A computer based study of the flora of highly acidic environments

Diaz, Bernard M.

A computer based study of the flora of highly acidic environments

by

Bernard M. Diaz (B.Sc. Dunelm)

A thesis submitted for the degree of Doctor of Philosophy in the University of Durham, England.

Department of Botany

December 1980

ABSTRACT

An account is given of the use of a computerised database approach to the study of the chemistry and botany of streams with a pH at and below 4.0. The SIEUR (Stream Information Entry, Update and Retrieval) system was designed and implemented on the Northumbrian Universities IBM System 370 computer running under the M.T.S. operating system. SIEUR has been used to store approximately 2500 water chemistry and 2700 biological records. Of these, 269 chemistries and 125 biologies were collected from 10 m reaches from four countries where the pH was at or below 4.0 on the day of sampling.

Principal component and cluster analyses of the acid chemistries suggested that they may be grouped together, based on the level of the various cations measured. This grouping follows closely a geographical breakdown of the sites and is probably a reflection of the mining associations of the majority of the sites. There was no similar grouping of the biological samples. Detailed examination of the distribution of 30 photosynthetic species which occurred live suggested that four patterns of reaction to low pH existed.

In the design and practical application of the SIEUR system to the investigation of the acid datasets several problems and some solutions were identified. The need to date, time and location stamp all data at all stages of the analyses was apparent. SIEUR provides an "expansion facility" to overcome the problem of chemical and biological sampling not necessarily occurring together on the same day. By the use of this facility the user can specify the leeway by which chemical and biological samples from the same site are linked to each other in terms of date of sampling. A major design consideration was the need to identify structure in the data and to be able to store and retrieve this as data. During the use of the system the need for comprehensive statistical and graphical facilities was apparent. The decision to provide interfaces to existing packages to do this, rather than provide integral facilities was fully justified in terms of the flexibility obained.

ACKNOWLEDGEMENTS

I have tremendously enjoyed doing this thesis and it is with pleasure that I acknowledge the many people who have helped me. I am of course expecially grateful to Dr B.A. Whitton, my research supervisor for his help and guidance throughout the work. Many colleagues were involved in discussion of all aspects of the project and gave valuable advice particularly Dr N.T.H. Holmes, Dr J.W. Hargreaves, Dr E.J.H Lloyd, Dr M. Hughes, Dr P.J. Say, Dr J.P.C. Harding, Dr A. Donaldson and Dr B.P. Jupp. I would also like to thank the many people involved in processing the data ready for the computer and in particular Mrs S. Harding, Mrs D. Kirby, Mr A. Waugh, Mr J.G. Grabham, Mrs M. Trinder, Mr P.J. Wright and Mr J.W. Simon. My thanks are also given to Mr W.B. Woodward and all the staff of the Science Library for their patience and help with all stages of the work. Special thanks also to Mr J.C. Carter for diatom identification and the Rev. G.G. Graham for help with the "bigger plants".

It is with pleasure that I acknowledge the advice and help of the various members of the University of Durham Computer Unit and particularly Mr R. Sheehan, Mr M. Munro, Dr A.M. Lamont, Mr A.A. Young, Mr B.R. Lander, Mr R. Gawley, Mr A.C. Heath, Dr N.F. Hall, Mr D. Lee, Dr P.S. Jones, Mrs M. Compton, Ms J. Rattenbury, Mr J.w. Steele and especially Mr G.R. Eadie who introduced me to the delights of graphics. Thanks are due also to the Durham Computer Operators Mr H. Baker, Mrs G. Orr, Miss A. West, Miss M. Crosby, Mr C. Iley and especially Mr S. Kay who wrote many of the graphical subroutines used in this thesis and Mrs H. Taylor and all the Durham Punch Operators for dealing so efficiently and accurately with all my "little" jobs. I would also like to thank Mr J.S. Roper (Department of Computing, University of Durham) for suggesting the relational approach and for help with the SIEUR Query Subsystem, Dr P.G. Barker (Teesside Polytechnic) for help and discussion of PL/1 problems and Dr B.K Aldred (IBM Research Centre, Peterlee) for discussion of database concepts.

Finally I would like to thank Prof. D. Boulter for the provision of research facilities within the Botany Department, the Natural Environmental Research Council and the University of Durham for research studentships and all Howlcrofters, especially Dr B.J. Cromarty for much moral support.

TABLE OF CONTENTS

LIST OF TABLES

CONVENTIONS AND ABBREVIATIONS

All tables (figures are refered to as tables) and equations are referenced by Section number and numerical sequence within sections. Tables are referenced - Table 3.14-5 and equations Equation 3.14:2. Each page of the thesis except multiple page tables, have a section and page number for ease in cross referencing.

Abbreviations (e.g. PCA) and acronyms (e.g. NUMAC) are expanded fully once when first used. The periods are usually not included and the letters are usually in capitals. Names in capitals refer to computer files; thus "TAXONINFO" is a data file and "SIEUR" is the file containing the system object code.

cond	conductivity
IBM	International Business Machines
MTS	Michigan Terminal System
nspec	number of species in a sample
NUMAC	Northumbrian Universities Multiaccess Computer
O.D.-420	optical density measured at 420 nm
P	Probability
RG	Report Generator
SIEUR	Stream Information Entry, Update and Retrieval system
SPIRES	Stanford Public Information Retrieval System
SD	Standard deviation
TAXIR	Taxonomic Information Retrieval system
km	kilometre
l	litre
m	metre
mg	miligramme
ml	millilitre
nm	nanometre
km	kilometre
>=	greater than or equal to
>	greater than
<=	less than or equal to
<	less than

Chapter 1 Introduction

1.1 Computing aspects

1.11 General Introduction

Computer technology is proving to be of increasing importance to biologists and legislators concerned with the ecology of aquatic environments. Cutbill (1971) noted, "there is an urgent need for an investigation of biological records - what size they are and how they are created, how they are structured, how they are linked together, how they are used and what their useful life is". These comments are still applicable; and although many databases now exist there are few complete studies which look at all the questions posed by Cutbill.

1.12 Computer technology

The introduction in 1960 of the IBM System/360 series of computers marked the beginning of an era of readily available fast serial processors. These had large instruction sets and the ability to handle simultaneously several peripheral devices and could address a very large memory (McKeag, 1972; IBM, 1980). The provision of this hardware resulted in considerable development in several areas of computer programming (Buckholz, 1963; Rosin, 1969; Denning, 1971). Of these developments the most spectacular was the invention of new programming languages (e.g. Rosen, 1964) and the concomitant algorithmic approach to computer application (Knuth, 1968). The availability of this hardware to commerce and industry resulted in new applications being developed which were concerned with the storage and retrieval of vast quantities of diverse types of data (IBM, 1960; Clippinger, 1961; General Electric Corp., 1961; Perstein, 1965). In the late 1960s and early 1970s these two separate strands of development began to come together to form the new technology of database management (Date, 1976).

1.13 Database technology

Despite the multitude of different origins and lines of development of database technology outlined by Fry et al. (1976), an underlying four stage pattern of evolution can be recognised (Martin, 1975). The first stage was characterised by the use of sequential files, usually on magnetic tape, where the logical view of the data stored was identical to the physical representation. The data were accessed by specific application programs which merely restructured the data records into an output format. The second stage, found predominantly in the late 1960s, was characterised by the use of sequential files to store bulk data and the use of direct access files on magnetic disks to store indices and keys. It was by the use of these keys and indices that the sequential bulk data were accessed. The database system was usually designed and optimised for one particular application. These early database systems are reviewed by Minker et al. (1967), CODASYL (1969) and Fry et al. (1973).

The late 1960s and early 1970s saw the ideas of McGee (1959), McGee et al. (1960) and Childs (1968) concerning "generalised data handling routines" and "a set theoretic approach to data" having an increasing influence on the evolution of database technology. The systems which were implemented using these techniques were independent of the data they contained and could be considered to be "database management systems" capable of handling any dataset. Thus stage three was characterised by the use of multiple files derived from the same physical data file (see Miller et al., 1960). These separate visualisations of the data were managed by data handling routines which removed redundant or replicated data and which ensured data integrity. Data retrieval was possible within these systems by the use of multiple keys and complex data reorganisations could be made without the need for extensive revisions of application programs. The fourth stage described also by Witney (1973) was marked by the development of a global view of the data, which was independent of the physical data organisation and the users view of it. This type of database could evolve by the addition of new elements to the global view of the data, and this addition did not adversely affect the maintenance of the database. Rapid search and optimisation occurred automatically within the database management system by the use of inverted files. Furthermore, the system software provided a data description language for a database administrator (see D'imperio, 1969; Canning, 1972; Everest, 1973; Taylor, 1974; Senko, 1975), a command language for the applications programmer (Dodd, 1969) and a query language for the various database users (see for example Hardgrave, 1974; McDonald et al., 1975 and Chamberlin et al., 1976).

The primary objectives of database design have been the concern of many reports (CODASYL, 1971; DBTG, 1971; Dean, 1971; Engles, 1971; Date et al. 1971; ANSI, 1975). Summarised briefly they are high performance in computing terms for both setting up and using the database; low cost of storing and using the data and making changes; accuracy and consistency especially while updates are made; guaranteed data integrity, privacy and security; mechanisms which ensure that the database can evolve easily and can be independent of the data structures allowed by the computer hardware or the original data structures; systems which ensure that unanticipated queries can be handled easily and efficiently such that the data can have multiple uses; and the implementation of a clear and simple logical data structure which can be used easily and effectively by both the expert and novice user. In addition the system needs to be "tunable" such that query types performed frequently can be optimised. Finally many authors (e.g. Martin, 1975; Date, 1976) consider it important that the database management system should be able to deal with data migration, such that infrequently used data can be placed on cheaper storage devices and more frequently used data moved in the opposite direction.

The data independence shown by modern database management systems stem from recognition of underlying data structures (Bachman, 1972). Three basic structures have been identified. These are hierarchical structures (see Bleier, 1967; IBM (IMS), 1975; Tsichritzis et al., 1976), network structures (see Earnest, 1974; Codd et al., 1974; Taylor et al., 1976) or relational structures (see Codd, 1970; Date, 1972; Hitchcock, 1974; Codd, 1974; Held, 1975; Chamberlin, 1976). Detailed comparisons between these methods of structuring the data have been made by Date et al. (1974), Nijssen (1974), McGee (1974), Sibley (1974), Held et al. (1975) and Michaels et al. (1976). In recent years the relational approach has received considerable attention because of its simple, tabular conceptualisation of the data. Furthermore, the method is susceptible to mathematical analysis (see Kent, 1973; Fadous, 1975; Hall et al., 1975). Much work has been done in identifying basic database problems and proposing solutions using the combined relational approach and mathematical treatment (e.g. Codd, 1971; Heath, 1971; Armstrong, 1971; Bernstein et al. 1975). Several implementations of relational database structures exist, of which INGRES (McDonald et al., 1974) is perhaps the most easily accessible example. Other systems are described by Goldstein et al. (1970), Notley (1972), Lorie (1974), Czarnik et al. (1975), Winslow (1975) and Astrahan et al. (1976). In addition, recent research has been undertaken into giving mini and micro-computers some relational database capabilities (Mcleod et al., 1975; Manacher, 1975). The relational approach to data storage has given impetus to research into user query languages. A query algebra has been proposed by Codd (1971) for relational and other databases where queries are expressed in terms of set theoretic operations (Childs, 1968) on the data relations. This approach requires skill and knowledge on the part of the user and while it is simpler to implement (however, see Smith (1975) for comments on optimisation) it has received only passing attention. In contrast the query calculus (Codd, 1971) has received considerable attention and has led to the concept of "structured English query languages" (Chamberlin et al., 1974). In these approaches the user specifies what he wishes to see in a final "goal relation", using boolean constructs, and the system performs the necessary operations to satisfy the query.

Within the data themselves, there may be considerable structure. Martin (1975) notes four types:

1. tree structures

2. plex structures or networks with a directional hierarchy

3. plex structures which are bidirectional with no hierarchy

4. loops (single level cycles) where the hierarchy is looped.

At present few systems exist which are geared to recognise these structures as "items of data" (Hall, 1975; Aldred 1975). Some of the more detailed hydrological studies have examined water networks and have proposes systems for encoding, storing and processing these structure data (Coffman, 1971).

Many proposals exist for "database hardware". Canaday et al. (1974) have suggested a dedicated "back-end" computer which would deal exclusively with database operations. Heacock et al. (1975) have suggested dedicated "front-end" computers and Su et al. (1973) and Lin et al. (1976) have suggested dedicated data handling facilities independent of, but linked to, the main processor. Much work has been done on the development of mass storage devices (Becker, 1966; Pickering, 1971; Houston, 1973) and associative memories (Miuker, 1971; Crick et al., 1970) and means of implementing these (Moulder, 1973; Feldman et al., 1969; Ash, 1969; Symonds, 1969; Crick et al., 1970). Geographically distributed databases (Marill et al., 1975) have also received much attention especially as a means of linking government and research data centres.

1.14 Biological and aquatic databases

Edwards (1971) makes the distinction in her paper between "biological information", meaning textual received data, in the form of abstracts, bibliographies, reports and reviews (see also Hersey et al., 1968; Steere, 1970 and Frost et al., 1974) and "biological data" which are the actual results of experimentation or analysis. Biological information databases are many and varied and usually well known and documented (see Edwards, 1971b for summary). In contrast biological data databases are few and inadequately documented (Cutbill, 1971). Several museums, herbaria and culture collections have database systems which help catalogue their holdings (Beschel et al. 1970; Shetler, 1971; Crovello, 1972; Cutbill, 1973; Gomez-Pompa et al., 1973; Morris et al., 1975). These systems are sometimes extended to cover related aspects such as taxonomy, systematics in general, paleontology, geology, patents, palynology, biodeterioration and often relate to biological information services as well (Cutbill, 1971b). In Britain the research councils, especially the Natural Environmental Research Council, have initiated several biological data database projects (Perring, 1971) and several universities now have their own ad hoc systems. Crovello (1972) and Shetler(1971) report that in the USA many data and information databases exist at national, state and private organisation level and in general there is a movement by international bodies such as the United Nations and European Economic Community to provide both research personel and the public with access to biological and information databases (e.g. Geiss, 1973; Persoone 1979).

For the ecologist interested in aquatic environments the literature on databases may be found in two types of report. Pilot study reports like those of Jonsson (1970), Marelius (1971), Pignatti (1976) and the British Water Data Unit (1976a, 1976b) which concern themselves with standardisation of data recording procedures and the practical aspects of the use of the databases in question. In contrast research study reports like those of Klasvik (1974) and Cairns et al. (1972) deal exclusively with the biological interpretation of data and only rarely mention details of the database design or statistics concerning database performance.

1.2 Biological aspects

1.21 Environments with low pH

1.211 Acid sources

Three main sources of acidity have been reported in the literature. Natural organic acidity is due mainly to the presence of humic acids produced by plant decomposition. Pearsall (1949) in an examination of English moorland concluded that the pH of the typical "moss" was in the region 3.80 - 4.00. He noted that when oxidation occurs the pH can fall to 3.45 and that for Eriophorum moorland under considerable oxidation, pH levels as low as 2.80 can be attained (Pearsall, 1930). Brock (1969) notes that where the pH is below 3.00, an inorganic source can usually be found to explain the observation.

The second source of acidity is due to inorganic acids produced naturally in thermal springs and volcanic lakes. The pH here is often below 3.0. The acidity is due to the oxidation of H_2S and/or SO_2 present in the volcanic gases. Such oxidation can lead to pH levels around 1.0, as reported by Satake and Saijo (1974).

A third source of acidity is due to industrial effluent which may also produce extremely low pH levels, below 1.0. Several sources for this acidity have been recognised (Klein, 1959) but the most common is the production of sulphuric acid from the oxidation of sulphurous material associated with coal. Acid may also be produced where mining activities have exposed iron disulphides usually in the form of pyrites, marcasite or pyrohottite to the effects of air, water and possibly bacteria. Such conditions are found associated with lignite, pyrite, baryte, zinc, copper, gold, silver and lead mines as well as coal mines (Temple and Koehler, 1954). The countries from which such acid sources have been reported include North America (e.g. Braley 1951-1954; Parsons, 1952-1975; Kinney, 1964, Hanna et al., 1965; Boyer, 1972), Australia (Blessing et al., 1974), New Zealand (Kaplan, 1956), South Africa (Harrison et al., 1958-1962) and several European countries including Britain (Glover, 1967; Hargreaves et al., 1975), Denmark (Dahl, 1963) and Czechoslovakia (Fott, 1956).

1.212 Acid production and hydrology

The oxidation of pyritic minerals occurs slowly in nature when outcropping coal seams become exposed by erosion. However, it is mining activity which accounts for the majority of such acid formation. Once this acid source has begun it does not necessarily decrease with the cessation of mining but may continue to run for many years. Brand and Moulton (1960) have identified three steps in the production of acid from pyrites. The first step is an oxidation to $FeSO_4$ and SO_2 in the absence of water, or $FeSO_4$ and H_2SO_4 if water is present. In step 2, $FeSO_4$ and H_2SO_4 oxidise to give $Fe_2(SO_4)_3$. The third step occurs only in an acid environment and is the precipitation of the orange yellow $Fe(OH)_3$ and the production of further H_2SO_4. Clarke (1967) has shown that the kinetics of these reactions requires aerobic conditions. The oxidation of marcasite and pyrohottite (Fe_nS_{n+1}) follow essentially similar reactions but at significantly different rates. Braley (1954) found that pyrohottite reacted at 18 times and marcasite 9 times the rate of pyrite. It has been demonstrated by Blessing et al. (1974) that Zn, Cd, Cu, Ni and Co sulphides are also attacked by $Fe_2(SO_4)_3$ although this occurs at low rates. At low pH other reactions occur which render K, Ca, Al, Mn and Si mobile. On raising the pH the characteristic coloured precipitate of these as well as ferric sulphate may be noted. The grey white precipitate sometimes noted at pH levels between 3.0 and 6.0 is due mainly to Al and colloidal Si (Blessing et al., 1974). It has been noted by Powell and Parr (1919) and many others that bacteria are often associated with acidic environments. Much work has been done on the acid production caused by these chemautotrophic bacteria (e.g. Colmer et al., 1947; Temple et al., 1951; Ashmead, 1955; Ehrlich, 1962; Leathern et al., 1956; Baker et al., 1970; Singer et al, 1970; Lundgren, 1971; Walsh et al., 1972; Manning et al., 1972). However, the actual role of these bacteria is still uncertain, although it is clear that the formation of acid is not entirely dependent on their presence.

The hydrological aspects of acid drainage have been considered by Brant et al. (1960), Collier et al. (1955) and Smith et al. (1971). Ground water is important both in the formation of acid and its dispersal. The flushing action of changes in water table caused by seasonal and other changes has been recognised as being an important criterion in the formation of acid and the rate at which it is formed (Braley, 1954).

1.213 Chemical and physical characters of acid mine drainage

Hawley (1971) described the typical acid mine drainage as characterised by low pH, high Fe and SO_4 concentrations and unusually high levels of heavy metals including Mn, Cu, Co, Zn and Ni, the relative levels of the various parameters depending closely on the geology. Lundgren et al. (1971) identify 4 classes of acid lake depending on the pH, acidity, Fe and turbidity level. While useful, no guide is given as to what levels of nutrients or heavy metals are to be expected. Parsons (1964) concluded that classification of strip-mine lakes could only be done on the basis of the physical and chemical parameters and that this was dependent on the level of progressive oxidation and iron oxide precipitation present. Van Everdingen (1969) reports extensively on the acid springs in Kootenay National Park, British Columbia and presents a range of sites from pH 2.5 to 5.5, with measurements on 11 cations, 6 anions plus CO_2 and O_2 concentrations and the physical parameters pH, Eh, temperature and conductivity. However, no information concerning P or N is present. The nutrient status of acid water is poorly reported, although Roback et al. (1969) and Bennett (1969) report the PO_4-P and NO_3-N levels at several sites.

1.22 Biological effect of low pH

1.221 Introduction

It is generally believed (Bennett, 1969; Besch et al., 1972) that pH (or perhaps acidity) rather than levels of other parameters such as heavy metals controls the presence of species at acid sites. However, Besch et al. (1972), Hargreaves et al. (1976,1977) and Wilkinson et al. (1980) suggest that heavy metals may play an important secondary role.

1.222 Terrestrial organisms

There is not an extensive literature on the effect of low pH on terrestrial organisms. Kinney (1964) notes the dramatic effect on population number of a sudden fall in pH caused by opening a new coal mine in USA but is unable to ascribe these effects wholly to the acid. Mammals, birds and reptiles which graze the biota of acid waters are usually the first to suffer acid pollution. Boccardy et al. (1958) studied the effect of acid water due to mining disturbance and concluded that removal of cover, and other factors were as important as the effect on the food chain.

1.223 Fish

There is an extensive literature on the effect of low pH on fish which has been summarised by Douderoff and Katz (1950) and more recently by Lloyd (1968). The lower pH tolerance limit is about 5.0 although some species can exist in waters at pH 4.2. Other factors such as CO_2 and hardness are also important in determining the lower pH limit for fish. The reproductive ability of species which can survive at pH levels below 4.0 is usually diminished. Furthermore, the egg hatching rate at pH levels below 5.0 may be severely affected.

The pathological effect of low pH is not well understood. Ellis (1937) suggested that suffocation occurred due to the precipitation of mucus onto the gills. Lloyd et al. (1964) found no mucus or apparent gill tissue damage and attributed death to acidemia. The combined effects of low pH and other parameters, notably the heavy metals, have not been well studied. The effects of Fe on fish are not well understood although it would appear that large ammounts of Fe affect fish because of precipitation on the benthic community on which they feed.

1.224 Benthic fauna

Harrison et al. (1958, 1960 and 1965) report that the benthic fauna was severely restricted under the acid conditions found in several South and West African streams associated with gold and coal mining. These workers found that the acid waters were usually dominated by chironomids and caddis species especially *Leptocerus harrisonii* and *Argyrobothius* sp. Often these species were more abundant in acid waters than neutral waters due, they suggest, to the absence of competition and predation. Jewell (1922) and Lackey (1938, 1939) noted that the fauna of waters with pH in the range 2.2 to 3.9 and reported 12 species. This conclusion was supported by Warner (1968) who noted in addition that the Chironomidae were frequently dominant. Roback et al. (1969) reported that species of Odonata, Ephemeroptera and Plecoptera were eliminated by low pH. Other major studies have been conducted by Harp et al. (1967), Stockinger et al. (1960), Dinsmore (1968) Patrick et al. (1974), and Henricks et al. (1972).

1.225 Zooplankton

The literature concerning the effect of low pH on zooplankton is especially sparse. Lackey (1938) noted that several macroinvertebrates including Distyla sp., Actionophyrys sol and Brachionus urceolaris were dominant in lakes with pH between 2.96 and 3.30. B. urceolaris was also present in some Japanese lakes as reported by Ueno (1955). Other genera noted were Cyclops, Daphnia, Diatomus and Scapholbeeris. Parsons (1968) noted the presence of the rotifers Brachionus urceolaris, B. havanaensis and Keratella quadrata in several strip mine lakes in the USA. He also noted that where the acid condition was intermittant the acid species dominated. Repopulation after cessation of acid flow was found to depend on the length of the life cycle of potential species.

1.226 Macrophytes

Lackey (1938) reported that Typha latifolia was the most abundant macrophyte species in acid streams, in conjunction with Isoetes spp. Patrick et al. (1974) reported Isoetes below 4.0 although Vallisneria sp. was the most abundant plant recorded. Heather (1951) reported Eleocharis palustris, Typha latifolia and Carex sp. in lakes at pH 2.3 - 3.8. Above pH 5.0 he noted 7 species. These species have now been confirmed by several authors including Ehrle (1960), Moore et al (1967) and Sand-Jensen et al. (1978).

Harrison (1958,1965) reported a slightly different flora in South and West African acid streams. This flora included Typha latifolia, Phragmites australis and Sphagnum truncatum at pH levels around 2.9. The species which he found at higher levels (3.7 - 4.5) included Scirpus fluitans, Juncus exsertus and J. oxycarpus. He concluded that strong acid pollution eliminated some species but encouraged acidophiles to colonise the extreme environments probably owing to the lack of competition.

1.227 Algae

Algae in acidic environments have been poorly studied. Lackey (1938) was among the first to note and describe in any detail the restricted flora of acid streams. Hargreaves et al. (1975) report that the algal flora of English acid waters below pH 3.0 consisted of 24 species. Comparison with algal lists compiled by Lackey (1938), Steinback (1966), Warner (1969) and Bennett (1969) for the USA, showed that of the 40 species listed for the two countries only four were certainly common to both. It was difficult to assess to what extent these differences were due to taxonomic problems. Bennett (1969) investigated some 17 stations (pH 2.69 - 7.0) and reported 107 species below 4.1. The abundant species he noted were Euglena mutabilis, Ulothrix subtilis, Pinnularia braunii, Eunotia tenella (which is possibly E. exigua - J.R. Carter pers. comm), Ulothrix sp., Frustulia rhomboides and Penium generi. Bennett also noted a direct relationship between species number and pH level. Weaver and Nash (1968) report 20 species from 6 stations on one stream at pH levels from 3.0 - 4.0 in Kentucky. The flora was dominated by filamentous algae and Euglena spp. Warner (1968) in a comparison between streams at pH between 2.8 - 3.8 and 4.5 found 19 taxa at the upper level, while at the lower level 33 species were identified. The most abundant species were Ulothrix tenerrima, Pinnularia termitina, Eunotia exigua and Euglena mutabilis. In South Africa Harrison et al. (1962) noted filamentous algae and diatoms as co-dominants. The diatoms occurring below 5.0 were Eunotia exigua, Frustulia spp., Pinnularia acoricola, P. subcapitata, Achnanthes microcephala and A. minutissima. These diatoms were considered by Cholnoky (1958) to be characteristic of acidic environments. Negoro (1944) reported several species from Japanese streams and lakes at pH levels below 4.0 and includes some excellent studies concerning the chemistry of these waters. Extremely low pH waters have been reported by Satake (1974) and Fott et al. (1964) and were found to have Pinnularia braunii and Chlamydomonas spp. (especially C. acidóphila). Satake also reported the presence of Cyanidium caldarium in a Japanese lake at pH 1.7. The presence of this species at Yellowstone National Park, USA has been reported by Brock (1978), and would appear to be associated with the thermal environment below pH 4.0.

1.23 Physiological studies and tolerance mechanisms

1.231 Low pH studies

There are few studies aimed specifically at understanding the physiological effect of low pH. Fott et al. (1964) and McCarthy et al. (1965) and Hargreaves (1975) have studied the nutrient requirements of the pH tolerant species Chlamydomonas acidophila and concluded that at reasonable levels of light and temperature most species can withstand a wide range of nutrients at pH 2.0. Work that has been done by these authors on Carteria acidophila has resulted in similar conclusions. In contrast C. turfosa was found to have an absolute requirement for some vitamins and was unable to tolerate as low a pH. Much work has also been done on Cyanidium caldarium by Doemal et al. (1971) and Ascione et al. (1966). Several species of Euglena have also been studied by Jahn (1931), Schoenborn (1950) and Moss (1973). Kersler (1965) reported 7 species of Chlorella that were able to withstand a pH range of 2.0 - 3.0. Oborn (1960), Hutner et al. (1950) and Cassins (1974) have studied the effect of varying Fe in natural waters but little is known of the effect of pH on its availability. Foy et al. (1972) examined the response of Chlorella pyrenoides to Al and low pH and showed an increased requirement below pH 4.6. They also reported increased tolerance to high levels of Al using stress techniques.

1.232 Heavy metals and low pH

Apart from the work by Besch et al. (1971) and Hargreaves (1975) little detailed work has been done on the combined effect of low pH and heavy metals. Both these authors concluded that acidity is the primary factor governing species tolerance although heavy metals may be playing an important secondary role at the pH limits for a particular species. A more detailed account of the general conclusions drawn by Hargreaves is presented in 1.3.

1.233 Mechanisms involved in pH tolerance

Little work has been done on the mechanisms involved in tolerance to low pH. Cassins (1974) has proposed that specialised membranes must be involved in acid resistant algae. Hargreaves (1977) suggested that an active process was in operation in algae which was similar to that proposed by Manning et al. (1972) for bacteria. Clymo (1968) showed that an ion exchange system existed in Sphagnum by which H^+ ions were released in exchange for other metal ions. He also noted that this release was proportional to the the amount of polyuronic acid present in the cell wall.

1.3 Thesis aims

The main aim of this thesis is to show how database technology can be used in the analysis of ecological questions concerning the aquatic environment. This will be achieved by examination of the experience gained in designing a database system and in performing a specific analysis using that system. A secondary aim is to investigate the statistical and other techniques which may be used in such analyses. This study was initiated because of the absence of a detailed examination of the database technology available to the biologist interested in problems concerning aquatic ecology and because of the authors twin interests in botany and computing. The choice of the acid dataset for detailed examination was governed by the availability of the data and the authors practical knowledge of the species and sites involved and his participation in the data collection and analysis.

The questions which the database approach to the acid environment will hope to examine stem from conclusions drawn by Hargreaves (1977):

1. the number of species at a site is directly proportional to the pH (p. 317) Lowering the pH results in fewer species although the biomass may not change materially (p. 316)

2. factors additional to pH have an important effect on the distribution of species especially near the pH limit for those species. These factors are both chemical and physical - for example current speed and substratum stability (p. 318)

3. at a global level the species of acid environments are geographically distributed (p. 316)

4. although species composition changes very little in absolute terms seasonal variations occur which are very similar from site to site (p. 316).

In addition it is hoped to show that the database approach will allow pattern in the distribution of the data to be discerned and classification of low pH environments and photosynthetic organisms to be made.

Chapter 2 Methods

2.1 Computing system

2.11 General introduction and background information

2.111 Introduction

In order to examine the use of database techniques the decision was made to design and implement such a system based on "The Durham Recording System" (Whitton et al., 1976). Two problems were immediately apparent, choice of computer hardware and software with which to make the implementation and system design criteria. Secion 2.11 provides an overview of the computing background and the Northumbrian Universities Multi-Access Computers which were used, while Section 2.12 provides an insight into the design adopted.

2.112 Hardware available : N.U.M.A.C. system

The Northumbrian Universities Multi-Access computers (NUMAC) are an IBM 370/168 and an IBM 360/67 located in the Claremont Tower, Newcastle University. These are connected to "front end processing" satellites which are usually Decsystem PDP 11/20 minicomputers. The various FEPs handle batch and terminal communications between the central site (Newcastle) and the Remote Job Entry stations. The Durham RJE station is connected to NUMAC via a leased Post Office telephone line. Terminal traffic to NUMAC is handled by a PDP 11/21 located in Durham. Batch traffic to both the 360 and 370 computers is handled by an IBM 1130 computer located in Durham.

Other hardware used were the Durham Computer Unit PDP 11/34 running under the UNIX operating system (for INGRES work and file transfer to NUMAC), the Cambridge IBM 370 (for some statistical packages not available on NUMAC) and the FR-80 laser plotter available via the Rutherford IBM 360/195 which was used for some graphical applications.

2.113 M.T.S. (Michigan Terminal System)

MTS is the preferred operating system for the NUMAC 370 and is the one on which most of the computing for this thesis was performed. MTS (Boetner & Alexander, 1976) is a terminal orientated time sharing system that offers both batch and terminal facilities simultaneously on computers having the IBM 360/370 architecture. Among its attractive features are virtual memory handling (Denning, 1971), and multiple central processor support (Alexander,

Table 2.11-1

Types and characteristics of peripheral devices supported by NUMAC

Item or Peripheral	Type	Medium	Number	Location
Main processor	IBM 370/168	-	1	Newcastle
	IBM 360/67	-	1	Newcastle
Channel interfaces	IBM 2870	-	2	Newcastle
	IBM 2860	-	2	Newcastle
	IBM 2880	-	2	Newcastle
Communications modem	Memorex 1270	-	1	Newcastle
Paging drum	IBM 2820	-	1	Newcastle
Fixed head disk	IBM 2835	-	1	Newcastle
Disk drive	IBM 2314	IBM 2314	-	Newcastle
	IBM 3830	IBM 3330	8	Newcastle
M/tape drive	IBM 3803	any 9 track	4	Newcastle
P/tape reader punch	IBM 1134	any 9 track	1	Newcastle / Durham Newcastle / Durham
Printers	IBM 1403	most forms	several	Newcastle / Durham
	CTL 1.36	most forms	1	Durham
Card reader	IBM 2540	80 column	2	Newcastle
punch	IBM 2501	80 column	several	Newcastle / Durham
Plot hardcopy	IBM 1627	11 inch paper	1	Durham
	Calcomp 563	30 inch paper	1	Durham
	Cilplot 2	30 inch paper	1	Durham (Geography)
screen	e.g. Tektronix 4013	-	several	Newcastle / Durham
Vdus	e.g. Newbury 724	-	several	Newcastle / Durham

(Much of the communications equipment is omitted from this table - however mention must be made of the special channel interface referred to as the the "NUMAC CA00" which connect to MTS the various Front End Processing Decsystem PDP 11/20s which deals with the majority of Durham peripherals.)

1972).

The user sees MTS as two subsystems, the terminal or multi-access subsystem in which the user has facilities for communicating with running programs and for editing his files on disc; and the other is the batch processing subsystem which off-lines its input and output (McKeag, 1972). Both subsystems use the same command language and a user may submit jobs from his terminal for execution by the batch subsystem. Both batch and terminal subsystems are able to access the full range of peripheral devices through Device Support Routines (DSRs). Each device has its own DSR which the system accesses via a common Input/Output (I/O) interface, and an unique four character device name. The user may access the devices through their Pseudo-Device Name (PDN), some of which are predefined, for example the terminal device in terminal mode has the PDN "*MSOURCE*" on input (master source) and the PDN "*MSINK*" on output (master sink), other PDNs may be user defined for example to deal with magnetic tapes.

2.114 MTS file structure

MTS has three file categories. Public files are maintained by NUMAC staff and are accessible to all system users. Examples of files in the Public domain are *PL1 which contains the PL/1 compiler and the *SORT utility program. Private files are the users responsibility and must be permitted explicitly to other users. The third file category is the temporary or scratch file. These are identical to private files except that they exist only for the duration of a terminal or batch session.

There are two types of file, sequential and line files. Sequential files have logical records of information arranged in a strictly sequential fashion. Line files, consist of logical records or lines and an associated line number. Line files may be accessed by their line number or sequentially. Because of the indexed nature of line files they are used extensively within the SIEUR system. At the MTS level the basic unit of a line file is the physical record which is 4096 bytes (1 page of storage) long. A line file contains two logical components, namely the line directory and the data section. The line directory consists of fixed length entries, for every line in the file ordered by line number. In addition there are entries for each available 'hole' in the data section. These entries contain the line number and indicate where the line is in terms of the logical page number within the file (maximum 32767 pages), and the displacement within the page. The data section contains the logical line preceded by the line length, or holes where lines that have been deleted were. Long lines are broken into pieces and stored with additional information which allow all the pieces to be found. Lines up to 32767 bytes long may be stored in the 32767 pages which a file may possess, however, in practice the size limit for line files is approximately 24000 pages of information or approximately 36 million characters (Pirkola, 1975), which is the limit for IBM compatible 3330 type disks.

2.115 MTS facilities

The facilities (programs) which MTS offers are divided into three parts, Command Load Subsystems (CLSs) (MTS volume 1, 1978), public files (MTS volume 2, 1977) and subroutines (MTS volume 3, 1976). The two most useful CLSs are the system loader and the file editor. The system loader permits the user to "run" his program and to monitor or trace what occurs. The file editor is a powerful text editor incorporating a Snobol like pattern matching facility (Grisewold et al., 1971). Extensive use was made of this editor for data preparation and program development. The debug CLS was used extensively to test and to monitor statistically program performance.

Many public files were used extensively in the development of the computing system. The IBM PL/1 F-level compiler as implemented in the *PL1 public file (MTS volume 7, 1977) was used to compile the code. The *SORT program and subroutines and other public utility subroutines (MTS volume 5, 1976) were also used extensively.

2.116 Choice of programming language

PL/1 (IBM system documentation, 1971) was chosen as the preferred language for the computing system for the following reasons.

1. it can handle a wide variety of data types, and has reasonable string handling capabilities. SNOBOL4 (and its MTS variant SPITBOL, MTS volume 9, 1975) were considered too cumbersome and inefficient for numeric aspects of the programming, and too inflexible especially in their inability to call external routines written in other languages
2. it is a block structured language with many similarities to ALGOL (MTS volume 16, 1978). FORTRAN (MTS volume 6, 1978), 360/370 Assembler (MTS volume 14, 1978) and BASIC (MTS volume 10, 1974) were discarded as inappropriate
3. it has flexible storage handling facilities built into the language
4. it has sophisticated input/output handling facilities quite unlike, and far superior to, most other languages
5. it has several built in interrupt facilities for exception handling. In addition it is possible to write code for user written exception handling
6. it has many useful and powerful built in functions for mathematical and string handling purposes
7. it has good debugging facilities that result in high programmer productivity
8. it is reasonably well documented, well supported and easy to use.

The disadvantages of PL/1 stem from the complexity of the language and its consequent size, and the inefficency of the IBM F-level compiler see 7.116. In order to overcome these, and to increase the speed of certain frequently used but inefficient program code, Assembler subroutines were written to complement the PL/1 subroutines.

2.117 Statistical packages

Comparisons of various statistical packages have been performed before. Where possible any statistical procedure has been repeated using more than one statistical package. Where differences in the results are encountered these are discussed (Section 3.43). The statistical packages used were SPSS version 8 (Nie et al., 1975), MIDAS (Fox et al., 1976), OSIRIS III release 2 (Institute for Social Research, 1973) and CLUSTAN 1C (Wishart, 1978). Choice of a particular package for a statistical procedure was based on the known characteristics of the packages, the facilities provided for dealing with missing data and the ease and flexibility of use.

A series of tabulation programs were written to augment the computer system programs in the preparation of data for the packages.

2.12 Recording system protocols

2.121 Introduction

The Durham Recording System (Whitton et al., 1976) establishes rules or protocols for the collection of biological, physical and chemical sample data from specified sub-sections of streams and rivers. At the lowest level the Recording System protocols establish the units in which the data are to be collected. The biological units are "species" and are dealt with by the Specieslist Protocol, physical and chemical parameters are dealt with as "catalogued variables" by the Catlist Protocol and the "stream geographical" units are dealt with by the Reachlist Protocol. The protocols which deal with the actual samples taken are the Rivdata Protocol and the Chemdata Protocol.

The various protocols attempt also to describe the data which reside in the system in computing terms. It is obvious that there will be considerably more structure to the data than is suggested by the Recording System protocols.

2.122 Specieslist Protocol

The SIEUR implementation of the Specieslist Protocol currently establishes a checklist of 3810 species numbers (June 1980). The species numbers are 6 digit identifiers in which the first pair of "phylum digits" identify the phylum or broad botanical grouping e.g. lichens, the second pair of "genus digits" identify the genus and the last pair of "species digits" identify the species. A species number may represent either a strict binomial or else a more or less broad "dumping ground". At present (June 1980) there are 27 possible phylum pairs and 606 genus pairs defined by the Specieslist Protocol.

The Specieslist Protocol therefore establishes a check-list of species which is by no means complete. It is however, a list of all the species which an above average aquatic botanist should be able to identify. In addition it is a list which allows every specimen found to be given a species number. The potential loss of information and specificity due to the list excluding rare or difficult species is alleviated by the provision of a system for recording notes about species "dumped".

The addition of a new species number to the list requires that several procedures be followed. The first is a check that no data are present in the system in a dumping ground that could be better described by the new species number. The second requires that all computer and other files be updated and that the addition be listed with a date and time. The third procedure is a systematic update of all data, on computer file and other media affected by the addition.

The following files are associated with the Specieslist Protocol:

TAXONINFO	computer line file, indexed by species number, containing the definitive current version of the Specieslist
SPECIESLIST	printable list of all the species currently recognised by the protocol. The Master Specieslist is the only definitive copy
TAXVERDAT	list of every version that has ever existed of TAXONINFO
TAXNEWDAT	date, time, version and species number of every addition.

Two programs exist to manage the Specieslist Protocol. One adds new species numbers, updating TAXVERDAT and TAXNEWDAT as necessary. The second produces a Master Specieslist from the new version of TAXONINFO.

2.123 Reachlist Protocol

The SIEUR implementation of the Reachlist Protocol defines a stream and the reach sub-unit of the stream. Streams are grouped according to arbitary political units (countries), and at present (June 1980) there are some 400 streams distributed between 8 countries. The stream is divided into an arbitary number of 10 m reaches each of which must conform to several specified criteria. Every reach has a six digit decimal identifier referred to as a "sarnumber" viz. "ssss.rr". The first four digits of the sarnumber are the "stream number"; the first digit of the stream number identifies the country. The last two digits of the sarnumber is the "reach number"; reaches are numbered from 01 to 99, the higher the reach number the further downstream the reach. Reach number 01 is reserved exclusively for the reach which issues directly from an underground source and which can be deemed to receive no upstream innoculum. The Reachlist Protocol establishes the 10 m reach as a standard unit within the concept of a stream. It is only from designated reaches that 100 mm^2 samples may be taken for biological analysis. Similarly water chemistry samples may only be taken from designated reaches.

The addition of new streams and reaches occurs from time to time, and only occurs after a detailed analysis of the effect of the addition on the data already held.

The following files are associated with the Reachlist Protocol:

REACHINFO computer line file, indexed by stream and reach (sarnumber), containing the definitive current version of the Reachlist

REACHLIST printable list of all the reaches currently recognised by the protocol. The Master Reachlist is the only reliable copy by definition

REACHVERDAT list of every version that has ever existed of REACHINFO

REACHNEWDAT date, time, version and sarnumber of every new addition.

Two programs exist to manage the Reachlist Protocol. One adds new streams and reaches, updating REACHVERDAT and REACHNEWDAT as necessary. The second produces a Master Reachlist from the new version of REACHINFO.

2.124 Catlist Protocol

The current implementation of the Catlist Protocol defines some 20 semi-quantitative multistate variables and some 200 quantitative continuous variables. The multistate variables are managed fully by the Catlist Protocol, the continuous variables are not integrated fully into the automatic maintenance by the protocol. They are managed on an ad hoc basis in files known as dictionaries; they are however rigidly bound by the precepts of the protocol. When sample data are added to the system, the individual values recorded are checked against the protocol to see if they exist (in the case of multistate variables) or for reasonableness (in the case of continous variables). If a multistate variable value does not exist that whole sample record will not be added into the computing system. If a continuous variable value is outside the bounds of reasonableness the user is warned and must over-ride the warning if the sample record is to be added to the system.

Automatic maintenance of the Catlist Protocol only occurs for the multistate variables as mentioned earlier. Addition of variables to the protocols is avoided though not impossible. Addition to the states that may exist occurs from time to time. The data files associated with this protocol are:

CATINFO computer file indexed by variable number and multistate value code

CATLIST printable list of the currently recognised variables and their acceptable values

CATVERDAT list of all the version of CATINFO

CATNEWDAT date, time and version of all new variables or values added to the

	protocol
RIV.DICT	dictionary of all variables found in the Rivdata Protocol
CHEM.DICT	dictionary of all variables found in the Chemdata Protocol
QSYS.DICT	dictionary concerned with all system variables used by the Query Subsystem (see 2.134).

Two programs exist to manage the multistate portion of the protocol. These update the CATVERDAT and CATNEWDAT files as appropriate and produce a Master Catlist from the new version of CATINFO.

2.125 Rivdata Protocol for biologically related data

The current implementation of the Rivdata Protocol establishes four sampling units for the collection and storage of biological and related data. Each unit constitutes a detailed data collection sub-protocol; two are currently in use; the "floating reach", and "half km stretch" are not yet implemented. The "one by one" sub-protocol recognises data collected from 100 mm^2 areas of the reach substratum (as defined by the Reachlist Protocol), and the "aggregate" sub-protocol which allows description of the complete flora of a reach on one specified day.

The Rivdata Protocol recognises not only the sampling units but also that the data collected has some structure and therefore provides common mechanisms for handling these structures (see Section 1.13). The structural aspects of the data and more user relevent details may be found in the documentation, however, the main structures recognised are summarised below:

1. Attribute type; usually obligatory variables which may be multistate and are always positive integers

2. n-list type; any list of positive integers e.g. the species numbers found in a biological sample preceded by n the number of such integers

3. Hierarchical type; where several n-lists may be anchored only on an item in another n-list e.g. epiphyte data

4. Textual type; an extension of the n-list to cover character data where n is the number of separate coherent pieces of text of m characters each (currently m is 80 corresponding to an IBM punched card image).

The addition, update and retrieval of data from this protocol form the basis of the SIEUR computing system and is described in Section 2.13.

The following files are associated with the Rivdata Protocol:

RIVDATA main data file, which consists for each record of 6 n-m lists which contain the record key, the raw data as entered, the list of species and live and dead scores, the list of epiphytes, the list of morphological forms and a list of notes see also 2.132

SPEDAT auxilliary file indexed by speciesnumber and live dead values containing entries consisting of the RIVDATA serial processing identifier

SARDAT auxilliary file indexed by sarnumber containing entries consisting of the RIVDATA serial processing identifier

DATEDAT auxilliary file indexed by date i.e. "YYMMDD" containing entries consisting of the RIVDATA serial processing identifier

RIV.DICT See 2.214

S.A.LOG SIEUR system activity log containing entries for every transaction performed on the data. The entries are statistics recorded about the transaction undertaken.

2.126 The Chemdata Protocol for chemically related data

The Chemdata Protocol establishes a mechanism for collecting and storing water chemistry and related data. Only one data structure which is similar to the attribute type mentioned above is available, the difference from the attribute type being that continuous variables are stored and are usually optionally present. The chemical data collection is bound by the Reachlist Protocol 2.123. Each chemical sample must come from a designated reach and must be collected and analysed in the standard ways associated with each variable being measured. The addition, update and retrieval of chemical data is described in Section 2.13.

The following files are associated with the Chemdata Protocol:

CHEMDATA main water chemistry data file which contains the relevant chemistry data. The CHEMDATA serial processing identifier is used as the auxilliary file entry. See 2.132

SARDAT See 2.126

DATEDAT See 2.126

CHEM.DICT See 2.124

S.A.LOG See 2.126.

2.127 Other protocols

The Auxilliary File Protocol is a computer orientated protocol which defines the nature of index auxilliary files and is not strictly part of the Recording System protocol mechanism. The files are the MTS line files in which the line number is used to represent a key, and the logical record is an n-list (2.125) of pointers into the relevent datasets. The maintenance of the auxilliary index (inverted) files is dealt with in Section 2.13. The files which form the protocol are suffixed -DAT and are described in terms of their data content in 2.126.

There are also provisions for separate protocols to deal with sediment and plant analyses, although at this time there exists no complete implementation of these.

2.13 Data organisation

2.131 Introduction

The Recording System protocols for sampled data (See 2.125 & 2.126) are implemented in such a way as to allow three operations to occur: data entry, managed by the "Testadd & Add Subsystem", data update managed by the "Update Subsystem" and data retrieval managed by the "Query subsystem".

This section deals with the theoretical aspects of the three subsystems. For more detailed user information the system documentation must be consulted.

2.132 Data entry - Testadd & Add Subsystem

The Testadd & Add Subsystem performs two operations. It checks the syntax and semantics of each datum (Testadd), and then updates the data and auxilliary files, and logs information about such updates (Add).

The syntax check tests the input fields to ensure that the data are of the expected type, and contains no inconsistancies. A semantics check then uses the Catlist Protocol files to ensure that the data are valid in terms of the recording system protocols. When these two operations are complete for the whole input record an internal image of the record is made. When all the records of any particular input dataset have been checked and a complete internal image of the data exists, the internal image is checked to ensure that no duplicate input record exists. This is done using a key, the composition of which is discussed later. Each input record key is checked for duplication against data already in the system using the stream and reach auxiliary file (SARDAT). If no duplicates exist and no other errors have been encountered the Testadd & Add Subsystem moves into its final phase which performs the following sequence of activities:

1. a note is made into a file that the system is vunerable

2. the system data files are locked for update, and all other users are denied access forcibly

3. the internal image of the input data is converted into its final form and deposited in the relevant file. Any auxilliary files affected are also updated

4. each input record is given a "serid" which consists of an unique

processing serial number referred to as a "seral" which includes a one character seral type identifier known as the "sid" which is ":" for water chemistry records, "*" for one by one biological records, and "#" for aggregate biological records.

When these operations have been completed successfully, the data files are unlocked and other users permitted access to them. The system vulnerable message is removed and system update logs written with information about the progress of the addition, and statistics about the input data and the time taken to complete the data addition.

If the system should fail in this final phase or if MTS fails for any reason, the system is deemed to have crashed. If the system has crashed irrevocably, the back-up copy of the system must be brought across from tape or private disk. If only a portion of the system is affected, there are recovery programs which can "mend" the data. However, a system crash has occurred only once - because of a MTS fault - and recovery took 54.2 seconds of computer processor time (CPU time).

The sequence of operations described above is true for both biological and chemical data. The differences which exist are due to the nature of the comparison keys, and the physical separation of the final data into biology and chemistry files. This difference is for convenience in handling the data and testing the system and arises from the different sampling methodologies. It has been envisaged, but may be impractical, to combine the two for computing purposes such that only one series of serals need be allocated.

The comparison key for water chemistry records consists of the stream and reach number, year, month, day and hour of the sampling. The comparison key for 100 mm^2 biological samples is the stream, reach and sample number, year, month and day. The same key is used for biological aggregates samplings except that the sample number is always assumed to be Ø since by definition there will be only one aggregate on any one day. One problem associated with this assumption is that several water chemistries for a reach may all refer to the same aggregate biology.

2.133 Data update - Update Subsystem

Update is possible only on an ad hoc basis, whereby the MTS text editor (see 2.114), is invoked from the Update Subsystem and used to modify stored data. Where update involves the auxilliary files great care needs to be taken to ensure that it has been correctly researched. At present update is performed one 'seral' at a time. It is envisaged that a future version of the program will work on collections of serals. This modification would prove extremely useful in updating biological data e.g. when taxonomic changes occur.

2.134 Data retrieval - Query Subsystem

The Query Subsystem is organised to allow queries of the general form:

RETRIEVE what FROM where WHEN query

where the capitalised words are more or less obligatory noise words.

The "what" clause specifies what subset of the retrieved data is required for display. The more usual course is to allow this clause to take its default value of "stack". The stack is a collections of serids (as described in 2.132) which may be stored in a file for future use or used immediately to make further queries.

The "where" clause is used to notify the subsystem whether a stack or raw data are to be searched. If raw data are specified the system performs various optimisation procedures in processing the data.

The "query" clause is of the general form:

V O N [log op V O N . . .]

"V O N" is referred to as a query atom in which the "V" is any of the recognised variables (to be found in the query system dictionary file QSYS.DICT), "O" is any of the operators (= , ~= , > , >= , < , <=), "N" is a constant and "log op" is either of the logical operators "and" or "or".

Included in the Query Subsystem are many features including a "HELP" command, "SET" command and "TABULATE" command.

The Query Subsystem contains one major problem associated with the cross reference of biological data and chemical data. This is due to the difficulty in dealing with a query of the form:

"RETRIEVE speciesnumbers FROM rawdata WHEN zinc > 0.5" for which there are many possible alternative interpretations. For example, this may mean :

> "retrieve speciesnumbers when the zinc level was definately greater than 0.5 mg l^{-1}"

that is a biological sample collected on the same date and at the same time as a chemical sample; or it may mean:

> "for every site which has ever had zinc greater than 0.5 mg l^{-1} retrieve speciesnumbers"

The problem is one of expansion from a seral to a stream and reach number, and then from that stream and reach number to further serals and stems from the separation of chemical and biological data both in terms of sampling strategy and computer storage (see 2.132). Expansion is clearly dependant on whether a dated, direct, defined link is desired between the aggregate biology and chemistry (or vice-versa) or whether an indirect, by stream/reach association link is required. These expansion needs are fulfilled by the provision of two parameters to an "EXPAND" command, embedded in the Query Subsystem. The "BYDATE" parameter allows the former defined expansion, the "BYSAR", the indirect stream and reach expansion. The "SET" command allows the default parameter (EXPANDHOW = BYSAR) to be changed. A further set option (AUTOEXPAND = ON) allows every query to be automatically expanded according to the current setting of EXPANDHOW.

The model for the Query Subsystem is the SEQUEL, relational calculus proposed by Chamberlin <u>et</u> <u>al</u>. (1972). Further details of this rather extensive subsystem are covered in the documentation.

2.2 Acid stream data

2.21 Data sources and site information

2.211 Introduction

The acid data analysed in this report are only part of the considerable body of data resident in the SIEUR system which were collected by various collegues in the Botany Department at Durham, either as part of major surveys or as exploratory site studies. Although the majority of the acid samples were analysed by the people collecting the data much work was needed by the author to ensure that the sample data were complete and were processed according to all the various data collection protocols. The sites from which these data were collected are sketched out in Part 6.1 using data available in SIEUR system files. Additional data have been included where these are relevant but have not been entered into the computing system.

Section 2.21 will indicate what data exist in the system as an adjunct to what data have been selected using the Query Subsystem for further analysis. All the data collected have been processed through the SIEUR entry subsystem, which implies that they conform to the system protocols described before (Section 2.12). The remaining sections in 2.2 will describe the data selection, processing and assumptions inherent in the acid study.

2.212 United Kingdom data

Acid data for the United Kingdom (UK) are in two sections, the largest and most important being the Hargreaves dataset (Hargreaves, 1977). The second section is made up of several smaller datasets from many sources. The number of streams and reaches in which the pH has been at or below various levels is summarised in Table 4.12-1. The reach information for reaches below pH 4.0 is presented in Table 5.12-1. The Hargreaves dataset consists of data collected from the Brandon Acid Stream complex (1972 - 1975), and two general surveys of acid sites performed once in late summer (1973) and once in spring (1974). The smaller datasets come from occasional and survey trips to various sites which were undertaken to provide a more complete picture of the distribution of acidity within the UK. They are referred to as the "Parys Mountain" and "Westfield" datasets.

2.213 Other data

Some additional data were collected from several European and American sites. The data for these sites in Table 5.12-1 are incomplete owing to the difficulties of obtaining maps of a comparable nature to the British Ordnance Survey series, and owing to the short duration of each visit. These various datasets are referred to as the "Avoca Basin" dataset the "US mining" dataset, the "Yellowstone" dataset and the "Belgian" dataset.

2.22 Data types and data processing

2.221 Introduction

This section will describe the physico-chemical and biological data in terms of the protocols used to collect them, and will justify the methods used with reference to the acid dataset. Statistical assumptions concerning the data and considerations of specific acid datasets are discussed in Section 2.24.

2.222 Water chemistry data

The water chemistry data were collected according to the Chemdata Protocol and generally follows the methods suggested by the American Public Health Association (1973). However, since the data which comprise the acid dataset were collected at different dates, the specific method of obtaining a measurement was frequently different. Each new fraction or method represents a refinement of technique resulting in either a lowering of the detectable limit, a widening of the detectable range or a more accurate measurement of the variable based on criteria such as the presence of contaminants. When data selection was performed (see Section 2.23) the substitutions outlined in Table 2.22-1 were performed to allow data from different sites and different surveys to be compared. The selection order is the reverse of the historical sequence in which the various methods were added to the Chemdata Protocol.

Table 2.22-1 deals with zinc although the same fractions apply to all the cations. In the presence of more than one method or fraction in the data, the value for the method first encountered in the table is used. Anions are dealt with in a similar manner. However, for the acid analysis no substitutions were required. Table 2.22-2 summarises the cation fractions and methods used in generating the acid dataset.

There has been no attempt to see if these substitutions were valid in statistical terms for the following reasons.

1. there are insufficient data to see if the statistical means of any two methods are significantly different, mainly because two or more methods have only rarely been employed together

2. the data have been collected from widely different sources and with different aims in mind, consequently the various methods that have been

Table 2.22-1

Cation fractions - preferred selection sequence

$Zn+HNO_3$	-- The standard methodology implies SINTA filtering
Zn+HCl	-- As above, with HCl rather than HNO_3 added
$Zn-N+HNO_3$	-- Nuclepore filtering, (rarely employed)
Zn-N+HCl	-- As above with HCl rather than HNO_3 added
Zn-N	-- Nuclepore filtered no acid added
Zn	-- SINTA filtered no acid added
$Zn-T+HNO_3$	-- (Total) Not used for acid data
Zn-T+HCl	-- (Total) Not used for acid data
Zn-T	-- (Total) Not used for acid data

Acid (where present) is always added in the field.
The last three methods are included for completeness.

Table 2.22.2

Fraction or methods used in the acid samples dataset analyses

Variable	Total valid	sinta $+HNO_3$	sinta +HCl	nuclepore $+HNO_3$	nuclepore +HCl	nuclepore acid	sinta	Total $+HNO_3$	Total +HCl	Total -acid
Na	264	Ø	(20)	(12)	(1)	(14)	(217)	13	17	12
K	264	Ø	(20)	(12)	(1)	(14)	(217)	13	14	12
Mg	265	Ø	(20)	(12)	(2)	(14)	(217)	13	15	12
Ca	265	Ø	(20)	(12)	(2)	(14)	(217)	13	14	12
Al	264	Ø	(20)	(12)	(1)	(14)	(217)	13	14	12
Mn	265	Ø	(20)	(12)	(2)	(14)	(217)	13	15	12
Fe	265	Ø	(20)	(13)	(2)	(14)	(217)	12	15	12
Ni	254	Ø	(14)	(9)	(1)	(13)	(217)	13	14	9
Co	253	Ø	(14)	(9)	(1)	(12)	(217)	13	14	12
Cu	252	Ø	(20)	(12)	(1)	(12)	(207)	13	14	7
Zn	265	Ø	(20)	(12)	(2)	(14)	(217)	13	15	12
Pb	267	Ø	(21)	(13)	(2)	15(14)	(217)	13	15	12

(Read figures in brackets (15) as 15 present 15 used. Pairs of figures 15(14) read as 15 present 1 redundant 14 used. Single figures not in brackets read as 15 present 15 redundant.)

used are those that would have been the most appropriate to the needs of the moment; to disentangle the aims and the different methods would require a considerable study in itself

3. the view has been taken that since each measurement is a sampling of the true value, and therefore liable to sampling error, each different method merely includes a greater or lesser degree of error. It is assumed that this error will be taken into consideration when conclusions are being made.

In the results and discussions that follow it is the variable parameter which is considered not the variable measurement method; however, these problems of comparability must be borne in mind. It is likely that they will continue to occur both with the SIEUR system, and more widely in published data. The Determinand Dictionary of the British Water Data Unit (1975) proposes no solution to the problem and no comparable studies have been found in the literature.

The water chemistry data frequently include values which are quoted as less than the detection limit. (e.g. "<0.003" where 0.003 is the detection limit for that variable). These values are valid data, but are difficult to incorporate into statistical analyses: Three methods of looking at such data have been considered.

1. to treat the data as missing. This will result in the variable mean being biased upwards and the loss of data
2. to consider the data as being at the detection limit value. Here there is no loss of data however, the mean will be generally a higher estimate than if the exact value were available
3. the compromise utilised has been to halve the detection limit value. This has the advantage that data are not lost but has the disadvantage that if a considerable number of detection limit values are present a completely false estimate of the mean will result. Part of this compromise methodology therefore has been to see if more than 10% of the measurements are at the detection limit, and if so to take half at the detection limit and half at half the detection limit, in effect this uses the detection limit value multiplied by 0.75. (Variables which include these detection limit values are indicated in the discussions which follow, as are the methods of dealing with such data.)

2.223 Biological data

The analysis of acid biological data has been simplified by the inclusion of aggregate biologies and the exclusion of one by one biologies (Section 2.12). Where an expansion step produced only one by one biologies, these were combined to produce "synthetic aggregates". This was only done when the one by one biologies had been collected from the same reach and on the same day. The Hargreaves dataset consisted originally of only one by one biologies so where possible these too were combined to form synthetic aggregates. Since many of the aggregates had missing attributes, attributes have been excluded from the discussion. No attempt has been made to cross-refer to the many 100 mm^2 samples which may also occur with the aggregates. Since only aggregates are considered, special care has been taken when a species absence is recorded. Consideration has been given to the possibility of species being washed into a reach from sites upstream, and conversly, weight has been given to the absence of a species at reach 01.

The use of the Specieslist Protocol in identifying all plant material in a reach means that considerable cross site comparison is possible. In contrast the relative nature of the 0-5 scale for species live and dead abundance poses several problems for inter-site comparison. These are due to the absence of any absolute figure on which the relative abundances may be anchored. However, since the Rivdata Protocol always establishes one dominant species giving it an abundance score of 5 (or several at abundance 4), we can consider the abundances as rankings of an original "species importance" variable, "species importance" being related either to biomass or chlorophyll a content by some linear function.

Neither the epiphyte nor morphological form mechanisms of the Rivdata Protocol have been considered in the acid study.

2.224 Data storage and the use of the Query Subsystem

The acid stream data are all in the SIEUR files. In order to facilitate data processing the data selected for special study were accessed using the "saved seral stack" mechanism of the SIEUR Query Subsystem. In exceptional cases the data for a study were extracted and placed in separate files. Special programs were written for both these operations. The incorporation of these programs into the general system is under active review.

It was found to be exceptionally easy to use the tabulation routine and consequently many more stacks were examined than are discussed. Many problems arise out of the need to store these data and the results of statistical operations on the data. Some progress toward a solution of these problems has been made by developing programs which display the data graphically. The resulting diagrams take up less space than the statistical tabulations and make an immediate visual impact. However, much additional work needs to be done to develop these techniques further.

2.23 Data selection criteria

2.231 Introduction

The data in SIEUR are extracted for analyses using the Query Subsystem as described in Section 2.13. Each successful query produces a results stack of serals which form the starting point for data tabulations and subsequent statistical analyses. Data selection is therefore the selection of the parameters which govern the query. These parameters are the various variables and constants which define a query atom and ultimately the seral stack. The choice of a particular set of parameters depend on the number of samples required for a particular study and the nature of the study. The requirements for the acid study were that approximately 100 biological aggregates collected on the same day as an acid water chemistry were available. To facilitate parameter selection stream and reach distribution histograms for pH, conductivity, acidity and redox were prepared. These were then used in conjunction, to select the various pH levels at which stacks would be prepared. The levels used were <=2.0, <=3.0, <=4.0 and <=5.0. The "less than or equal" criterion was used to ensure that data selected were comprehensive and spanned the magnitude jump inherent in the logarithmic nature of pH. The 269 water chemistries found at or below pH 4.0 were found to yield 125 biological aggregates on expansion BYDATE, 124 were unique (Only two chemistries had been collected on the same date from the same reach, thereby yielding the same aggregate on expansion). The 269 chemistries constitute the "269 dataset" or acid samples dataset and the 125 chemistries and 124 biologies the acid species or "125 dataset". The "acid sites dataset" was deemed to be the BYSAR expansion of the 269 dataset to obtain aggregates followed by a second expansion BYSAR of the aggregates to obtain more chemistries. These operations yielded an additional 10 chemistries and 9 aggregates.

Data additional to the acid stream data have been used to clarify or extend what is known. This happens for example when knowledge of the chemical range of a species is required (Section 6.22).

The initial criteria for data selection were heavily dependent on the final aim of the analysis and it is the purpose of this section to outline these analysis aims and the route by which the analysis may be performed. Section 2.24 contains a description of the statistical approach to the acid analysis.

2.232 Acid studies and aims

The specific methods by which the aims outlined in Part 1.3 may be fullfilled are:

1. To identify those reaches which fall within the acid site definition and to summarise on a reach by reach basis all data held for these reaches (this implies BYSAR expansion)

2. to study intervariable relationships in the chemical dataset produced and to analyse any trends displayed

3. to group together reaches with similar chemistries and to provide means by which such similarities may be measured and discussed.

Similar approaches exist for the analysis of the biological data these are:

1. To identify the species present at reaches defined as being acid and to summarise on a species by species basis all the data held for those species

2. to study similarities in species assemblages for various reaches using a variety of techniques and to account for any differences

3. to group together species possessing similar distributions and to design ways of looking at and contrasting these distributions.

These methods were used to examine the three datasets described in 2.231. The acid samples chemistry dataset was used to study the chemistry of the acid condition and to identify groups among the chemistry samples. The acid sites chemistry dataset was used to provide complete information about the sites identified as being acid and to allow a more detailed grouping of the samples to occur such that site groupings could be postulated.

The initial query used to identify the acid samples was:

pH_fld <= 4.0

details of other possible values and the results of test queries using these are presented in Chapter 4. The results stacks produced were expanded as required and a list of species obtained. These are presented in Chapter 6.

Univariate statistics were produced (Table 4.13-1 and Table 4.23-1) to describe the variables and to act as a prelude to the multivariate study of the data. A principal aim of the univariate study being to identify variables or their transforms which were distributed symmetrically - these aims and their rationale are described in Section 2.24.

2.233 Species studies

In these studies the aim was to ellucidate some fact or facts concerning individual species. The reasons for making such a study are twofold. There may be an a priori reason why a particular species or group of species is of interest - or alternatively species thrown up by a specific study (acid biological studies) - may be of further interest. In either case data from reaches not included in the acid sites description may be used and there will be no attempt to describe these additional sites in this thesis.

Species which occur in more than 3 reaches with field pH values less than or equal to 4.0 on the date of sampling are automatically of interest. SIEUR has been entered using the Query Subsystem to extract all chemical data collected on the same date for a reach which on that date had the species recorded as present and live. The chemical data are presented variable by variable with the maximum, minimum and mean for each of these species in Table 6.12-1.

Two specific studies were performed to assess the differences in distribution of two filamentous genera (Hormidium and Zygogonium) and two diatom species (Eunotia exigua and Pinnularia acoricola). These were chosen because of their wide occurrence in acid waters and because in the case of the filamentous species of their suspected different tolerance to heavy metals especially copper (B.A. Whitton pers. comm.).

2.24 General assumptions and statistical approach

2.241 Introduction

Section 2.24 draws together the basic assumptions the database approach used and discusses the theoretical aspects of the methods and the inherent limitations of SIEUR and the acid stream dataset. A subsidiary aim will be to keep this introduction to a general level so as to allow comparisons of subsequent analyses using the SIEUR system. The data from such analyses will follow the same general pattern of a large table with areas of missing data and may be summarised by the diagram in Table 2.24-1.

The major problem of the database approach is data comparability and the validity of methods which because of their heuristic nature cannot be heavily relied upon. The aim will be to present those methods used which were robust statistically and which made few limiting assumptions in themselves. The main approach of the analyses was to extrapolate from the smaller complete datasets to cover conclusions drawn about the larger dataset. (The numbers in Table 2.24-1 refer to the acid datasets defined in Section 2.23 and will be dealt with in Section 4.15)

2.242 Assumptions made for acid data analysis

Assumptions are made about the data at several stages:

1. Collection

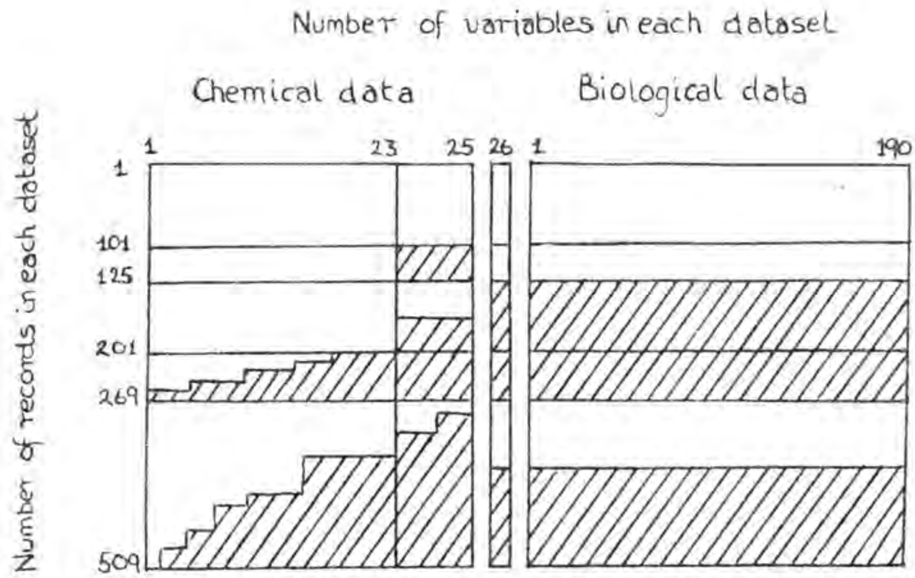
Assumptions concerning the chemical data analysis are concerned with the method used, the degree of replication, the means of taking the sample and the actual choice of sampling point both spatially and temporally. Similar problems exist with the biological data, where there are, however, additional ones; completeness of the species complement, subjectivity of the abundance measure, difficulties in identification, including levels of competence by the user and assumptions about species comparison across phyla, genus and species boundaries.

2. Collation

At the data collation level are the assumptions inherent in the various protocols (Section 2.12). Because of the need to encode the data certain broad assumptions are made which will be more or less accurate. Further, data which are highly relevant may be simply in note form because

Table 2.24-1

Schematic table of the acid datasets



Notes

1. the horizontal lines represent the various datasets
2. the shaded areas represent missing data
3. variable 26 is NSPEC, which is considered in many of the analyses to be a chemistry variable

of the absence of mechanisms to deal with them.

3. Entry

These assumptions include the encoding problems, data validification and all the assumptions made implicitly by the Testadd and Add system about what are "correct" data (see 2.132).

4. Selection

At this level are the assumptions concerning detectable limits and missing data, the validity of performing expansions and the assumptions about data exclusion made implicitly on performing a query (see 2.133). These, along with the data collection assumptions form the major areas of weakness of the computerised approach.

5. Synthesis for statistical analysis

The assumptions at this level concern the treatment of missing data of various types as dealt with in Section 2.22, the methods of dealing with detectable limits and the problems of inter and intra reach comparison. Of these the latter is probably of greater concern.

6. Statistical analyses

See 2.243.

2.243 Statistical assumptions

The statistical assumptions made are described with each method in 2.245, 2.246 and 2.247. The general assumptions are:

1. that the individual chemistry samples are all statistically independent. This implies that there is no link, for the purposes of the statistics between reaches or between multiple samplings of the same reaches. However, attempts have been made to ensure that few reaches have been included which are multiply sampled or which have been sampled closely together with regards to date

2. that the variables measured for each sample do indeed describe the reach chemistry and biology and that the underlying pattern of variable inter-relationships within a sample is the same for all samples across all the reaches. This implies that each sample is a snapshot of the reach which is in a more or less stable state in some general stream continuum

3. a general assumption about the chemical variables is that they are distributed symmetrically about a mean. If this is not the case a suitable transformation of the data is found which produces a symmetry. This symmetric distribution may or may not approach statistical normality; however, it is assumed that this is of no consequence and that all a statistical method will require is a symmetric distribution (J. Besag pers. comm.) This technique is justified since no probability values are associated with any of the methods under consideration and only an heuristic analysis of the data is being performed.

4. it is assumed that the species abundance values in an aggregate will follow the "J distribution" described by Jaccard (1923) and others. There is no attempt to analyse the implications of this distribution with reference to the data under consideration. It must however be borne in mind that there is a bias in the data because of the Rivdata Protocol requirement for a dominant at rank 5 (or several at rank 4) always to be present.

2.244 Data validity checks

The data are checked for validity by the Add Subsystem. However, there are no ways of checking whether the data are correct. This implies that erroneous species may be added or that a chemical measurement may be incorrect. Checks are performed visually to spot errors, and some machine tests are also possible; however, a small percentage of errors do escape detection. There have been several attempts to trap these errors. None has been able to deal automatically with the extreme condition which is the basis of these studies. It has proved necessary therefore to examine the data statistically looking for values which do not correspond to subjective a priori assumptions about probable values. In the three cases detected, the errors were due to incorrect transcription of data from notebook to computer entry sheet. Values which are statistically valid but neverthelss incorrect, can in fact escape detection.

2.245 Chemistry variable distribution analyses and transformations

The purpose of looking at the distribution of the chemistry variables is to ascertain if those distributions are symmetrical. If they are not symmetrical, various transformations were applied to obtain the required symmetry (Bartlett, 1947). It is usually these transformed variables which are discussed in this thesis.

An histogram technique was used initially to look at the distribution of chemical variables. These are presented in Table 4.22-1 with the transformation which produced the most reasonable symmetrical distribution. The choice of the number of histogram bins was based on the square root of the number of valid values (Fox, 1978). The various transformations which were tried are shown in Table 2.24-2.

No attempt has been made to check if any of the variables fitted normal distributions except insofar as symmetry statistics and variances were calculated. It was assumed that since only an heuristic approach was being adopted a subjective symmetrical distribution for a variable was all that need be presented to any statistic requiring normally distributed variables.

Non-parametric versions of the main statistics were used to overcome variance heteroscedasticity. It was assumed that the variables would be particularly susceptible to variance heteroscedasticity owing to the cross

Table 2.24-2

Possible variable transformations and their effects

Transformation Y = f(X)	Range of variable	Characteristics of distribution	Examples of such distributions
SQRT(X) + SQRT(X+1)	0<=X<=inf	Variance proportional to the mean (Distribution skewed to the right)	Distributions arising from "count" data e.g. Poisson
Log(X)	0<=X<=inf	Standard deviation is proportional to mean	Distributions of s2 for normal samples
arcsin SQRT(X)	0<=X<=1	Mean = m Variance = k*m*(1-m)	Distributions of proportions e.g. binomial

site pooling of data; however, this cannot be tested since for most sites only one sample exists. It was further assumed that variance heterogeneity would occur owing to the different accuracies of variable estimation. Where the results of parametric statistics are under discussion these problems must be given consideration.

2.246 Biology variable distribution and transformations

The distributions of species abundances as measured by the subjective 1-5 scale is shown in Table 6.14-1. The order of species in this table was determined by the lowest pH level at which that species was found. The order of reaches is that obtained by the ordination of sites using the multivariate methods discussed in 2.467.

Cross comparison of biological samples is made simpler by the existence of the Rivdata Protocol requirement whereby there must be a dominant species at level 5 (or two or more at level 4) to which all the levels of other species are related. Species diversity is assessed by use of "nspec" the simple count of species found in a sample. This is a highly variable measure because of possible changes in taxonomy resulting in species being merged or otherwise changed, and because of possible additions of new species on more detailed examination of preserved or cleared (diatom) material. This is a special problem when extreme sites like those of the acid study are considered because of the inherent low nspec values. A method by which this could be alleviated is by summary of ranks or some simple functional combination based on weighting the species or ranks on subjective criteria. It has been found that the statistic obtained is of of little additional value to the simple count of species when sites with high NSPEC are considered but can help to smooth out perturbations when sites with low NSPEC are considered - but only when the abundance value for the "new" species is low (1 or 2). This methodology has been rejected for the current study because of the considerable subjectivity of the decision on weightings and the tendency for such decisions to be correct only for the limited analyses under consideration. When comparisons between species from different sites are required the most appropriate method would be to translate the abundance scores, using some measure (or estimate) of the aggregate sample biomass, into absolute biomass values for each species. The clear goal of the Rivdata Protocol abundance mechanism is to facilitate such translations when methods of measuring biomass become available. This is of particular importance for botanical studies of extreme environments where the primary concern is the species' ability to cope with the prevailing conditions.

Transformations of the species data are performed either to simplify the data or as a means of making comparisons with other data or other methods. The simplest transformation of the abundance data is to a simple presence absence measure. This transformation is also used in the calculation of various similarity measures between species assemblages. It is used frequently in an implicit manner in the Query Subsystem, and the calculation of species chemistry ranges as presented in Table 3.41-3. The view has been taken that since the abundance scale is a relative function of the sample biomass, it is difficult to compare a rank 2 (say) with a 2 from a species rich site unless the biomass, however measured, is known. Where only acid sites are under consideration, the assumption is made that the total biomass of samples are sufficiently similar to allow this restriction to be relaxed.

The distribution of a species over the various chemistry variables ranges is presented. These assume that a live presence absence transformation has been performed.

2.247 Multivariate methods - Principal Components Analysis (PCA)

Many methods have been suggested in the literature for the simultaneous analysis of several variables (Cairns et al., 1972), both for continuous (chemistry like) data and binary (species presence/absence) or categorical (species abundance) data (e.g. Lambert et al., 1966; Lance et al., 1969). In all these methods the process involves reducing the information in a Table of q rows and r columns into a more manageable form. The methods usually commence with an estimation of the similarity in the matrix between rows or columns. If some statistically meaningful, precise method exists e.g. a product/moment correlation coefficient between chemistry variables in a matrix of measurements of several variables sampled at several sites, the method can be more or less exact (Kim, 1975). Where this is not so (e.g. in estimating the similarity of sites in the chemistry data matrix), other statistics or transformations must be performed first. Where the data matrix consists of simple presence/absence binary records or non-continuous discrete data (e.g. in the aggregate biologies dataset where the matrix consists of subjective abundances for species at several sites), other usually less precise methods of computing similarities must be examined.

The usual aims of these methods are:

1. Exploration of the data - the heuristic approach
2. Confirmation of hypotheses - the probablistic approach
3. Simultaneous variable measurement - the combinatorial approach.

Of the several multivariate methods which could be used for examining similarity in the chemistry data, Principal Components Analysis (PCA) was chosen for presentation here for the following reasons:

1. the method provides an exact transformation of the original q by q or r by r similarity matrix with no loss of variance information
2. the components are orthogonal (uncorrelated) linear combinations of the original variables where the first component is the best single summary of linear relationship exhibited in the data, the second component the second best and so on
3. no particular assumptions about the underlying structure of the variables are required - although the use of standardised symmetrical variables provides a component solution which is easier to interpret
4. the method is sufficiently general for most purposes and sufficiently robust statistically to be used in most multivariate analysis situations.

An heuristic PCA was performed on the acid chemistry datasets (Table 4.15-1) to see if there was any underlying pattern to the inter-correlation of the chemical variables. The aim being to produce a few synthetic components which would describe more exactly than the variables themselves any pattern in the variation present. To perform these analyses the OSIRIS III computer package was used for the reasons outlined in 3.217. The use of PCA for the investigation of between site similarity and between species similarity was not undertaken except as confirmation of the results of cluster analysis (q.v), which was deemed to be a far more suitable technique for these types of analysis.

2.248 Multivariate methods - Cluster Analysis

The acid datasets (Table 3.217) were subjected to various cluster analysis methods which grouped the samples, species or variables together based on similarities in their species assemblages or chemistries (see also 2.247). The resulting "clusters" of samples, species or variables are discussed, compared and contrasted.

The method requires two decisions to be made. The first is the measure of similarity to be used between the cases (sample chemistries, species or sample aggregates). Although several measures are available the one thought to be the most suitable for all situations (Wishart, 1975) was the Euclidean distance measure calculated as:

$$d^2_{ik} = 1/M \sum (x_{ji} - x_{jk})^2 \qquad \ldots 1$$

For continous data where:

d^2_{ik} = the squared dissimilarity between case i and case k.

x_{ji} = the ith measure on variable j.

M is the number of variables.

and:

$$d^2_{ik} = (B + C) / M$$

for binary data where:

B = the number of attributes present in case i and absent in case k.

C = the number of attributes present in case k and absent in case i.

M = the total number of attributes.

The second decision is the method of clustering, whether it should be divisive or agglomerative and monothethic or polythetic. Eignt agglomerative methods were compared using the cophenetic correlation method of comparing their results dendrograms. This was achieved using the method described by Sokal & Rohlf (1972) using a specially written computer program. Using the two most dissimilar methods - one which produced "tight spherical clusters" and one which produced "natural clusters" (Wishart, 1978) the binary Table of sites against species were analysed to produce species clusters and site clusters. The site clusters were compared with those obtained from a similar analysis of the chemistry dataset.

2.249 Other statistical methods

An attempt was made to assess the importance of the absence (or over representation) of a species in the aggregates associated with the pH bins in the histograms to be found in Table 4.52-1. This was achieved by use of a chi-square ratio (Visvalingham, 1979) which takes the general form:

$$X^2_{ij} = (Op_{ij} - Ep_{ij})^2/Ep_{ij} + (Oa_{ij} - Ea_{ij})^2/Ea_{ij} \qquad ...1$$

where:

X^2_{ij} is the chi-square value for the jth species at the ith bin.

Op_{ij} is the observed number of aggregates with the jth species present in the ith bin.

Ep_{ij} is the expected number of aggregates with the jth species present in the ith bin.

Oa_{ij} is the observed number of aggregates with the jth species absent in the ith bin.

Op_{ij} is the expected number of aggregates with the jth species present in the ith bin.

but because:

$$Op_{ij} + Oa_{ij} = Ep_{ij} + Ea_{ij} = R_i \qquad ...2$$

$$Op_{ij} - Ep_{ij} = Oa_{ij} - Ea_{ij} \qquad ...3$$

(where R_i is the number of aggregates associated with the ith bin)

Thus:

$$X^2_{ij} = (Op_{ij} - Ep_{ij})^2 * (1/Ep_{ij} + 1/Ea_{ij}) \qquad ...4$$

where:

$$Ep_{ij} = R_i * C_j / S \qquad ...5$$

$$Ea_{ij} = 1 - C_j/S \qquad ...6$$

and:

C_j is the total number of aggregates with species j present in them.

S is the total number of aggregates in the analysis.

This ratio will vary from:

Ø indicating no difference between observed and expected values and

$(R_i * S / C_j) - R_i$ indicating maximum difference.

Clearly as S increases or the ratio R_i / C_j increases the numerical value of the maximum chi-square will increase, further underlining any difference in bins. When plotted as a histogram those bins at which the presence or absence of a species is noteworthy will have the higher chi-square values. Furthermore, the larger numerically the maximum chi-square value, the more reliable the histogram as a measure of species presence/absence importance. Thus, for example, the absence of Euglena mutabilis in the 3.61 - 3.81 pH bin is noteworthy, indicating perhaps missing data, similarly its absence below a pH of 1.Ø9 is noteworthy suggesting perhaps a lower pH limit for the species. In contrast the occurrence (count of 4) of Lepocinclis ovum in pH bin 2.64 - 2.85 is of note, suggesting a disproportionate number of this species at this bin. The respective chi-square values for these two species are 8.42 and 7.4Ø both of which are relatively high when compared with the other maximum chi-square values. The highest chi-square values are (52.Ø8) for Pinnularia braunii (4 occurrences) and 21.85 for Gomphonema parvulum (4 occurrences) both in pH bin 3.Ø2 - 3.22, and indicates perhaps an artefact in the data or alternatively a preferred optimum pH. More detailed discussion and examples of the use of this statistic are presented in Section 4.52.

Chapter 3 The computing system

3.1 SIEUR size and generation time

3.11 Introduction

The subject of Chapter 3 is the SIEUR system produced as a result of implementing the Recording System protocols described in Section 2.12. Part 3.1 will deal with its development in terms of the computer time and other resources used in adding the data and the time and other resources used in developing the computer system itself (3.125). Part 3.2 and 3.3 deal with the restrictions and limitations inherent in the implementation while 3.3 presents the statistics concerning the computing aspects of the acid study and the extraction of the acid data. Chapter 4 will then present the results of the acid data selection procedure and Chapter 5 and 6 the results of the analyses performed on the acid sites and acid species respectively.

In presenting the the estimates of database size, the values used in 3.1 refer to the size of the data files and all index and dictionary files but excludes files containing programs and files associated with these programs - for example "help" files. This estimate is clearly neither the database size in the classical sense of all the data and programs associated with it, nor a real estimate of the quantity of sampled data; the latter could be better expressed as a count of the number of samples and the number of valid measurements made with each sample. The estimate does however give some indication of the quantity of data in practical terms and affords a reference measure for other users. Although some data compression is inherent in the system, no specific techniques have been used and therefore the file sizes quoted include much unused space.

Tables presented in this section contain figures for file generation time which are extracted from the Sieur Activity Log (S.A.LOG). This log was written by programs which have changed considerably as system modifications were implemented making it unlikely that such figures would be obtained by the current versions of the same programs.

3.12 Files associated with the -LIST protocol

Table 3.12-1 shows the size and time taken to generate versions of each file. The development of the -LIST protocols was initially rather arbitary, being geared to the needs of the biological recorder. Once field data had been collected and the problems associated with their processing identified,

Table 3.12-1

Current size and generation time for files associated with the -LIST protocols

Filename	Date of 1st	Number of lines	Size of current	Time taken last update	Total time to generate	Number of versions	Number of editions
TAXONINFO	Aug:75.28	3850	42	11.76	423	7	35
SPECIESLIST	Mar:74.03	3539	200	8.91	423	9	40
TAXVERDAT	Sep:77.01	40	1	.	.	3	20
TAXNEWDAT	Sep:77.01	1063	6	.	.	3	20
REACHINFO	Sep:77.01	1101	24	4.91	154	3	22
REACHLIST	Sep:77.01	3079	249	5.94	175	3	24
RCHVERDAT	Jan:78.01	22	1	.	.	1	6
RCHNEWDAT	Jan:78.01	1310	2	.	.	1	6
CATINFO	Jan:78.01	161	3	1.23	75	4	14
CATLIST	Jun:79.23	799	8	1.60	75	3	13
CATVERDAT	Jun:79.23	13	1	.	.	1	2
CATNEWDAT	Jun:79.23	13	1	.	.	1	2

The current edition of the latest version (number 302) of the Specieslist protocol was produced on May:80.01

The current edition of the latest version (number 120) of the Reachlist protocol was produced on Jun:79.01

The current edition of the latest version (number 100) of the Catlist protocol was produced on Jun:79.23

the implementations of the protocols were improved. The greater number of SPECIESLIST and REACHLIST versions may be explained in terms of these improvements. The earlier first edition dates for SPECIESLIST and REACHLIST files reflect the fact that the protocol approach was not utilised until after work on lists of species and reaches had begun. A limited list of species existed on computer file before the project began but needed considerable modification to allow expansion to occur. The Catlist Protocol was the last to be designed and implemented (see 7.125).

Most versions of the files mentioned in Table 3.12-1 are preserved on magnetic computer tapes. Only the files suffixed -INFO are permanently available. All other files are stored in the SIEUR System Archive on magnetic tape. These files are not included in any of the estimates of database size.

It can be seen that the processor time taken to develop these protocols and generate the files is negligible costing well under £100. The time taken in biological discussions to decide the protocol design both in terms of the Recording System and computer system exceed by several orders of magnitude this total processor time. The main problems in handling these protocols was caused not by difficulties in computing but by problems associated with the Recording System definition of basic items such as "species" and "reach". Clearly for computing purposes these problems had to be dealt with rapidly. Some of the main questions which were thrown up are introduced later in Section 3.24.

3.13 Size and generation time for -DATA files

Table 3.13-1 indicates the size and time taken to generate versions of each data file.

Table 3.13-1

Current size and generation time for files associated with the -DATA protocols

Filename	Date of 1st	Number of lines	Size of current (pages)	Time taken last update (seconds)	Records added last update	Total time to generate (seconds)
CHEMDATA	Oct:78.05	2704	1063	10.47	18	1117
RIVDATA	Aug:77.20	2548	579	37.58	44	2665
DATEDAT	Aug:79.18	414	15	-	-	-
SARDAT	Jun:79.16	648	17	-	-	-
SPEDAT	Jun:79.16	3571	106	-	-	-
CHEM.DICT	Jul:77.11	172	4	-	-	-
RIV.DICT	Aug:77.30	61	2	-	-	-
S.A.LOG	May:78.12	223	6	-	-	-

Table 3.13-2 and 3.13-3 shows the time taken by each phase of the addition program used to generate CHEMDATA and RIVDATA from their constituent raw data files.

The highly variable program loading times measured as:

Total time - (Phase I time + Phase II time) ...1

may be explained by the varying program size due to system development work occurring side by side with data processing. It is also influenced by the time of day at which the program was loaded into the computer.

Much time was spent in ensuring that the data entered were correct. Four Testadd runs were required on average to process the biological raw data files and two to process the chemical files. The most frequent sources of error were the mis-transcription of species records, and format errors in the chemical data.

Table 3.13-2

Timing statistics for the generation of the CHEMDATA file from the constituent raw data files using the Testadd Subsystem

Input filename	Size	Date	Time of day	Phase 1 time	Phase II time	Total time
BTK8:CW.AA	52	Oct:78.05	17:49:11	6.754	2.319	12.799
BTK8:CW.AB	51	Oct:78.05	17:51:36	7.158	2.292	13.104
BTK8:CW.AD	67	Oct:78.06	16:36:56	9.737	3.528	17.069
BTK8:CW.AE	69	Oct:78.06	16:54:16	10.701	4.584	19.191
BTK8:CW.AF	80	Oct:78.06	17:00:58	13.397	5.696	23.365
BTK8:CW.AC	61	Oct:78.09	13:06:57	9.743	4.197	17.638
BTK8:CW.AH	50	Oct:78.09	14:19:31	10.351	2.949	18.020
BTK8:CW.AI	13	Oct:78.11	17:03:49	3.011	1.213	8.319
BTK8:CW.AJ	12	Oct:78.11	17:08:12	3.047	1.040	8.813
BTK8:CW.AK	19	Oct:78.12	15:56:09	4.196	1.569	10.300
BTK8:CW.AL	17	Oct:78.12	15:57:19	4.178	1.424	10.531
BTK8:CW.AM	15	Oct:78.13	15:08:07	4.212	1.678	10.776
BTK8:CW.AN	11	Oct:78.13	16:59:29	2.958	1.105	8.895
BTK8:CW.AO	29	Oct:78.13	17:02:10	6.384	2.233	13.702
BTK8:CW.AP	25	Oct:78.17	15:16:14	6.346	2.625	14.033
BTK8:CW.AQ	36	Oct:78.18	11:39:50	8.777	3.174	17.159
BTK8:CW.AR	21	Oct:78.18	11:43:45	4.194	2.424	11.030
BTK8:CW.AS	30	Oct:78.18	16:56:39	7.818	2.630	15.304
BTK8:CW.AT	25	Oct:78.18	16:59:53	5.201	2.453	12.451
BTK8:CW.AU	10	Oct:78.20	11:49:47	2.623	1.002	8.314
BTK8:CW.AV	44	Oct:78.20	14:54:54	8.885	4.598	18.351
BTK8:CW.AW	19	Oct:78.23	13:14:36	3.694	1.300	10.030
BTK8:CW.AX	11	Oct:78.23	13:14:37	2.817	1.285	8.831
BTK8:CW.AG	85	Jun:79.05	18:57:05	14.325	5.595	23.431
BTK8:CW.AHALF	57	Jun:79.05	19:02:35	11.088	4.038	17.394
BTK8:CW.BA	10	Jun:79.07	13:25:41	2.324	1.015	6.876
BTK8:CW.BB	24	Jun:79.07	13:26:56	3.800	1.831	9.048
BTK8:CW.BD	22	Jun:79.08	13:07:19	3.856	1.970	9.399
BTK8:CW.BE	19	Jun:79.08	13:10:02	3.448	1.803	8.799
BTK8:CW.BF	19	Jun:79.08	13:13:59	3.688	1.525	8.720
BTK8:CW.BG	55	Jun:79.08	13:15:54	8.521	4.468	16.796
BTK8:CW.BH	57	Jun:79.08	13:19:51	9.457	4.309	17.516
BTK8:CW.BI	19	Jun:79.13	13:35:46	4.239	1.850	9.998
BTK8:CW.BJ	32	Jun:79.13	14:06:48	7.121	3.157	14.416
BTK8:CW.BK	36	Jun:79.13	14:25:50	7.333	3.467	14.824
BTK8:CW.BL	30	Jun:79.13	18:15:07	6.228	2.445	12.644
BTK8:CW.BM	25	Jun:79.13	18:24:43	4.341	2.351	10.145
BTK8:CW.BC	28	Jun:79.14	09:41:36	6.447	3.439	13.930
BTK8:CW.BN	29	Jun:79.15	00:38:19	5.624	2.707	11.942
BTK8:CW.BO	49	Jun:79.15	00:40:34	7.629	4.304	15.702
BTK8:CW.BP	49	Jun:79.15	00:43:30	8.230	4.294	16.297
BTK8:CW.BQ	49	Jun:79.15	00:46:18	8.031	4.245	16.048
BTK8:CW.BR	49	Jun:79.15	00:49:06	7.962	4.332	16.023
BTK8:CW.BS	49	Jun:79.15	00:51:41	8.439	4.301	16.493
BTK8:CW.BT	40	Jun:79.21	18:38:27	6.524	3.769	14.396
BTK8:CW.BU	28	Jun:79.21	18:40:54	5.291	2.735	11.630
BTK8:CW.BW	10	Jun:79.21	18:49:47	46.596	17.013	68.834
BTK8:CW.BX	88	Jun:79.21	19:01:55	16.886	7.652	29.813
BTK8:CW.BV	180	Jun:79.23	10:45:43	38.438	16.516	60.292
BTK8:CW.BBE	16	Jul:79.31	12:54:19	3.807	2.416	9.979
BTK8:CW.BBF	17	Jul:79.31	16:23:27	6.187	3.222	13.909
BTK8:CW.BBH	4	Jul:79.31	16:42:34	2.240	1.505	8.898
BTK8:CW.BBI	36	Jul:79.31	16:48:04	5.423	6.114	15.468
BTK8:CW.BBG	45	Jul:79.31	17:05:57	9.425	6.589	20.057
BTK8:CW.BBA	29	Aug:79.03	11:29:02	22.347	19.502	45.923
BTK8:CW.BBB	49	Aug:79.03	10:57:43	9.533	7.931	21.549
BTK8:CW.BBC	49	Aug:79.03	11:07:49	9.719	7.975	21.624
BTK8:CW.BBD	28	Aug:79.03	11:17:50	6.741	5.091	15.935
BTK8:CW.BBJ	34	Aug:79.03	12:21:05	23.184	16.470	43.621
BTK8:CW.BBK	64	Aug:79.03	12:34:44	11.317	9.421	24.537
BTK8:CW.BBL	10	Aug:79.03	17:24:26	2.469	2.170	8.421
BTK8:CW.BBM	18	Aug:79.03	17:56:26	4.023	3.147	10.462

Table 3.13-2

Timing statistics for the generation of the RIVDATA file from the constituent raw data files using the Testadd Subsystem

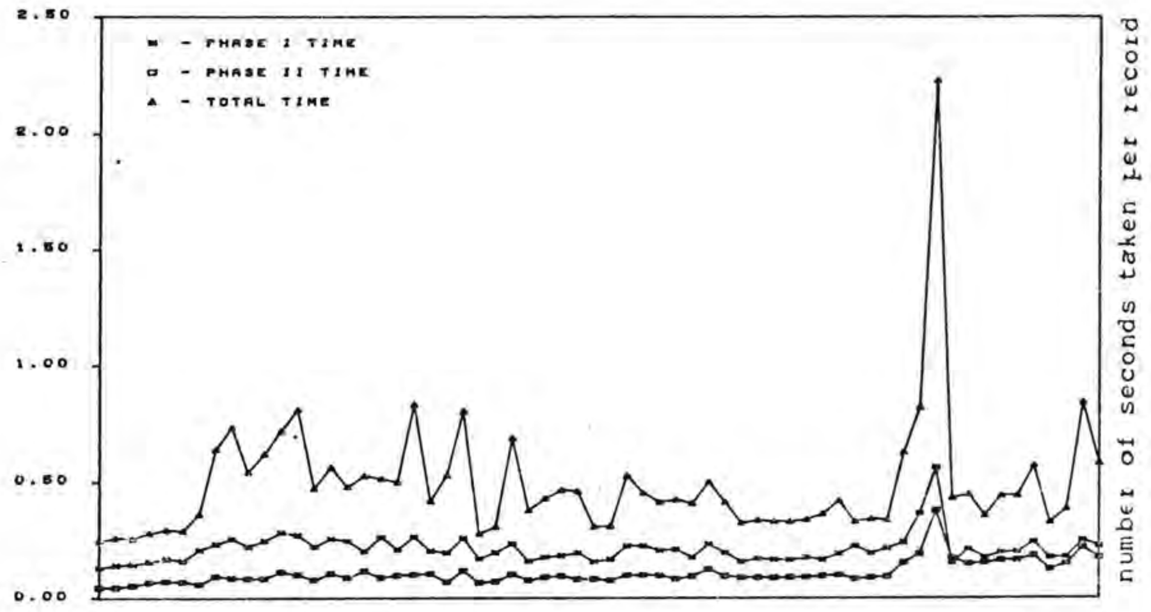
Input filename	Size	Date	Time of day	Phase 1 time	Phase II time	Total time
BTK8:BS.1.GODOT	120	APR:79.05	18:07:34	32.936	52.329	-
BTK8:BS.2.GODOT	118	APR:79.05	18:36:28	38.807	57.951	-
BTK8:BS.3.GODOT	115	APR:79.11	18:36:28	34.021	62.435	-
BTK8:BS.4.GODOT	121	APR:79.11	18:36:28	42.846	72.024	-
BTK8:BS.5.GODOT	100	APR:79.11	18:36:28	31.871	72.291	-
BTK8:BS.6.GODOT	92	APR:79.12	18:36:28	30.071	66.084	-
BTK8:BS.7.GODOT	68	APR:79.12	18:36:28	23.567	86.305	-
BTK8:BS.8.GODOT	67	APR:79.12	18:36:28	24.989	93.461	-
BTK8:BS.1.TWEED	88	APR:79.12	18:36:28	34.200	115.871	-
BTK8:BS.2.TWEED	91	APR:79.12	18:36:28	41.595	145.266	-
BTK8:BS.3.TWEED	96	APR:79.12	18:36:28	39.788	71.482	-
BTK8:BS.4.TWEED	64	APR:79.25	14:09:28	34.506	134.408	-
BTK8:BS.5.TWEED	55	APR:79.25	18:25:23	31.508	120.182	-
BTK8:BS.NIG1	89	APR:79.27	18:30:09	47.799	43.335	91.134
BTK8:BS.NIG2	84	MAY:79.04	18:19:09	43.330	43.882	89.207
BTK8:BS.NIG3	71	MAY:79.05	02:40:50	34.531	38.027	74.289
BTK8:BS.NIG4	19	MAY:79.09	06:30:37	9.345	10.573	21.574
BTK8:BS.WHT1	104	MAY:79.14	18:09:27	31.616	24.414	58.030
BTK8:BS.SAY1	65	MAY:79.15	19:49:31	16.160	15.007	32.020
BTK8:BS.SAY2	80	MAY:79.17	06:34:40	20.948	21.876	44.521
BTK8:BS.JWH1	132	MAY:79.23	01:58:21	27.981	18.441	48.283
BTK8:BS.JWH2	151	MAY:79.23	02:04:41	36.777	21.187	59.757
BTK8:BS.JPH1	16	MAY:79.23	12:22:50	3.383	5.164	10.371
BTK8:BA.NAG1	24	MAY:79.24	23:16:58	14.563	22.571	38.966
BTK8:BA.NAG2	24	MAY:79.24	23:28:31	17.897	28.422	48.133
BTK8:BA.NAG3	18	MAY:79.24	23:52:30	12.975	20.010	34.637
BTK8:BA.NAG4	6	MAY:79.25	00:56:33	4.315	6.938	12.909
BTK8:BA.WAG1	122	MAY:79.25	15:23:28	42.242	44.800	88.816
BTK8:BA.SAG1	98	MAY:79.25	16:06:13	27.877	37.490	67.184
BTK8:BA.SAG2	83	MAY:79.26	10:42:58	25.800	33.890	61.531
BTK8:BA.HAG1	45	MAY:79.26	10:57:47	7.436	8.094	17.228
BTK8:BA.HAG2	63	MAY:79.26	11:00:52	12.435	12.995	27.108
BTK8:BA.PHAG1	10	MAY:79.26	11:06:03	2.193	3.853	7.765
BTK8:BA.SAG3	49	AUG:79.28	19:13:21	15.480	20.121	37.481

3.14 Per record processing statistics

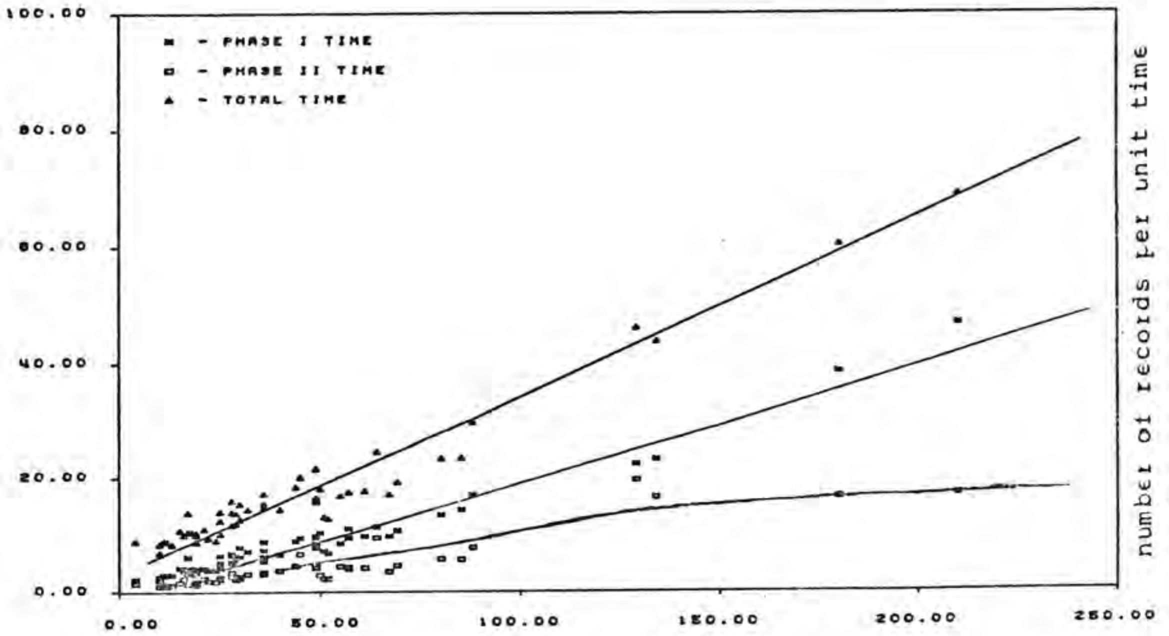
The data in Table 3.13-2 and 3.13-3 are presented in figures 3.14-1 (for CHEMDATA) and 3.14-2 (for RIVDATA) in graphical form. The linear relationship of the various time functions against number of CHEMDATA records processed may be clearly discerned. In contrast the figure for RIVDATA record additions is considerably more complex and shows no clear pattern. Reasons for these departures from a linear relationship are discussed in 7.122. The first peak (A) in Table 3.14-2 was due to the extensive sequential file searching required to ensure no duplication of data existed (see 2.132). To overcome this problem, the Datedat Protocol (see 2.125 and 2.126) was conceived and implemented. This system modification almost certainly accounts for the improvement in the statistics thereafter (note peak (B) is total time not Phase II time).

Table 3.14-1

Per record processing statistics for chemical data



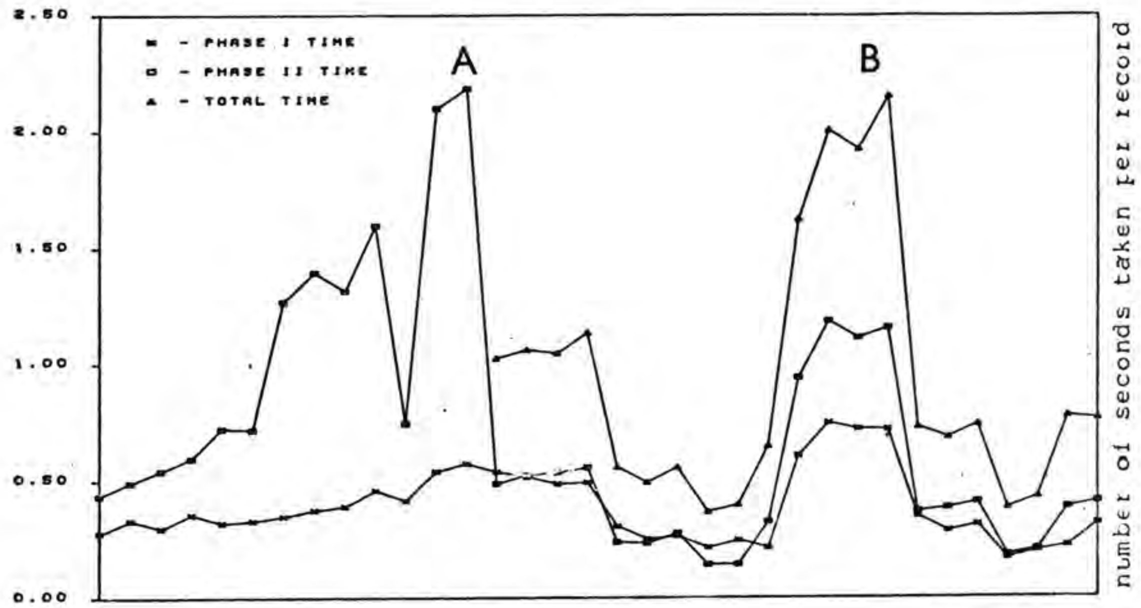
Processing sequence



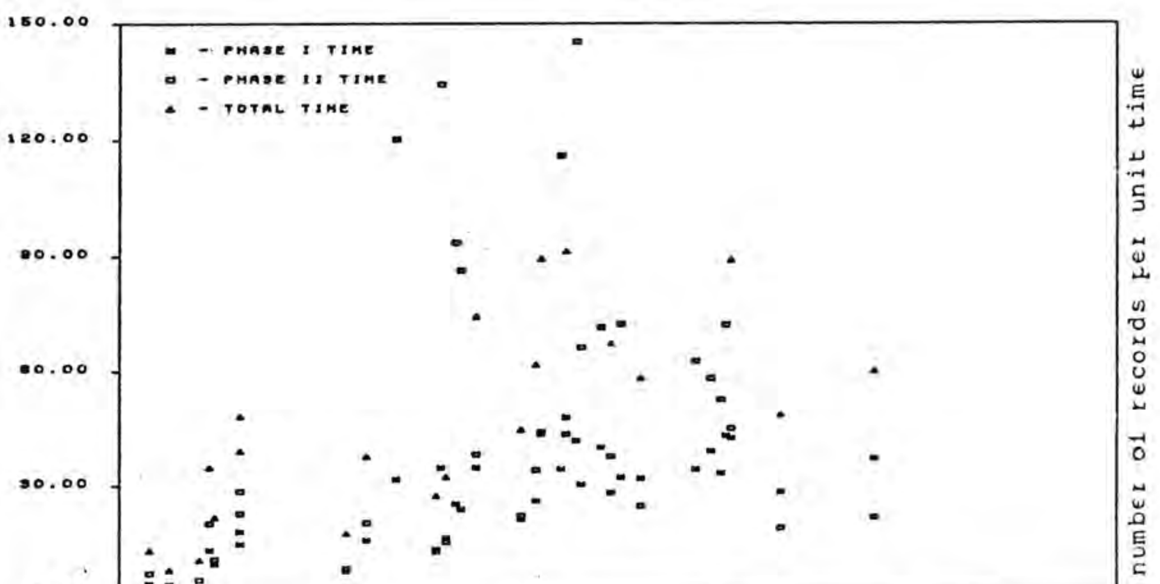
Total number of records added per run

Table 3.14-2

Per record processing statistics for biological data



Processing sequence



Total number of records added per run

3.15 System program development, security and integrity

Table 3.15-1 shows the approximate times taken to devise and implement the major protocols. In each case considerable development work has gone into the implementation which is not necessarily shown. Furthermore, the longest phase - the design and planning phase was very difficult to quantify accurately. A more detailed breakdown of the implementation of the Datedat Protocol is pesented in Section 3.22. This was the only protocol which was studied in any detail from the point of view of development and implementation.

Table 3.15-1

Protocol implementation statistics

Protocol	Design	Programming	Testing	Update	Documentation
SPECIESLIST	3 weeks	1 week	1 week	1 week	2 weeks
REACHLIST	2 weeks	1 week	1 day	2 days	1 week
CATLIST	1 week	1 day	15 mins	-	-
RIVDATA	20 months	2 months	1 month	3 weeks	?
CHEMDATA	6 months	1 month	1 month	-	?
SPEDAT	4 weeks	1 week	3 days	4 days	?
SARDAT	3 weeks	1 week	1 day	-	?
DATEDAT	1 week	3 days	2 days	-	?

Both the integrity and security of the system are dependant on the excellent facilities afforded by MTS. However, in addition five magnetic tapes contain copies of all the data and programs structured in a grandfather - father - son system. Archive copies of all raw datasets are kept on tape and other media in case of system failure. One copy of the data and programs is kept separately in Durham in case of a catastrophe occurring in the Newcastle tape library.

3.2 Protocol mechanisms and the database approach

3.21 Introduction

The Recording System protocols form both encoding rules and an implementation specification (see 2.122). The implementation specification is in terms of the relationships with the other elements of the system. The exploitation and maintenance of this network of relationships constitute the SIEUR system. In contrast the database management approach (as opposed to the database system approach - see Section 1.13) would be to examine the relationships in the data and to express these in simpler terms. Although this approach was not the one implemented, such a visualisation and its terminology can be used effectively to describe the problems encountered and the solutions employed. Furthermore, use of such a visualisation of the data forms a theoretical framework which could be used for development of the system as described in Section 7.12. The simplest database management visualisation is the relational one (Section 1.13) which considers the data to be two dimensional tables or "relations" which have the following properties (Martin, 1975).

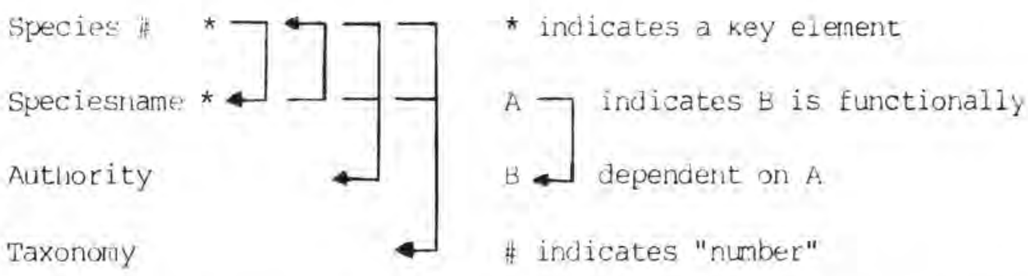
1. each entry in a table represents one data item; there are no repeating groups
2. they are column homogeneous; that is in any column all items are of the same kind
3. each column is assigned a distinct name
4. all rows are distinct, having an unique key consisting of one or more columns. Duplicate rows are not allowed
5. both the rows and columns can be viewed in any sequence at any time without affecting either the information content or the semantics of any function using the table.

This "first normal form" of the data (Codd, 1971) may then give rise to "second normal form" where each non-key column contains items which are wholly dependent on the entire key of a row. This usually implies further simplification of the data structure. The second normal form may then be further normalised to give the "third normal form", where each non-key column is independent of every other non-key column in that relation. It can be seen therefore that the goal of the relational approach would be to render the data into relations which are in third normal form and then to implement a method of handling these relations. Several such implementations have been reviewed (Section 1.13).

3.22 -LIST protocols

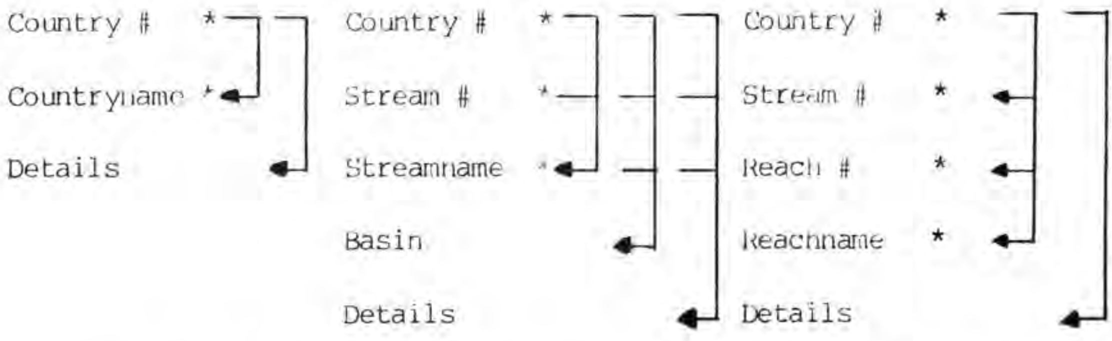
The -LIST protocols (Section 2.12) may be represented schematically by the following diagram (after Martin 1975).

For the Specieslist Protocol (2.122):



To convert this schematic representation into a relation read the items as column headings in a two dimensional table of entries where each separate row represents a species. An actual database management implementation may "see" each relation as several relations e.g. one for each non-key column. This in the example above would give rise to two additional relations one for authorities and one for taxonomic details. The currently implemented version of this protocol has no details concerning species authorities or taxonomy although provision has been made for their inclusion.

For the Reachlist Protocol (2.123):



Three relations are required in order that countries and streams can be added to the system without the necessity of inventing streams and reaches (that is making entries in the relation) to go with the country or stream being added. The SIEUR implementation overcomes this problem by considering that each country is a stream (that is it has been allocated a stream number) and that each stream within that country, and reaches on that stream, will be related in some numerical sequence. This mechanism has worked well with the limited number of countries and streams which SIEUR deals with.

For the Catlist Protocol (2.124):

```
Attribute number *─┐  Attribute number       * ─┐
Attribute name   *◄┘  Attribute value number * ◄┤
                      Attribute value name     ◄┘
```

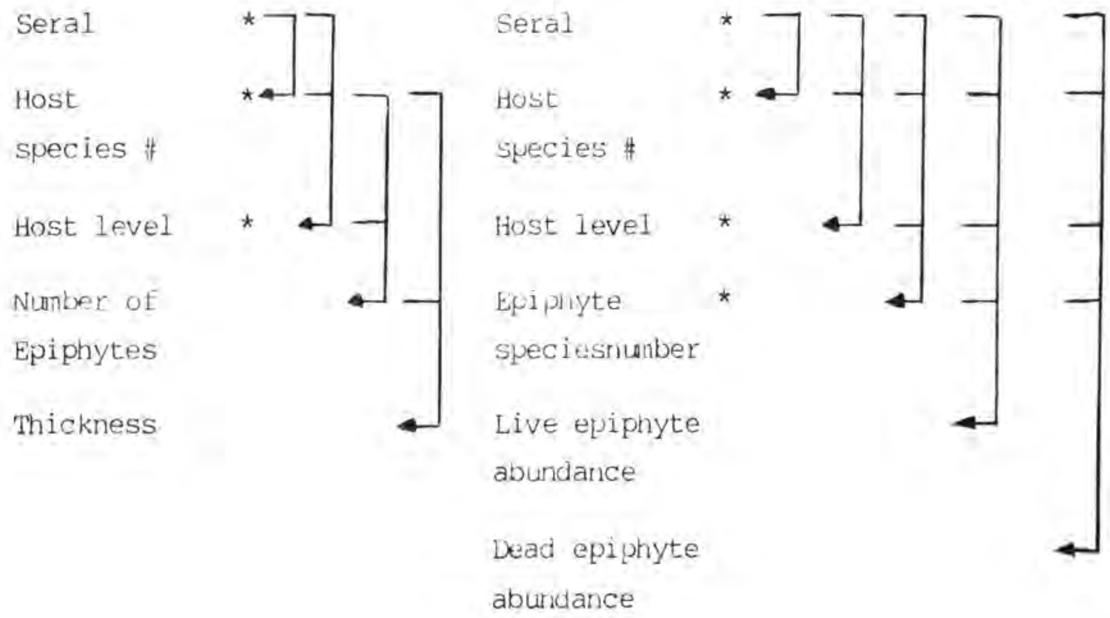
Two relations are needed to represent the data structure. The first defines what attributes are recognised, the second, for each attribute delimits the values and value details of every level of the attribute. In the SIEUR implementation, this mechanism is employed for the multistate attributes used in RIVDATA records but is not employed for the chemical data (7.125).

It can be seen that the -LIST protocols can be rendered easily into a relational visualisation and that this probably accounts for the ease with which the SIEUR implementation was made (Table 4.52-1). However, adding new -LIST protocols in the SIEUR implementation requires new program code to be written, tested and implemented (see 3.321). In contrast in a database management implementation, adding a new protocol would merely require defining the relational structure of it and adding the definition to the system.

3.23 -DATA protocols

The -DATA protocols consist of two schematic types, those whose data have inherent structure in themselves and those which do not. The inherent structure is at the sub-record level. Thus the epiphyte or morphological form data (2.125), recorded for one biological sample has inherent structure which is in itself data meriting storage. The data subrecords which do not contain structures may be mapped easily into the relational schema in ways identical to those proposed for the -LIST protocols. The data structural elements of a subrecord are processed in two stages. The first separates these elements from the simpler data elements. The second renders into data, the structures found. Sieur recognises only simple plex structures (see Section 1.13) which it processes by rendering them into an hierarchy which it stores in conjunction with the data. Thus the epiphyte species form the data and the information about which epiphytes are found growing on what hosts forms the structural information. Storage of more complex structures than this (see Section 1.13), and their retrieval, will need considerably more sophisticated mechanisms than those currently used.

To express the epiphyte structure in a relational form consistent with the descriptions of the -LIST protocols, two relations must be used. It may be seen that neither of these relations is in third normal form.



Some of the problems associated with this representation are:

1. need for additional structural concepts (e.g. host levels)

2. determination of keys. For example is host level a key ?
(Clearly yes since without it one would not be able to distinguish the same species acting as a host at different levels)

3. hosts and epiphytes are species, and updates, especially deletions, in the species relation must be reflected in the hosts and epiphytes relations

4. does deletion of an epiphyte necessarily mean the deletion of the species, in the species relation ?
(Clearly no unless the only record of the species is as that epiphyte, which implies more data (structural data) than exists)

5. maintaining dynamically the count of epiphytes in the host relation, the epiphyte number in the epiphytes relation and the host level in general.

Because of these problems and the major problem of recognising the initial structure in the data, both the current implementation and a database management implementation are complex. An immediate result is that data update within SIEUR of these structural data is only possible in a laborious manual fashion.

3.24 Problems encountered in using the Recording System protocols

Although the design of the recording system protocols were influenced by thoughts of how they could be implemented, few restrictions were placed on their design. In performing the implementation the more important problems which were identified were:

"What is a stream and conversely what is a bank?"

"What is the optimal size of the stream sampling unit?"

"How do we deal with diatom species which have to be cleared before full binomial identification can occur?"

"How do we deal with variable chemical detectable limits?"

3.3 Current SIEUR limitations

3.31 Introduction

The SIEUR system has both theoretical and physical limitations. The theoretical limitations were introduced by the nature of the protocols, the design of the programs and the limitations of PL/1 and MTS. The physical limitations of disk space and computer size and complexity were for the most part expressed by limitations in the MTS operating system. Thus disk size limitations were exercised through the MTS file handling mechanisms.

Because of the difficulty of separating operating system limitations from real physical limitations - the term "absolute physical limitation" will be restricted to the limitations imposed on the system by MTS in terms of absolute file sizes, absolute memory sizes and relative efficiencies of program execution. The trade off between reasonable execution time and SIEUR system complexity was a physical limitation which was relieved or aggravated as the NUMAC system changed or got more heavily loaded with users.

The term "physical limitation" will be restricted to those actual maximum sizes which the operating system will allow SIEUR to expand to; these are described in Section 3.33.

3.32 Theoretical limitations

The optimal sizes for files are well within the absolute limits as indicated in Table 3.32-1. For example the absolute maximum line length for an MTS line file is 32767 bytes, however, a length of greater than 1000 bytes incurs time penalties when input or output of it must occur. Table 5.22-1 attempts to summarise some of these optimal values for MTS files, although it must be stressed that these are conservative estimates based on no more than rules of thumb.

The file limitations were the major limitation on the SIEUR system. Many of the protocols were designed to exploit the advantages of the MTS file system (see 3.1); however, they also take on the disadvantages. None of these disadvantages has as yet severely affected the system; however, a tenfold increase in the data may cause some of the limitations to become critical.

The protocols themselves posed several limitations on the system. The complexity of the system was a reflection on the need to ensure that data

Table 3.32-1

Optimal MTS line file characteristics

Quantity	Stated limit	Actual limit	Optimum	Value used	Notes
Start line number	-99999.999	-2147483.647	-	-99999.000	-
End line number	99999.999	2147483.647	-	99999.999	-
Number of lines	~20 m	~4000 m	-	-	(i)
Length of lines	32767	32767	<4094	-	(ii)
Number of buffers	100	100	100	-	(iii)
Expansion factor	32767 32767	<120 <100%	- 20%	20 p -	(iv)
Size in pages	32767	3500	see notes	-	(v)

Notes:
(i) Optima are dependant on file and line length.
(ii) Each line consumes 2 additional bytes to store its length 4096 is the physical record length, records longer than this need 6 additional bytes per physical record.
(iii) Files larger than 100 pages incur some buffer problems.
(iv) Top line refers to absolute expansion sizes, bottom line to relative expansion sizes. In either case expansion is highly dependent on size - a large file is best expanded using absolute size expansion factors, small files using relative expansion factors (which are measured in pages).
(v) Optima are dependent on line length, and number of lines. Files with lines of less than 4095 bytes per line occupy one page per line and are usually optimal in terms of size. (large files are >100 pages, small files <80 pages)

(all notes assume the IBM 3330 type disk. MTS cannot support one single file spread over several disks (of any type). There is no facility within MTS for communication with the line directory, making the application specific optimisation of files impossible.)

integrity was maintained; however, it also placed much computer processing between the data and the system user which was not necessarily desirable because it made system expansion more difficult than it needed to be. The addition of a new protocol becames a major step unless it already fitted into an existing protocol type. Table 3.32-2 summarises the various parameters associated with the addition of the Datedat Protocol mentioned in Section 3.14 to the other existing auxilliary file protocols.

Table 3.32-2

Estimated computer resources used in adding the DATEDAT file to the auxilliary file protocols

Item	Elapsed real time (days)	CPU time (seconds)	Computing costs ($)	Overlap
Design	2	50	8.40	Pro & doc
Programming	1	4	2.4[illegible]	Des & tes
Testing	1	32	9.60	Pro & upd
Update	2	10	2.[illegible]6	Tes & doc
Documentation	1	-	-	Des, pro, upd & doc

An estimate is presented in Table 3.32-3 of the addition of a plant analysis protocol to the existing -DATA protocols. This constitutes a new protocol type which would have to be fully implemented.

Table 3.32-3

Projected resource utilisation for the addition of a plant analysis protocol to the existing -DATA protocols

Item	Elapsed real time (days)	Elapsed Terminal time (mins)	CPU time (seconds)	Computing costs ($)	Overlap
Design	7-10	30	50	8-10	Pro & Doc
Programming	2	60	20	8-10	Des, Tes & Upd
Testing	3-4	240+	100-200	120-130	Pro & Upd
Update	5+	120+	90-120	10-15	Des, Tes & Doc
Documentation	2-3	120+	30	5	Des, Tes, Upd & Doc

In addition to the the protocols, the choice of programming language (or more correctly the compiler used) and the structures of the programs themselves contributed to the limitations and the lack of speed of the system. The speed at which the system operates is a reflection of a series of decisions concerning user or programming convenience. For example, for user convenience, the decision was made to maintain all data (except data structural elements) in character form needing translation to internal floating point representation every time those data were required by the system. This decision was made more difficult because the programming language chosen, namely PL/1, was known to have a slow and complex translation mechanism. Similarly various decisions were made which resulted in only one data record at a time residing in main memory - a decision which further sacrificed time efficiency, in this case in favour of space efficiency and convenience.

The SIEUR program documentation must be referred to for more detailed discussions of the program limitations and more complex examples of protocol to file to program interactions.

3.33 Total cost and current physical limitations

The total computing cost of the system and its development is summarised by protocol in Table 3.33-1

Additional costs were the purchase of tapes, disks and manuals of various kinds.

The current physical limitations are summarised in Table 3.33-2.

These limitations are somewhat arbitrary owing to the ease with which for example more disk space may be obtained. This ease of expansion is summarised also in the table. Modifications to the MTS system, which occur from time to time, allows major re-evaluation of values quoted in the table, for example with the concept of a "project identifier" which has been proposed recently G.R. Eadie (pers. comm.), sign-on identifiers become meaningless.

Table 3.33-1

Estimated total costs of setting up and maintaining the SIEUR system

Protocol or Subsystem	Total development costs (NUMAC $)	Total current filespace (pages)	Other real costs (sterling)
-INFO protocols (including data input)	6000	50	20.00
-DATA protocols (including data input)	12000	350	360.00
Testadd Subsystem	9000	20	-
Query Subsystem	7000	15	-
Update Subsystem	4000	10	-
TOTALS	35000	440	380.00

1. The value of the NUMAC $ has fluctuated considerably over the period of the study. Multiplication by 0.5 gives an approximate sterling estimate.
2. Filespace estimates are for the basic system with no allowance for auxilliary files, dictionaries, system backups or logs or additional miscellaneous files such as "help" files.
3. Real costs exclude labour.

Table 3.33-2

Current physical limitations of the SIEUR system (June, 1980)

Item	Current number	Total size	Ease of expansion	Notes
Disks	1	8 k pages	nil	(i)
Tapes	8	30 k pages	3-4 days	(ii)
Filespace	4240 pages	3 k pages	1-2 days	(iii)
Ration	$750/week	-	-	(iv)
Signon ids	3	-	-	(v)
Availability	8 hrs/day	-	-	(vi)

(i) IBM 2314 disks are no longer supported by NUMAC.
(ii) Tapes are easily purchased through NUMAC.
(iii) Special requests must be made for permanent public filespace (> 120 pages). Databases are considered a special case and extra filespace is usually easily acquired.
(iv) Ration is increased by special request. The normal ration is 360 units per week. Increase for limited periods is possible.
(v) One id holds the database. The other two are used for possible development work.
(vi) This fluctuates on a day by day basis according to published timetables. A special service is sometimes available at weekends - for 5-6 hrs.

3.4 Acid study considerations

3.41 Introduction

The SIEUR query language was used as the starting point for all the analyses which are presented. The special programs which were written to manipulate the data could easily be built into the system as integral parts. This would however require some work to ensure that they were general enough to be used in all conditions.

This section summarises the costs in performing these analyses as a guide to the costs that would be expected if similar analyses were performed. An hidden element in these presentations is the development of the technique, be that in terms of the query specifications or in terms of the program development. All the results presented in this section are approximations based on single runs. No attempt has been made to interpret them statistically or to remove any bias due to runs being at different times of day.

3.42 Average query times and costs

The types of query may be classified loosely into 'chemical queries', 'biological queries' and 'mixed queries' involving both chemical and biological variables.

Examples are:

a. pH-fld <= 4.0

b. SPECIESNUMBER = 102069

c. pH-fld <= 3.0 & SPECIESNUMBER = 102069

The first involves no auxilliary index files and requires a full search of the chemical dataset, and is therefore dependant on the dataset size. The second is a fully indexed search of the species index file. The third, is optimised by the system such that the indexed biological query is performed first, expansion then occurs and redundant chemistries are discarded in a final pass over the expansion chemistries. A summary is given (Table 3.42-1) of various resource utilisations when the queries above were presented to the query system.

Table 3.42-1

Computing resources used in performing different query types
(on (autoexpand = on) ; off (autoexpand = off))

Type	Query Variable	Query Constant	Item measured	= on	= off	¬= off	> off	>= off	< off	<= off
Simple Chemical search	pH-fld	4.Ø	T (seconds) n (records)	17.7 3	37.2 2	37.6 1387	64.Ø 1221	110.0 1389	37.3 26	37.5 28
Simple indexed 1 search	SPN	2312Ø2	T n	7.7 5	1.2 3	- -	- -	- -	- -	- -
Simple indexed >1 search	Genus	1Ø19	T n	- -	- -	- -	47.4 1Ø7	- -	- -	- -
Mixed	SPN (=) pH-fld	231ØØ2 4.Ø	T n	- -	9.3 5	- -	9.4 7	12.2 12	11.3 Ø	15.6 Ø

It was soon noted that the best way to perform an analysis was to generate a stack using an indexed variable where possible, and then to expand that stack explicitly. This also means that there was much greater control over the expansion step which could be used in many ways. It was noted that simple queries were better than complex multiple queries, and that in general the biological data were easier to access because of the inherent indexed nature of the speciesnumber concept. Multiple queries of the biological data were usually counterproductive since it was frequently not obvious that queries were being specified erroneously. For example:

Phylum-pair = 1Ø & genus-pair = 2Ø ...1

is not equivalent to

Genusnumber = 1Ø2Ø ...2

because the query system has no way of knowing whether the two query atoms are linked or not. Hence query (1) would retrieve all samples where there existed a species in the phylum 1Ø, and a species in any genus 2Ø. Thus a sample with only 1Ø2ØØ1 in it would be retrieved as would any with 1Ø1Ø1Ø and Ø12Ø5Ø, which may or may not be what was intended.

Special attractions of the system were that it was verbose enough to allow the hard copy trace of a session to give complete documentation in itself; and the stack storage mechanism which allowed various subsets to be stored efficiently.

The many disadvantages of the system were due usually to the mechanisms which were not implemented. These ranged from more comprehensive optimisation techniques in the query system to the absence of cross links to the TAXONINFO and REACHINFO files. Several minor program bugs were encountered and corrected, and many possible directions for expansion of the system were identified.

The cost of setting up the acid dataset was not assessed since it was considered not necessary to maintain separate records for the addition of acid data.

3.43 Data processing and file storage costs

Data processing and file storage costs may be separated into the additional programs needed to manipulate the acid data and the cost of using the various statistical and other packages. Analyses are presented in terms of performing one particular type of statistical analysis on all packages. This does not take into consideration the known strengths and weaknesses of the packages tried.

Presented first in Table 3.43-1 are the estimated additional resources expended in writing, testing and running the additional programs used by the acid data analyses. These are presented in terms of one example program which was used to plot the species histograms presented in 4.5.

Present in Table 3.43-2 is a summary of the resources expended running a principal component analysis on the four different packages, 1 SPSS, 2 MIDAS, 3 OSIRIS and 4 CLUSTAN.

The four packages refer to different matrices as the Principal components matrix. CLUSTAN and MIDAS refer to the characteristic vectors (eigenvectors) extracted from the correlation matrix with unities in the main diagonal as the principal component matrix. SPSS and OSIRIS produce a matrix in which the eigenvectors are multiplied by the square root of the corresponding eigenvalue. This matrix is referred to as the "factor matrix" and it is implied that this is the principal component matrix although no mention is made of a principal component matrix as such. OSIRIS provides the eigenvectors and eigenvalues in addition to the factor matrix referring to them as the characteristic vectors and characteristic roots respectively. Throughout this thesis the term "principal component matrix" will be used to refer to the "factor matrix" of OSIRIS terminology which is produced from the correlation matrix with unities in the main diagonal.

When comparing the results for the different packages it was noted that

Table 3.43-1

Estimated computer resources used in setting up and using an acid data manipulation program involving plotting

Item	Development	Testing	Running (cost of 1 run)	Update	Comments
Filespace	4 pages	6 pages	2 pages	2 pages	(iii),(ii)
CPU time	150 secs	150 secs	25 secs	50 secs	(iii)
Elapsed time	5 days	1 day	~12 mins	1 day	(iii)
Plotting time	50 mins	40 mins	~20 mins	20 mins	(i),(iii)
Total cost	15	10	< 1	5	(iv)

(i) Plotting time is batch plotting of seven plot frames.
(ii) Filespace estimates do not include space required by data.
(iii) Development and testing times are difficult to separate exactly.
(iv) Cost is quoted in NUMAC accounting units which correspond roughly to pounds sterling.

although the weights for a component were usually numerically identical (allowing for different numbers of significant digits) the signs were frequently reflected (i.e. for a whole component multiplied by -1). Harman (1976) points out that the sign is arbitary and may, without affecting the solution, be reflected at will. This is because the ellipsoid which represents the component can, without loss of generalisation, be rotated about its midpoint in any axis (the midpoint will be at 0.0 because of standardisation.)

Table 3.43-2

Principal Component Analysis - resource utilisation

Item	SPSS	MIDAS	OSIRIS	CLUSTAN
Time (CPU seconds)	2.81	1.29	4.50	2.83
Virtual memory (pages)	68	110	75	42
Elapsed time (mins)	4	2	7	5
Pages printed	12	2	20	6

(The recommended default parameters with regards output generated and statistics performed were used)

3.44 Software use limitations and considerations

All the SIEUR software is fully available to any research worker. The whole system is heavily MTS dependant because of the intimate use of the file system and would involve considerable modification if it were to be moved in its entirity to another computer running under a different operating system. However, many parts of the system, for example the additional data manipulation programs, may be moved to other computers or operating systems with little difficulty.

Many of the program products used by the system and in performing the analysis are NUMAC specific and use of them is restricted - although it is likely that similar packages exist at most computer installations (for example a "sort" utility or general statistical program). In addition NUMAC is charged for use of certain packages like SPSS and CLUSTAN and further restrictions are placed on who may use them. For example SPSS costs $3000 per annum and may be used only by registered academic users. An additional problem is that the source code for many packages is not distributed. If problems do arise, they usually take a long time to put right (especially CLUSTAN).

Chapter 4 Acid samples

4.1 Data selection

4.11 Introduction

Chapter 4 presents the results of the analysis of the acid data present in SIEUR (4.1) and describes these data in chemical and biological terms using simple univariate (4.2), bivariate (4.3) and multivariate (4.3 and 4.4) statistics. Finally in 4.5 an attempt is made to relate the occurrence of species to the environmental pH. At several points in this chapter discussion material will be introduced where this helps to clarify the results. These additional points, which are secondary to the mainstream of this thesis will not be taken up further.

The aim of 4.1 is to present the results of the analyses of the 269 water chemistries of the acid samples dataset which were obtained on issuing the simple query:

pH-fld <= 4.0 ...1

which assumes that all lab pH data were ignored.

In generating and examining this basic query, several potential pH ranges, and the ranges of other variables were examined for the whole dataset. These results are included in this section. The relevant analyses of the 125 water chemistries of the acid species dataset found on expansion "bydate" of the 269 acid samples into aggregate biologies are also presented. The results of the limited analyses of the 509 water chemistries of the acid sites dataset found on expanding the acid sample dataset "bysar" are given in 5.1.

The underlying problem with the use of the acid samples dataset is that in taking a specified range of data, bias is introduced immediately into the study which may render statistical analysis invalid or very difficult. Furthermore the multiple samplings of some reaches will be given equal weight when compared with the single samplings of other reaches. These problems are ignored in 4.1, which aims to present the results with only some interpretive comments.

4.12 Other pH ranges examined and stacks produced

The other pH ranges examined in this analysis are compared and summarised in Table 4.12-1. Presented in Table 4.12-2 are distribution

Table 4.12-1

Comparison of various query results stacks

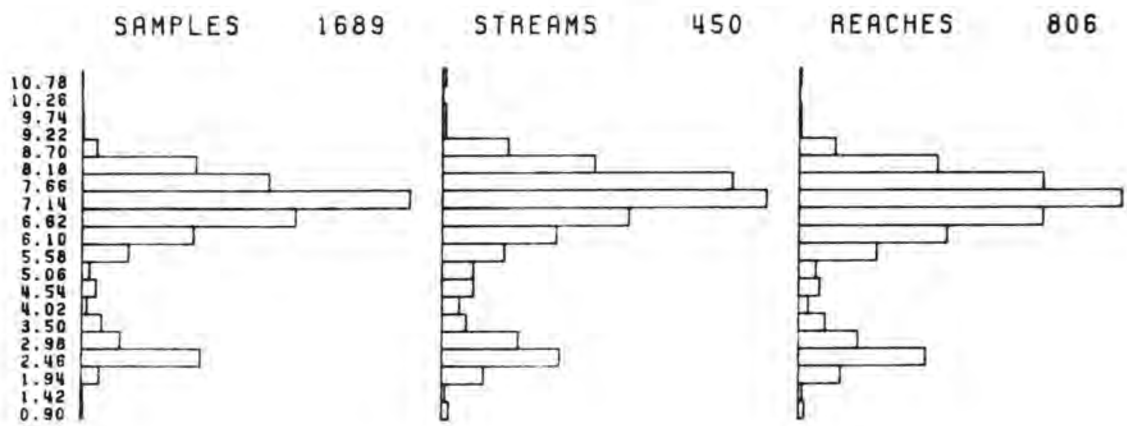
pH range	chemistries (:)			1 by 1 samples (*)			aggregates (#)		
<=2.0	6	6	4	13	3	2	6	5	4
<=3.0	221	95	45	303	83	45	124	82	41
<=4.0	284	115	59	319	95	52	136	93	50
<=5.0	320	128	71	517	100	57	138	97	54

The first figure in each triple is the number of samples, the second the number of reaches from which the samples came and the third the number of streams from which the reaches came.

histograms of water chemistry samples, reaches and streams in terms of their pH across the whole of the SIEUR dataset.

Table 4.12-2

Distribution of chemical samples, streams and reaches in terms of their pH values



The over representation of acid streams can be seen in Table 4.12-2. This pH profile of the SIEUR dataset will change with time as more non-acid data are added to the system.

4.13 Statistics of variables in acid stacks

Table 4.13-1 summarises the ranges, means and standard deviations of the variables associated with the field pH <= 5.0, <=4.0 and <= 3.0 acid stacks. In generating Table 4.13-1 values at the detectable limit were ignored (see 2.22). The intervariable statistics for this dataset are presented in Section 4.32 as part of the principal component analysis results.

Table 4.13-1

Comparison of the physical and chemical parameters drawn from SIEUR (June 1980) when the queries pH <= 3.0, <= 4.0 and <= 5.0 were used to define the datasets

Key

min max
mean
(n) SD
(all elements as mg l^{-1})

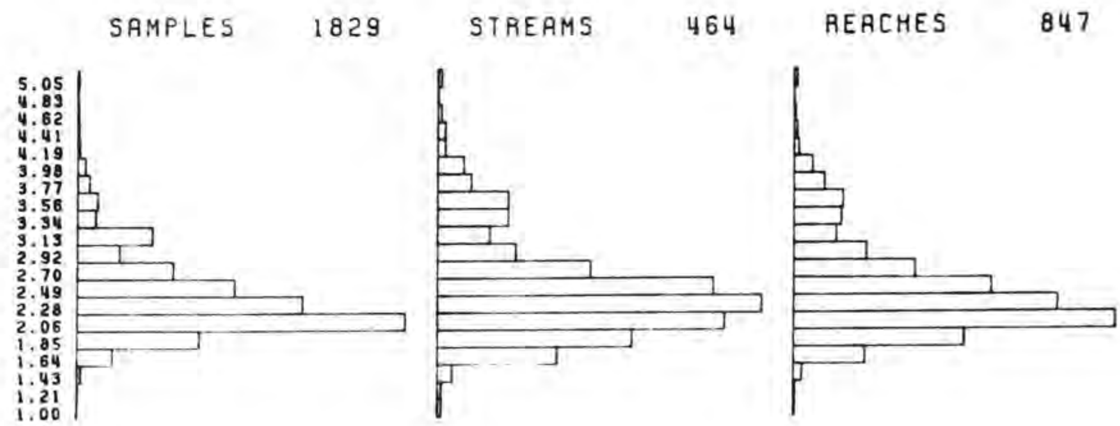
	pH <= 3.0	pH <= 4.0	pH <= 5.0
O.D.-420	0.002 0.99 0.11 (204) 0.18	0.001 0.99 0.10 (252) 0.17	0.001 0.99 0.09 (281) 0.16
cond ($\mu S\ cm^{-1}$)	600.000 80000.00 5615.14 (198) 16135.46	63.700 80000.00 4795.99 (245) 14598.95	34.000 80000.00 4317.69 (274) 13872.20
pH-fld	0.900 3.00 2.63 (212) 0.29	0.900 4.00 2.82 (269) 0.46	0.900 6.30 3.00 (297) 0.73
acidity ($mg\ l^{-1}\ CaCO_3$)	10.600 64000.00 4037.13 (202) 6841.74	10.600 64000.00 3248.20 (256) 6264.76	8.300 64000.00 3016.76 (276) 6089.43
Eh-fld	440.000 560.00 512.50 (12) 39.95	382.000 560.00 493.13 (15) 55.60	260.000 560.00 441.72 (22) 97.90
Na	0.270 3550.00 162.35 (219) 335.65	0.270 3550.00 148.42 (279) 301.63	0.270 3550.00 137.10 (315) 293.47
K	0.030 63.00 6.32 (219) 11.55	0.030 123.00 8.91 (279) 18.01	0.030 484.00 9.87 (315) 32.01
Mg	0.620 2340.00 226.58 (219) 312.51	0.250 2340.00 192.48 (280) 284.89	0.250 2340.00 173.54 (316) 274.72
Ca	1.200 556.00 171.26 (219) 159.42	0.540 556.00 156.70 (280) 149.30	0.540 596.00 143.04 (316) 149.50
Zn	0.025 303.00 15.19 (219) 45.03	0.025 303.00 12.72 (280) 40.77	0.004 3610.00 22.76 (316) 206.08
Cu	0.002 536.00 13.54 (213) 58.69	0.002 536.00 10.89 (267) 52.66	0.001 536.00 9.73 (299) 49.86
Mn	0.014 544.00 34.03 (219) 60.03	0.014 544.00 27.88 (280) 54.45	0.014 544.00 25.72 (316) 54.11
Fe	0.990 23000.00 704.93 (219) 2068.76	0.040 23000.00 558.37 (280) 1850.05	0.030 23000.00 494.83 (316) 1750.15
Al	0.320 3130.00 149.79 (219) 342.48	0.270 3130.00 120.70 (279) 308.38	0.015 3130.00 107.69 (313) 293.48
Pb	0.001 4.95 0.20 (220) 0.47	0.001 260.00 1.10 (282) 15.48	0.001 260.00 1.00 (318) 14.58
Co	0.001 20.00 0.98 (209) 1.93	0.001 20.00 0.81 (261) 1.76	0.001 20.00 0.79 (276) 1.73
Ni	0.020 50.40 2.24 (209) 4.53	0.004 50.40 1.84 (262) 4.12	0.004 50.40 1.74 (278) 4.02
PO_4-P	0.010 76.00 1.10 (185) 6.47	0.010 76.00 0.92 (223) 5.90	0.010 76.00 0.86 (241) 5.68
NH_4-N	0.030 10.80 1.56 (185) 1.94	0.030 10.80 1.39 (224) 1.81	0.009 10.80 1.31 (246) 1.75
NO_3-N	0.050 4.00 0.59 (180) 0.58	0.020 4.00 0.63 (225) 0.61	0.010 4.00 0.62 (246) 0.60
SO_4-S	68.000 8580.00 1428.70 (184) 1439.25	0.840 8580.00 1224.86 (228) 1361.48	0.840 8580.00 1124.12 (250) 1340.29
Cl	10.500 1800.00 53.39 (202) 131.22	7.680 1800.00 49.16 (251) 119.03	3.600 1800.00 45.87 (280) 113.15
Si	0.500 114.00 36.17 (194) 18.61	0.100 114.00 32.84 (240) 19.57	0.025 114.00 29.82 (268) 20.58
O_2-satn (%)	0.000 106.00 74.71 (172) 21.47	0.000 106.00 76.53 (207) 20.07	0.000 106.00 76.88 (214) 19.84
temperature (°C)	0.000 49.00 11.87 (208) 7.01	0.000 49.00 11.37 (265) 6.95	0.000 49.00 11.21 (292) 6.80

4.14 Other variable ranges examined as potential stack generators

In Tables 4.14-1 and 4.14-2 are presented the results when variables other than pH were analysed as potential dataset generators. The major potential variables were acidity, conductivity and redox potential. Redox potential was rejected because of the considerable quantity of missing data associated with it. Alkalinity figures were ignored as inappropriate.

Table 4.14-1

Distribution of SIEUR chemistry samples, streams and reaches in terms of Conductivity



To generate approximately 250 sample chemistries using Conductivity or Acidity the values pConductivity ($\log_{10}$ conductivity) >= 2.02 and pAcidity ($\log_{10}$ acidity) >= 2.55 would have to be used. These yield almost identical numbers of streams and reaches to the pH <= 4.0 level with considerable similarity in the actual streams and reaches retrieved. The degree of correlation between the three potential stack generating variables is summarised for the total SIEUR dataset in Table 4.14-3.

Many other methods of generating the initial acid dataset were examined. These were based on both chemical and biological criteria and used complex inclusion formulae where sites were included if on average their pH was below various levels or their modal leval of other parameters was above specified criteria. None of these methods was adopted because they were considered to be too subjective or based on criteria which could not be reconciled with a generalised database approach (See 7.312).

Table 4.14-2

Distribution of SIEUR chemistry samples, streams and reaches in terms of aciditity

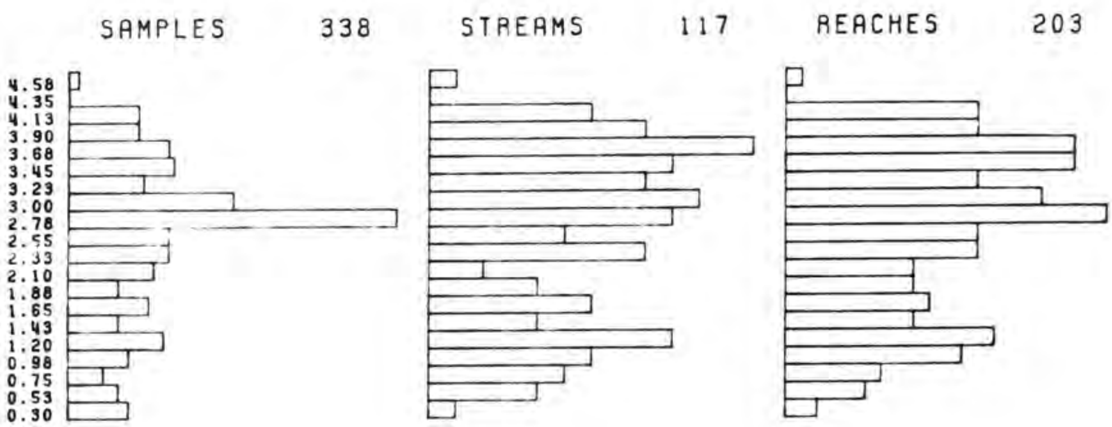


Table 4.14-3

Correlations between potential stack generating variables accross the whole SIEUR database

	pH	conductivity	log_{10} conductivity	acidity	log_{10} acidity	alkalinity	log_{10} alkalinity
pH	X X X X	-.2905 (1243) -10.69 ***	-.6558 (1243) -30.60 ***	.4953 (575) 13.64 ***	.6542 (552) 20.29 ***	-.3358 (382) -6.94 ***	-.8374 (315) -27.10 ***
conductivity	6.46 1.9 1198. 6725.	X X X X	X X X X	.3021 (620) 7.88 ***	.2521 (613) 6.44 ***	.7453 (331) 20.28 ***	.6090 (294) 13.12 ***
log_{10} conductivity	6.46 1.9 2.57. 0.53.	X X X X	X X X X	.6695 (620) 22.41 ***	.6487 (613) 21.07 ***	.4994 (331) 10.46 ***	.7910 (294) 22.09 ***
acidity	7.36 0.7 75.9 67.9	326.02 586. 80.1 68.3	2.36 0.32 80.0 68.3	X X X X	X X X X	- - - -	- - - -
log_{10} acidity	7.36 0.7 1.65 0.5	326.02 586. 1.71 0.47	2.36 0.32 1.71 0.47	X X X X	X X X X	- - - -	- - - -
alkalinity	3.68 1.7 2584. 5790.	2728.0 1197. 2747. 5960.	3.20 0.49 2747. 5960.	- - - -	- - - -	X X X X	X X X X
log_{10} alkalinity	3.68 1.7 2.629 1.03	2728.0 3397. 2.66 1.04	3.20 0.49 0.49 1.04	- - - -	- - - -	X X X X	X X X X

(Top half diagonal matrix contains correlation coefficient, (n), "Z" statistic and significance (*** = 99 %). Bottom half diagonal contains maxima and minima values for the horizontal then vertical parameters in each case)

4.15 Raw data table - data subsets

A raw data table (Table 4.15-1) has been produced for the pH <= 4.0 stack using the fraction substitution technique outlined in 2.222. The raw data table (Table 4.15-1) may be represented schematically by Table 2.24-1 which also presents the other biological and chemical datasets. Table 4.15-2 presents a summary of these various datasets in terms of missing data and other parameters.

Table 4.15-1

Chemical and physical parameters of water chemistry samples to be found in SIEUR (June 1980) identified by the query pH-fld <= 4.0

(Units as in Table 4.13-1 p 93)

Stack is pH-fld <= 4.0 (269 dataset)

Sheet A.1

serid	samnumber	date time file	O.D.-420	cond.	pH-fld	acidity	Eh-fld	Na	K	Mg
000456:	000152.10	19750812144ØCD.AK	0.1570	2450.0000	2.30	901.00	540.00	14.800	3.8000	126.5200
000457:	000152.20	19750812131ØCD.AK	0.0060	2800.0000	2.70	1121.00	510.00	15.400	3.9000	153.5000
000458:	000153.01	19750812152ØCD.AK			2.60	1121.00	444.00			
000459:	000154.01	19750812144ØCD.AK	0.0220	3100.0000	3.25	1150.00	405.00	14.500	5.3000	206.0000
000492:	000066.01	19751208172ØCD.AM	0.0210	1670.0000	2.70	542.00	560.00	2.100	0.3700	11.1000
000493:	002003.01	19751007094ØCD.AM	0.0320	2620.0000	2.70	1558.00	450.00	11.300	1.8500	72.5000
000494:	002004.01	19751007160ØCD.AM	0.0330	2650.0000	2.70	1572.00	455.00	6.600	1.6000	72.5000
000495:	002006.01	19751007172ØCD.AM	0.2400	4900.0000	2.60	4896.00	550.00	5.900	0.2900	107.5000
000496:	002007.01	19751007140ØCD.AM	0.2600	5900.0000	2.55	6052.00	540.00	4.600	0.3200	157.5000
000497:	002007.02	19751007130ØCD.AM	0.2700	5900.0000	2.55	5740.00	540.00	5.100	0.4300	185.0000
000498:	002007.03	19751007141ØCD.AM	0.2310	5800.0000	2.50	5750.00	545.00	5.000	0.2700	162.5000
000505:	002014.01	19751007105ØCD.AM	0.1170	4450.0000	2.60	4202.00	508.00	5.600	1.2900	230.0000
000506:	002015.10	19751007100ØCD.AM	0.0100	840.0000	3.25	290.00	460.00	10.500	2.4000	21.7000
000605:	001004.01	19770209124ØCD.AQ			2.90			710.000	53.0000	448.0000
000608:	009001.01	19770514150ØCD.AR			3.05	340.00		10.100	10.0000	96.0000
000612:	009004.01	19770514150ØCD.AR			2.85	800.00		28.500	2.6400	70.2000
000613:	009005.01	19770515113ØCD.AR			3.30	350.00		37.800	0.4500	134.0000
000614:	009006.01	19770515140ØCD.AR			3.20	500.00		96.100	1.5200	124.0000
000615:	009007.10	19770517162ØCD.AR			2.52	1640.00		16.500	1.4500	175.0000
000616:	009008.45	19770611140ØCD.AR			2.40	420.00		5.500	5.5000	0.7400
000617:	009008.48	19770611160ØCD.AR			2.30	800.00		7.800	9.9000	0.7000
000618:	009008.50	19770611169ØCD.AR			2.40	810.00		8.800	10.5000	0.7400
000619:	009008.52	19770612110ØCD.AR			2.25			11.200	11.8000	0.8400
000620:	009009.20	19770611110ØCD.AR			3.10					
000621:	009009.20	19770611101ØCD.AR			3.10			151.000	40.4000	0.4000
000622:	009010.02	19770611115ØCD.AR			3.00					
000623:	009011.01	19770612070ØCD.AR			2.85	110.00		36.000	26.2000	0.7200
000624:	009012.01	19770612073ØCD.AR			2.60	230.00		61.200	35.0000	0.6200
000625:	009013.03	19770612080ØCD.AR			3.10	90.00		46.600	34.6000	0.6500
000628:	009016.99	19770623134ØCD.AR			4.00	20.00		134.000	39.6000	47.5000
000746:	003016.01	19780630130ØCD.AW			3.80			2.000	2.3000	12.5000
000747:	003016.03	19780630131ØCD.AW			3.70					
000757:	000213.01	19780531110ØCD.AX	0.0010	2320.0000	3.25			16.700	6.0000	115.0000
000758:	000214.21	19780531152ØCD.AX	<0.001	1400.0000	2.85			15.400	1.6600	22.3000
000759:	000215.01	19780531153ØCD.AX	0.1000	42000.0000	2.10			11.500	0.5500	125.0000
000760:	000215.15	[illegible]CD.AX	0.8500	4500.0000	2.20			11.600	0.6600	120.0000
000761:	000216.15	[illegible]CD.AX	0.2500	4300.0000	2.20			16.400	1.6200	74.5000
000762:	000216.50	[illegible]CD.AX	0.0500	1700.0000	2.40			17.700	1.7100	25.8000
000763:	000217.01	[illegible]CD.AX	0.0300	1500.0000	2.40			18.800	2.3000	59.5000
000764:	000218.01	[illegible]CD.AX	0.0050	900.0000	2.60			16.200	1.4600	28.3000
000765:	000220.03	[illegible]CD.AX	0.3230	8000.0000	0.90			102.000	8.7000	120.0000
000766:	000220.15	[illegible]CD.AX	0.1760	30000.0000	0.90			895.000	37.0000	14.1000
000767:	000221.05	[illegible]CD.AX	0.9880	31000.0000	1.41			47.600	4.5000	6.2000
000891:	000100.30	1975080817[illegible]CW.AHHA1	0.0050	760.0000	2.70	146.00	482.00	3.800	1.3500	11.1000
000895:	000100.01	1975080816[illegible]CW.AHHA1	0.0140	392.0000	3.90	74.00	382.00	2.900	1.3600	5.3000
001198:	000119.01	1974091919[illegible]CW.BK	0.1440	65.0000	4.00			3.300	0.3000	0.7600
001199:	000120.01	1974091919[illegible]CW.BK	0.1120	80.0000	3.90			3.700	0.3000	1.1000
002120:	000100.30	[illegible]CW.BBE			3.90					[illegible]
002224:	000127.01	1973070410[illegible]CW.BBB	0.0280	2000.0000	2.60	1390.00		12.300	0.5700	68.[illegible]
002225:	000127.02	1973070410[illegible]CW.BBB	0.0250	1800.0000	2.60	1428.00		12.600	0.7400	[illegible]
002226:	000127.03	1973070411[illegible]CW.BBB	0.0150	1800.0000	2.60	1196.00		13.200	1.9400	65.7000
002227:	000127.04	1973070411[illegible]CW.BBB	0.0040	2500.0000	2.90	1476.00		119.000	7.6000	152.[illegible]
002228:	000128.01	1973070810[illegible]CW.BBB	0.0420	4200.0000	2.90	2080.00		169.200	14.5000	551.[illegible]
002229:	000128.02	1973070810[illegible]CW.BBB	0.1010	7000.0000	3.00	3320.00		785.000	5.0100	783.[illegible]
002230:	000128.03	1973070811[illegible]CW.BBB	0.0410	5500.0000	2.90	2400.00		45.300	7.8000	6[illegible]2.[illegible]
002231:	000130.01	1973070610[illegible]CW.BBB	0.0140	17000.0000	2.30	13800.00		455.000	2.5500	920.[illegible]
002232:	000130.02	1973070610[illegible]CW.BBB	0.1200	7500.0000	2.30	3320.00		380.000	2.5000	715.[illegible]
002233:	000130.03	1973070611[illegible]CW.BBB	0.1700	17500.0000	3.00	13000.00		500.000	3.6000	324.[illegible]
002234:	000131.01	1973070612[illegible]CW.BBB	0.6000	3750.0000	2.70	1656.00		17.600	0.6700	165.[illegible]
002235:	000131.02	1973070612[illegible]CW.BBB	0.1400	3000.0000	2.70	1620.00		19.800	0.8200	1[illegible]
002236:	000131.03	1973070612[illegible]CW.BBB	0.0610	4500.0000	3.00	1650.00		40.000	1.8900	23.[illegible]
002237:	000132.01	1973091010[illegible]CW.BBB	0.0580	10000.0000	2.70	15650.00		581.000	15.6000	720.[illegible]
002238:	000132.02	1973091010[illegible]CW.BBB	0.0910	10000.0000	2.50	15520.00		392.000	1.4200	724.[illegible]
002239:	000132.03	1973091011[illegible]CW.BBB	0.2310	12000.0000	2.10	8565.00		232.000	1.0000	810.[illegible]
002240:	000132.04	1973091011[illegible]CW.BBB	0.0900	6500.0000	2.50	9200.00		395.000	1.3500	730.[illegible]
002241:	000132.05	1973091011[illegible]CW.BBB	0.0310	6500.0000	2.50	8000.00		355.000	1.4700	766.[illegible]
002242:	000132.06	1973091011[illegible]CW.BBB	0.0650	6500.0000	2.50	6560.00		402.000	3.4000	755.[illegible]
002243:	000132.07	1973091012[illegible]CW.BBB	0.0780	6000.0000	2.50	5360.00		403.000	3.3000	717.[illegible]
002244:	000133.01	1973091714[illegible]CW.BBB	0.0470	2700.0000	3.00	768.00		13.600	16.0000	36.[illegible]
002247:	000134.03	1973091511[illegible]CW.BBB	0.0460	6500.0000	2.70	2920.00		738.000	11.1000	350.[illegible]
002249:	000135.01	1973071810[illegible]CW.BBB	0.0610	13500.0000	2.20	2224.00		690.000	13.2000	144.[illegible]
002250:	000135.02	1973071810[illegible]CW.BBB	0.2600	18000.0000	2.00	16970.00		505.000	4.4000	980.[illegible]
002251:	000135.03	1973071811[illegible]CW.BBB	0.0910	11500.0000	2.10	16760.00		585.000	1.6500	1220.[illegible]
002252:	000135.04	1973071811[illegible]CW.BBB	0.1200	2500.0000	3.00			640.000	1.0800	122.[illegible]
002253:	000136.01	1973091110[illegible]CW.BBB	0.0830	14200.0000	2.10	4560.00		11.400	4.4000	285.[illegible]
002254:	000137.01	1973091111[illegible]CW.BBB	0.0430	4700.0000	2.50	1840.00		385.000	5.0000	[illegible]
002255:	000137.02	1973091111[illegible]CW.BBB	0.0420	4000.0000	2.50	2000.00		378.000	5.6000	206.[illegible]
002256:	000137.03	1973091111[illegible]CW.BBB	0.0410	4500.0000	2.60	2000.00		330.000	5.1200	210.[illegible]
002257:	000138.01	1973101310[illegible]CW.BBB	0.0460	1000.0000	3.00	740.00		8.400	8.4000	28.0000
002258:	000138.02	1973101310[illegible]CW.BBB	0.0070	1300.0000	2.90	740.00		8.000	8.0000	28.[illegible]
002259:	000138.03	[illegible]CW.BBB	0.0040	1500.0000	2.90	928.00		3.900	3.9000	2[illegible]
002260:	000139.01	1973091112[illegible]CW.BBB	0.4110	7500.0000	2.30	15200.00		213.000	[illegible]	[illegible]
002261:	000139.02	1973091112[illegible]CW.BBB	0.3100	30000.0000	2.10	64000.00		856.000	[illegible]	379.[illegible]
002262:	000139.03	[illegible]CW.BBB	2.6300	9000.0000	2.10	16000.00		3550.000	[illegible]	[illegible]
002263:	000140.01	1973091111[illegible]CW.BBB	0.0420	6500.0000	2.90	5700.00		650.000	10.3000	[illegible]
002264:	000140.02	1973091112[illegible]CW.BBB	0.0710	4500.0000	2.90	3680.00		440.000	3.4000	[illegible]
002265:	000140.03	1973091112[illegible]CW.BBB	0.0730	4500.0000	2.50	3200.00		395.000	3.5000	2[illegible]
002266:	000157.05	1973070611[illegible]CW.BBB	0.1000	20000.0000	1.50	20000.00		405.000	0.3100	1[illegible]
002267:	000157.06	1973070611[illegible]CW.BBB	0.1400	5500.0000	2.80	2200.00		77.000	1.5000	116.0000

Stack is pH-fld <= 4.0 (269 dataset)

serid	samnumber	date time file	O.D.-420	cond.	pH-fld	acidity	Eh-fld	Na	K	Mg
002268:	000157.07	197307061130CW.BBB	0.6500	12500.0000	1.80	10600.00		1250.000	9.2500	905.0000
002269:	000158.08	197309101200CW.BBB	0.1500	5000.0000	2.80	3760.00		180.000	21.2000	870.0000
002270:	000159.09	197309101230CW.BBB	0.0120	12000.2000	2.80	7360.00		278.000	52.8000	398.0000
002271:	000160.05	197307181130CW.BBB	0.4300	9000.0000	3.00	13800.00		530.000	33.0000	925.0000
002272:	000160.06	197307181130CW.BBB	0.1410	10000.0000	2.50	13800.00		550.000	29.0000	1125.0000
002273:	000127.01	197402141030CW.BBC	0.2300	2000.0000	2.60	1500.00		11.200	0.4200	77.0000
002274:	000127.02	197402141030CW.BBC	0.0160	1500.0000	2.60	1500.00		12.000	0.5700	80.1000
002275:	000127.03	197402141100CW.BBC	0.0070	1500.0000	2.60	1500.00		13.000	0.3100	82.0000
002276:	000127.04	197402141130CW.BBC	0.0060	1750.0000	2.70	1900.00		76.500	5.0000	150.0000
002277:	000128.01	197402101030CW.BBC	0.0110	3500.0000	3.00	4100.00		130.000	15.5000	390.0000
002278:	000128.02	197402101030CW.BBC	0.0690	3500.0000	2.90	3800.00		621.000	3.1400	678.0000
002279:	000128.03	197402101100CW.BBC	0.0770	4000.0000	2.75	7800.00		62.800	5.0000	545.0000
002280:	000129.01	197402111030CW.BBC	0.1400	2750.0000	2.60	6000.00		7.900	0.8100	29.0000
002281:	000129.02	197402111030CW.BBC	0.2460	3700.0000	2.40	8300.00		7.900	0.3900	76.0000
002282:	000130.01	197402121030CW.BBC	0.7750	7000.0000	2.40	12500.00		77.500	0.3900	325.0000
002283:	000130.02	197402121030CW.BBC	0.2190	3200.0000	2.80	5900.00		106.000	2.5000	110.0000
002284:	000130.03	197402121100CW.BBC	0.4400	4700.0000	3.00	11000.00		139.000	2.2000	240.0000
002285:	000130.04	197402121100CW.BBC	0.3820	4500.0000	2.60	15500.00		380.000	12.2000	535.0000
002286:	000131.01	197402121130CW.BBC	0.2760	3700.0000	2.60	8300.00		17.300	12.0000	153.0000
002287:	000131.02	197402121130CW.BBC	0.1870	3000.0000	2.70	6200.00		26.200	0.8900	120.0000
002288:	000131.03	197402121200CW.BBC	0.2000	3250.0000	2.70	5900.00		23.300	0.6600	130.0000
002289:	000132.01	197402201030CW.BBC	0.1170	3250.0000	2.70	8900.00		159.000	2.9000	340.0000
002290:	000132.02	197402201030CW.BBC	0.1120	3250.0000	2.70	4100.00		155.000	2.9000	330.0000
002291:	000132.03	197402201100CW.BBC	0.6050	10000.0000	2.50	2950.00		40.900	0.1400	475.0000
002292:	000132.04	197402201100CW.BBC	0.1220	5000.0000	2.50	8600.00		173.000	2.6000	368.0000
002293:	000132.05	197402201130CW.BBC	0.0500	4500.0000	3.00	3100.00		319.000	14.1000	575.0000
002294:	000132.06	197402201130CW.BBC	0.1450	4500.0000	2.80	5500.00		289.000	8.4000	595.0000
002295:	000132.07	197402201200CW.BBC	0.1340	4500.0000	2.80	4600.00		281.000	8.2000	575.0000
002296:	000133.01	197402201200CW.BBC	0.0400	4250.0000	3.00	2300.00		166.000	24.7000	620.0000
002297:	000134.01	197402121200CW.BBC	0.0910	3000.0000	2.60	4000.00		82.500	4.3000	110.0000
002298:	000134.02	197402121200CW.BBC	0.0880	3000.0000	2.60	4000.00		87.600	4.1000	133.0000
002299:	000134.03	197402121230CW.BBC	0.0260	4500.0000	3.00	4600.00		603.000	13.6000	310.0000
002300:	000134.04	197402121230CW.BBC	0.0970	3500.0000	2.80	3500.00		227.000	7.7000	193.0000
002301:	000135.01	197402221030CW.BBC	0.5720	8000.0000	2.90	14100.00		518.000	19.9000	934.0000
002302:	000135.02	197402221030CW.BBC	0.5280	7500.0000	2.90	13400.00		358.000	18.8000	885.0000
002303:	000135.03	197402221100CW.BBC	0.4900	7500.0000	2.90	13900.00		395.000	21.7000	905.0000
002304:	000135.04	197402221100CW.BBC	0.2080	4000.0000	3.00	7100.00		141.000	7.5000	340.0000
002305:	000136.01	197402181030CW.BBC	0.0230	1500.0000	2.40	5000.00		23.600	5.2000	278.0000
002306:	000137.01	197402181030CW.BBC	0.0420	3000.0000	2.90	2100.00		329.000	6.5000	208.0000
002307:	000137.02	197402181100CW.BBC	0.0410	3000.0000	2.90	2100.00		330.000	7.2300	195.0000
002308:	000137.03	197402181100CW.BBC	0.0340	3500.0000	2.90	1800.00		318.000	6.3200	190.0000
002309:	000138.01	197403031030CW.BBC	0.0320	1250.0000	3.00	1020.00		20.400	7.7000	31.5000
002310:	000138.02	197403031030CW.BBC	0.0080	800.0000	2.90	744.00		8.200	1.6000	6.6000
002311:	000138.03	197403031100CW.BBC	0.0070	1000.0000	3.00	844.00		20.100	7.4000	30.8000
002312:	000139.01	197402191030CW.BBC	0.2410	4250.0000	2.60	10000.00		151.000	0.2500	163.0000
002313:	000139.02	197402191030CW.BBC	0.7860	1500.0000	2.50	46000.00		942.000	1.5700	480.0000
002314:	000139.03	197402191100CW.BBC	0.7630	9000.0000	2.60	20000.00		1640.000	6.0000	1310.0000
002315:	000140.01	197402181130CW.BBC	0.0140	4500.0000	3.00	3800.00		608.000	10.8000	335.0000
002316:	000140.02	197402181130CW.BBC	0.1440	6500.0000	3.00	5300.00		720.000	3.0000	428.0000
002317:	000140.03	197402181200CW.BBC	0.0390	5200.0000	3.00	4500.00		623.000	9.1000	323.0000
002318:	000158.08	197402201200CW.BBC	0.0050	3000.0000	3.00	1900.00		51.700	16.5000	505.0000
002319:	000159.09	197402201230CW.BBC	0.4470	5200.0000	2.70	7900.00		53.700	16.0000	425.0000
002320:	000160.05	197402271130CW.BBC	0.4450	600.0000	2.80	13500.00		233.000	6.0000	565.0000
002321:	000160.06	197402271130CW.BBC	0.5430	650.0000	2.80	13100.00		211.000	3.7000	640.0000
002322:	000127.01	197212101033CW.BBD	0.0380	1700.0000	2.68	553.00		10.400	0.1400	62.0000
002323:	000127.01	197301101033CW.BBD	0.0320	1650.0000	2.65	652.00		11.000	0.2000	57.0000
002324:	000127.01	197302101033CW.BBD	0.0240	1650.0000	2.62	608.00		11.500	0.1500	58.0000
002325:	000127.01	197303131033CW.BBD	0.0200	1750.0000	2.65	748.00		11.200	0.1300	66.0000
002326:	000127.01	197304121033CW.BBD	0.0120	1800.0000	2.65	744.00		11.300	0.3000	62.4000
002327:	000127.01	197305111033CW.BBD	0.0180	1950.0000	2.60	620.00		11.100	0.1800	58.0000
002328:	000127.01	197306141033CW.BBD	0.0130	1900.0000	2.65	618.00		11.300	0.0800	64.5000
002329:	000127.01	197307251033CW.BBD	0.0140	2000.0000	2.65	1390.00		12.300	0.0600	68.3000
002330:	000127.01	197308241033CW.BBD	0.0140	2000.0000	2.65	700.00		11.300	0.1000	75.0000
002331:	000127.01	197309201033CW.BBD	0.0160	2000.0000	2.65	785.00		11.400	0.1000	73.9000
002332:	000127.01	197310181033CW.BBD	0.0210	1950.0000	2.65	843.00		10.300	0.0300	67.0000
002333:	000127.01	197311161033CW.BBD	0.0320	1850.0000	2.65	1283.00		11.300	0.1200	64.8000
002334:	000127.01	197312121033CW.BBD	0.0160	1800.0000	2.65	985.00		11.400	0.4800	58.6000
002335:	000127.01	197401151033CW.BBD	0.0120	1450.0000	2.65	825.00		11.700	0.8200	52.5000
002336:	000127.01	197402181033CW.BBD	0.0230	2000.0000	2.65	1300.00		11.200	0.4300	77.0000
002337:	000127.01	197403121033CW.BBD	0.0180	1850.0000	2.65	1000.00		11.100	0.3300	60.5000
002338:	000127.01	197404221033CW.BBD	0.0240	1600.0000	2.60	850.00		11.900	0.9200	54.3000
002339:	000127.01	197405201033CW.BBD	0.0190	1800.0000	2.65	790.00		14.600	0.1100	78.0000
002340:	000127.01	197406171033CW.BBD	0.0310	1900.0000	2.65	800.00		11.400	0.2300	76.8000
002341:	000127.01	197407281033CW.BBD	0.0230	1750.0000	2.65	772.00		11.300	0.1000	72.0000
002342:	000127.01	197408131033CW.BBD	0.0310	1800.0000	2.65	790.00		11.200	0.1500	73.0000
002343:	000127.01	197409121033CW.BBD	0.0410	2000.0000	2.65	1200.00		11.100	0.2400	66.1000
002344:	000127.01	197410141033CW.BBD	0.0170	2000.0000	2.65	1000.00		12.100	0.2600	78.0000
002345:	000127.01	197411161033CW.BBD	0.0200	1350.0000	2.70	1200.00		12.200	0.3600	68.0000
002346:	000127.01	197412171033CW.BBD	0.0250	700.0000	2.65	1100.00		11.300	1.0200	66.1000
002347:	000127.01	197501181033CW.BBD	0.0240	1750.0000	2.65	1300.00		11.600	0.1400	71.2000
002348:	000127.01	197502191033CW.BBD	0.0170	1000.0000	2.65	1400.00		10.500	0.7300	62.0000
002349:	000127.01	197503141033CW.BBD	0.0670	1750.0000	2.65	1350.00		10.700	0.2400	60.0000
002379:	000125.01	197304111130CW.BBA	0.0020	1500.0000	2.70	608.00		203.000	6.9000	151.0000
002380:	000125.01	197307251130CW.BBA	0.0040	2500.0000	2.90	780.00		119.000	7.6000	152.0000
002381:	000125.01	197312061130CW.BBA	0.0020	1000.0000	2.80	700.00		148.000	8.1000	151.0000
002382:	000125.01	197404231130CW.BBA	0.0030	2000.0000	2.70	689.00		130.000	7.2000	152.0000
002383:	000125.01	197409311130CW.BBA	0.0030	2000.0000	2.90	324.00		185.000	8.3000	120.0000
002384:	000125.01	197410141130CW.BBA	0.0030	2100.0000	2.90	400.00		185.000	9.3000	132.0000
002385:	000125.01	197501181130CW.BBA	0.0020	2000.0000	3.30	500.00		365.000	13.5000	149.0000
002386:	000125.01	197503141130CW.BBA	0.0010	1750.0000	3.30	450.00		258.000	15.3000	133.0000

Stack is pH-fld <= 4.0 (269 dataset) Sheet A.3

serid	samnumber	date time file	O.D.-420	cond.	pH-fld	acidity	Eh-fld	Na	K	Mg
002387:	000127.01	197301241030CW.BBA	0.0150	1700.0000	2.60	852.00		11.300	2.3000	56.8000
002388:	000127.01	197304111030CW.BBA	0.0120	1800.0000	2.60	744.00		11.700	0.3000	62.4000
002389:	000127.01	197312061030CW.BBA	0.0200	1450.0000	2.65	800.00		11.700	0.8200	52.5000
002390:	000127.02	197301241030CW.BBA	0.0170	1700.0000	2.65	824.00		10.800	0.5000	58.6000
002391:	000127.02	197304111030CW.BBA	0.0200	1850.0000	2.65	736.00		11.800	0.5000	62.1000
002392:	000127.02	197307251030CW.BBA	0.0300	1800.0000	2.60	1284.00		12.300	0.3600	69.3000
002393:	000127.02	197312161030CW.BBA	0.0260	1200.0000	2.65	680.00		10.500	1.6900	44.0000
002394:	000127.02	197404221030CW.BBA	0.0250	1500.0000	2.65	846.00		11.800	0.4600	53.2000
002395:	000127.03	197304111030CW.BBA	0.0150	1750.0000	2.65	680.00		11.800	0.5000	61.5000
002396:	000127.03	197307251030CW.BBA	0.0250	1800.0000	2.65	1200.00		12.200	0.4000	68.4000
002397:	000127.03	197312061030CW.BBA	0.0240	1200.0000	2.65	640.00		16.000	2.5000	46.4000
002398:	000127.03	197401241030CW.BBA	0.0160	1600.0000	2.65	820.00		11.000	0.7000	58.9000
002399:	000127.03	197404221030CW.BBA	0.0230	1400.0000	2.65	810.00		11.000	1.5000	42.0000
002400:	000127.03	197408131030CW.BBA	0.0210	1950.0000	2.65	879.00		15.600	0.9800	68.1000
002401:	000127.03	197410141030CW.BBA	0.0210	1400.0000	2.25	1200.00		11.700	0.5400	67.7000
002402:	000127.03	197501181030CW.BBA	0.0150	1500.0000	2.70	932.00		11.400	1.0900	59.9000
002403:	000127.03	197503141030CW.BBA	0.0040	1800.0000	2.65	950.00		11.500	0.9900	62.0000
002404:	000127.04	197301241100CW.BBA	0.0160	1650.0000	2.60	820.00		11.100	0.9300	57.2000
002405:	000127.04	197304111100CW.BBA	0.0170	2000.0000	2.65	488.00		11.900	0.9000	62.0000
002406:	000127.04	197307251100CW.BBA	0.0260	1800.0000	2.65	1100.00		12.600	0.7400	68.4000
002407:	000127.04	197312061100CW.BBA	0.0280	1200.0000	2.65	641.00		16.000	2.4000	45.4000
002408:	000127.04	197404221100CW.BBA	0.0210	1500.0000	2.65	785.00		11.000	1.5000	42.0000
002409:	000127.04	197408131100CW.BBA	0.0210	1700.0000	2.70	404.00		11.100	3.5000	54.3000
002410:	000127.04	197410141100CW.BBA	0.0240	1500.0000	2.70	879.00		11.700	0.5500	67.0000
002411:	000127.04	197501181100CW.BBA	0.0180	1300.0000	2.65	1150.00		16.900	0.5300	67.0000
002412:	000127.04	197503141100CW.BBA	0.0140	1800.0000	2.70	900.00		11.500	1.0900	58.9000
002413:	000127.05	197304111100CW.BBA	0.0200	1800.0000	2.70	688.00		12.000	1.7000	57.0000
002414:	000127.05	197307251100CW.BBA	0.0190	1000.0000	2.70	1000.00		13.100	1.8500	66.9000
002415:	000127.05	197312161100CW.BBA	0.0170	1400.0000	2.70	520.00		16.500	3.6000	41.3000
002416:	000127.05	197404221100CW.BBA	0.0140	1700.0000	2.65	780.00		11.000	1.4000	40.1000
002417:	000127.05	197410141100CW.BBA	0.0230	1300.0000	2.70	843.00		11.100	3.6000	52.4000
002418:	000127.05	197501181100CW.BBA	0.0160	1050.0000	2.75	1100.00		17.200	6.1000	68.5000
002419:	000127.05	197503141100CW.BBA	0.0150	1750.0000	2.70	750.00		11.700	0.7500	65.6000
002420:	000127.06	197304111100CW.BBA	0.0100	1650.0000	2.70	600.00		15.000	1.1300	58.8000
002421:	000127.06	197307251100CW.BBA	0.0150	1800.0000	2.70	900.00		13.200	1.9500	65.7000
002422:	000127.06	197312061100CW.BBA	0.0140	1050.0000	2.80	512.00		16.400	3.4000	40.6000
002423:	000127.06	197404221100CW.BBA	0.0130	1400.0000	2.70	780.00		10.800	1.3000	40.0000
002424:	000127.08	197301241130CW.BBA	0.0160	1800.0000	3.90	321.00		277.000	19.0000	156.0000
002425:	000127.08	197304111130CW.BBA	0.0200	1800.0000	3.60	264.00		176.000	11.8000	92.0000
002426:	000127.08	197307251130CW.BBA	0.0400	1800.0000	3.20	744.00		117.000	10.0000	122.0000
002427:	000127.08	197312061130CW.BBA	0.0180	1350.0000	3.10	700.00		111.000	9.8000	78.2000
002428:	000127.08	197404221130CW.BBA	0.0200	1600.0000	3.70	390.00		133.000	11.2000	110.0000
002431:	000127.08	197501181130CW.BBA	0.0170	1500.0000	2.90	800.00		69.000	4.8000	72.9000
002432:	000127.08	197503141130CW.BBA	0.0160	1700.0000	2.85	750.00		51.000	4.4000	64.6000
002433:	000127.10	197301241130CW.BBA	0.0020	2000.0000	3.20	220.00		257.000	15.2000	142.0000
002434:	000127.10	197304111130CW.BBA	0.0010	1450.0000	3.30	212.00		120.000	11.8000	92.8000
002435:	000127.10	197307251130CW.BBA	0.0030	1800.0000	3.00	550.00		74.400	6.2000	121.0000
002436:	000127.10	197312061130CW.BBA	0.0020	1100.0000	3.30	260.00		127.000	7.6000	59.9000
002437:	000127.10	197404221130CW.BBA	0.0030	1350.0000	3.30	218.00		195.000	8.3000	87.0000
002438:	000127.10	197408131130CW.BBA	0.0230	1800.0000	3.10	244.00		190.000	12.8000	103.0000
002439:	000127.10	197410141130CW.BBA	0.0000	2000.0000	3.80	240.00		67.500	13.0000	110.0000
002440:	000127.10	197501181130CW.BBA	0.0700	1000.0000	3.10	608.00		51.800	5.4000	71.1000
002441:	000127.10	197503141130CW.BBA	0.0170	1400.0000	2.80	600.00		176.000	5.2000	62.0000
002442:	000127.11	197301241200CW.BBA	0.0020	2000.0000	3.50	250.00		120.000	8.1000	82.0000
002443:	000127.11	197304111200CW.BBA	0.0020	850.0000	3.90	100.00		67.000	5.6000	58.0000
002444:	000127.11	197307251200CW.BBA	0.0020	1800.0000	3.30	200.00		62.000	3.5000	62.5000
002445:	000127.11	197312061200CW.BBA	0.0040	1850.0000	3.40	232.00		54.300	6.0000	51.7000
002446:	000127.11	197404221200CW.BBA	0.0030	1400.0000	3.50	202.00		106.000	4.2000	56.0000
002447:	000127.11	197408131200CW.BBA	0.0100	1650.0000	3.80	201.00		170.000	11.2000	97.3000
002448:	000127.11	197410141200CW.BBA	0.0140	1600.0000	3.90	200.00		182.000	12.3000	86.4000
002449:	000127.11	197501181200CW.BBA	0.0540	700.0000	3.10	388.00		51.200	4.9000	62.4000
002450:	000127.11	197503141200CW.BBA	0.0120	1200.0000	3.10	450.00		44.800	4.6000	57.9000
002453:	000127.12	197307251200CW.BBA	0.0010	1250.0000	3.70	187.00		38.400	4.2000	42.1000
002458:	000127.13	197307251230CW.BBA	0.0030	1250.0000	3.70	180.00		40.000	4.0000	44.5000
002463:	000127.13	197501181230CW.BBA	0.0220	400.0000	3.70	236.00		36.000	4.0000	52.9000
002464:	000127.13	197503141230CW.BBA	0.0180	900.0000	3.50	250.00		35.100	4.0000	50.9000
002668:	000022.01	197303011400CW.BBK	0.0100	1750.0000	2.60			11.000	0.2400	59.0000
002669:	000022.01	197306301330CW.BBK	0.0130	2000.0000	2.60			11.200	0.0800	61.0000
002670:	000022.01	197310101330CW.BBK	0.0120	1800.0000	2.50			11.100	0.0400	76.9000
002671:	000022.01	197401191330CW.BBK	0.0200	1450.0000	2.60			11.700	0.8200	52.5000
002677:	000127.07	197301241030CW.BBL	0.0410	1900.0000	3.90	123.00		272.000	18.6000	152.0000
002678:	000127.07	197304111030CW.BBL	0.0320	1900.0000	3.40	112.00		320.000	20.8000	161.0000
002679:	000127.07	197304221030CW.BBL	0.0310	1500.0000	3.80	426.00		198.000	16.2000	110.0000
002680:	000127.07	197307251030CW.BBL	0.0280	2050.0000	3.70	484.00		194.000	15.7000	158.0000
002681:	000127.07	197312061030CW.BBL	0.0300	1600.0000	3.70	400.00		197.000	15.9000	132.0000
002682:	000127.09	197301241030CW.BBL	0.0110	2000.0000	3.50			282.000	19.1000	152.0000
002683:	000127.09	197304111030CW.BBL	0.0120	1700.0000	3.80	208.00		132.000	16.6000	132.0000
002684:	000127.09	197304221030CW.BBL	0.0020	1400.0000	3.60	210.00		130.000	10.1000	93.0000
002685:	000127.09	197307251030CW.BBL	0.0050	1800.0000	3.00	236.00		121.000	11.0000	124.0000
002686:	000127.09	197312061030CW.BBL	0.0120	1100.0000	3.20	292.00		75.000	7.8000	60.0000
002691:	000127.02	1972102311000CW.BBM			2.60	10.60		0.270	61.0000	0.0600
002692:	000127.03	1972102311000CW.BBM			2.60	10.80		0.560	63.0000	0.6800
002693:	000127.04	1972102311000CW.BBM			2.60	10.80		0.560	63.0000	0.6800
002694:	000127.05	1972102311000CW.BBM			2.60	10.80		0.700	62.0000	0.6500
002695:	000127.06	1972102311000CW.BBM			2.60	11.00		0.840	61.0000	0.6700
002696:	000127.07	1972102311000CW.BBM			3.20	159.00		10.700	123.0000	0.3600
002697:	000127.08	1972102311000CW.BBM			3.80	145.00		9.600	119.0000	0.3900
002698:	000127.09	1972102311000CW.BBM			3.80	137.00		9.600	114.0000	0.3300
002699:	000127.10	1972102311000CW.BBM			3.30	135.00		8.800	115.0000	0.3900
002700:	000127.11	1972102311000CW.BBM			3.60	102.00		8.600	97.0000	0.2500
MINIMUM			0.0010	65.0000	0.90	10.60	382.00	0.270	0.0300	0.2500
MAXIMUM			0.9880	80000.0000	4.00	64000.00	500.00	3550.000	123.0000	2340.0000
MEAN			0.0978	4919.5252	2.82	3354.56	493.13	156.473	9.2733	200.4349
STD DEVIATION			0.1697	14794.5219	0.46	6412.41	55.60	308.150	18.4415	200.6390

Stack is pH-fld <= 4.0 (269 dataset)

Sheet B.1

serid	sarnumber	date time file	Ca	Zn	CU	Mn	Fe	Al	Pb	Co
000456:	000152.10	197508121440CD.AK	385.000	2.1400	0.0740	4.7000	99.0000	19.0000	0.0050	0.1950
000457:	000152.20	197508121310CD.AK	445.000	3.6000	0.0880	6.4000	96.0000	28.3000	0.0060	0.2450
000458:	000153.01	197508121520CD.AK							0.0040	
000459:	000154.01	197508121400CD.AK	441.000	1.6400	0.0300	8.6000	293.0000	32.8000	0.0020	0.3500
000492:	000066.01	197510081720CD.AM	38.500	2.4100	0.2900	1.1400	87.5000	20.9000	0.0040	0.1000
000493:	002003.01	197510070940CD.AM	6.700	141.0000	135.0000	3.6200	113.0000	115.5000	0.8400	0.6100
000494:	002004.01	197510071000CD.AM	6.500	140.0000	135.0000	3.7700	111.0000	112.0000	0.8400	0.6100
000495:	002006.01	197510071200CD.AM	11.000	158.0000	220.0000	8.5000	455.0000	505.0000	0.4100	1.0000
000496:	002007.01	197510071400CD.AM	1.500	303.0000	320.0000	9.3800	725.0000	522.5000	1.7500	1.1000
000497:	002007.02	197510071300CD.AM	7.700	280.0000	293.0000	9.0000	675.0000	522.5000	1.4300	1.1000
000498:	002007.03	197510071415CD.AM	1.200	270.0000	325.0000	9.2500	600.0000	520.0000	0.9400	1.1000
000505:	002014.01	197510071030CD.AM	6.700	260.0000	200.0000	9.7500	388.0000	375.0000	0.8100	1.1000
000506:	002015.10	197510071600CD.AM	17.700	18.9000	3.6000	5.3000	8.6000	24.3000	0.3400	0.1100
000605:	001004.01	197702091045CD.AQ	528.000	43.5000	0.4300	387.0000	5450.0000	346.0000	0.1170	8.7000
000608:	009001.01	197705051500CD.AR	213.000	8.4000	1.6900	22.0000	360.0000	52.0000	0.0050	0.1900
000612:	009004.01	197705141500CD.AR	204.000	1.9300	0.3100	3.9700	507.0000	51.8000	0.0180	0.4200
000613:	009005.01	197705151130CD.AR	321.000	0.8400	0.0870	17.0000	73.4000	30.0000	0.0070	0.2500
000614:	009006.01	197705151400CD.AR	413.000	1.1300	0.1030	20.9000	303.0000	39.7000	0.0040	0.3900
000615:	009007.10	197705171620CD.AR	342.000	5.0900	0.9700	44.5000	715.0000	159.5000	0.0090	1.4000
000616:	009008.45	197706111400CD.AR	2.980	0.0400	0.0020	0.0210	0.9900	6.7000	0.0060<0.01	
000617:	009008.48	197706111600CD.AR	2.780	0.0250	0.0020	0.0180	1.4300	13.0000	0.0040	0.0200
000618:	009008.50	197706111650CD.AR	2.860	0.0340	0.0040	0.0140	1.6500	14.0000	0.0040	0.0300
000619:	009008.52	197706121100CD.AR	3.400	0.0430	0.0050	0.0180	1.5500	15.4000	0.0060	0.0300
000620:	009009.20	197706111100CD.AR							260.0000	
000621:	009009.20	197706111101CD.AR	3.280	0.0280<0.002		0.0980	1.4500	3.2600	0.0040	0.0300
000622:	009010.02	197706111150CD.AR								
000623:	009011.01	197706120700CD.AR	2.880	0.0480<0.002		0.0560	4.7400	8.4000	0.0040	0.0400
000624:	009012.01	197706120730CD.AR	2.680	0.0410	0.0050	0.0360	3.6900	23.5000	0.0030	0.0320
000625:	009013.03	197706120800CD.AR	2.480	0.0480	0.0080	0.0560	4.1200	8.0000	0.0060<0.01	
000628:	009016.99	197706231345CD.AR	328.000	5.5000	0.0800	5.0300	1.1500	2.6800	2.3500	0.2600
000746:	003016.01	197806301300CD.AW	66.300	44.1000	0.2000	1.5400	0.0400	11.1000	0.0330	0.1200
000747:	003016.03	197806301315CD.AW								
000757:	000213.01	197805201100CD.AX	197.500	118.0000	9.3000	54.2000	247.0000	13.6000	0.6100	0.4300
000758:	000214.21	197805201500CD.AX	27.700	15.8000	8.8000	5.4000	16.8000	23.5000	0.3300	0.0800
000759:	000215.01	197805311500CD.AX	41.000	61.5000	44.5000	19.2000	662.0000	147.5000	0.1300	0.5400
000760:	000215.15	197806011200CD.AX	39.000	62.0000	44.5000	18.7000	598.0000	142.5000	0.1700	0.4800
000761:	000216.15	197806011500CD.AX	41.000	41.0000	26.5000	11.1000	242.0000	83.0000	0.7200	0.2900
000762:	000216.50	197806021100CD.AX	28.000	11.8000	5.7000	3.5000	52.0000	20.0000	0.1400	0.0800
000763:	000217.01	197806031400CD.AX	85.000	17.4000	11.7000	10.2000	103.0000	58.0000	0.0800	0.2700
000764:	000218.01	197806031100CD.AX	20.000	12.5000	3.2900	8.8000	9.2000	16.5000	0.6200	0.0700
000765:	000220.03	197806151500CD.AX	61.500	0.0730	0.4300	0.1440	4.1000	0.3200	0.0280	
000766:	000220.15	197806151530CD.AX	23.400	0.2000	0.6700	0.2300	11.4000	1.0000	0.1030	
000767:	000221.05	197806151600CD.AX	16.900	0.4400	0.1000	0.1150	4.6000	1.4200	4.9500	
000891:	000100.30	197508081740CW.ABBA1	53.500	1.1700	0.0570	4.0100	16.0000	4.8200	1.0100	0.0700
000895:	000109.01	197508081800CW.ABBA1	16.100	0.1350	0.0050	2.5100	11.8500	0.5800	0.0150	0.0200
001198:	000119.01	197409191920CW.BK	0.540	0.0390	0.0100	0.0450	1.2100	0.4400	0.0250	
001199:	000120.01	197409191930CW.BK	0.710	0.0510	0.0280	0.0480	0.7500	0.4400	0.0250	
002120:	000100.30	197831301330CW.BBE	8.950	0.3200		0.5900	7.8000		0.1440	
002224:	000127.01	197307041430CW.BBB	64.200	1.1400	0.6900	6.6000	110.0000	31.2000	0.0130	0.2000
002225:	000127.02	197307041430CW.BBB	67.800	1.1000	0.6300	7.1000	100.0000	31.9000	0.0120	0.0000
002226:	000127.03	197307041100CW.BBB	72.000	1.0700	0.5700	6.9000	85.0000	29.0000	0.0140	0.0070
002227:	000127.04	197307041100CW.BBB	148.000	1.3700	0.3700	9.1000	21.0000	29.1000	0.0060	0.0100
002228:	000128.01	197307081030CW.BBB	556.000	0.9600	0.0230	45.2000	193.0000	149.0000	0.0100	1.8500
002229:	000128.02	197307081030CW.BBB	510.000	1.1000	0.2400	88.7000	70.0000	325.0000	0.0180	2.4800
002230:	000128.03	197307081100CW.BBB	524.000	1.4000	1.2400	113.0000	105.0000	411.0000	0.0200	2.6200
002231:	000130.01	197307061030CW.BBB	490.000	10.2000	3.9800	160.0000	4300.0000	1570.0000	0.3300	0.8400
002232:	000130.02	197307061100CW.BBB	405.000	3.7900	0.9800	126.0000	4200.0000	320.0000	0.0950	4.0500
002233:	000130.03	197307061130CW.BBB	428.000	4.7900	0.0800	60.0000	420.0000	280.0000	0.0340	1.6800
002234:	000131.01	197307061200CW.BBB	300.000	2.5500	1.2400	27.4000	431.0000	94.5000	0.0070	0.7000
002235:	000131.02	197307061200CW.BBB	260.000	2.6700	1.7600	28.7000	420.0000	100.0000	0.0080	0.8000
002236:	000131.03	197307061230CW.BBB	383.000	4.5400	0.5400	89.0000	390.0000	190.0000	0.0150	2.1500
002237:	000132.01	197309101030CW.BBB	399.000	1.0900	0.1800	85.6000	1310.0000	169.0000	0.0260	2.0000
002238:	000132.02	197309101030CW.BBB	236.000	0.6100	0.1300	41.4000	353.0000	142.0000	0.0120	1.0000
002239:	000132.03	197309101100CW.BBB	520.000	1.3200	3.1000	40.1000	1080.0000	400.0000	0.2100	2.3400
002240:	000132.04	197309101100CW.BBB	240.000	1.1000	0.7900	41.9000	371.0000	142.0000	0.0050	1.2600
002241:	000132.05	197309101130CW.BBB	237.000	0.6000	0.1400	41.2000	370.0000	141.0000	0.0130	1.2500
002242:	000132.06	197309101130CW.BBB	250.000	0.6100	0.1300	42.0000	357.0000	132.0000	0.0130	1.2000
002243:	000132.07	197309101200CW.BBB	248.000	0.6200	0.1300	41.5000	352.0000	132.0000	0.0070	1.2500
002244:	000133.01	197309171030CW.BBB	138.000	0.0900	0.0800	0.3700	12.5000	23.0000	0.2400	0.0500
002247:	000134.03	197303151100CW.BBB	400.000	0.9600	0.4200	73.5000	1020.0000	136.0000	0.0160	0.1300
002249:	000135.01	197307181030CW.BBB	444.000	6.7000	0.1100	192.0000	4900.0000	179.0000	0.0830	2.9400
002250:	000135.02	197307181030CW.BBB	426.000	5.7100	0.0500	133.0000	2700.0000	154.0000	0.0400	1.6500
002251:	000135.03	197307181100CW.BBB	445.000	6.4000	0.1100	163.0000	3600.0000	172.0000	0.0570	2.0000
002252:	000135.04	197307181130CW.BBB	41.000	0.7600	0.0200	16.4000	290.0000	19.5000	0.0060	0.4300
002253:	000136.01	197309111030CW.BBB	338.000	88.0000	3.1000	114.0000	23000.0000	425.0000	0.0800	3.2000
002254:	000137.01	197309111030CW.BBB	340.000	0.8500	0.0180	50.1000	83.0000	4.8000	0.0200	0.8000
002255:	000137.02	197309111100CW.BBB	330.000	0.8300	0.0190	51.0000	85.0000	4.7600	0.0010	0.8100
002256:	000137.03	197309111100CW.BBB	325.000	0.8600	0.0170	50.0000	76.2000	4.8300	0.0120	0.6100
002257:	000138.01	197310131030CW.BBB	68.000	54.5000	0.0060	29.9000	112.0000	4.7000	1.8000	0.1900
002258:	000138.02	197310131030CW.BBB	67.000	193.0000	2.3700	44.8000	18.7000	10.4000	0.9500	0.2900
002259:	000138.03	197310131100CW.BBB	61.200	57.0000	0.1340	31.0000	34.5000	5.6000	1.5200	0.1800
002260:	000139.01	197409111030CW.BBB	212.000	9.7400	2.9100	51.3000	809.0000	96.0000	0.0160	6.0000
002261:	000139.02	197409111200CW.BBB	300.000	76.5000	4.0000	110.0000	9450.0000	3150.0000	0.0190	20.0000
002262:	000139.03	197409111300CW.BBB	385.000	8.4800	2.9800	544.0000	1782.0000	775.0000	0.2900	6.4000
002263:	000140.01	197409111400CW.BBB	535.000	24.6000	0.0520	112.0000	186.0000	169.0000	0.0130	2.0120
002264:	000140.02	197409111230CW.BBB	303.000	18.0000	0.1400	67.2000	112.0000	111.0000	0.0190	0.3000
002265:	000140.03	197409111300CW.BBB	290.000	20.5000	0.1300	62.3000	96.0000	100.0000	0.0130	1.0000
002266:	000157.05	197307061100CW.BBB	480.000	17.1000	8.3600	285.0000	4950.0000	2500.0000	0.1100	5.8700
002267:	000157.06	197307061130CW.BBB	400.000	2.7400	0.7400	55.0000	454.0000	152.0000	0.0040	0.8300

Stack is pH-fld <= 4.0 (269 dataset)

Sheet B.2

serid	sarnumber	date time file	Ca	Zn	CU	Mn	Fe	Al	Pb	Co
002268:	000157.07	197307061130CW.BBB	420.000	7.4100	2.8700	151.0000	345.0000	890.0000	0.0570	2.4800
002269:	000158.08	197309101200CW.BBB	377.000	0.5400	0.1400	40.6000	155.0000	81.3000	0.0080	1.2300
002270:	000159.09	197309101230CW.BBB	389.000	0.3200	0.1400	77.0000	679.0000	35.0000	0.0170	1.2500
002271:	000160.05	197307101130CW.BBB	415.000	5.5000	0.0500	128.0000	2900.0000	114.0000	0.0770	1.4000
002272:	000160.06	197307181130CW.BBB	435.000	5.9000	0.0800	132.0000	2750.0000	124.0000	0.0960	1.5200
002273:	000127.01	197402141030CW.BBC	60.100	0.9700	0.6500	6.3000	95.0000	32.0000	0.0500	0.2100
002274:	000127.02	197402141030CW.BBC	57.200	0.9600	0.6300	6.3000	90.0000	30.5000	0.0500	0.0800
002275:	000127.03	197402141100CW.BBC	58.600	0.9500	0.6300	6.2000	89.0000	30.1000	0.0400	0.0400
002276:	000127.04	197402141100CW.BBC	133.000	1.5000	0.4800	10.6000	36.0000	38.5000	0.0800	0.2300
002277:	000128.01	197402101030CW.BBC	389.000	0.8500	0.0700	34.2000	225.0000	119.0000	0.2100	1.3100
002278:	000128.02	197402101030CW.BBC	410.000	1.2000	0.1200	72.4000	111.0000	206.0000	0.1000	2.3100
002279:	000128.03	197402101100CW.BBC	435.000	1.2900	2.3200	76.9000	255.0000	400.0000	0.3200	2.1500
002280:	000129.01	197402111030CW.BBC	272.000	0.4100	2.1000	5.1000	480.0000	149.0000	0.1700	2.1500
002281:	000129.02	197402111030CW.BBC	420.000	0.7100	2.6700	8.7000	810.0000	295.0000	0.2200	1.1000
002282:	000130.01	197402121030CW.BBC	433.000	4.3800	4.4000	54.4000	2500.0000	608.0000	0.3700	1.5200
002283:	000130.02	197402121030CW.BBC	176.000	1.5900	0.6600	29.3000	230.0000	87.5000	0.1300	0.4100
002284:	000130.03	197402121100CW.BBC	256.000	2.8600	1.3000	48.1000	1400.0000	294.0000	0.2500	0.8500
002285:	000130.04	197402121100CW.BBC	361.000	3.9200	0.4500	84.8000	3650.0000	380.0000	0.4300	1.1000
002286:	000131.01	197402121130CW.BBC	415.000	4.3300	3.1900	31.5000	840.0000	146.0000	0.1700	1.0000
002287:	000131.02	197402121130CW.BBC	268.000	4.6400	2.8400	42.4000	610.0000	176.0000	0.1700	1.1200
002288:	000131.03	197402121200CW.BBC	249.000	4.9900	3.0200	40.8000	593.0000	212.0000	0.1900	1.3000
002289:	000132.01	197402201030CW.BBC	164.000	0.7900	0.3800	22.9000	275.0000	101.0000	0.1500	0.7200
002290:	000132.02	197402201030CW.BBC	157.000	0.7500	0.3600	21.7000	260.0000	101.0000	0.1000	0.7100
002291:	000132.03	197402201100CW.BBC	437.000	2.9700	16.0000	26.1000	4700.0000	836.0000	0.4500	3.2000
002292:	000132.04	197402201100CW.BBC	165.000	0.8600	0.5100	22.7000	315.0000	109.0000	0.1400	0.7200
002293:	000132.05	197402201130CW.BBC	311.000	1.3800	0.3600	40.2000	155.0000	72.0000	0.1600	0.8200
002294:	000132.06	197402201130CW.BBC	221.000	1.1200	0.8100	31.5000	415.0000	126.0000	0.1700	1.0000
002295:	000132.07	197402201200CW.BBC	215.000	1.1000	0.8000	33.2000	400.0000	126.0000	0.1700	1.0100
002296:	000133.01	197402201230CW.BBC	453.000	0.5800	0.0700	11.5000	25.0000	36.5000	0.2100	0.3000
002297:	000134.01	197402121200CW.BBC	379.000	3.8000	0.8800	11.2000	265.0000	71.5000	0.1700	2.7100
002298:	000134.02	197402121200CW.BBC	450.000	3.7500	0.8700	11.4000	270.0000	74.5000	0.1700	0.4100
002299:	000134.03	197402121230CW.BBC	336.000	2.5000	0.0400	34.0000	535.0000	73.5000	0.2000	0.4000
002300:	000134.04	197402121230CW.BBC	422.000	3.0100	0.5900	15.8000	365.0000	76.0000	0.3200	0.5500
002301:	000135.01	197402221030CW.BBC	415.000	5.4000	0.2400	100.0000	3250.0000	149.0000	0.3600	1.6100
002302:	000135.02	197402221030CW.BBC	391.000	7.6000	0.0900	96.1000	2000.0000	201.0000	0.1900	1.6100
002303:	000135.03	197402221100CW.BBC	473.000	6.8100	0.1500	100.0000	2850.0000	173.0000	0.1300	1.8200
002304:	000135.04	197402221100CW.BBC	175.000	2.6000	0.1100	38.8000	770.0000	82.5000	0.3200	0.7200
002305:	000136.01	197402181030CW.BBC	401.000	62.0000	3.2100	120.0000	10000.0000	501.0000	0.1000	4.1000
002306:	000137.01	197402181030CW.BBC	265.000	0.7200	0.0300	45.4000	195.0000	3.7200	0.1300	0.4100
002307:	000137.02	197402181100CW.BBC	249.000	0.7100	0.0300	45.2000	110.0000	3.7200	0.2300	0.4000
002308:	000137.03	197402181100CW.BBC	239.000	0.6400	0.0300	44.4000	95.0000	3.9200	0.2200	0.4100
002309:	000138.01	197403181030CW.BBC	73.300	67.0000	0.1400	33.0000	132.0000	5.1000	1.2900	0.2300
002310:	000138.02	197403181030CW.BBC	4.000	117.0000	1.8800	21.5000	7.5000	6.7000	0.5100	0.3300
002311:	000138.03	197403181100CW.BBC	72.000	64.0000	0.2900	32.5000	91.5000	12.6000	1.7500	0.2600
002312:	000139.01	197403191030CW.BBC	130.000	7.8000	1.5600	36.9000	660.0000	620.0000	0.9100	2.7300
002313:	000139.02	197402161030CW.BBC	398.000	32.1000	2.7000	106.0000	8000.0000	2000.0000	0.5600	11.3000
002314:	000139.03	197402191100CW.BBC	298.000	14.9000	2.6000	211.0000	2400.0000	656.0000	1.5000	5.5000
002315:	000140.01	197402181130CW.BBC	435.000	41.4000	0.0400	99.8000	23.5000	138.0000	0.2500	2.1000
002316:	000140.02	197402181130CW.BBC	425.000	22.9000	0.2800	100.0000	465.0000	175.0000	0.2600	2.2100
002317:	000140.03	197402181200CW.BBC	359.000	37.9000	0.2400	103.0000	120.0000	138.0000	0.2300	1.9100
002318:	000158.08	197402201200CW.BBC	397.000	0.7700	0.2200	28.9000	10.1000	55.0000	0.1800	0.5500
002319:	000159.09	197402201230CW.BBC	425.000	1.1600	0.8100	81.0000	1250.0000	133.0000	0.3400	1.2100
002320:	000160.05	197402211130CW.BBC	295.000	6.1000	0.6900	83.3000	1925.0000	306.0000	0.3000	1.2600
002321:	000160.06	197402221130CW.BBC	340.000	5.8000	0.6300	79.1000	1725.0000	292.0000	0.2610	1.7100
002322:	000127.01	197212101033CW.BBD	68.200	1.1000	0.6700	6.2000	95.0000	24.8000	0.0100	0.1700
002323:	000127.01	197301101033CW.BBD	60.500	1.0500	0.7000	6.3000	111.0000	27.7000	0.0010	0.2100
002324:	000127.01	197302101033CW.BBD	57.700	1.1200	0.7100	6.0000	102.0000	26.5000	0.0100	0.5000
002325:	000127.01	197303131033CW.BBD	60.500	1.1300	0.7700	6.3000	120.0000	28.8000	0.0120	0.2000
002326:	000127.01	197304121033CW.BBD	58.000	1.0200	0.8300	6.3000	108.0000	29.8000	0.0090	0.2300
002327:	000127.01	197305111033CW.BBD	53.400	1.0000	0.7900	6.3000	102.0000	28.5000	0.0080	0.1500
002328:	000127.01	197306141033CW.BBD	67.500	1.1400	0.6800	5.5000	94.0000	28.8000	0.0130	0.1800
002329:	000127.01	197307251033CW.BBD	64.200	1.0800	0.6900	6.6000	110.0000	31.2000	0.0170	0.1600
002330:	000127.01	197308241033CW.BBD	67.200	1.2400	0.7500	7.1000	116.0000	34.6000	0.0030	0.1100
002331:	000127.01	197309301033CW.BBD	64.300	1.3200	0.8300	7.6000	126.0000	36.7000	0.0130	0.1200
002332:	000127.01	197310181033CW.BBD	67.000	1.2300	0.8500	7.6000	144.0000	36.5000	0.0180	0.2500
002333:	000127.01	197311161033CW.BBD	65.200	1.3000	0.7800	7.2000	105.0000	34.2000	0.0500	0.2100
002334:	000127.01	197312121033CW.BBD	62.100	1.3100	0.6900	7.2000	98.5000	32.3000	0.0510	0.2400
002335:	000127.01	197401151033CW.BBD	53.100	1.2300	0.5700	6.2000	82.0000	25.0000	0.0140	0.2000
002336:	000127.01	197402181033CW.BBD	60.000	0.9700	0.6500	5.3000	95.0000	32.0000	0.0600	0.2300
002337:	000127.01	197403121033CW.BBD	66.300	1.0800	0.7300	6.8000	103.0000	33.2000	0.0030	0.2100
002338:	000127.01	197404221033CW.BBD	61.000	1.0600	0.5900	6.1000	100.0000	23.0000	0.0200	0.2300
002339:	000127.01	197405201033CW.BBD	100.000	1.5800	1.0000	8.9000	140.0000	43.2000	0.0100	0.6500
002340:	000127.01	197406171033CW.BBD	81.200	1.3400	0.6800	6.6000	115.0000	36.2000	0.0030	0.3700
002341:	000127.01	197407281033CW.BBD	74.000	1.2400	0.8200	7.6000	116.0000	33.8000	0.0030	0.4700
002342:	000127.01	197408131033CW.BBD	72.600	1.2300	0.8400	7.2000	122.0000	34.4000	0.0060	0.2500
002343:	000127.01	197409121033CW.BBD	66.500	1.2100	0.8600	7.6000	129.0000	37.9000	0.0080	0.1600
002344:	000127.01	197410141033CW.BBD	81.000	1.2100	0.8200	7.1000	125.0000	35.1000	0.0090	0.1800
002345:	000127.01	197411161033CW.BBD	70.000	1.0100	0.6600	5.4000	95.0000	26.5000	0.0093	0.1700
002346:	000127.01	197412171033CW.BBD	62.400	1.2300	0.8300	6.9000	129.0000	51.8000	0.0170	0.2400
002347:	000127.01	197501101033CW.BBD	66.200	1.2900	0.8700	7.8000	131.0000	38.3000	0.0070	0.2100
002348:	000127.01	197502151033CW.BBD	58.600	1.0500	0.6900	6.5000	100.0000	28.1000	0.0020	0.0400
002349:	000127.01	197503141033CW.BBD	56.500	1.1300	0.7500	6.9000	109.0000	30.4000	0.0180	0.2400
002379:	000125.01	197404301130CW.BBA	120.000	1.2200	0.3200	9.2000	14.7000	30.4000	0.0130	0.1700
002380:	000125.01	197407051130CW.BBA	148.000	1.3400	0.3600	9.1000	21.0000	29.1000	0.0060	0.3000
002381:	000125.01	197312011130CW.BBA	147.000	1.2700	0.3700	8.4000	25.0000	30.3000	0.0090	0.3100
002382:	000125.01	197404221130CW.BBA	144.000	1.2300	0.3400	9.0000	18.0000	30.2000	0.0100	0.3000
002383:	000125.01	197408131130CW.BBA	122.000	1.0000	0.2600	7.8000	8.6000	24.6000	0.0110	0.1900
002384:	000125.01	197410141130CW.BBA	124.000	0.7500	0.1600	7.2000	3.9000	20.6000	0.0070	0.1600
002385:	000125.01	197501101130CW.BBA	150.000	0.7300	0.1200	7.0000	2.0000	22.1000	0.0050	0.1600
002386:	000125.01	197503141130CW.BBA	152.000	0.6400	0.1000	7.0000	2.0100	22.1000	0.0050	0.1600

Stack is pH-fld <= 4.0 (269 dataset) Sheet B.3

serid	sarnumber	date time file	Ca	Zn	CU	Mn	Fe	Al	Pb	Co
002387:	000127.01	197301241030CW.BBA	59.900	1.0500	0.7200	6.0000	108.0000	27.6000	0.0120	0.2400
002388:	000127.01	197304111030CW.BBA	58.000	1.1300	0.8300	6.3000	108.0000	29.8000	0.0120	0.2200
002389:	000127.01	197312061030CW.BBA	53.000	1.1000	0.5700	5.3000	82.0000	25.0000	0.1040	0.2000
002390:	000127.02	197301241030CW.BBA	61.800	1.0500	0.7300	6.1000	114.0000	28.3000	0.0090	0.2100
002391:	000127.02	197304111030CW.BBA	58.000	1.1000	0.8100	6.4000	110.0000	30.1000	0.0120	0.2000
002392:	000127.02	197307251030CW.BBA	64.100	1.1000	0.6500	6.6000	110.0000	30.6000	0.0090	0.1900
002393:	000127.02	197312161030CW.BBA	44.500	0.7800	0.4200	4.3000	63.0000	19.7000	0.0190	0.1800
002394:	000127.02	197404221030CW.BBA	60.100	0.9800	0.4800	5.2000	95.0000	22.0000	0.0180	0.1900
002395:	000127.03	197304111030CW.BBA	58.900	1.1000	0.8100	6.5000	110.0000	29.2000	0.0180	0.2000
002396:	000127.03	197307251030CW.BBA	64.100	1.0800	0.6500	6.6000	110.0000	31.9000	0.0120	0.0800
002397:	000127.03	197312061030CW.BBA	50.400	0.7600	0.3900	4.7000	58.0000	18.9000	0.0180	0.2100
002398:	000127.03	197401241030CW.BBA	63.300	1.1000	0.7200	6.2000	110.0000	29.2000	0.0160	0.1900
002399:	000127.03	197404221030CW.BBA	39.600	0.9500	0.4800	5.1000	93.0000	18.1000	0.0170	0.1600
002400:	000127.03	197408131030CW.BBA	69.200	1.1100	0.7300	6.8000	113.0000	32.4000	0.0050	0.2700
002401:	000127.03	197410141030CW.BBA	63.700	1.2000	0.8000	7.4000	124.0000	36.1000	0.0030	0.2300
002402:	000127.03	197501181030CW.BBA	57.900	1.0800	0.6600	6.7000	98.0000	27.5000	0.0040	0.2200
002403:	000127.03	197503141030CW.BBA	58.700	1.0600	0.7200	7.1000	101.0000	28.4000	0.0030	0.2300
002404:	000127.04	197301241100CW.BBA	59.400	1.0500	0.7200	6.3000	110.0000	29.2000	0.0160	0.1900
002405:	000127.04	197304111100CW.BBA	60.200	1.1300	0.8100	6.7000	110.0000	29.2000	0.0180	0.2000
002406:	000127.04	197307251100CW.BBA	67.800	1.1000	0.6300	7.1000	100.0000	31.9000	0.0120	[illegible]
002407:	000127.04	197312061100CW.BBA	50.300	0.7600	0.3800	4.7000	58.0000	18.7000	0.0180	0.2100
002408:	000127.04	197404221100CW.BBA	38.500	0.9300	0.4600	5.0000	93.0000	18.1000	0.0170	0.1600
002409:	000127.04	197408131100CW.BBA	64.200	0.8600	0.4500	6.1000	62.0000	21.8000	0.0070	0.1400
002410:	000127.04	197410141100CW.BBA	63.400	1.1900	0.7800	5.7000	68.0000	23.1000	0.0110	0.1800
002411:	000127.04	197501181100CW.BBA	64.200	0.8300	0.4700	7.3000	123.0000	36.0000	0.0030	0.2300
002412:	000127.04	197503141100CW.BBA	58.000	1.0600	0.6400	6.7000	96.0000	26.5000	0.0040	0.2200
002413:	000127.05	197304111100CW.BBA	56.500	1.0700	0.7100	6.1000	97.0000	26.1000	0.0140	0.1900
002414:	000127.05	197307251100CW.BBA	71.300	1.0700	0.5800	6.9000	89.0000	29.2000	0.0130	0.1800
002415:	000127.05	197312161100CW.BBA	51.300	0.6100	0.3000	4.6000	38.0000	16.6000	0.0140	0.0800
002416:	000127.05	197404221100CW.BBA	38.000	0.9200	0.4100	5.0000	93.0000	18.1000	0.0160	0.1400
002417:	000127.05	197410141100CW.BBA	63.200	0.8600	0.4300	6.1000	62.0000	22.0000	0.1600	0.1500
002418:	000127.05	197501181100CW.BBA	64.200	0.7100	0.4500	5.6000	52.1000	24.1000	0.0100	0.1700
002419:	000127.05	197503141100CW.BBA	62.400	1.1600	0.7500	7.1000	116.0000	34.2000	0.0030	0.2200
002420:	000127.06	197304111100CW.BBA	59.500	1.0000	0.6300	6.3000	91.0000	25.4000	0.0210	0.1400
002421:	000127.06	197307251100CW.BBA	72.000	1.0700	0.5700	6.9000	85.0000	29.0000	0.0140	0.0010
002422:	000127.06	197312061100CW.BBA	49.200	0.6100	0.2800	4.6000	33.0000	16.1000	0.0060	0.0800
002423:	000127.06	197404221100CW.BBA	39.000	0.9000	0.4200	5.0000	92.0000	18.1000	0.0170	0.1400
002424:	000127.08	197301241130CW.BBA	131.000	1.0500	0.2600	6.5000	3.1800	24.0000	0.0060	0.2100
002425:	000127.08	197304111130CW.BBA	100.000	0.7300	0.1900	5.0000	37.0000	16.0000	0.0070	0.1000
002426:	000127.08	197307251130CW.BBA	120.000	1.0700	0.4100	7.3000	37.5000	25.5000	0.0110	0.3000
002427:	000127.08	197312061130CW.BBA	102.000	0.7600	0.2600	5.0000	6.9000	20.4000	0.0060	0.1000
002428:	000127.08	197404221130CW.BBA	112.000	0.9000	0.3700	5.1000	1.5600	21.0000	0.0040	0.2000
002431:	000127.08	197501181130CW.BBA	71.500	1.0400	0.6500	6.6000	98.0000	31.1000	0.0020	0.1700
002432:	000127.08	197503141130CW.BBA	71.600	0.9100	0.5600	6.1000	80.0000	24.2000	0.0040	0.1000
002433:	000127.10	197301241130CW.BBA	130.000	1.0000	0.1900	6.6000	6.9000	25.5000	0.0090	0.2100
002434:	000127.10	197304111130CW.BBA	100.000	0.7300	0.3900	5.0000	3.1800	16.0000	0.0070	0.1000
002435:	000127.10	197307251130CW.BBA	123.000	1.0300	0.1900	7.1000	26.0000	24.8000	0.0150	0.3800
002436:	000127.10	197312061130CW.BBA	87.100	0.4400	0.1800	3.1000	6.2000	11.2000	0.0070	0.0900
002437:	000127.10	197404221130CW.BBA	104.000	0.6300	0.1900	3.3000	6.4000	12.0000	0.0060	0.1300
002438:	000127.10	197408131130CW.BBA	115.000	0.8500	0.2200	5.2000	11.4000	16.4000	0.0250	0.1000
002439:	000127.10	197410141130CW.BBA	93.000	0.8300	0.0400	2.8000	2.6400	1.1000	0.0040	0.0700
002440:	000127.10	197501181130CW.BBA	61.000	1.4000	0.5800	6.2000	80.0000	28.8000	0.0090	0.1900
002441:	000127.10	197503141130CW.BBA	65.200	0.4400	0.5100	5.6000	67.0000	22.0000	0.0050	0.1700
002442:	000127.11	197301241200CW.BBA	102.000	0.4200	0.1100	6.8000	11.6000	13.8000	0.0360	0.1400
002443:	000127.11	197304111200CW.BBA	80.000	0.2300	0.0600	4.8000	4.7600	11.3000	0.0080	0.0900
002444:	000127.11	197307251200CW.BBA	85.000	0.2400	0.0600	3.1000	1.8100	4.7800	0.0060	0.0700
002445:	000127.11	197312061200CW.BBA	73.000	0.3400	0.1100	4.2000	15.8000	16.8000	0.0120	[illegible]
002446:	000127.11	197404221200CW.BBA	73.000	0.3200	0.1200	2.4700	2.6800	7.8000	0.0070	[illegible]
002447:	000127.11	197408131200CW.BBA	110.000	0.4600	0.1100	2.4000	3.1400	10.0000	0.0030	[illegible]
002448:	000127.11	197410141200CW.BBA	93.400	0.0300	0.0100	2.6000	2.5600	0.4100	0.0030	[illegible]
002449:	000127.11	197501181200CW.BBA	70.800	0.6700	0.3600	4.7000	39.0000	19.2000	0.0210	0.1100
002450:	000127.11	197503141200CW.BBA	72.600	0.6100	0.3500	4.4000	41.0000	15.8000	0.0080	0.1400
002453:	000127.12	197307251200CW.BBA	65.200	0.2400	0.0780	2.2000	0.1200	1.1000	0.0060	0.0500
002458:	000127.13	197307251230CW.BBA	66.300	0.0800	0.0250	1.2500	1.0500	3.0500	0.0120	0.0600
002463:	000127.13	197501181230CW.BBA	61.000	0.4000	0.2200	2.9000	7.3000	11.9000	0.0270	0.0900
002464:	000127.13	197503141230CW.BBA	65.200	0.4400	0.2500	3.2000	18.3000	11.5000	0.0110	0.1000
002668:	000022.01	197303011455CW.BBK	58.600	1.0800	0.7500	6.0000	110.0000	28.0000	0.0090	
002669:	000022.01	197306301300CW.BBK	57.500	1.0400	0.7000	6.0000	98.0000	30.8000	0.0070	
002670:	000022.01	197310121300CW.BBK	68.400	1.2000	0.8200	7.5000	126.0000	36.7000	0.0040	
002671:	000022.01	197401191300CW.BBK	53.000	1.3100	0.5200	5.3000	82.0000	25.0000	0.1040	
002677:	000127.07	197301241030CW.BBL	133.000	1.0500	0.2800	6.5000	2.1300	24.3000	0.0050	0.2200
002678:	000127.07	197304111030CW.BBL	123.000	0.7100	0.1350	5.4000	4.5000	9.9000	0.0110	0.1000
002679:	000127.07	197404221030CW.BBL	124.000	1.0100	0.1400	5.2000	4.1000	8.2000	0.0030	0.2100
002680:	000127.07	197307251030CW.BBL	145.000	1.0700	0.2900	7.1000	6.3000	22.7000	0.0100	0.2000
002681:	000127.07	197312061030CW.BBL	138.000	0.8300	0.2300	5.4000	7.8000	20.6000	0.0070	0.0900
002682:	000127.09	197301241030CW.BBL	132.000	1.0000	0.2500	6.4000	12.6500	23.4000	0.0080	0.2200
002683:	000127.09	197304111030CW.BBL	110.000	0.7400	0.2500	5.2000	2.4100	17.5000	0.0060	0.1200
002684:	000127.09	197304221030CW.BBL	108.000	0.7100	0.2100	4.1000	5.3000	12.6000	0.0030	0.1400
002685:	000127.09	197307251030CW.BBL	120.000	1.0900	0.4050	7.2000	32.0000	25.2000	0.0140	0.1800
002686:	000127.09	197312061030CW.BBL	86.900	0.4600	0.1570	3.1000	7.4000	11.2000	0.0070	0.0800
002691:	000127.02	197210231100CW.BBM	56.400	0.9900		6.1500	91.6000	22.4000	0.1100	0.2000
002692:	000127.03	197210231100CW.BBM	60.700	1.0200		6.5000	95.0000	21.4000	0.2400	0.1700
002693:	000127.04	197210231100CW.BBM	60.700	1.0200		6.5000	95.0000	21.4000	0.2400	0.1700
002694:	000127.05	197210231100CW.BBM	59.800	1.0100		6.5500	93.0000	21.2000	0.0300	0.1700
002695:	000127.06	197210231100CW.BBM	60.700	1.0000		6.5000	87.0000	22.2000	0.2300	0.1400
002696:	000127.07	197210231100CW.BBM	113.000	1.2000		7.4000	21.6000	20.2000	0.1900	0.2000
002697:	000127.08	197210231100CW.BBM	108.000	1.2000		7.3000	28.0000	20.6000	0.0200	0.2100
002698:	000127.09	197210231100CW.BBM	108.000	1.1800		7.3000	25.2000	20.4000	0.0200	0.1700
002699:	000127.10	197210231100CW.BBM	107.000	1.1500		7.3000	17.4000	20.2000	0.0300	0.1900
002700:	000127.11	197210231100CW.BBM	100.000	0.7800		6.8000	14.6000	13.8000	0.0300	0.1400
MINIMUM			0.540	0.0250	0.0020	0.0140	0.0400	0.3200	0.0010	0.0010
MAXIMUM			556.000	303.0000	323.0000	544.0000	23000.0000	3130.0000	260.0000	20.0000
MEAN			164.418	12.8039	7.7594	29.2071	583.1935	121.1724	1.1393	0.8257
STD DEVIATION			149.729	41.8276	40.3100	55.6714	1898.1586	314.6458	15.9078	1.7051

Stack is pH-fld <= 4.0 (269 dataset)

Sheet C.1

serid	sarnumber	date time file	Ni	PO_4-P	NH_4-N	NO_3-N	SO_4-S	Cl	Si	O_2-satn
000456:	000152.10	197508121440CD.AK	0.4500	0.0140	1.8500	1.2100	1794.000	40.0000	34.5000	
000457:	000152.20	197508121310CD.AK	0.6000	<.01	2.1600	1.3900	900.000	32.0000	36.0000	
000458:	000153.01	197508121520CD.AK								
000459:	000154.01	197508121400CD.AK	0.8100	0.0810	3.0600	1.1500	975.000	26.5000	19.5000	
000492:	000066.01	197510081720CD.AM	0.2200	0.3670	<.01	1.1200	221.600	10.5000	0.5000	98.00
000493:	002003.01	197510070940CD.AM	0.3400	0.0100	<.01	<.01	534.000	15.5000	1.7000	
000494:	002004.01	197510071000CD.AM	0.3600	<.01	<.01	<.01	539.000	13.1000	2.0000	
000495:	002006.01	197510071200CD.AM	0.7000	0.1630	0.2400	0.8400	1496.800	15.0000	11.2000	
000496:	002007.01	197510071400CD.AM	0.5900	0.0470	0.1800	<.01	1903.700	19.7000	18.4000	
000497:	002007.02	197510071300CD.AM	0.5900	0.0820	0.0300	0.1000	1816.500	18.4000	15.2000	
000498:	002007.03	197510071415CD.AM	0.6200	0.0500	0.2600	0.3300	1852.800	19.8000	9.0000	
000505:	002014.01	197510071030CD.AM	0.3800	0.0230	<.01	<.01	1467.000	14.1000	7.8000	
000506:	002015.10	197510071600CD.AM	0.0500	<.01	<.01	0.9100	130.800	16.9000	0.5000	
000605:	001004.01	197702091045CD.AO	6.0000	1.5000			3116.000	1800.0000	4.8000	
000608:	009001.01	197705051500CD.AR	0.0800							
000612:	009004.01	197705141500CD.AR	0.7800							
000613:	009005.01	197705151130CD.AR	0.3600							
000614:	009006.01	197705151400CD.AR	0.6400							
000615:	009007.10	197705171620CD.AR	2.2300							
000616:	009008.45	197706111400CD.AR	<0.01					70.0000		
000617:	009008.48	197706111600CD.AR	0.0200					60.0000		
000618:	009008.50	197706111650CD.AR	0.0300					60.0000		
000619:	009008.52	197706121120CD.AR	0.0300					60.0000		
000620:	009009.20	197706111100CD.AR								
000621:	009009.20	197706111101CD.AR	0.0200					265.0000		
000622:	009010.62	197706111150CD.AR						110.0000		
000623:	009011.01	197706120720CD.AR	0.0400					100.0000		
000624:	009012.01	197706120730CD.AR	0.0300					110.0000		
000625:	009013.03	197706120800CD.AR	0.0100					60.0000		
000628:	009016.99	197706231345CD.AR	0.1800							
000746:	003016.01	197806301300CD.AW	0.4500							
000747:	003016.03	197806301315CD.AW								
000757:	000213.01	197805281100CD.AX	0.2600							
000758:	000214.21	197805281500CD.AX	0.0800							
000759:	000215.01	197805311500CD.AX	0.3000							
000760:	000215.15	197806011200CD.AX	0.2900							
000761:	000216.15	197806011500CD.AX	0.1700							
000762:	000216.50	197806021100CD.AX	0.0700							
000763:	000217.01	197806021400CD.AX	0.1500							
000764:	000218.01	197806031100CD.AX	0.0800							
000765:	000220.03	197806151500CD.AX								
000766:	000220.15	197806151530CD.AX								
000767:	000221.05	197806151600CD.AX								
000891:	000100.30	197506081700CW.ABBA1	0.1200	<.01	<.01	0.4300	68.000	81.3000	12.9500	
000895:	000109.01	197506081805CW.ABBA1	0.0400	<.01	<.01	0.4500	27.600	72.3000	106.5000	
001198:	000119.01	197409191920CW.BK		0.0250	<.01	0.6300	0.840	11.2500	1.0000	
001199:	000120.01	197409191930CW.BK		0.0900	<.01	0.4200	1.140	23.7500	0.9000	
002120:	000100.30	197803301330CW.BBE		<.010						
002224:	000127.01	197307041030CW.BBB	0.4800	0.4300	0.6400	0.5000	412.000	86.0000	39.0000	85.00
002225:	000127.02	197307041030CW.BBB	0.0500	0.4500	0.4600	0.5000	385.000	74.0000	35.0000	95.00
002226:	000127.03	197307041100CW.BBB	0.0300	0.1700	0.3300	0.5000	387.000	52.0000	36.0000	86.00
002227:	000127.04	197307041100CW.BBB	0.0400	0.0100	0.2500	0.5000	605.000	23.0000	44.0000	95.00
002228:	000128.01	197307081030CW.BBB	2.7800	0.1000	0.6000	0.5000	1265.000	27.0000	50.0000	55.00
002229:	000128.02	197307081030CW.BBB	3.3200	0.2400	3.0000	0.5000	2383.000	21.0000	67.0000	65.00
002230:	000128.03	197307081100CW.BBB	3.8400	0.0400	2.6000	0.5000	1950.000	42.0000	69.0000	55.00
002231:	000130.01	197307061030CW.BBB	7.4500	7.2000	0.1500	0.5000	2145.000	155.0000	97.0000	60.00
002232:	000130.02	197307061030CW.BBB	2.2400	1.0000	9.3000	0.5000	4080.000	80.0000	55.0000	40.00
002233:	000130.03	197307061100CW.BBB	3.2600	44.0000	10.3000	0.5000	2028.000	65.0000	47.0000	50.00
002234:	000131.01	197307061200CW.BBB	2.7200	1.4300	0.7300	0.5000	4875.000	28.0000	33.5000	90.00
002235:	000131.02	197307061200CW.BBB	2.8400	1.2000	0.6500	0.5000	5304.000	27.0000	35.0000	80.00
002236:	000131.03	197307061230CW.BBB	3.3000	0.3100	1.0500	0.5000	3978.000	32.0000	35.0000	81.00
002237:	000132.01	197309101030CW.BBB	3.0400	0.0500	2.5000	0.5000	4485.000	18.0000	28.5000	30.00
002238:	000132.02	197309101030CW.BBB	1.9000	0.0500	2.4800	0.5000	2535.000	14.0000	4.5000	90.00
002239:	000132.03	197309101100CW.BBB	8.4100	0.8600	2.1000	0.7000	3413.000	18.0000	6.2000	80.00
002240:	000132.04	197309101100CW.BBB	2.0300	0.1200	2.4000	0.5000	2730.000	14.0000	5.5000	90.00
002241:	000132.05	197309101130CW.BBB	1.9500	0.1400	2.5100	0.5000	2418.000	31.0000	4.5000	79.00
002242:	000132.06	197309101130CW.BBB	1.8800	0.1300	2.3000	0.5000	2340.000	21.0000	6.0000	84.00
002243:	000132.07	197309101200CW.BBB	1.8000	0.1000	2.1000	0.5000	2730.000	25.0000	6.5000	95.00
002244:	000133.01	197309171030CW.BBB	0.0500	0.0500	0.3100	0.5000	429.000	40.0000	6.5000	59.00
002247:	000134.03	197303151100CW.BBB	1.1500	0.1400	2.1000	0.5000	3042.000	44.0000	44.0000	60.00
002249:	000135.01	197307181030CW.BBB	4.2800	0.2600	6.7500	0.5000	4455.000	27.0000	21.0000	70.00
002250:	000135.02	197307181030CW.BBB	3.9100	0.1500	7.1000	0.5000	3660.000	27.0000	29.5000	75.00
002251:	000135.03	197307181100CW.BBB	4.0500	0.1500	3.7500	0.5000	4200.000	29.0000	26.0000	80.00
002252:	000135.04	197307181100CW.BBB	0.4500	0.0100	1.6000	0.8000	685.000	33.0000	9.0000	85.00
002253:	000136.01	197309111030CW.BBB	5.1000	0.2500	1.0500	0.5000	3042.000	44.0000	44.5000	95.00
002254:	000137.01	197309111100CW.BBB	0.6300	0.1000	2.6000	0.5000	1028.000	14.0000	13.5000	90.00
002255:	000137.02	197309111100CW.BBB	0.5600	0.0800	2.3800	0.5000	1250.000	18.0000	12.8000	90.00
002256:	000137.03	197309111130CW.BBB	0.4200	0.1200	2.4100	0.5000	1450.000	21.0000	14.3000	92.00
002257:	000138.01	197310131030CW.BBB	0.5300	0.1000	0.1400	0.5000	936.000	29.5000	11.6000	10.00
002258:	000138.02	197310131030CW.BBB	0.8600	0.0800	0.1000	0.5000	468.000	33.7000	4.6000	70.00
002259:	000138.03	197310131100CW.BBB	0.5300	0.0850	0.0500	2.9000	429.000	33.7000	7.6000	70.00
002260:	000139.01	197309111230CW.BBB	9.7000	0.8200	1.5000	0.5000	4075.000	151.0000	37.5000	56.00
002261:	000139.02	197309111230CW.BBB	50.4000	0.5800	10.8000	0.5000	5304.000	168.0000	35.5000	40.00
002262:	000139.03	197309111300CW.BBB	11.4000	0.5000	2.1800	0.5000	3378.000	225.0000	114.0000	60.00
002263:	000140.01	197309111130CW.BBB	4.6000	0.1300	1.9000	1.0000	2738.000	88.0000	57.0000	37.00
002264:	000140.02	197309111200CW.BBB	2.4000	0.0500	5.1000	3.2000	2028.000	149.0000	36.5000	43.00
002265:	000140.03	197309111200CW.BBB	2.0000	0.6800	4.4000	4.0000	1677.000	170.0000	34.5000	80.00
002266:	000157.05	197307061100CW.BBB	11.3000	76.0000	0.2300	0.5000	2730.000	145.0000	111.0000	30.00
002267:	000157.06	197307061130CW.BBB	1.2000	1.8000	10.5000	0.5000	2028.000	60.0000	40.5000	40.00

Stack is pH-fld <= 4.0 (269 dataset)

Sheet C.2

serid	samnumber	date time file	Ni	PO_4-P	NH_4-N	NO_3-N	SO_4-S	Cl	Si	O_2-satn
002268:	000157.07	197307061130CW.BBB	4.6400	1.0500	0.6600	0.5000	1677.000	28.0000	65.5000	40.00
002269:	000158.08	197309101200CW.BBB	1.6800	0.0500	2.0500	0.5000	2145.000	21.0000	37.5000	89.00
002270:	000159.09	197309101230CW.BBB	1.0200	0.0600	2.2500	0.5000	4480.000	25.0000	24.0000	18.00
002271:	000160.05	197307181130CW.BBB	3.3400	0.0700	5.1000	0.5000	2065.000	56.0000	20.0000	54.00
002272:	000160.06	197307181130CW.BBB	3.5000	0.1300	3.9100	0.5000	1570.000	132.0000	25.0000	68.00
002273:	000127.01	197402141030CW.BBC	1.3000	0.2600	0.3000	0.4000	312.000	27.8000	32.5000	95.00
002274:	000127.02	197402141030CW.BBC	1.2100	0.3000	0.4000	0.2100	273.000	26.9000	32.0000	95.00
002275:	000127.03	197402141100CW.BBC	1.1400	0.2800	0.4200	0.3100	250.000	26.8000	31.2000	98.00
002276:	000127.04	197402141100CW.BBC	0.8000	0.1400	0.5500	0.3200	429.000	24.6000	37.0000	96.00
002277:	000128.01	197402101030CW.BBC	4.5200	0.1400	3.9000	0.5600	1794.000	59.1000	53.0000	60.00
002278:	000128.02	197402101030CW.BBC	4.1100	0.1500	2.1000	0.4300	1632.000	60.2000	56.0000	82.00
002279:	000128.03	197402101100CW.BBC	6.2100	0.8800	3.0800	0.4000	2106.000	29.5000	77.0000	80.00
002280:	000129.01	197402111030CW.BBC	2.5300	0.3800	1.8000	0.2600	1638.000	33.7000	43.5000	26.00
002281:	000129.02	197402111030CW.BBC	6.5200	0.0800	0.7900	0.2600	1833.000	35.9000	56.0000	42.00
002282:	000130.01	197402121030CW.BBC	5.8000	1.6400	1.3000	2.0400	5889.000	128.0000	77.5000	85.00
002283:	000130.02	197402121030CW.BBC	2.2200	1.3600	0.6800	0.5500	3549.000	25.3000	25.0000	85.00
002284:	000130.03	197402121100CW.BBC	3.4300	1.6500	0.8000	0.3700	2340.000	141.0000	43.5000	87.00
002285:	000130.04	197402121100CW.BBC	5.7100	0.2800	4.4800	0.3800	1404.000	46.4000	37.0000	80.00
002286:	000131.01	197402121130CW.BBC	4.2000	1.5400	1.3700	0.1500	3705.000	21.1000	50.5000	48.00
002287:	000131.02	197402121130CW.BBC	4.5100	1.6000	1.5000	0.6000	1833.000	23.2000	53.5000	85.00
002288:	000131.03	197402121200CW.BBC	4.3200	1.5600	1.1000	0.5600	1638.000	18.9000	46.5000	92.00
002289:	000132.01	197402201030CW.BBC	2.2300	0.1500	0.8200	2.1500	897.000	25.3000	15.0000	89.00
002290:	000132.02	197402201030CW.BBC	3.3100	0.1200	1.4400	0.6500	887.000	27.4000	13.5000	100.00
002291:	000132.03	197402201100CW.BBC	12.5000	3.7000	1.7800	2.1000	3939.000	27.4000	11.0000	73.00
002292:	000132.04	197402201100CW.BBC	3.1000	0.2900	1.6600	1.5300	1131.000	23.2000	17.0000	97.00
002293:	000132.05	197402201130CW.BBC	3.3200	0.0500	1.8000	0.3000	975.000	25.3000	38.0000	60.00
002294:	000132.06	197402201130CW.BBC	3.9100	0.1400	0.7800	0.7000	702.000	29.5000	23.0000	42.00
002295:	000132.07	197402201200CW.BBC	3.2600	0.2500	0.9200	2.9000	975.000	29.5000	24.0000	98.00
002296:	000133.01	197402201230CW.BBC	2.0100	0.3100	3.1000	0.4000	936.000	27.4000	25.0000	0.00
002297:	000134.01	197402121200CW.BBC	6.5000	0.2100	1.4100	2.5000	2416.000	29.5000	47.0000	48.00
002298:	000134.02	197402121200CW.BBC	2.7000	0.0600	1.3000	0.3000	1090.000	50.7000	50.0000	60.00
002299:	000134.03	197402121230CW.BBC	2.3300	0.0600	3.5000	0.2600	5460.000	92.8000	21.0000	62.00
002300:	000134.04	197402121230CW.BBC	2.5200	0.0400	2.0000	0.2600	3744.000	50.6000	40.5000	78.00
002301:	000135.01	197402221030CW.BBC	7.5400	0.2500	2.6000	0.2200	3822.000	75.9000	26.6000	85.00
002302:	000135.02	197402221030CW.BBC	8.0500	0.3500	2.7800	0.5600	2184.000	65.4000	27.5000	60.00
002303:	000135.03	197402221100CW.BBC	7.0300	0.1000	0.8200	0.4000	2148.000	46.4000	27.0000	72.00
002304:	000135.04	197402221100CW.BBC	3.8000	0.0200	1.5700	0.2800	1365.000	37.9000	13.0000	90.00
002305:	000136.01	197402181030CW.BBC	3.2100	0.8200	1.0000	0.3100	3642.000	32.0000	31.0000	87.00
002306:	000137.01	197402181030CW.BBC	2.5200	0.0600	4.0000	0.3800	1209.000	130.0000	11.0000	97.00
002307:	000137.02	197402181100CW.BBC	2.0100	0.0700	4.1800	0.2600	658.000	255.0000	11.5000	100.00
002308:	000137.03	197402181100CW.BBC	1.9100	0.1700	5.3000	0.2600	936.000	175.0000	18.5000	90.00
002309:	000138.01	197403031030CW.BBC	2.1200	0.1500	0.3000	0.2400	508.000	32.1000	26.0000	0.00
002310:	000138.02	197403031030CW.BBC	0.8100	0.1300	0.3600	0.3100	507.000	26.1000	25.0000	0.00
002311:	000138.03	197403031100CW.BBC	1.0200	0.1300	0.5100	0.2600	390.000	27.3000	25.2000	0.00
002312:	000139.01	197402191030CW.BBC	7.6200	1.4200	0.7000	0.2500	2340.000	67.5000	21.5000	95.00
002313:	000139.02	197402191030CW.BBC	27.0000	8.3000	4.3600	0.5000	8580.000	223.0000	72.0000	87.00
002314:	000139.03	197402191100CW.BBC	14.0000	5.6000	3.5000	2.0000	4680.000	80.2000	45.5000	92.00
002315:	000140.01	197402181130CW.BBC	4.6000	0.1000	4.3600	0.7600	2379.000	173.0000	59.0000	92.00
002316:	000140.02	197402181130CW.BBC	5.9300	0.0300	4.5500	0.4000	3080.000	156.0000	67.5000	12.00
002317:	000140.03	197402181200CW.BBC	5.3100	0.0600	5.2000	0.5500	1950.000	173.0000	62.0000	89.00
002318:	000158.08	197402201200CW.BBC	3.0500	0.2900	1.1000	0.3800	1014.000	29.5000	26.0000	101.00
002319:	000159.09	197402201230CW.BBC	5.5000	0.5700	1.4100	0.5800	1443.000	29.0000	41.0000	106.00
002320:	000160.05	197402221130CW.BBC	6.7100	1.9000	0.6000	0.2200	1599.000	23.2000	31.5000	88.00
002321:	000160.06	197402221130CW.BBC	6.2800	2.1000	2.2400	3.0000	1755.000	23.2000	32.5000	77.00
002322:	000127.01	197212121033CW.BBD	0.4200	0.3600	0.3500	0.5000	283.000	20.0000	59.5000	65.00
002323:	000127.01	197301101033CW.BBD	0.4500	0.3700	0.4800	0.5000	323.000	18.8000	55.5000	60.00
002324:	000127.01	197302101033CW.BBD	0.5000	0.3500	0.3800	0.5000	313.000	22.5000	48.5000	74.00
002325:	000127.01	197303131033CW.BBD	0.5400	0.3800	0.3700	0.5000	311.000	17.5000	44.5000	65.00
002326:	000127.01	197304121033CW.BBD	0.5000	0.3800	0.3700	0.5000	460.000	22.5000	42.0000	65.00
002327:	000127.01	197305111033CW.BBD	0.4100	0.4300	0.3500	0.5000	308.000	21.5000	36.0000	73.00
002328:	000127.01	197306141033CW.BBD	0.4600	0.3600	0.2800	0.5000	107.000	22.0000	33.5000	84.00
002329:	000127.01	197307251033CW.BBD	0.4300	0.4300	0.5400	0.5000	321.000	20.2000	39.0000	85.00
002330:	000127.01	197308241033CW.BBD	0.4700	0.5400	0.6800	0.6500	546.000	14.7000	37.7000	66.00
002331:	000127.01	197309201033CW.BBD	0.4600	0.5100	0.6200	0.7900	451.000	27.4000	38.2000	68.00
002332:	000127.01	197310181033CW.BBD	0.5700	0.5600	0.4100	0.8100	378.000	25.4000	38.4000	70.00
002333:	000127.01	197311161033CW.BBD	0.4300	0.6200	0.1900	1.0600	305.000	29.5000	40.1000	70.00
002334:	000127.01	197312121033CW.BBD	0.5100	0.2400	0.6800	0.8100	361.000	24.6000	56.8000	65.00
002335:	000127.01	197401151033CW.BBD	0.5200	0.1000	2.3000	0.7500	468.000	25.4000	61.5000	63.00
002336:	000127.01	197402181033CW.BBD	0.4100	0.2600	0.4000	0.4500	312.000	27.8000	32.5000	70.00
002337:	000127.01	197403121033CW.BBD	0.4600	0.3300	0.4200	0.5300	410.000	26.0000	36.6000	79.00
002338:	000127.01	197404221033CW.BBD	0.4100	0.6300	0.3000	0.6300	415.000	23.0000	52.0000	74.00
002339:	000127.01	197405201033CW.BBD	0.2700	0.2000	0.6500	0.5400	554.000	33.8000	38.5000	80.00
002340:	000127.01	197406171033CW.BBD	0.5000	0.4100	0.4000	0.5200	601.000	25.3000	30.5000	77.00
002341:	000127.01	197407281034CW.BBD	0.2300	0.1900	0.7200	0.4200	638.000	26.0000	32.4000	81.00
002342:	000127.01	197408131033CW.BBD	0.5200	0.1800	0.8200	0.6500	554.000	24.0000	43.5000	87.00
002343:	000127.01	197409121233CW.BBD	0.4100	0.2500	0.2000	0.4700	851.000	32.1000	40.5000	76.00
002344:	000127.01	197410141233CW.BBD	0.4400	0.3100	0.1400	0.6000	633.000	30.0000	41.0000	82.00
002345:	000127.01	197411161033CW.BBD	0.4200	0.2900	0.4100	0.4500	330.000	26.6000	40.0000	84.00
002346:	000127.01	197412171033CW.BBD	0.5200	0.4200	0.3500	0.3500	357.000	27.2000	39.0000	80.00
002347:	000127.01	197501181033CW.BBD	0.5800	0.3200	0.6500	0.7000	575.000	25.0000	36.0000	78.00
002348:	000127.01	197502151033CW.BBD	0.0400	0.1000	0.8000	0.2500	500.000	21.0000	33.0000	80.00
002349:	000127.01	197503141033CW.BBD	0.0500	0.1400	0.7500	0.2500	495.000	20.2000	33.5000	87.00
002379:	000125.01	197304111130CW.BBA	0.4300	0.0100	0.1500	0.4100	605.000	25.0000	30.5000	91.00
002380:	000125.01	197307251130CW.BBA	0.2700	0.0200	0.2500	0.0600	605.000	23.5000	44.0000	95.00
002381:	000125.01	197312061130CW.BBA	0.5200	[illegible]	0.5000	0.4000	600.000	27.5000	61.5000	80.00
002382:	000125.01	197404221130CW.BBA	0.4300	0.0100	0.2400	0.4100	600.000	25.1000	52.0000	89.00
002383:	000125.01	197408131130CW.BBA	0.3700	0.0100	0.1200	0.3000	504.000	24.1000	60.0000	82.00
002384:	000125.01	197410141130CW.BBA	0.3700	0.0300	1.1000	0.4500	850.000	25.2000	23.5000	91.00
002385:	000125.01	197501181130CW.BBA	0.3200	0.0200	1.1000	0.9000	825.000	24.2000	34.5000	79.00
002386:	000125.01	197503141130CW.BBA	0.3200	0.0300	0.6000	0.6000	550.000	26.1000	32.0000	82.00

Stack is pH-fld <= 4.0 (269 dataset)

Sheet C.3

serid	samnumber	date time file	Ni	PO_4-P	NH_4-N	NO_3-N	SO_4-S	Cl	Si	O_2-satn
002387:	000127.01	197301241030CW.BBA	0.4900	0.4200	0.2400	0.2800	450.000	25.0000	53.0000	60.00
002388:	000127.01	197304111030CW.BBA	0.5000	0.3600	0.2700	0.2700	440.000	24.1000	42.0000	65.00
002389:	000127.01	197312061030CW.BBA	0.5200	0.1000	2.3000	0.7500	420.000	22.5000	61.5000	63.00
002390:	000127.02	197301241030CW.BBA	0.4700	0.4100	0.2100	0.1500	452.000	24.0000	54.0000	85.00
002391:	000127.02	197304111030CW.BBA	0.4500	0.3700	0.5500	0.2300	443.000	40.0000	41.0000	86.00
002392:	000127.02	197307251030CW.BBA	0.4200	0.4600	0.4600	0.1500	400.000	19.8000	36.0000	95.00
002393:	000127.02	197312161030CW.BBA	0.5100	0.0750	1.5000	0.7500	410.000	15.0000	50.0000	85.00
002394:	000127.02	197404221030CW.BBA	0.5100	0.0120	0.2900	0.6000	420.000	20.1000	51.4000	85.00
002395:	000127.03	197304111030CW.BBA	0.4200	0.3900	0.2600	0.2600	445.000	25.5000	40.0000	85.00
002396:	000127.03	197307251030CW.BBA	0.2600	0.3900	0.3300	0.1300	371.000	26.8000	35.0000	92.00
002397:	000127.03	197312061030CW.BBA	0.5200	0.0800	1.5000	0.9000	390.000	17.5000	52.0000	86.00
002398:	000127.03	197401241030CW.BBA	0.4900	0.4100	0.2000	0.1000	451.000	23.0000	52.0000	86.00
002399:	000127.03	197404221030CW.BBA	0.5000	0.1000	0.2300	0.4100	430.000	14.2000	48.2000	85.00
002400:	000127.03	197408131030CW.BBA	0.6300	0.2400	0.1200	0.5500	620.000	29.8000	37.5000	87.00
002401:	000127.03	197410141030CW.BBA	0.5100	0.3200	0.7000	0.3000	495.000	19.5000	29.5000	83.00
002402:	000127.03	197501181030CW.BBA	0.4600	0.1200	0.6000	0.7000	480.000	20.1000	33.5000	85.00
002403:	000127.03	197503141030CW.BBA	0.5200	0.1400	0.5600	0.4600	455.000	22.1000	38.6000	89.00
002404:	000127.04	197301241100CW.BBA	0.4900	0.0900	0.2400	0.1500	450.000	23.0000	51.0000	86.00
002405:	000127.04	197304111100CW.BBA	0.4200	0.0500	0.4000	0.2900	440.000	35.5000	40.5000	86.00
002406:	000127.04	197307251100CW.BBA	0.2600	0.3500	0.3600	0.0800	385.000	32.5000	35.0000	70.00
002407:	000127.04	197312061100CW.BBA	0.5200	0.0700	0.8500	0.2500	400.000	16.8000	51.0000	80.00
002408:	000127.04	197404221100CW.BBA	0.5000	0.1000	0.1000	0.2100	389.000	14.1000	57.0000	84.00
002409:	000127.04	197408131100CW.BBA	0.3600	0.0800	1.0000	0.2200	440.000	19.3000	31.0000	84.00
002410:	000127.04	197410141100CW.BBA	0.4400	0.1800	0.5800	0.5500	577.000	17.2000	27.0000	88.00
002411:	000127.04	197501181100CW.BBA	0.5000	0.3300	0.7000	0.3100	465.000	12.5000	34.1000	87.00
002412:	000127.04	197503141100CW.BBA	0.4500	0.1100	0.5600	0.6500	445.000	18.2000	33.5000	88.00
002413:	000127.05	197304111100CW.BBA	0.4100	0.4000	0.3400	0.1000	440.000	25.5000	34.5000	87.00
002414:	000127.05	197307251100CW.BBA	0.2000	0.1700	0.3300	0.0500	385.000	24.8000	36.0000	85.00
002415:	000127.05	197312161100CW.BBA	0.5100	0.0300	0.7000	1.0000	390.000	17.5000	39.5000	84.00
002416:	000127.05	197404221100CW.BBA	0.4900	0.1000	0.1100	0.1800	380.000	14.0000	56.0000	85.00
002417:	000127.05	197410141100CW.BBA	0.3200	0.0500	0.8000	0.2300	430.000	16.6000	52.1000	81.00
002418:	000127.05	197501181100CW.BBA	0.4300	0.1900	0.8200	0.3200	500.000	19.3000	24.0000	82.00
002419:	000127.05	197503141100CW.BBA	0.4900	0.4100	0.4000	0.3800	412.000	12.5000	34.5000	84.00
002420:	000127.06	197304111100CW.BBA	0.4000	0.1000	0.1900	0.0700	429.000	25.0000	35.5000	91.00
002421:	000127.06	197307251100CW.BBA	0.2100	0.0400	0.3400	0.0500	389.000	26.8000	37.0000	95.00
002422:	000127.06	197312061100CW.BBA	0.5000	0.0300	0.8000	1.0800	401.000	27.5000	39.0000	85.00
002423:	000127.06	197404221100CW.BBA	0.4900	0.0800	0.1000	0.1200	380.000	14.0000	56.0000	84.00
002424:	000127.08	197301241130CW.BBA	0.3600	0.0300	0.1300	0.5100	510.000	21.0000	34.0000	82.00
002425:	000127.08	197304111130CW.BBA	0.2100	<.01	0.3700	1.8600	495.000	35.0000	20.0000	93.00
002426:	000127.08	197307251130CW.BBA	0.1800	0.0200	0.6000	0.2600	520.000	46.2000	33.5000	85.00
002427:	000127.08	197312061130CW.BBA	0.3600	0.0300	0.3100	0.8700	500.000	20.0000	18.0000	83.00
002428:	000127.08	197404221130CW.BBA	0.2900	0.0300	0.2400	0.9200	510.000	37.0000	22.0000	84.00
002431:	000127.08	197501181130CW.BBA	0.4300	0.3000	0.6500	0.3000	410.000	24.5000	31.5000	85.00
002432:	000127.08	197503141130CW.BBA	0.3900	0.1600	0.4500	0.2200	302.000	26.8000	31.0000	86.00
002433:	000127.10	197301241130CW.BBA	0.3600	0.0400	0.3100	0.4100	495.000	26.0000	20.0000	85.00
002434:	000127.10	197304111130CW.BBA	0.2100	<.01	0.3700	1.8600	540.000	35.0000	20.0000	82.00
002435:	000127.10	197307251130CW.BBA	0.1400	0.0420	0.3800	0.1800	413.000	25.0000	30.0000	43.00
002436:	000127.10	197312061130CW.BBA	0.3600	0.0100	0.9000	3.2000	420.000	22.0000	13.5000	85.00
002437:	000127.10	197404221130CW.BBA	0.2000	0.0200	0.4500	0.9800	425.000	33.0000	16.0000	84.00
002438:	000127.10	197408131130CW.BBA	0.2500	0.2300	1.0500	0.3200	644.000	31.0000	12.4000	78.00
002439:	000127.10	197410141130CW.BBA	0.0700	0.2000	0.4100	0.5400	590.000	32.0000	14.3000	82.00
002440:	000127.10	197501181130CW.BBA	0.4300	0.1900	0.6500	0.7500	495.000	28.3000	16.0000	84.00
002441:	000127.10	197503141130CW.BBA	0.3700	0.1200	0.6300	0.4300	301.000	26.4000	16.8000	87.00
002442:	000127.11	197301241200CW.BBA	0.3000	0.0200	0.3200	0.4700	532.000	25.0000	22.0000	88.00
002443:	000127.11	197304111200CW.BBA	0.1700	0.0100	0.1600	1.4400	291.000	26.0000	8.5000	85.00
002444:	000127.11	197307251200CW.BBA	0.1000	0.0350	0.3300	0.0800	400.000	21.0000	22.0000	90.00
002445:	000127.11	197312061200CW.BBA	0.0500	0.1500	0.9000	2.5000	320.000	17.5000	11.0000	84.00
002446:	000127.11	197404221200CW.BBA	0.1700	0.0100	0.3100	1.1000	365.000	28.5000	15.2000	88.00
002447:	000127.11	197408131200CW.BBA	0.1200	0.1300	0.8500	0.4600	388.000	29.1000	9.9000	80.00
002448:	000127.11	197410141200CW.BBA	0.0500	0.2500	0.5400	0.6100	330.000	24.3000	12.4000	87.00
002449:	000127.11	197501181200CW.BBA	0.2700	0.0500	0.7100	0.6000	301.000	25.3000	11.3000	84.00
002450:	000127.11	197503141200CW.BBA	0.2600	0.1800	0.7300	0.5000	420.000	26.4000	14.0000	95.00
002453:	000127.12	197307251230CW.BBA	0.0900	0.0200	0.7000	0.0500	304.000	29.4000	19.1000	88.00
002458:	000127.13	197307251230CW.BBA	0.0040	0.0100	0.4400	0.0500	275.000	27.2000	15.0000	93.00
002463:	000127.13	197501181230CW.BBA	0.1800	0.0600	0.6200	0.7000	302.000	23.3000	8.1000	82.00
002464:	000127.13	197503141230CW.BBA	0.1000	0.0200	0.2500	0.6000	170.000	20.1000	10.2000	98.00
002668:	000022.01	197303011400CW.BBK		0.2450	0.2800	<0.40		22.1000	42.0000	
002669:	000022.01	197306301300CW.BBK		0.4020	0.3400	<0.40		20.2000	39.5900	
002670:	000022.01	197310101300CW.BBK		0.2020	0.4000	<0.4		27.3000	38.0000	
002671:	000022.01	197401191300CW.BBK		0.1000	2.3000	0.7500		20.0000	61.5000	
002677:	000127.07	197301241030CW.BBL	0.3700	0.3100	0.1800	0.1800	461.000	22.0000	53.0000	80.00
002678:	000127.07	197304111030CW.BBL	0.2700	0.2100	<.05	0.2200	357.000	55.0000	15.5000	86.00
002679:	000127.07	197304221030CW.BBL	0.2200	0.2300	0.1600	0.1200	382.000	41.0000	18.5000	87.00
002680:	000127.07	197307251030CW.BBL	0.1600	0.0500	0.2800	0.1000	398.000	18.9000	38.0000	94.00
002681:	000127.07	197312061030CW.BBL	0.4900	0.0100	0.7000	0.9500		30.0000	38.0000	86.00
002682:	000127.09	197301241030CW.BBL	0.3400	0.0200	0.1200	0.5200	567.000	23.0000	21.0000	
002683:	000127.09	197304111030CW.BBL	0.2200	<.01	0.2100	1.0900	555.000	30.0000	19.5000	80.00
002684:	000127.09	197304221030CW.BBL	0.2100	0.0200	0.2400	0.8600	410.000	30.0000	14.0000	85.00
002685:	000127.09	197307251030CW.BBL	0.1300	0.0400	0.2600	0.1800	385.000	22.0000	29.0000	90.00
002686:	000127.09	197312061030CW.BBL	0.3700	0.0100	1.0000	2.2500		20.0000	13.5000	84.00
002691:	000127.02	197210231100CW.BBM	0.4300					25.3000	32.0000	
002692:	000127.03	197210231100CW.BBM	0.4200					24.3000	31.5000	
002693:	000127.04	197210231100CW.BBM	0.4200							
002694:	000127.05	197210231100CW.BBM	0.4200							
002695:	000127.06	197210231100CW.BBM	0.4100							
002696:	000127.07	197210231100CW.BBM	0.3600							
002697:	000127.08	197210231100CW.BBM	0.4000							
002698:	000127.09	197210231100CW.BBM	0.4000							
002699:	000127.10	197210231100CW.BBM	0.4100							
002700:	000127.11	197210231100CW.BBM	0.3000							
MINIMUM			0.0040	0.0100	0.0300	0.0500	0.840	10.5000	0.5000	0.00
MAXIMUM			50.4000	76.0000	10.8000	4.0000	8580.000	1800.0000	114.0000	106.00
MEAN			1.8841	0.9346	1.3477	0.6285	1256.399	49.3038	33.6054	76.53
STD DEVIATION			4.1741	5.9430	1.7841	0.6051	1365.970	121.5152	19.3834	20.07

Stack is pH-fld <= 4.0 (269 dataset)

serid	sarnumber	date time file	temp	flow-past	flow-pres
000456:	000152.10	197508121440CD.AK	23.00	1.0	1.0
000457:	000152.20	197508121310CD.AK	26.50	1.0	1.0
000458:	000153.01	197508121520CD.AK	19.20	1.0	1.0
000459:	000154.01	197508121400CD.AK	17.00	1.0	1.0
000492:	000066.01	197510081720CD.AM	10.30	0.0	0.0
000493:	002003.01	197510070940CD.AM	10.60	4.0	4.0
000494:	002004.01	197510071000CD.AM	10.60	0.0	0.0
000495:	002006.01	197510071200CD.AM	11.50	4.0	4.0
000496:	002007.01	197510071400CD.AM	9.00	4.0	4.0
000497:	002007.02	197510071300CD.AM	10.50	4.0	4.0
000498:	002007.03	197510071415CD.AM	13.00	4.0	4.0
000505:	002014.01	197510071030CD.AM	11.90	3.0	3.0
000506:	002015.10	197510071600CD.AM	13.00	4.0	4.0
000605:	001004.01	197702091045CD.AQ	6.50	4.0	4.0
000608:	009001.01	197705051500CD.AR	15.70	3.0	3.0
000612:	009004.01	197705141500CD.AR	11.50	3.0	3.0
000613:	009005.01	197705151130CD.AR	13.50	3.0	3.0
000614:	009006.01	197705151400CD.AR	15.50	3.0	3.0
000615:	009007.10	197705171620CD.AR	24.00	3.0	3.0
000616:	009008.45	197706111400CD.AR	33.00	4.0	4.0
000617:	009008.48	197706111600CD.AR	34.00	4.0	4.0
000618:	009008.50	197706111650CD.AR	26.00	4.0	4.0
000619:	009008.52	197706121120CD.AR	21.00	4.0	4.0
000620:	009009.20	197706111100CD.AR	34.50	3.0	3.0
000621:	009009.20	197706111101CD.AR	29.50	3.0	3.0
000622:	009010.02	197706111150CD.AR	49.00	3.0	3.0
000623:	009011.01	197706120700CD.AR	41.00	3.0	3.0
000624:	009012.01	197706120730CD.AR	49.00	3.0	3.0
000625:	009013.03	197706120800CD.AR	25.00	3.0	3.0
000628:	009016.99	197706231345CD.AR	23.00	3.0	4.0
000746:	003016.01	197806301300CD.AW	10.50	3.0	3.0
000747:	003016.03	197806301315CD.AW	10.50	3.0	3.0
000757:	000213.01	197805201100CD.AX	10.70	3.0	3.0
000758:	000214.21	197805201500CD.AX	19.80	3.0	3.0
000759:	000215.01	197805311500CD.AX	10.00	3.0	2.0
000760:	000215.15	197806011200CD.AX	21.50	3.0	2.0
000761:	000216.15	197806011500CD.AX	29.50	3.0	2.0
000762:	000216.50	197806021100CD.AX	19.50	3.0	2.0
000763:	000217.01	197806021400CD.AX	9.50	3.0	2.0
000764:	000218.01	197806031200CD.AX	12.00	3.0	3.0
000765:	000220.03	197806151500CD.AX	21.50	3.0	3.0
000766:	000220.15	197806151530CD.AX	19.00	3.0	3.0
000767:	000221.05	197806151600CD.AX	19.00	3.0	3.0
000891:	000100.30	197506081740CW.AHHA1	19.50	3.0	3.0
000895:	000109.01	197506081805CW.AHHA1	9.00	1.0	1.0
001198:	000119.01	197409191920CW.BK	11.00	2.0	2.0
001199:	000120.01	197409191930CW.BK	10.00	2.0	2.0
002120:	000100.30	197803301330CW.BHE	5.50	3.0	3.0
002224:	000127.01	197307041030CW.BBB	11.30		
002225:	000127.02	197307041030CW.BBB	9.50		
002226:	000127.03	197307041100CW.BBB	11.20		
002227:	000127.04	197307041100CW.BBB	11.40		
002228:	000128.01	197307081000CW.BBB	18.60		
002229:	000128.02	197307081030CW.BBB	12.60		
002230:	000128.03	197307081100CW.BBB	15.30		
002231:	000130.01	197307061030CW.BBB	19.00		
002232:	000130.02	197307061030CW.BBB	20.00		
002233:	000130.03	197307061100CW.BBB	21.00		
002234:	000131.01	197307061200CW.BBB	20.10		
002235:	000131.02	197307061200CW.BBB	18.50		
002236:	000131.03	197307061230CW.BBB	20.00		
002237:	000132.01	197309101030CW.BBB	20.20		
002238:	000132.02	197309101030CW.BBB	19.30		
002239:	000132.03	197309101100CW.BBB	20.30		
002240:	000132.04	197309101100CW.BBB	18.60		
002241:	000132.05	197309101130CW.BBB	17.90		
002242:	000132.06	197309101130CW.BBB	17.30		
002243:	000132.07	197309101200CW.BBB	19.00		
002244:	000133.01	197309171030CW.BBB	8.50		
002247:	000134.03	197303151100CW.BBB	10.80		
002249:	000135.01	197307181030CW.BBB	16.50		
002250:	000135.02	197307181030CW.BBB	18.00		
002251:	000135.03	197307181100CW.BBB	15.20		
002252:	000135.04	197307181100CW.BBB	12.30		
002253:	000136.01	197309111030CW.BBB	10.30		
002254:	000137.01	197309111100CW.BBB	21.00		
002255:	000137.02	197309111100CW.BBB	21.00		
002256:	000137.03	197309111130CW.BBB	19.50		
002257:	000138.01	197310131030CW.BBB	14.60		
002258:	000138.02	197310131030CW.BBB	10.80		
002259:	000138.03	197310131100CW.BBB	10.70		
002260:	000139.01	197309111230CW.BBB	20.10		
002261:	000139.02	197309111230CW.BBB	18.50		
002262:	000139.03	197309111300CW.BBB	20.00		
002263:	000140.01	197309111130CW.BBB	11.60		
002264:	000140.02	197309111200CW.BBB	14.80		
002265:	000140.03	197309111200CW.BBB	20.00		
002266:	000157.05	197307061100CW.BBB	11.60		
002267:	000157.06	197307061130CW.BBB	14.80		

Stack is pH-fld <= 4.0 (269 dataset)

serid	sarnumber	date time file	temp	flow-past	flow-pres
002268:	000157.07	197307061130CW.BBB	20.00		
002269:	000158.08	197309101200CW.BBB	20.00		
002270:	000159.09	197309101230CW.BBB	21.20		
002271:	000160.05	197307181130CW.BBB	13.30		
002272:	000160.06	197307181130CW.BBB	12.40		
002273:	000127.01	197402141030CW.BBC	6.10		
002274:	000127.02	197402141030CW.BBC	7.00		
002275:	000127.03	197402141100CW.BBC	7.20		
002276:	000127.04	197402141100CW.BBC	7.40		
002277:	000128.01	197402101030CW.BBC	15.60		
002278:	000128.02	197402101030CW.BBC	10.80		
002279:	000128.03	197402101100CW.BBC	8.30		
002280:	000129.01	197402111030CW.BBC	7.40		
002281:	000129.02	197402111030CW.BBC	11.30		
002282:	000130.01	197402121030CW.BBC	8.80		
002283:	000130.02	197402121030CW.BBC	6.50		
002284:	000130.03	197402121100CW.BBC	7.60		
002285:	000130.04	197402121100CW.BBC	12.00		
002286:	000131.01	197402121130CW.BBC	7.60		
002287:	000131.02	197402121130CW.BBC	7.10		
002288:	000131.03	197402121200CW.BBC	6.60		
002289:	000132.01	197402201030CW.BBC	9.10		
002290:	000132.02	197402201030CW.BBC	7.40		
002291:	000132.03	197402201100CW.BBC	10.10		
002292:	000132.04	197402201100CW.BBC	8.90		
002293:	000132.05	197402201130CW.BBC	8.20		
002294:	000132.06	197402201130CW.BBC	8.80		
002295:	000132.07	197402201200CW.BBC	8.30		
002296:	000133.01	197402201230CW.BBC	0.00		
002297:	000134.01	197402121200CW.BBC	8.50		
002298:	000134.02	197402121200CW.BBC	7.20		
002299:	000134.03	197402121230CW.BBC	7.10		
002300:	000134.04	197402121230CW.BBC	8.50		
002301:	000135.01	197402221030CW.BBC	7.30		
002302:	000135.02	197402221030CW.BBC	7.60		
002303:	000135.03	197402221100CW.BBC	7.40		
002304:	000135.04	197402221100CW.BBC	6.60		
002305:	000136.01	197402181030CW.BBC	7.40		
002306:	000137.01	197402181030CW.BBC	7.80		
002307:	000137.02	197402181100CW.BBC	7.10		
002308:	000137.03	197402181100CW.BBC	7.30		
002309:	000138.01	197403031030CW.BBC			
002310:	000138.02	197403031030CW.BBC			
002311:	000138.03	197403031100CW.BBC			
002312:	000139.01	197402191030CW.BBC	7.40		
002313:	000139.02	197402191030CW.BBC	7.80		
002314:	000139.03	197402191100CW.BBC	7.10		
002315:	000140.01	197402181130CW.BBC	8.20		
002316:	000140.02	197402181130CW.BBC	10.80		
002317:	000140.03	197402181200CW.BBC	8.30		
002318:	000158.08	197402201200CW.BBC	9.30		
002319:	000159.09	197402201230CW.BBC	9.60		
002320:	000160.05	197402221130CW.BBC	6.80		
002321:	000160.06	197402221130CW.BBC	5.90		
002322:	000127.01	197212181033CW.BBD	8.30		
002323:	000127.01	197301101033CW.BBD	9.50		
002324:	000127.01	197302101033CW.BBD	8.80		
002325:	000127.01	197303131033CW.BBD	8.20		
002326:	000127.01	197304121033CW.BBD	7.90		
002327:	000127.01	197305111033CW.BBD	9.50		
002328:	000127.01	197306141033CW.BBD	9.00		
002329:	000127.01	197307251033CW.BBD	9.90		
002330:	000127.01	197308241033CW.BBD	9.80		
002331:	000127.01	197309201033CW.BBD	8.90		
002332:	000127.01	197310181033CW.BBD	8.40		
002333:	000127.01	197311161033CW.BBD	4.50		
002334:	000127.01	197312121033CW.BBD	4.70		
002335:	000127.01	197401151033CW.BBD	5.50		
002336:	000127.01	197402181033CW.BBD	4.40		
002337:	000127.01	197403121033CW.BBD	7.80		
002338:	000127.01	197404221033CW.BBD	8.10		
002339:	000127.01	197405201033CW.BBD	8.40		
002340:	000127.01	197406171033CW.BBD	8.80		
002341:	000127.01	197407281033CW.BBD	8.70		
002342:	000127.01	197408131033CW.BBD	9.90		
002343:	000127.01	197409121033CW.BBD	9.10		
002344:	000127.01	197410141033CW.BBD	8.70		
002345:	000127.01	197411161033CW.BBD	9.00		
002346:	000127.01	197412171033CW.BBD	7.90		
002347:	000127.01	197501181033CW.BBD	7.60		
002348:	000127.01	197502151033CW.BBD	7.40		
002349:	000127.01	197503141033CW.BBD	6.40		
002379:	000125.01	197304111130CW.BBA	4.70		
002380:	000125.01	197307251130CW.BBA	11.60		
002381:	000125.01	197312061130CW.BBA	8.80		
002382:	000125.01	197402221130CW.BBA	6.20		
002383:	000125.01	197408131130CW.BBA	12.10		
002384:	000125.01	197410141130CW.BBA	9.30		
002385:	000125.01	197501181130CW.BBA	4.00		
002386:	000125.01	197503141130CW.BBA	6.00		

Stack is pH-fld <= 4.0 (269 dataset)

serid	sarnumber	date time file	temp	flow-past	flow-pres
002387:	000127.01	197301241030CW.BBA	7.50		
002388:	000127.01	197304111030CW.BBA	7.90		
002389:	000127.01	197312061030CW.BBA	8.50		
002390:	000127.02	197301241030CW.BBA	7.60		
002391:	000127.02	197304111030CW.BBA	7.20		
002392:	000127.02	197307251030CW.BBA	9.50		
002393:	000127.02	197312161030CW.BBA	8.50		
002394:	000127.02	197404221030CW.BBA	7.80		
002395:	000127.03	197304111030CW.BBA	4.50		
002396:	000127.03	197307251030CW.BBA	10.50		
002397:	000127.03	197312061030CW.BBA	8.00		
002398:	000127.03	197401241030CW.BBA	6.10		
002399:	000127.03	197404221030CW.BBA	8.20		
002400:	000127.03	197408131030CW.BBA	9.10		
002401:	000127.03	197410141030CW.BBA	6.80		
002402:	000127.03	197501181030CW.BBA	6.50		
002403:	000127.03	197503141030CW.BBA	6.40		
002404:	000127.04	197301241100CW.BBA	6.10		
002405:	000127.04	197304111100CW.BBA	4.60		
002406:	000127.04	197307251100CW.BBA	11.30		
002407:	000127.04	197312061100CW.BBA	8.20		
002408:	000127.04	197404221100CW.BBA	8.30		
002409:	000127.04	197408131100CW.BBA	10.60		
002410:	000127.04	197410141100CW.BBA	10.20		
002411:	000127.04	197501181100CW.BBA	5.50		
002412:	000127.04	197503141100CW.BBA	6.30		
002413:	000127.05	197304111100CW.BBA	5.50		
002414:	000127.05	197307251100CW.BBA	12.90		
002415:	000127.05	197312161100CW.BBA	7.40		
002416:	000127.05	197404221100CW.BBA	8.30		
002417:	000127.05	197410141100CW.BBA	10.90		
002418:	000127.05	197501181100CW.BBA	3.80		
002419:	000127.05	197503141100CW.BBA	6.00		
002420:	000127.06	197304111100CW.BBA	6.20		
002421:	000127.06	197307251100CW.BBA	14.90		
002422:	000127.06	197312061100CW.BBA	7.20		
002423:	000127.06	197404221100CW.BBA	8.30		
002424:	000127.08	197301241130CW.BBA	6.30		
002425:	000127.08	197304111130CW.BBA	5.20		
002426:	000127.08	197307251130CW.BBA	13.20		
002427:	000127.08	197312061130CW.BBA	7.20		
002428:	000127.08	197404221130CW.BBA	6.80		
002431:	000127.08	197501181130CW.BBA	3.80		
002432:	000127.08	197503141130CW.BBA	6.00		
002433:	000127.10	197301241130CW.BBA	4.80		
002434:	000127.10	197304111130CW.BBA	5.20		
002435:	000127.10	197307251130CW.BBA	12.80		
002436:	000127.10	197312061130CW.BBA	6.50		
002437:	000127.10	197404221130CW.BBA	6.70		
002438:	000127.10	197408131130CW.BBA	12.20		
002439:	000127.10	197410141130CW.BBA	9.30		
002440:	000127.10	197501181130CW.BBA	3.80		
002441:	000127.10	197503141130CW.BBA	5.00		
002442:	000127.11	197301241200CW.BBA	4.90		
002443:	000127.11	197304111200CW.BBA	4.50		
002444:	000127.11	197307251200CW.BBA	12.80		
002445:	000127.11	197312061200CW.BBA	6.90		
002446:	000127.11	197404221200CW.BBA	6.80		
002447:	000127.11	197408131200CW.BBA	12.20		
002448:	000127.11	197410141200CW.BBA	10.10		
002449:	000127.11	197501181200CW.BBA	3.80		
002450:	000127.11	197503141200CW.BBA	5.10		
002453:	000127.12	197307251200CW.BBA	7.70		
002458:	000127.13	197307251230CW.BBA	12.90		
002463:	000127.13	197501181230CW.BBA	3.80		
002464:	000127.13	197503141230CW.BBA	5.20		
002668:	000022.01	197303011400CW.BBK			
002669:	000022.01	197306301300CW.BBK			
002670:	000022.01	197310101300CW.BBK			
002671:	000022.01	197401191300CW.BBK			
002677:	000127.07	197301241030CW.BBL	4.60		
002678:	000127.07	197304111030CW.BBL	2.90		
002679:	000127.07	197304221030CW.BBL	3.90		
002680:	000127.07	197307251030CW.BBL	11.40		
002681:	000127.07	197312061030CW.BBL	7.20		
002682:	000127.09	197301241030CW.BBL			
002683:	000127.09	197304111030CW.BBL	4.00		
002684:	000127.09	197304221030CW.BBL	6.80		
002685:	000127.09	197307251030CW.BBL	13.20		
002686:	000127.09	197312061030CW.BBL	6.50		
002691:	000127.02	197210231100CW.BBM			
002692:	000127.03	197210231100CW.BBM			
002693:	000127.04	197210231100CW.BBM			
002694:	000127.05	197210231100CW.BBM			
002695:	000127.06	197210231100CW.BBM			
002696:	000127.07	197210231100CW.BBM			
002697:	000127.08	197210231100CW.BBM			
002698:	000127.09	197210231100CW.BBM			
002699:	000127.10	197210231100CW.BBM			
002700:	000127.11	197210231100CW.BBM			
MINiUMUM			0.00	0.0	0.0
MAXIMUM			49.00	4.0	4.0
MEAN			11.46	2.8	2.7
STD DEVIATION			7.08	1.0	1.0

Table 4.15-2

Names and descriptions of the acid data subsets related to the pH <= 4.0 query

Name	Alternative name	Description	Number missing	Creation criteria
269 dataset	Acid sampler	27 variables 269 cases (OD-420, Conductivity, pH, Acidity, Na, K, Mg, Ca, Zn, Cu, Mn, Fe, Al, Pb, Co, Ni, PO_4-P, NH_3-N, NO_3-N, SO_4-S, Cl, Si, O_2-SATN, temp, flow-past, flow-present)	1290	Query pH-fld <= 4.0
217 dataset	-	24 variables 217 cases (OD-420, Conductivity, pH, Acidity, Na, K, Mg, Ca, Zn, Cu, Mn, Fe, Al, Pb, Co, Ni, PO_4-P, NH_4-N, NO_3-N, SO_4-S, Cl, Si, temp)	19	269 subset where samples with >= 20 variables included. Variables with >= 195 of samples are included. Missing values replaced with the means for the variable.
125 dataset	Acid species	24 variables 125 cases includes 124 aggregate biologies from which [illegible] the count of the number of species is extracted and becomes the 24 th variable. The chemistry variables are the same as for the 217 dataset.	395	Bydate expansion of 269 dataset.
101 [illegible]		23 variables [illegible] cases includes 101 aggregate biologies. Variables as for 125 dataset.	12	Samples with >= [illegible] of variables present. Variables with >= [illegible] of samples.
301 dataset	Acid [illegible]	27 variables 301 cases No biologies. Variables as for 269 dataset. On expansion bysar into aggregates yields 134 aggregates. On expansion bydate yields 128 aggregates.	1521	The bysar expansion of 269 dataset - chemistries into further chemistries.

4.2 Descriptive statistics

4.21 Introduction

The distributions of pH, conductivity and acidity were presented for the total SIEUR dataset in Tables 4.12-2, 4.14-1 and 4.14-2. Although there are no a priori reasons to believe that they should follow any statistical distribution it can be seen that there is a tendency toward a median value similar to that for a normal distribution. Further it can be seen that at certain other values there are smaller secondary peaks for example see the peak at pH 2.5.

The reasons for the main and secondary peaks may lie in the nature of the data collection methodology. For example the acid stream study is likely to provide a disproportionate number of low pH sites. Alternatively it may point to some real complex distribution which may be present in stream waters.

Two problems exist which are ignored in this section. These are the bias introduced because of the selection of a portion of the total dataset based on one criterion variable. And the degree of statistical independance of the variables. Any bias is wholly justified in terms of the study aim, that is the survey of low pH data available in SIEUR. Only those variables weakly correlated with pH will be affected by this selection; strongly correlated variables will probably remain strongly correlated even in the subset. The presence or absence of strong correlation between variables in this subset may be artificial or again may be some indication of relationship, which may or may not be reflected over the total dataset. The search for transformations, to render the subset variables symmetric remain valid in the context of the search for statistically independant variables. However, note must be made of the artificiality of the method.

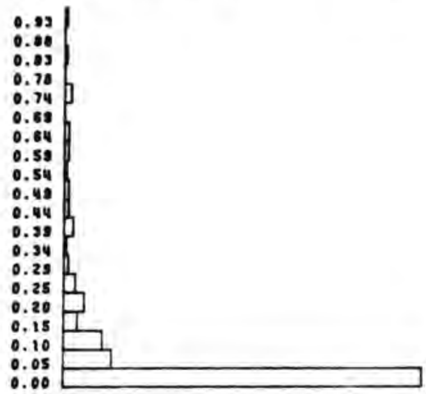
4.22 Distributions and transformations

In the following figures, the data distribution and the transformation applied are presented for all the chemical variables concerned in this study. Data at the detectable limit have been ignored (Section 2.22). The choice of the number of histogram bins was based arbitarily on the square root of the number of non-missing pH data points. The histograms display method follows that suggested by Fox (1976).

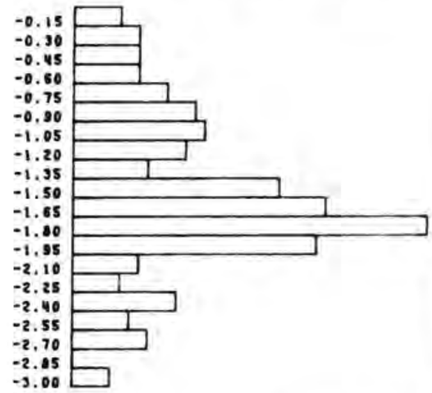
Table 4.22-1 (pp 112 - 120)

Distribution histograms for the chemical and physical parameters and their $\log_{10}$ transformations associated with samples where the pH <= 4.0
(Units as in Table 4.13-1 p 93)
All bins are scaled relative to the longest which is set at 100%

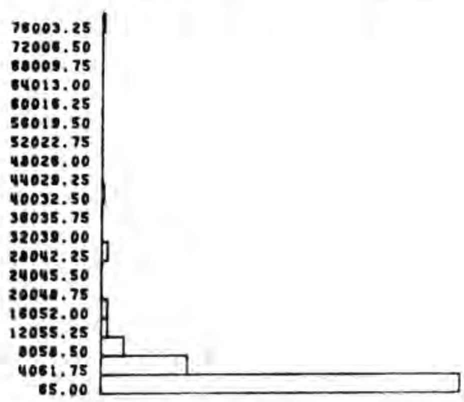
O.D._420 237



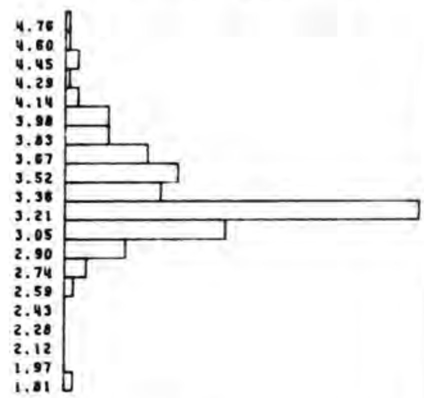
COMMON LOG 237



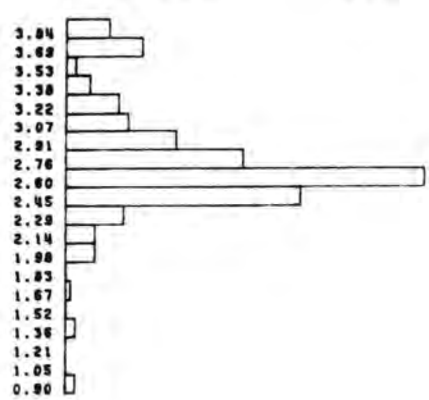
COND 238



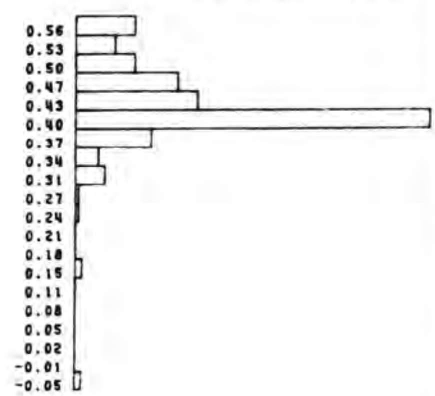
COMMON LOG 238



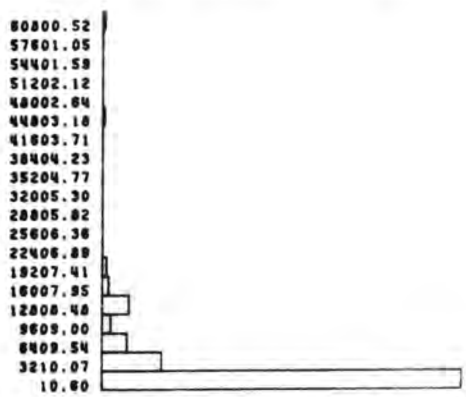
PH_FLD 269



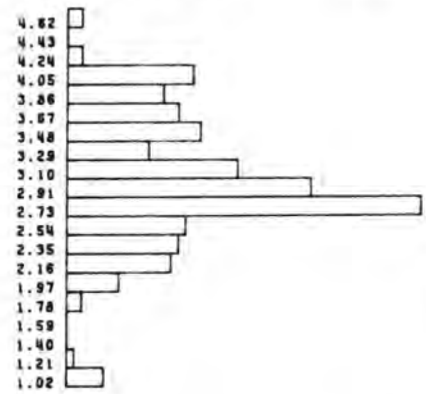
COMMON LOG 269



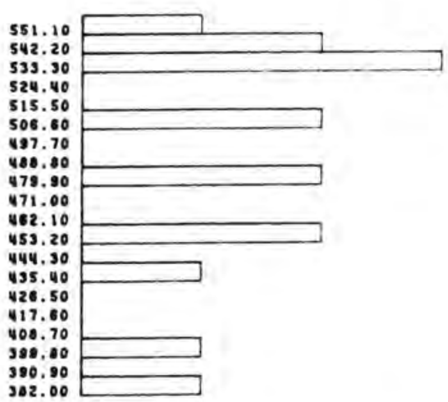
ACIDITY 242



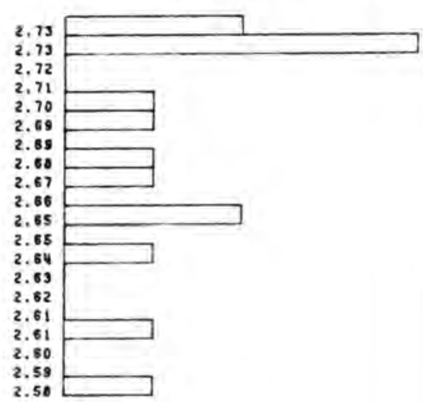
COMMON LOG 242



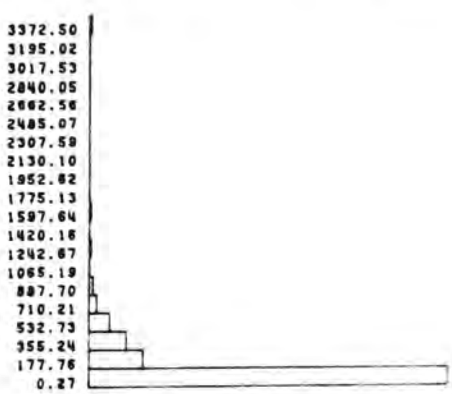
EH_FLD 15



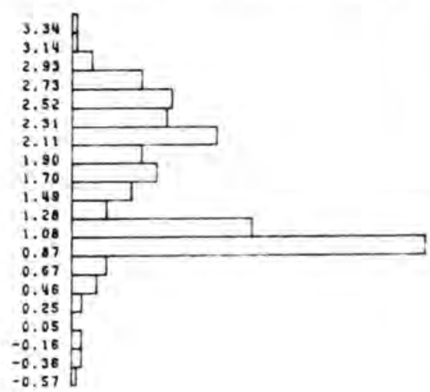
COMMON LOG 15

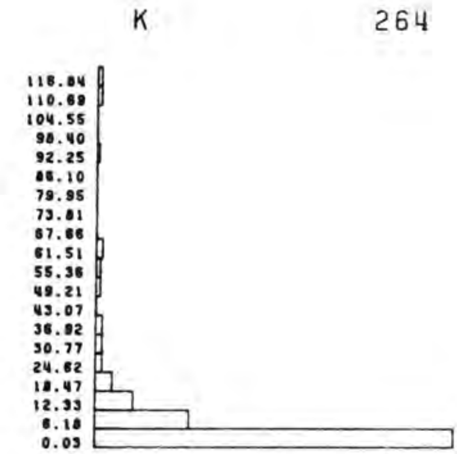


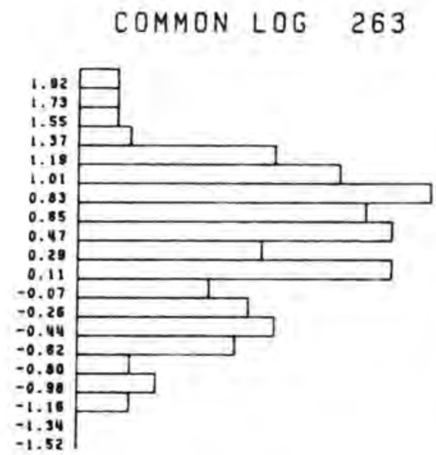
NA 264

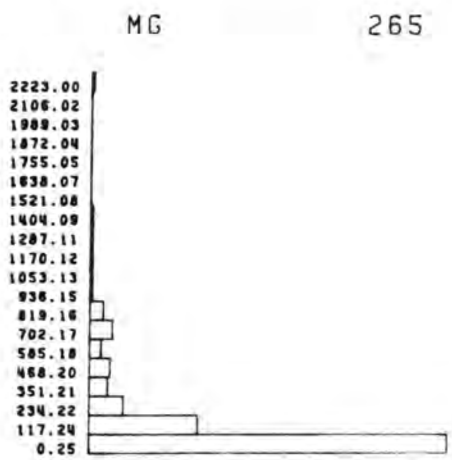


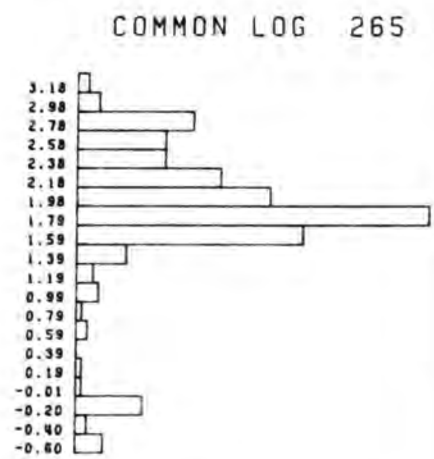
COMMON LOG 264

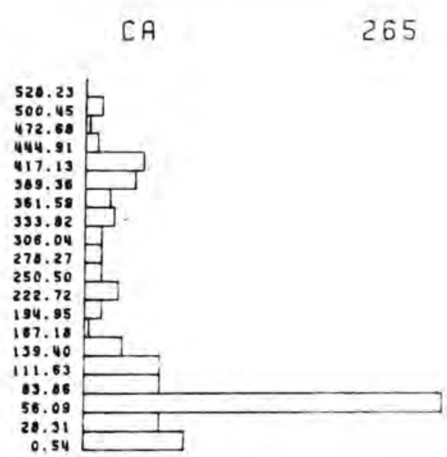


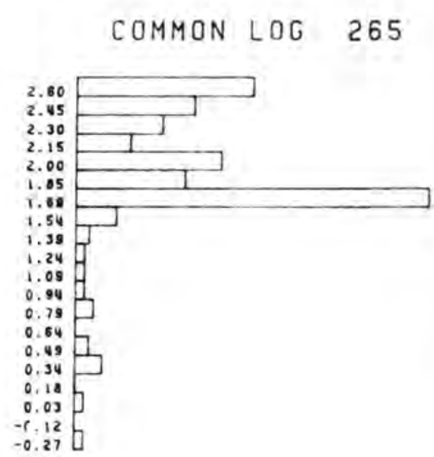

K 264


COMMON LOG 263

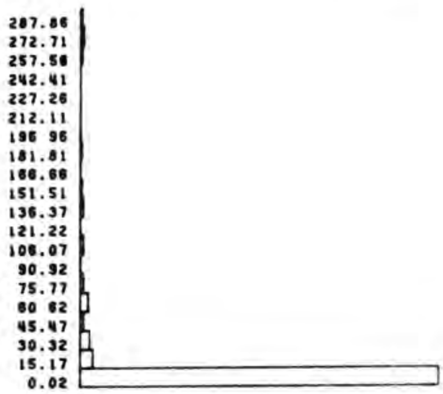

MG 265


COMMON LOG 265

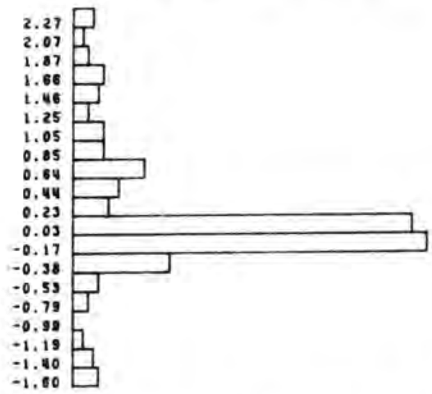

CA 265


COMMON LOG 265

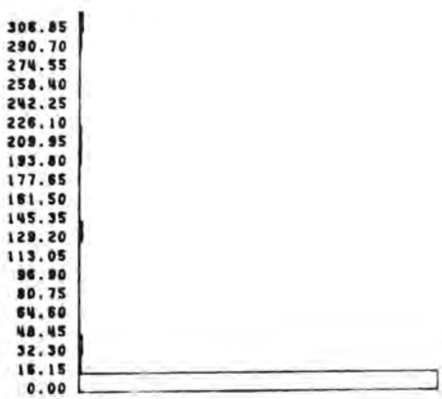
ZN 265



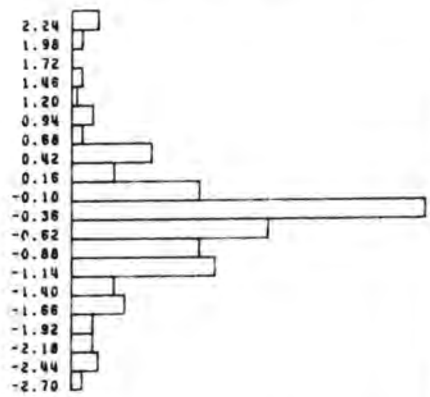
COMMON LOG 260



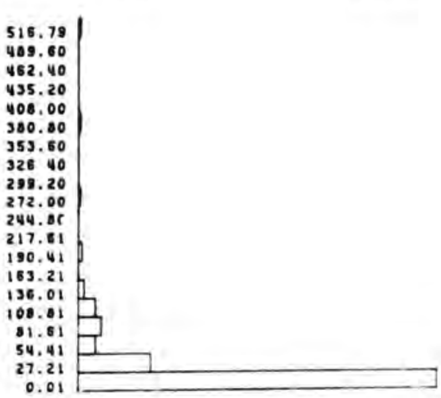
CU 252



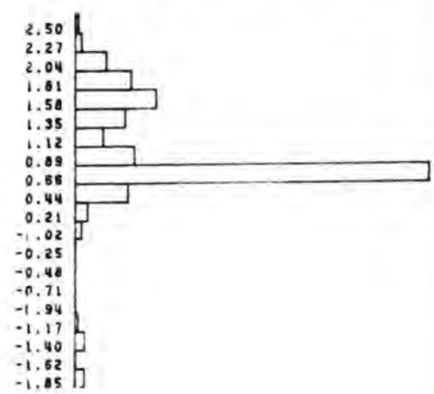
COMMON LOG 251

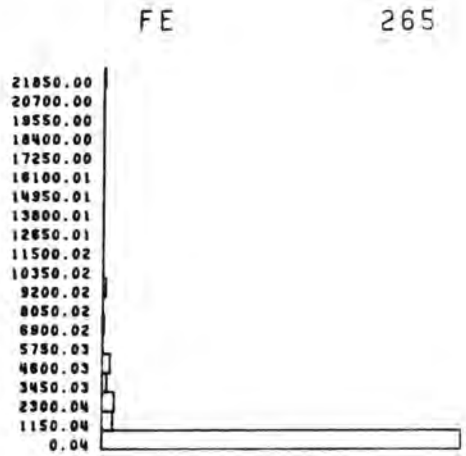


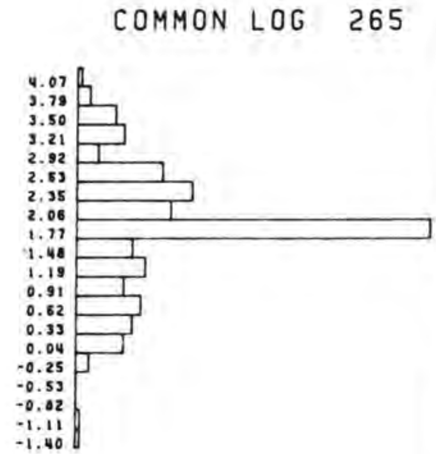
MN 265

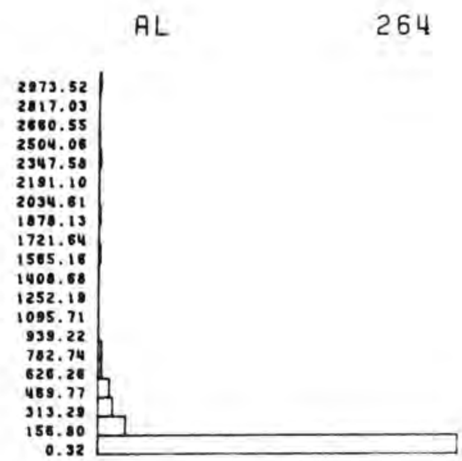


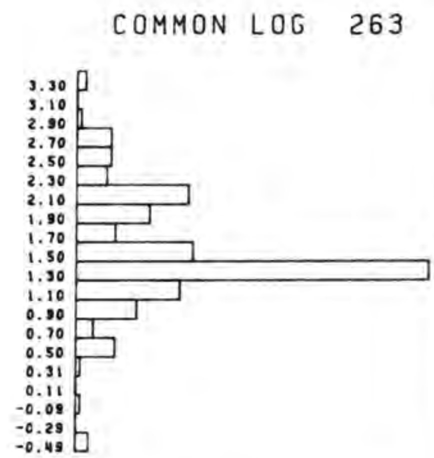
COMMON LOG 265

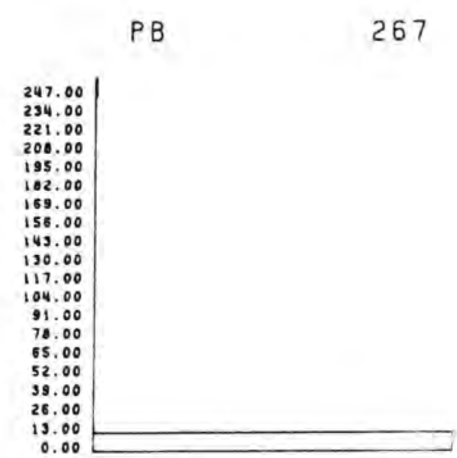


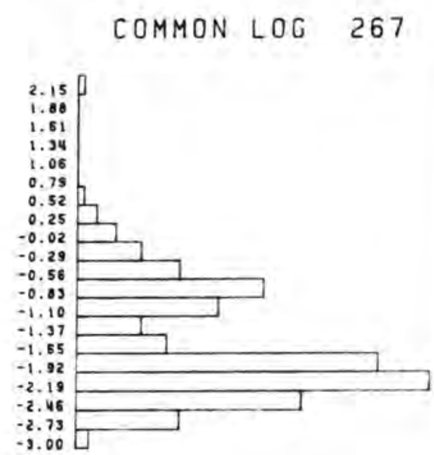

FE 265
21850.00
20700.00
19550.00
18400.00
17250.00
16100.01
14950.01
13800.01
12650.01
11500.02
10350.02
9200.02
8050.02
6900.02
5750.03
4600.03
3450.03
2300.04
1150.04
0.04


COMMON LOG 265
4.07
3.79
3.50
3.21
2.92
2.63
2.35
2.06
1.77
1.48
1.19
0.91
0.62
0.33
0.04
-0.25
-0.53
-0.82
-1.11
-1.40

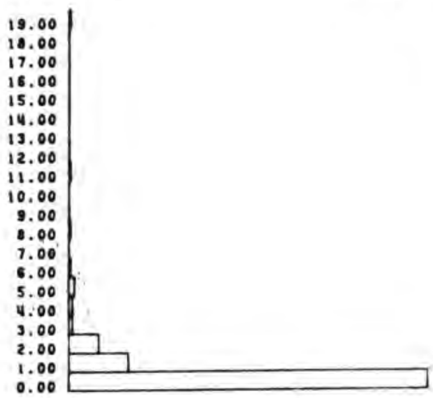

AL 264
2973.52
2817.03
2660.55
2504.06
2347.58
2191.10
2034.61
1878.13
1721.64
1565.16
1408.68
1252.19
1095.71
939.22
782.74
626.26
469.77
313.29
156.80
0.32


COMMON LOG 263
3.30
3.10
2.90
2.70
2.50
2.30
2.10
1.90
1.70
1.50
1.30
1.10
0.90
0.70
0.50
0.31
0.11
-0.09
-0.29
-0.49

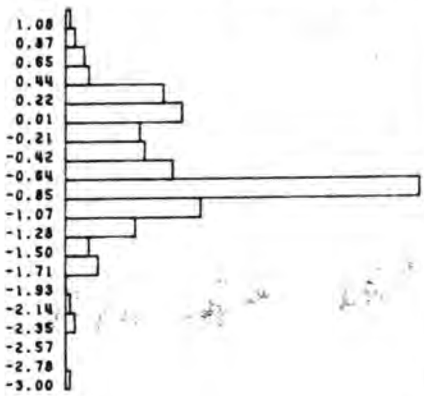

PB 267
247.00
234.00
221.00
208.00
195.00
182.00
169.00
156.00
143.00
130.00
117.00
104.00
91.00
78.00
65.00
52.00
39.00
26.00
13.00
0.00


COMMON LOG 267
2.15
1.88
1.61
1.34
1.06
0.79
0.52
0.25
-0.02
-0.29
-0.56
-0.83
-1.10
-1.37
-1.65
-1.92
-2.19
-2.46
-2.73
-3.00

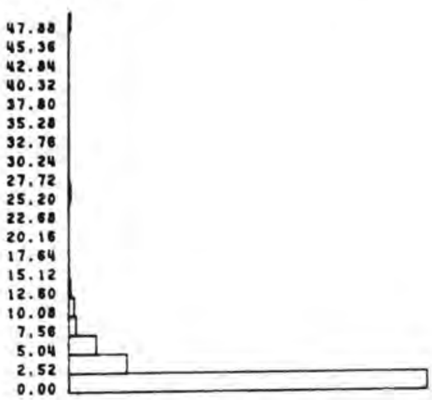
CO 253



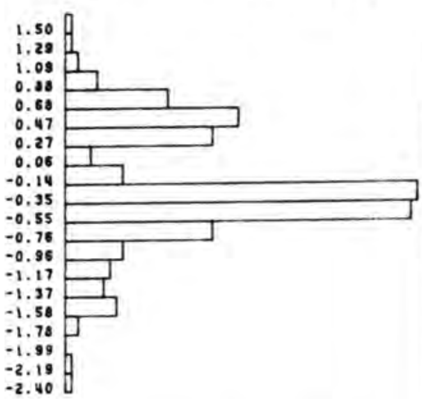
COMMON LOG 250



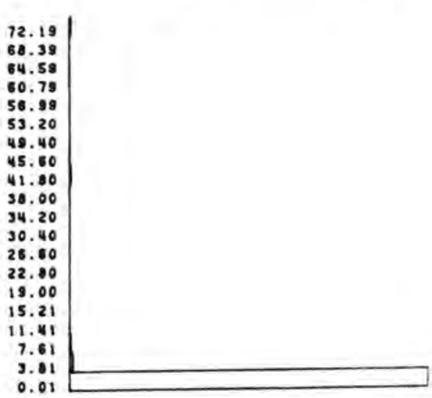
NI 254



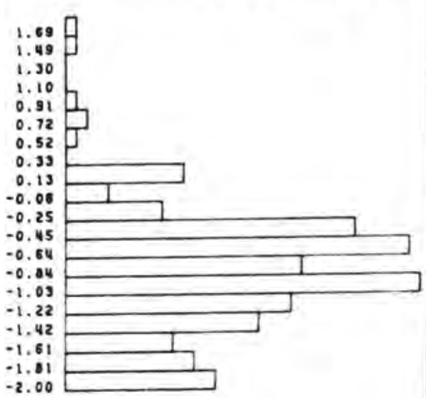
COMMON LOG 254

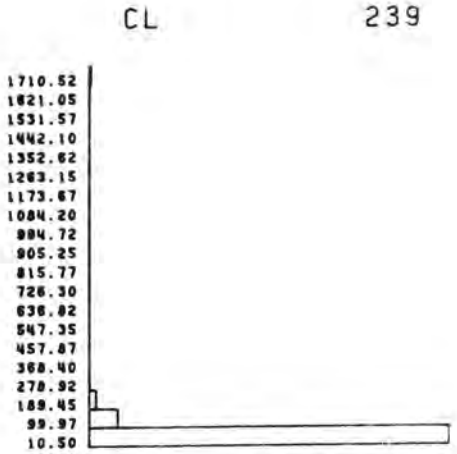


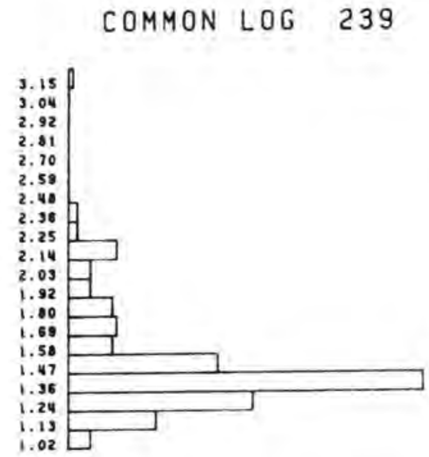
PO4_P 220

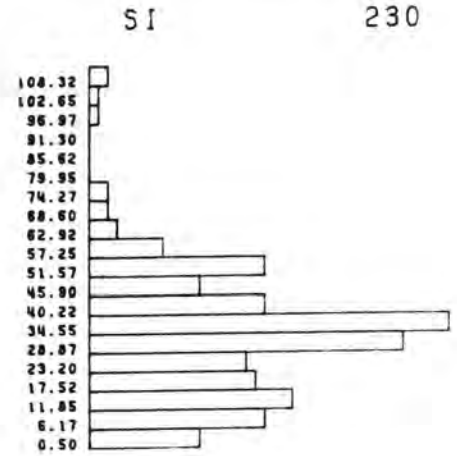


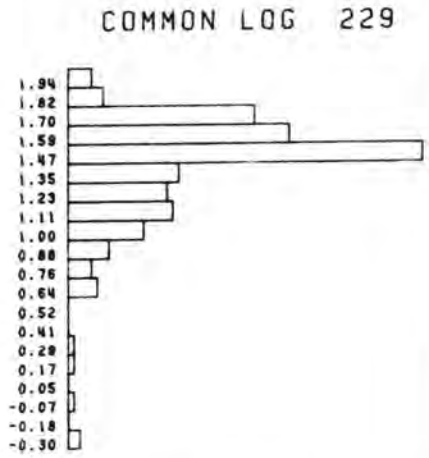
COMMON LOG 219

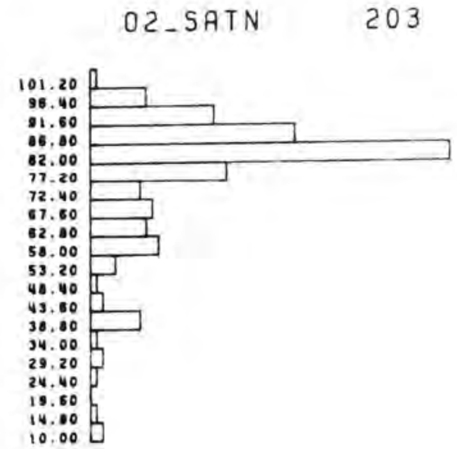


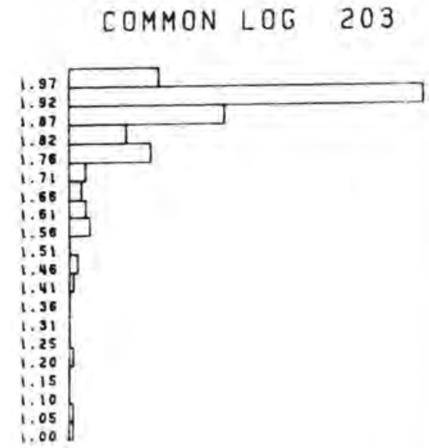

CL 239
1710.52
1621.05
1531.57
1442.10
1352.62
1263.15
1173.67
1084.20
994.72
905.25
815.77
726.30
636.82
547.35
457.87
368.40
278.92
189.45
99.97
10.50


COMMON LOG 239
3.15
3.04
2.92
2.81
2.70
2.59
2.48
2.36
2.25
2.14
2.03
1.92
1.80
1.69
1.58
1.47
1.36
1.24
1.13
1.02

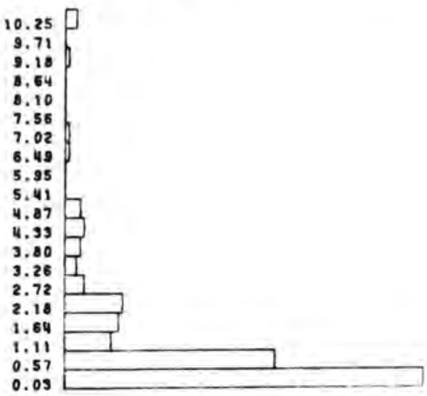

SI 230
108.32
102.65
96.97
91.30
85.62
79.95
74.27
68.60
62.92
57.25
51.57
45.90
40.22
34.55
28.87
23.20
17.52
11.85
6.17
0.50


COMMON LOG 229
1.94
1.82
1.70
1.59
1.47
1.35
1.23
1.11
1.00
0.88
0.76
0.64
0.52
0.41
0.29
0.17
0.05
-0.07
-0.18
-0.30

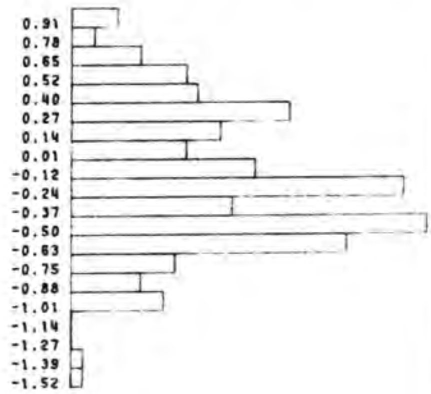

O2_SATN 203
101.20
96.40
91.60
86.80
82.00
77.20
72.40
67.60
62.80
58.00
53.20
48.40
43.60
38.80
34.00
29.20
24.40
19.60
14.80
10.00


COMMON LOG 203
1.97
1.92
1.87
1.82
1.76
1.71
1.66
1.61
1.56
1.51
1.46
1.41
1.36
1.31
1.25
1.20
1.15
1.10
1.05
1.00

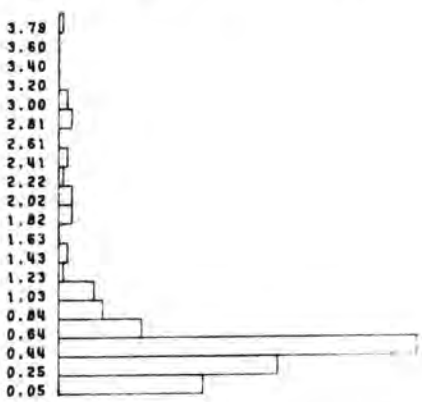
NH4_N 217



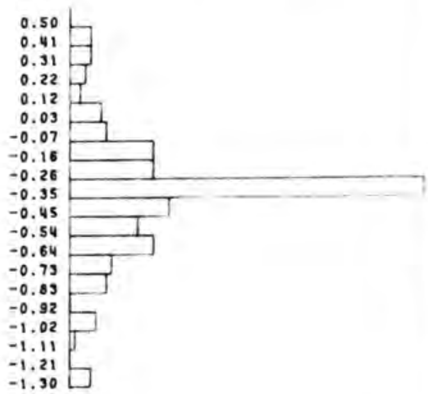
COMMON LOG 214



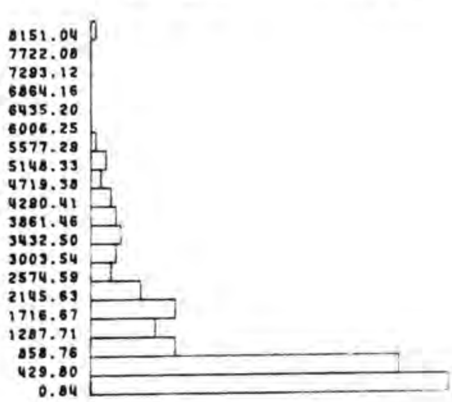
NO3_N 220



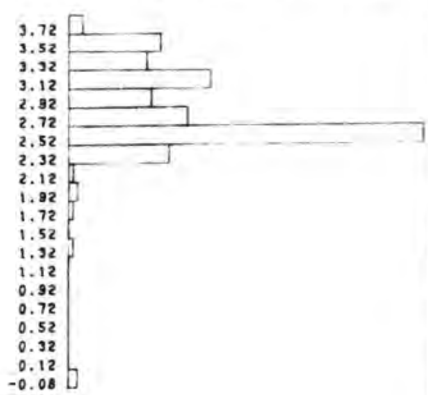
COMMON LOG 219

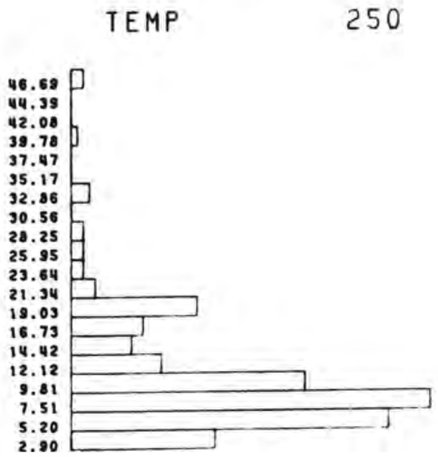


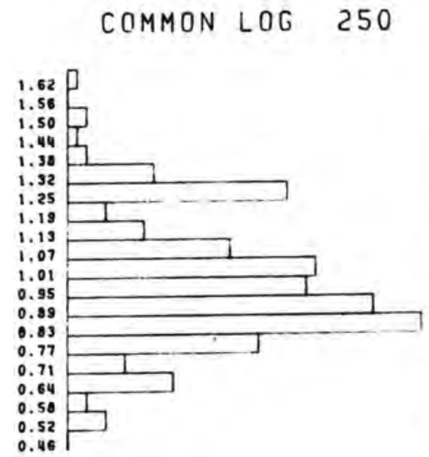
SO4_S 222



COMMON LOG 222




TEMP 250
46.69
44.39
42.08
39.78
37.47
35.17
32.86
30.56
28.25
25.95
23.64
21.34
19.03
16.73
14.42
12.12
9.81
7.51
5.20
2.90


COMMON LOG 250
1.62
1.56
1.50
1.44
1.38
1.32
1.25
1.19
1.13
1.07
1.01
0.95
0.89
0.83
0.77
0.71
0.64
0.58
0.52
0.46

It is clear from Tables 4.22-1 and 4.23-1 that for most variables a logarithmic transformation ($\log_{10}$ X) rendered the data symmetric about some single modal value. Furthermore logarithmic transformations seem most appropriate even when departures from symmetry are detected for the following reasons:

1. The data may span several orders of magnitude and include many extreme values, which may result in only visual apparent departures from symmetry.

2. One of the major variables under consideration (pH) is already logarithmically scaled.

3. Logarithmic transformations are readily understood and applied.

4. With reference to the 269 acid sample chemistries under consideration the departures from symmetry of the logarithmically transformed variables can be explained in terms consistent with the nature of the extreme environment.

4.23 Skewness and kurtosis of the distributions

The variable distributions (Table 4.22-1) can show three possible departures from symmetry, modality (the tendancy to have many peaks in the distribution), skewness (the tendancy for the modal peak to be displaced to one end of the distribution) and kurtosis (the peakedness of the distribution). Modality cannot be easily removed by transformation; variables which display bimodality are Na, Mg, Ca, Mn and Pb. Skewness is displayed by Mg, Ca, Pb and PO_4-P. It is probable that the cause of this skewness is the occasional extreme value, a fact borne out by the numerically low skewness figures to be found for these variable in Table 4.23-1. Kurtosis is displayed to a certain extent by all the variables, but is more clearly discernable in Zn, Cu, Fe and Al and to a lesser extent by Ni.

Presented in Table 4.23-1 are various shape statistics calculated as though the variables and their transformations were normally distributed. If the curves were true approximations of the normal distribution, then it would be expected that the skewness would be zero (positive values indicate clustering to the right of the mode) kurtosis too would be zero (positive values indicating more pointed distributions). The statistics in Table 4.23-1 are presented merely as quantitative confirmation that the transformations yield more symmetric, normal looking distributions for the variables. Where

Table 4.23-1

Symmetry statistics of the chemical and physical parameters for sample chemistries where the pH <= 4.0

(second line is for $\log_{10}$ transformed data)

Variable	n	minimum	maximum	mean	s.d.	skewness	kurtosis
O.D.-420	237	.01000	.98800	.09777	.16973	2.823	8.153
	237	-3.0000	-.00524	-1.4820	.65429	.116	-.291
conductivity	238	65.000	80000.	4079.2	7018.3	6.827	61.433
	238	1.8129	4.9031	3.4000	.38266	.398	2.851
pH-fld	269	.90000	4.0000	2.8207	.46512	.105	2.355
acidity	242	10.600	64000.	3354.6	6412.4	5.270	39.646
	242	1.0253	4.8062	3.0674	.66461	-.261	.687
Eh-fld	15	382.00	560.00	493.13	55.600	-.586	-.786
	15	2.5821	2.7482	2.6902	.05109	-.742	-.506
Na	264	.27000	3550.0	156.47	308.15	6.080	56.662
	264	-.56864	3.5502	1.6333	.74298	.089	-.632
K	264	.03000	123.00	9.2739	18.442	4.113	19.017
	264	-1.5229	2.0899	.43114	.74534	-.212	-.426
Mg	265	.25000	2340.0	200.43	290.64	3.072	13.248
	265	-.60206	3.3692	1.8904	.76382	-1.371	2.342
Ca	265	.54000	556.00	164.42	149.73	.957	-.516
	265	-.26761	2.7451	1.9788	.55260	-1.297	2.578
Zn	265	.02500	303.00	12.804	41.828	4.992	26.848
	265	-1.6021	2.4814	.22804	.75597	.772	1.357
Cu	252	.00200	323.00	7.7594	40.310	6.447	42.334
	252	-2.6990	2.5092	-.38100	.83874	.542	2.384
Mn	265	.01400	544.00	29.207	55.671	4.944	34.417
	265	-1.8539	2.7356	.98622	.74520	-1.090	3.112
Fe	265	.04000	23000.	583.19	1898.2	7.655	76.672
	265	-1.3979	4.3617	1.9135	.92174	-.285	.433
Al	264	.32000	3130.0	121.17	314.65	6.380	48.412
	264	-.49485	3.4955	1.5723	.64792	-.052	1.137
Pb	267	.00100	260.00	1.1393	15.908	16.229	261.589
	267	-3.0000	2.4150	-1.5074	.81187	.909	1.029
Co	253	.00100	20.000	.82574	1.7851	6.557	57.646
	253	-3.0000	1.3010	-.50281	.59501	-.033	1.125
Ni	254	.00400	50.400	1.8841	4.1741	7.451	75.489
	254	-2.3979	1.7024	-.19304	.64891	-.040	.252
PO_4-P	220	.01000	76.000	.93464	5.9430	10.978	126.417
	220	-2.0000	1.8808	-.81445	.64537	.572	1.536
NH_4-N	217	.03000	10.800	1.3477	1.7841	2.927	10.380
	217	-1.5229	1.0334	-.13421	.47473	.204	-.301
NO_3-N	220	.05000	4.0000	.62850	.60511	2.885	9.303
	220	-1.3010	.60206	-.33475	.33944	-.168	1.191
SO_4-S	222	.84000	8580.0	1256.4	1366.0	1.982	4.420
	222	-.07572	3.9335	2.8762	.49255	-1.551	9.163
Cl	239	10.500	1800.0	49.304	121.52	12.695	178.939
	239	1.0212	3.2553	1.5102	.30310	1.835	4.822
Si	230	.50000	114.00	33.605	19.383	.942	2.187
	230	-.30103	2.0569	1.4234	.36496	-1.924	5.456
O_2-SATN	207	0.	106.00	76.536	20.071	-1.897	3.804
	203	1.0000	2.0253	1.8756	.14008	-3.243	13.693
temperature	251	0.	49.000	11.463	7.0810	2.188	6.863
	250	.46240	1.6902	.99960	.22244	.489	.106

this is not the case, the untransformed variable was used for example temperature.

4.3 Intervariable statistics & principal component analysis

4.31 Introduction

This section presents the results of the intervariable analysis of the various acid chemistry datasets. Also presented are the principal components extracted from these datasets.

The results of the two principal component analyses on the 23 variables (or transformed variables) and 217 chemistries of the 269 data subset and the 24 variables (or transformed variables) and 101 chemistries of the 125 data subset are presented. In these subsets, data at the detectable limit were treated using the methods described in 2.222. Missing data for the principal components analysis were replaced by the variable means as outlined in Table 4.15-2. The two sets of component results are presented together because the aim is to compare the pattern of variation and to generalise the results for the chemical data as a whole.

4.32 Correlation analysis

Table 4.32-1 presents the Pearson, Spearman and Kendall correlation coefficients for all the transformed variables of the acid samples 269 dataset with the missing data ignored. Care must be taken in interpreting these correlations for reasons explained in 4.21.

Table 4.32-1 (pp 125 - 127)

Pearson (P), Spearmans (S) and Kendals (K) correlation coefficients among the chemical and physical parameters of the 269 acid samples dataset

		O.D.-420	cond	pH-fld	acidity	Eh	NSPEC	Na	K	Mg
O.D.-420	(P)	1.0000	0.3975	-0.3696	0.6429	0.5754	-0.3138	0.2183	-0.2234	0.3034
	(S)	1.0000	0.4317	-0.3683	0.6509	0.4687	-0.3540	0.1651	-0.1963	0.2900
	(K)	1.0000	0.3198	-0.2643	0.4939	0.3761	-0.2647	0.1069	-0.1518	0.2030
		110	110	110	102	13	110	110	109	110
cond	(P)	0.3975	1.0000	-0.5580	0.5710	0.3451	-0.0186	0.4201	0.0454	0.4202
	(S)	0.3198	1.0000	-0.3381	0.6241	0.1788	-0.0809	0.5298	0.1145	0.6022
	(K)	0.4317	1.0000	-0.2424	0.4902	0.0811	-0.0614	0.3943	0.0841	0.4761
		110	110	110	102	13	110	110	109	110
pH-fld	(P)	-0.3696	-0.5580	1.0000	-0.4922	-0.7067	0.2250	-0.0449	0.2944	-0.0564
	(S)	-0.2643	-0.2424	1.0000	-0.3690	-0.5368	0.2802	0.0259	0.4018	-0.0597
	(K)	-0.3683	-0.3381	1.0000	-0.2702	-0.4027	0.2119	0.0346	0.2809	-0.0382
		110	110	126	113	14	126	122	121	122
acidity	(P)	0.6429	0.5710	-0.4922	1.0000	0.3697	-0.4536	0.5063	-0.1973	0.7284
	(S)	0.4939	0.4902	-0.2702	1.0000	0.1604	-0.4185	0.5619	-0.1035	0.7423
	(K)	0.6509	0.6241	-0.3690	1.0000	0.1150	-0.3107	0.3989	-0.0693	0.5626
		102	102	113	113	14	113	112	111	112
Eh	(P)	0.5754	0.3451	-0.7067	0.3697	1.0000	0.1188	-0.4665	-0.6017	-0.0355
	(S)	0.3761	0.0811	-0.4027	0.1150	1.0000	0.1391	-0.4770	-0.6444	-0.1201
	(K)	0.4687	0.1788	-0.5368	0.1804	1.0000	0.0728	-0.3492	-0.4835	-0.0946
		13	13	14	14	14	14	13	13	13
NSPEC	(P)	-0.3138	-0.0186	0.2250	-0.4536	0.1188	1.0000	-0.2109	0.1189	-0.4616
	(S)	-0.2647	-0.0614	0.2119	-0.3107	0.0728	1.0000	-0.1793	0.0960	-0.3309
	(K)	-0.3540	-0.0809	0.2802	-0.4185	0.1391	1.0000	-0.1259	0.0696	-0.2391
		110	110	126	113	14	126	122	121	122
Na	(P)	0.2183	0.4201	-0.0449	0.5063	-0.4665	-0.2109	1.0000	0.3116	0.6342
	(S)	0.1069	0.3943	0.0346	0.3989	-0.3492	-0.1259	1.0000	0.3360	0.7666
	(K)	0.1651	0.5298	0.0259	0.5619	-0.4770	-0.1793	1.0000	0.2234	0.5589
		110	110	122	112	13	122	122	121	122
K	(P)	-0.2234	0.0454	0.2944	-0.1973	-0.6017	0.1189	0.3116	1.0000	-0.0472
	(S)	-0.1518	0.0841	0.2809	-0.0693	-0.4835	0.0696	0.2234	1.0000	0.1923
	(K)	-0.1963	0.1145	0.4018	-0.1035	-0.6444	0.0960	0.3360	1.0000	0.1323
		109	109	121	111	13	121	121	121	121
Mg	(P)	0.3034	0.4202	-0.0564	0.7284	-0.0355	-0.4616	0.6342	-0.0472	1.0000
	(S)	0.2030	0.4761	-0.0382	0.5626	-0.0946	-0.2391	0.5589	0.1323	1.0000
	(K)	0.2900	0.6022	-0.0597	0.7423	-0.1201	-0.3309	0.7666	0.1923	1.0000
		110	110	122	112	13	122	122	121	122
Ca	(P)	0.0794	0.1754	0.1111	0.4095	-0.2101	-0.3710	0.6024	0.1555	0.7238
	(S)	0.0683	0.2739	0.0665	0.2522	-0.0676	-0.2021	0.4207	0.1951	0.4740
	(K)	0.0974	0.3742	0.0877	0.3694	-0.1006	-0.2742	0.5872	0.2758	0.6515
		110	110	122	112	13	122	122	121	122
Zn	(P)	0.0963	-0.0445	0.0171	0.3075	0.1658	-0.1793	-0.1418	-0.3425	0.2885
	(S)	0.1229	0.0279	-0.0378	0.1941	0.1217	-0.0840	-0.0316	-0.2027	0.0319
	(K)	0.1707	0.0305	-0.0632	0.2771	0.2025	-0.1113	-0.0686	-0.2940	0.0434
		110	110	122	112	13	122	122	121	122
Cu	(P)	0.2720	0.0443	-0.2435	0.3231	0.2036	-0.1724	-0.3161	-0.5852	0.2049
	(S)	0.2349	-0.0054	-0.2651	0.2303	0.2569	-0.1393	-0.2204	-0.4162	-0.0245
	(K)	0.3278	-0.0048	-0.3477	0.3365	0.3464	-0.1908	-0.2691	-0.5925	-0.0387
		110	110	121	111	13	121	121	120	121
Mn	(P)	0.1744	0.1935	0.0751	0.6709	-0.1007	-0.3947	0.5717	-0.0883	0.8720
	(S)	0.1457	0.4120	-0.0023	0.4851	0.0000	-0.1958	0.5477	0.1108	0.5939
	(K)	0.2247	0.5266	-0.0167	0.6605	0.0251	-0.2718	0.7372	0.1521	0.7678
		110	110	122	112	13	122	122	121	122
Fe	(P)	0.5403	0.3365	-0.2707	0.7854	0.3151	-0.4914	0.3815	-0.2338	0.6938
	(S)	0.4685	0.4136	-0.2252	0.6035	0.1074	-0.3884	0.2402	-0.1506	0.4215
	(K)	0.5941	0.5159	-0.3044	0.7738	0.1841	-0.5211	0.3530	-0.1906	0.5795
		110	110	122	112	13	122	122	121	122
Al	(P)	0.4220	0.2327	-0.1335	0.7420	0.2718	-0.2893	0.2737	-0.3278	0.5620
	(S)	0.3766	0.3837	-0.1818	0.5438	0.0811	-0.2374	0.2305	-0.2015	0.4154
	(K)	0.5234	0.4872	-0.2493	0.7425	0.1117	-0.3266	0.3148	-0.2889	0.5843
		110	110	122	112	13	122	122	121	122
Pb	(P)	0.2203	-0.0295	0.0728	0.3055	0.2122	0.1519	-0.0189	-0.0972	0.1888
	(S)	0.1672	-0.0078	0.0664	0.2654	0.1272	-0.0358	-0.0225	-0.0508	0.0498
	(K)	0.2349	-0.0342	0.0097	0.3593	0.2297	-0.0499	-0.0342	-0.0746	0.0799
		110	110	125	113	14	125	122	121	122
Co	(P)	0.5552	0.5562	-0.2268	0.6976	0.0593	-0.3345	0.5441	-0.1070	0.7001
	(S)	0.3582	0.4817	-0.1421	0.5252	0.0000	-0.2509	0.4113	-0.0321	0.5036
	(K)	0.4968	0.6065	-0.2073	0.7163	-0.0145	-0.3334	0.5628	-0.0409	0.6814
		105	105	115	107	12	115	115	114	115
Ni	(P)	0.5093	0.4442	-0.1695	0.7868	0.1769	-0.4679	0.5680	-0.1100	0.7581
	(S)	0.3752	0.3378	-0.0600	0.5855	0.2163	-0.3325	0.3982	0.0224	0.4971
	(K)	0.5163	0.4586	-0.0986	0.7679	0.1676	-0.4503	0.5861	0.0170	0.6827
		108	108	119	111	13	119	119	118	119
PO_4-P	(P)	0.3915	0.1028	-0.3420	0.4067	0.1547	-0.2840	0.1248	-0.3307	0.1605
	(S)	0.2650	0.0374	-0.1703	0.2229	0.5069	-0.2217	0.0365	-0.2060	0.0678
	(K)	0.3842	0.0432	-0.2464	0.3258	0.5833	-0.3038	0.0476	-0.3133	0.0795
		98	98	98	98	10	98	98	97	98
NH_4-N	(P)	0.1755	0.3371	0.0367	0.2834	-0.5084	-0.0652	0.6351	0.3060	0.4636
	(S)	-0.0062	0.2629	0.0711	0.1630	-0.5744	-0.0114	0.4759	0.2170	0.3128
	(K)	0.0193	0.3615	0.0859	0.2373	-0.6380	-0.0164	0.6440	0.3185	0.4470
		96	96	96	96	8	96	96	95	96
NO_3-N	(P)	0.0145	-0.0292	0.0385	-0.0877	-0.2519	-0.0020	0.0277	-0.0202	-0.0011
	(S)	0.0174	0.0420	-0.0054	-0.0498	-0.2060	0.0653	0.0208	-0.0377	0.0257
	(K)	0.0296	0.0604	-0.0143	-0.0657	-0.3575	0.0914	0.0153	-0.0563	0.0250
		98	98	98	98	9	98	98	97	98
SO_4-S	(P)	0.5592	0.6060	-0.3111	0.6610	0.3798	-0.2223	0.5093	0.0195	0.5382
	(S)	0.3525	0.4924	-0.1975	0.4257	0.3223	-0.1702	0.3542	-0.0151	0.3060
	(K)	0.4878	0.6357	-0.2865	0.5909	0.4268	-0.2333	0.4971	-0.0276	0.4403
		102	102	102	102	13	102	102	101	102
Cl	(P)	0.1231	0.2134	-0.0090	0.1366	0.0167	0.0551	0.3080	0.0520	-0.0057
	(S)	0.0560	0.0843	0.0648	0.0472	0.0000	-0.0249	0.2108	0.1076	0.0218
	(K)	0.0767	0.1277	0.0823	0.0776	-0.0223	-0.0327	0.3240	0.1542	0.0486
		102	102	108	107	13	108	108	107	108
Si	(P)	0.2437	0.2311	-0.1501	0.3811	0.1989	-0.0965	0.3370	0.0669	0.3782
	(S)	0.1311	0.1013	-0.0074	0.1749	0.01[illegible]	-0.0299	0.1749	0.0175	0.1348
	(K)	0.1917	0.1310	-0.0163	0.2[illegible]	0.05[illegible]	-0.0421	0.2695	0.0380	0.1874
		102	102	102	102	13	102	102	101	102
O_2-SATN	(P)	-0.0467	-0.2099	0.0162	-0.0977	0.0000	0.0417	-0.0670	-0.1178	0.0031
	(S)	-0.1068	-0.2[illegible]	-0.0148	-0.15[illegible]	0.0000	0.0472	-0.1478	-0.1210	-0.1239
	(K)	-0.1588	-0.3[illegible]	-0.0372	-0.2[illegible]	0.0000	0.0684	-0.2[illegible]	-0.1[illegible]	-0.1775
		89	89	89	89	1	89	89	88	89
temp	(P)	-0.0056	0.3610	-0.1894	-0.3576	[illegible]	0.41[illegible]	[illegible]	0.1353	-0.3309
	(S)	-0.0105	0.24[illegible]	-0.1[illegible]	-0.1107	[illegible]	0.1909	[illegible]	0.3673	[illegible]
	(K)	-0.0126	0.3574	-0.2271	-0.2[illegible]	-0.2765	0.2513	[illegible]	[illegible]	[illegible]
		109	109	12[illegible]	[illegible]	14	125	[illegible]	120	121

		Ca	Zn	Cu	Mn	Fe	Al	Pb	Co	Ni
O.D.-420	(P)	0.0794	0.0963	0.2720	0.1744	0.5403	0.4220	0.2203	0.5552	0.5093
	(S)	0.0974	0.1707	0.3278	0.2247	0.5941	0.5234	0.2349	0.4968	0.5163
	(K)	0.0683	0.1229	0.2349	0.1457	0.4685	0.3766	0.1672	0.3582	0.3752
		110	110	110	110	110	110	110	105	108
cond	(P)	0.1754	-0.0445	0.0443	0.1935	0.3365	0.2327	-0.0295	0.5562	0.4442
	(S)	0.3742	0.0305	-0.0048	0.5266	0.5159	0.4872	-0.0342	0.6065	0.4586
	(K)	0.2739	0.0279	-0.0054	0.4120	0.4136	0.3837	-0.0078	0.4817	0.3378
		110	110	110	110	110	110	110	105	108
pH-fld	(P)	0.1111	0.0171	-0.2435	0.0751	-0.2707	-0.1335	0.0728	-0.2268	-0.1695
	(S)	0.0877	-0.0632	-0.3477	-0.0167	-0.3044	-0.2493	0.0697	-0.2073	-0.0986
	(K)	0.0665	-0.0378	-0.2651	-0.0023	-0.2252	-0.1818	0.0664	-0.1421	-0.0600
		122	122	121	122	122	122	125	115	119
acidity	(P)	0.4095	0.3075	0.3231	0.6709	0.7854	0.7420	0.3055	0.6976	0.7868
	(S)	0.3694	0.2771	0.3365	0.6605	0.7738	0.7425	0.3593	0.7163	0.7679
	(K)	0.2522	0.1941	0.2303	0.4851	0.6035	0.5438	0.2654	0.5252	0.5855
		112	112	111	112	112	112	113	107	111
Eh	(P)	-0.2101	0.1658	0.2036	-0.1007	0.3151	0.2718	0.2122	0.0593	0.1769
	(S)	-0.1006	0.2025	0.3464	0.0251	0.1841	0.1117	0.2297	-0.0145	0.1676
	(K)	-0.0676	0.1217	0.2569	0.0000	0.1074	0.0811	0.1272	0.0000	0.2163
		13	13	13	13	13	13	14	12	13
NSPEC	(P)	-0.3710	-0.1793	-0.1724	-0.3947	-0.4914	-0.2893	0.1519	-0.3345	-0.4679
	(S)	-0.2742	-0.1113	-0.1908	-0.2718	-0.5211	-0.3266	-0.0499	-0.3334	-0.4503
	(K)	-0.2021	-0.0840	-0.1393	-0.1958	-0.3884	-0.2374	-0.0358	-0.2509	-0.3325
		122	122	121	122	122	122	125	115	119
Na	(P)	0.6024	-0.1418	-0.3161	0.5717	0.3815	0.2737	-0.0189	0.5441	0.5680
	(S)	0.5872	-0.0686	-0.2891	0.7372	0.3530	0.3148	-0.0342	0.5628	0.5861
	(K)	0.4207	-0.0316	-0.2204	0.5477	0.2402	0.2305	-0.0225	0.4113	0.3982
		122	122	121	122	122	122	122	115	119
K	(P)	0.1555	-0.3425	-0.5852	-0.0883	-0.2338	-0.3278	-0.0972	-0.1070	-0.1100
	(S)	0.2758	-0.2940	-0.5925	0.1521	-0.1906	-0.2889	-0.0746	-0.0409	0.0170
	(K)	0.1951	-0.2027	-0.4162	0.1108	-0.1506	-0.2015	-0.0508	-0.0321	0.0224
		121	121	120	121	121	121	121	114	118
Mg	(P)	0.7238	0.2885	0.2049	0.8720	0.6938	0.5620	0.1888	0.7001	0.7581
	(S)	0.6515	0.0434	-0.0387	0.7678	0.5795	0.5843	0.0799	0.6814	0.6827
	(K)	0.4740	0.0319	-0.0245	0.5939	0.4215	0.4154	0.0498	0.5036	0.4971
		122	122	121	122	122	122	122	115	119
Ca	(P)	1.0000	-0.0359	-0.2318	0.7471	0.4892	0.2796	-0.0756	0.5145	0.6868
	(S)	1.0000	0.0232	-0.1568	0.6662	0.4490	0.4155	-0.0351	0.5808	0.6557
	(K)	1.0000	0.0228	-0.1161	0.5087	0.3261	0.2974	-0.0158	0.4120	0.4655
		122	122	121	122	122	122	122	115	119
Zn	(P)	-0.0359	1.0000	0.6638	0.4580	0.3798	0.4463	0.5174	0.3560	0.2190
	(S)	0.0228	1.0000	0.5452	0.2993	0.3934	0.4508	0.4721	0.2835	0.1814
	(K)	0.0232	1.0000	0.4173	0.2218	0.2929	0.3212	0.3292	0.2136	0.1678
		122	122	121	122	122	122	122	115	119
Cu	(P)	-0.2318	0.6638	1.0000	0.1996	0.3906	0.5445	0.4827	0.2302	0.1091
	(S)	-0.1161	0.4173	1.0000	0.0020	0.4319	0.5515	0.4334	0.2152	0.1615
	(K)	-0.1568	0.5452	1.0000	-0.0012	0.3131	0.3911	0.3224	0.1493	0.1441
		121	121	121	121	121	121	121	114	118
Mn	(P)	0.7471	0.4580	0.1996	1.0000	0.7363	0.5776	0.2231	0.7565	0.7906
	(S)	0.5087	0.2218	-0.0012	1.0000	0.6376	0.6010	0.1249	0.7854	0.7463
	(K)	0.6662	0.2993	0.0020	1.0000	0.4676	0.4593	0.0946	0.6085	0.5517
		122	122	121	122	122	122	122	115	119
Fe	(P)	0.4892	0.3798	0.3906	0.7363	1.0000	0.7464	0.2062	0.6891	0.6924
	(S)	0.3261	0.2929	0.3131	0.4676	1.0000	0.7977	0.2541	0.6916	0.6854
	(K)	0.4490	0.3934	0.4319	0.6376	1.0000	0.6058	0.1785	0.5175	0.4994
		122	122	121	122	122	122	122	115	119
Al	(P)	0.2796	0.4463	0.5445	0.5776	0.7464	1.0000	0.2326	0.6994	0.6420
	(S)	0.2974	0.3212	0.3911	0.4593	0.6058	1.0000	0.3178	0.8060	0.6996
	(K)	0.4155	0.4508	0.5515	0.6010	0.7977	1.0000	0.2436	0.6120	0.5222
		122	122	121	122	122	122	122	115	119
Pb	(P)	-0.0756	0.5174	0.4827	0.2231	0.2062	0.2326	1.0000	0.2553	0.2930
	(S)	-0.0158	0.3292	0.3224	0.0946	0.1785	0.2436	1.0000	0.2028	0.3125
	(K)	-0.0351	0.4721	0.4334	0.1249	0.2541	0.3178	1.0000	0.1372	0.2511
		122	122	121	122	122	122	125	115	119
Co	(P)	0.5145	0.3560	0.2302	0.7565	0.6891	0.6994	0.2553	1.0000	0.8637
	(S)	0.4120	0.2136	0.1493	0.6085	0.5175	0.6120	0.1372	1.0000	0.8311
	(K)	0.5808	0.2835	0.2152	0.7854	0.6916	0.8060	0.2028	1.0000	0.6462
		115	115	114	115	115	115	115	115	115
Ni	(P)	0.6868	0.2190	0.1091	0.7906	0.6924	0.6420	0.2930	0.8637	1.0000
	(S)	0.4655	0.1678	0.1441	0.5517	0.4994	0.5222	0.2511	0.6462	1.0000
	(K)	0.6557	0.1814	0.1615	0.7463	0.6854	0.6996	0.3125	0.8311	1.0000
		119	119	118	119	119	119	119	115	119
PO_4-P	(P)	0.2589	0.0472	0.2654	0.3576	0.4694	0.4823	0.1708	0.3363	0.5030
	(S)	0.0893	0.1065	0.3295	0.1680	0.2779	0.2806	0.0895	0.2262	0.3612
	(K)	0.1281	0.1230	0.4366	0.2496	0.3973	0.4040	0.1099	0.3298	0.5116
		98	98	98	98	98	98	98	95	98
NH_4-N	(P)	0.6370	-0.2645	-0.4389	0.4342	0.2332	0.0677	-0.2655	0.4286	0.4191
	(S)	0.3381	-0.0407	-0.3068	0.3687	0.1151	0.0588	-0.0869	0.2569	0.1784
	(K)	0.4708	-0.0679	-0.4361	0.5139	0.1686	0.0657	-0.1333	0.3693	0.2642
		96	96	96	96	96	96	96	93	96
NO_3-N	(P)	0.0937	0.0847	-0.0967	-0.0037	-0.0765	-0.0123	-0.1409	0.0422	-0.0075
	(S)	0.0160	0.1074	-0.0413	0.0007	-0.0604	-0.0308	-0.1382	0.0395	0.0069
	(K)	0.0237	0.1480	-0.0566	-0.0049	-0.0775	-0.0316	-0.1733	0.0570	-0.0005
		98	98	98	98	98	98	98	95	98
SO_4-S	(P)	0.4617	0.1298	0.0488	0.6395	0.6814	0.5712	0.0010	0.7496	0.6739
	(S)	0.3353	0.1569	0.1069	0.4112	0.4908	0.3875	-0.0151	0.5452	0.4139
	(K)	0.4823	0.2327	0.1577	0.5765	0.6824	0.5720	-0.0165	0.6995	0.5788
		102	102	102	102	102	102	102	99	102
Cl	(P)	0.1621	-0.0054	-0.2399	0.1001	0.0619	0.0672	0.0560	0.1229	0.2091
	(S)	0.1405	0.0681	-0.1246	0.1813	0.0533	0.0052	0.0445	0.1011	0.2199
	(K)	0.2138	0.0720	-0.1866	0.2870	0.0814	0.0148	0.0612	0.1483	0.3105
		108	108	107	108	108	108	108	103	107
Si	(P)	0.5371	-0.0874	-0.0849	0.4611	0.3155	0.3703	-0.0382	0.2953	0.4866
	(S)	0.3701	0.1043	0.1549	0.2501	0.1502	0.2739	-0.0186	0.2778	0.3729
	(K)	0.5362	0.1224	0.1731	0.3809	0.2077	0.3763	-0.0338	0.4003	0.6128
		102	102	102	102	102	102	102	99	102
O_2-SATN	(P)	-0.1912	-0.2053	0.0316	-0.1441	-0.1329	-0.1118	-0.1189	-0.2056	-0.1074
	(S)	-0.3293	-0.1987	-0.0539	-0.1970	-0.2002	-0.1967	-0.0584	-0.2609	-0.1620
	(K)	-0.4608	-0.2770	-0.0843	-0.2816	-0.3089	-0.2913	-0.0618	-0.3739	-0.2365
		89	89	89	89	89	89	89	87	89
temp	(P)	-0.2743	-0.2736	-0.2775	-0.3470	-0.2952	-0.2436	-0.3067	-0.1115	-0.3696
	(S)	-0.0200	-0.1293	-0.1602	-0.0619	-0.1391	-0.1431	-0.2966	-0.0312	-0.2264
	(K)	-0.3209	-0.1643	-0.2354	-0.1219	-0.2048	-0.2113	-0.4189	-0.0311	-0.3404
		121	121	120	121	121	121	124	114	118

		PO_4-P	NH_4-N	NO_3-N	SO_4-S	Cl	Si	O_2-SATN	temp
O.D.-420	(P)	0.3915	0.1755	0.0145	0.5592	0.1231	0.2437	-0.0467	-0.0056
	(S)	0.3842	0.0193	0.0296	0.4878	0.0767	0.1917	-0.1588	-0.0126
	(K)	0.2650	-0.0062	0.0174	0.3525	0.0560	0.1311	-0.1068	-0.0105
		98	96	98	102	102	102	89	109
cond	(P)	0.1028	0.3371	-0.0292	0.6060	0.2134	0.2311	-0.2099	0.3630
	(S)	0.0432	0.3615	0.0604	0.6357	0.1277	0.1310	-0.3348	0.3578
	(K)	0.0374	0.2629	0.0420	0.4924	0.0843	0.1013	-0.2352	0.2453
		98	96	98	102	102	102	89	109
pH-fld	(P)	-0.3420	0.0367	0.0385	-0.3111	-0.0090	-0.1501	0.0162	-0.1894
	(S)	-0.2464	0.0859	-0.0143	-0.2865	0.0825	-0.0163	-0.0322	-0.2273
	(K)	-0.1703	0.0711	-0.0054	-0.1975	0.0648	-0.0074	-0.0148	-0.1553
		98	96	98	102	108	102	89	125
acidity	(P)	0.4067	0.2834	-0.0877	0.6610	0.1366	0.3811	-0.0977	-0.3576
	(S)	0.3258	0.2373	-0.0657	0.5909	0.0776	0.2358	-0.2312	-0.2563
	(K)	0.2229	0.1630	-0.0498	0.4257	0.0472	0.1749	-0.1553	-0.1807
		98	96	98	102	107	102	89	112
Eh	(P)	0.1547	-0.5084	-0.2518	0.3798	0.0167	0.1989	0.0000	-0.1450
	(S)	0.5833	-0.6380	-0.3575	0.4268	-0.0223	0.0559	0.0000	-0.2765
	(K)	0.5069	-0.5744	-0.2060	0.3223	0.0000	0.0135	0.0000	-0.2197
		10	8	9	13	13	13	1	14
NSPEC	(P)	-0.2840	-0.0652	-0.0020	-0.2223	0.0551	-0.0965	0.0417	0.4167
	(S)	-0.3038	-0.0164	0.0914	-0.2333	-0.0327	-0.0421	0.0684	0.2515
	(K)	-0.2217	-0.0114	0.0653	-0.1702	-0.0249	-0.0299	0.0472	0.1809
		98	96	98	102	108	102	89	125
Na	(P)	0.1248	0.6351	0.0277	0.5053	0.3880	0.3370	-0.0670	-0.0213
	(S)	0.0476	0.6440	0.0153	0.4971	0.3240	0.2695	-0.2137	-0.0102
	(K)	0.0365	0.4759	0.0208	0.3542	0.2108	0.1749	-0.1478	-0.0226
		98	96	98	102	108	102	89	121
K	(P)	-0.3307	0.3060	-0.0202	0.0195	0.0520	0.0669	-0.1178	0.1353
	(S)	-0.3133	0.3185	-0.0563	-0.0276	0.1542	0.0380	-0.1695	0.0985
	(K)	-0.2060	0.2170	-0.0377	-0.0151	0.1076	0.0175	-0.1210	0.0673
		97	95	97	101	107	101	88	120
MG	(P)	0.1605	0.4636	-0.0011	0.5382	-0.0057	0.3782	0.0031	-0.3309
	(S)	0.0795	0.4470	0.0250	0.4403	0.0486	0.1874	-0.1775	-0.1060
	(K)	0.0678	0.3128	0.0257	0.3060	0.0218	0.1348	-0.1239	-0.0667
		98	96	98	102	108	102	89	121
Ca	(P)	0.2589	0.6370	0.0937	0.4617	0.1621	0.5371	-0.1912	-0.2243
	(S)	0.1281	0.4708	0.0237	0.4823	0.2138	0.5362	-0.4608	-0.0209
	(K)	0.0893	0.3381	0.0160	0.3353	0.1405	0.3781	-0.3293	-0.0200
		98	96	98	102	108	102	89	121
Zn	(P)	0.0472	-0.2645	0.0847	0.1298	-0.0054	-0.0874	-0.2053	-0.2736
	(S)	0.1230	-0.0679	0.1480	0.2327	0.0720	0.1224	-0.2770	-0.1643
	(K)	0.1065	-0.0407	0.1074	0.1569	0.0681	0.1043	-0.1987	-0.1203
		98	96	98	102	108	102	89	121
Cu	(P)	0.2654	-0.4389	-0.0967	0.0488	-0.2399	-0.0849	0.0316	-0.2775
	(S)	0.4366	-0.4361	-0.0566	0.1577	-0.1866	0.1731	-0.0843	-0.2354
	(K)	0.3295	-0.3068	-0.0413	0.1069	-0.1246	0.1549	-0.0539	-0.1602
		98	96	98	102	107	102	89	120
Mn	(P)	0.3576	0.4342	-0.0037	0.6395	0.1001	0.4611	-0.1441	-0.3470
	(S)	0.2496	0.5139	-0.0049	0.5765	0.2870	0.3809	-0.2816	-0.1219
	(K)	0.1680	0.3687	0.0007	0.4112	0.1813	0.2581	-0.1970	-0.0819
		98	96	98	102	108	102	89	121
Fe	(P)	0.4694	0.2332	-0.0765	0.6814	0.0619	0.3155	-0.1329	-0.2852
	(S)	0.3973	0.1686	-0.0775	0.6824	0.0814	0.2077	-0.3089	-0.2098
	(K)	0.2779	0.1151	-0.0604	0.4908	0.0533	0.1502	-0.2082	-0.1391
		98	96	98	102	108	102	89	121
Al	(P)	0.4823	0.0677	-0.0123	0.5712	0.0672	0.3703	-0.1118	-0.2436
	(S)	0.4040	0.0657	-0.0316	0.5720	0.0148	0.3763	-0.2913	-0.2113
	(K)	0.2806	0.0588	-0.0308	0.3875	0.0052	0.2729	-0.1967	-0.1431
		98	96	98	102	108	102	89	121
Pb	(P)	0.1708	-0.2655	-0.1409	0.0010	0.0560	-0.0382	-0.1189	-0.3067
	(S)	0.1099	-0.1333	-0.1733	-0.0165	0.0612	-0.0338	-0.0818	-0.4189
	(K)	0.0895	-0.0869	-0.1382	-0.0151	0.0445	-0.0186	-0.0584	-0.2886
		98	96	98	102	108	102	89	124
Co	(P)	0.3363	0.4286	0.0422	0.7496	0.1229	0.2953	-0.2056	-0.1115
	(S)	0.3298	0.3693	0.0570	0.6995	0.1483	0.4003	-0.3739	-0.0511
	(K)	0.2262	0.2569	0.0395	0.5452	0.1011	0.2778	-0.2609	-0.0312
		95	93	95	99	103	99	87	114
Ni	(P)	0.5030	0.4191	-0.0075	0.6739	0.2091	0.4866	-0.1074	-0.3696
	(S)	0.5116	0.2642	-0.0005	0.5788	0.3105	0.5128	-0.2365	-0.3404
	(K)	0.3612	0.1784	0.0069	0.4139	0.2199	0.3729	-0.1628	-0.2264
		98	96	98	102	107	102	89	118
PO_4-P	(P)	1.0000	-0.0448	0.0770	0.2861	0.2202	0.3378	0.0937	-0.1968
	(S)	1.0000	-0.0966	0.0859	0.2340	0.1519	0.3049	0.1109	-0.2385
	(K)	1.0000	-0.0565	0.0599	0.1588	0.1008	0.2201	0.0846	-0.1664
		98	94	95	98	98	98	88	97
NH_4-N	(P)	-0.0448	1.0000	0.1105	0.4553	0.2995	0.1433	-0.0168	0.1909
	(S)	-0.0565	1.0000	0.0841	0.3700	0.2966	0.0609	-0.1423	0.1898
	(K)	-0.0966	1.0000	0.0539	0.2476	0.1998	0.0410	-0.0971	0.1112
		94	96	95	96	96	96	87	95
NO_3-N	(P)	0.0770	0.1105	1.0000	-0.0351	0.0093	-0.0905	0.0446	0.1625
	(S)	0.0599	0.0539	1.0000	-0.0116	-0.0742	-0.0468	-0.0147	0.2668
	(K)	0.0859	0.0841	1.0000	-0.0102	-0.0537	-0.0355	-0.0134	0.2047
		95	95	98	98	98	98	89	97
SO_4-S	(P)	0.2861	0.4553	-0.0351	1.0000	0.2731	0.4356	-0.2143	0.2398
	(S)	0.1588	0.2476	-0.0102	1.0000	0.2279	0.3113	-0.3484	0.2214
	(K)	0.2340	0.3700	-0.0116	1.0000	0.1576	0.2252	-0.2369	0.1535
		98	96	98	102	102	102	89	101
Cl	(P)	0.2202	0.2995	0.0093	0.2731	1.0000	0.4818	-0.1397	0.0004
	(S)	0.1008	0.1998	-0.0537	0.1576	1.0000	0.4892	-0.2297	-0.0392
	(K)	0.1519	0.2966	-0.0742	0.2279	1.0000	0.3590	-0.1528	-0.0259
		98	96	98	102	108	102	89	107
Si	(P)	0.3378	0.1433	-0.0905	0.4356	0.4818	1.0000	-0.2186	-0.0713
	(S)	0.2201	0.0410	-0.0355	0.2252	0.3590	1.0000	-0.3065	-0.0634
	(K)	0.3049	0.0609	-0.0468	0.3113	0.4892	1.0000	-0.2076	-0.0486
		98	96	98	102	102	102	89	101
O_2-SATN	(P)	0.0937	-0.0168	0.0446	-0.2143	-0.1397	-0.2186	1.0000	-0.2465
	(S)	0.0846	-0.0971	-0.0134	-0.2369	-0.1528	-0.2076	1.0000	-0.3603
	(K)	0.1109	-0.1423	-0.0147	-0.3484	-0.2297	-0.3065	1.0000	-0.2571
		88	87	89	89	89	89	89	89
temp	(P)	-0.1968	0.1909	0.1625	0.2398	0.0004	-0.0713	-0.2465	1.0000
	(S)	-0.1664	0.1112	0.2047	0.1535	-0.0259	-0.0486	-0.2571	1.0000
	(K)	-0.2385	0.1898	0.2668	0.2214	-0.0392	-0.0634	-0.3603	1.0000
		97	95	97	101	107	101	89	125

4.33 Principal components analysis

The results of the two PCA's performed on the Pearson product moment correlation matrix of the standardised transformed variables are presented in Table 4.33-1. The matrix presented is the principal component loading matrix produced by multiplying each eigenvector value by the square root of the corresponding eigenvalue; the resulting matrix is also known as the "factor matrix" (Harman, 1977). The components were chosen for display using the rule of thumb suggested by Kaiser (1970) whereby only those components where the corresponding eigenvalue is greater than unity are included. In the discussion that follows a variable which loads negatively on a component is distinguished with a minus sign placed before the variable name thus -Zn. The weights of some components have been multiplied by -1 (Harman, 1977) to ensure that the signs coincide. This modification is valid, as explained in Section 3.43, because the ellipsoid which represents the component can, without loss of generalisation, be rotated about its midpoint in any axis (the effect of multiplying all the weightings of a component by -1). Where this modification to the weightings has been employed the relevant component is marked in Table 4.33-1 with an asterisk.

Of immediate interest is the marked similarity in the variables which load on the first two components of the two datasets. In both analyses the first component appears to be a general acid component with most of the cations, SO_4-S, OD-420, total acidity and conductivity loading on it at relatively high levels. Conspicuous by its low value is the loading of pH. In both analyses the second component is identified by the high loading of Cu, -Na, -K and -Ca. Also loading on this component are Fe, Al, Pb and -pH although the levels of these loadings are low. The two analyses are different in their variable composition for components 3 and 4. Component 3 is characterised by high loadings of -Pb, -Zn, and Si for the 217 dataset and -Pb, temperature and -pH for the 101 dataset. Component 4 is characterised by temperature and -Si and to lesser extent -Cl in the 217 dataset and PO_4-P and NSPEC in the 101 dataset. The 5th component is identical in both analyses and is identified by the high NO_3-N loading. The greater variation in the smaller 101 dataset as represented by the extra component is explained in terms of an Si, NSPEC, -NO_3-N component with the signs for NO_3-N and Si suggesting that they act in antipathy.

These results suggest that the major source of variation is the relative level of acid. What additional variation is present is due to the high value of certain specific variables. Of these Cu is perhaps the most interesting

Table 4.33-1

Principal components analysis of the 217 and 101 chemical datasets

Component	1		2		3		4		5		6	
Dataset	217	101	217	101	217	101	217	101	217	101	217	101
Rotation			*					*				
Eigenvalue	9.8	8.2	3.6	4.1	1.7	2.0	1.3	1.5	1.0	1.2	0.7	1.1
% variation	42.8	34.3	15.6	17.1	7.3	8.2	5.8	6.4	4.5	4.8	3.3	4.7
cumulated %	42.8	34.3	58.5	51.3	65.7	59.6	71.5	66.0	76.0	70.9	79.3	75.5
CO-42E	0.771	0.651	0.196	0.363	0.046	-0.017	0.041	-0.103	0.081	-0.129	-	-0.271
acidity	0.834	0.692	-0.122	-0.042	-0.029	0.350	0.211	0.272	-0.177	-0.136	-	0.166
pH	-0.544	-0.448	-0.542	-0.366	0.358	-0.524	-0.193	0.206	0.095	0.068	-	-0.221
conductivity	0.854	0.865	0.164	0.176	-0.035	-0.062	0.001	0.035	-0.073	-0.136	-	-0.060
Na	0.517	0.695	-0.737	-0.536	0.094	-0.067	-0.054	0.146	-0.011	-0.092	-	-0.001
K	0.037	0.018	-0.756	-0.662	0.389	-0.318	-0.144	0.295	-0.218	-0.224	-	-0.186
Mg	0.799	0.773	-0.373	-0.258	0.005	0.024	0.042	0.183	-0.095	-0.183	-	-0.035
Ca	0.615	0.560	-0.619	-0.668	-0.091	-0.064	-0.079	-0.325	0.101	0.082	-	-0.014
Zn	0.502	0.095	0.449	0.648	0.522	-0.116	-0.103	0.479	-0.117	0.282	-	0.095
Cu	0.199	0.042	0.828	0.912	0.196	0.062	-0.005	0.106	-0.061	-0.043	-	0.070
Mn	0.903	0.813	-0.174	-0.236	0.026	-0.094	-0.093	0.165	-0.054	0.125	-	0.119
Fe	0.833	0.796	0.346	0.321	-0.132	0.071	0.055	-0.112	-0.032	-0.133	-	-0.119
Al	0.839	0.713	0.296	0.510	0.271	0.023	-0.019	0.041	-0.010	-0.015	-	0.029
Pb	0.543	0.079	0.262	0.545	0.557	-0.649	-0.298	0.264	-0.097	-0.066	-	-0.051
Co	0.634	0.637	0.022	0.130	0.105	0.031	0.018	0.271	0.035	0.086	-	-0.086
Ni	0.895	0.867	0.005	0.010	0.013	-0.228	-0.169	-0.050	0.114	0.059	-	-0.067
PO_4-P	0.465	0.435	0.362	0.337	-0.277	-0.069	-0.121	-0.534	0.337	0.373	-	-0.049
NH_4-N	0.636	0.486	-0.377	-0.541	-0.022	0.040	0.225	0.093	0.194	0.099	-	-0.276
NO_3-N	0.176	-0.036	-0.030	0.039	0.331	0.179	0.379	0.191	0.757	0.713	-	-0.482
SO_4-S	0.799	0.833	-0.053	-0.004	-0.005	0.112	0.347	0.096	-0.138	-0.021	-	0.017
Cl	0.541	0.419	-0.271	-0.178	0.012	-0.411	-0.419	-0.105	0.166	0.416	-	0.361
Si	0.199	0.540	0.055	-0.149	-0.029	-0.265	-0.599	-0.306	0.014	0.110	-	0.462
Temperature	0.476	0.153	-0.151	-0.254	-0.328	0.813	0.631	0.137	-0.258	0.160	-	0.132
NSPEC		-0.327		-0.213		-0.014		0.519		0.133		0.442

(* indicates that a component has been reflected through 180 degrees as described in Section 3.42)

and its exceptionally high loading on component 2 is of considerable importance. The absence of an "ameliorating component" as reported by Say et al. (1976) for the high Zn environment is also noted. The low loading of NSPEC and PO_4-P on component 4 is not easily explained and may well be an artefact of the data subset since there is no similar corresponding component in the 217 dataset. Component 5 stands out as anomalous since only NO_3-N loads on it at any appreciable level - although NO_3-N does also load on the general acid component 1. Component 6 in the 101 dataset is of special interest since it includes loadings for NSPEC and Si with $-NO_3$-N. This suggests that a weak link exists between the number of species at a site and the Si level. This observation is of importance because of the diatom requirement for Si. The loading of Si on components 1, 4 and 6 for the 101 dataset and 3 and 4 for the 217 dataset suggest a highly complex variation pattern for this variable. The complex variation pattern for pH suggests that it is acting independently of the other variables. It is highly probable that this independence is due to the nature of the dataset specification mechanism as outlined in Section 4.21 and therefore reflects the bias introduced by the use of pH as the sole data selecting criterion variable.

4.4 Cluster analysis

4.41 Introduction

In 4.4 the results of the various cluster analyses are presented. Section 4.42 presents a comparison of the 8 different hierarchical cluster techniques used. The similarity in results suggest that the rationale of examining only the two most dissimilar methods when other data are being examined is correct (see 2.248). The cluster levels and clusters obtained as displayed by the various dendrograms presented are discussed both for the chemical and biological analyses in Section 4.43. Finally the species clusterings are presented (Section 4.43).

4.42 Comparison of cluster techniques

The eight hierarchical methods used have the following characteristics which have been observed before (Wishart, 1978):

1. Single linkage

   Defines the similarity between clusters P and Q to be the highest similarity coefficient between the two individuals one from each cluster. Single linkage tends to find "straggling" clusters and often ends merely in chaining if large populations are being investigated (for example the 269 chemistries dataset).

2. Complete linkage

   Defines the similarity between clusters as the smallest similarity coefficient between individuals one from each cluster. Complete linkage finds spherical clusters but is liable to produce irregular results because the similarity criterion is determined for only two individuals and does not measure group structure.

3. Centroid sorting

   Defines the similarity as the average of all the similarity coefficients for pairs of individuals one from each cluster. This method found spherical clusters and is therefore preferable to the complete linkage method since it does attempt to take account of group structure.

4. Centroid sorting

Computes the centroid of clusters P and Q as the mean vector for the two clusters then calculates the Euclidean distance between these in the usual way. This method resulted in chaining of the 269 data.

5. Median method

Defines the similarity of R to P+Q as the distance from the centroid of R to the midpoint of the line joining the centroids of P and Q. The trivial two cluster case is clearly the midpoint of the line joining their centroids. This method like the median method is limited to distance similarity coefficients and chained both the 125 and 269 datasets.

6. Ward's method

Defines the error sum of squares as the sum of the distances from each individual to the centroid of the cluster. The technique merges the two clusters which yield the least increase in the error sum of squares. The method is highly to be recommended since it finds minimum variance spherical clusters.

7. Flexible BETA

The similarity between R and P+Q is defined as the variable transformation:

$$s(R,P+Q) = ((s(R,P)+s(R,Q)) * (1-BETA)/2 + s(P,Q) * BETA \quad ...1$$

where BETA is a parameter having a value less than 1 (-0.25 was used as suggested by Lance & Williams, 1967). The results here show a marked similarity to those obtained from Ward's method.

8. McQuitty's method

Uses the transformation:

$$s(R,P+Q) = (s(R,P) + s(R+Q)) / 2 \quad ...2$$

This method produced similar results to Ward's method both for the 125 data and the 269 datasets. However Wishart observed that this method will chain with large populations!

These 8 methods were used on the transformed 125 chemistries dataset where missing data were replaced by the mean as outlined in Table 4.15-2. They produced the dendrograms of sites which are compared in Table 4.42-1.

It can be seen that the group of techniques consisting of Ward's method, the flexible BETA method and McQuitty's method produce the most cross

Table 4.42-1

Cophenetic correlations among 8 different cluster techniques

		V1	V2	V3	V4	V5	V6	V7	V8
V1	single linkage	1.0000	.5197 .6721	.7164 .8687	.7822 .9157	.4397 .5359	.5851 .7344	.5845 .7335	.6459 .7995
V2	complete linkage	.7185	1.0000	.7340 .7967	.6708 .7523	.4263 .4252	.7830 .8394	.7763 .8293	.7784 .8283
V3	average linkage	.8853	.8031	1.0000	.8848 .9502	.5878 .6244	.8344 .8863	.8363 .8913	.8830 .9323
V4	centroid method	.8781	.6989	.9473	1.0000	.5862 .6068	.7540 .8494	.7591 .8441	.8216 .9014
V5	median method	.4292	.4154	.5019	.4275	1.0000	.5760 .6254	.5978 .6210	.6112 .6263
V6	Ward's method	.7854	.8354	.8718	.7665	.5044	1.0000	.9442 .9735	.8907 .9509
V7	flexible BETA	.7857	.8902	.8742	.7671	.5131	.9930	1.0000	.9191 .9722
V8	McQuitty's method	.8172	.8729	.8993	.7978	.5431	.9511	.9766	1.0000

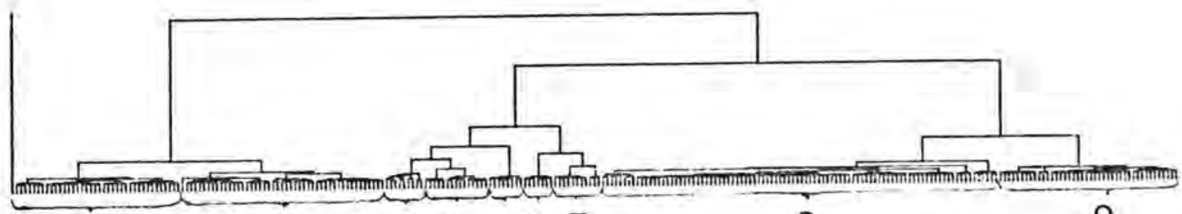
(Top half diagonal matrix contains Kendals and Spearmans rank correlation coefficients bottom half contains Pearsons correlation coefficient. n = 7749)

correlated dendrograms. The method which is least cross correlated is the median method. The use of Ward's method, the median method and one of the linkage methods (average linkage) would seem to the give a range of possible clusterings as a basis for discussion. However, although all these methods were used, only the results of the use of Ward's method are presented. This is because of the degree of chaining which the Median method introduced, and the considerable similarity of the dendrograms.

Table 4.43-1

Dendrogram produced by Wards Method of the 269 acid chemistries

ssss.rr stream and reach numbers are as in Table 5.12-1 p. 158



	1	2	3	4	5	6	7	8		9
left ↓	0152.10	1004.01	0153.01	9001.01	9008.45	2003.01	0213.01	0066.01	0127.01	0127.04
	0152.20	0130.02	3016.03	9004.01	9013.03	2014.01	0214.21	2015.10	0127.01	0125.01
	0154.01	0157.06	0100.30	9005.01	9008.48	2004.01	0216.50	0100.30	0127.01	0125.01
	0135.04	0130.03	9016.99	9006.01	9008.50	2006.01	0218.01	0133.01	0127.03	0125.01
	0137.01	0132.03	3016.01	9007.10	9008.52	2007.01	0215.01	0109.01	0127.01	0125.01
	0137.03	0157.07	9009.20	0127.02	9012.01	2007.03	0217.01	0127.02	0127.01	0125.01
	0137.02	0135.01	9010.02	0127.03	9009.20	2007.02	0215.15	0127.03	0127.01	0127.10
	0128.01	0135.02	0119.01	0127.04	9011.01		0216.15	0127.06	0127.01	0127.09
	0132.01	0135.03	0120.01	0127.06			0220.03	0127.01	0127.01	0127.08
	0158.08	0160.05		0127.05			0220.15	0127.01	0127.01	0127.11
	0159.09	0160.06		0127.07			0221.05	0127.04	0127.01	0127.12
	0128.02	0130.04		0127.08				0127.01	0127.01	0127.13
	0128.02	0135.01		0127.09				0127.01	0127.01	0125.01
	0128.03	0135.02		0127.10				0127.01	0127.01	0125.01
	0128.01	0135.03		0127.11				0127.02	0127.02	0125.01
	0132.02	0140.01						0127.03	0127.03	0127.10
	0132.04	0140.01						0127.01	0127.05	0127.11
	0132.05	0140.03						0127.01	0127.06	0127.08
	0132.06	0140.02						0127.01	0127.05	0127.08
	0132.07	0140.02						0127.03	0127.08	0127.09
	0132.01	0140.03						0127.02	0127.08	0127.08
	0132.04	0130.01						0127.02	0127.10	0127.10
	0132.07	0157.05						0127.01	0127.11	0127.09
	0132.02	0139.01						0127.01	0127.10	0127.11
	0132.06	0139.03						0127.03	0127.11	0127.11
	0134.01	0139.02						0127.04	0127.02	0127.13
	0134.03	0130.01						0127.05	0127.03	0127.13
	0132.05	0132.03						0127.03	0127.05	0127.10
	0135.04	0160.06						0127.04	0127.04	0127.11
	0134.03	0139.02						0127.03	0127.06	0127.07
	0134.02	0139.03						0127.02	0127.03	0127.09
	0134.04	0131.01						0127.03	0127.04	0127.08
	0127.04	0131.02						0127.04	0127.05	0127.09
	0158.08	0131.03						0127.04	0127.04	0127.10
	0133.01	0129.01						0127.04	0127.05	0127.10
	0137.01	0129.02						0127.05	0127.06	0127.11
	0137.02	0136.01						0127.01	0022.01	0127.11
right	0137.03	0136.01						0127.01	0022.01	0127.10
		0128.03						0127.01	0022.01	0127.07
		0159.09						0127.01	0022.01	0127.07
		0131.01						0127.01	0138.01	0127.07
		0130.02						0127.01	0138.01	0127.07
		0131.02						0127.01	0138.03	
		0131.03						0127.01	0138.02	
		0130.03						0127.01	0138.02	
		0160.05						0127.01	0138.03	
		0139.01								

4.43 Cluster levels and clusters obtained

The dendrogram of Ward's method using the 269 chemistry samples dataset is presented in Table 4.43-1. This contains all the important features of the analysis which will be brought out and displays the advantages of the use of Ward's method in clustering these data:

1. tight (spherical) clusters with an absence of chaining
2. low initial similarity levels suggesting clustering based on geographical proximity of sample sites
3. high final levels of inter-cluster similarity suggesting that the geographical units (inter site) similarity is significantly greater than any intra site similarity
4. the long tail of Brandon sample/sites.

Closer inspection of the dendrogram in Table 4.43-1 suggests that the 9 cluster level gives the most meaningful separation of sites based on the clustering of their chemistries. The streams and (where relevant) reaches which this represent are:

1. Ø154 & Ø152 Streams at Westfield, Ø135 Denby A, Ø137 Polesworth, Ø128 Walkmill, Ø132 Chisnall Hall A, Ø158 Chisnall Hall B, Ø159 Chisnall Hall C, Ø134 Gibfield, Ø133 Welsh Whittle, Ø127.Ø4 Brandon (Ø4)
2. 1ØØ4 Bois de Cornillion, Ø13Ø Rowley A, Ø157 Rowley B, Ø132 Chisnall Hall A, Ø135 Denby A, Ø16Ø Denby B, Ø14Ø Kingsbury, Ø139 Birch Coppice, Ø131 Deerplay, Ø129 Oatlands, Ø136 Cannock, Ø128 Walkmill, Ø159 Chisnall Hall C
3. Ø153 Westfield, 3Ø16 Soussu Seepage, Ø1ØØ Dowgang, 9Ø16 West Fork USA, 9ØØ9 Tantalus Creek USA, 9Ø1Ø Root Pool USA, Ø119 Rake Sike A, Ø12Ø Rake Sike B
4. 9ØØ1 Burra-Burra Creek Trib., 9ØØ4 Shinniston , 9ØØ5 Lumberport, 9ØØ6 Lamberts Run, 9ØØ7 East Morgantown , Ø127 Brandon (reaches Ø2,Ø3,Ø4,Ø5,Ø6,Ø7,Ø8,Ø9,1Ø,11)
5. 9ØØ8 Lemonade Creek, 9ØØ9 Tantalus Creek, 9Ø11 Trib. Near Monarch Geyser, 9Ø12 Near Monarch Geyser
6. Ø213 Parys Iron Flush, Ø214 Parys lagoon A, Ø215 Parys lagoon B, Ø216 Afon Goch East, Ø217 Dyffryn Adda adit , Ø218 Parys South West adit, Ø22Ø Adeer effluent A, Ø221 Adeer effluent B

7. 2003 Tigroney 850 A, 2004 Tigroney 850 B, 2006 Tigroney Grass level, 2007 Tigroney Barrack level, 2014 Tigroney 850 C

8. 0127 Brandon (reaches 01,02,03,04,05,06,07,10,11), 2015 Red Road Stream, 0066 Lowlands site, 0022 Red Burn, 0138 Bridford, 0100 Dowgang

9. 0127 Brandon (reaches 04,07,08,09,10,11,12,13), 0125 Brandon Pithouse Stream B.

Inspection of the dendrogram in Table 4.43-2 of the 125 chemistry samples datset suggests a 7 cluster separation of sites based on the clustering of the chemistry data. The clusters are:

1. Westfield, Polesworth, Chisnall Hall A, Chisnall Hall B, Chisnall Hall C, Gibfield, Welsh Whittle, Brandon (04), Rowley A, Rowley B, Kingsbury, Denby A, Denby B.

   Which corresponds well with clusters 1 and 2 of the 269 data

2. Rowley A, Rowley B, Birch Coppice, Chisnall Hall A, Chisnall Hall C, Deerplay, Oatlands, Cannock, Walkmill, Denby B,

   Clusters 1 and 2 seperate the Spring and Autumn data for the same sites This fine separation is obscured by the large quantity of data present in the 269 dataset

3. Lowlands, Red Road Stream, Brandon (01,02,03,04), Bridford, Welsh Whittle,

   This cluster identifies with cluster 8 of the 269 dataset. The inclusion of the Welsh Whittle site suggests a link with cluster 1

4. All Avoca sites.

   This cluster is identical to cluster 6 of the 269 dataset

5. Westfield, US mining , Soussu Seepage,

   This cluster is similar to cluster 4 of the 269 dataset however, the Westfield site and the French Soussu Seepage sites are anomalous

6. Parys Mountain, Adeer effluent

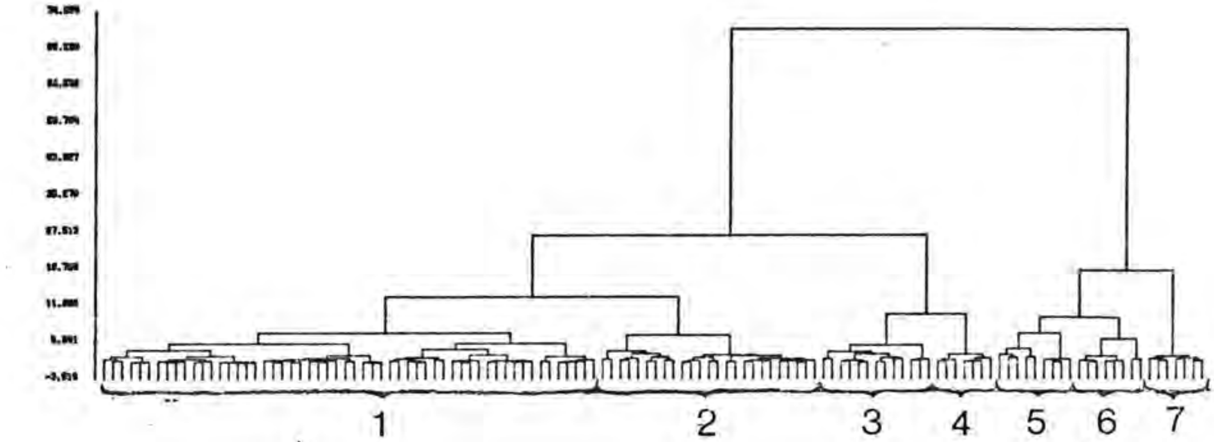
   This cluster is identical to cluster 7 of the 269 dataset

7. US Yellowstone sites

   This is identical to cluster 5 of the 269 dataset.

Table 4.43-2

Dendrogram produced by Wards Method of the 125 acid chemistries
ssssrr stream and reach numbers are as in Table 5.12-1 p.158



left ↓

0152.10
0152.20
0154.01
0137.01
0137.03
0137.02
0128.01
0132.01
0158.08
0159.09
0128.02
0128.02
0128.03
0132.02
0132.04
0132.05
0132.06
0132.07
0134.03
0132.02
0132.06
0134.02
0134.04
0128.01
0132.05
0134.03
0133.01
0127.04
0158.08
0137.01
0137.02
0137.03
0130.02
0157.06
0135.01
0135.02
0135.03
0132.03
0157.07
0140.01
0140.01
0140.03
0140.02
0160.05
0160.06
0130.04
0135.01
0135.02
0135.03
0140.02
0140.03
0132.01
0132.04
0132.07
0134.01
0160.06

0130.01
0157.05
0139.01
0139.03
0139.02
0130.01
0132.03
0139.02
0139.03
0131.01
0131.02
0131.03
0129.01
0129.02
0136.01
0136.01
0128.03
0159.09
0131.01
0160.05
0130.02
0130.03
0131.02
0131.03
0139.01

0066.01
2015.10
0127.01
0127.01
0127.02
0127.03
0127.02
0127.03
0127.04
0133.01
0138.01
0138.02
0138.03

2003.01
2014.01
2004.01
2006.01
2007.01
2007.03
2007.02

0153.01
9009.20
3016.03
9016.99
3016.01
9001.01
9005.01
9006.01
9007.10

0215.01
0215.15
0216.15
0216.50
0217.01
0218.01
0220.03
0221.05

9008.45
9008.48
9008.52
9009.20
9013.03
9011.01
9012.01

right

The marked similarity in the clusters of the two datasets especially in the separation of the Brandon sites, the USA mining, USA Yellowstone, Avoca and Parys Mountain sites although striking is not altogether surprising since the chemical data on which the clusterings are based are identical. The English non-Brandon sites cluster into three groups in the 269 dataset which coalesce at levels above that selected for cluster identification and separation. The reasons for this three cluster separation are not clear. In the 125 dataset these same sites cluster into two groups which may be loosely identified as a spring data cluster and a late summer/autumn data cluster. In both data subsets these clusters coalesce before being joined by the non-English sites. It is probable that the three cluster separation of the non-Brandon English sites in the 269 dataset follows some complex pattern probably similar to the spring/autumn separation of the 125 dataset but that this is being masked by the large number of sites considered and probably also the large number of multiple Brandon samplings.

Inspection of the three main clusters containing Brandon sites in the 269 dataset (clusters 4,8 & 9) suggests a three way division of these sites. The division is into reach Ø1, reaches Ø2 - 11 and reaches 12 & 13 and represent an upstream - downstream chemical polarity with reaches Ø2 - 11 being capable of moving toward either pole. Always associated with the "downstream pole" is the Brandon Pithouse stream B (see also Section 5.22).

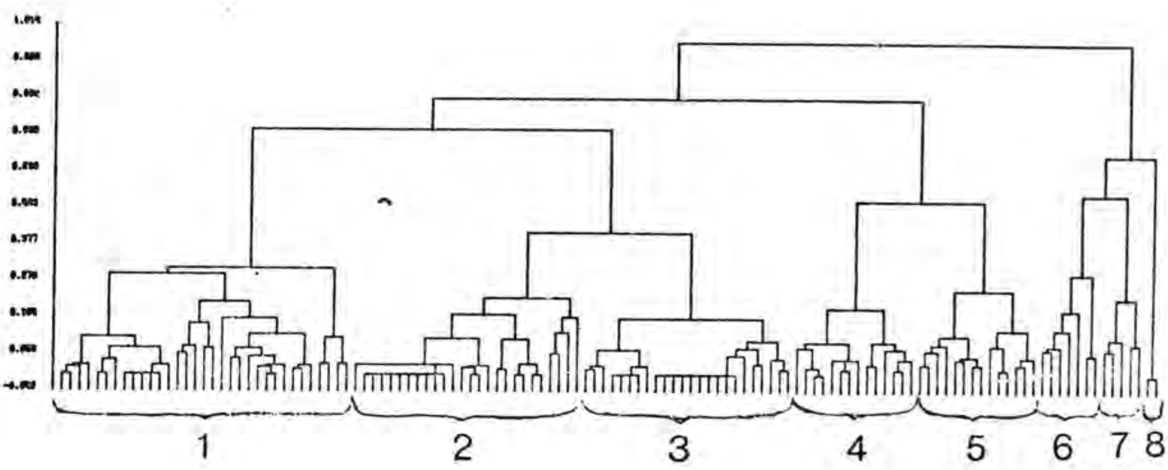
The 269 dataset cannot be relied on for further analyses because of the large amount of variation introduced into the dataset by the quantity of data and the number of multiple Brandon samplings. One of the major facets of the analysis masked by this problem in the 269 dataset is the clustering together of the Avoca and Brandon sites which can be seen easily in the 125 dataset. This new cluster, then clusters with the non-Brandon English sites to give a British data cluster - which then clusters with the American/Parys Mountain data. Although these higher level clusters are loosely defined they do suggest affinities which should be examined further.

The dendrogram in Table 4.43-3 is of sites based on the biological presence/absence data found in the 125 dataset and was produced using Ward's method of clustering an Euclidean distance similarity matrix. The 8 cluster level was chosen for analysis. The sites which form these clusters do not follow a geographical proximity pattern as was discernable in the 125 chemistries dataset. Comparison with Table 6.14-1 which contains the linearisation of sites afforded by the dendrogram in Table 4.43-2 against species abundances where the order of species is the minimum pH at which that

Table 4.43-3

Dendrogram produced by Wards Method of the 125 sites based on the presence/absence biological data

ssss.rr stream and reach numbers are as in Table 5.12-1 p158



left ↓	0152.10	2004.01	0130.01	0215.01	0132.06	9001.01	9006.01	9009.20
	0136.01	0127.03	0135.03	0134.03	0158.08	0218.01	9008.48	9009.20
	0153.01	0139.03	0157.07	0134.03	0140.01	0216.50	9008.52	
	0140.02	0127.03	0130.02	0134.04	0140.01	0216.15	9011.01	
	0152.20	0130.04	0132.05	0128.02	0135.02	9013.03	9012.01	
	0154.01	0132.05	0135.02	0132.01	0158.08	3016.01		
	2014.01	0134.02	0130.02	0132.02	0140.02	3016.03		
	0066.01	0135.01	0132.04	0134.01	0140.03			
	0131.01	0135.03	0132.03	0128.01	0137.01			
	0135.01	0136.01	0132.07	0128.03	0137.03			
	0160.05	0139.02	0136.01	0128.03	0137.03			
	0160.06	0129.01	0139.02	0133.01	0137.02			
	0132.01	2007.01	0159.09	0128.01	0137.02			
	2003.01	0128.02	0129.02	0132.02	0137.01			
	2006.01	0133.01	0130.01					
	2015.10	0132.03	0130.03					
	2007.03	2007.02	0131.01					
	0138.03	0138.02	0139.03					
	9005.01	0139.01	0157.05					
	9007.10	0139.01	0140.03					
	0215.15	0160.05	0132.04					
	0217.01	0160.06	0132.06					
	0127.02	9008.45	0132.07					
	0157.06	0220.03	0159.09					
	0131.03	0221.05						
	0127.02	9016.99						
	0131.02							
	0131.02							
	0131.03							
	0127.01							
	0127.01							
	0127.04							
right	0127.04							

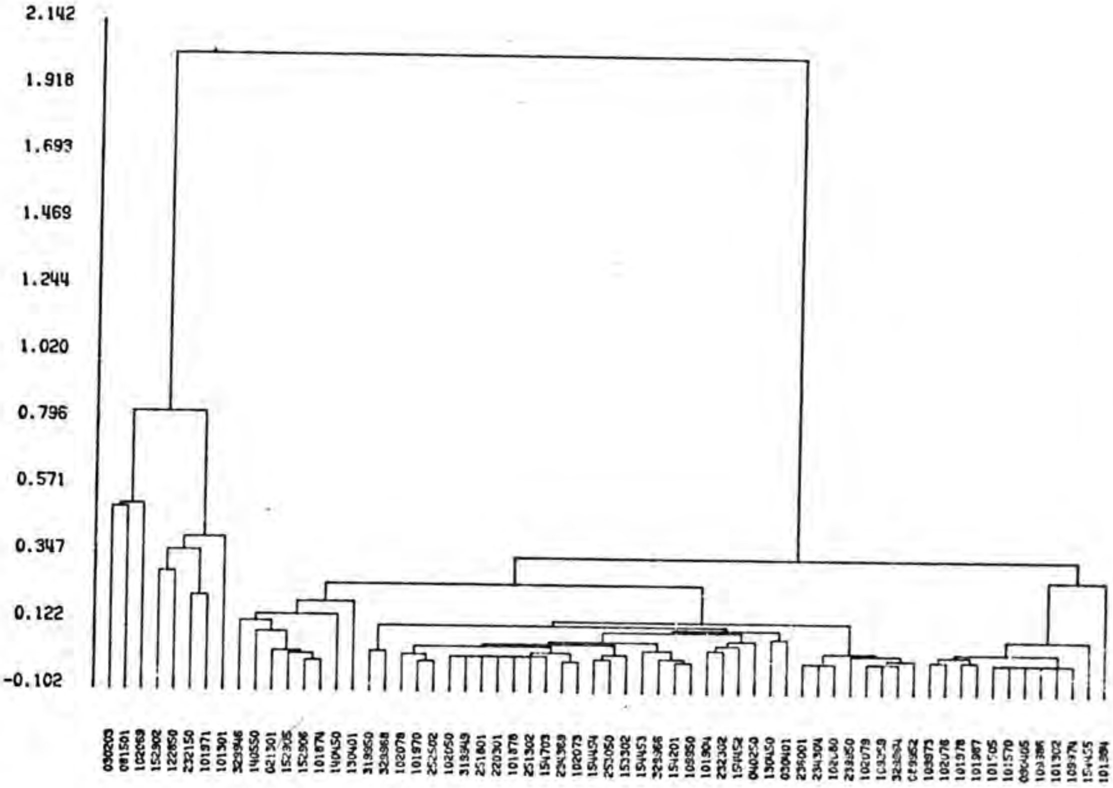
had been found (minimum occurrence pH) suggests that the clusters are in some way linked to pH or variables associated with pH. In general sites with lower pH cluster together before being joined by sites (which may already be in clusters) at higher pH. Thus in Table 6.14-1 sites near a right cluster boundary have higher pH in general than those on the left cluster boundary. (It must; however, must be borne in mind that sites (or clusters) within a cluster are free to rotate about join points and examination of the two tables in conjunction must be made to verify the observation.)

Table 4.43-4 contains the dendrogram of the species based on the biological presence/absence data. If the 13 cluster level is examined a two way division of clusters may be observed. This division is into the single species clusters (which are the more or less ubiquitously distributed species with regards to pH) and multiple species clusters. Further examination reveals that the more sites that two species occur at, the more likely those two species will be dissimilar in their distribution and hence the less likely they are to occur in the same cluster. This suggests that those species which occur together only in species poor sites (by chance or naturally) will be strikingly well clustered together. This observation will probably be true of any extreme environment where a range of species tolerance is present and probably represents an artefact of the equal weight placed on a species absence which may or may not be valid. Alternatively it may be an artefact of the greater care taken in establishing the completeness of the species complement of species poor sites (which is another way of saying that extra weight is being placed on species absence).

Table 4.43-4

Dendrogram of species produced by Wards method using the 125 presence/absence data

PPggss species numbers are as in Table 6.11-1 p 172.


2.142
1.918
1.693
1.469
1.244
1.020
0.796
0.571
0.347
0.122
-0.102

4.5 Species pH sensitivity and related statistics

4.51 Introduction

The 91 species (63 live & 28 dead) found on performing a "bydate" expansion of the acid samples are more fully dealt with in Section 3.4. However, of particular interest is the reaction of a species to a lowering of pH, and more especially the pH at which a species ceases to be viable. The pH at which a species is "significantly" absent or over abundant is represented using the novel chi-square ratio described in 2.249.

Only the 30 photosynthethic species which were present live in 3 or more samples are presented here although all photosynthetic species present live were used to generate the statistics.

4.52 Reaction types recognised and the inviability level

The Table of figures 4.52-1 presents for each species three histograms representing the absolute level (count of occurences), relative level (percentage of total count) and species chi-square ratio for the species considered. Each histogram has the same axis divisions for pH. The bin lengths are, where relevant, scaled proportional to the largest bin length and this maximum figure is noted in the diagram. The units for the histogram bin lengths are respectively number of individuals, percentage of individuals in a bin verses the total number of species in that bin and an arbitary chi-square unit. The first histogram shows the distribution of the species over the restricted pH range, the second the relative importance of that species as a contributor to the species complement of the respective pH bin and the third attempts to attach some measure of importance to the presence or absence of the species from a bin.

Considering the relative percentage histograms, 4 types of distribution can be recognised, which will be described here using examples drawn from Table 4.52-1. Type 1 is characterised by Hormidium rivulare and depicts a characteristic steady fall off of a species as the pH is lowered. This is shown in Table 4.52-2 which is the relevant diagram in Table 4.52-1 repeated with the relative percentage histogram drawn to a larger scale. Above pH 4.0, it is assumed that the species is healthy and in its more normal form. The species disappear below some more or less exact cut off pH level - or pH range. If the disappearance is due to an exact cut off level the line "A" (Table

Table 4.52-1 (pp 142-149)

Species pH sensitivity over the pH range 0.9 - 4.0

Bin lengths may be calculated from MAXn where:

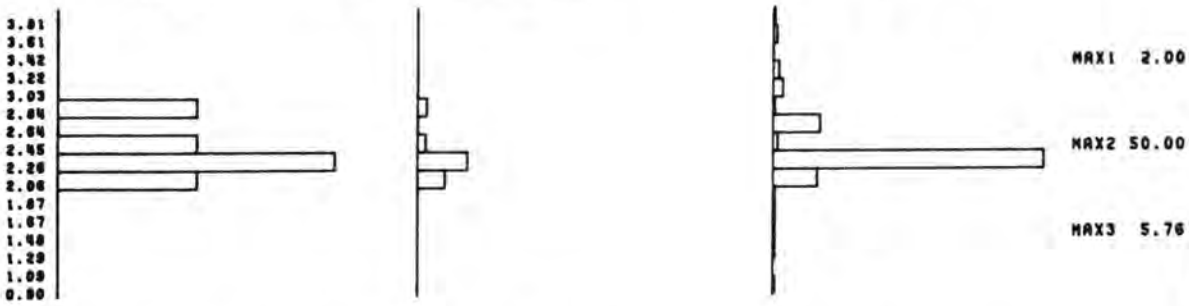
MAX1 (is a variable count of occurences) - and is the longest bin; scaled to 100%.

MAX2 (is fixed at 50%) - bins are scaled to double the relative %; values > 50% are scaled to 50%.

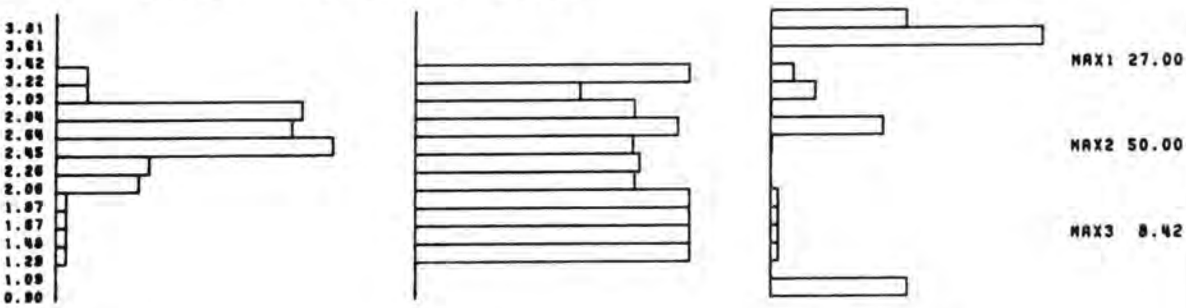
MAX3 (is a variable chi-square) - and is the longest bin; scaled to 100%.

(for further description see Section 4.52 p. 141)

021301 CYANIDIUM CALDARIUM NS = 5



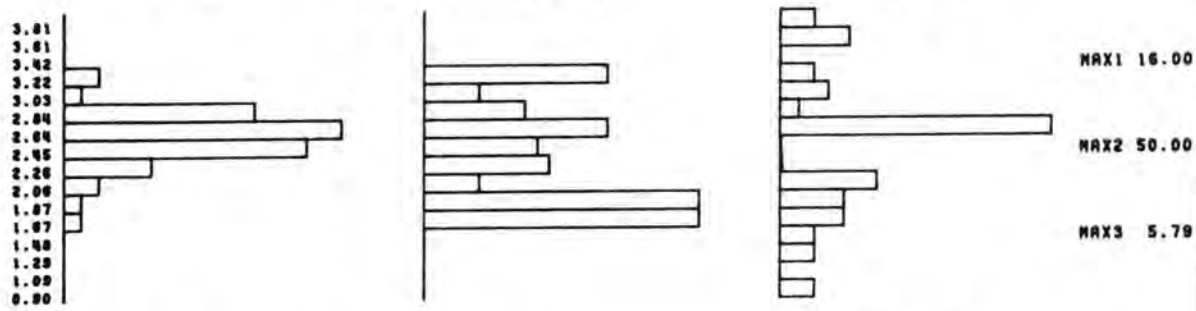
030203 EUGLENA MUTABILIS NS =101



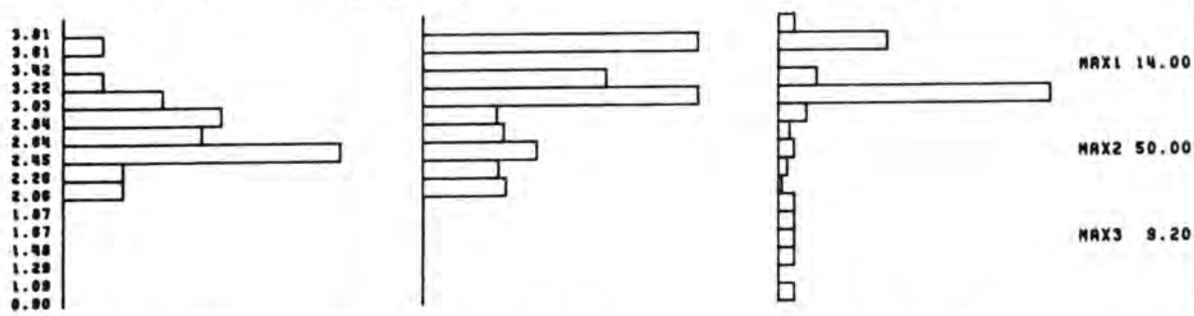
030401 LEPOCINCLIS OVUM NS = 6

040250 CRYPTOMONAS SP. NS = 5

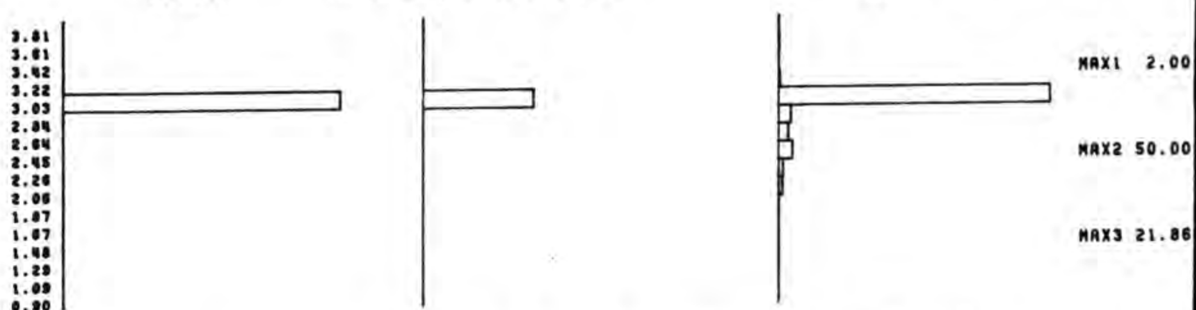
081501 GLOEOCHRYSIS TURFOSA NS = 53



101301 EUNOTIA EXIGUA NS = 44

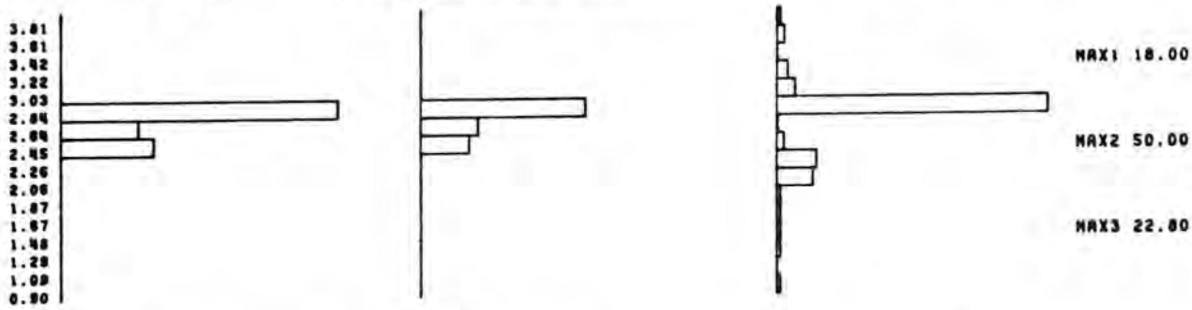


101570 GOMPHONEMA PARVULUM NS = 4

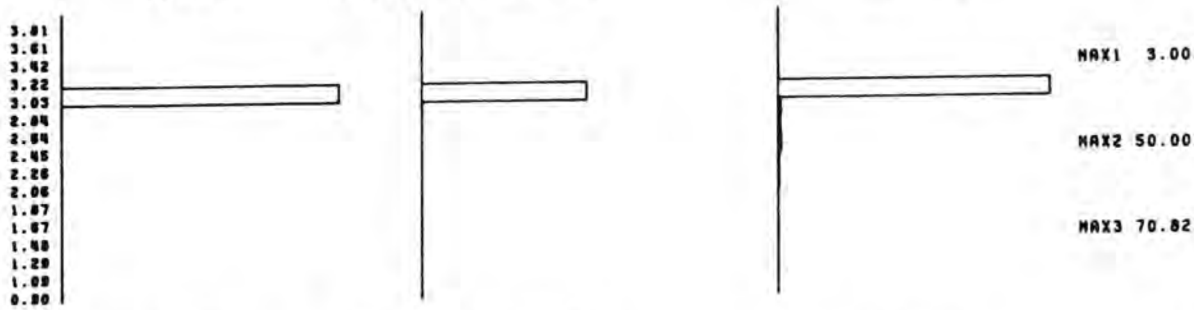


101904 NITZSCHIA PALEA + PALEACEA NS = 3

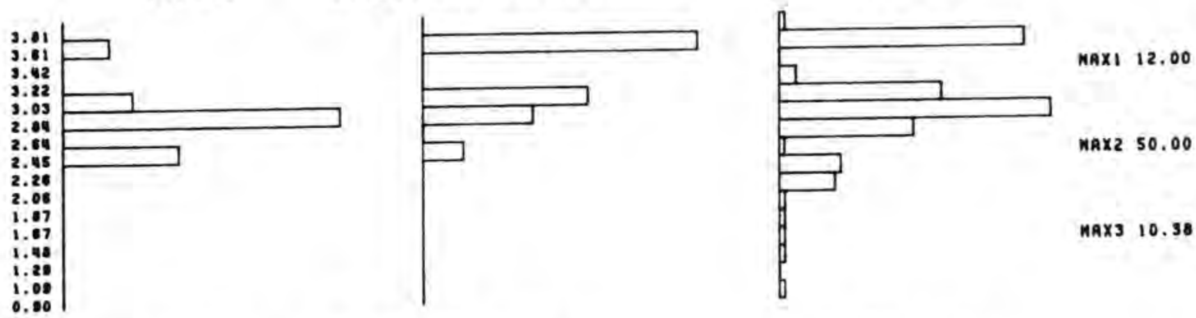
101971 NITZSCHIA SUBCAPITELLATA NS = 29



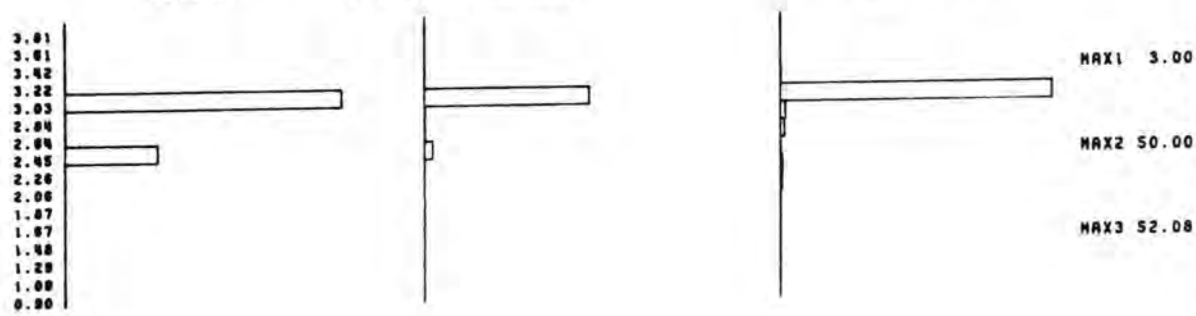
101978 NITZSCHIA UMBONATA NS = 3



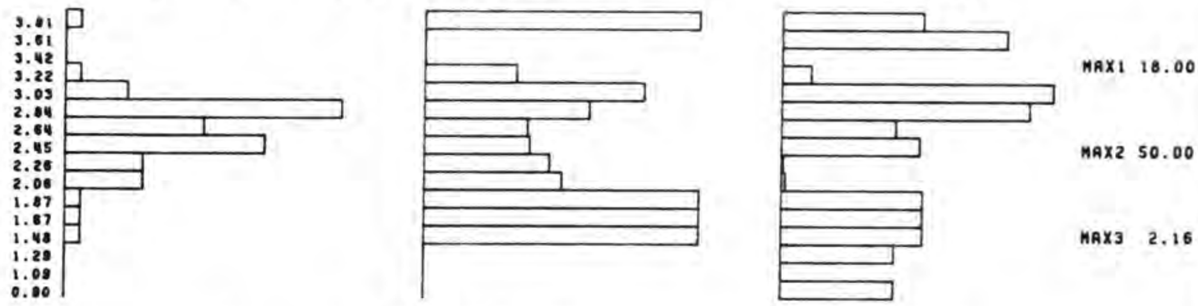
101984 NITZSCHIA PUSILLA NS = 22



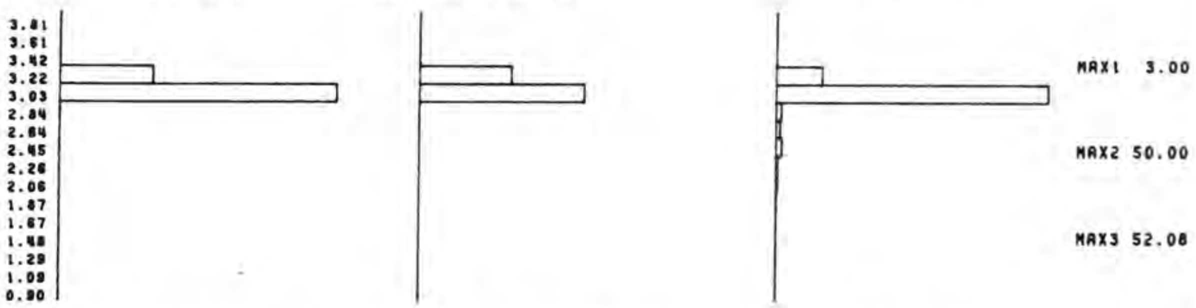
101987 NITZSCHIA OBTUSA NS = 4



102069 PINNULARIA ACORICOLA NS = 59

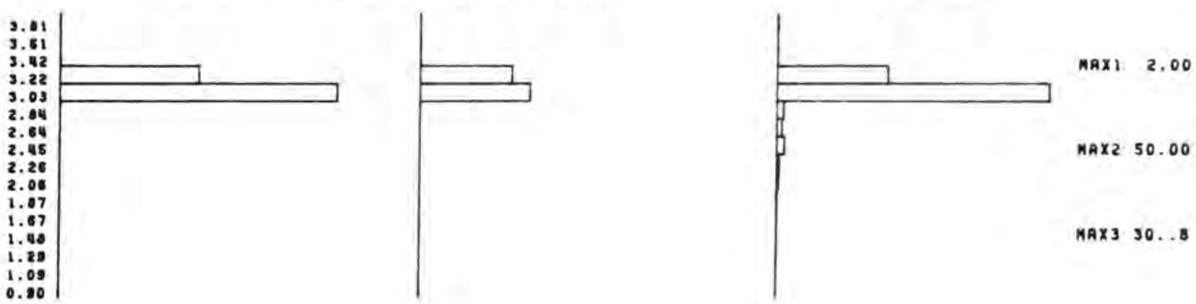


102076 PINNULARIA BRAUNII NS = 4

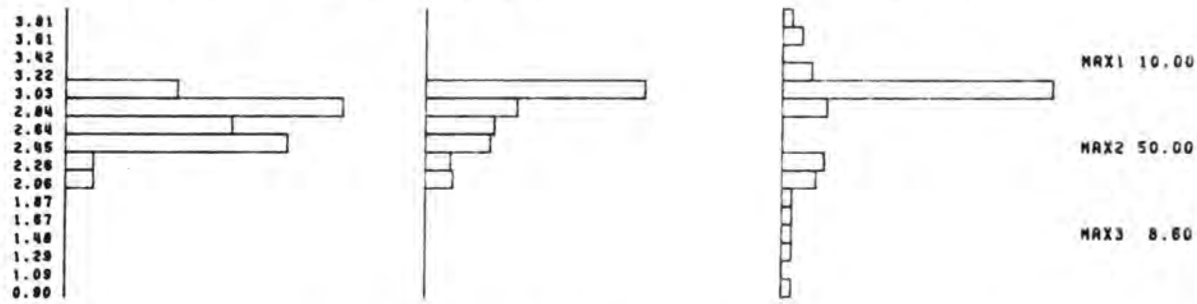


10_078 PINNULARIA MICROSTAURON NS = 3

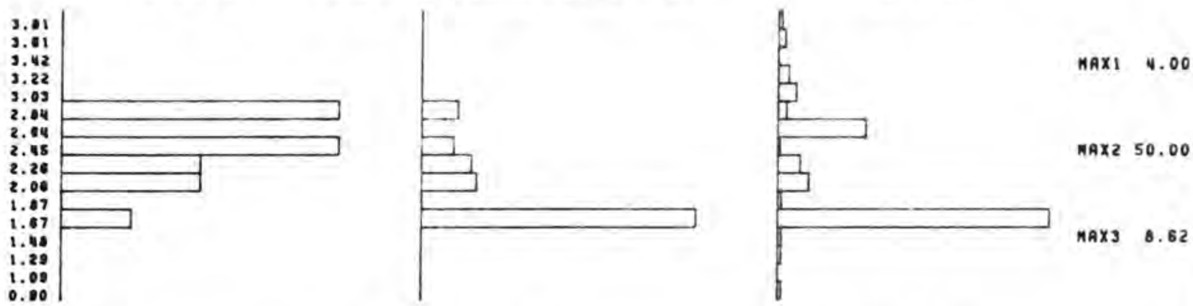
109973 ANOMOEONEIS EXILIS NS = 3



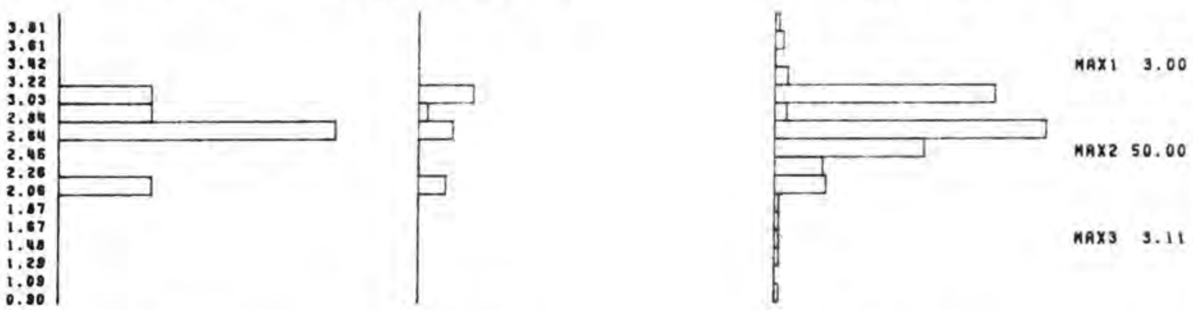
122850 ZYGOGONIUM SP. NS = 30



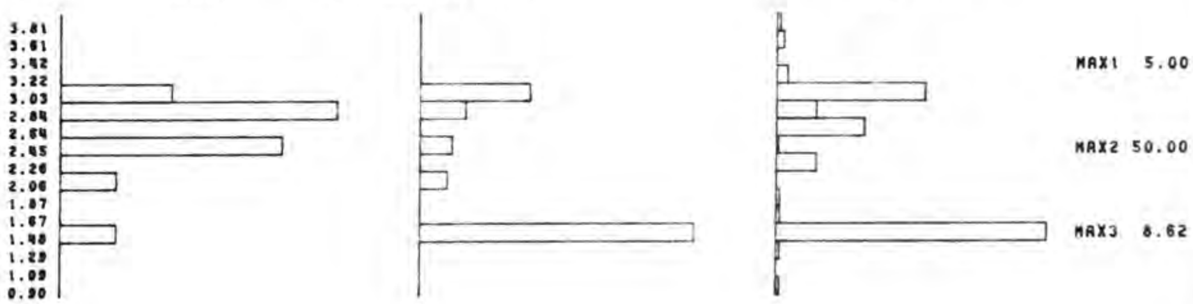
130401 CHLAMYDOMONAS ACIDOPHILA NS = 13



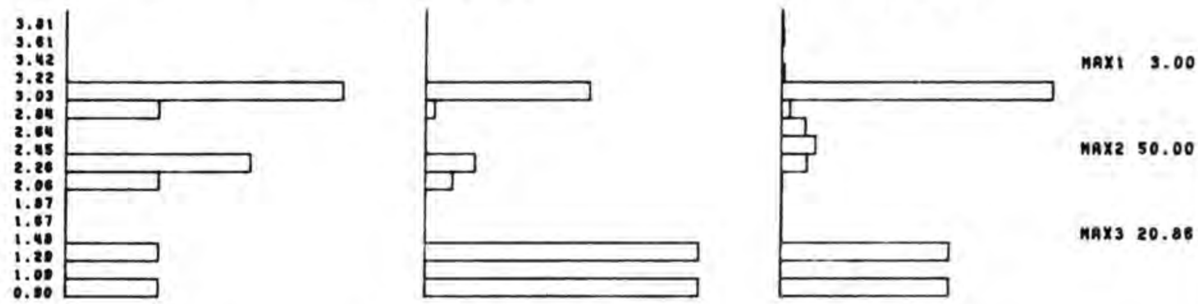
130450 CHLAMYDOMONAS SP. NS = 6



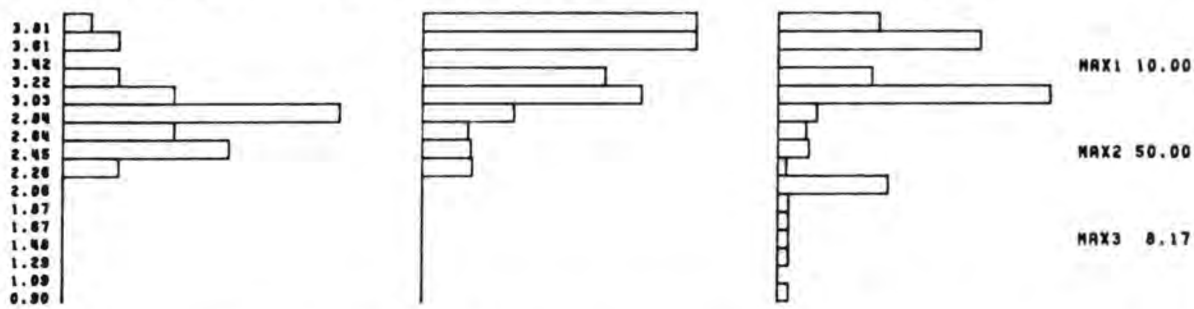
140450 CHARACIUM SP. NS = 13



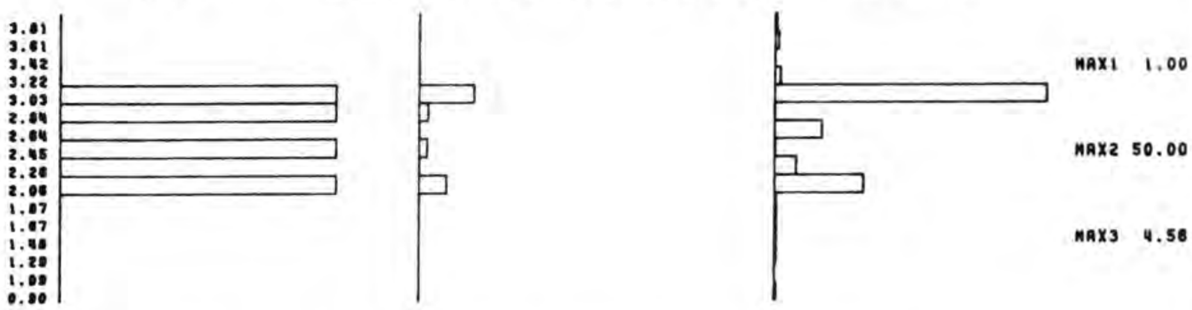
140550 CHLORELLA SP. NS = 9



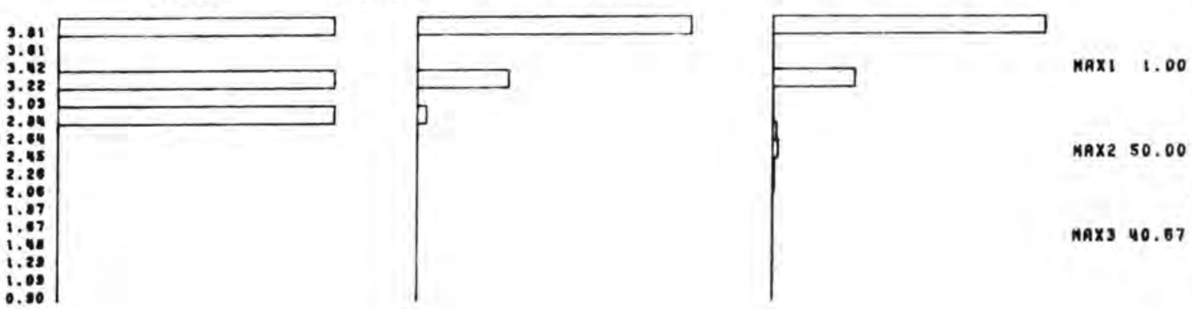
152902 HORMIDIUM RIVULARE NS = 31



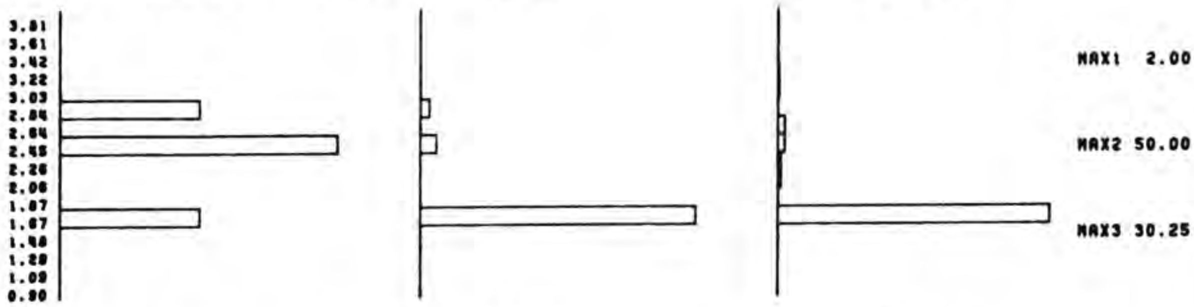
152905 HORMIDIUM PSEUDOSTICHOCOCCUS NS = 4



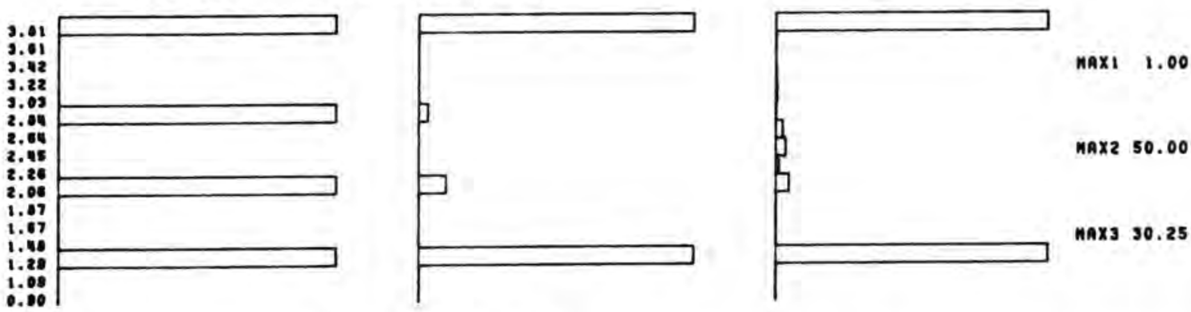
153302 MICROTHAMNION KUTZINGIANUM NS = 3



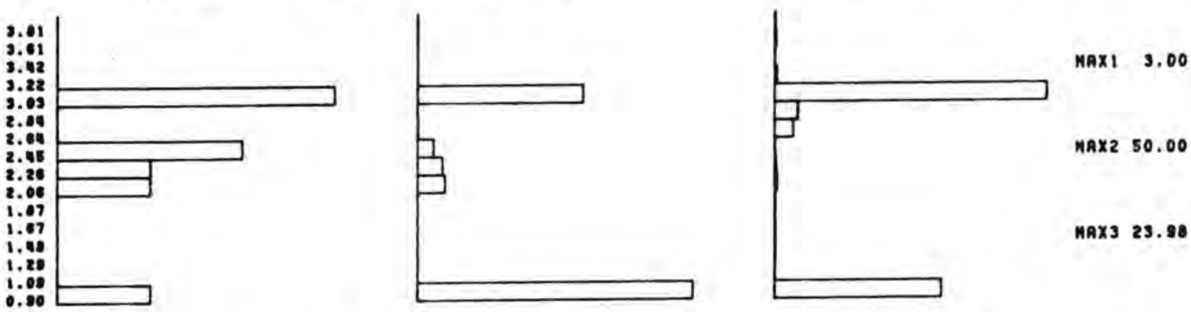
154452 STICHOCOCCUS SP. NS = 4



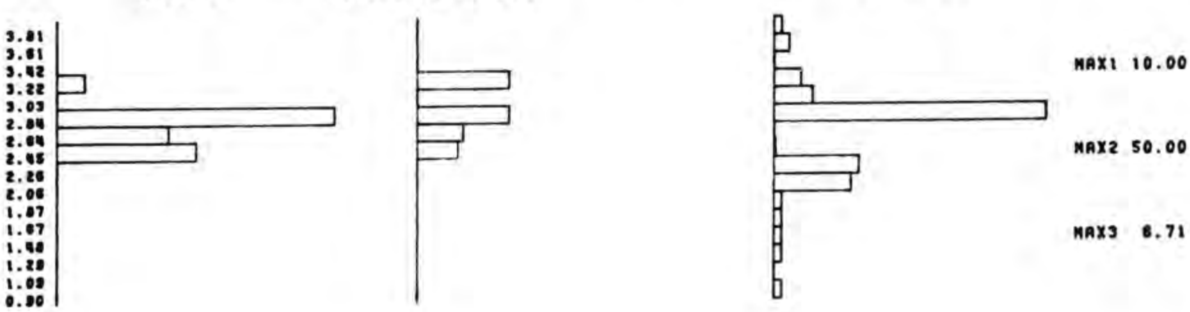
154453 STICHOCOCCUS SP. NS = 4



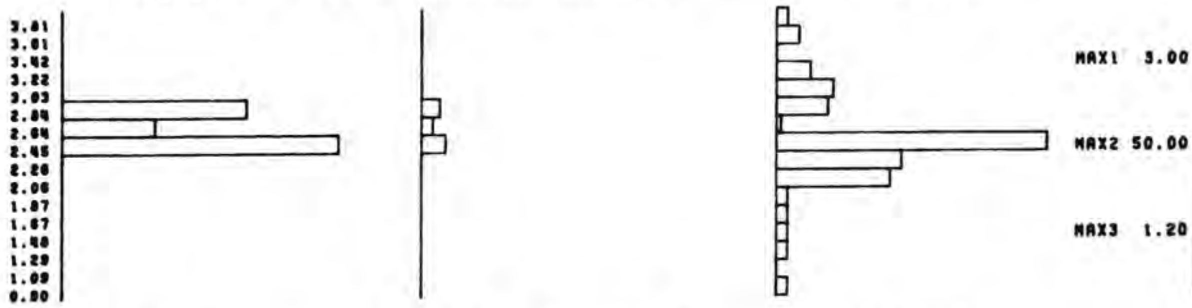
154455 STICHOCOCCUS SP. NS = 8



232150 DICRANELLA SP. NS = 20



232302 DREPANOCLADUS FLUITANS NS = 6



239950 MOSS, GENUS NOT KNOWN SP. NS = 4

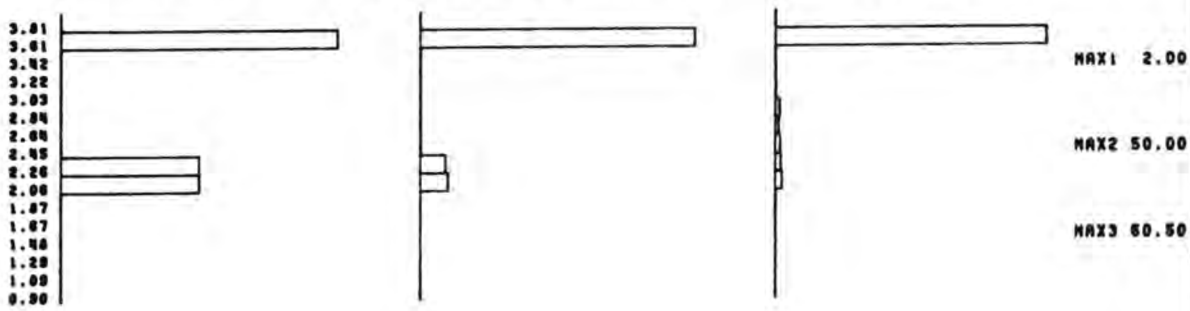
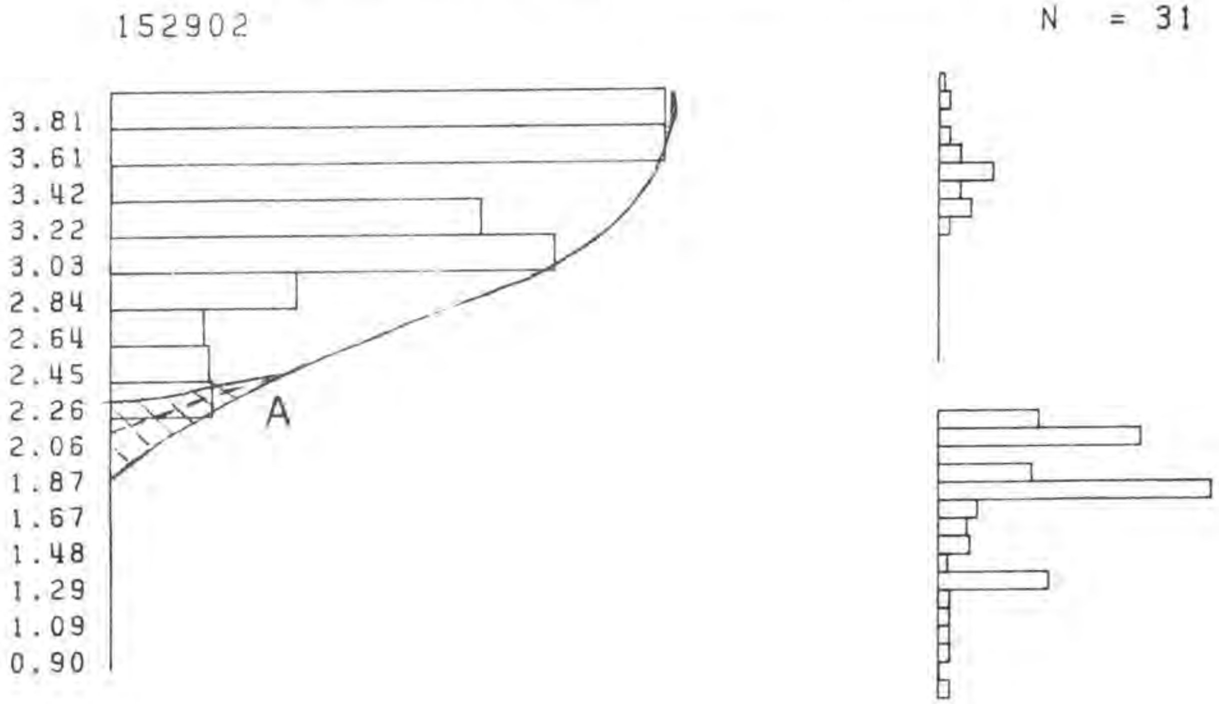


Table 4.52-2

Species reaction type 1 (Hormidium rivulare)



4.52-2) probably best describes the reaction type, if a pH range is expected then the bottom of the range would splay out as shown by the shaded ares. Reasons for a range being the true representation of this reaction type would be amelioration or synergism due to other variables.

Type 2 is characterised by Euglena mutabilis. This shows a sudden fall off at a specified pH level. In reality this fall-off may be a compressed fall off of type 1 which is being masked by a lack of data and the width of the pH bins. Above pH 4.Ø (Section 6.12) the species may or may not be present. This type of distribution is taken to indicate that the species is a dominant in that it is always associated with the pH level of the peak. Further more detailed studies based on the species abundance value confirms that this is usually the case.

Types 3 and 4 may be expressions of the same distribution. Type 3 is characterised by Cyanidium caldarium and may be described as a gentle rise to some peak and then Type 1 fall off. The curve above pH 4.Ø may or may not be present, may or may not rise again but is probably very important. The interpretation, which because of the paucity of data is extremely tentative is that there is some preferred low pH value at which the species is capable of existing. Another related explanation is that the species is truely acidic in nature and can thus be expected to show a slight rise at the tail (low pH end) of its distribution because of the disappearance of other species as the pH

Table 4.52-3

Species reaction type 2 (Euglena mutabilis)

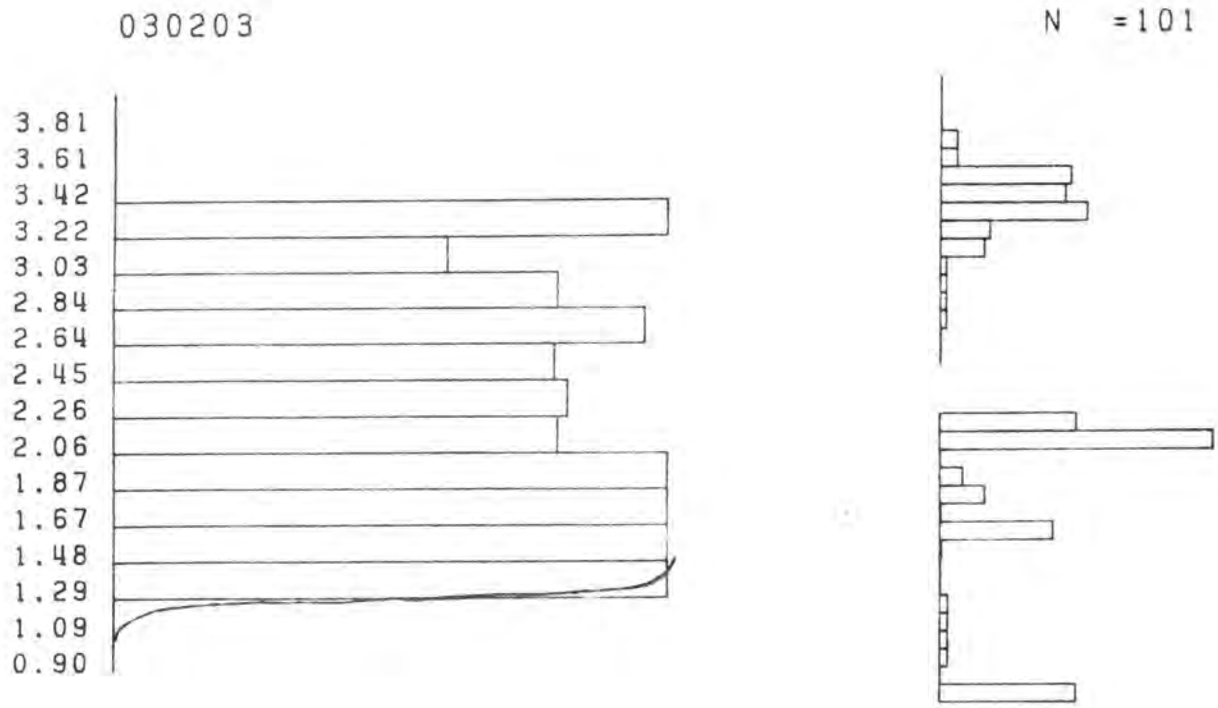
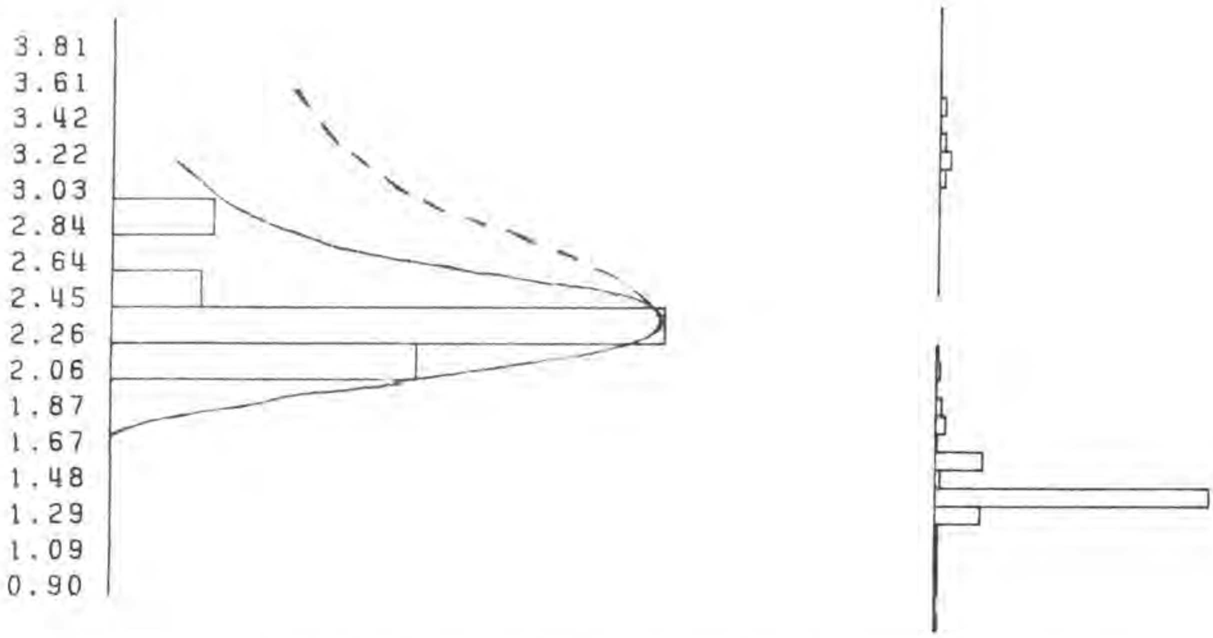


Table 4.52-4

Species reaction type 3 (Cyanidium caldarium)

021301 N = 5

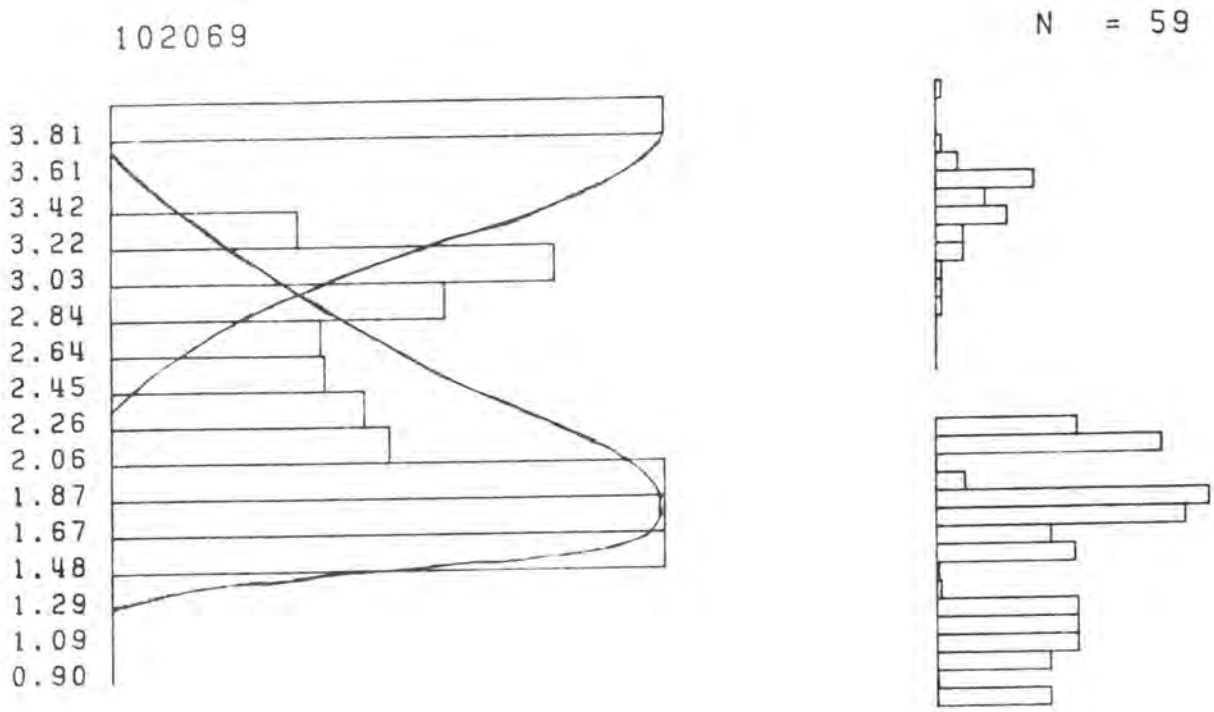


falls. This "tail peaking" in the relative (main) histogram will be seen as a "flat tail" in the absolute (top right) histogram - a feature visible in the tail of the histogram for Euglena mutabilis for example. If, because of the paucity of data at a higher pH only the "flat tail" of low pH data were present then only the "tail peaking" would be expected and reaction type 3 generated.

Reaction type 4 is characterised by Pinnularia acoricola which shows two or more peaks. One peak may be explained by "tail peaking" at the lower end of the pH range. The peak at the upper pH limit suggest that the species is present (abundently - not shown) above pH 4.0 and the third, middle peak may suggest some preferred low pH level. If this analysis is correct then the absolute abundance curve should show two peaks and the flat tail which can be seen to be the case for Pinnularia acoricola. With this sort of reaction type where two pH optima are present an important possible explanation of the distribution is that there is an overlap of two varieties of the same species. The "low pH" variety accounting for the tail peaking and half of the middle peak in the relative histogram and the "normal" variety showing type 1 fall off which overlaps the "low pH" curve at the middle peak. Their combined curves giving rise to the middle peak on addition.

Table 4.52-5

Species reaction type 4 (Pinnularia acoricola)



4.53 Species reactions (specieslist order)

Cyanidium caldarium follows a type 3 distribution. However, with only 5 live occurrences few conclusions may be made. It is noticeable that it peaks within the range 2.26 - 2.45 and does not occur below pH 2.06.

Euglena mutabilis shows the characteristic dominance of the low pH environment in this analysis its adaptation as a true acid species may be recognised by the tail peaking between pH 2.06 and 1.29. Its absence is significant and occurs at both ends of the distribution as characterised here. Its steady fall-off in absolute terms may be a reflection of the absence of data in terms of low pH sites between 2.45 and 1.29.

Lepocinclis ovum fits a Type 3 distribution but too few data are present to allow further comments to be made. It does not occur below a pH of 2.45 in this dataset.

If the cryptomonads are considered together they show a type 3 distribution - disappearing below a pH of 2.45.

Gloeochrysis turfosa shows three peaks when the relative histogram is examined. However, this is not a type 4 reaction because the absolute histogram only shows one peak - and a flat tail which will result in tail peaking. Examination of bin 3.22 - 3.42 suggests that two factors need consideration, the low number of aggregates for this bin and the high fidelity of Gloeochrysis turfosa to this bin - and the lack of fidelity of other species. Given this explanation Gloeochrysis turfosa can be seen to follow a type 2 reaction, with the peak in bin 2.64 - 2.84 and disappearing below pH 1.67.

Eunotia exigua shows steady fall off from pH 4.0 down. This is entirely consistant and suggests a pH limit for the species between 1.87 and 2.06.

All the Nitzschia spp. can be seen to be acid species showing type 3 distributions. However, for all the species the data are scanty.

Pinnularia acoricola shows three peaks in the percentage histogram. The absolute data shows a steep rise to two peaks and a gradual fall off. The peak in the percentage histogram at 2.84 - 3.04 is important when the chi ratio value is examined. If the trough between pH 3.22 and 3.61 is real and not an artefact of scanty data over this pH range then it would be a strong reason to believe that two varieties are interacting or that the records of species found above 3.03 must be treated with caution.

The Pinnularia data suggest that these species follow a type 3 distribution with a gradual rise in numbers to an optimum and then type 1 fall off.

Anomoeoneis exilis shows a type 3 reaction although the data are too sparse to place any confidence on the interpretation.

The distribution of Zygogonium sp. shows typical type 1 distribution.

Analysis of the histograms of Clamydomonas acidophila and Clamydomonas sp. are difficult because of the absence of these species from the pH bins 1.87 - 2.06 and 2.26 - 2.64 respectively and the general absence of the species above a pH of 3.04. However, a tentative Type 3 distribution may be proposed for these species.

Too few data exist to allow an analysis of Characium sp. however a type 3 reaction type is probably appropriate.

The higher pH Chlorellas (those found above pH 4.0) and the lack of data (below pH 4.0) mask what could be a type 3 reaction.

Hormidium rivulare demonstrates type 1 distribution with gradual fall off. The minimum pH appearing to lie somewhere in the 2.06 - 2.26 bin.

Dicranella sp. While this is not an abundant species, the histograms are entirely consistant with a type 1 pH reaction with a minimum pH in the bin 2.26 - 2.45.

Too few data exist to allow comment on the reaction types of Hormidium pseudostichococcus and Stichococcus sp. However, the data available are not inconsistant with a type 3 reaction to pH.

Drepanocladus fluitans appears to follow a typical type 3 reaction to pH however, the data are insufficient for more detailed comments.

Comparisons of these reaction types with the data in Table 6.11-1 suggests that reaction type 3 coincides with those species with a curtailed distribution restricted to pH values below 4.0 and that reaction type 1 is true for the majority of the more ubiquitous species as identified in the table. Reaction types 2 and 4 may be exhibited by species belonging to either of these two main reaction groups.

Chapter 5 Acid sites

5.1 Site list production

5.11 Introduction

Chapter 5 contains the results of an analysis of the sites defined as being acidic. In 5.1 these sites will be presented in geographical terms using the information contained in SIEUR. The results of the various analyses performed on the data associated with these sites will be presented in 5.2.

The acid site dataset was produced by performing a "bysar" expansion of the 269 acid water chemistries found on issuing the basic query pH-fld <= 4.Ø. The resulting water chemistries have been used to describe the chemistry of the reaches in more detail and to give some idea of the range of chemistries it is possible to obtain even with as severe a restriction as that imposed by the basic query.

5.12 Geographical notes

Table 5.12-1 summarises the geographical data to be found in SIEUR concerning the acid sites.

Table 5.12-1 (pp 158-159)

Geographical details present in SIEUR (June 1980) associated with the acid sites

Table 5.12-1

United Kingdom (0nnn)

```
0022 Red burn Wear basin
     01  NZ 999999, 99.99 99.99

0066 "Lowlands acid site" Wear basin
     01  NZ 133252, 54.37 01.48

0100 Dowgang burn Tyne basin
     30  NY 780435, 54.47 02.20

0109 Dowgang burn adit Tyne basin
     01  NY 776424, 54.46 02.21

0119 "Rake sike A" Wear basin
     01  NY 808435, 54.47 02.18

0120 "Rake sike B" Wear basin
     01  NY 808434, 54.47 02.18

0125 "Brandon Pithouse acid stream B" Wear basin
     01  NZ 215404, 54.45 01.39
     01  NZ 212404, 54.45 01.40

0127 "Brandon Pithouse acid stream A" Wear basin
     02  NZ 214406, 54.45 01.40
     03  NZ 215406, 54.45 01.40
     04  NZ 215406, 54.45 01.39
     05  NZ 216406, 54.45 01.40
     06  NZ 215407, 54.45 01.40
     07  NZ 215407, 54.45 01.40
     08  NZ 216408, 54.45 01.40
     09  NZ 217409, 54.45 01.40
     10  NZ 217410, 54.45 01.40
     11  NZ 216411, 54.45 01.40
     12  NZ 218415, 54.45 01.40
     13  NZ 218416, 54.46 01.40

0128 "Walkmill acid stream" Ehen basin
     01  NY 007188, 54.33 03.31
     02  NY 007187, 54.33 03.31
     03  NY 006187, 54.33 03.31

0129 "Oatlands acid stream" Ehen basin
     01  NY 025216, 54.34 03.30
     02  NY 025215, 54.34 03.30

0130 "Rowley acid stream A" Ribble basin
     01  SD 861333, 53.47 02.12
     02  SD 860334, 53.47 02.12
     03  SD 859335, 53.47 02.12
     04  SD 857337, 53.47 02.12

0131 "Deerplay acid stream" Humber basin
     01  SD 869267, 53.44 02.10
     02  SD 868266, 53.44 02.11
     03  SD 867264, 53.44 02.12

0132 "Chisnall Hall acid stream A" Ribble basin
     01  SD 553126, 53.36 02.40
     02  SD 553128, 53.36 02.40
     03  SD 552126, 53.36 02.40
     04  SD 546127, 53.36 02.41
     05  SD 542127, 53.36 02.41
     06  SD 545127, 53.36 02.41
     07  SD 548125, 53.36 02.41

0133 "Welsh Whittle acid stream" Ribble basin
     01  SD 545135, 53.37 02.41

0134 "Gibfield acid stream" Mersey basin
     01  SD 661023, 53.30 02.30
     02  SD 660022, 53.30 02.31
     03  SD 659022, 53.30 02.30
     04  SD 659023, 53.30 02.30

0135 "Denby acid stream A" Trent basin
     01  SD 392483, 53.55 01.25
     02  SD 393484, 53.55 01.25
     03  SD 394483, 53.55 01.25
     04  SD 395483, 53.55 01.25

0136 "Cannock open caste acid stream" Trent basin
     01  SJ 990083, 52.40 02.00

0137 "Polesworth acid stream" Trent basin
     01  SK 257038, 52.38 02.42
     02  SK 255039, 52.38 02.42
     03  SK 253042, 52.38 02.42

0138 "Bridford acid stream" Teign basin
     01  SX 816854, 50.41 03.39
     02  SX 817855, 50.41 03.39
     03  SX 817856, 50.41 03.39

0139 "Birch Coppice acid stream" Trent basin
     01  SP 255001, 52.36 02.38
     02  SP 253001, 52.36 02.38
     03  SP 254001, 52.36 02.38

0140 "Kingsbury acid stream" Trent basin
     01  SP 233986, 52.35 01.38
     02  SP 232985, 52.35 01.38
     03  SP 232984, 52.35 01.38

0152 "Westfield: stream south of main slurry lagoon" bbbbbb basin
     10  ZZ 999999, 99.99 99.99
     20  ZZ 999999, 99.99 99.99
     01  ZZ 999999, 99.99 99.99

0154 "Westfield: l.h. flush S.E. of main slurry lagoon" bbbbbb basin
     01  ZZ 999999, 99.99 99.99
```

0157 "Rowley acid stream B" Ribble basin
05 SD 858331, 53.47 02.12
06 SD 857332, 53.47 02.12
07 SD 856334, 53.47 02.13

0158 "Chisnall Hall acid stream B" Ribble basin
08 SD 549125, 53.36 02.40

0159 "Chisnall Hall acid stream c" Ribble basin
09 SD 549126, 53.36 02.40

0160 "Denby acid stream B" Trent basin
05 SD 392485, 53.55 01.25
06 SD 392485, 53.55 01.25

0213 "Parys iron flush" Afon Goch "East" basin
01 SH 443899, 53.22 04.20

0214 "Parys lagoons B" Afon Goch "East" basin
21 SH 443899, 53.22 04.20

0215 "Parys lagoons A" Afon Goch "East" basin
01 SH 448905, 53.22 04.20
15 SH 449905, 53.22 04.20

0216 Afon Goch "East" Afon Goch "East" basin
15 SH 454898, 53.22 04.19
50 SH 886446, 53.21 15.20

0217 Dyffryn Adda adit Afon Goch "north" basin
01 SH 437913, 53.23 04.21

0218 "Parys south west adit" Afon Goch "East" basin
01 SH 438897, 53.22 04.21

0220 "Adeer factory effluent 8" Garnock basin
03 NS 298396, 55.37 04.42
15 NS 298396, 55.37 04.42

0221 "Adeer factory effluent 9" Garnock basin
05 NS 298396, 55.37 04.42

Belgium (1nnn)

1004 Bois de cornillon stream A Maas basin
01 2550/56092, 50.35 5.32

Ireland (2nnn)

2003 "Tigroney 850 A" Avoca basin
01 32031826, 52.52 06.12

2004 "Tigroney 850 B" Avoca basin
01 32031826, 52.52 06.12

2006 Tigroney grass level Avoca basin
01 32041831, 52.52 06.12

2007 Tigroney barrack level Avoca basin
01 32041831, 52.52 06.12
02 32041831, 52.52 06.12
03 32041831, 52.52 06.12

2014 "Tigroney 850 c" Avoca basin
01 32031826, 52.52 06.12

2015 "Red road stream" Avoca basin
10 31961808, 52.51 06.13

France (3nnn)

3016 "Soussu seepage stream" Adour basin
01 4023-750, 43.40 01.00
03 4023-750, 43.40 01.00

United States of America (9nnn)

9001 Tributary of Burra Burra creek bbbbbb basin
01 99999/99999, 99.99 99.99

9004 "Shinnston big elm trickle" bbbbbb basin
01 99999/99999, 99.99 99.99

9005 "Lumberport bridge road trickle" bbbbbb basin
01 99999/99999, 99.99 99.99

9006 "Lambert's run watershed" bbbbbb basin
01 99999/99999, 99.99 99.99

9007 "East Morgantown stream" bbbbbb basin
10 99999/99999, 99.99 99.99

9008 Lemonade creek bbbbbb basin
45 99999/99999, 44.48 110.43
48 99999/99999, 44.48 110.43
50 99999/99999, 44.48 110.43
52 99999/99999, 44.48 110.43

9009 Tantalus creek bbbbbb basin
20 99999/99999, 44.44 110.42

9010 Root pool outflow bbbbbb basin
02 99999/99999, 44.44 110.42

9011 "Tributary of near Monarch Geyser stream B" bbbbbb basin
01 99999/99999, 44.44 110.42

9012 "Near Monarch Geyser stream A" bbbbbb basin
01 99999/99999, 44.44 110.42
03 99999/99999, 44.44 110.42

9016 West fork r.h. tributary bbbbbb basin
99 99999/99999, 99.99 99.99

5.2 Chemistry summaries and exceptional sites

5.21 Introduction

In this section the acid sites obtained by the procedures described in Section 4.15 are outlined in terms of their chemistries using the total data present in SIEUR. Sites for which there is only one water chemistry have been ignored.

5.22 Acid site data chemistry summaries

The common logarithms of the maximum, minimum and mean for all the variables which have been measured at any site more than twice is presented as the sequence of diagrams labelled Table 5.22-1. The variable substitution technique described in Section 2.22 has been employed in generating the data and values found at the detection limit have been ignored. The sites have been sorted into decreasing order of range maxima except for pH which is ordered by increasing range minima.

In general the diagrams for the cations show wide ranges for sites placed at the two ends of the diagrams, with the sites at the lower end (to the right in all diagrams except that for pH) having the widest range (e.g. K). In contrast the sites show a constant wide range for anions, with sites sampled more frequently having the largest ranges (e.g. PO_4-P and SO_4-S). The wide range for anions at sites have been attributed by Hargreaves to seasonal flow variations, so much so that the sites in summer may be considered eutrophic (Hargreaves, 1977). The diagrams for conductivity, pH and acidity follow the pattern for cations.

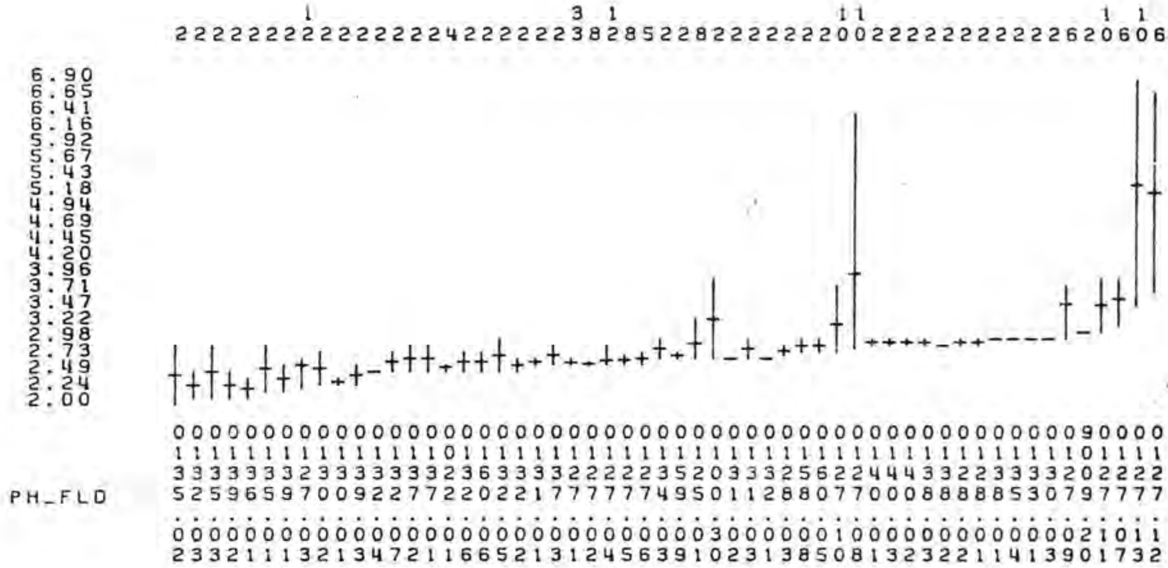
The diagram for conductivity suggests that 6 sites (0139 Birch Coppice, 0136.01 Cannock, 0160.05 & 0160.06 Denby, 0127.01 Brandon and to a lesser extent 0130.01 & 0130.03 Rowley) have large range variations. Other reaches at these sites show no such wide variations, suggesting that these data may be anomalous. The high mean value for the Brandon 01 site indicates the negligible effect of the one low conductivity value, which is almost certainly an error in the data since Hargreaves does not comment on it and it plays no part in his Table 5.1 entry for minimum conductivity which is quoted as being 1350 although the minimum value to be seen in the table is the value found in SIEUR namely 700 (pconductivity 2.8). The only site which cannot be checked in these ways is Denby Stream B (0160), the high values here are associated

Table 5.22-1 (pp 161-168)

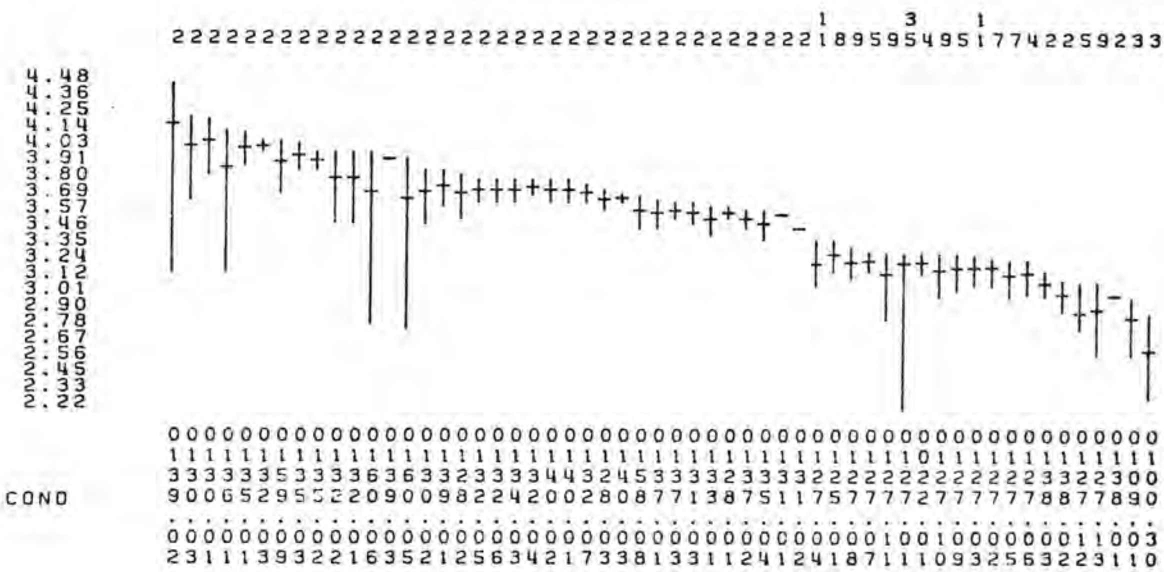
Chemical and physical parameters drawn from SIEUR (June 1980) irrespective of pH for the acid sites

Notes:

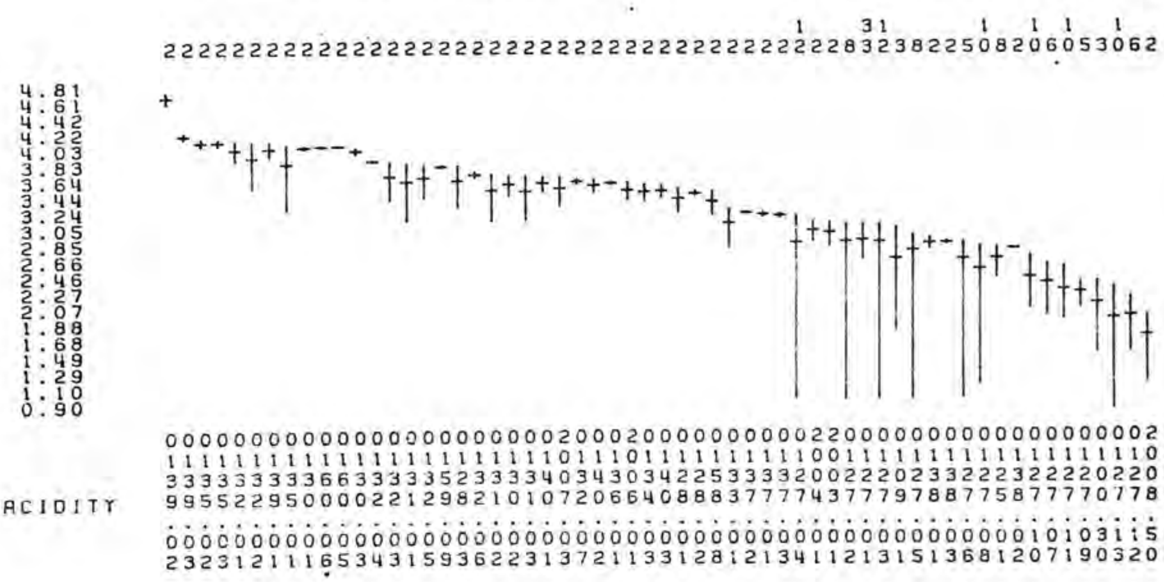
1. all scales are $\log_{10}x$ where x is in the units of Table 4.13-1 p. 93.
2. number of records are at the top.
3. ssss.rr stream and reach numbers are at the bottom.



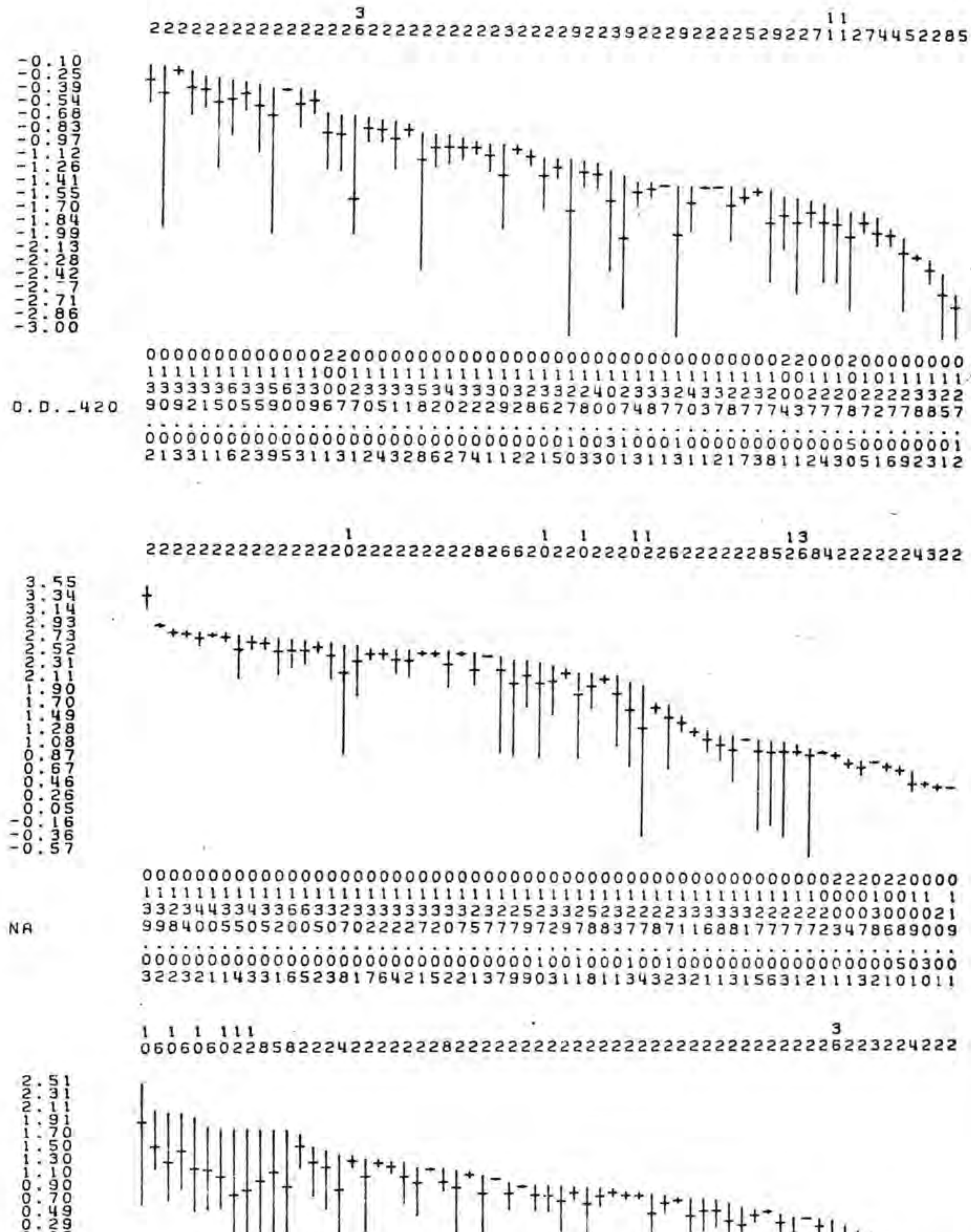
1 3 1 1 1 1 1
22222222222222224222222382852282222222002222222222222620606
6.90
6.65
6.41
6.16
5.92
5.67
5.43
5.18
4.94
4.69
4.45
4.20
3.96
3.71
3.47
3.22
2.98
2.73
2.49
2.24
2.00
000000000000000000000000000000000000000000000000000000090000
111111111111111101111111111111111111111111111111111111101111
333333323333333236333322222352033325622444332233333202222
525965970092277220221777777495011288077000888885307977777
PH_FLD
000000000000000000000000000003000000010000000000000021011
23321113213472116652131245639102313850813232211413901732

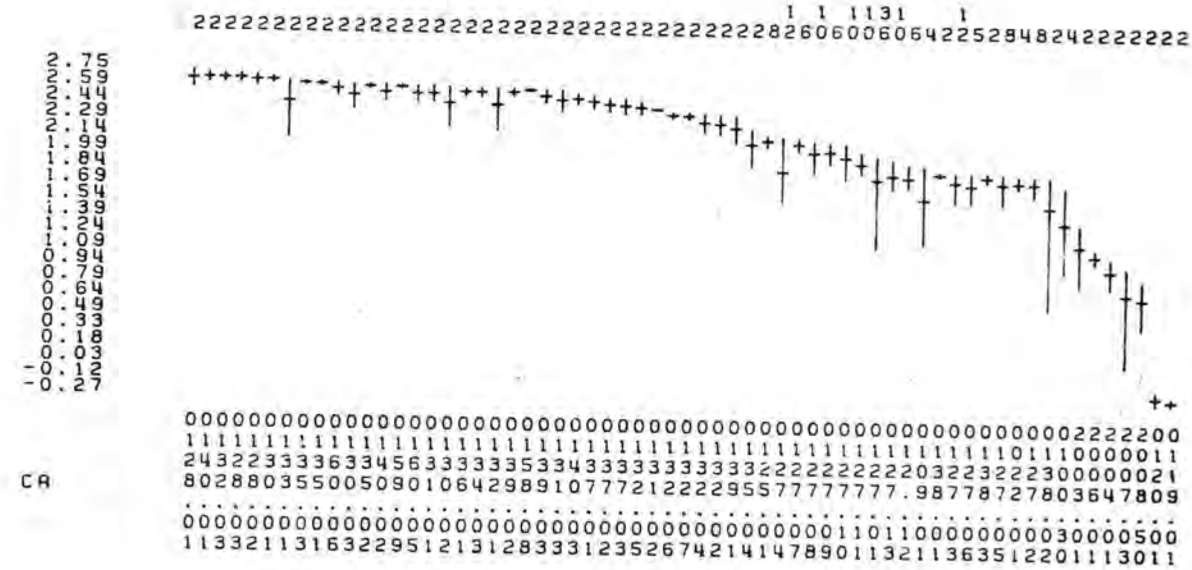


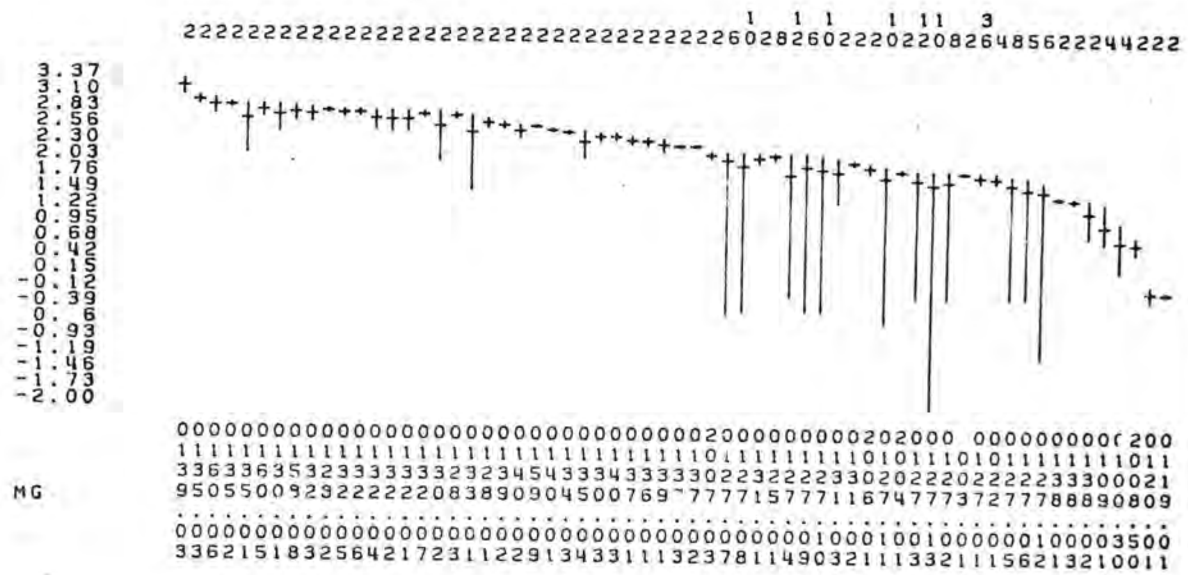
1 3 1
2222222222222222222222222222222222221895954951774225923 3
4.48
4.36
4.25
4.14
4.03
3.91
3.80
3.69
3.57
3.46
3.35
3.24
3.12
3.01
2.90
2.78
2.67
2.56
2.45
2.33
2.22
0000000000000000000000000000000000000000000000000000000000
1111111111111111111111111111111111111111111111011111111111111
33333353333363633232333443245333323333222222222222223322300
9006529552209009822420028087713875117577772777777788778 90
COND
000000000000000000000000000000000000000001001000000011003
2311139322163521256342173381331124124187111093256322 3110

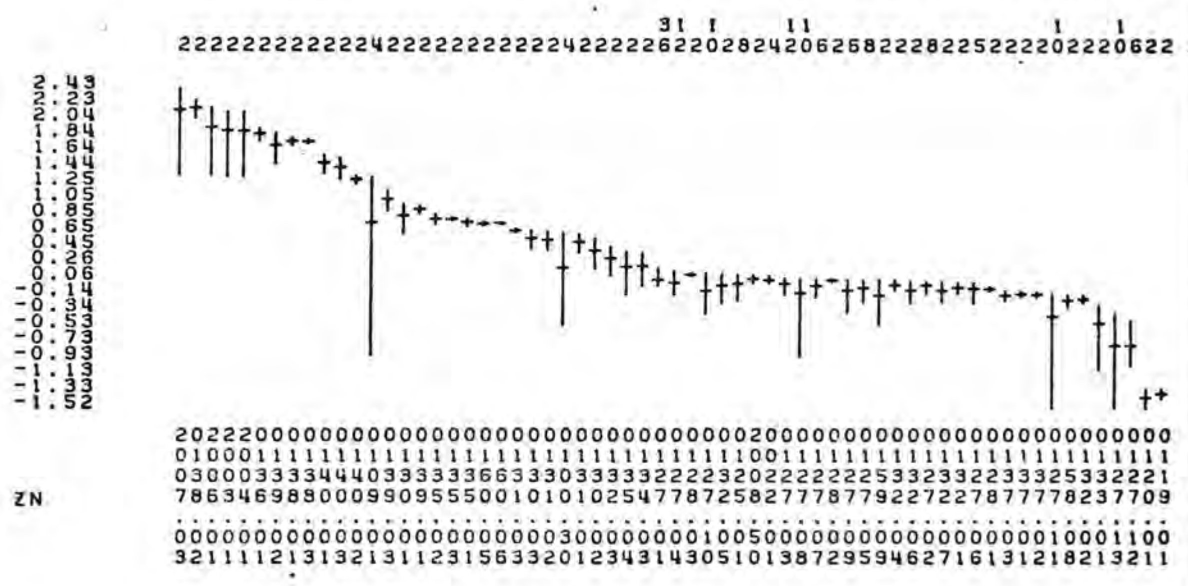


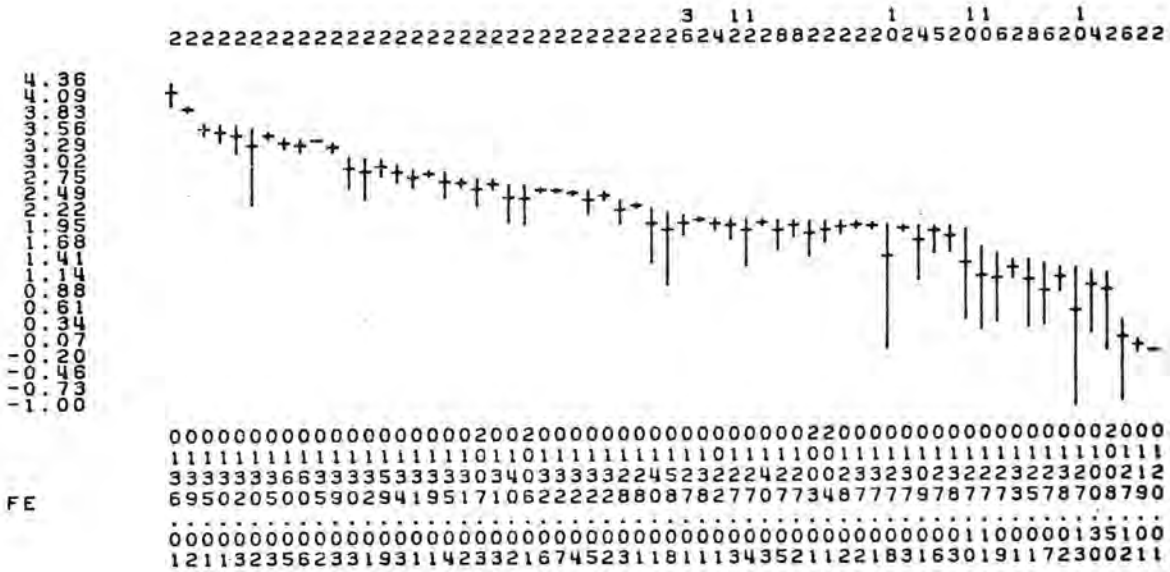
1 31 1 1 1 1
222222222222222222222222222222222222283238225082060530 62
4.81
4.61
4.42
4.22
4.03
3.83
3.64
3.44
3.24
3.05
2.85
2.66
2.46
2.27
2.07
1.88
1.68
1.49
1.29
1.10
0.90
0000000000000000000000020002000000000000220000000000000000002
1111111111111111111111111011101111111111100111111111111111110
33333333366333335233334034303422533332002220233222322220220
99552295000022129821010720664088837777437779788775877770778
ACIDITY
000000000000000000000000000000000000000000000000000010103115
2323121116534315936223137211331281213411213151368120719032 0


O.D._420
NA
K

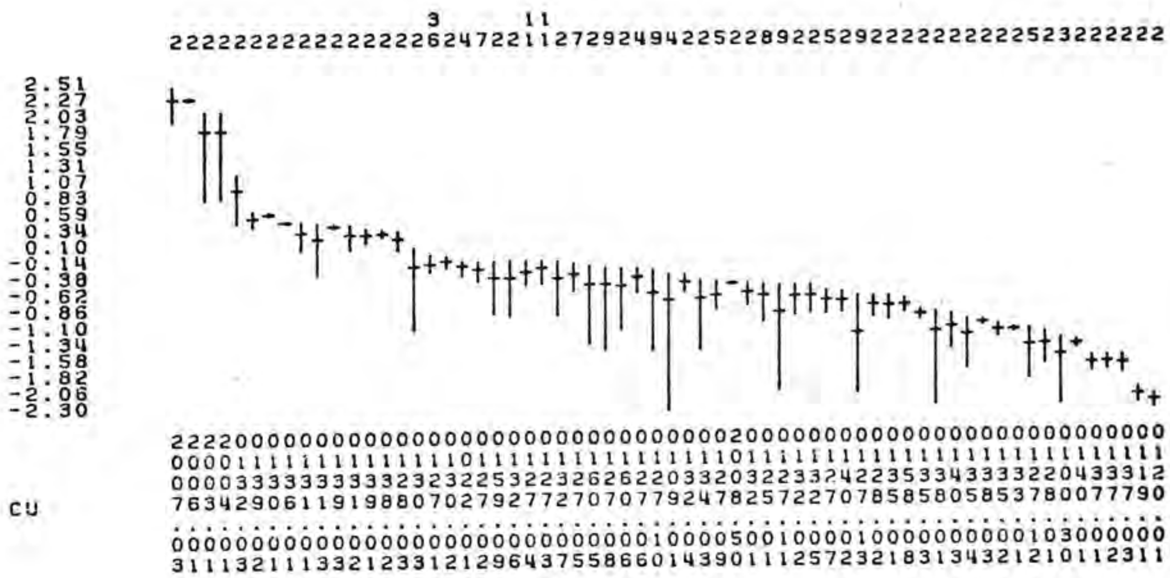

CA


MG

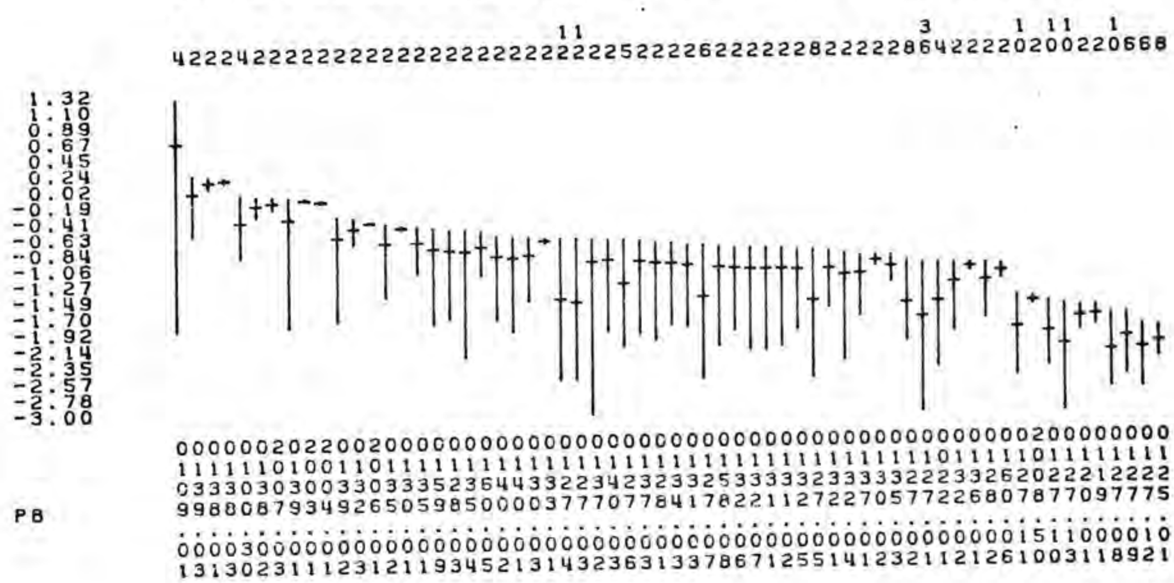

ZN


4.36
4.09
3.83
3.56
3.29
3.02
2.75
2.49
2.22
1.95
1.68
1.41
1.14
0.88
0.61
0.34
0.07
-0.20
-0.46
-0.73
-1.00

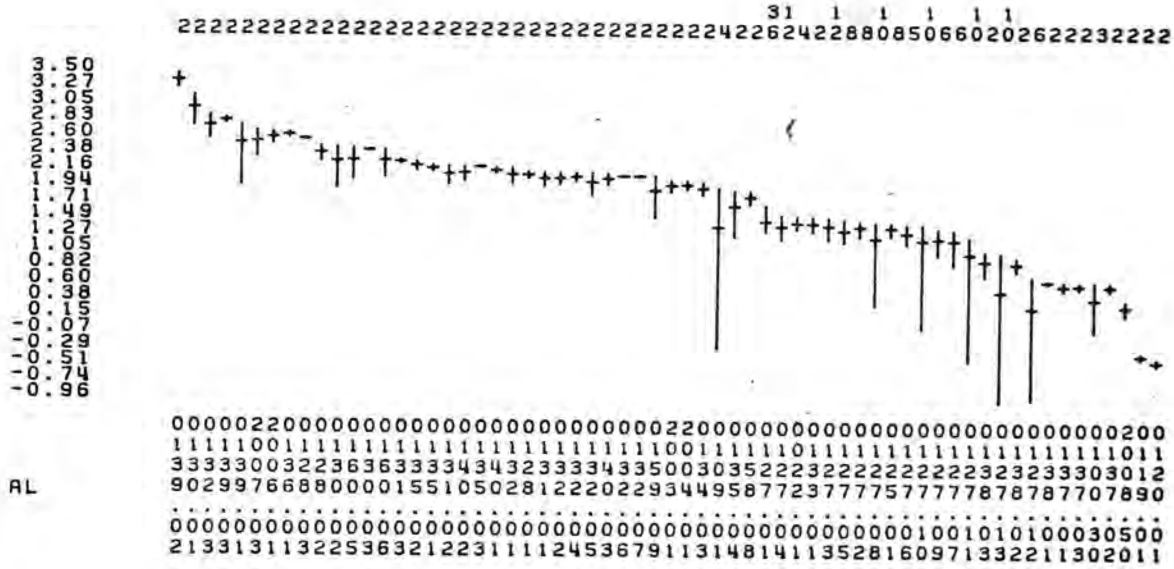
FE

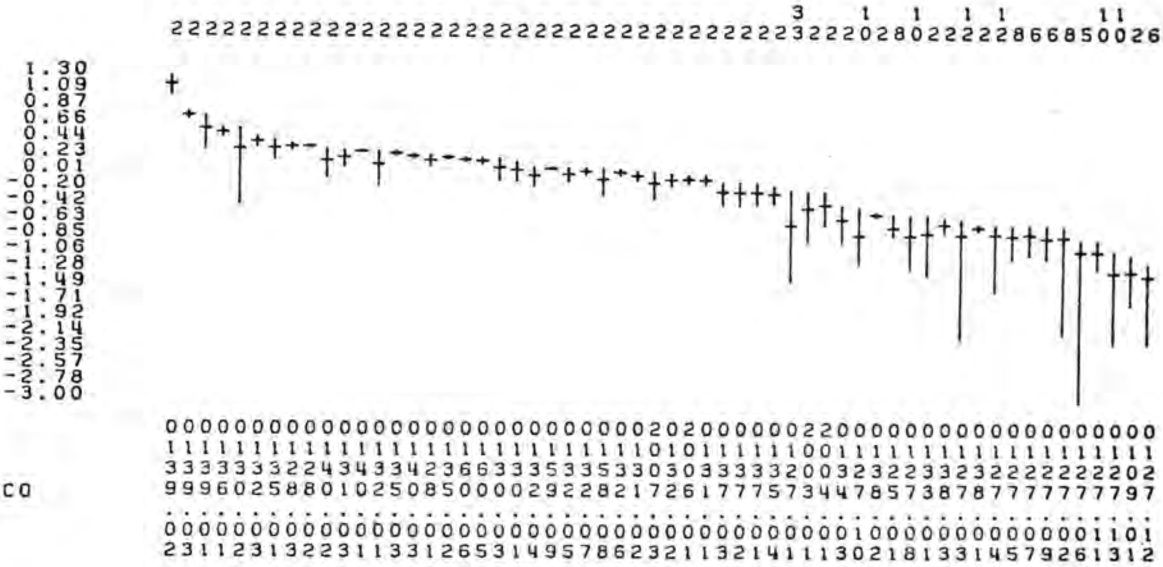

2.51
2.27
2.03
1.79
1.55
1.31
1.07
0.83
0.59
0.34
0.10
-0.14
-0.38
-0.62
-0.86
-1.10
-1.34
-1.58
-1.82
-2.06
-2.30

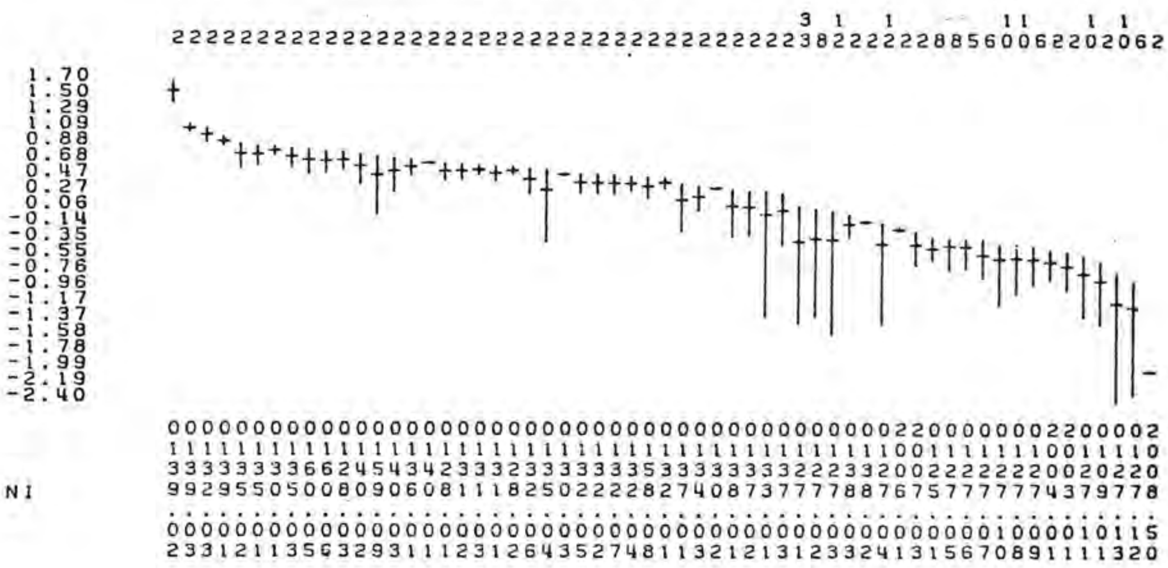
CU

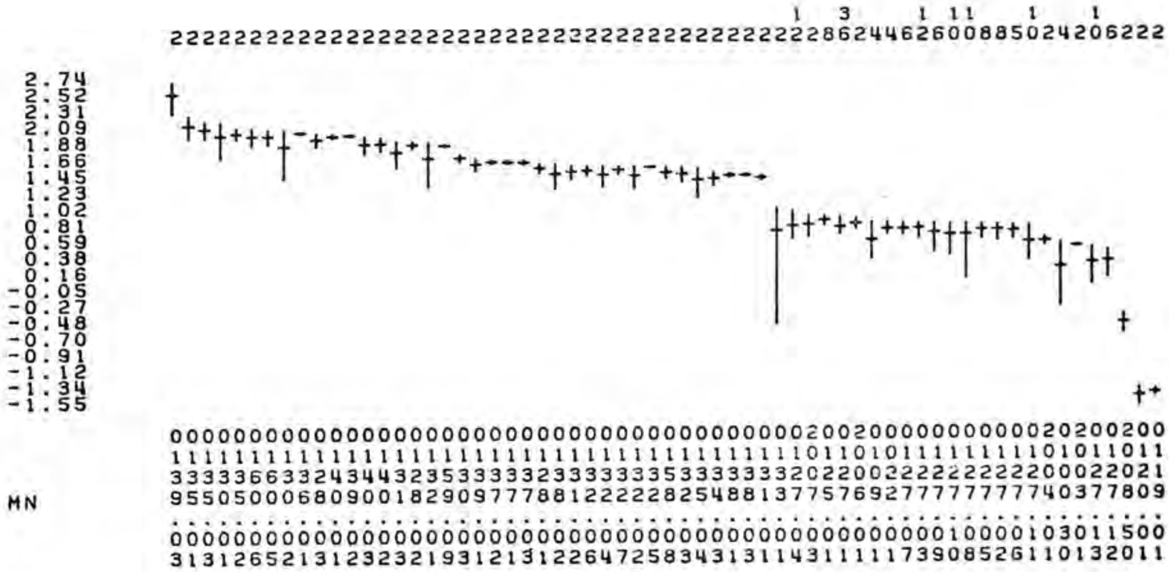

1.32
1.10
0.89
0.67
0.45
0.24
0.02
-0.19
-0.41
-0.63
-0.84
-1.06
-1.27
-1.49
-1.70
-1.92
-2.14
-2.35
-2.57
-2.78
-3.00

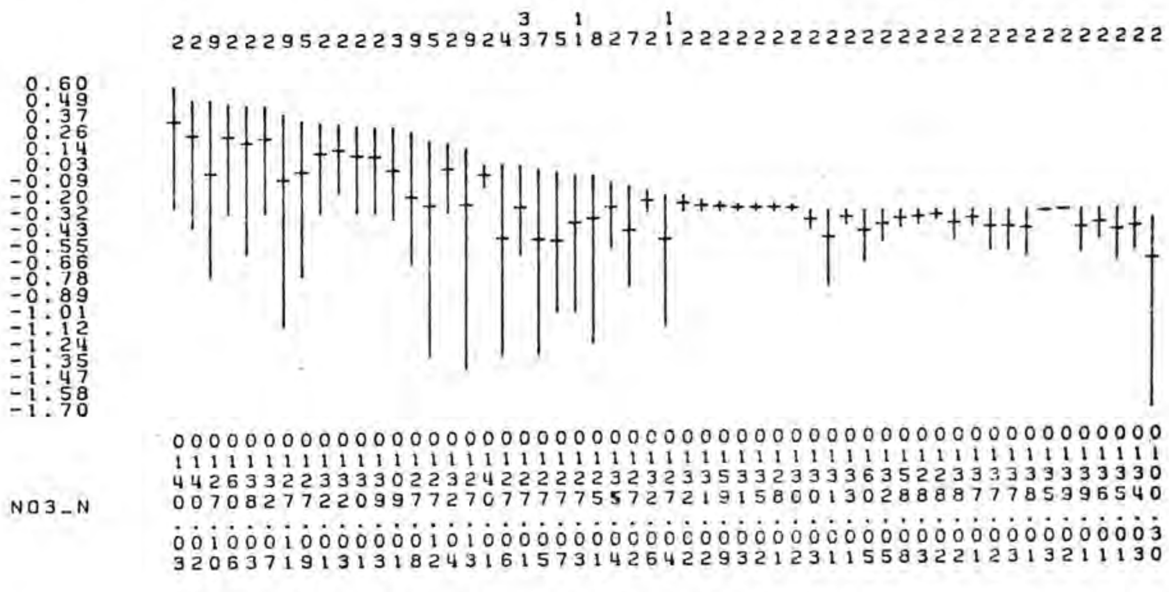
PB

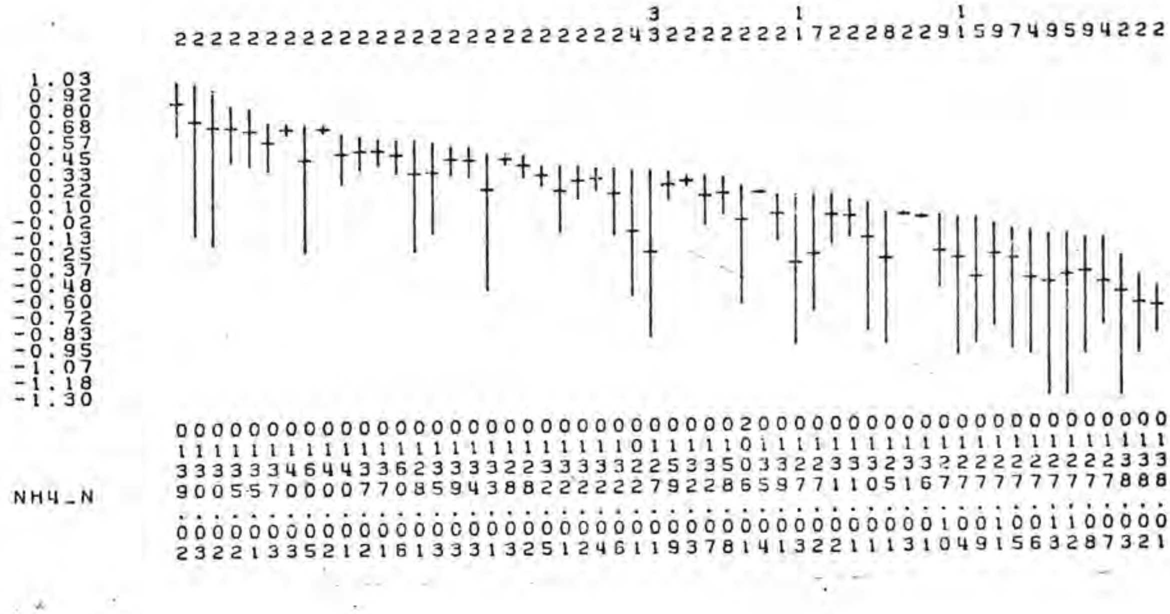

AL
3.50
3.27
3.05
2.83
2.60
2.38
2.16
1.94
1.71
1.49
1.27
1.05
0.82
0.60
0.38
0.15
-0.07
-0.29
-0.51
-0.74
-0.96

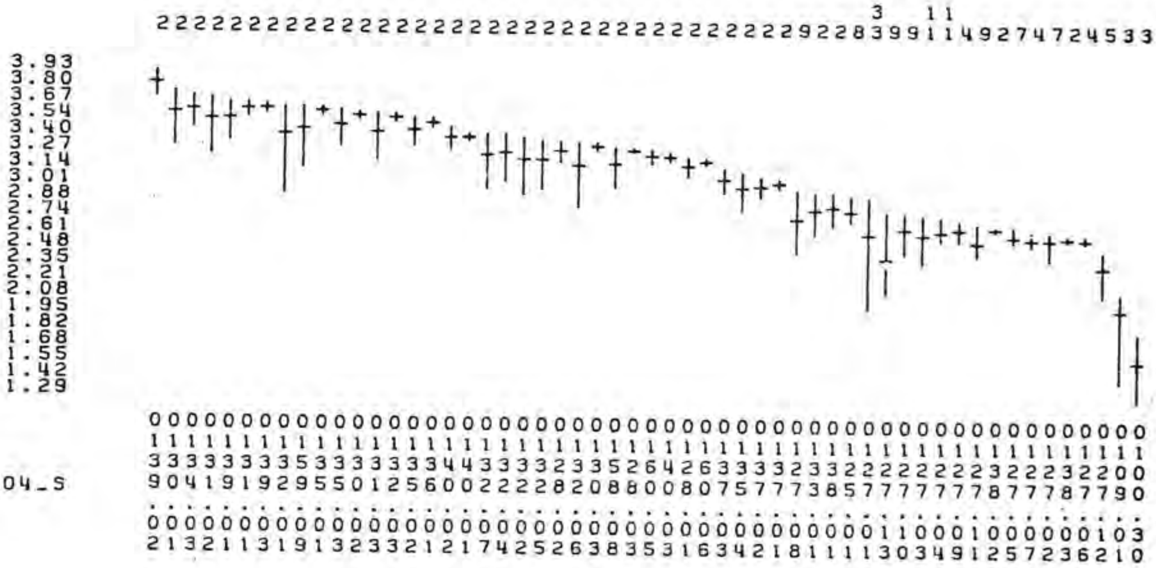

CO
1.30
1.09
0.87
0.66
0.44
0.23
0.01
-0.20
-0.42
-0.63
-0.85
-1.06
-1.28
-1.49
-1.71
-1.92
-2.14
-2.35
-2.57
-2.78
-3.00


NI
1.70
1.50
1.29
1.09
0.88
0.68
0.47
0.27
0.06
-0.14
-0.35
-0.55
-0.76
-0.96
-1.17
-1.37
-1.58
-1.78
-1.99
-2.19
-2.40

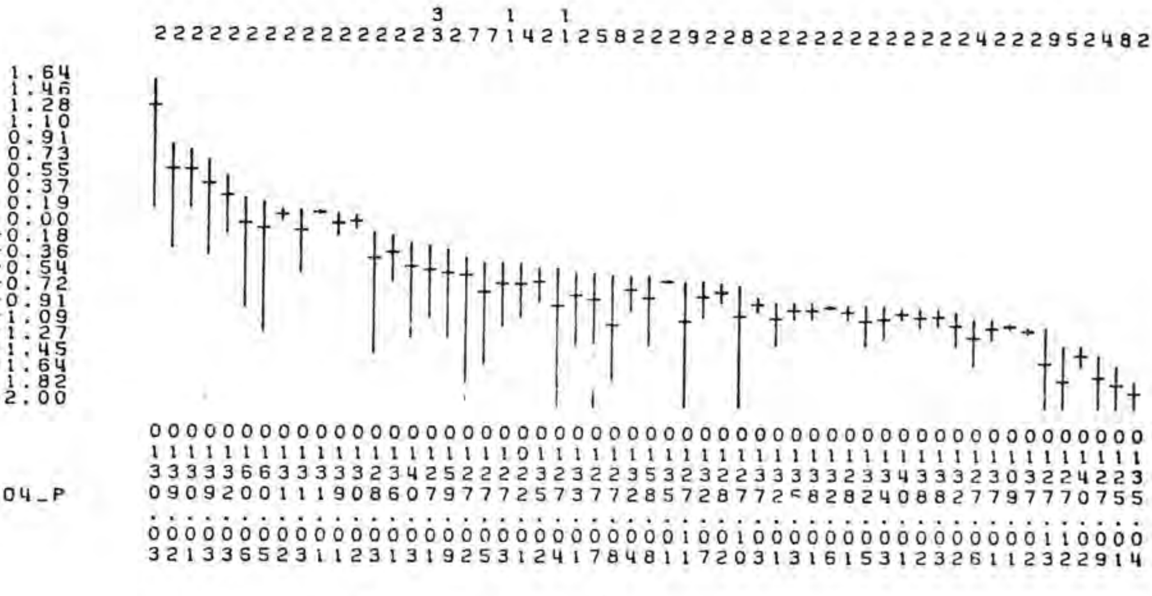

MN
2.74
2.52
2.31
2.09
1.88
1.66
1.45
1.23
1.02
0.81
0.59
0.38
0.16
-0.05
-0.27
-0.48
-0.70
-0.91
-1.12
-1.34
-1.55


NO3_N
0.60
0.49
0.37
0.26
0.14
0.03
-0.09
-0.20
-0.32
-0.43
-0.55
-0.66
-0.78
-0.89
-1.01
-1.12
-1.24
-1.35
-1.47
-1.58
-1.70

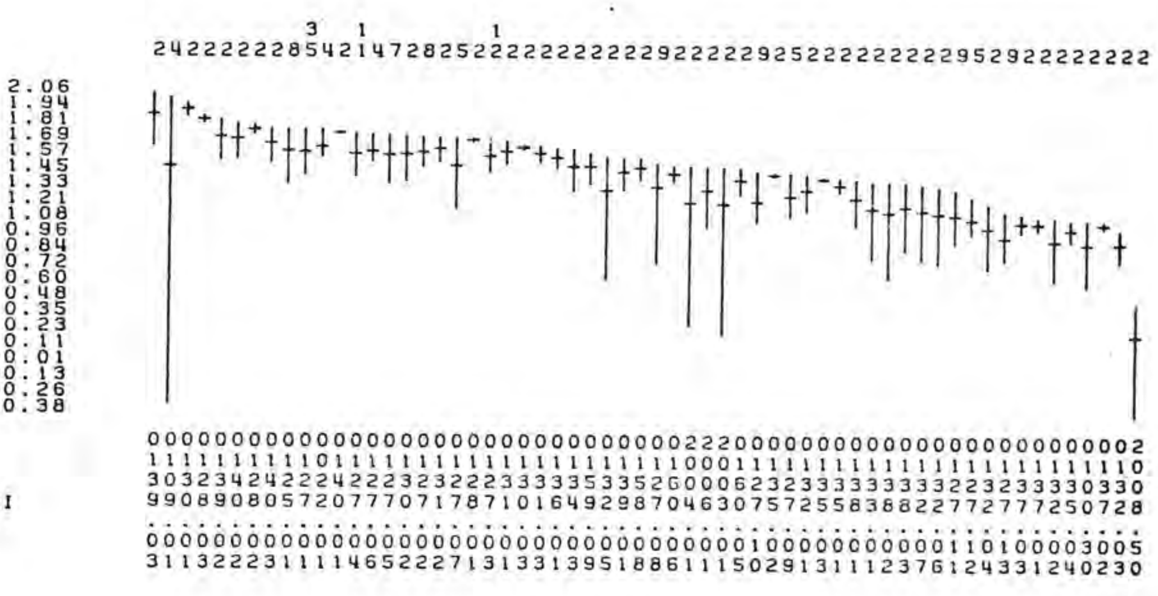

NH4_N
1.03
0.92
0.80
0.68
0.57
0.45
0.33
0.22
0.10
-0.02
-0.13
-0.25
-0.37
-0.48
-0.60
-0.72
-0.83
-0.95
-1.07
-1.18
-1.30



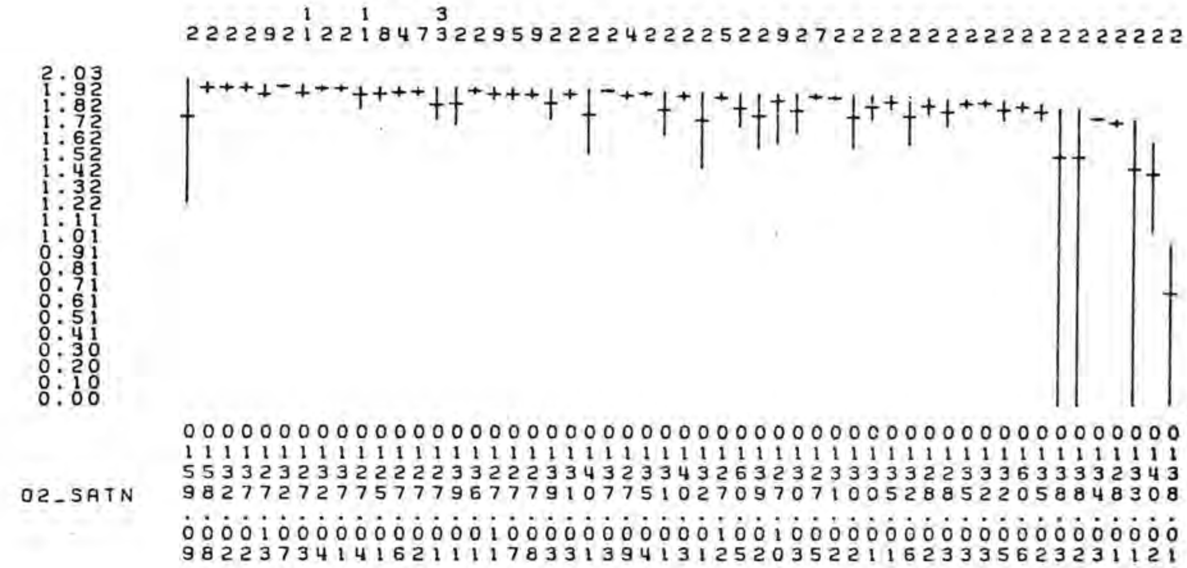
SO4_S

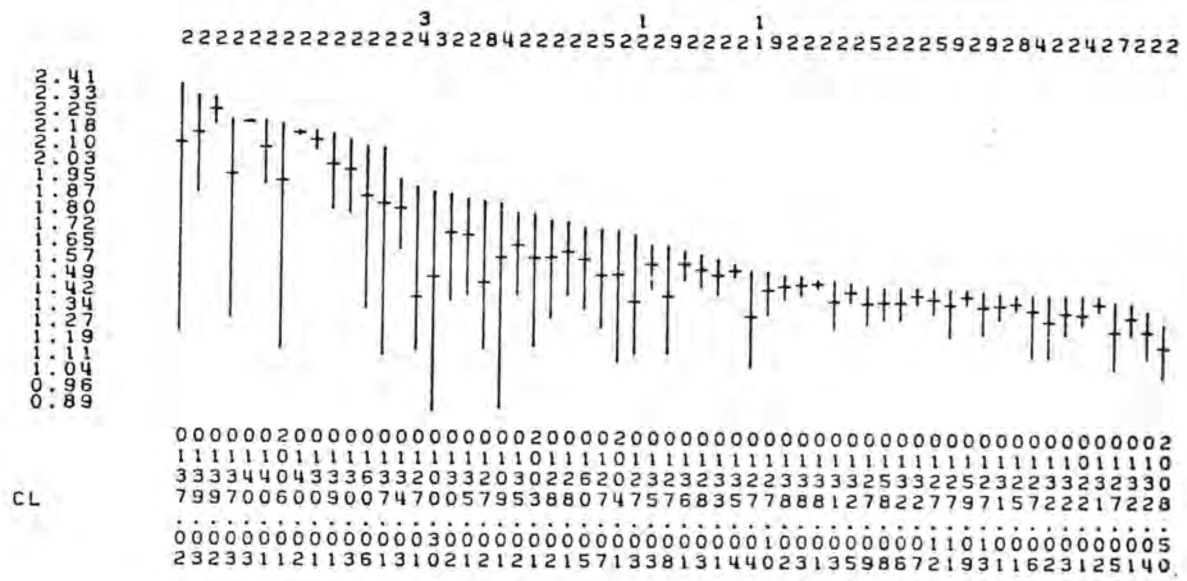


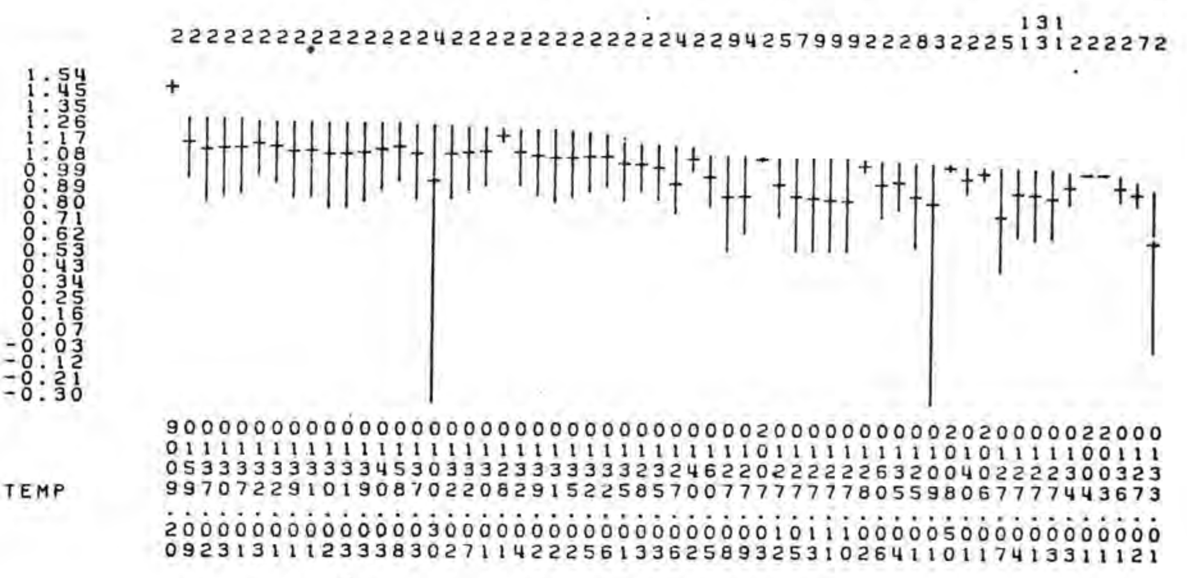
PO4_P



SI


2.03
1.92
1.82
1.72
1.62
1.52
1.42
1.32
1.22
1.11
1.01
0.91
0.81
0.71
0.61
0.51
0.41
0.30
0.20
0.10
0.00
O2_SATN


2.41
2.33
2.25
2.18
2.10
2.03
1.95
1.87
1.80
1.72
1.65
1.57
1.49
1.42
1.34
1.27
1.19
1.11
1.04
0.96
0.89
CL


1.54
1.45
1.35
1.26
1.17
1.08
0.99
0.89
0.80
0.71
0.62
0.53
0.43
0.34
0.25
0.16
0.07
-0.03
-0.12
-0.21
-0.30
TEMP

the summer sampling and the low value with the winter sampling and is almost certainly affected by the high flow regime of winter, a fact which in general may account for the variable range phenomenon for conductivity. The diagram for pH show 9 sites which cross unit boundaries, with Ø127.Ø8 being remarkable in that it ranges from below pH 3.Ø to above pH 6.Ø. Of these 9 sites 8 are at Brandon (reaches Ø8 - 13 and Ø125.Ø1 Brandon Pithouse Stream B) and one at Dowgang (Ø1ØØ.3Ø). The diagram for K is interesting in that it shows 12 Brandon Pithouse Stream A sites at the left of the diagram, exhibiting in addition very wide range variations. The only Brandon sites not there are Ø125.Ø1 and Ø127.Ø1 although these appear to have the characteristic feature of a wide range variation for K. Whether this is a feature of the Brandon site or an artefact requires additional research. Inspection of the diagram for Mg shows that the sites with the largest ranges are the Brandon sites (Ø2 - 13). Absent from this group are Ø127.Ø1 and Ø125.Ø1. The high mean values for all these sites suggests that the occasional low values have in reality a negligible effect on the overall high Mg regime. The diagram for Ca is made remarkable by the rapid fall in Ca maxima below pCa 1.39, and the sites which characterise the low Ca are the Avoca Basin streams 2ØØ3, 2ØØ4, 2ØØ6 and 2ØØ7. The diagrams for Zn and Cu show the Avoca sites at the left indicating high levels for these elements, with generally wide range variations as well. The Mn diagram shows discontinuities over unit boundaries with little range variation outside an order of magnitude. The diagrams for Fe, Co and Ni are unremarkable except for the relatively high values and wider variations to be found at the lower end of the variable ranges. The diagram for Pb is unremarkable except that the range variations are considerable and the low mean for the Brandon sites suggest that high Pb is the exception rather than the rule for Brandon. In contrast to the rest of the anions the variation in the ranges of SO_4-S are relatively small, rarely spanning an order of magnitude. Birch Coppice acid stream (Ø139) has the highest SO_4-S regime and also the highest levels for a majority of the cations and other variables. The Avoca sites are marked out by their high Cu and Zn and low Ca values. The Brandon sites are especially marked out by their high K and Mg values.

Chapter 6 Acid species

6.1 Species list generation and descriptive statistics

6.11 Introduction

Acid species are those found in aggregate biologies on performing a "bydate" expansion of acid chemistries, produced by the query fld-pH <= 4.0.

This section presents these 91 species in Table 6.11-1 and summarises some of them in terms of their associated environmental chemistries and gives a brief description of them in terms of their importance to the acid study.

The chemical data, which are used to describe these species are drawn from SIEUR irrespective of pH. There are no descriptions presented in this thesis of the sites above pH 4.0 which contribute to this synthesis.

Table 6.11-1

Minimum pH and other details of species found in waters with pH <= 4.0

Table 6.11.1

Speciesnumber W.D.U.	Durham	n	min. pH	Species binomial and authorities	Durham authorities
02——	021301	5	2.25	Cyanidium caldarium (Tilden) Geitler	.
030206	030203	101	1.41	Euglena mutabilis Schmitz	
030401	030401	6	2.50	Lepocinclis ovum (Ehrenberg) Lemmermann	
040200	040250	5	2.50	Cryptomonas sp.	
081301	081501	53	1.80	Gloeochrysis turfosa (Pascher) Bourrelly	
080000	089950	0	4.00	Chrysophyta, genus not known, sp.	
0800—	089952	1	3.70	Chrysophyta, palmelloid, > 8 mum long	
110404	090302	0	2.52	Cyclotella meneghiniana Kutzing	
110609	090404	0	2.52	Melosira varians Agardh	
1106—	090405	1	3.10	Melosira distans (Ehrenberg) Kutzing	
120107	100101	0	2.30	Achnanthes ilexella (Kutzing) J. Brunnthaler	
120100	100150	0	2.90	Achnanthes sp.	
120114	100171	0	2.10	Achnanthes minutissima Kutzing	(P.J.S.,J.R.C.)
120113	100175	0	2.30	Achnanthes microcephala (Kutzing) Grunow	(J.R.C.)
1211—	100602	0	2.52	Cocconeis placentula Ehrenberg	
121324	100870	0	2.52	Cymbella ventricosa Kutzing	(P.J.S.,J.R.C.)
121309	100871	0	2.10	Cymbella deliculata Kutzing	(J.R.C.)
121301	100873	0	2.30	Cymbella affinis Kutzing	(J.R.C.)
121321	100876	0	2.52	Cymbella sinuata Gregory	(J.R.C.)
121305	100877	0	2.30	Cymbella cesatii (Rabenhorst) Grunow	(J.R.C.)
121501	101071	0	2.10	Diatoma elongatum (Lyngbye) Agardh	(J.R.C.)
122002	101301	44	2.20	Eunotia exigua (Brebisson) Grunow	
1220—	101302	1	3.10	Eunotia lunaris (Ehrenberg) Grunow	
122100	101450	0	2.52	Fragilaria sp.	
123902	101472	0	2.40	Fragilaria (Synedra)] amphicephala Kutzing	(J.R.C.)
122311	101505	2	3.10	Gomphonema olivaceum (Lyngbye) Kutzing	
122311	101570	4	3.10	Gomphonema parvulum (Kutzing) Grunow	(P.J.S.,J.R.C.)
1227—	101801	0	3.10	Navicula avenacea Brebisson	
1227—	101802	0	2.52	Navicula cryptocephala Kutzing + gregaria Donkin	
122700	101850	1	2.40	Navicula sp	
1227—	101874	2	2.52	Navicula mutica Kutzing	(J.R.C.)
1227—	101875	0	2.30	Navicula mediocris Krasske	(J.R.C.)
1227—	101878	1	3.00	Navicula nivalis Ehrenberg	(J.R.C.)
1230—	101904	3	2.52	Nitzschia palea (Kutzing) W. Smith + paleacea Grunow	
123000	101950	0	3.05	Nitzschia sp.	
1230—	101970	1	2.50	Nitzschia ovalis Arnott	(J.W.H.,J.R.C.)
1230—	101971	29	2.50	Nitzschia subcapitellata Hustedt	(J.W.H.,J.R.C.)
123003	101974	0	2.52	Nitzschia amphibia Grunow	(P.J.S.,J.R.C.)
1230—	101978	2	3.10	Nitzschia umbonata (Ehrenberg) Lange-Bertalot	(J.R.C.)
1230—	101984	22	2.50	Nitzschia pusilla Kutzing	(J.R.C.)
1230—	101987	4	2.60	Nitzschia obtusa Ehrenberg	(J.R.C.)
123300	102050	1	3.05	Pinnularia sp	
123301	102069	59	1.50	Pinnularia acoricola Hustedt	(J.W.H.,J.R.C.)
1233—	102070	0	2.52	Pinnularia appendiculata (Agardh) Cleve	(P.J.S.,J.R.C.)
1233—	102071	0	3.10	Pinnularia borealis Ehrenberg	(P.J.S.,J.R.C.)
1233—	102073	1	3.30	Pinnularia subcapitata Gregory	(P.J.S.,J.R.C.)
1233—	102076	4	3.10	Pinnularia braunii (Grunow) Cleve	(J.R.C.)
123311	102078	3	2.50	Pinnularia microstauron (Ehrenberg) Cleve	(J.W.H.,J.R.C.)
1233—	102079	2	3.70	Pinnularia variabilis A. Cleve	(J.R.C.)
1233—	102080	1	3.80	Pinnularia termitina (Ehrenberg) Patrick	(J.R.C.)
123501	102201	0	2.52	Rhoicosphenia curvata (Kutzing) Ehrenberg	
123908	102403	0	2.52	Synedra ulna (Nitzsch) Ehrenberg	
124002	102502	0	2.10	Tabellaria flocculosa (Roth) Kutzing	
120000	109950	1	1.41	Pennales, genus not known, sp.	
1229—	109969	2	3.70	Neidium alpinum Hustedt	(P.J.S.,J.R.C.)
1208—	109972	0	2.52	Caloneis lagerstedtii (Lagerstedt) Cholnoky	(P.J.S.,J.R.C.)
120501	109973	3	3.10	Anomoeoneis exilis (Kutzing) Cleve	(P.J.S.,J.R.C.)
1205—	109974	1	3.10	Frustulia rhomboides Ehrenberg not var saxonica	(P.J.S.,J.R.C.)
1205—	109986	0	2.10	Anomoeoneis serians (Brebisson) Cleve	(J.R.C.)
213000	122850	30	2.10	Zygogonium sp	
1506—	130401	13	1.80	Chlamydomonas acidophila Negoro	
150600	130450	6	2.25	Chlamydomonas sp.	
160500	140450	13	1.50	Characium sp	
160600	140550	9	0.90	Chlorella sp.	
171923	152902	31	2.30	Hormidium rivulare Kutzing	
1719—	152905	4	2.25	Hormidium pseudostichococcus Heering	
1719—	152906	2	2.60	Hormidium vulcanum Negoro	
172201	153302	3	3.25	Microthamnion kuetzingianum Nageli	
172901	154201	1	1.41	Rosenvingiella polyrhiza (Rosenvinge) Silva	
1731—	154452	4	1.80	Stichococcus cells rounded, > 1 <= 2 mum	
1731—	154453	4	1.41	Stichococcus cells rounded, > 2 mum	
1731—	154454	2	2.50	Stichococcus cells + cylindrical <= 2 mum	
1731—	154455	8	0.90	Stichococcus cells + cylindrical > 2 <= 4 mum	
173410	154703	1	3.00	Ulothrix zonata (Weber & Mohr) Kutzing	
341001	220301	1	2.60	Cephalozia bicuspidata (Linnaeus) Dumortier	
322600	232150	20	2.50	Dicranella sp.	
3229—	232302	6	2.60	Drepanocladus fluitans (Hedwig) Warnstein	
326109	234304	1	3.80	Pohlia nutans (Hedwig) Lindberg	
3261—	234369	1	3.30	Pohlia cruda (Hedwig) Lindberg	(B.A.W.)
322801	236031	1	3.80	Ditrichum cylindricum (Hedwig) Grout	
320000	239950	4	2.20	Moss, genus not known sp.	
383010	251302	1	2.70	Juncus effusus Linnaeus	
383801	251801	1	2.70	Phragmites australis (Cav.) Trinnius ex Steud.	
384902	252502	1	2.50	Typha latifolia Linnaeus	
380200	253050	1	4.00	Agrostis sp	
—	319949	1	2.40	Prokaryote heterotrophs, genus not above	
—	319950	6	1.50	Prokaryote heterotrophs, genus not known, sp.	
—	329946	12	2.75	Eukaryote colourless heterotrophic ascomycete,	
—	329964	2	3.70	Eukaryote colourless heterotrophic aquatic hyphomycete	
—	329966	2	1.41	Eukaryote colourless heterotrophic ascomycete, not yeast	
—	329968	4	2.10	Eukaryote colourless heterotrophic yeast	