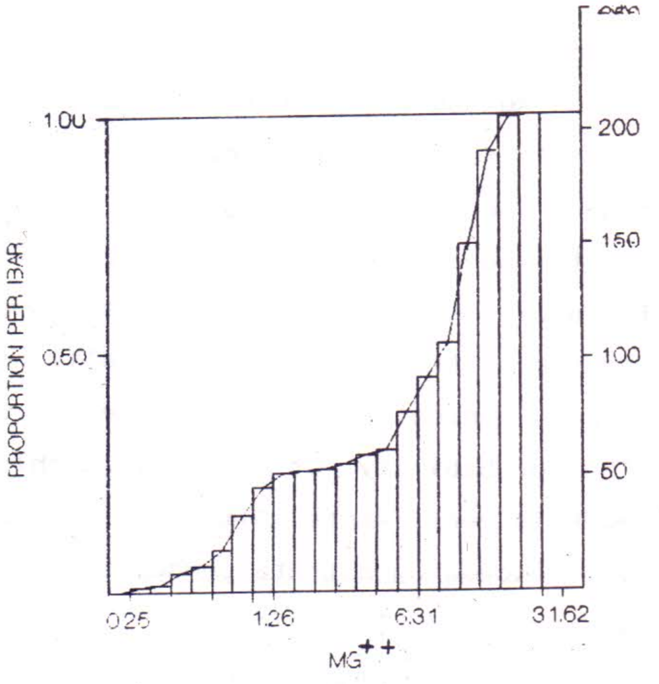


FIG 42 :Frequency distribution of Magnesium ion

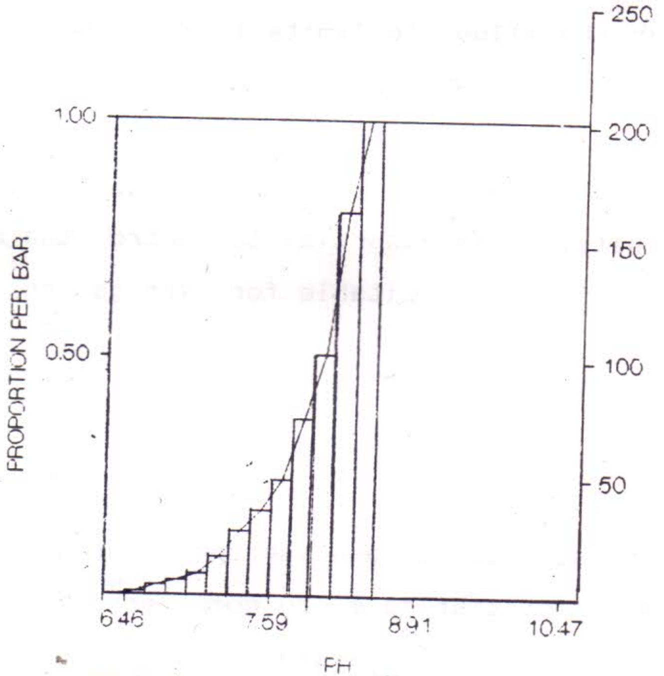


FIG.43 : Frequency distribution of p^H

most plants under different soil conditions. From Fig. 32, the total dissolved solids concentration remain 40 mg/L to 200 mg/L, which is suitable for most of beneficial uses of water. Hardness is also with in the limits (Fig. 33) of 250 mg/L, and the probability of exceeding hardness 250 mg/L is rare. However, for the specific uses like textile industries and power generation etc, some degree of treatment is essentially required. From Fig. 34, the chloride concentration is also well with in the limit of 250 mg/L. The probability of exceeding chloride concentration beyond 80 mg/L is very remote. Similarly Fig. 35 shows that the probability of exceeding nitrate concentration beyond 25 mg/L is very less whereas, the allowable limit according to BIS (Buero of Indian Standards) i9s 45 mg/L. The water quality parameters like So_4 , turbidity, HCo_3^-etc are also well within the prescribed limits of BSI and the probability of violation are very less. The concentration of various cations e.g. Na , K , Ca , Mg are also well within the allowable limits for various designated uses of water (Figs. 36 to 42). Fig. 43 shows that p^H is more than 7 for most of the times as it varies from 6.9 to 8.9 which can be considered safe at present but seeing the increasing trend towards the alkalinity, it is essential to control the alkalinity because p^H more than 9 is not suitable for irrigation and for most of other uses of water.

9.3 Regression Based Modelling

All the statistical analyses procedures are performed using the system for statistics micro computer software package (SYSTAT 1985). In order to generate proper and relevant multiple regressions, simple Pearson coefficients were calculated for the independent variables. If a strong correlation was exhibited among any two independent variables, then it would be statistically incorrect to use both of these variables in the same regression model. The results of the statistical analysis produced the following models:

$$\text{Hardness} = 57.81 - 0.73(Na^{+}) - 0.12(\text{Turbidity}) + 0.24(\text{Conductivity}) \quad (R^{2}=0.54)$$

$$Cl = 8.18 + 0.06(\text{Conductivity}) + 0.90(Ca^{++}) - 0.277(HCo_3^{-}) - 0.23(So_4^{--}) \quad (R^{2}=0.62)$$

$$HCo_3^{-} = 41.4 + 3.78(Ca^{++}) - 1.56(Cl^{-}) - 2.16(No_3^{-}) \quad (R^{2}=0.54)$$

$$K^{+} = 2.76 + 0.016(Ca^{++}) - 0.43(NH_4^{+}) - 0.29(p^{H}) + 0.014(PO_4^{--}) \quad (R^{2}=0.29)$$

$$Na^{+} = -6.15 + 0.15(No_3^{-}) + 0.03(HCo_3) + 0.056(TDS) \quad (R^{2}=0.42)$$

$$No_3^{-} = -0.35 - 2.22(NH_4^{+}) + 0.14(Na^{+}) + 2.21(Fe^{++}) \quad (R^{2}=0.43)$$

$$TDS = -21.93 + 6.67(p^{H}) + 0.50(\text{Conductivity}) - 2.92(PO_4^{--}) \quad (R^{2}=0.72)$$

$$\text{Conductivity} = 61.51 + 1.59(Cl^{-}) + 6.22(Mg^{++}) - 0.38(\text{Turbidity}) + 0.8(TDS) \quad (R^{2}=0.80)$$

$$\text{Turbidity} = 46.33 + 0.047(Q) - 0.33(HCo_3^{-}) - 9.28(NH_4^{+}) \quad (R^{2}=0.83)$$

The above models have been presented in tabular form (Tables: 4 t0 12) to giving various statistics of the models e.g. multiple R, squared multiple R, adjusted multiple R, and standard error of estimate etc. The model coefficients for the concerned model variables have been given with their respective t-value and standard errors. The analysis of variance is also presented giving F-value of the overall model.

Table 4: Detailed Model for Hardness

DEP VAR: HARD N: 51 MULTIPLE R: 0.733 SQUARE MULTIPLE R:0.538

ADJUSTED SQUARED MULTIPLE R: .508 STANDARD ERROR OF ESTIMATE: 21.270

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	57.813	10.939	0.000	--	5.285	0.000
NA	-0.727	0.694	-0.132	0.616	-1.047	0.300
TUR	-0.119	0.086	-0.155	0.790	-1.387	0.172
CON	0.242	0.044	0.727	0.570	5.538	0.000

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	24725.159	3	8241.720	18.217	0.000
RESIDUAL	21263.495	47	452.415		

Table 5: Detailed Model for Chloride

DEP VAR: CL N: 62 MULTIPLE R: 0.787 SQUARED MULTIPLE R: 0.619
ADJUSTED SQUARED MULTIPLE R: .592 STANDARD ERROR OFESTIMATE: 7.755

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	8.184	5.566	0.000	--	1.470	0.147
CON	0.060	0.018	0.435	0.403	3.378	0.001
CA	0.903	0.238	0.473	0.431	3.796	0.000
HCO3	-0.277	0.037	-0.751	0.655	-7.427	0.000
SO4	-0.232	0.107	-0.193	0.840	-2.168	0.034

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	5563.714	4	1390.928	23.129	0.000
RESIDUAL	3427.835	57	60.137		

Table 6: Detailed Model for Bicarbonate

```
DEP VAR: HCO3  N: 61   MULTIPLE  R: 0.736  SQUARED MULTIPLE  R: 0.542
ADJUSTED SQUARED MULTIPLE R: .518  STANDARD ERROR OF ESTIMATE: 22.929

VARIABLE  COEFFICIENT  STD ERROR  STD COEF TOLERANCE    T   P(2 TAIL)

CONSTANT     41.397       14.472    0.000       --      2.861   0.006
CA            3.779        0.535    0.734     0.744     7.061   0.000
CL           -1.561        0.276   -0.577     0.770    -5.649   0.000
NO3          -2.156        1.014   -0.195     0.958    -2.126   0.038


                      ANALYSIS OF VARIANCE

SOURCE         SUM-OF-SQUARES    DF  MEAN-SQUARE       F-RATIO       P

REGRESSION         35443.943      3    11814.648       22.472     0.000
RESIDUAL           29967.107     57      525.739
```

Table 7: Detailed Model for Pottacium

```
DEP VAR: K   N: 46    MULTIPLE   R: 0.541    SQUARED  MULTIPLE  R: 0.293
ADJUSTED SQUARED MULTIPLE R: .205  STANDARD ERROR OF ESTIMATE: 0.313

VARIABLE  COEFFICIENT  STD ERROR  STD COEF TOLERANCE    T   P(2 TAIL)

CONSTANT    2.759         1.052      0.000               2.622     0.012
TUR        -0.000         0.003     -0.026     0.414   -0.123     0.902
CA          0.016         0.009      0.304     0.573    1.731     0.091
NH4        -0.434         0.212     -0.280     0.946   -2.048     0.047
PH         -0.297         0.123     -0.400     0.643   -2.413     0.020
PO4         0.014         0.008      0.252     0.940    1.836     0.074


                      ANALYSIS OF VARIANCE

SOURCE         SUM-OF-SQUARES    DF  MEAN-SQUARE       F-RATIO       P

REGRESSION             1.627      5        0.325        3.314     0.013
RESIDUAL               3.928     40        0.098
```

Table 8: Detailed Model for Sodium

```
DEP VAR: NA   N:  55  MULTIPLE R: 0.645  SQUARED  MULTIPLE R: 0.416
ADJUSTED SQUARED MULTIPLE R:  .370 STANDARD ERROR OF ESTIMATE: 4.412

VARIABLE  COEFFICIENT  STD ERROR  STD COEF TOLERANCE    T    P(2 TAIL)

CONSTANT    -6.149       2.887      0.000      --     -2.130    0.038
NO3          0.154       0.219      0.086    0.775     0.702    0.486
HCO3         0.030       0.022      0.177    0.708     1.380    0.174
TDS          0.056       0.014      0.512    0.740     4.078    0.000
OXY          1.129       0.608      0.220    0.832     1.857    0.069

                       ANALYSIS OF VARIANCE

SOURCE          SUM-OF-SQUARES   DF  MEAN-SQUARE      F-RATIO       P

REGRESSION          694.564      4     173.641        8.921       0.000
RESIDUAL            973.230     50      19.465
```

Table 9: Detailed Model for Nitrate

```
DEP VAR: NO3  N: 46   MULTIPLE  R: 0.653  SQUARED MULTIPLE R: 0.426
ADJUSTED SQUARED MULTIPLE R: .385 STANDARD ERROR OF ESTIMATE: 2.627

VARIABLE  COEFFICIENT  STD ERROR  STD COEF TOLERANCE    T    P(2 TAIL)

CONSTANT     -0.347      0.927      0.000       --     -0.374    0.710
NH4          -2.223      1.743     -0.150     0.985   -1.275     0.209
NA            0.136      0.078      0.216     0.889    1.739     0.089
FE            2.211      0.402      0.677     0.900    5.497     0.000

                       ANALYSIS OF VARIANCE

SOURCE          SUM-OF-SQUARES   DF  MEAN-SQUARE      F-RATIO       P

REGRESSION          215.356      3      71.785        10.406      0.000
RESIDUAL            289.740     42       6.899
```

Table 12: Detailed Model for Total dissolved Solids

```
DEP VAR: TDS   N: 50   MULTIPLE R: 0.850   SQUARED  MULTIPLE R: 0.723
ADJUSTED SQUARED MULTIPLE R: .705 STANDARD ERROR OF ESTIMATE: 30.999

VARIABLE  COEFFICIENT  STD ERROR  STD COEF TOLERANCE    T    P(2 TAIL)

CONSTANT    -21.925      69.830     0.000      --     -0.314    0.755
PH            6.672       9.449     0.056     0.960    0.706    0.484
CON           0.503       0.053     0.760     0.948    9.534    0.000
PO4          -2.919       0.819    -0.279     0.985   -3.563    0.001

                      ANALYSIS OF VARIANCE

SOURCE          SUM-OF-SQUARES   DF  MEAN-SQUARE      F-RATIO       P

REGRESSION         115354.327     3    38451.442      40.015      0.000
RESIDUAL            44202.253    46      960.919
```

Table 11: Detailed Model for Conductivity

```
DEP VAR: CON    N: 49    MULTIPLE  R: 0.892   SQUARED MULTIPLE R: 0.798
ADJUSTED SQUARED MULTIPLE R: .778  STANDARD ERROR OF ESTIMATE: 39.637

VARIABLE  COEFFICIENT  STD ERROR  STD COEF TOLERANCE    T    P(2 TAIL)

CONSTANT    61.508       21.353     0.000              2.881    0.006
CL           1.594        0.473     0.256     0.804    3.373    0.002
MG           6.219        1.691     0.286     0.763    3.679    0.001
TUR         -0.379        0.155    -0.181     0.847   -2.447    0.018
TDS          0.803        0.121     0.541     0.696    6.631    0.000

                      ANALYSIS OF VARIANCE

SOURCE          SUM-OF-SQUARES   DF  MEAN-SQUARE      F-RATIO       P

REGRESSION         270642.851     4    67660.713      43.065      0.000
RESIDUAL            69129.108    44     1571.116
```

Table 12: Detailed Model for Turbidity

```
DEP VAR:   TUR    N: 46    MULTIPLE R: 0.909    SQUARED MULTIPLE R: 0.827
ADJUSTED SQUARED MULTIPLE R: .815  STANDARD ERROR OF ESTIMATE: 10.096

VARIABLE  COEFFICIENT   STD ERROR   STD COEF   TOLERANCE     T    P(2 TAIL)

CONSTANT    46.335        6.802       0.000        --      6.812     0.000
Q            0.047        0.006       0.556      0.732     7.415     0.000
HCO3        -0.333        0.053      -0.474      0.734    -6.329     0.000
NH4         -9.278        6.675      -0.090      0.992    -1.390     0.172


                         ANALYSIS OF VARIANCE

SOURCE         SUM-OF-SQUARES   DF  MEAN-SQUARE      F-RATIO       P

REGRESSION        20480.431      3     6826.810       66.978     0.000
RESIDUAL           4280.879     42      101.926
```

9.0 CONCLUSIONS

From the statistical analysis of water quality data observed at Manot site of the upper Narmada River basin following conclusions have been drawn:

(i) The quality-quantity relationships have been identified for various water quality determinands. It is found that the magnitude of both suspended sediment concentration and solute concentrations may be closely related to the magnitude of stream flow flow at a given site in the stream. The results have been presented in both graphically and analytically.

(ii) The cumulative probability distribution plots have been drawn for various water quality parameters highlighting the probability of non-exceedence for different level of concentrations of water quality parameters. It is found that at the Manot site, the Narmada water is safe for most of the beneficial purposes and hence at present there is no threat of problems associated with river pollution as the probability of violation of allowable water quality levels prescribed by the BIS for most of the water quality parameters are very very small. However, alkalinity is found to have an increasing trend as p^H varies from 6.9 to 8.9 and at p^H more than 9, the water looses its utility for most of the beneficial uses. Therefore, it is necessary to carry out the investigation to identify the sources of alkalinity so that the increasing trend of alkalinity can be controlled.

(iii) Multiple regression models have been developed for various water quality parameters explaining a significant portion of variability of water quality data. The developed model may be used for filling the missing data values which is an important step for further analyses such as time series analysis, stream quality forecasting, pollutant concentration frequency analysis etc.

REFERENCES

Biswas, A.K. (1974w). System approach to water management., McGraw-Hill, Inc., New York, USA.

Draper and Smith (1981). Applied regression analysis. Second edition, John Wiley and Sons, New York, N.Y.w

Douglas, I.,(1968). The effects of precipitation chemistry and catchment area lithology on the quality of river water in selected catchments in Eastern Australia. Earth Sci. J., 2, 126-44.

Durum, W.H., (1953).Relationship of the mineral constituents in solution to streamflow, Saline River near Russell, Kansas. Trans.,AFU, Vol. 34, No. 3.

Gibbs. R.,(1970).Mechanisms controlling world water chemistry. Science, 170, 1088-90.

Hall, F.R., (1970). Dissolved solids-discharge relationships. 1 Mixing Models. Water Resour. Res., 6, 845-850.

Hall, F.R., (1971). Dissolved solids-discharge relationships. 2 Application to field data. Water Resour. Res., 7, 591-601.

Johnson, N.M., Likens, G.E., Bormann, F.H., Fisher, D.W., and Pierce, R.S.(1969). A working model for the variation of stream water chemistry at the Hubbard Brook Experimental Forest, New Hampshire. Water Resour. Res., 5, 1353-1363.

Ledbetter, J.O., and Gloyna, E.F., (1964). Predictive techniques for water quality inorganics. J. Sanit. Engrg. Div., ASCE, Vol.

90, No. SA1, 127-151.

Lenz, A.T.,and Sawyer, C.N., (1944).Estimation of stream flow from alkalinity Determinations. Trans., AGU, Vol. 25.

Langbein, W.B., and Schumm, S.A., (1958). Yield of sediment in relation to mean annual precipitation. Trans. Am. Geophys. Union, 39, 1076-1084.

Schumm, S.A., (1954). The relation of drainage basin relief to sediment loss. Int. Assoc. Hydrol. Sci. Publ., 36, 216-9.

Striffler, W.D., (1965). Suspended sediment concentrations in a Michigan trout stream as related to watershed characteristics. Proceeding of the federal Inter-agency Sedimentation Conference, US Dept. of Agriculture, Miscellaneous Publication 970, 144-50. Washington DC.

Walling, D.E., (1974). Suspended sediments and solute yields from a small catchment prior to urbanization. In Geography, K.J., and Walling D.E. (eds). Fluvial Processesin Instrumented Watersheds, special PublicationNo. 6, 169-92, Institute of British Geographers, London.

Ward, J.C.,II (1958). Correlation of stream flow quality with quantity. thesis presented to the university of Oklahoma, at Norman, Okla.,in partial fulfilment of the requirements for the degree of Master of Science.

Webb, B.W., and Walling, D.E., (1974). Local variation in background water quality. Sci. Total Environ., 3, 141-53.

DIRECTOR : DR. S. M SETH

DIV HEAD : DR. K.K.S. BHATIA

STUDY GROUP : ADITYA TYAGI AND DR. K.K.S. BHATIA