AN ASSESSMENT OF THE USE OF *BACILLARIOPHYCEAE* AS BIOLOGICAL MONITORS OF HEAVY METAL POLLUTION IN AUSTRALIAN TROPICAL STREAMS

Louise von Falkenhayn

Ph.D. dissertation

Department of Geographical and Environmental Studies

The University of Adelaide

December 2007

CHAPTER 1 - Introduction

1.1 Australian water resources and water quality

Australia faces a daunting problem regarding its water quality and resources. With the knowledge that Australia is the driest continent after Antarctica (Boulton & Brock 1999) and the country with the most variable precipitation (Smith 1999), water is one of Australia's most precious resources. As water resource issues become more important there is increasing public and governmental pressure for better environmental management and monitoring of water resources to retain and acquire critical quality levels. Water quality in streams is a measure of the health of an aquatic environment. The quality of drinking water, conservation of aquatic ecosystems and the aesthetics of surrounding waters are all directly linked to water quality.

Detrimental impacts on water quality can potentially affect all aquatic ecosystem processes, and thereby detract from the economic, social, recreational and natural values of the catchment. Awareness of water quality issues has provoked public and government bodies to establish water quality guidelines. An example is the Australian New Zealand Environment and Conservation Council (ANZECC) water quality guidelines. They aim 'to protect all forms of aquatic life and all aspects of the aquatic life cycle' (Fairweather *et al.* 1998, p. 2). However, one of the critical issues facing water quality is the impact of water utilisation by industries such as agriculture and mining which cause water pollution (Gouldie 1993). Heavy metals are one of the prime causes of toxic pollution and serious water quality deterioration. Australia is a nation that currently relies upon the exploitation of its natural resources through mining. This naturally leads to conflict between water management and development interests as resource exploitation commonly pollutes aquatic systems. Monitoring is therefore important for solving these water quality issues.

1.2 Mining in the Northern Territory

Acid water is known to drain from mines under all climatic conditions, ranging from arctic to tropical. The resulting acidic and heavy metal rich waters released from mines are a major source of aquatic pollution to receiving waters (Ramsey & Brannon 1988). In the tropics, local climate conditions exacerbate the acid mine drainage problem as the high temperature and humidity levels significantly accelerate the rate of acid production (Noller *et al.* 1997). Acid mine drainage is principally associated with the mining of sulphide ores, the most commonly associated minerals being sulphur, copper, zinc, silver, gold, lead and uranium (Gray 1998).

Mining has a long and established history in the Northern Territory (N.T.) of Australia. Mining is the leading industry, ahead of agriculture, tourism and defence, and accounts for an average $3.2 billion per annum (GDP) (www.nrm.gov.au/nrm/nt.html). However, as a result of fluctuating markets, there are many abandoned mines scattered throughout the Territory. Acid mine drainage (AMD) and its associated environmental degradation is linked, not only with working mines, but also with closed and disused ones. The pollution generation mechanisms (or AMD) at each of the closed mines are the same in all waste rock heaps and tailings. Mining of metallic ore deposits exposes pyrites and sulphidic material to the atmosphere or oxygenated groundwater. As a consequence, the sulphides oxidise to produce acid water laden with sulphate, heavy metals and metaloids. The term ´heavy metal´ includes most metals with an atomic number greater than 20, and excludes alkali metals, alkaline earths, lanthanides and actinides (Mason 1991). Heavy metal pollution to surface waters causes problems ranging from toxification of drinking water to biological problems related to pH imbalance, heavy metal toxicity, food chain disruption and interruption of growth and reproductive patterns of aquatic organisms (Fostner & Whittmann 1981; Lindegaard 1995). The acidification of streams to a pH lower than five results in the dissolution of metals from mineral deposits in streams (Planas 1996) compounding acid mine drainage effects.

In contrast to temperate regions, where sulphide oxidation in a wet climate produces a continuous supply of acid mine water, in the seasonally wet–dry climate of northern Australia, sulphidic materials are exposed to an irregular wet–dry cycle, which can lead to ephemeral acid mine drainage (Harris *et al.* 2003). At the onset of the wet season, sulphide oxidation and mineral dissolution processes will generate acid mine drainage waters that may enter local surface waters and aquifers. During the dry season, mobilisation of acid mine drainage, produced from the sulphidic material, will cease. Changed redox conditions commonly lead to the formation of Fe, Mg, $MgSO_4$-rich precipitates. Evaporation causes the formation of mineral efflorescence (i.e. post mining minerals that form due to evaporation as surface encrustations: Jambor *et al.* 2000) in pore spaces of sulphidic rocks, on stream beds and at seepage points and surfaces of waste-rock dumps, ore stockpiles and tailings repositories.

1.3 Monitoring change in aquatic systems

Aquatic ecosystems are dynamic and are characterised by great complexity, varying over space and time (Palmer *et al.* 1997; Stevenson & Pan 1999). However, they are increasingly being impacted by human activities occurring within catchments (Hart *et al.* 1999). Consequently, the study of aquatic ecological variation (heterogeneity) has increased due to recognition that variance may reveal a great deal about a system. Particularly, the ability to detect human- induced change in ecosystems depends on our ability to quantify natural heterogeneity (Li & Reynolds 1994).

Aquatic ecosystems are largely structured by climatic regimes, the physical (e.g. light, temperature, mixing, flow, habitat) and chemical (e.g. organic and inorganic, oxygen and nutrients) environment with which they interact, and the biological interactions (e.g. grazing and predation) that occur within them. Variation in these physical and chemical factors occurs naturally due to droughts, floods and erosion events. Climate variations, and consequent variations in rainfall, runoff and river flow, are particularly marked in Australia as they are strongly linked to inter-annual climate variability through mechanisms such as the El Niño Southern Oscillation. Longer climatic cycles, such as El Niño events and the ice ages, and geological/geomorphological events, such as faulting, eruptions and continental drift, are also major agencies of natural variation in ecological processes in rivers. However, generally rivers and streams undergo fairly predictable daily and seasonal changes in physico-chemistry and biota (Townsend & Riley 1999).

Another driver of change in aquatic ecosystems is human activities. Humans have almost universal roles in shaping and disturbing stream ecosystems (Resh & Grodhaus 1983; Petersen *et al.* 1987). A wide range of human-related stressors can impact upon aquatic ecosystems and modify their 'health'. For instance, pollution from industrial, urban, agricultural and mining sources, regulation of rivers through construction of dams or weirs, salinisation from land clearance, forestry, clearance of riparian zones and the introduction of exotic animal or plant species are a few examples of stressors. The severity and relative importance of human impact on streams varies across spatial scales, from affecting extremely localised microhabitats to affecting large regions. The temporal span of these disturbances can also vary greatly, ranging from days (e.g. some toxins) to centuries (e.g. dam construction) (Resh *et al.* 1988).

There are both natural and anthropogenic-induced changes in aquatic systems which occur at different rates. What is pertinent to the rate of change is the resilience of an aquatic system to disturbance. In dynamic environments, stasis may represent a disturbance or, as previously mentioned, in a stable environment disturbance may occur from human impact. The resilience of a stream to disturbance will depend on the 'health' of the stream. The general concept of 'stream health' is loosely defined as the degree to which a natural ecosystem is unimpaired by human impact, sustainable and resilient to environmental perturbations, while maintaining its ecological structure and function over time (Scrimgeour & Wicklum 1996). The resilience then of a natural 'healthy' stream is very high. However, consensus on what constitutes a 'healthy' ecosystem, and agreements on definitions, is contentious and various definitions have been generated by researchers (i.e. Steedman 1994; Karr 1995; Scrimgeour & Wicklum 1996; Meyer 1997). One attribute of a 'healthy' stream is its ability to recover from disturbance (its resilience). Resilience has long been recognised as a component of ecosystem health (Rapport 1989; Arrow *et al.* 1995), yet it is difficult to define and quantify (Karr & Thomas 1996). Ecological resilience has been defined as the amount of disturbance that an ecosystem can withstand without changing its self-

organising processes and variables that control its structures, for instance, without shifting to an alternative stable state (Holling 1973; Gunderson *et al.* 2002). According to Karr (1999), biota, for instance, are resilient when faced with normal environmental variation (within the range of its evolutionary experience) even when the variation is large, but the biota may not be able to withstand even the smallest disturbance outside its evolutionary experience. While an aquatic ecosystem is resilient to natural disturbance regimes it may be very sensitive to human activities (Bornette *et al.* 1998; Ward 1998).

In order to determine human impacts on an aquatic ecosystem it is important to distinguish impact (environmental change) from natural variation. It has already been determined that natural 'healthy' systems are resilient to the normal variation in an environment. The boundary to which a system changes from normal conditions, as a result of natural events, are narrow in comparison with the changes that result from human actions such as urbanisation or agriculture (Karr 1999). Normal or expected conditions, constituting 'health' or 'integrity', vary geographically as every river's biota evolves in the context of local and regional constraints and opportunities. Understanding this baseline can be the foundation for assessing anthropogenic change caused by humans, enabling assessments of health to be made and decisions made about which levels of change is acceptable. Humans may alter biological systems in a river by altering physical habitat, modifying seasonal water flow, changing the system's food base, changing interactions among stream organisms and contaminating the water with chemicals. When human disturbance is minimal the aquatic biota is determined by the interaction of biogeographic and evolutionary processes in the regional climatic and geological context. Severe human disturbance causes changes which alter the river's biota and thus the entire biological context of the river. Changes may be minor or they may eliminate all or most biota in a river. By measuring biological condition, and evaluating the result as a divergence from baseline biological natural conditions, one can assess whether change has occurred.

The 'health' of a system, or, as referred to by Karr (1995) the 'integrity' of a system, can be judged with reference to sites shaped by evolutionary and biogeographic forces and little human influence. RIVPACS, a system for stream health assessment in Britain, and AusRivas in Australia, have taken a similar approach, using the least disturbed streams to establish one end of the scale of reference; the other end of a scale can be established by studying highly disturbed sites (Wright 1995). Changes in environmental conditions can also be identified using either a 'top-down' or a 'bottom-up' approach. In a 'top-down' approach, changes at the level of community and ecosystem are assessed in the natural environment followed by identification of the cause of change (Cairns *et al.* 1993). In contrast, a 'bottom-up' approach usually relies on data produced from laboratory experiments, often at small temporal and spatial scales, to model changes in natural ecosystems (Cairns *et al.* 1993).

One tool for identifying and assessing biological condition and change in aquatic ecosystems is by using biological monitoring. Biological monitors integrate their responses to the environmental changes on a short-term, as well as a long-term, basis and reflect a change in ecosystem dynamics at many spatial, temporal and organisational scales (Burgman & Lindemayer 1998). Virtually all biota tell us something about their environment. However, it is generally impossible to study the entire biota present in a sampling area (Stevenson & Pan 1999) because of the constraints of time and the wide variety of sampling methods required for different groups of organisms. Consequently, strategically selected biological indicators are used whose interaction with the ecosystem makes them especially informative about the quality of habitat, communities and ecosystem processes. Some biota are more useful as monitors than others; specific biota may be suitable for monitoring (John 1998) specific conditions. Two important arguments in favour of biological monitoring are that; as the organisms have an integrating response to their environment, fluctuation in water quality which may have been missed by intermittent chemical analyses are recorded and; secondly, to maintain health of diverse biological communities it is more appropriate to monitor the aquatic community rather than physico-chemical variables (Cox 1991).

The attributes of ideal biomonitors in streams are summarised by several authors (Cox 1991; John 1998). The organisms selected should be close to the transfer of nutrients and energy in the food web, they should be simple and not complicated by life-cycle stages and be consistently and unequivocally identifiable to the species level (John 1998). They should be sensitive to even fine changes in the environment and, ideally, have precise ecological limits to their range of tolerance and preference to environmental variables. These ranges should be identifiable (Cox 1991). The biota should be distributed widely throughout the water body, preferably universal in distribution or should have a high degree of universality in distribution. They should be present in abundance and easily quantifiable. These criteria are met by a range of organisms, however most attention to date has focused on benthic macroinvertebrates, fish and algae.

Awareness of the detrimental effects of human stress on streams has resulted in a long history of monitoring using biological indicators (Rosenberg & Resh 1993). For example, the use of invertebrates in bioassessment began as early as the 1900s. Internationally the most utilised biotas for biomonitoring are macroinvertebrates, fish and algae. Fish have traditionally been used as indicators of habitat degradation, flow regulation and pollution (e.g. Gorman & Karr 1978; Bain *et al.* 1988; Belpaire *et al.* 2000). Benthic macroinvertebrates have been used to assess the effects of multiple stressor types such as organic pollution, acidification and general stress in aquatic systems (e.g. Karr & Chu 1999; Cao *et al.* 2003; Sandin *et al.* 2004). Algae have also been used for assessing nutrient enrichment, salinity and acidity (Round 1993).

Biological monitoring can be used to evaluate the success of management decisions because most restoration efforts aim at explicit biological goals, for instance, the return of fish. Thus biological endpoints can provide both a guide and a goal for ecological restoration. Additionally approaches such as before-after-control-impact (BACI) designs (Underwood 1994) and randomised intervention analysis (RIA) (Carpenter *et al.* 1989) determine whether biological communities have been influenced by human activity through the comparison of communities at the potentially impacted 'test' site and the minimally impacted 'reference' sites.

1.4 Monitoring and biomonitoring of Australian aquatic systems

One of the most significant impediments to aquatic monitoring in Australia is the poor ecological understanding of most rivers and streams (Hart & Campbell 1991; Boulton & Brock 1999). This complicates the ability to successfully implement management strategies for Australian waters. However, monitoring of Australian waters is an essential effort directed at assessing the impacts of development, and the success and usefulness of management of an environment (New 2000).

Monitoring of Australian aquatic ecosystems has traditionally been by means of assessment of water chemistry. However, the current trend in ecosystem health assessment is to directly monitor biological components of the ecosystem (Breen 2000). The move away from chemical monitoring of streams is due to a variety of reasons, but mainly because chemical monitoring has frequently underestimated or even failed to detect many important kinds of environmental degradation (Karr 1991). For instance, intermittent pollution may be missed in chemical sampling programs (Planas 1996). Additionally, chemical monitoring does not show whole ecosystem effects. As all living functions of aquatic organisms are carried out in aquatic systems they are naturally more sensitive and reflective of changes in water quality.

The role of a bioindicator has been well established in aquatic ecosystems (Kelly *et al.* 1995). There is a diverse range of organisms and techniques that have been used for biomonitoring. In Australia, benthic macroinvertebrates are the focus of most biomonitoring work (Norris & Norris 1995; Breen *et al.* 2000). Both the Australian National River Health Program and the Australian River Assessment System emphasise the use of aquatic macroinvertebrate fauna as water quality indicators. In assessing mine pollution, macroinvertebrates can show reduced abundance and richness at metal impacted sites (Winner *et al.* 1975; Winner *et al.* 1980; Hirst *et al.* 2002). However, in a study by Ferris and Vyverman (1996) of heavy metal impact on the Finnis River (N.T. Australia), there was an absence of macroinvertebrates at the most polluted sites. This limits macroinvertebrates to presence and absence studies, reducing the available information on how heavy metal pollution impacts ecosystems at highly polluted sites. Newall *et al.* (2006) compared diatoms and macroinvertebrates in the assessment of stream condition of the Kiewa River

(Australia). They found that diatoms were more closely related to water quality variables, whereas macroinvertebrates were primarily related to catchment and habitat features. While macroinvertebrate bioindicators continue to dominate monitoring studies in Australia there has been a move to incorporate algae into aquatic assessment programs. In 1995, under the Monitoring River Health Initiative, an algae bioassessment protocol was developed for Australian rivers (Hotzel & Croome 1999).

In the year 2000, the ANZECC guidelines were revised to incorporate biological indicators as well as bioavailable fractions of chemicals. The bioavailable fraction of chemicals is increasingly being seen as an important complementary addition to biological monitoring of aquatic systems. Heavy metals such as aluminium, cadmium, uranium, zinc, cobalt, nickel, lead, manganese and copper have all been identified as priority metals of potential ecotoxicological concern in aquatic systems of tropical Australia, largely as a consequence of mining activities (Markich 1997). Metal toxicity to aquatic organisms is not simply a measure of metal concentrations in systems. The toxicity of a metal is dependant upon metal speciation and its bioavailability (the ability of the metal to interact with a biological cell membrane). The ultimate bioavailability of a metal will depend on a number of factors including solution and solid phase speciation, sediment water partitioning, solution transport, passage across a cell membrane and absorption into organisms and accumulation in a target organ. Metals present as free metal ions (or as weak complexes that are able to dissociate at a cell or gill membrane) are more bioavailable than metals in strong complexes or adsorbed to colloidal and/or particulate matter (Batley 2002). Metal uptake is controlled by competition between metal species, protons, and calcium and magnesium for binding sites on the cell surface (Batley 2002). Ecological risk assessment computer codes such as HARPHRQ (Brown *et al.* 1994) and AquaRisk (Twining *et al.* 1999) predict the degree of ecological detriment associated with metal pollution in streams by determining the bioavailable fractions of metals. AquaRisk has been positively applied in risk-based assessment studies of the impact of aluminium in South Australia's Dawsley Creek (Brown & Ferris 2004) from Brukunga mine and copper effluent from Rum Jungle mine (Twining 2002a, Ferris *et al.* 2002). HARPHRQ has been used by Brown *et al.* (1994) to predict environmental impact of uranium on receiving waters from a uranium mine in northern Australia. It is increasingly seen as important to incorporate bioavailable fractions of metals into aquatic biological monitoring studies.

1.5 Diatoms as biological indicators

Because aquatic organisms continuously respond to their own environments, the use of biota as biological indicators provides substantial information on changing aquatic conditions (Dixit *et al.* 1992a). In addition, biota respond to the interaction and cumulative impacts of a variety of factors or habitat characteristics, which can not be interpreted individually by chemical or physical

analysis. Algae sit at the base of many food webs, hence their immense ecological importance. Any perturbation to the environment, physical or chemical, which seriously impairs algal communities, may then seriously affect numbers of other biotic communities dependent on the algae. Diatoms are single-celled algae belonging to the class *Bacillariophyceae*. The cell walls are composed of silica (SiO_2), which are divided into two valves (one frustule) joined by girdle bands. They are abundant and ubiquitous, occurring worldwide in lakes, oceans, wetlands, rivers and soils. Diatoms are the most species rich group of algae, with over 11,000 species documented (Mann 1999).

Diatoms are almost ideal biological monitors (Dixit *et al.* 1992b). They satisfy all criteria to qualify as suitable biological indicators as they are ecologically diverse and inhabit all types of habitats such as; open waters (planktonic), plants (epiphytes), rocks (epilithon), sand (episammon), mud (epipelon), wood (xylophilic) and animals (epizootic). Due to their high sensitivity and their quick response (short cell cycles), diatom assemblages can reveal abrupt changes in environmental factors (Dixit *et al.* 1992b), which are difficult to detect by chemical means. They are relatively easy to sample and only small samples are required for reliable assessments of community composition (Round 1993). Also, diatom taxonomy is well established (Whitton 1991). Diatom assemblages can be used to infer quantitatively environmental variables that are strongly directly, or indirectly, correlated with them through their optima and tolerances for each variable (Smol & Douglas 1996).

As a result of their importance as primary producers in freshwater ecosystems and their rapid response to environmental stress (Stoermer & Smol, 1999), diatoms have long been used to assess and monitor ecological conditions and monitor environmental change in streams and rivers of Europe, North America, Japan and sporadically in temperate Australia and New Zealand (i.e. Chessman *et al.* 1999; Prygiel *et al.* 1999; Stevenson & Pan 1999; Hill *et al.* 2000; Potapova & Charles 2002). Diatoms have been used as indicators in the northern hemisphere to monitor eutrophication (Descy & Coste 1990; Kelly & Whitton 1995; Coring *et al.* 1999; Stevenson & Pan 1999), organic pollution (Watanabe *et al.*1988) and human disturbance (Fore & Grafe 2002), acidification (Dixit & Smol 1995), salinisation (Round 1993), thermal effluents, forest fires and land use changes (Dixit *et al.* 1992b) and metal contamination (Stevenson 1984), and are now widely applied during routine water quality surveys in the northern hemisphere (i.e. the Water Framework Directive in the United Kingdom). By contrast, there are comparatively few studies using diatoms as indicators of pollution in regions of the sub-tropics and tropics and, in most cases, the type of pollution monitored is organic (Nather Khan 1991; Lobo *et al.* 1996; Silva-Benavides 1996; Michels 1998; Gomez & Licursi 2001; Asai *et al.* 2002). In tropical African and Brazilian streams diatoms have been used to infer eutrophication (Bellinger *et al.* 2006; Salomoni *et al.* 2006). In Malaysia diatoms have been used to indicate domestic and industrial waste (Maznah & Mashhor 2002).

One reason for the scarcity of studies in tropical regions is that diatom geographic distribution, floristic studies, ecological requirements, species tolerances and community responses are poorly known in many of these areas. This raises the question of whether diatom taxonomy, pollution indices and studies developed in the northern hemisphere can be applied in tropical contexts. The evidence of endemism of Australian tropical species would theoretically limit the applicability of indices and calibration sets developed overseas for temperate regions. This is confirmed by studies such as Newall *et al.* (2006) who found that locally derived bioassessment models and indices provided a more accurate assessment of the sites than the overseas-derived diatom index.

1.6 Diatom taxonomic assumptions

Diatoms are increasingly used in ecological monitoring and paleoecological reconstruction, in which accurate identification of species and knowledge of geographic distributions is essential. Over the last 15 years many new studies on diatom floristics and biogeography have been published (i.e. Krammer 1992; Lange-Bertalot & Metzeltin 1996; Kociolek & Spaulding 2000). Studies such as these are particularly important as diatom flora are increasingly being used to monitor changes in aquatic systems, through the development of indices (Kelly & Whitton 1995; Prygiel *et al.* 1999) and transfer functions (Philibert *et al.* 2006). For the application of diatoms, ecological studies based on taxonomy rely on consistency and correctness (Kociolek & Stoermer 2001). However, taxonomic classification of species for ecological studies is almost always reliant on morphological features for identification and floristic texts focussing on temperate regions.

Many diatom studies rely on the belief that the distribution of diatoms is cosmopolitan (i.e. Taylor 1929). However, increasingly, researchers believe that there is greater rarity and endemism among species than previously thought (Tyler 1996; Mann 1999). If diatoms were cosmopolitan it would be acceptable to conclude that the same floras and keys used in Europe could also be applied in Australia. However, studies on the regional distribution of diatoms, herbarium voucher material, local diatom populations and eolian dispersal do not support cosmopolitanism (Kociolek & Spaulding 2000). For instance, despite similar ecological conditions shared by the Arctic and Antarctic, out of 897 taxa of the West Antarctic and the Arctic only 80 taxa were reported as common to both polar regions (Hamilton *et al.* 1994). Within the cosmopolitan view, endemism is thought to occur infrequently, however there are increasing examples of endemism on the island of New Caledonia (Moser *et al.* 1998), Madagascar (Kociolek & Rhode 1998) and Australia (Vyverman *et al.* 1997 & 1998; Sabbe *et al.* 2001). In New Caledonia, over 40% of the species are considered endemics (Moser *et al.* 1998). Recent palaeolimnological studies, and studies of water chemistry relationships in Tasmania (Vyverman *et al.* 1995 a & b), have encountered many undescribed diatom species and provide further evidence that there is a distinct and unique

biogeographic element in the diatom flora of Australia. The discovery of new species was not unexpected given the many records of endemism in Australian flora (i.e. Tyler & Wickham 1988; Tyler 1992, 1995; Bowling *et al.* 1993). Vyverman *et al.* (1996, 1997) discovered new freshwater species in the genus *Biremis* and *Navicula* as well as a new genus, *Eunophora*, from Tasmania. However, determining whether these species are actually endemic is difficult to demonstrate.

In light of this, it is important to be wary of applying floras and taxonomies beyond their original regional scope (Edlund & Jahn 2000). With the limited number of floristic studies and taxonomic books mainly focusing on regions from the northern hemisphere, studies are repeatedly force fitting species into existing taxa. Mann (2000) has termed this 'amalgamation' or the clumping or grouping of morphologically similar taxa into fewer taxa. This leads to misclassification, masking rare or new species and highlights the importance of studies on regional distribution and taxonomic notation of species. Misclassification of species becomes especially problematic when cross-comparison between studies occurs.

Not only have diatoms as biological monitors been underutilised in Australia, taxonomic research in Australia has been sparse. Consequently an outcome of the prevailing view of the cosmopolitan nature of diatoms and the limited number of taxonomic studies in Australia, is the continual reliance of Australian diatom studies on European (Hustedt, Krammer & Lange-Bertalot) and North American (Patrick & Reimer) taxonomic texts. This reliance on northern hemisphere texts is compounded by the difficulty of researchers to consult type material held in overseas museum collections. The production of an iconograph specific to the wet tropics of Australia is essential to reduce the amount of 'force fitting' and 'masking' of endemic or new species. This will ensure future consistency between studies and add to the knowledge of diatom biogeography.

The success of diatoms for ecological studies is partly derived from the high level of physiological divergance that exists between individual species. However, there is the perception that phenotypic variation within diatom populations reflects equally high levels of genetic variation (Brand 1990; Wood & Leatham 1992). Hence, traditionally, the systematics of diatoms has been almost exclusively based upon frustule (silica shell) characteristics, for instance frustule shape, size and structure (Battarbee *et al.* 2001). Electron microscopy has revealed finer detail of the cell wall and enabled the use of a much larger range of ultrastructoral characters in identification. The trend of current taxonomic studies is to place greater emphasis on living material and on molecular techniques (i.e. Jahn 1986; Medlin *et al.* 1993; Cox 1996; Mann 2001) and, over the last 25 years the number of culture based diatom systematics studies have increased. These techniques are showing the limitations of the previous morphological based classification systems, and leading to the creation of new taxa. For instance, genetic studies have improved the taxonomy of particular

groups such as *Bacillariaceae* and have given insights into the evolution of the major lineages of the diatoms.

Through molecular techniques Gallagher (1980) and Brand *et al.* (1981) were the first to confirm directly that both physiological and genetic variation existed within diatom populations. With protein techniques Gallagher was able to detect two, essential mutually exclusive, summer and winter genotypes of the species *Skeletonema costatum*. In recent years a variety of more sensitive DNA-based techniques have been used to document intraspecific genetic variation in diatoms (i.e. Medlin *et al.* 1991; Stabile *et al.* 1992; Lewis *et al.* 1997). It is possible that organisms within classically defined taxa may differ morphologically along a pollution gradient and even subtle morphological differences can define distinct species. This morphological change may be attributed to either environmental factors or genetic differences, which could indicate a new deme (a closely related assemblage within species) or species and can often only be determined by molecular analysis (Mann 1999). The method is also essential in revealing the actual level of endemicity and number of new species in Australia through genetic comparison of morphologically similar populations in different geographical areas. DNA-based techniques are becoming increasingly important, especially as taxonomic precision, which is aided by genetic distinction of species and morphotypes, is crucial for consistency and monitoring studies using indicator species or indices. However, there is a large divide between taxonomic precision derived by professional diatom taxonomists concerned primarily with issues such as species concepts (Mann & Droop 1996) and ecologists seeking to apply taxonomic concepts to wider problems.

Genetic diversity within a species, especially a species with wide a geographic range, is advantageous because it ensures plasticity and, ultimately, the survival of a particular species in a variety of niches and a changing environment (Rynearson & Armbrust 2000). Molecular methods have greatly increased the ability to estimate genetic variation within communities and ensure accurate identification of species. The rRNA genes of diatoms have been employed to determine close phylogenetic relationships within diatoms (Beszteri *et al.* 2001; Lundholm *et al.* 2002). The nuclear-encoded small subunit (SSU) and large subunit (LSU) ribosomal RNA genes have been the most widely used for studying molecular systematics in diatoms (Medlin *et al.* 1993; Sorhannus *et al.* 1995; Van Der Auwera & De Wachter 1998; Medlin *et al.* 2000; Beszteri *et al.* 2001; Lundholm *et al.* 2002). Analyses of the ribosomal RNA genes have also provided useful insights into diatom evolution including the establishment of major clades (Medlin *et al.* 1993; Sorhannus *et al.* 1995). Additionally, variation in the non-coding internal transcriber spacer (ITS) have been used to separate isolates of diatoms from different geographic regions (Zechman *et al.* 1994; Penna *et al.* 2005) and have revealed several lineages within *Pseudo-nitzschia delicatissima*-like species in the Mediterranean Sea (Orsini *et al.* 2004). Though each of these molecular methods have been used successfully, there remains some discussion of which method is most suitable for certain species

and which method is best for determining genetic variation among species and within individual species. For instance, Godhe *et al.* (2006) suggests that, as the coding internal transcriber spacer (ITS) region is higher than LSU rDNA, it may be a better tool for intraspecific variation analyses.

1.7 Methods of diatom monitoring

Assessment of environmental conditions with diatoms can be based on single species (Raschke 1993), a group of indicator species (Ndiritu *et al.* 2003) or whole assemblages (Stevenson 1984; Jüttner *et al.* 1996). Transfer functions and biodiversity indices (e.g. species richness, diversity and evenness) are whole assemblage or community-based methods which have been used to monitor environmental conditions of aquatic systems. Two methods based on single species assessment are the definition of indicator species based on their autecology and, secondly, the use of presence and absence studies of teratological diatoms.

Diatom community-based transfer functions are commonly used techniques for diatom assessments of aquatic environments but were initially used primarily in lake studies, particularly for the reconstruction of historical and pre-historical environments (Pienitz & Smol 1993; Battarbee *et al.* 2001). Recently however, transfer functions have been established for streams to infer a variety of environmental variables, for instance, pH and conductivity (van Dam & Mertens 1995; Philibert *et al.* 2006; Walker & Pan 2006). However studies by Philibert and Prairie (1999) and Pipp (2001) have suggested that diatom-based indices and transfer functions that were developed within one region are inaccurate in other regions. For instance, Philibert and Prairie (1999) found that geographical location was significant even though environmental conditions were similar. It is important therefore, to develop site specific transfer functions for aquatic monitoring.

Most commonly, transfer functions are developed by using datasets with diatoms identified to the species level. However, the use of datasets with diatoms identified to the genus rather than species level has the advantage of being considered ‘rapid assessment’ (Chessman 1999). Monitoring using datasets of genus level identification is considered rapid assessment as large numbers of sites can be assessed without a high level of taxonomic expertise. The time and money saved would be attractive to mine managers. Genus level identification has been applied to various water quality monitoring studies (Prygiel 1991; Chessman *et al.* 1999; Growns 1999) both with positive and negative results. Growns (1999) used genus level identification of periphytic diatoms to test river regulation in eastern Australia. However, he found broadly similar results whichever taxonomic level was used. Chessman (1999) assessed the use of a predictive model based on periphytic diatoms identified at genus level for assessing ecological health over a broad area of eastern Australia. However, the predictive model was not able to demonstrate ecological disturbance at the majority of test sites. Success in application of the genus level to monitoring studies may be due to

the differing responses of species within a genus which may mask response if only genus level information is used. Lowe (1974) showed that species of *Navicula* have different ecological requirements, some are 'acidophilic' while others are 'basiophilic'.

Two methods applied for assessing the presence and extent of water pollution are species richness and diversity indices. Patrick (1963) suggested that the best means for detecting shifts in the environment was by considering shifts in numbers of species (species richness) and in evenness of species abundances. In urban and rural streams in Australia a decline in species richness with increasing nutrient concentrations has been reported (Sonneman *et al.* 2001). In an Argentinean stream, species richness was found to decrease at sites receiving effluents from chemical, textile, leather and metal industries (Gomez 1998), and in Idaho rivers, the total number of diatom taxa declined at mining impacted sites (Fore & Grafe 2002). However, an important aspect of the numerical structure of the community is ignored when the composition of the community is described simply in terms of the number of species present. It misses the information that some species are rare and others are common. Pollution in aquatic systems can create optimal conditions for some species increasing their abundance while making other species relatively rarer. Consequently, changes in the abundance, relative abundance and diversity of benthic diatoms have been commonly used to survey water quality (i.e. Marcus 1980; Tent 1981; John 1993). Low species diversity is often associated with stressful conditions such as polluted waters (Sullivan 1986; Rott & Pfister 1988) and diversity indexes, therefore, may be considered a reasonable measure of pollution impact. However, diversity can decrease or increase with pollution (van Dam 1982; Izsak *et al.* 2002), or changes can vary depending on the type of pollution (Jüttner *et al.* 1996; Hillebrand & Sommer 2000; Jüttner *et al.* 2003). Diatom diversity responses to different types of pollution are evidently not fully known. Archibald (1972) found that regardless of the diversity index used, diversity of diatom associations in some South African rivers could not be directly related to water quality.

One method of diatom assessment of water quality by single species is based on autecological information of the optima and tolerance of species to environmental variables such as pH. Species with well defined optima and narrow tolerance ranges can be used as potential key indicator species for monitoring programs. Battarbee (1999) states that diatoms are, perhaps the most useful and powerful indicators of water chemistry as their ecological optima and tolerances can be quantitatively defined. Although diatoms are ecologically sensitive to a wide range of parameters, their relationships to pH, total phosphorus and salinity are perhaps the best known and most established. Lacking still are the ecological responses of diatom species to heavy metals for example, copper and aluminium. Some of the most comprehensive investigations on the autecology of diatoms are the projects that examined the effect of changing pH on diatom communities: PIRLA-I and PIRLA-II (Palaeoecological Investigation of recent Lake Acidification – Charles & Whitehead 1986; Charles & Smol 1990) in eastern North America and SWAP (Surface Water

Acidification Project – Stevenson *et al.* 1991) in Europe. In the tropical north of Australia there exists almost no autecological information on the diatom species found in the region.

A second and rarely used method of single species indicators is the assessment of the presence and absence of teratological (aberrant) forms of diatom species. Teratological forms of diatoms have been observed most often in 'extreme' environments for instance in waters containing high heavy metal concentrations (Sunda & Guillard 1976; William *et al.* 1980; Adshead-Simonsen *et al.* 1981; Barber & Carter 1981; Fisher *et al.* 1981; Carter 1990; Dickman *et al.* 1990; Andresen & Tuchman 1991; McFarland *et al.* 1997; Dickman 1998; Yang & Duthie 1999). Teratological forms have been correlated with the presence of genotoxic and teratogenic chemicals (i.e. copper) in aquatic systems in a number of countries (William *et al.* 1980; Barber & Carter 1981; Dickman *et al.* 1990; Andresen & Tuchman 1991; Yang & Duthie 1993). However, there have been few studies on the use of teratological diatoms as indicators of metal pollution. Dickman (1998) doubts diatom deformity frequency is a useful tool as the amount of time and effort required to identify a statistically meaningful number of deformed diatoms is so great, and suggests that an automated system of assessing diatom deformity frequency needs to be developed. However, Cattaneo *et al.* (2004) successfully used morphological changes as indicators of metal pollution and recovery in Lac Dufault (Canada).

1.8 Diatom studies in Australia and the tropics

Up until the 1980´s, with the exception of Foged (1978), diatom studies were almost nonexistent in Australia, particularly on inland waters. In 1983 Thomas compiled an illustrated species list with emphasis on the ecology of diatoms from the Alligator Rivers Region (Northern Territory). In the same year John (1983) produced a similar treatise of diatoms from the Swan River Estuary in Western Australia. Diatom monitoring studies were initiated by Chessman in 1985. After a period of being underutilised in Australia (Reid *et al.* 1995) diatoms are increasingly being valued in Australia as biomonitoring organisms (Tibby *et al.* 2003; Philibert *et al.* 2006). Diatom datasets have been established in Australia to infer lake and reservoir salinity (Gell 1997), pH (Tibby *et al.* 2003) and phosphorus (Tibby 2004). Though stream diatoms have been applied to the AUSRIVAS model with limited success (Chessman *et al.* 1999) they have been used to demonstrate the impact of urbanisation (John 2000a, 2000b; Sonneman *et al.* 2001), river regulation (Growns & Growns 2001), nutrient (Chessman 1985a; Philibert *et al.* 2006), and thermal (Chessman 1985b, 1986) pollution, and general human modification (Chessman *et al.* 2007) in Australia.

Despite their increased use in temperate Australia (i.e. 1993; Gell 1997; Chessman *et al.* 1999), and despite the wide use of diatoms as indicators of various forms of pollution in Europe and North America, there has been few water quality monitoring studies in the wet/dry tropics of northern

Australia (Ferris & Vyverman 1996). One of the first studies conducted in the Northern Territory was by Thomas (1983) who developed a taxonomic and regional description of the Alligator Rivers Region. Townsend *et al.* (2002) investigated the response of diatoms to reduced dry season flows in the Daly River. More recently Townsend and Gell (2005) investigated the role of substratum type on benthic diatom assemblages in the Daly and Roper rivers. The only published diatom water quality monitoring studies, focusing on mine pollution, were conducted by Ferris & Vyverman (1996) on the Rum Jungle mine, Chaney *et al.* (1979) on Magela creek and Schultz *et al.* (2002) on impacts of the Union Reef and Moline mine impacts on the Mary River. Additionally, most studies on benthic diatom in the tropical waters of northern Australia are applied studies rather than taxonomic or floristic investigations. Apart from the following investigations listed, the diatom flora in Australian tropical rivers and their autecology is almost unknown. In 1987, Thomas described two new species (*Achnanthes pseudohungarica* and *Eunotia didyma* var. *maxima* f. *tumida*) from his diatom survey between 1979 and 1981 of the Alligator Rivers region of the Northern Territory. Even within the whole of Australia relatively little research has been undertaken on diatom taxonomy compared to other parts of the world. Apart from the descriptions of new species such as those by Foged (1978); John (1980, 1981 a, b, c); Haworth & Tyler (1993); Hodgson *et al.* (1997) and Vyverman *et al.* (1997, 1998), and regional flora descriptions such as Sonneman (*et al.* 2000), most Australian researchers rely heavily upon the comprehensive taxonomic texts produced in the northern hemisphere to identify Australian species.

1.9 Diatom monitoring of heavy metal pollution

It could be expected that diatoms are responsive to heavy metal pollution as benthic diatoms have been widely used for detection of changes in water quality in rivers due to their species specific sensitivities to a wide range of ecological conditions such as pH, salinity and nutrients (Lange-Bertalot 1979; ter Braak & Van Dam 1989; Van Dam *et al.* 1994; Kelly & Whitton 1995; Pan *et al.* 1996). However, their use as biological monitors of water quality has been predominantly established in Europe and North America. Owing to the climate and nature of the environmental pollution problems faced by these regions, monitoring has mainly been on the eutrophication (Kelly *et al.* 1995) of rivers and lake acidification caused by acid rain (van Dam *et al.* 1994; Patrick *et al.* 1996; Battarbee *et al.* 1999). While diatom responses to acid water caused by acid rain is well documented (i.e. the North American Paleolimnological Investigation of Recent Lake Acidification and in Europe the Surface Water Acidification Program), low pH due to mine drainage affects water chemistry differently (Kelly 1988) and only recently have attempts been made to quantify the effects of metal pollution using diatoms (Dixit *et al.* 1992b; Patrick *et al.* 1996). There has been little in the way of characterising diatom community response to acid mine drainage. The general assumption is that species sensitivities to metals may be reflected in changes in relative abundances of species in communities (Genter *et al.* 1987) or the replacement of sensitive algal species by

more tolerant ones, which ultimately induces adaptation of individual species (Blanck *et al.* 1988). Of the limited number of published studies on diatoms in highly acidic mining waters, most have been undertaken in the USA and England (i.e. Patrick 1974; Whitton & Diaz 1981; DeNicola 2000; Hirst *et al.* 2002) with only sporadic research around Europe (Sabater 2000). Isolated studies have been conducted in Antarctica (Cunningham *et al.* 2005) and Malaysia (Douglas *et al.* 1998).

There have been few studies on the affect of acidification or metal pollution on diatom species and community structure in Australian waters. One researcher who has attempted to rectify this is John (1993) through his studies in Western Australia on acidic and metal contaminated lakes. Stauber *et al.* (1996) incorporated diatoms into her mine water quality studies of the Mt Lyell mine remediation project. Since then Sincock *et al.* (ANSTO media release 1999) have used diatoms for similar purposes at the Brukunga mine site in South Australia and the Rum Jungle mine in the Northern Territory (Ferris *et al.* 2000). Diatoms have also been used in studies on metal impact on the Daly River (Townsend *et al.* 2002), Mary River (Schultz *et al.* 2002), Magela Creek (Chaney *et al.* 1979) and Settlement Creek (Mutton, unpublished thesis 2001).

Owing to the lack of studies on diatoms and acid mine drainage, the responses of diatom species and communities to heavy metals are not well described. Several studies on metal polluted rivers have shown that metal contamination can exert a selection pressure that might modify species composition of algal communities, affecting genetic composition of algal populations (Harding & Whitton 1976; Foster 1982; Kelly & Whitton 1989). Metal contamination may drive succession in algal communities towards more tolerant species (Gustavson & Wängberg 1995; Sabatar 2000), resulting in an increased tolerance of communities (Blanck *et al.* 1988), but may also pose a loss of species diversity (Leland & Carter 1984; Medley & Clements 1998; Sabatar 2000). However, Patrick *et al.* (1996) found that toxic pollutants could increase evenness and that severe pollution could decrease species numbers while a study by Hirst *et al.* (2002) found that diatom diversity, richness and evenness did not vary in association with changes in metal concentrations. Diatom morphology (i.e. cell size reduction) has also been found to change in response to metal pollution (Cattaneo *et al.* 2004) and teratological diatom forms have been noted (Sunda & Guillard 1976; William *et al.* 1980; Adshead-Simonsen *et al.* 1981; Barber & Carter 1981; Fisher *et al.* 1981; Carter 1990; Andresen & Tuchman 1991; Dickman *et al.* 1990; McFarland *et al.* 1997; Dickman 1998; Yang & Duthie 1993).

1.10 Study aims

The overall goal of my study is to test the utility and suitability of diatom taxa and community structure as indicators and monitors of heavy metal pollution on streams in the Northern Territory. Suitability was tested by using indices of community structure (e.g. diversity, evenness, and

richness), diatom transfer functions and the establishment of possible indicator species with weighted averaging and the presence of teratologies. In parallel, the response of diatoms to heavy metal pollution at both community and species level will be assessed. Additionally, this study addresses the lack of documented diatom floristic studies from the tropical Northern Territory of Australia. It provides one of the first detailed inventories of freshwater diatoms from the region. Finally the study aims to assess diatom classification techniques, morphological and genetic, for the morphologically variable diatom taxon *Nitzschia palea,* and to determine if there is any significant genetic difference between *Nitzschia palea* isolates from the Northern Territory and from sequences documented in GenBank. Taxa identified genetically as species other than *Nitzschia palea* are morphologically described.

1.11 Thesis Structure

Chapter 2

Chapter two describes the main environmental characteristics of tropical rivers in the Northern Territory and provides a description of each of the four mine sites studied, in terms of their location, geography and mining history.

Chapter 3

This chapter describes the methodology used for the field and laboratory collection and analysis of diatoms, water chemistry and physical observations. This includes the culturing and sequencing techniques for *Nitzschia palea*-like cells. Additionally, this chapter describes the descriptive and quantitative statistical analysis of the diatom and environmental datasets. This is composed of two main steps, calibration and regression. Calibration aims at assessing the relation between the abundance of diatoms and environmental variables. Once the correlated environmental variables to diatom assemblage are identified models are developed with weighted averaging techniques. Weighted averaging is employed together with Gaussian logit curves to identify possible indicator species for mine polluted environments. The final section of this chapter outlines various diversity indices and the regression methods by which diatoms are statistically related to environmental variables.

Chapter 4

In chapter four the chemical, physical and biological data collected from four catchments in tropical northern Australia are described. The chemical gradients at sites downstream from the pollution sources are summarised as well as the general chemical characteristics of control sites.

Additionally, the relationships between each of the chemical and physical variables are assessed with the development of a correlation matrix. The bioavailability of each metal is determined by AquaRisk and their relative importance at each mine site is described. The biological data, in terms of diatom composition at species and genus level of identification, is discussed in terms of abundance and occurrence within the entire dataset. Finally, as the freshwater diatom flora from the tropical region of northern Australia remains poorly known, this chapter improves the knowledge on diatom distribution by compiling a diatom iconograph from the flora from the four catchments sampled.

Chapter 5

The classification methods for morphologically different *Nitzschia palea* cultures are compared in this section. Optical and genetic classification techniques are used to ascertain whether the classified groups of *N. palea* differ according to the method used. Finally, this section aims to outline any genetic variation between sequenced cultures and to determine if the genetic grouping indicates that the *N. palea* cultures are, in fact, different species or morphotypes. Lastly, the species, which are identified by genetic classification, are described morphologically.

Chapter 6

In this chapter each of the various methods of using diatoms as monitors of water quality are assessed to determine which method is optimal for monitoring tropical rivers impacted by heavy metal pollution. The first section in this chapter explores the relationship between diatom assemblages and environmental variables. It identifies the environmental variables that would best explain the diatom species distribution. Transfer functions are developed using three different datasets; species with an occurrence greater than one, genera with an occurrence greater than one, and the bioavailable fraction of metals. The outcomes are compared to determine which type of dataset is optimal for a transfer function method of diatom monitoring. The next section within this chapter evaluates the manner in which acid mine drainage, represented by pH and copper, affects species diversity and richness of benthic diatom communities. In particular, it aims to determine if changes in diversity and species richness can infer heavy metal pollution and possibly be used to monitor acid mine drainage. Subsequently, I sought to identify diatom indicators by studying the optima and tolerances of diatom species to pH levels in streams of the Northern Territory. The final section of this chapter seeks to determine which diatom species display abnormal cell morphology and to describe the characteristics of the deformities. This final section aims to determine whether the presence of teratological diatom cells can be used as an indicator of heavy metal pollution to reliably infer copper concentrations or pH levels.

Chapter 7

This section outlines the main results concluded from the research and provides future research perspectives.

CHAPTER 2 - Characteristics of study sites

2.1 Characteristics of Australian tropical rivers

Year round equable temperatures and strong rainfall gradients make Australian tropical rivers quite different from otherwise comparable systems in southern Australia (Pearson 2004). Knowledge of northern tropical river systems is scant and largely limited to studies at the periphery of the region, for example, the Alligator Rivers region (Pearson 2004). Studies on regions such as the Alligator Rivers region are often outcomes of environmental oversight of resource development, for instance, uranium mining. Consequently, the river systems of tropical northern Australia are poorly understood in comparison to Australia's temperate freshwater and tropical marine systems (Hamilton & Gehrke 2005). In recent years, forums such as the 'Sustainable Futures for Australia's Tropical Rivers' forum, held at Charles Darwin University in 2004, have been held to assemble and synthesise existing scientific knowledge on Australian tropical river systems and to identify gaps in knowledge or future research perspectives.

2.1.1 Hydrology

Tropical Australian rivers tend to have highly seasonal flow regimes (McMahon *et al.* 1991) with much of their annual discharge occurring between November and May. Consequently the surface hydrology of these rivers are a reflection of the intensity and duration of the wet season. More than 95% of rivers' annual flow volume falls during the wet season (Townsend 2003). Surface drainage occurs during the December to March wet season and most rivers and streams then cease to flow during the dry season, when rainfall rarely exceeds 20 mm each month. In the dry season, except for large river systems which are fed by groundwater, the rivers reduce to either a series of permanent or temporary pools, or dry riverbeds. The variability of annual flows and annual floods exhibits a similar spatial pattern to rainfall (Erskine *et al.* 2005).

2.1.2 Climate

The climate of all four mine regions is classified as monsoonal, having two well defined seasons, commonly known as the wet and the dry season. The wet season lasts from November to April. The climatic characteristics vary greatly between the two seasons. During the monsoonal summer months an average of 1400 mm of rain falls which accounts for approximately 95 % of the annual rainfall. Seasonal temperature ranges from 21-38 ºC during the dry season to 23-44 ºC during the wet season. This is coupled with humidity levels of between 20-30 % in winter rising dramatically to between 50-100 % preceding, and during, the summer monsoon.

2.1.3 Water chemistry

Large alkalinity ranges exist for Northern Territory rivers, from very soft water with little carbonate buffering capacity (i.e. 5 mg/L $CaCO_3$), to rivers supplied during the dry season by dolomite and limestone aquifers with alkalinity levels reaching as high as 500 mg/L (Townsend 2003). Alkalinity is primarily a function of catchment geology however it can be affected by acidic drainage from mine sites, the most commonly occurring point source of pollution in the Northern Territory. An additional effect of the higher rainfall and consequent leaching of soluble materials from soil is that many tropical soils tend to have low nutrient concentrations (Brodie & Mitchell 2005). First flows in ephemeral streams in the Northern Territory are composed of high levels of soluble and physical constituents (i.e. high turbidity and suspended matter) due to the 'flushing' effect of the first rains on the catchment following the months of the dry season. Stream water quality typically improves through the wet season, with best water quality usually encountered during the peak runoff events.

2.1.4 Biota

Tropical inland waters are notably more biologically diverse than waters typical of the northern temperate zone (Williams 1994) and many species are regionally restricted in distribution. As in other tropical river systems, aquatic food webs appear to be largely algal driven (Jolly 2004). However, threats to these tropical systems include poor management of current activities such as grazing, river regulation, mining and invasive plants and animals which disrupt normal biophysical processes.

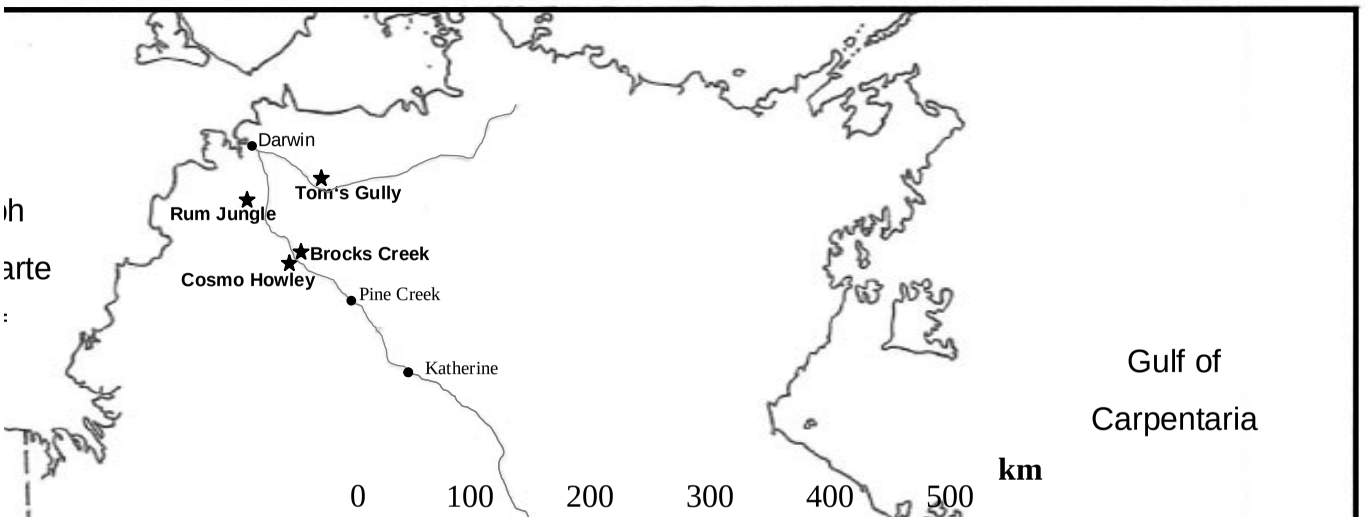
2.2 Mining in the Northern Territory

Mining in the Northern Territory (N.T.) started as early as the 1870's with the discovery of gold by workers on the Adelaide to Darwin Overland Telegraph. Over successive years, major discoveries were made for manganese, bauxite, uranium, tin and copper. Today, the minerals industry in the N.T. continues to be one of the main economic contributors. However, with the fluctuation of mineral prices and the exhaustion of supply, a mine may be abandoned or decommissioned. Throughout the N.T. there are many such mines that the N.T. government have placed in a management condition known as 'care and maintenance'. 'Care and maintenance' defines a mine that is not currently operating but is maintained in such state as to be capable of re-opening. For example, Tom's Gully mine was placed in 'Care and Maintenance' but due to higher gold prices the mine was reopened early in 2007. Ideally, when the decision is made to place a mine in 'care

and maintenance' (C&M), it is recognised that there are still ongoing environmental obligations to be met such as monitoring of any continued impact. Illustrated in Figure 2.1 are the locations of the four decommissioned mines, in tropical N.T., from which streams in their associated catchments were sampled. Each of the mines were under a 'care and maintenance' status at the time of sampling.

The management and monitoring of impacts from mine wastewater is addressed under the Northern Territory Mine Management Act (2001). The Department of Business, Industry and Resource Development issue guidelines to provide guidance on matters referred to in the Act, including, wastewater, decommissioning and rehabilitation of mines. The guidelines are not mandatory (there is no binding legislation) and where possible they are developed using policies from other states. The legislation emphasises that the responsibility for outcomes is borne by operators and the duty of care by all parties (self-regulation) rather than the traditional control approach to regulation (Fox & Jan 2000; McGill 2002). It is prudent then that monitoring programs are developed which are effective and attractive to mine managers.

Figure 2.1. Location of mine sites (★) in the Northern Territory.



2.2.1 Rum Jungle Mine

Rum Jungle is an abandoned uranium/copper open cut mine located 85 km south of Darwin in the Northern Territory on the East Finnis River. Mining was carried out between 1954 and 1971. Initially discharge of tailings was unconstrained, draining into the old tailings creek and then into the East Branch of the Finnis River. Later, tailings from the treatment plant were discharged into Dyson's opencut until the cessation of mining in 1971. From 1971 tailings were discharged into White's opencut. Contamination occurred by natural leaching of heavy metals from the waste rock dumps and the heap leach pile and, to a lesser extent, from the opencuts and tailings dam. The major pollutants to the river were copper, manganese, zinc and sulphate; copper being the most significant of these (Kraatz *et. al.* 1998). By the early 1970s the generation of sulphuric acid in the overburden heaps resulted in the East Branch becoming biologically 'dead' from the mine site to its confluence with the main branch of the Finnis River 8.5 km downstream from the mine site (Kraatz 1992; Kraatz *et al.* 1998). Rehabilitation was attempted between 1982 and 1986.

The topography of the Rum Jungle mine area is characterised by low relief and drainage patterns (Wills 1984) and gently undulating land drained by the Finnis River Catchment. The Rum Jungle mine is located on the East Branch which has no flow from approximately July to December whereas the Finnis River has permanent flow from its junction with Banyan Creek, 26 km upstream from its junction with East Branch. The Rum Jungle area is located in the north western part of the Pine Creek geosyncline. The mineralisation of the area, which besides uranium and copper, also includes small concentrations of lead, cobalt and nickel, occurs in highly sheared and faulted slates. The vegetation of Rum Jungle is characterised by Casuarina woodlands with species such as *Casuarina cunninghamiana* and *Casuarina obesa* interspersed with eucalypts such as *Eucalyptus camaldulensis* (River gum). Additionally, there is an infestation of the exotic plant *Lycium ferocissimum* (South African Boxthorn).

2.2.2 Tom's Gully Mine

Tom's Gully is a gold mine located 100 km south east of Darwin and 1.6 km west of the Arnhem Highway on Mount Bundy Station. Gold mining commenced at this mine in 1988 and ceased in 1991 (Sirocco 1998). The majority of drainage from the mine site is within the Mount Bundy Creek Catchment, which forms part of the Mary River Catchment. There are numerous ephemeral creeks in the area, of which the main is Mount Bundy Creek. Any discharge from the mine (i.e. from the tailings dam or sulphide or oxide waste dumps) should flow through an oxbow wetland and then into Mount Bundy Creek, which continues into Hardies Creek and then into the perennial Mary River.

The majority of the catchment contains outcropping rock with skeletal soil cover (Sirocco 1999). Vegetation is characterised by eucalypt woodlands with tropical grass understoreys. The mine lease lies in the north central part of the Pine Creek geosyncline. The deposit is hosted by the Wildman Siltstone, which is poorly exposed throughout the region occurring as low, undulating, rubbly rises north of the Arnhem Highway. The host rocks to the mineralisation are graphitic siltstone and sandstones of the Wildman Siltstone (Carpentaria Gold 1988).

2.2.3 Redbank Mine

The Redbank mine is located 900 kilometres southeast of Darwin and 30 kilometres west of the Queensland border near the Gulf of Carpentaria. The mine is situated partway along Hanrahan Creek. Any leachate from the ore stockpile flows into Hanrahan Creek, which then flows into Settlement Creek extending to the Queensland border. Redbank has been mined for copper from 1916 to 1996 (Redbank Copper 1992). The mine site consists of a pit, waste rock dump, tailings dam and crushed ore piles. There are two major sources of contamination. The first is the tailings dam and the second source is the waste rock dump adjacent to the tailings dam. The mine is located on the Gold Creek Volcanics, overlying the Wollogorang Formation, at the south-eastern end of the McArthur River Basin. The geomorphology of the area is sandstone and rhyolite (Masterton sandstone and hobblechain). The soils are low in fertility and low in exchangeable cations, particularly Ca^{2+} and Mg^{2+}, due to rapid decay of organic matter and thorough flushing of the soil system by consecutive wet seasons. Flora is dominated by eucalypts, speargrass and spinifix grasses.

2.2.4 Cosmo Howley Mine

The sites of Cosmo Howley are affected by effluent from both Cosmo Howley Mine and Brocks Creek Mine. These two mines are situated 40 kilometres from one another. Effluents from these two mines flow into the Howley Creek system.

2.2.4.1 Cosmo Howley Mine

Cosmo Howley mine is located approximately 130 kilometres south east of Darwin and approximately 40 kilometres south east of the township of Adelaide River. Gold and silver was mined at Cosmo Howley from 1986 until 1995. The main heavy metal contaminants from the mine are copper, iron, aluminium, manganese, cobalt, nickel, zinc and arsenic. The primary source of these metals is runoff and seepage from a waste rock dump (Jones *et al.* 2002) into an unnamed tributary of Bridge Creek which flows north eastwards into Bridge Creek and then northwards to

join Howley Creek, 19 kilometres to the north. Howley Creek drains into the Margaret and Adelaide Rivers. Bridge Creek has a short recessional flow period at the end of the wet season while Howley Creek flows throughout the wet season. The mine site area is located on the watershed between two major catchments and at the intersection between three sub-catchments. Bridge Creek and Howley Creek sub-catchments flow into the Adelaide River system to the north. The Adelaide River Catchment drains low lying land to the north and debouches into Van Diemen Gulf.

The physical setting of the project area is characterised by gently undulating terrain comprising hills/ridges of low relief, flats and valleys. The site occurs in Precambrian (Proterozoic) rocks of the Pine Creek geosyncline (the Koolpin Formation) (Minerals & Energy Group 2003). The vegetation of the area is described as low to tall open mixed eucalypt woodland and grasses.

2.2.4.2 Brocks Creek Mine

The Brocks Creek mine is located about 170 kilometres southeast of Darwin. Gold was first discovered and mined at Brocks Creek in 1872 and mining has been carried out intermittently ever since. Drainage from the project area is to the southwest into the westward flowing Howley Creek, a tributary of the Adelaide River system, where the soils are alluvial and acidic. Two small ephemeral creeks drain the immediate project area: Burgan Creek which runs between the Alligator and Faded Lily orebodies (waste dump pits), and Brocks Creek draining the site of the tailings storage.

CHAPTER 3 - Field, laboratory, genetic and statistical techniques

3.1 Field methods

3.1.1 Selection of sites

The study region includes four open-cut mines and their associated water catchments, from each of which a maximum of fifteen sites were sampled. Of these sites one third were control sites, unimpacted by any mine pollution. At the time of sampling, each of the four open-cut mines, Cosmo Howley, Redbank, Rum Jungle and Tom's Gully (Figure 2.1), were decommissioned and had 'Care and Maintenance' status imposed**.** The pollution generation mechanisms at each of the closed mines were identical in all main waste rock heaps and tailings. Mining of metallic ore deposits expose to the atmosphere pyrites and sulphide material in rock heaps. As a consequence, the sulphide will oxidise to produce acid water laden with sulphate, heavy metals and metaloids (Harris *et al.* 2003). During the intense wet seasons, tailing dams may overflow or leachate from waste rock dumps may flow passively into receiving streams, negatively affecting aquatic systems.

Initially, in the preliminary field trip (2003), eight mine sites were viewed but four were chosen based on the criteria of a strong pollution gradient, ease of accessibility, presence of a riffle or river run habitat, decommissioned status, open-cut mining method, presence of acid mine drainage and the geographical position of the mine (aiming for a large representation of the sub-tropical region of the N.T.). Together with pH, dissolved oxygen and conductivity, the presence of mineral efflorescence/salt crusts along rivers were used as an indicator of acid mine drainage (AMD) as, at a pH of 2.8-3.0, an aluminium precipitate tends to form. The positions of the sites were chosen based on topographic maps of the areas and an initial visit to gain familiarity with the river and drainage patterns. A maximum of ten sites were chosen downstream of each mine to examine the effects of mine pollution on the water bodies based on the diatom assemblages. The first few sampling sites closest to the mine pollution source were closely spaced until there was physical indication that some metals had precipitated out from the water column. This was indicated by (iron) orange floc and (magnesium sulphate) white salt crusts. Sites were deliberately spaced closely together to reduce any skewed effect of metal concentrations along the river gradient as these decrease dramatically when orange floc or salt crusts are evident. However, positioning of sites close together prior to 'drop-out' was not always possible due to constraints on site access.

Sites were sampled in May, 2004, during the recessional flow period. This period marks the beginning of the dry season when sites are most accessible and still have substantial water flow from the wet season. This period is also when the mine tailing dams are under greatest stress and

therefore show maximum impact downstream of creeks in terms of pollution concentration. Dry conditions at some sites, access difficulties and other logistical considerations prevented sampling from all 15 sites at each mine. Additionally, the presence of crocodiles (*Crocodilus porosus*) and feral pigs (*Sus scrofa*) at some sites made sampling too dangerous to attempt. Diatom and water chemistry samples from the Redbank mine were collected by an employee of the Northern Territory Department of Business, Industry and Resource Development. In total 50 sites, 17 control sites and 33 impacted sites were sampled. Impacted sites are those which were situated downstream of the pollution source, control sites were located upstream of each mine, upstream of a convergent stream or from an unimpacted stream within the same catchment.

At Rum Jungle mine ten impacted sites were sampled. The first two samples closest to the mine were taken along the Finnis River into which effluent from Whites overburden heap (Figure 3.2), and the intermediate overburden heap, flow. This creek flows into the East Branch along which seven sites are situated. The East Branch flows into the Finnis River on which the final site is situated, over 30.5 km away from the pollution source. As seen in Figure 3.1, five control sites were sampled, three from Finnis River (RJC1, RJC3, RJC4), one from Hanna's Spring (RJC2) and one from Rum Jungle Creek (RJC5).

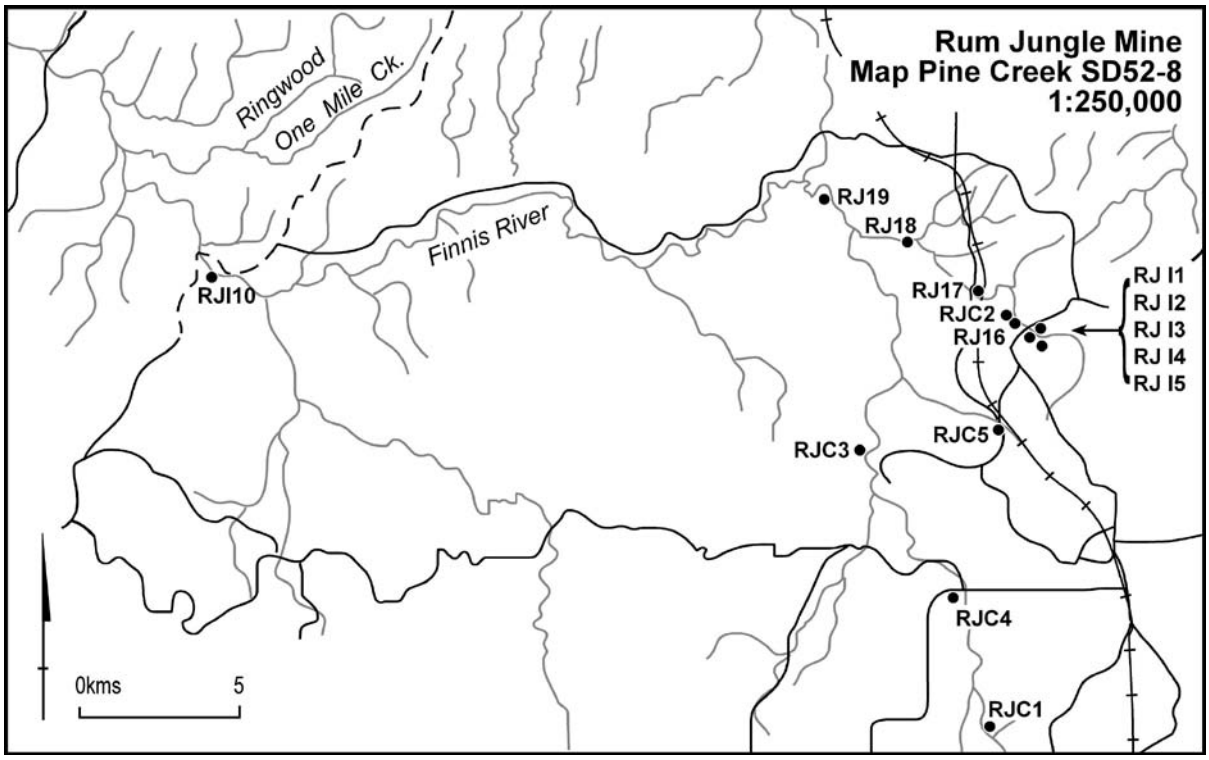


Figure 3.1. Location of monitoring sites at Rum Jungle mine.

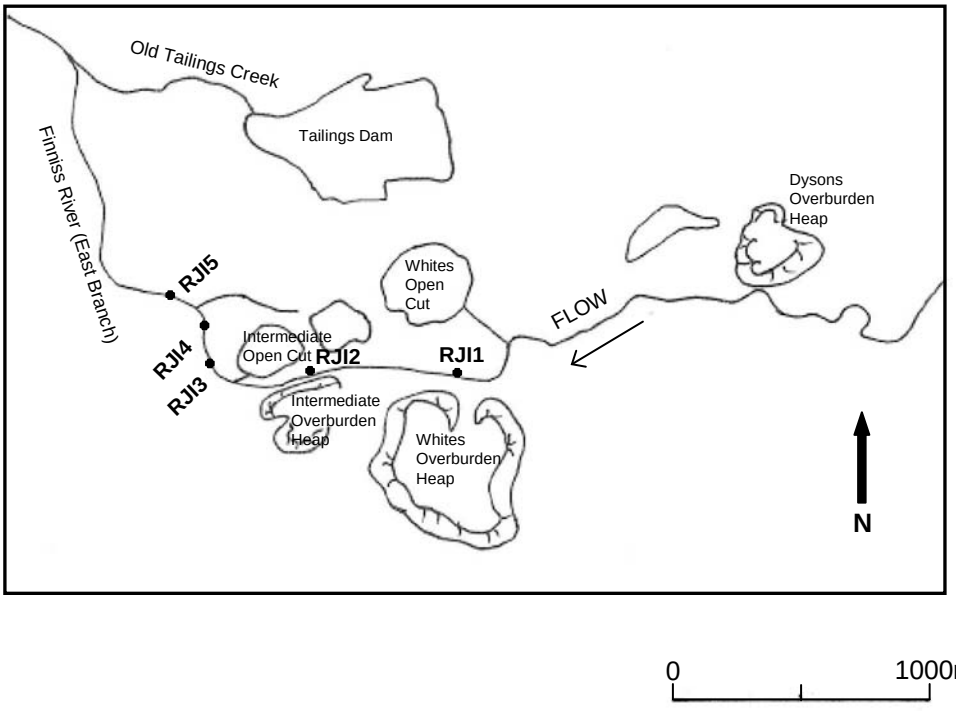


Figure 3.2. Location of monitoring sites RJI1 – RJI5 at the Rum Jungle mine site.

From Cosmo Howley mine (Figure 3.3), ten sites were sampled downstream from the pollution sources of Cosmo Howley mine and Brocks Creek mine. Creeks flow through each of these mine sites, converge and flow into Howley Creek. At Brocks Creek mine, site BKI1, situated along a branch of Howley Creek, receives passive mine waste from the Faded Lily waste dump while site BKI2, situated along Brocks Creek, receives waste from the Alligators waste dump. At Cosmo Howley mine two sites (CHI1, CHI2) are situated downstream of the Cosmo Howley wetlands, which filters water coming from the mine pit along an unnamed creek which converges with Howley Creek. Three sites, downstream from the mine water dam along an unnamed creek (which converges with Howley creek), were sampled. Site CHI7, situated along Howley Creek, represents the convergence of streams from Cosmo Howley and Brocks mine. The furthest site from any of the pollution sources is site CHI8 situated along Howley Creek, over 16 km downstream. Four control sites were sampled, one upstream of Cosmo Howley mine (CHC1) along Howley Creek and the rest at unnamed creeks off Howley Creek and at other unnamed creeks within the catchment.

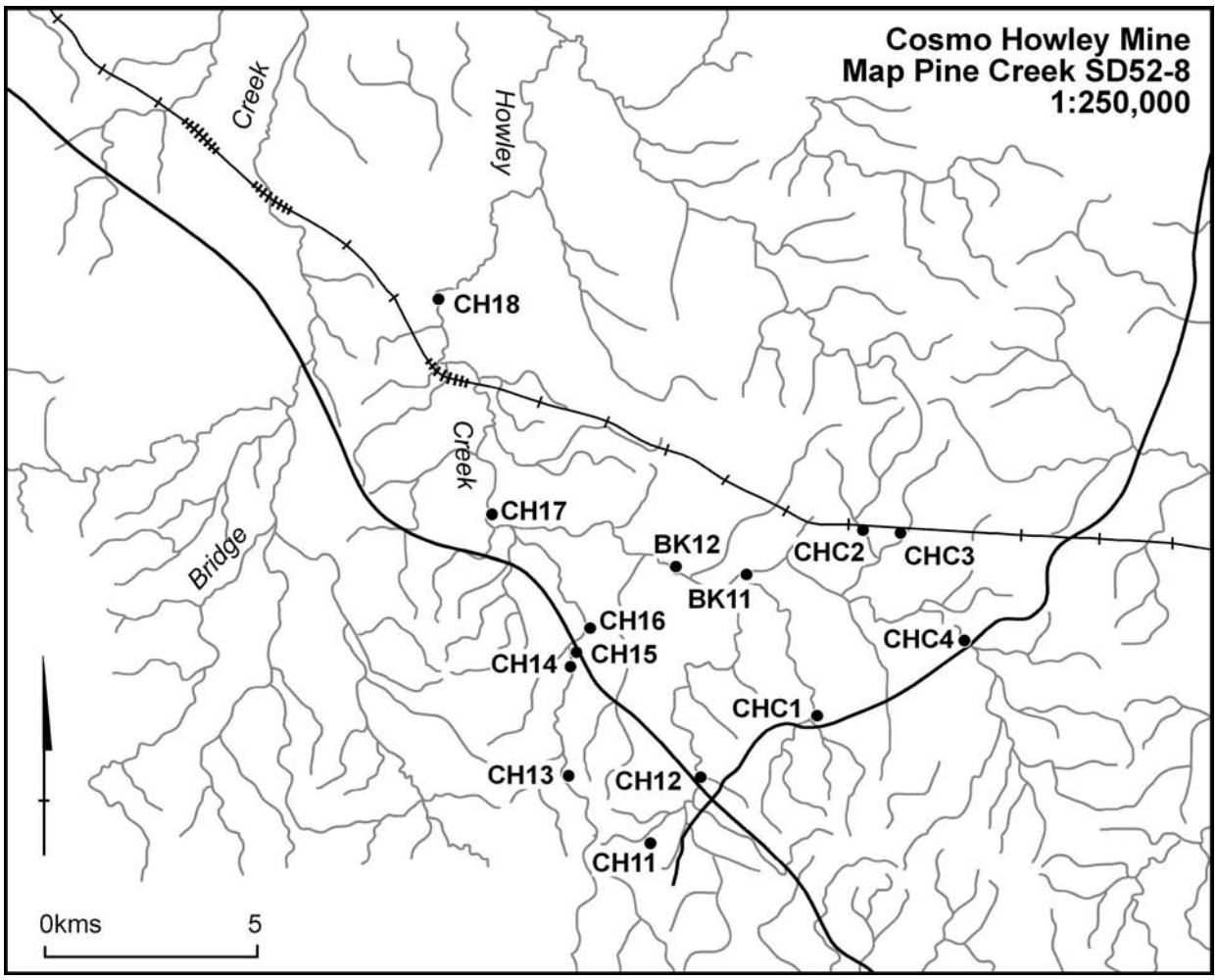


Figure 3.3. Location of monitoring sites at Cosmo Howley mine.

At Tom's Gully mine, six impacted sites were sampled downstream of the mine along Mount Bundy Creek (Figure 3.4). Site TGI1 receives water from the sulphide waste dump. Further down Mount Bundy Creek site TGI2 receives water from the evaporation pond. Site TGI4, and then TGI3, receives water flowing from an oxbow wetland into which tailings dam water flows. The site furthest away from the initial point source is site TGI6. This site is located along the Mount Bundy Creek close to the National Park boundary, 7.7 km downstream from site TGI1. Only two sites from the original five control sites were sampled due to the presence of wild pigs and the lack of water in creeks. The two control sites sampled were site TGC2, along Scot Creek, and site TGC1 at Mount Bundy Creek upstream of the Tom's Gully mine.

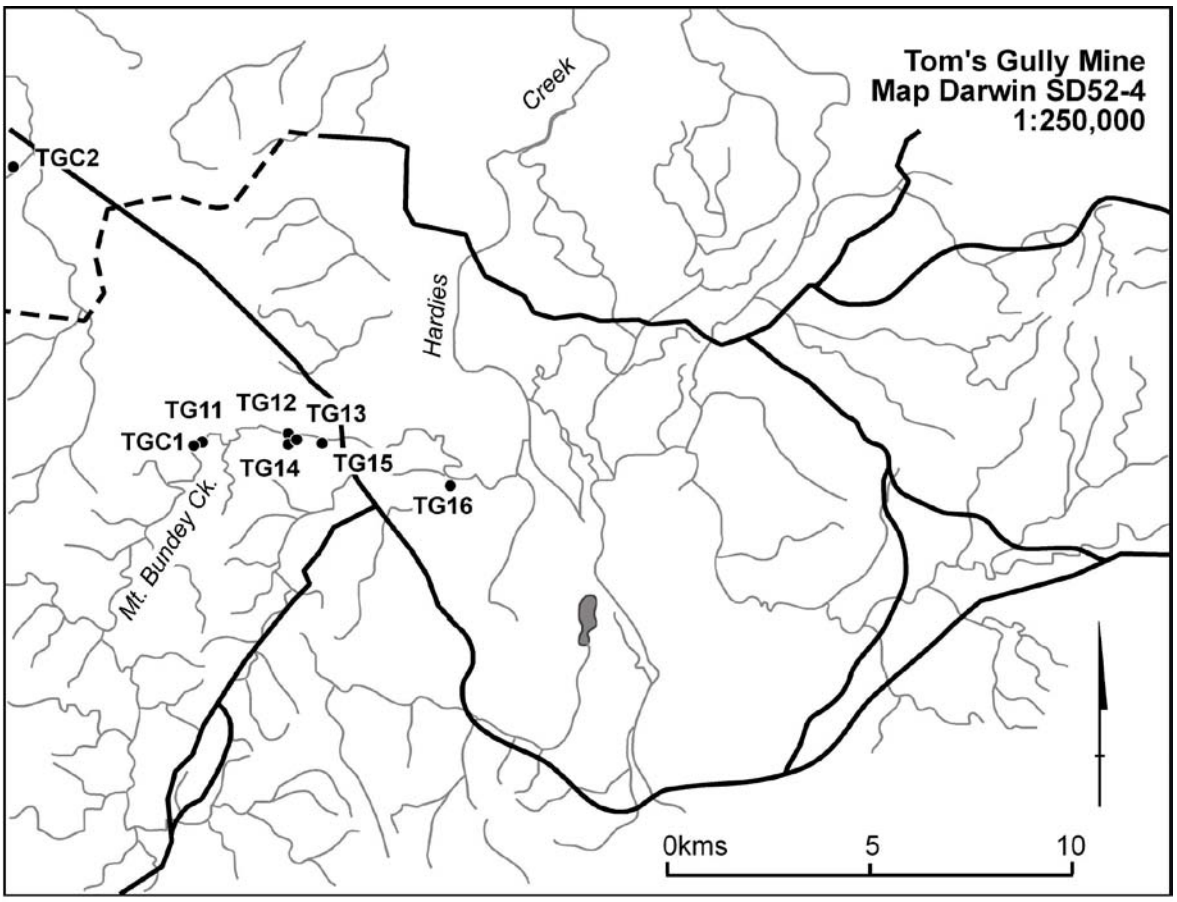


Figure 3.4. Location of monitoring sites at Tom's Gully mine.

At Redbank mine 13 sites were sampled by an employee of the Northern Territory Department of Business, Industry and Resource Development. Site HCUSEC, situated along Hanrahan's Creek (Figure 3.5), receives effluent from the tailings dam and waste rock heap. The next sites, ECDSHC and ECUS12MC are situated on Echo Creek downstream of the confluence with Hanrahan's Creek. Sites 12MCDSEC and 12MCUSSC are downstream of Echo Creek along 12 Mile Creek. Downstream of Echo Creek, sites SCDS12MC and SC@BRX are located along Settlement Creek. The site furthest downstream of the point source is site SC@BRX, over 32.5 km to the east. Six control sites were sampled, one each from 12 Mile Creek, Echo Creek and Camp Creek and three sites at Settlement Creek. Each control site was sampled upstream of the confluence with the polluted water source.

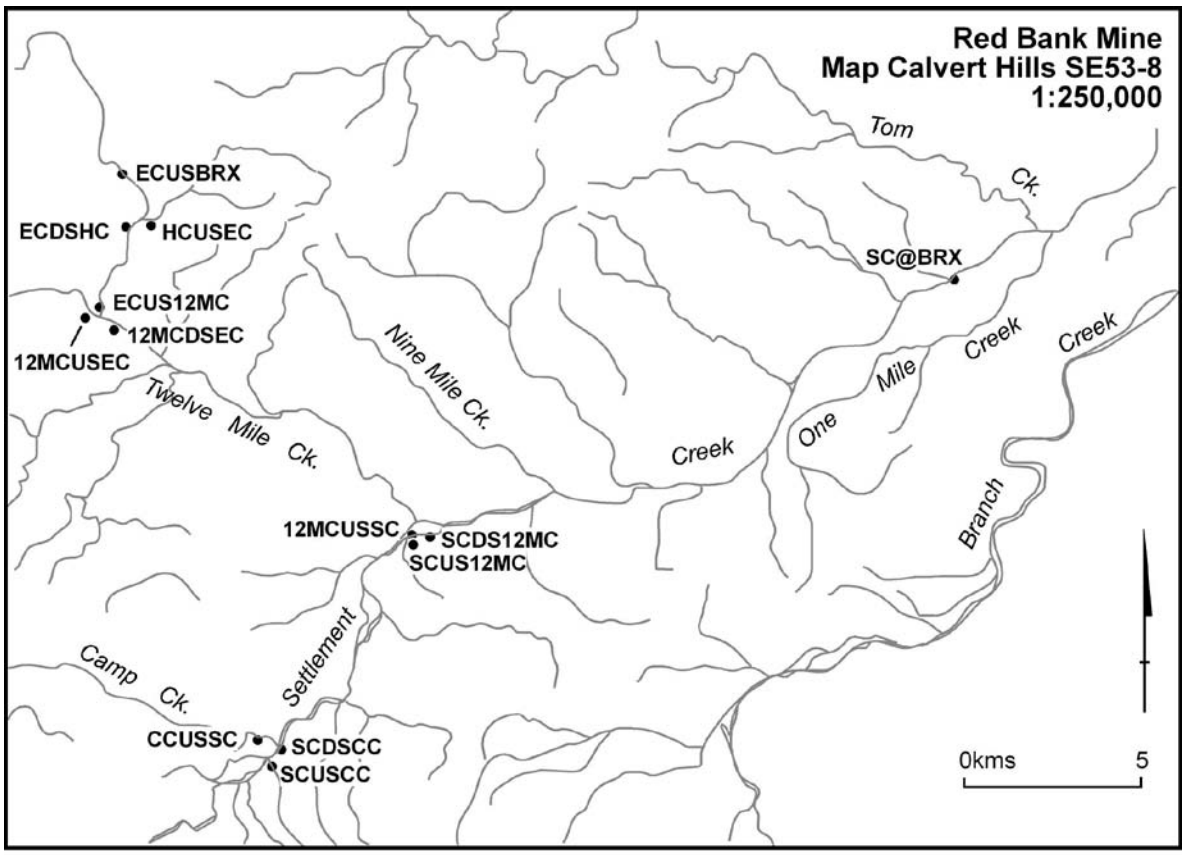


Figure 3.5. Location of monitoring sites at Redbank mine.

Each of the sites used in the dataset were assigned a unique alphanumeric code made up of three letters and two digits, for example, CHI2 or RJC3. The codes were used in the computer programs CANOCO, JMP and CALIBRATE which allow site (sample) codes of no more than eight characters. The site name is readily identifiable as the first two letters represent the name of the mine i.e. CH stands for Cosmo Howley mine. The third letter refers to whether the site is a control site or an impacted sites (I = impacted and C = control site). The numbers in the code refer to the sites positioning downstream from the point source. Sites from Redbank mine were labelled by an employee of the Dept. of Business, Industry and Resource Development and it was thought best to retain these for ease of comparison with other studies.

3.1.2. Selection and measurement of environmental variables

Diatom samples, physico-chemical site data and chemical samples were collected from creeks and rivers associated with four open-cut mines, Rum Jungle mine, Cosmo Howley mine, Redbank mine and Tom's Gully mine. A total of 24 water quality parameters were measured at each site. The water quality parameters selected (Appendix 1, 2) for analysis were those known to correlate with shifts in diatom communities (Bennion *et al.* 1994) and heavy metals were those which occurred in concentrations well above those found in surrounding waters (Cyrus Edwards, pers. comm., N.T.

Dept. of Business, Industry and Resource Development). For instance pH was chosen as it is highly influential in determining diatom species composition as it exerts direct physiological stress on diatom cells (Hall & Smol 1992). High water temperatures decrease the amount of silica in the cell walls of diatoms and affect the diffusion rates of chemicals. This leads to changes in rates of reproduction and metabolism (Sudhakar *et al.* 1994). The electrical conductivity of water (EC) provides an estimate of the total dissolved salts in aquatic systems. Nutrients, such as total nitrogen (TN), influence photosynthetic biomass and composition while nitrate (NO_3 or N) has been recorded as a limiting nutrient for autotrophic organisms (APHA, 1992). Physico-chemical data, cadmium (Cd), lead (Pb) and nutrient measurements were not included for Redbank sites as these sites were sampled by an employee from the Department of Business, Industry and Resource Development, N.T.

Measurements of pH, dissolved oxygen (DO), conductivity (EC), temperature (Temp) and turbidity (Turb) were made using a Horiba multi-parameter probe (Horiba Pty Ltd, Kyoto). The Horiba meter was placed in flowing water and was gently moved to maintain a continual supply of oxygenated water over the probe (the probe consumes oxygen whilst in use).

Nutrient and metal samples were collected in acid washed, plastic bottles. Before sampling the bottles were rinsed in creek water three times before being filled to capacity. The bottle mouth was faced upstream and submerged to a depth of about 20 cm and filled. Water samples for metals were filtered, as collected, in the field through a Millipore cellulose nitrate membrane with a pore size of 0.45 μm and stored in high density plastic bottles. Added to the water bottles was 4 ml of nitric acid. All samples were stored at less than 4 °C and were sent within 48 hours for analysis of nutrients, major ions and trace metals which was undertaken by the Australian Nuclear, Science and Technology Organisation (ANSTO) laboratory, following standard methods.

3.1.3. AUSRIVAS observations

Physico-chemical observations of stream characteristics were chosen and measured in accordance with the Northern Territory AUSRIVAS (Australian River Assessment Scheme) guidelines which are part of the National River Health Program (NRHP). Certain environmental variables such as width between banks and detailed observations of the riparian zones and erosion, although part of the ASURIVAS protocol, were not measured. These factors were either deemed of low importance to diatom communities or were simply too difficult to measure because of dangers from wildlife and time constraints. Stream reach is defined by AUSRIVAS as 100 m around the sampling point. In this study, because of the proximity of the sites, the river characteristics were observed in a five meter radius around the sampling point. General stream conditions (i.e. water clarity, surface oils etc.) were measured by selection of categories (1-4) outlined in AUSRIVAS. Field sheets used in

documenting physico-chemical observations are included in Appendix 3. A full list of the environmental parameters measured is included in Appendix 1.

3.1.4. Diatom sampling

Benthic diatom samples were collected at each site from submerged rocks at a consistent depth of less than 15 cm, from riffle/run sections of streams. Diatoms were removed from three rocks by scraping with a cleaned knife. The scraped material was combined and preserved in ethanol. Similar to findings of Watanabe *et al.* (1986), epilithon (rock) (the most common substratum sampled (Gell *et al.* 1999)) substrata were chosen as they were present at all sites, whereas epipsammon (sand), epiphyton (plants) and epipelon (mud and silt) substrata were absent at many sites.

Samples for genetic analysis were scraped from three rocks, combined and stored in cool dark conditions before being sent to Gent University (Belgium) where they were cultured. Samples were collected from each of the 50 sites.

3.2. Laboratory methods

3.1.3. Chemical Analysis

The analysis of the metals, aluminium, cadmium, calcium, cobalt, iron, potassium, magnesium, manganese, sodium, nickel, lead, zinc with concentrations above 500 µg/L were analysed at the ANSTO laboratories with Inductive Coupled Plasma Atomic Emission Spectrocscopy (ICPAES) and followed the methods outlined in the following references:

- Standard Methods for the Examination of Water and Wastewater., 18th Edition 1992, APHA, Washington, DC.
- Test Methods for Evaluating Solid Waste, Physical/Chemical Methods, SW-846, 3rd Edition, Final Update II, 1995, US Environmental Protection Agency, Washington DC.
- CRC Handbook of Inductively Coupled Plasma Atomic Emission Spectroscopy., Asha Varma, 1991, CRC Press, Boca Raton, Florida, USA.

For metals with concentrations below 500 µg/L analysed with the Inductive Coupled Plasma Mass Spectrometry (ICPMS) the following analytical methods were followed:

- US EPA methods 6020 and 200.8.

- HP 4500 Chemstation Operator's Manual

The measurement of alkalinity by ANSTO laboratories followed the method outlined in the following reference:

- Eaton A.D., Clesceri L.S., and Greenburg A.E. (1995). Standard methods for the Examination of Water and Wastewater. Method 2320 page 2-25

Nutrients were analysed by the ALS Laboratory Group Environment Division and followed the methods outlined in 'Standard Methods for the Examination of Water and Wastewater, 21st Edition 2004, APHA, Washington, DC:

Nitrate and Nitrite (N);

- Combined oxidised Nitrogen (NO_2+NO_3) is determined by Cadmium Reduction and direct colourimetry by FIA.

Total Kjeldahl Nitrogen (TKN);

- 100 mL water samples are digested using a traditional Kjeldahl digestion followed by determination by FIA.

Total Nitrogen (TN);

- TN is combined TKN and nitrate and nitrite

Total Phosphorus (TP);

- This procedure involves sulphuric acid digestion of a 100 mL sample to convert all phosphorus to orthophosphate. The orthophosphate reacts with ammonium molybdate and antimony potassium tartrate to form a complex which is then reduced and its concentration measured at 880 nm using FIA.

Total Organic Carbon (TOC);

- The automated TOC analyser determines total and inorganic carbon by IR cell. TOC is calculated as the difference between TC and TIC.

3.2.2. Diatom preparation techniques

In the process of collecting diatoms from substrata, organics, clastics and carbonates are inevitably included in samples. In order for identification to be achieved, this material needs to be removed. The preparation and analysis of diatom samples followed standard techniques (Battarbee 1986). To digest the organic material within the samples, 30 ml of dilute hydrogen peroxide (10% H_2O_2) was added to the samples. To accelerate the reaction, samples were heated at simmering point. The samples were then washed three times with distilled water and left for at least six hours each time to allow the disaggregated samples to settle before decanting. This process was repeated with dilute hydrochloric acid in order to dissolve any carbonates.

The resulting siliceous material is dried on coverslips and mounted on slides in a resin (naphrax) for counting with a light microscope (Nikon Eclipse E600) with Differential Inference Contrast at 1500x magnification using immersion oil.

3.2.3. Diatom counting procedure and taxonomy

For each sample 400 diatom valves were counted and identified along random transects at a magnification of 1500x using a Nikon Eclipse (E600) microscope. A total of 400 valves were counted per slide for each site. This number was determined by the relationship between the number of species identified and counting effort (section 4.5) for 600 valves at Rum Jungle sites. This number concurs with Battarbee (1986) and Gell *et al.* (2002) who needed to count 400 valves from the Katherine and Daly Rivers to obtain 90% of species.

Diatom valve identifications were made on the basis of comparison of diatom valve morphology with photomicrographs and descriptions contained in diatom taxonomic books. Diatom taxonomy and nomenclature were based primarily upon Hustedt (1937-1939); Krammer (1992-2002); Krammer and Lange-Bertalot (1986-2000); Lange-Bertalot and Krammer (2000-2002); Lange-Bertalot (1996-2001), Lange-Bertalot and Metzeltin (1996) and Reichardt (1999). Species were identified, where possible, to varietal level. Special reference was given to south-east Asian, South American and Australian tropical floras (i.e. Thomas 1983; Moser *et al.* 1995; Lange-Bertalot 1998; Moser *et al.* 1999).

A number of diatom taxa could not be easily identified as published species. These taxa fell into two categories. Firstly, valves which confidently could only be identified to genus level. Taxa in this category were assigned epithets such as *Nitzschia* species (sp.) 1. Valves which generally had morphological characteristics similar to recorded species, but had features (such as higher striae density or more capitate valve ends) which precluded them from being confidently ascribed to a

published taxon were labelled “cf.” to acknowledge that it could be a published species but that some features were different.

After identification each taxon was given a code name as the computer programs CANOCO and CALIBRATE require taxon names of less than eight characters. The short code name reduces crowding on ordination plots. In the coding system used in this study, the code for a given taxon is generally the first two letters of both the species and genus e.g. *Nitzschia nana* becomes Ni.na. When taxa were identified to varietal level then a “v” and the first letter of the varietal epithet was added. When taxa had the same three letters for both genus and species, then taxa which would otherwise have had the same code name had an additional letter. In the development of codes the abbreviation “cf.” was not used. Where taxa could only be identified to genus level, the first two letters of the genus were used together with an abbreviation of species (sp) followed by a numeral, e.g. Achsp1. A list of taxa recorded in this study and their code names in included in Appendix 4.

3.2.4. Selection of Nitzschia palea-like cells for culturing

Using an inverted microscope, living cells of *Nitzschia palea*-like species were found in eight samples (Table 3.1). Each of the other 40 samples either contained no *Nitzschia palea*-like frustules or the frustules were empty or very rare. Frustules selected from samples for culturing were those with typical *N. palea* characteristics, but which varied morphologically.

Table 3.1. Geographic and ecological information of culture sample sites.

Mine Location	Site	Sample	pH	Copper (mg/L)
Cosmo Howley (CH)	**BKI2**	Impacted	6.9	0.010
	CHI8	Impacted	7.0	0.004
Rum Jungle (RJ)	**RJI1**	Impacted	3.4	5.990
	RJI5	Impacted	6.2	0.200
	RJI10	Impacted	7.0	0.003
Redbank	**SCUSCC**	Control	8.4	0.002
Tom's Gully (TG)	**TGI1**	Impacted	6.3	0.005
	TGI6	Impacted	6.8	0.002

3.2.5. Culturing

Using an inverted microscope and a pipette *N. palea* cells were taken from the live samples and transferred into a Petri dish with culture medium. Low magnification (20 µm) was used to select the cells based on the shape of the frustule, consequently, some error incurred during cell selection. The medium used was based on the WC medium (Guillard & Lorenzen 1972) recipe. A total of 69 clonal cultures were established by isolation of single cells by micropipette into Repli dish wells containing culture media (47 of the 69 were used as some did not grow or were too similar or too

sparse). Each culture was given a study name i.e. (BKI2)B1, the letter and numerals in the brackets indicating the mine and site from which the sample was collected. Subsequently, each monoclonal culture was transferred to polystyrene 50 mm Petri dishes with 15-20 ml of culture medium and maintained in an incubator at 18 ºC with a 12:12 hour light:dark period and 25-30 µmol photons m^{-2} s^{-1} from cool-white fluorescent lights. Reinoculation of cells into fresh medium was performed approximately every three weeks. In addition, subcultures of the clones were stored in small plastic containers, with 10 ml of culture medium, in 10:14 h light: dark period but at a lower temperature (6-7 ºC) and illumination (c. 5 µmol photons m^{-2} s^{-1}), where the cells multiplied much more slowly.

3.2.6. Image capturing of cultured N. palea cells

For image capturing and analysis of live cells, cultures were initiated in Petri dishes in which cover slips had been placed. A cover slip could then be removed, wiped clean on one side and mounted on a drop of the medium (resin was used to stop the culture and medium from escaping) for examination. For light microscopy (LM) and scanning electron microscopy (SEM) analysis of cultures, frustules were killed and cleaned by oxidation with hydrogen peroxide (3 ml of H_2O_2) and nitric acid (1 ml HNO_3 at 69%). They were then washed three times with distilled water before being mounted on a slide with Naphrax for LM. For SEM analysis, two to three drops of culture samples were air-dried onto aluminium stubs. The stubs were sputter- coated with approximately 20 nm of Au-Pd. The most effective oxidation method for the greatest number of culture samples was found to be hydrogen peroxide. Light microscopical studies were carried out using a Zeiss Universal microscope (x100 oil immersion objective) equipped with a Hamamatsu (VIP III) digital camera connected to a computer. Measurements of linear dimensions and stria densities of valves of *Nitzschia* were made with the aid of Studio Lite software. Scanning electron microscopy was performed using a JOEL JSM5600LV (JOEL, Tokyo, Japan). The Image J program was used to document size and fibulae density. Up to 20 specimens per culture were photographed and measured. Scanning electron micrographs of some cultured frustules were not suitable as the valves were found to be weakly silicified.

3.2.7. Morphometric measurements of N. palea clones

The following parameters were assessed: valve length, width, number of fibulae and striae in 10 µm along the length of the frustule. Measurements were made mostly from light microscope (LM) images with the program ImageJ 1.38 (Abramoff *et al.* 2004). Only a few electron microscopical (SEM) images were useful for measurements because images were predominantly of valve exteriors and oxidation was not effective enough for high resolution viewing as undissolved material obscured valvular details.

3.2.8. DNA extraction, amplification and sequencing of N. palea clones

The hypervariable region of nuclear small subunit rDNA (SSU rDNA) has proved useful for examining species interrelationships in *Skeletonema* (Sarno *et al.* 2005), and this technique was used for this study of *Nitzschia palea.* For rDNA sequence analysis, cells were harvested from a culture in exponential growth phase and concentrated by centrifugation. DNA was isolated and extracted from the culture cells following the protocol developed by Muyzer *et al.* (1998).

The small subunit (SSU) rDNA gene was amplified by PCR, using the primers P4 and P21. Genomic DNA was amplified in a 50 µL reaction mix containing 1-5 µL of template DNA, primers at a concentration of 0.5 µM, deoxynucleoside triphosphates at 200 µM each, bovine serum albumin (BSA) at 0.4 µg μL^{-1}, 5 µL of 10 X PCR buffer [Tris-HCl, $(NH_4)_2SO_4$, KCl, 15 mM $MgCl_2$, pH 8.7 at 20 ºC; "Buffer 1", Applied Biosystems, Foster City, USA] and 2.5 units of Taq polymerase (AmpliTaq, Perkin-Elmer, Wellesley, USA); mixtures were adjusted to a final volume of 50 µL with sterile water (Sigma, St-Louis, USA). The polymaise chain reaction (PCR) amplification conditions were as follows: one initial denaturation of 94 °C for 5 min; followed by 35 cycles, each consisting of 1 min at 94°C, 1 min at 55 °C, and 1 min at 72 °C; and a final step of 10 min at 72 °C. Amplified products were purified using Qiaquick PCR purification kit (Qiagen, Hilden, Germany) and directly sequenced with the aid of a Big Dye™ Terminator Cycle Sequencing Ready reaction Kit (Applied Biosystems, Foster City, California), using the same forward and reverse primers as applied for PCR amplifications. Sequencing was done with capillary sequences (AB13100, Applied Biosystems, Foster City, California). Sequences ran in both directions were overlapping. The 20 cultures sequenced are shown in Table 3.2.

Table 3.2. Sequenced cultures of *Nitzschia palea*-like cells

SSU
(SCUSCC)D2
(SCUSCC)F4
(SCUSCC)B3
(SCUSCC)B2
(TGI16)E5
(BKI2)D6
(CHI8)B7
(CHI8)B9
(CHI8)E9
(CHI8)D7
(CHI8)B8
(TGI1)C2
(TGI1)C3
(TGI1)B2
(TGI6)G4
(TGI6)G5
(RJI1)D4
(RJI10)C44
(RJI5)F10
(RJI5)E7

3.2.9. Image capturing and descriptions of teratological diatoms

From each site slide, 200 diatom frustules were counted and observed using a using a Nikon Eclipse (E600) light microscope (LM) (x1500 oil immersion objective) equipped with a Nikon digital camera (Q imaging 3.3RTV) connected to a computer. The imaging software program used was Olysia bioReport (5.0) 2004 Soft Imaging System GmbH. LM pictures of each frustule were used for measurements of frustule length, width and fibula count for each frustule, determined with the ImageJ 1.38 computer program. Scanning electron microscopy (SEM) was performed using a field emission Philips (XL30) SEM. Morphologic descriptions of frustules followed the terminology used by Krammer and Lange-Bertalot (1986-1991).

3.3. Statistical methods

3.3.1. Introduction

Multivariate statistical analyses were used to determine the distribution of species and environmental variables and to determine the relationship between both. Forty chemical and physiological variables were measured and integrated into the statistical analysis to explain the diatom species variance (distribution).

3.3.2. Data screening

Data screening was applied to reduce multivariate data into a manageable number of variables. The species and environmental data and samples were screened to identify and eliminate redundant environmental variables, species and outlier samples. Diatom taxa not occurring at every site or having a relative abundance of greater than or equal to 1% were eliminated from the dataset. A matrix of correlation produced in JMP 5.1 (SAS Inc. 2003) was used to determine relationships between environmental variables and the statistical significance of these correlations.

Additionally, CALIBRATE 1.3 (Juggins 2003) was used to determine which environmental variables had a skewed distribution. Variables which were proven to have a skewed distribution, such as copper, were normalised using log-transformations. All metals, as well as sulphate (SO_4), alkalinity ($CaCO_3$) and electrical conductivity (EC), were log transformed.

3.3.3. AquaRisk, metal ranking and bioavailability

The measured and modelled water quality data from each of the four mines were used in AquaRisk, (version 3) (Twining 1999) in conjunction with National Water Quality Guidelines (2000), to determine metals of potential concern whose concentrations exceed the ANZECC guideline values. AquaRisk was employed to determine the bioavailable fraction of each metal. The data comprised the filtered heavy metals, cobalt (Co), cadmium (Cd), copper (Cu), aluminium (Al), zinc (Zn), manganese (Mn), nickel (Ni) and lead (Pb). In addition information was included on major cations, calcium, magnesium, sodium and hardness, together with pH, temperature, sulphate (SO_4), conductivity and total organic carbon (TOC) which were needed to determine bioavailability of metals. These parameters represent only a proportion of those that are likely to reduce bioavailability of metals, the final estimates are higher than in reality, and hence allow for a reasonably conservative risk assessment (Twining 2002).

Tier 1 analysis of AquaRisk is a screening level analysis by which the chemicals of potential concern (COPCs) are identified and ranked. This is done by comparing the estimated maximum value of the observed concentration of the set of water quality data with National Water Quality Guidelines (WQG) and identifies and ranks each metal on the degree to which it exceeds the water quality guideline levels. Hazard quotients are calculated for each of the chemicals of interest as follows:

Hazard Quotient = Maximum Concentration / WQG concentration.

If the hazard quotient is greater than 1.0, the chemical is identified as a COPC.

Bioavailable fractions (the concentrations of the metal species present that are able to cross biological membranes) of metals are estimated using a geochemical speciation code. This code, MODPHRQ (a modular version of HARPHRQ, Brown *et al.* 1991), is part of AquaRisk and calculates the complex array of chemical species formed by various metals in water of a given chemistry based on knowledge of pH and concentrations of major ions and total organic carbon. To facilitate these calculations, detailed information of the water sample (beyond the metal concentration of interest) is required. It is necessary to ascertain how the organic and inorganic complexants in the water sample affect the bioavailability of the metal.

Once the bioavailable fraction of each metal is calculated by AquaRisk, the bioavailable water quality data for each metal are fitted to both lognormal and Burr III distributions. Burr Type III distribution has been adopted by ANZECC Committee on water quality for the derivation of water quality guidelines (Twining *et al.* 2002). The distributions are tested for goodness-of-fit to assess whether the measured concentrations follow the fitted distributions. Biological dose response data are then accessed for contaminants that exceed the guideline values. It is possible to filter the data to suit the habitat and species of the area of interest. For the data selected, a curve is fitted for each metal assuming a lognormal or log-logistic distribution. The goodness-of-fit of each distribution is derived using the data available and the measured biological-effect concentrations can be compared graphically against the distribution. Once the distribution parameters and their uncertainties have been determined, critical values based on the selected ecotoxicology data are also determined. These values are the median hazardous concentration affecting 5% of the species (HC5) and its 95% lower confidence limit (HC5; 95).

3.3.4. Multivariate Analysis

Multivariate statistical techniques were used to explore the relationships between environmental variables and diatom taxa to identify variables most likely to produce the most reliable inference models.

3.3.4.1. Detrended Correspondence Analysis (DCA)

A detrended correspondence analysis (DCA) was tested to determine if the diatoms followed a unimodal or linear distribution (length of gradients) and to determine the amount of variation in the species data (ter Braak 1985). DCA is an eigen-analysis based ordination technique derived from

correspondence analysis but detrends the data to counteract the arch effect which is common in correspondence analysis (ter Braak & Prentice 1991). A DCA plots samples according to their similarities and dissimilarities in their diatom assemblages, graphically representing variation in the biological data.

3.3.4.2. Canonical Correspondence Analysis (CCA)

As the diatoms' response (in the DCA) was unimodal a canonical correspondence analysis (CCA) was used to explore the relationships between diatom assemblages and environmental variables (ter Braak 1987). Forward selection of the environmental variables, based on a Monte Carlo permutation test with 999 permutations ($p < 0.05$), was used to select environmental variables which were most related to diatom communities. According to Gasse *et al.* (1995) a Monte Carlo permutation test, which determines the influence that a variable exercises over the diatom data, is the best test of whether an environmental variable can be used. They argue that any variable that explains a significant proportion of the species data can be reliably used in a transfer function. The suitability of a CCA can be partly assessed by the eigenvalues for each axis. Eigenvalues measure the importance of each of the axes (values between 0 and 1) for the explanation of the dataset; they indicate how much of the total dispersion of all data, within the multi-dimensional space, is determined by the particular CCA axis (see ter Braak 1987). Additionally, species-environment correlations indicate how well the recorded environmental data explain the floristic structure of the dataset. Cumulative percentage variance of the species data or of species-environment correlation gives the cumulative percentage of the data dispersion depicted on the particular CCA axis. Diagrams generated by CCA show ordination axes scores for samples and for species, and vectors for the environmental variables. The vectors indicate the direction of maximum variation of each environmental variable, and their length is directly proportional to their influence in explaining variation in the dataset. All multivariate analyses were performed using CANOCO version 4.0 (ter Braak & Šmilauer 1998).

3.3.4.3. Detrended Canonical Correspondence Analysis (DCCA)

Prior to the development of the transfer function, a series of detrended canonical correspondence analyses (DCCA), constrained solely to each environmental variable, were performed, firstly to determine the species gradient lengths with respect to the selected variable, and to determine if unimodal or linear based inference models were most appropriate (Birks 1995). Secondly, the DCCA ordinations were used to establish the significance of selected variables (taken individually) in the explanation of diatom distributions and, therefore, to choose the variables that are likely the

best predictors of species changes (Birks 1995). This was achieved using the computer program CANOCO version 4.0.

3.3.4.4. Weighted averaging regression and calibration of diatom data

Quantitative calibration models (transfer functions) are developed for the modeling of selected variables from diatom assemblages from the dataset. The Weighted Average Partial Least Squares (WA-PLS) technique was used when the DCCA determined which environmental variables are most strongly (and significantly) correlated with diatom taxa distribution. The relationships between variables and diatoms must be unimodal. Once this is established the statistical technique of weighted averaging (WA) regression is used to determine the environmental optima of individual diatom taxa. WA regression assumes that the environmental optimum of a taxon occurs at, or close to, its maximum abundance (ter Braak & van Dam 1989), and that this optimum can be accurately estimated by a simple abundance – weighted average of the taxon abundance along the environmental gradient (Anderson 1997). The WA-PLS technique takes into account both the unimodal response of diatoms along environmental gradients, and the information contained in the residuals, thereby diminishing bias (ter Braak & Juggins 1993). The performance of the models was assessed using the apparent coefficient of determination (R^2) between measured and diatom-inferred values, and the apparent root mean squared error (RMSE). Estimates, based on jackknife resampling, $R^2_{jackknife}$ and $RMSE_{jackknife}$, were used as they represent more realistic measures of the predictive power than the apparent R^2 or RMSE. The analysis was carried out with the program C2 version 1.3 (Juggins 2003).

The weighted averaging (WA) technique adopts the simple principle that a taxon relative abundance increases to an optimum before declining, in a unimodal fashion, across environmental gradients. WA-PLS techniques (ter Braak & Juggins 1993) were implemented to improve WA. WA-PLS uses structures in the diatom-environment residuals to increase the predictive power of WA diatom models, particularly by shifting the optima of taxa that we abundant at sites with high residuals. In the PLS iteration, components are selected which maximally correlate diatom-environment residuals with variable x. ter Braak (1987) states that all taxa tend to occur over a characteristic, but limited, environmental range, with their environmental optimum occurring within this range, with this relationship being termed the Gaussian response.

There are many methods for selecting outliers, the removal of which results in greater predictive power of models (ter Braak & Juggins 1993; Gasse *et al.* 1995, Jones & Juggins 1995). Regardless of the method used to select outliers, the process of outlier removal results in a loss of ecological information. In this study outliers were removed on the basis of poor ecological rather than

numerical fit, and so included samples which were dominated by taxa which did not occur in other samples in the dataset.

3.3.5. Statistical Analysis of morphometric data of cultured diatoms

Morphometric data of cultured *N. palea*-like cells was explored using principal components analysis (PCA) using JMP 5.1 (SAS Inc. 2003). A hierarchical cluster analysis based on Ward's method was used in order to group cultured individuals based on their different morphological characteristics (length, width, fibulae) (JMP 5.1, SAS Inc. 2003). Once clusters were made, their reliability was tested by running a discriminant analysis. The discriminant analysis tested whether the morphological variables chosen were each significantly responsible for significant differences between clusters. Discriminant analysis also tests whether the individuals were correctly grouped.

3.3.6. Phylogenetic analysis

The resulting SSU sequences of the clones were used to search the GenBank database (BLAST), and the closest matches were downloaded and aligned (by Griet Casteleyn, Gent University, Belgium) with the clones to provide reference sequences for phylogenetic analysis. Sequences were aligned using Clustal X (Thompson *et al.* 1999) and edited manually in BIOEDIT (Hall 1999). The dataset, including the sequences of the twenty clones with the 35 taxa downloaded from GenBank (ten were known *Nitzschia palea*), was analysed using Bayesian analysis.

Bayesian inference (BI) was performed with MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003). Selection of the model of nucleotide substitution that best fits the data was performed with PAUP/MrModeltest 1.0b (Nylander, 2004), under the Akaike information criterion (AIC). For the SSU rDNA alignments, a general time-reversible model with rate variation across sites and a proportion of invariable sites (GTR+I+G) was selected. All analyses were run for three million generations, with two parallel runs of four chains each, the default settings for MrBayes 3.1.2. Summary statistics and a phylogenetic tree were generated using the last one million generations, well beyond the point at which convergence of parameter estimates had taken place. The phylogenetic tree was rooted with *Navicula phyllepta.* Distance analyses were performed with parsimony as the optimality criterion. Bootstrapping was performed for the maximum parsimony (MP) and distance analyses. The mean distance and standard deviation were calculated from the matrix with Microsoft Excel.

3.3.7. Criteria for selection of diatom indicator species

For the selection of diatom indicator species, taxa with a strong statistical relationship to the environmental variable of interest, a well defined optimum, and a narrow tolerance to the variable of interest were selected, adapting criteria outlined by Pienitz *et al.* (1995).

These criteria comprise;

1. The taxon must occur at ≥ 50% of highly impacted or control sites.
2. Presence and absence: the species is present at one or two sites at an abundance of ≥ 50%.
3. The taxon must have an optimum and tolerance range within the mean plus the standard deviation range of the variable from the pollution category.
4. More than 10% of the variance of the taxon must be explained by the environmental variable.
5. Species response curves should be close to unimodel because the calculation of optima is based on weighted averaging.

To fulfil the first two criteria, the diatom dataset was screened for occurrence and abundance. Any species with an occurrence of less than six was excluded from analysis except if the species had an abundance of greater than 50% at one site or more. The site value of seven was chosen based on the DCA separation of sites in section 6.2.1.

For criterion three, the tolerances of the taxa to pH and copper were calculated by weighted averaging (WA) with each tolerance adjusted for the effective number of occurrences (N2) of the taxon. A single indicator value was obtained for each diatom species by using the optimum derived from weighted average inference models. Weighted average modeling was conducted with the CALIBRATE program version 0.3 (Juggins & ter Braak 1993).

A canonical correspondence analysis (CCA) [CANOCO 4.0 program (ter Braak & Šmilauer 1998)] constrained solely to each variable, pH and copper, was run to determine the significance of each in accounting for the diatom species variance in the dataset. The proportion of variance of each taxon explained by CCA axis 1 (ter Braak 1990b) was examined to fulfil criterion four. If the proportion of species variance explained by the environmental variable was less than 10% the variable was not used for analysis

For each species which satisfied criteria 1, 2, 3 and 4, species response curves were developed in Canodraw (4.12). For evaluation of species response curves, criterion five, species responses to pH and Cu, were studied by fitting a parametric regression model using a Generalised Linear Modelling (GLM) approach. GLM is an extension of the classical linear regression model. A

significant fit to a quadratic model implied a symmetric bell shaped (Gaussian) response curve. A statistically significant linear model indicated monotonically decreasing or increasing species abundance along a pH gradient. A generalised linear model however, does not imply a straight-line response. It is linear in that it consists of a linear combination of explanatory variables and can describe non-linear relationships between response and explanatory variables (Potapova *et al.* 2004). With Canodraw models were created with Poisson error distributions and log link functions that are appropriate for count data, and tested for significance of parameter estimates. When both linear and quadratic models were significant, the most parsimonious model was selected by estimating the drop in residual deviance, which occurred after adding the quadratic term. If the drop was larger than 3.84, which equals $\chi 0.05$ at one degree of freedom, then the quadratic model was considered as providing significant improvement over the linear model in explaining species distribution (ter Braak & Looman 1995).

3.3.8. Diversity indexes

The Simpson's index of diversity, the Shannon-Wiener diversity index (Shannon & Weaver 1949) and species richness values were calculated for each site with species having more than 1% occurrence at at least one site. The higher the value of the indices, the more diverse is the site. Copper concentration and values of Simpson's and Shannon-Wiener indices were log-transformed because of negative values. The formulae of each index are as follows;

3.3.8.1. Simpson's index of diversity (D)

The Simpson's index of diversity is considered a dominance index because it weights towards the abundance of the most common species. The Simpson's index is the probability that two individuals, randomly selected from a sample, will belong to the same species and is a measure of the species diversity in a community. The value of this index ranges between 0 and 1, with the greater the value, the greater the sample diversity.

$$1\text{-}D=\sum_{i=1}^{s} Pi^2$$

Pi = is the fraction of all organisms which belong to the *i*-th species

3.3.8.2. Shannon-Wiener diversity index (H);

The Shannon-Wiener Diversity Index is one of the most widely used species diversity indices for examining overall community characteristics. The index utilises taxon richness and their relative

abundance. It assumes that all species are represented in a sample and that the sample was obtained randomly. The value will always result in a diversity value (H') ranging between 0 (indicating low community complexity) and 4 (indicating high community complexity). The Shannon-Wiener diversity scores were calculated for each sample as;

$$H = \sum_{i=1}^{s} Pi * \ln Pi$$

Pi = is the proportion of individuals found in the *i*th species and ln is the natural logarithm.
S= species richness

3.3.8.3. Species Richness

Species richness is the number of species in a sample.
= Total N

3.3.9. Relationships between diversity indexes, pH and copper

There are two obvious factors affecting the diatom distribution in mine affected streams, these are acidity and heavy metals. Copper and pH were identified in section 6.2.1. as the two environmental variables that best explain species variance. Hence, copper and pH were chosen to represent the heavy metal pollution of the streams and to assess whether changes in diatom diversity can be used to infer heavy metal pollution. The relationship between the diversity indices with pH and copper was explored using bivariate regression analysis in JMP 5.1 (SAS Inc. 2003). Bivariate analysis explains how the distribution of one continuous variable is related to another variable. Linear regression fits a straight line through the data points using a partial least squares regression.

Sample outliers were determined by creating a distance plot using jackknifed distances in JMP. The distance for each observation is calculated with estimates of the mean, standard deviation and correlation matrix. Any sample which plots above the distance of 2.5 was considered an outlier and excluded from bivariate analysis.

3.3.10. Statistical analysis of teratological diatoms

The relationship between teratological valves and pH and copper was explored using bivariate regression analysis in JMP 5.1 (SAS Inc. 2003). Bivariate analysis explains how the distribution of one continuous variable is related to another variable. Linear regression fits a straight line through

the data points using a partial least squares regression. The relationship between the variables was again analysed in SigmaPlot 6.0 (SPSS Inc. 2000) using rational parametric 2 analyses.

CHAPTER 4 - Water chemistry and diatom flora descriptions

4.1 Introduction

Measurements of water samples, diatom and physico-chemical samples were taken from 50 sites in the Northern Territory. This section outlines the chemical gradients of metals and nutrients at each mine. The multivariate dataset is reduced to a manageable number of variables by screening the environmental variables for significant correlations, using a matrix of correlation. The metal component of the dataset is further added to by the determination of the bioavailable fraction of the heavy metals. Diatom genera and species abundances and dominance in the dataset are described. Lastly, this section includes an iconograph of the 52 most abundant species in the dataset.

4.2 Chemical gradients

In all 40 variables were measured and included in the dataset (raw data is presented in Appendix 1 and 2). Of these variables, 23 were chemical. For each mine a summary of the pH measurements for the sites are given along with summaries for selected nutrients and metals. Diagrams 4.1 to 4.11 illustrate the range for each environmental variable and the spread of samples downstream from the mine.

4.2.1 Rum Jungle Mine

At Rum Jungle sites the pH measurements followed a trend of increasing pH with increased distance downstream from the mine (Figure 4.1). The lowest pH measurement, pH 3.4, was recorded at site RJI1, the highest pH 7.3 at site RJI9. Control site pH values ranged from 7.2 (RJC1) to 8.2 (RJC5), varying by only one pH unit.

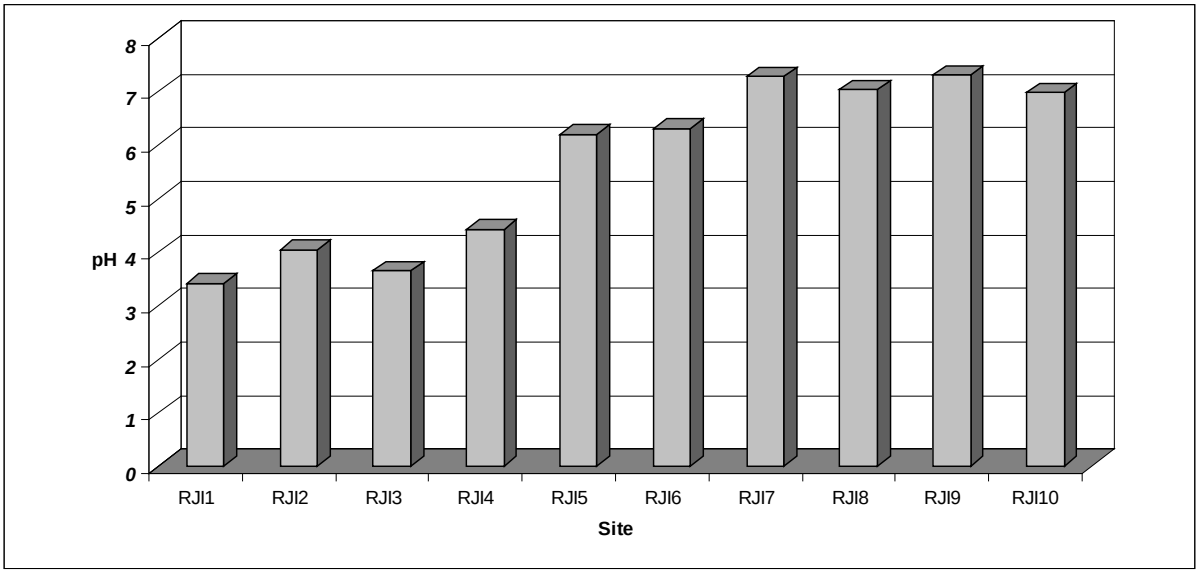


Figure 4.1. Spatial trends of pH at Rum Jungle sites.

Figure 4.2 shows the general trends of selected metals at each site downstream of Rum Jungle mine. Metal measurements decreased at each site away from the mine. At the site closest to the mine (RJI1) copper (Cu) and aluminium (Al) concentrations were measured at 6.0 and 18.6 mg/L respectively. At this site white efflorescence salts and orange iron flocs were present along the creek (Plate 4.1 and 4.2). At the furthest site from the mine (RJI10) the concentrations of these two metals decreased to 0.003 and 0.005 mg/L. At control sites aluminium and copper concentrations showed very little variation, ranging from 0.002 to 0.018 g/L for copper and 0.005 to 0.06 mg/L for aluminium. These values were comparable to the sites furthest from the mine. Site RJI2 had the highest concentration of copper and aluminium values as it receives polluted waters from the intermediate overburden heap. From site RJI2 to site RJI3 there was a substantial decrease in all measured metals, for instance, iron (Fe) drops from 23.4 to 1.2 mg/L.

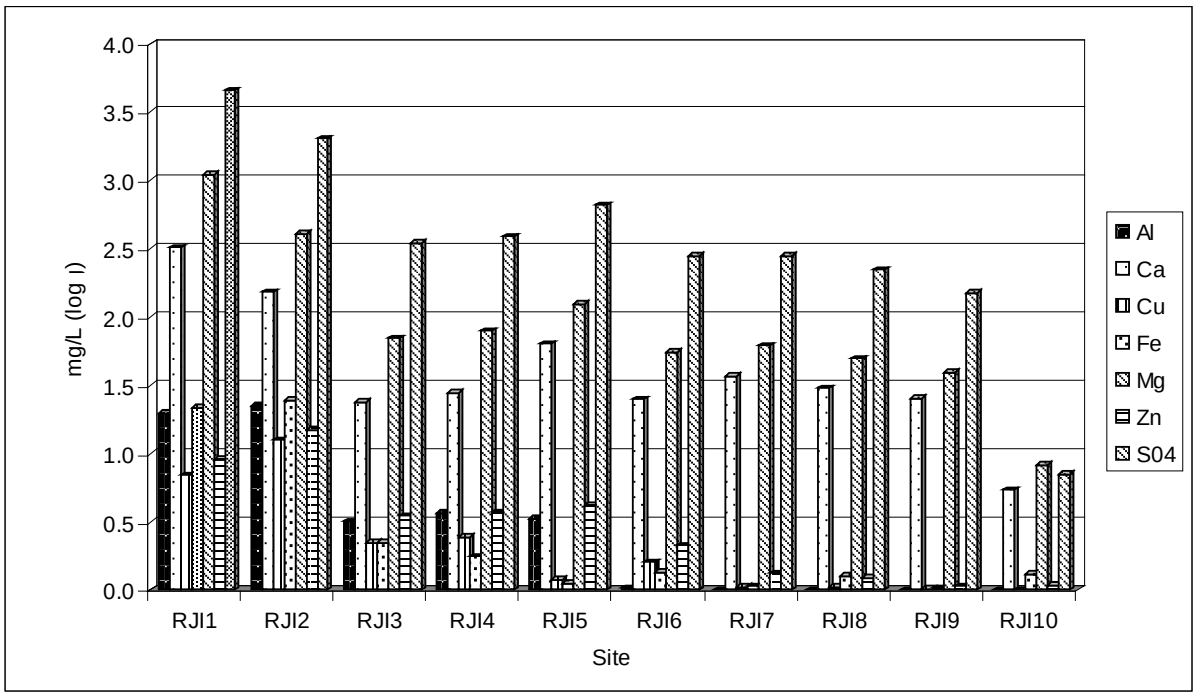


Figure 4.2. Spatial trends of selected metals (Al, Cu, Zn, Ca, Mg, Fe, SO_4) at Rum Jungle sites (mg/L).

Plate 4.1. White efflorescence salts at Rum Jungle site RJI1.

Plate 4.2. Orange iron floc at Rum Jungle site RJI1.

Nutrient measurements at sites downstream of Rum Jungle mine were highly variable. Total nitrogen (TN) and total kjeldahl nitrogen (TKN) showed uniformity in their concentrations at each site. From site RJI1 to RJI2 there were large decreases in TKN and TN concentrations of 1.2 to 0.5 mg/L and 1.3 to 0.5 mg/L for TKN and TN respectively. There was little variation in total phosphorus (TP) measurements across sites.

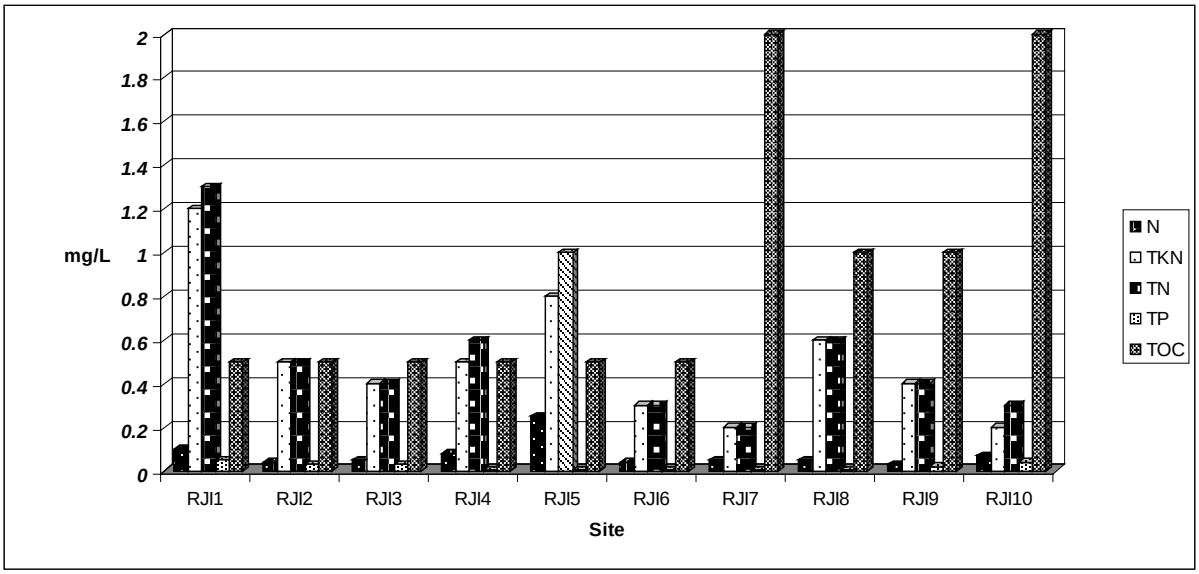


Figure 4.3. Nutrient trends at Rum Jungle sites (mg/L).

4.2.2 Redbank Mine

pH measurements at Redbank sites (Figure 4.4) followed a trend of increasing pH with distance away from the pollution source. The lowest pH measured was at site HCUSEC, closest to the mine, with the highest being at site SCDS12MC (pH 8.7). The control sites ranged from pH 6.0 at

ECUSBRX to 8.4 at three of the six control sites (CCUSSC, SCUSCC, SCUS12MC). Site ECUSBRX had a substantially lower pH than the majority of the control sites, almost 2.5 pH units lower. This may be due to unknown passive effluence from a nearby mine site.

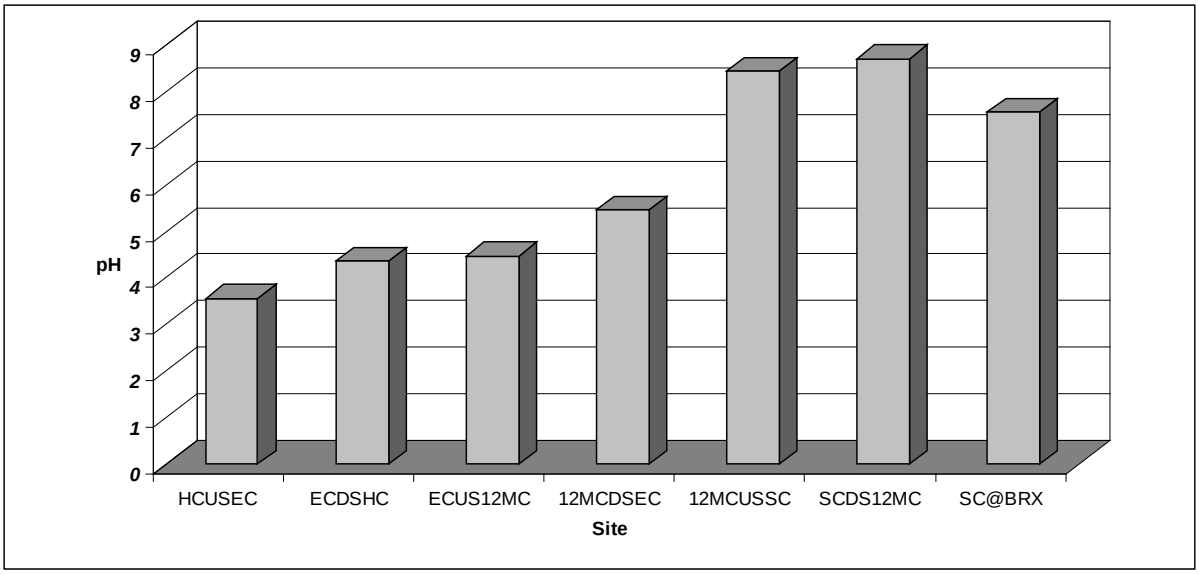


Figure 4.4. Spatial trends of pH at Redbank sites.

Spatial trends of metal concentrations at Redbank mine sites followed the inverse of the pH trend; they decreased with distance away from the mine. For instance copper (Cu) was measured at 97.8 mg/L at site HCUSEC closest to the mine and decreased to 0.05 mg/L at site SC@BRX, furthest downstream from the mine (Figure 4.5). The control site concentrations for this metal ranged from 0.28 mg/L at site ECUSBRX to 0.0008 mg/L at site SCUS12MC. Four of the control sites had magnesium (Mg) concentrations greater than 34 mg/L. For instance, site CCUSSC had a magnesium level of 57.1 mg/L. These concentrations were greater than all but the closest site to the pollution source, site HCUSEC.

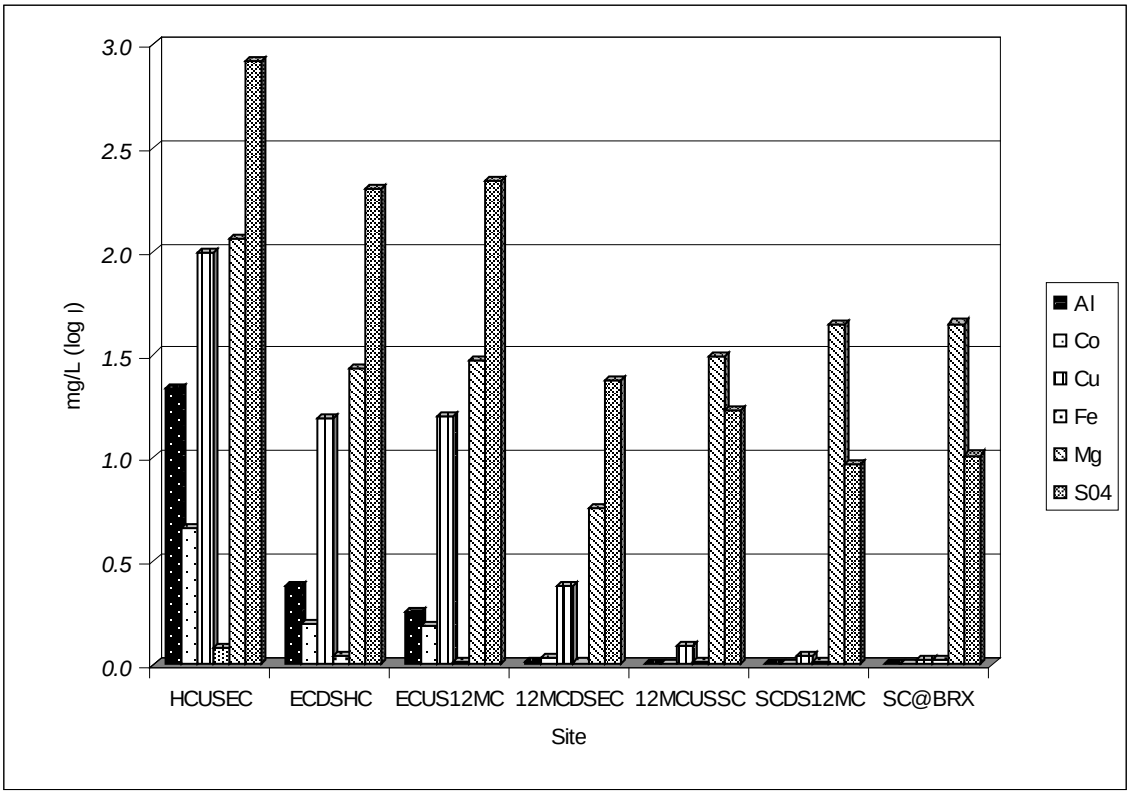


Figure 4.5. Spatial trends of selected metals (Al, Co, Cu, Mg, Fe, SO_4) at Redbank sites (mg/L).

4.2.3 Cosmo Howley Mine

At Cosmo Howley sites, pH ranged from 3.0 at site CHI1 to 7.3 at site CHI7, the site second furthest from the pollution source (Figure 4.6). There are three potential pollution sources. The sites BKI1 and BKI2 are both situated downstream from Brocks mine and drain the Faded Lily and Alligators waste dumps. The pH of both of these sites was neutral (approx. pH 7.0). Sites CHI1 and CHI2 receive mine dam water from Cosmo Howley mine which flows via a wetland system into the creek. The water at these sites was very acidic (pH = 3.0 and 4.0 respectively). Site CHI3 (pH 4.9) receives mine water dam from Cosmo Howley mine via an unnamed creek with flows into sites CHI4, CHI5 and CHI6 before converging with Howley Creek. Upstream from this point, along Howley Creek, sites CHI1, CHI2 and BK1 and BKI2 are situated. The control sites range from pH 7.2 at site CHC3 to 7.7 at site CHC2.

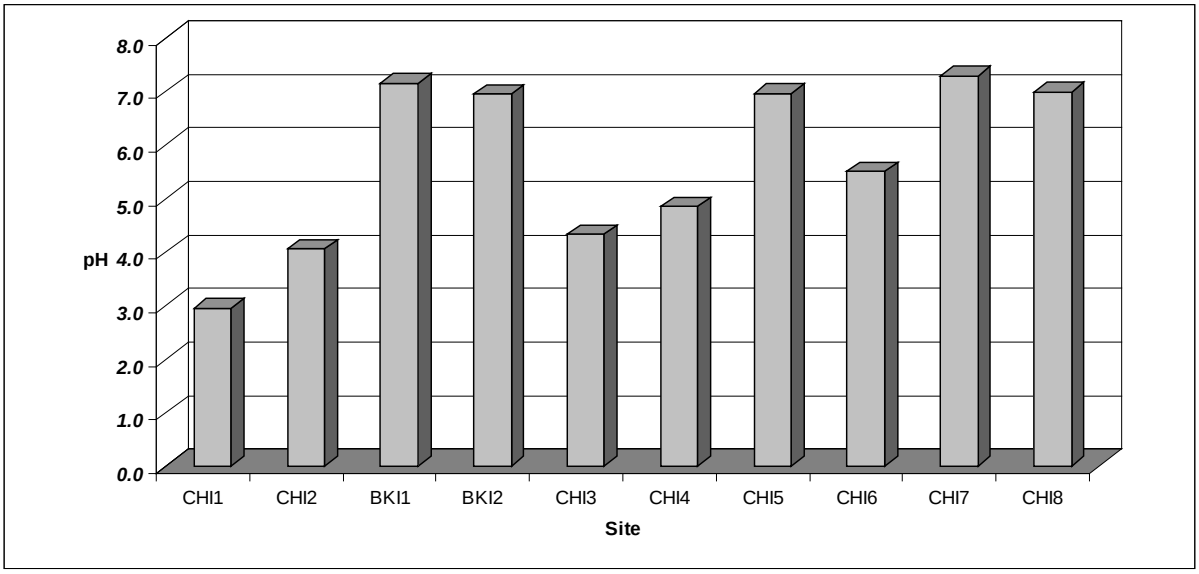


Figure 4.6. Spatial trends of pH at Cosmo Howley sites.

Metal concentrations at sites influenced from Brocks mine are relatively low in concentration compared to sites influenced by Cosmo Howley mine (Figure 4.7). For example, copper and aluminium concentrations measure 0.003 and 0.1 mg/L respectively at site BKI1. The highest concentrations of metals were measured at CHI1. Aluminium and sulphate (SO_4) measured 847.0 and 26000 mg/L respectively, decreasing to 0.005 and 130 mg/L at site CHI8. Control sites for all metals, except SO_4, showed metal concentrations similar to that of site CHI8, the site furthest from the pollution sources. Aluminium for instance, ranged from 0.005 to 0.2 mg/L at control sites.

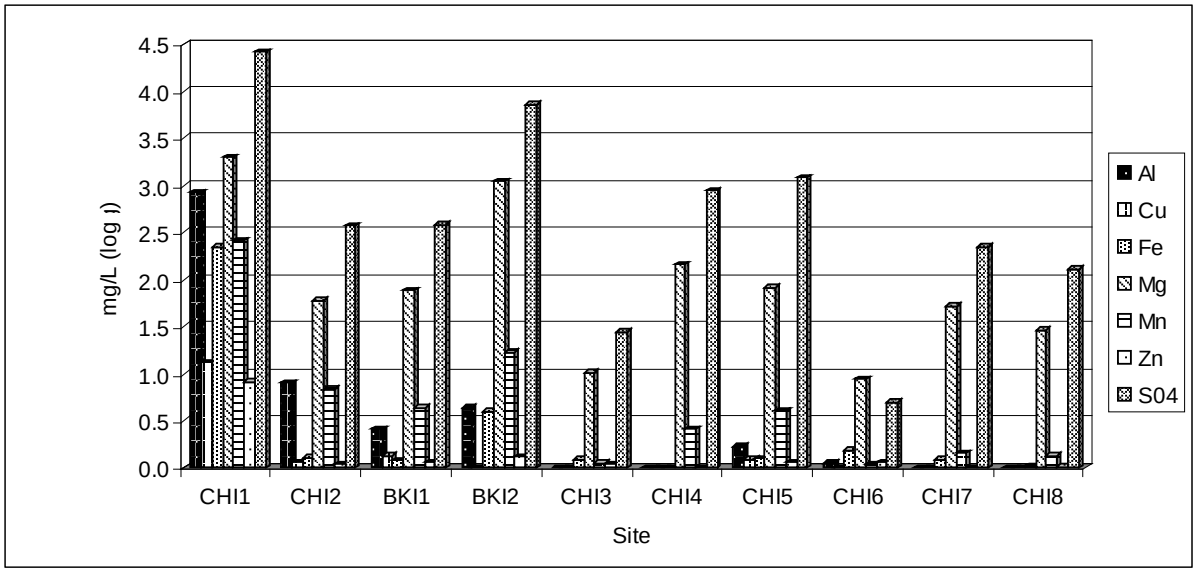


Figure 4.7. Spatial trends of selected metals (Al, Cu, Mg, Mn, Fe, SO_4, Zn) at Cosmo Howley sites (mg/L).

Nutrient measurements were highest for total organic carbon (TOC) and total nitrogen (TN) at sites CHI1 (0.8 & 4.3 mg/L) and BKI1 (7.0 & 0.8 mg/L), the sites closest to Cosmo Howley and Brocks Creek mines. Downstream, nutrient measurements generally decrease. Measurements of nitrate (N) and total phosphorus (TP) did not show much variability between sites, ranging from 0.01 and 0.04 mg/L at site BKI1 to 0.005 mg/L at CHI7 (Figure 4.8). The nitrate and total organic carbon measurements at control sites range from 0.1 to 0.008 mg/L and 6.0 to 1.0 mg/L at sites CHI1 and CHI2 respectively.

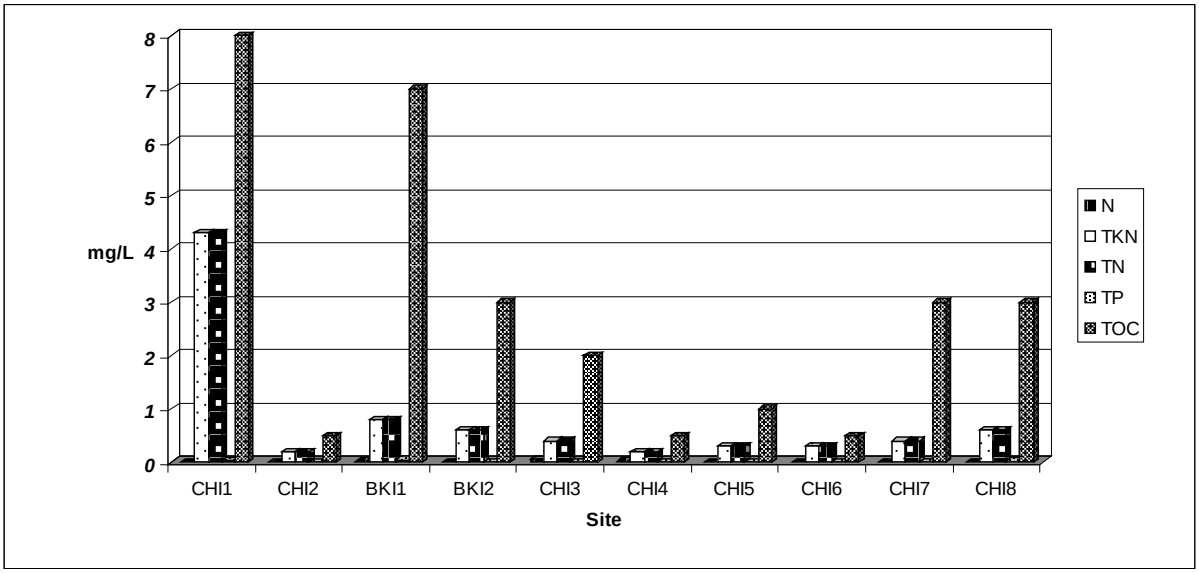


Figure 4.8. Nutrient trends at Cosmo Howley sites (mg/L).

4.2.4 Tom's Gully Mine

All sites downstream of Tom's Gully mine along Mount Bundy Creek, with the exception of TGI5 (pH 4.9), had pH values above 6.0 (Figure 4.9). Site TGI5 receives water draining the mine wetlands and is one of the sites where the presence of orange iron floc was widespread. The highest pH value, 7.6, was measured at site TGI4. The pH values of the control sites ranged from 7.0 at site TGC2 to 7.1 at site TGC1.

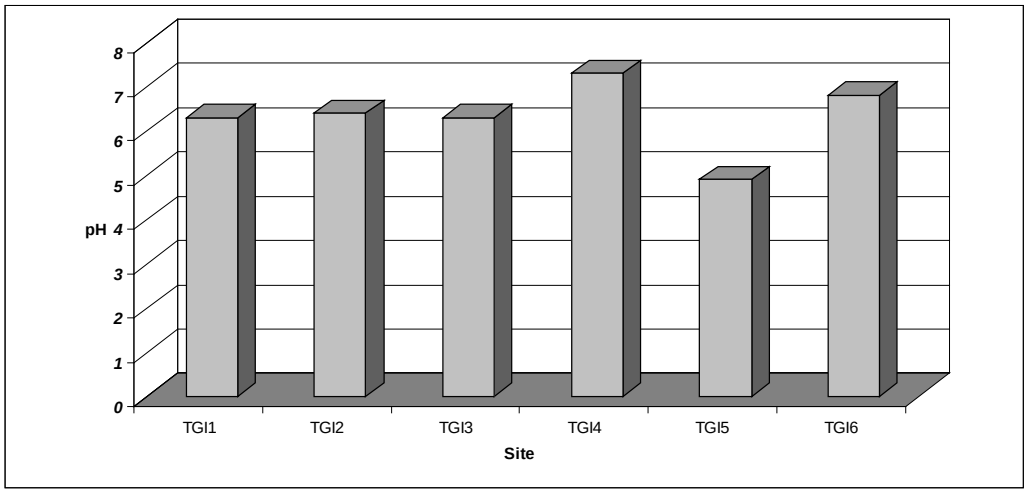


Figure 4.9. Spatial trends of pH at Tom's Gully sites.

The measurements of metal concentrations followed the same trend as those for pH (Figure 4.10). Magnesium (Mg), calcium (Ca) and sulphate (SO_4) measured 1.5, 1.8 and 2.0 mg/L respectively at site TGI1 and increased to their highest values at site TGI4 (57.6, 122, 440 mg/L) before decreasing again. Aluminium (Al) and copper (Cu) both measured 0.005 at site TGI1 and peaked at TGI5 (0.9, 0.1 mg/L) before decreasing. Sites TGI4 and TGI5 both receive waters from the wetlands. Control sites ranged from 3.0 to 0.05 mg/L for sulphate and 0.005 to 0.07 mg/L for aluminium.

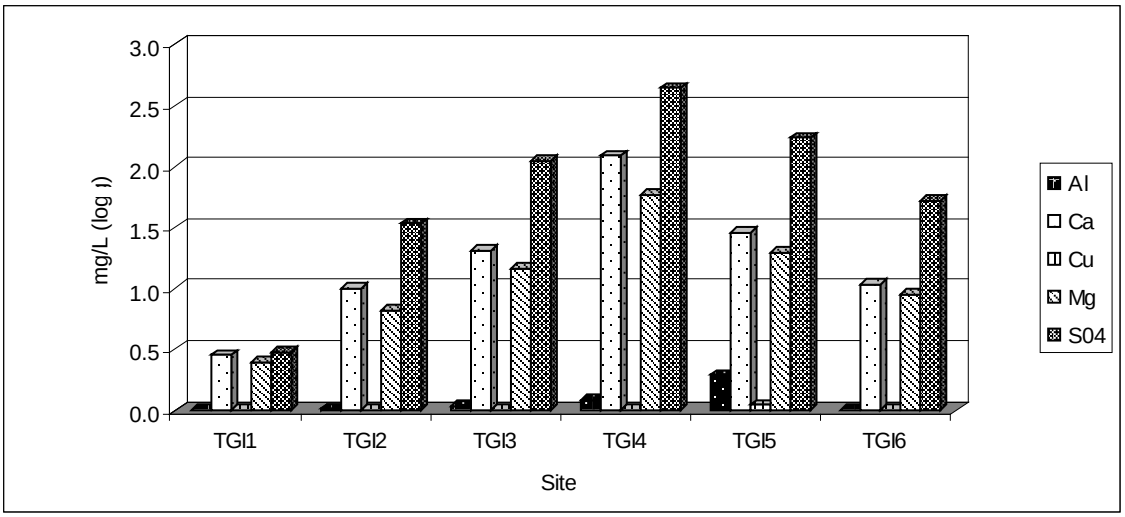


Figure 4.10. Spatial trends of selected metals (Al, Cu, Ca, Mg, SO_4) at Tom's Gully sites (mg/L).

The nutrient levels at sites downstream of Tom's Gully mine showed very little variability between sites (Figure 4.11). The largest variability was in the total organic carbon (TOC) between the sites TGI3 (0.5 mg/L) and TGI4 (1 mg/L). However, higher measurements were obtained from control sites. For instance, at site CHC2, total organic carbon measurements were much higher at 8.0 mg/L.

The highest total kjeldahl nitrogen and total nitrogen levels measured at impacted sites were 0.5 mg/L. This compares to 0.9 and 1.1 mg/L respectively at control site CHC2.

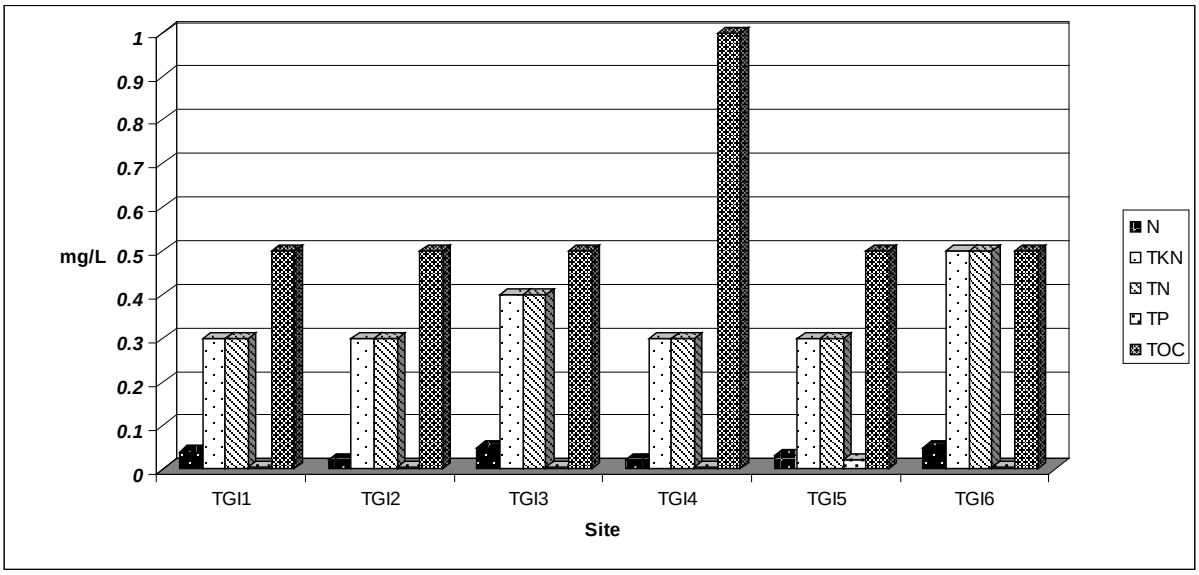


Figure 4.11. Nutrient trends at Tom's Gully sites (mg/L).

4.3 Correlation matrix of environmental variables

A correlation matrix of environmental data from 50 sites was used to determine the independent variables to be used in regression equations and transfer functions and to reduce the multivariate dataset to a manageable size. The matrix (Table 4.1) shows that there is a strong degree of covariance between many variables, especially between metals. Many published transfer function studies provide no information about covariance between variables, however multiple correlations between environmental variables are a feature of calibration datasets (Birks, 1998). Shaded cells indicate significant negative or positive correlation of $p \geq 0.5$.

Environmental variables such as dissolved oxygen (DO), total phosphorus (TP), nitrate (N) and temperature (Temp) are not significantly correlated with any other variable within the dataset. Consequently, these variables were retained for further analysis. Although many of the metal variables are strongly correlated to each other, for instance, copper (Cu) and aluminium (Al) ($p = 0.87$), both these variables were retained because of their importance in acid mine pollution. The percentage of trees, shrubs and grass in the riparian zone were correlated to one another (> 0.5), with shrubs having a correlation of 0.53 with sodium (Na). Consequently, sodium was retained to represent these variables. Salt and fish presence, together with % shade, was correlated with pH. The variables retained for use in the transfer function study were pH, aluminium, copper, iron (Fe), magnesium (Mg), sodium, lead (Pb), sulphate (SO_4), nitrate, total phosphorus, total organic carbon

(TOC), dissolved oxygen, temperature (Temp), conductivity (EC) and turbidity (Turb). The excluded variables were correlated with one or more of the retained variables.

Table 4.1. A correlation matrix of environmental variables.

	pH	Do	Temp	Turb	Flow	Foam	Oil	Clarity	Salts	Shade	Fish	Sand	Rock	Algae	bTrees	Trees	Shrubs	Grass	Tailings	N	TKN	TN	TP	TOC	EC	Al	Caco3	Ca	Cd	Co	Cu	Fe	K	Mg	Mn	Na	Ni	Pb	Zn	S04
pH	1	-0.07	-0.32	0.22	-0.16	-0.16	-0.42	0.37	-0.66	0.51	0.64	0.23	0.16	-0.49	0.28	0.39	0.23	-0.13	-0.35	-0.02	-0.25	-0.24	-0.01	0.17	-0.56	-0.79	0.79	-0.49	-0.57	-0.72	-0.71	-0.64	-0.28	-0.59	-0.86	-0.15	-0.73	-0.37	-0.71	-0.66
Do	-0.07	1	0.11	-0.35	0.20	0.12	0.03	-0.47	0.09	-0.36	-0.29	0.06	-0.16	0.14	-0.08	0.09	-0.38	0.13	-0.01	0.07	-0.17	-0.16	-0.15	-0.24	0.34	0.00	-0.11	0.22	-0.21	-0.04	-0.17	-0.24	-0.03	0.33	0.13	0.12	-0.04	-0.12	-0.03	0.29
Temp	-0.32	0.11	1	0.08	-0.20	0.12	0.19	-0.12	0.21	-0.36	-0.23	0.14	-0.20	0.11	-0.23	-0.11	-0.11	0.04	0.21	-0.25	0.03	0.01	0.06	-0.09	0.28	0.23	-0.11	0.26	0.12	0.14	0.20	0.22	0.37	0.25	0.23	0.30	0.16	0.06	0.14	0.36
Turb	0.22	-0.35	0.08	1	-0.54	0.17	-0.01	0.45	-0.21	0.27	0.19	-0.09	0.22	-0.18	-0.07	-0.36	0.05	0.23	-0.13	-0.20	0.01	0.00	-0.05	0.27	-0.13	-0.17	0.19	-0.28	-0.13	-0.20	-0.17	-0.13	0.15	-0.22	-0.22	0.34	-0.20	-0.06	-0.22	-0.21
Flow	-0.16	0.20	-0.20	-0.54	1	-0.01	-0.02	-0.43	0.23	-0.11	-0.11	0.08	-0.27	0.23	0.21	0.22	0.00	-0.28	0.16	0.07	-0.01	-0.01	-0.25	-0.43	0.39	0.14	-0.01	0.43	0.17	0.22	0.18	0.13	-0.25	0.46	0.27	-0.23	0.22	0.10	0.23	0.34
Foam	-0.16	0.12	0.12	0.17	-0.01	1	0.57	0.54	-0.06	0.09	-0.14	-0.21	0.19	-0.01	0.13	0.02	-0.15	0.05	-0.12	-0.18	-0.10	-0.14	-0.19	0.02	0.04	-0.06	0.02	0.03	-0.02	-0.01	-0.02	-0.04	0.03	0.05	0.10	0.24	-0.02	0.06	0.00	0.24
Oil	-0.42	0.03	0.19	-0.01	-0.02	0.57	1	-0.55	0.15	0.12	-0.22	-0.16	0.11	0.05	-0.07	-0.05	-0.04	0.08	-0.01	0.01	-0.05	-0.06	-0.28	-0.09	0.01	0.09	-0.36	0.00	0.02	0.02	-0.02	-0.03	0.22	0.03	0.26	0.26	0.04	0.07	0.04	0.34
Clarity	0.37	-0.47	-0.12	0.45	-0.43	0.54	-0.55	1	-0.40	0.49	0.39	0.02	0.12	-0.19	0.40	-0.21	0.02	0.07	-0.31	0.01	0.05	0.05	0.17	0.44	-0.53	-0.32	0.32	-0.56	-0.18	-0.34	-0.31	-0.20	-0.13	-0.59	-0.44	0.06	-0.34	-0.08	-0.32	-0.51
Salts	-0.66	0.09	0.21	-0.21	0.23	-0.06	0.15	-0.40	1	-0.55	-0.41	-0.19	-0.19	0.51	-0.35	-0.28	-0.13	0.03	0.42	0.22	0.51	0.52	0.07	-0.07	0.56	0.75	-0.56	0.51	0.66	0.82	0.77	0.68	0.08	0.60	0.68	-0.07	0.82	0.47	0.80	0.51
Shade	0.51	-0.36	-0.36	0.27	-0.11	0.09	0.12	0.49	-0.55	1	0.61	-0.16	0.53	-0.53	0.15	0.13	0.27	0.01	-0.31	-0.18	-0.17	-0.18	-0.12	0.08	-0.60	-0.55	0.42	-0.55	-0.30	-0.60	-0.54	-0.45	-0.18	-0.59	-0.57	-0.04	-0.58	-0.16	-0.58	-0.51
Fish	0.64	-0.29	-0.23	0.19	-0.11	-0.14	-0.42	0.39	-0.41	0.61	1	-0.09	0.30	-0.32	0.29	0.35	0.40	-0.19	-0.32	-0.01	-0.12	-0.11	-0.02	0.08	-0.61	-0.55	0.58	-0.55	-0.36	-0.52	-0.45	-0.40	-0.52	-0.56	-0.67	-0.30	-0.52	-0.20	-0.53	-0.58
Sand	0.23	0.06	0.14	-0.09	0.08	-0.21	-0.16	0.02	-0.19	-0.16	-0.09	1	-0.64	-0.28	0.26	0.05	-0.21	-0.06	-0.03	-0.24	0.00	-0.01	0.00	0.15	-0.11	-0.14	0.25	-0.10	-0.06	-0.22	-0.17	-0.10	0.11	-0.08	-0.17	0.03	-0.19	0.04	-0.27	-0.08
Rock	0.16	-0.16	-0.20	0.22	-0.27	0.19	0.11	0.12	-0.19	0.53	0.30	-0.64	1	-0.53	-0.19	0.00	0.26	0.11	-0.14	-0.02	-0.06	-0.07	0.08	0.01	-0.37	-0.20	0.06	-0.29	-0.11	-0.25	-0.16	-0.15	-0.06	-0.35	-0.25	-0.11	-0.25	-0.10	-0.25	-0.28
Algae	-0.49	0.14	0.11	-0.18	0.23	-0.01	0.05	-0.19	0.51	-0.53	-0.32	-0.28	-0.53	1	-0.11	-0.09	-0.10	-0.03	0.21	0.30	0.08	0.10	-0.08	-0.17	0.62	0.45	-0.40	0.52	0.24	0.62	0.45	0.33	-0.02	0.56	0.54	0.09	0.58	0.12	0.66	0.48
Btrees	0.28	-0.08	-0.23	-0.07	0.21	0.13	-0.07	0.40	-0.35	0.15	0.29	0.26	-0.19	-0.11	1	0.28	0.17	-0.57	-0.11	0.21	-0.16	-0.13	-0.15	0.13	-0.31	-0.38	0.44	-0.29	-0.30	-0.32	-0.33	-0.29	-0.46	-0.29	-0.41	-0.36	-0.34	-0.18	-0.28	-0.34
Trees	0.39	0.09	-0.11	-0.36	0.22	0.02	-0.05	-0.21	-0.28	0.13	0.35	0.05	0.00	-0.09	0.28	1	0.51	-0.58	-0.10	0.16	-0.10	-0.08	-0.13	-0.04	-0.34	-0.31	0.41	-0.07	-0.19	-0.24	-0.24	-0.24	-0.37	-0.22	-0.33	-0.45	-0.25	-0.07	-0.23	-0.32
Shrubs	0.23	-0.38	-0.11	0.05	0.00	-0.15	-0.04	0.02	-0.13	0.27	0.40	-0.21	0.26	-0.10	0.17	0.51	1	-0.49	-0.11	0.28	-0.09	-0.07	-0.19	-0.02	-0.32	-0.24	0.15	-0.11	-0.11	-0.20	-0.16	-0.16	-0.06	-0.33	-0.32	-0.53	-0.21	-0.14	-0.16	-0.39
Grass	-0.13	0.13	0.04	0.23	-0.28	0.05	0.08	0.07	0.03	0.01	-0.19	-0.06	0.11	-0.03	-0.57	-0.58	-0.49	1	-0.54	-0.38	0.09	0.05	0.11	0.08	0.14	0.14	-0.25	-0.01	0.07	-0.03	-0.01	0.02	0.37	0.13	0.23	0.57	0.01	0.17	-0.11	0.20
Tailings	-0.35	-0.01	0.21	-0.13	0.16	-0.12	-0.01	-0.31	0.42	-0.31	-0.32	-0.03	-0.14	0.21	-0.11	-0.10	-0.11	-0.54	1	0.20	0.07	0.10	0.10	-0.21	0.34	0.36	-0.27	0.33	0.28	0.50	0.46	0.39	0.02	0.29	0.29	-0.14	0.46	-0.01	0.57	0.32
N	-0.02	0.07	-0.25	-0.20	0.07	-0.18	0.01	0.01	0.22	-0.18	-0.01	-0.24	-0.02	0.30	0.21	0.16	0.28	-0.38	0.20	1	0.1322	0.24	-0.02	0.25	-0.03	0.05	-0.24	0.08	-0.01	0.18	0.02	0.04	-0.13	-0.07	0.01	-0.29	0.15	-0.08	0.26	-0.15
TKN	-0.25	-0.17	0.03	0.01	-0.01	-0.10	-0.05	0.05	0.51	-0.17	-0.12	0.00	-0.06	0.08	-0.16	-0.10	-0.09	0.09	0.07	0.13	1	0.99	0.21	0.49	0.05	0.68	-0.14	0.16	0.81	0.54	0.53	0.73	0.12	0.23	0.48	0.06	0.61	0.89	0.41	0.07
TN	-0.24	-0.16	0.01	0.00	-0.01	-0.14	-0.06	0.05	0.52	-0.18	-0.11	-0.01	-0.07	0.10	-0.13	-0.08	-0.07	0.05	0.10	0.24	0.99	1	0.21	0.51	0.04	0.67	-0.16	0.15	0.79	0.54	0.52	0.71	0.10	0.20	0.47	0.02	0.61	0.86	0.42	0.04
TP	-0.01	-0.15	0.06	-0.05	-0.25	-0.19	-0.28	0.17	0.07	-0.12	-0.02	0.00	0.08	-0.08	-0.15	-0.13	-0.19	0.11	0.10	-0.02	0.21	0.21	1	0.30	-0.20	0.09	0.00	-0.24	0.11	0.08	0.13	0.17	-0.09	-0.22	-0.06	-0.08	0.08	0.02	0.09	-0.26
TOC	0.17	-0.24	-0.09	0.27	-0.43	0.02	-0.09	0.44	-0.07	0.08	0.08	0.15	0.01	-0.17	0.13	-0.04	-0.02	0.08	-0.21	0.25	0.49	0.51	0.30	1	-0.33	0.09	0.12	-0.27	0.19	-0.03	0.00	0.24	0.19	-0.24	-0.07	0.21	0.01	0.33	-0.10	-0.33
EC	-0.56	0.34	0.28	-0.13	0.39	0.04	0.01	-0.53	0.56	-0.60	-0.61	-0.11	-0.37	0.62	-0.31	-0.34	-0.32	0.14	0.34	-0.03	0.05	0.04	-0.20	-0.33	1	0.54	-0.33	0.87	0.35	0.60	0.50	0.41	0.36	0.95	0.71	0.43	0.58	0.16	0.61	0.84
Al	-0.79	0.00	0.23	-0.17	0.14	-0.06	0.09	-0.32	0.75	-0.55	-0.55	-0.14	-0.20	0.45	-0.38	-0.31	-0.24	0.14	0.36	0.05	0.68	0.67	0.09	0.09	0.5381	1	-0.60	0.55	0.89	0.89	0.87	0.90	0.28	0.65	0.92	0.16	0.92	0.75	0.79	0.55
CaC03	0.79	-0.11	-0.11	0.19	-0.01	0.02	-0.36	0.32	-0.56	0.42	0.58	0.25	0.06	-0.40	0.44	0.41	0.15	-0.25	-0.27	-0.24	-0.14	-0.16	0.00	0.12	-0.33	-0.60	1	-0.22	-0.40	-0.59	-0.53	-0.37	-0.27	-0.28	-0.64	-0.02	-0.59	-0.22	-0.61	-0.46
Ca	-0.49	0.22	0.26	-0.28	0.43	0.03	0.00	-0.56	0.51	-0.55	-0.55	-0.10	-0.29	0.52	-0.29	-0.07	-0.11	-0.01	0.33	0.08	0.16	0.15	-0.24	-0.27	0.8699	0.5501	-0.223	1	0.46	0.57	0.51	0.51	0.42	0.90	0.71	0.32	0.58	0.27	0.59	0.78
Cd	-0.57	-0.21	0.12	-0.13	0.17	-0.02	0.02	-0.18	0.66	-0.30	-0.36	-0.06	-0.11	0.24	-0.30	-0.19	-0.11	0.07	0.28	-0.01	0.81	0.79	0.11	0.19	0.35	0.89	-0.40	0.46	1	0.80	0.82	0.90	0.23	0.50	0.77	0.07	0.86	0.88	0.70	0.41
Co	-0.72	-0.04	0.14	-0.20	0.22	-0.01	0.02	-0.34	0.82	-0.60	-0.52	-0.22	-0.25	0.62	-0.32	-0.24	-0.20	-0.03	0.50	0.18	0.54	0.54	0.08	-0.03	0.60	0.89	-0.59	0.57	0.8028	1	0.96	0.85	0.09	0.65	0.83	-0.02	0.99	0.57	0.97	0.55
Cu	-0.71	-0.17	0.20	-0.17	0.18	-0.02	-0.02	-0.31	0.77	-0.54	-0.45	-0.17	-0.16	0.45	-0.33	-0.24	-0.16	-0.01	0.46	0.02	0.53	0.52	0.13	0.00	0.50	0.87	-0.53	0.51	0.823	0.9559	1	0.90	0.12	0.59	0.78	-0.05	0.95	0.57	0.91	0.49
Fe	-0.64	-0.24	0.22	-0.13	0.13	-0.04	-0.03	-0.20	0.68	-0.45	-0.40	-0.10	-0.15	0.33	-0.29	-0.24	-0.16	0.02	0.39	0.04	0.73	0.71	0.17	0.24	0.41	0.90	-0.37	0.51	0.90	0.85	0.90	1	0.26	0.56	0.79	0.11	0.88	0.74	0.76	0.40
K	-0.28	-0.03	0.37	0.15	-0.25	0.03	0.22	-0.13	0.08	-0.18	-0.52	0.11	-0.06	-0.02	-0.46	-0.37	-0.06	0.37	0.02	-0.13	0.12	0.10	-0.09	0.19	0.36	0.28	-0.27	0.42	0.23	0.09	0.12	0.2554	1	0.30	0.39	0.71	0.13	0.18	0.07	0.40
Mg	-0.59	0.33	0.25	-0.22	0.46	0.05	0.03	-0.59	0.60	-0.59	-0.56	-0.08	-0.35	0.56	-0.29	-0.22	-0.33	0.13	0.29	-0.07	0.23	0.20	-0.22	-0.24	0.95	0.65	-0.28	0.90	0.50	0.65	0.59	0.5579	0.30	1	0.80	0.38	0.66	0.35	0.62	0.84
Mn	-0.86	0.13	0.23	-0.22	0.27	0.10	0.26	-0.44	0.68	-0.57	-0.67	-0.17	-0.25	0.54	-0.41	-0.33	-0.32	0.23	0.29	0.01	0.48	0.47	-0.06	-0.07	0.71	0.92	-0.64	0.71	0.77	0.83	0.78	0.79	0.39	0.80	1	0.34	0.86	0.63	0.76	0.76
Na	-0.15	0.12	0.30	0.34	-0.23	0.24	0.26	0.06	-0.07	-0.04	-0.30	0.03	-0.11	0.09	-0.36	-0.45	-0.53	0.57	-0.14	-0.29	0.06	0.02	-0.08	0.21	0.43	0.16	-0.02	0.32	0.07	-0.02	-0.05	0.11	0.71	0.38	0.34	1	0.02	0.14	-0.07	0.46
Ni	-0.73	-0.04	0.16	-0.20	0.22	-0.02	0.04	-0.34	0.82	-0.58	-0.52	-0.19	-0.25	0.58	-0.34	-0.25	-0.21	0.01	0.46	0.15	0.61	0.61	0.08	0.01	0.58	0.92	-0.59	0.58	0.86	0.99	0.95	0.88	0.13	0.66	0.86	0.02	1	0.65	0.94	0.56
Pb	-0.37	-0.12	0.06	-0.06	0.10	0.06	0.07	-0.08	0.47	-0.16	-0.20	0.04	-0.10	0.12	-0.18	-0.07	-0.14	0.17	-0.01	-0.08	0.89	0.86	0.02	0.33	0.16	0.75	-0.22	0.27	0.88	0.57	0.57	0.74	0.18	0.35	0.63	0.14	0.7	1	0.40	0.28
Zn	-0.71	-0.03	0.14	-0.22	0.23	0.00	0.04	-0.32	0.80	-0.58	-0.53	-0.27	-0.25	0.66	-0.28	-0.23	-0.16	-0.11	0.57	0.26	0.41	0.42	0.09	-0.10	0.61	0.79	-0.61	0.59	0.70	0.97	0.91	0.76	0.07	0.62	0.76	-0.07	0.94	0.40	1	0.54
SO4	-0.66	0.29	0.36	-0.21	0.34	0.24	0.34	-0.51	0.51	-0.51	-0.58	-0.08	-0.28	0.48	-0.34	-0.32	-0.39	0.20	0.32	-0.15	0.07	0.04	-0.26	-0.33	0.84	0.55	-0.46	0.78	0.41	0.55	0.49	0.40	0.40	0.84	0.76	0.46	0.56	0.28	0.54	1

Table 4.2 outlines the range of measured variables for the reduced dataset. Overall, the sites are warm (25.1 - 35 ºC) and well oxygenated (mean > 82.5%) in both impacted and control sites. Sites are also characterised by low nutrient (mean N and TP < 0.7 mg/L) levels.

Table 4.2. Summary of environmental variable ranges for combined impacted and control sites after data screening.

	pH	Al mg/L	Cu mg/L	Fe mg/L	Mg mg/L	Na mg/L	Pb mg/L	SO_4 mg/L	N mg/L	TP mg/L	TOC mg/L	DO %	Temp ^oC	EC µS/cm	Turb
CONTROL															
Min	6.00	0.005	0.001	0.010	0.600	2.100	0.015	0.050	0.500	0.005	0.50	26.90	25.10	15.80	0.00
Max	8.40	0.200	0.300	1.000	57.000	28.600	0.015	44.000	1.100	0.100	8.00	101.80	75.00	767.00	764.00
Mean	7.99	0.043	0.032	0.243	20.782	11.104	0.015	4.691	0.700	0.020	2.77	82.51	33.10	288.95	198.68
Median	7.39	0.008	0.007	0.138	12.000	5.980	0.015	0.300	0.600	0.005	2.00	80.90	27.80	170.00	8.40
Stndev	0.65	0.064	0.069	0.270	19.827	9.475	0.000	10.497	0.210	0.029	2.40	22.96	11.91	272.12	302.94
IMPACT															
Min	2.95	0.005	0.002	0.020	1.520	2.500	0.015	2.000	0.200	0.005	0.50	6.70	25.50	38.00	0.20
Max	8.70	847.000	97.800	223.000	2000.000	58.600	0.700	26000.000	4.300	0.050	8.00	130.70	32.00	5635.00	1000.00
Mean	5.89	28.228	4.947	8.396	179.948	17.167	0.028	1439.473	0.608	0.013	1.58	87.87	28.94	923.99	114.49
Median	6.20	0.040	0.050	0.237	43.600	7.220	0.015	219.000	0.400	0.005	0.50	90.00	28.80	433.00	25.00
Stdn	1.56	147.101	17.244	38.887	418.944	28.290	0.060	4641.137	0.795	0.014	1.95	27.73	1.98	1327.86	196.69

The main difference between control and impacted sites were the heavy metal, conductivity and pH levels. Impacted sites had a mean copper level of 4.95 mg/L compared to a mean of 0.03 mg/L at control sites. This phenomenon was repeated for sulphate and heavy metals such as aluminium, iron and magnesium. The most impacted sites were characterised by extreme concentrations of heavy metals, for instance copper (max 97.8 mg/L), aluminium (max 847 mg/L), iron (max 223 mg/L) and magnesium (max 2000 mg/L). These sites also had high conductivity levels (max 5635 µS/cm). The variability within impacted sites, especially of metals, is indicated by the large standard deviation (stdn) values. For instance, the standard deviation value for aluminium is 147.1 and the coefficient of variance for this variable is 5.2 units.

4.4 AquaRisk

The measured and modelled water quality data from each of the four mines were used in AquaRisk (version 3), in conjunction with National Water Quality Guidelines (2000), to determine metals of potential concern. The bioavailable fraction of each metal was determined and similarly ranked in Tier 1 analysis by AquaRisk to identify the variables of potential concern. Bioavailable fractions of pollutant metals (i.e. copper) were estimated by AquaRisk using the geochemical speciation code, MOPDPHRQ. Table 4.3 illustrates the metals of potential concern at each mine in both total and bioavailable fraction measurements. Each of the metals are of potential concern (values > 1).

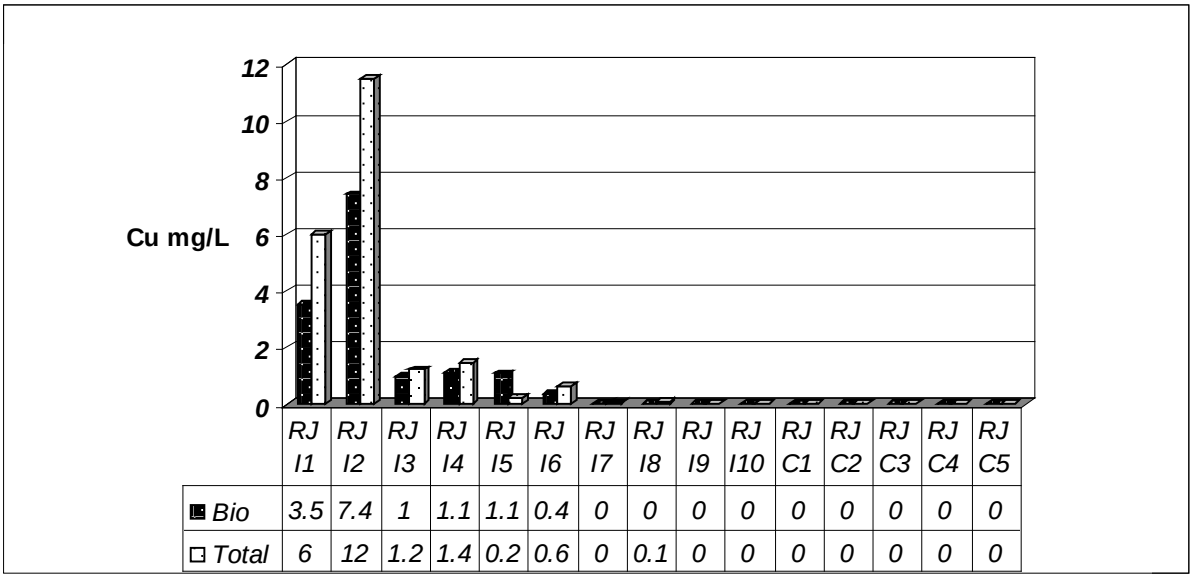
Of the metals, which were ranked according to their total values (mg/L), aluminium and copper were the two metals which exceeded the water quality guidelines (WQG) the most. For example, the average concentration of copper at Cosmo Howley mine was 6100 times greater than the guideline levels. At Cosmo Howley mine, aluminium ranked as the highest metal of potential concern followed by cobalt and copper. The ANZECC water quality guideline level for aluminium in aquatic systems (95% protection level) is 55 µg/L (0.055 mg/L). 95% is the trigger value (guideline) concentration of each metal applying to moderately disturbed systems to protect 95% of species. At Rum Jungle mine cobalt was the highest ranking metal (87,200) followed by copper (33,200) and aluminium (26,300). There are no values for lead as there was less than 0.015 mg/L detected at any one of the sites. The metal contaminants at Tom's Gully mine all exceed the WQG. In order of ranking, aluminium, zinc and then cadmium ranked as the greatest metals of potential concern. For Redbank mine, the metals copper, cobalt and then aluminium ranked in order of potential concern. There are no cadmium or lead values for Redbank as these metals were not sampled at this mine.

Table 4.3. Contaminants of potential concern at each mine (mg/L).

	Como Howley		**Tom's Gully**		**Rum Jungle**		**Redbank**	
	Total	Bio	Total	Bio	Total	Bio	Total	Bio
Al	169000	503000000	554	123	26300	112000000	478	1170000
Cd	445	33	143	110	244	127	-	-
Co	7500	2320	59.1	49.2	87200	58800	1760	8880
Cu	6100	6140	58.8	75.2	33200	102000	54400	3570000
Mn	5530	2400	96.8	80.5	1030	748	64.1	214
Ni	1330	-	12.8	-	2120	-	325	-
Pb	-	-	-	-	-	-	-	-
Zn	1450	206	413	334	16100	10200	37	294

Each of the metals were ranked again in tier one analysis of AquaRisk, according to how much their bioavailable fractions exceeded the WQG (Table 4.3). For Cosmo Howley mine the bioavailable fraction changed the order of ranking of metals. Aluminium remained the metal of most concern but exceeded the WQG at an even greater level, 503,000,000 mg/L instead of 169, 000 mg/L. Instead of cobalt ranking next copper was now more important. Similarly, for the Rum Jungle mine, aluminium remained as highest potential concern but was now followed by copper. Copper remained the highest of potential concern for Redbank mine followed by aluminium and cobalt. The WQG were developed for water at pH 6.5. Most of these sites were more acidic than this. If pH was less than 6.5 a greater proportion of the metal i.e. aluminium, will be bioavailable as the acid dissolves some of the organically bound aluminium. Hence toxicity of the aluminium is actually greater at lower pH and will exceed the WQG to an even greater extent.

Figure 4.12 plots the total and bioavailable fractions of each metal illustrating the difference between the two measurements (for a list of the bioavailable and total measurements for each metal see Appendix 5). For instance, the bioavailable copper concentration at site RJI1 is 3.5 mg/L compared to 6.0 mg/L of the total concentration. This illustrates the significant disparity between the two measurements.



	RJ I1	RJ I2	RJ I3	RJ I4	RJ I5	RJ I6	RJ I7	RJ I8	RJ I9	RJ I10	RJ C1	RJ C2	RJ C3	RJ C4	RJ C5
Bio	3.5	7.4	1	1.1	1.1	0.4	0	0	0	0	0	0	0	0	0
Total	6	12	1.2	1.4	0.2	0.6	0	0.1	0	0	0	0	0	0	0

Figure 4.12. Measurements of bioavailable and total copper at Rum Jungle sites.

Using the bioavailable metal results of Rum Jungle as an example, the cumulative probability distributions of measured and modelled copper concentrations are graphed (Figure 4.13) illustrating the number of sites exceeding the WQG for bioavailable concentrations.

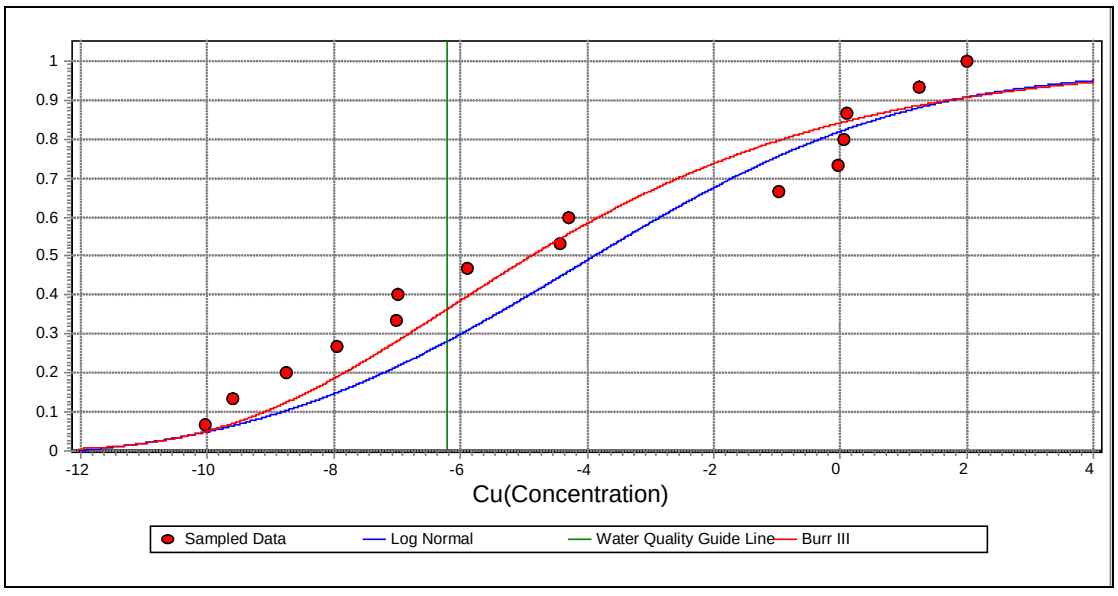


Figure 4.13. Cumulative probability distribution of measured and modelled (bioavailable) copper (mg/L) concentrations in Rum Jungle samples (n= 15).

Of the 15 Rum Jungle sites plotted in Figure 4.13, nine sites had bioavailable concentrations above national water quality guideline values. This indicated that all but one site sampled downstream from the mine had bioavailable metal measurements which exceeded the Australian Water Quality Guidelines (WQG).

4.5 Diatom genus and species floristic descriptions

For Rum Jungle sites, 600 diatom valves were counted, and the relationship between the number of species identified and counting effort was determined (Figure 4.14). By the 400 valve count, 80% of the species has been encountered. At nine sites, 90% of the species had been recorded. By the 500 valve count 90% of the species were recorded at eleven sites. 400 valves were counted per slide to achieve good representation of species at each site. Overall, from the entire dataset, 267 diatom species and 45 genera were recorded. Of the 267 species many were difficult to identify. A total of 51 species which were morphologically similar, but divergent enough to preclude them from being identified as a specific taxon, were given an "cf ". 42 diatoms were unidentifiable and so were given a study name (spec.) rather than force fitting them into species. There were six morphological sub groups were defined which were thought to be morphotypes of published species.

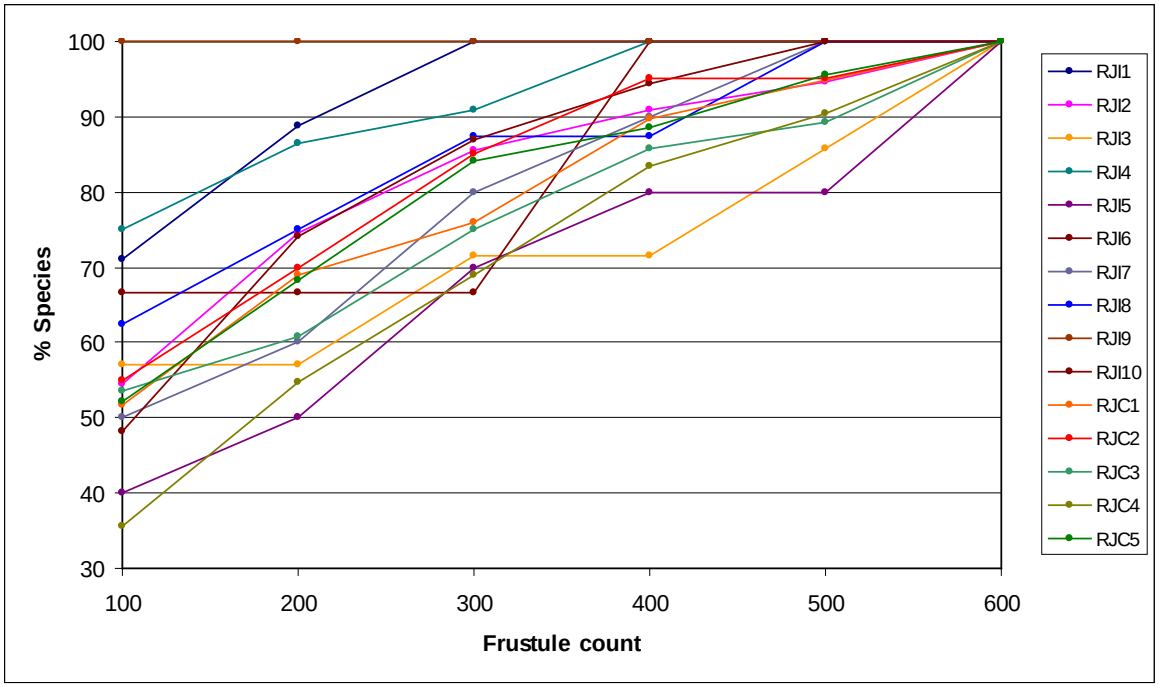


Figure 4.14. Number of diatom species per 100 frustule count at Rum Jungle sites.

For the purpose of statistical analysis all taxa which did not occur at more than one site were eliminated. In total, 74 species occurred at one site only. When these species were excluded the dataset included 193 diatom taxa. Appendix 6 contains descriptive details regarding the taxa. A list of authorities and synonyms for all diatom taxa mentioned in the text or illustrations are presented in Appendix 4.

A total of 161 species were found at control sites compared to 175 at impacted sites. 18 species were found only at control sites, for instance, *Neidium ampliatum, Rhopalodia gibba, Cocconeis placentula* and *Navicula veneta*. The five most dominant species at the combined control sites were, in order of dominance, *Achanthidium minutissimum* v. *minutissimum, Epithemia* cf. *cistula, Navicula notha, Encyonema minutum* and *Achnanthidium minutissimum* v. *exilis*. The five dominant species at combined impacted sites were *Achnanthidium minutissimum* v. *minutissimum, Nitzschia paleaeformis, Nitzschia palea* v. *palea, Nitzschia vasta* and *Brachysira neoexilis*. The dominant species at sites with a pH < 5 were *Nitzschia paleaeformis, Achnanthidium minutissimum* v. *minutissimum, Nitzschia vasta, Nitzschia nana, Pinnularia subcapitata* and *Nitzschia palea* v. *palea*. *Nitzschia vasta* dominated the only site with a pH below 3.0 (site CHI1). Species such as *Pinnularia joculata* and *Nitzschia paleaeformis* were found only at impacted sites. The species with the highest occurrence rates were *N. palea* v. *palea* and *A. minutissmum* v. *minutissimum* (38). Examples of species occurring at just two sites are *Pinnularia rupestris* and *Navicula variostriata.*

Hill's N2, an index of evenness which takes into account both the number of occurrences and the relative abundance of species in samples, gave higher weights to the following taxa: *Achnanthidium minutissimum* v. *minutissimum (*13.3), *Nitzschia palea* v. *palea* (9.3), *Encyonema minutum* (6.6), *Navicula notha* (6.3), *Brachysira neoexilis* (6), *Nitzschia paleaeformis* (5.8), *Nitzschia archibaldii* (3.2), *Synedra ulna* (3.1), *Nitzschia vasta* (3). *Eunotia bilunaris* and *Planothidium frequentissimum* are two examples of species with low Hill's N2 values (≤ 0.1).

The dataset was further reduced for the purpose of statistical analysis to include only taxa with a minimum occurrence across all samples of more than three. Across all sites a total of 112 taxa occurring at least at three sites were retained. The Hill's N2 index gave higher weights to the following taxa; *Synedra ulna* (14.5), *Encyonema minutum* (14.3), *Achnanthidium minutissimum* v. *minutissimum* (14.1), *Navicula notha* (12*), Nitzschia palea* v. *palea* (11.8), *Nitzschia archibaldii* (9.7), *Luticola mutica* (9.3), *Brachysira neoexilis* (8.8) and *Sellaphora pupula* v. *pupula* (8.8). The dominant taxa were *Aachnanthidium minutissimum* v. *minutissimum*, *Nitzschia paleaeformis*, *Nitzschia palea* v. *palea*, *Nitzschia vasta* and *Brachysira neoexilis*. Examples of taxa with the lowest N2 values are *Nitzschia pumila* (1.3) and *Pinnularia joculata* (1.2).

From the 50 sites sampled, 45 genera were identified (Appendix 7), although five genera, *Placoneis, Adlafia, Mastogloia, Lemnicola* and *Bacillaria* were excluded from analysis as they

occurred at only one site. Dominant genera include *Nitzschia*, *Achnanthidium*, *Navicula*, *Pinnularia* and *Gomphonema*. Hill's N2 gave higher weights to the following genera: *Nitzschia* (25.4), *Gomphonema* (18.1), *Achnanthidium* (16.2) and *Navicula* (15.7). The genera *Nupela* and *Stauroneis* did not occur at control sites. Control sites were dominated by the genus *Achnanthidium*, *Nitzschia* and *Navicula*. Combined, these genera accounted for over fifty percent of the counts at control sites. Over half of the genera had a Hill's N2 value of less than five. The most dominant genera at impacted sites are *Nitzschia*, *Achnanthidium* and *Navicula*, which account for sixty percent of generic abundance. At the most impacted sites (pH < 5) there are 34 genera and the sites are dominated by *Nitzschia*, *Achanthidium*, *Pinnularia* and *Eunotia* (> 68%). Genera with the highest number of occurrences were *Nitzschia* (50) and *Navicula* (46). Examples of genera which occurred only twice in the dataset are *Adlafia* and *Bacillaria*.

4.5.1 Iconograph

4.5.1.1 Introduction

Diatoms have been widely used in water quality monitoring and continue to be applied (Dixit *et al.* 1992b) because of their largely cosmopolitan distribution and their species-specific environmental requirements. However, one hurdle in using diatoms in the Northern Territory of Australia is the lack of documentation of diatoms from this tropical area. So far, few studies have been completed in tropical Australia and the diatom flora of this region remains poorly known. This work provides one of the first overviews of freshwater diatom floristic diversity in the area and should facilitate quantitative environmental studies based on diatom assemblages. In ecological monitoring and palaeoecological reconstructions accurate identification and knowledge of diatom geographic distributions are essential and it is crucial then that studies on the taxonomy and ecology of diatoms are linked (Kociolek & Stoermer 2001). This study has tried to address this by including the optima and tolerance of species for copper and pH.

Notes on the taxonomy, occurrence (#occ) and the effective number of occurrences (Hill's N2 1973 diversity measure) within the dataset, and autecological notes (optima and tolerances), accompanied by light microscope pictures (micrograph), are given for the most important 52 species (importance determined by high occurrence and Hill's N2 values). This is not a comprehensive list of all species noted (267 in total) but, an iconograph of the most common species in the region it is a useful first step. This will help future monitoring in keeping a consistent and correct identification and allow greater evaluation between studies. The scale for each micrograph is given so that measurements of the cell dimensions and the stria density can be made. For each species, one or two pertinent references are cited that contain illustrations that best

matched these specimens. Additionally, the occurrence of the taxa at control (C), impacted (I) and very impacted (VI) sites is given.

The estimated optima (opt.) and tolerances (tol.) were calculated for each species for copper and pH, by weighted averaging (WA). Ecological information of tolerance of species to copper and pH were chosen to represent heavy metal pollution and acidity. In WA regression, the optimum and tolerance for each taxon are estimated from the dataset based on their abundance in samples and the measured copper or pH data. These were developed with the program CALIBRATE version 3.0 (Juggins & ter Braak 1993). The full list of optima and tolerances for each species can be found in chapter 6, section 6.

Each of the 21 genera included within this iconograph are pennate diatoms and are divided into araphid, monoraphid, pseudoraphid and biraphid sections. The genera and species within these sections are displayed in alphabetical order. Araphid diatoms (e.g. *Fragilaria*) have no raphe and are generally bilaterally symmetrical. The monoraphid group (e.g. *Achnanthidium; Planothidium*) of diatoms are heterovalvar (two valves with a different structure) and have a raphe present in only one valve. The pseudoraphids (e.g. *Eunotia*) are diatoms with a short raphe present only at the poles. The biraphid group includes all diatoms with both valves containing a raphe system (i.e. *Chamaepinnularia, Craticula, Nitzschia, Encyonema, Encyonopsis, Epithemia, Gomphonema, Navicula, Naviculadicta, Nitzschia, Pinnularia, Sellaphora, Stenopterobia, Surirella, Eolimna, Brachysira*).

Araphid

Genus: ***Fragilaria***

Fragilaria capucina Desmazières 1925

Floras:	Krammer & Lange-Bertalot, 1991a (pl. 108, f. 110-113)	
Copper:	opt. ± tol.: 1.28 ± 3.99 mg/L	
pH:	opt. ± tol.: 6.8 ± 1.4	
#occ.:	10	
N2:	3.6	
Sites:	C, I, VI	(pl. 1: f. 5, 6)

Fragilaria tenera (W. Smith) Lange-Bertalot 1980

Floras:	Krammer & Lange-Bertalot, 1991a (pl. 114, f. 12-16) Lange-Bertalot & Metzeltin, 1996 (pl. 7, f. 1-3)	
Copper:	opt. ± tol.: 7.06 ± 17.72 mg/L	
pH:	opt. ± tol.: 5.1 ± 1.7	
#occ.:	10	
N2:	5.5	
Sites:	C, I, VI	(pl. 1: f. 1, 2, 3, 4)

Genus: ***Synedra***

Synedra ulna (Nitzsch) Ehrenberg 1836

Floras:	Krammer & Lange-Bertalot, 1991a (p. 143 fig. 119-122)	
Copper:	opt. ± tol.: 0.81 ± 2.46 mg/L	
pH:	opt. ± tol.: 6.7 ± 1.2	
#occ.:	31	
N2:	14.5	
Sites:	C, I, VI	(pl. 1: f. 2)

Monoraphid

Genus: ***Achnanthidium***

Achnanthidium minutissimum v. ***minutissimum*** Kützing 1833

Floras:	Krammer & Lange-Bertalot, 1991b (pl. 32-35)	
Copper:	opt. ± tol.: 0.88 ± 4.86 mg/L	
pH:	opt. ± tol.: 6.53 ± 1.45	
#occ.:	38	
N2:	4.1	
Sites:	C, I, VI	(pl. 6: f. 1, 2, 3, 4)

Achnanthidium minutissimum v. ***exilis*** Kützing 1833

Floras:	Krammer & Lange-Bertalot, 2004 (pl. 33, f. 23-31)	
Copper:	opt. ± tol.: 0.03 ± 0.04 mg/L	
pH:	opt. ± tol.: 7.6 ± 1.1	
#occ.:	6	
N2:	1.7	
Sites:	C, VI	(pl. 6: f. 5, 6, 7, 8)

Genus: ***Achnanthes***

Achnanthes exigua Grunow in Cleve & Grunow 1880

Floras:	Krammer & Lange-Bertalot, 1991a (pl. 38, f. 6 (4) f. 23 (1-27))	
Copper:	opt. ± tol.:1.98 ± 5.51 mg/L	
pH:	opt. ± tol.: 7.34 ± 1.45	
#occ.:	12	
N2:	5.9	
Sites:	C, I	(pl. 6: f. 9, 10, 11)

Genus: ***Planothidium***

Planothidium rostratum (Østrup) Round & Bukhtiyarova 1996

Floras:	Krammer & Lange-Bertalot, 2004 (pl. 43, f. 1-26)	
Copper:	opt. ± tol.: 1.66 ± 4.42 mg/L	
pH:	opt. ± tol.: 7.3 ± 1.8	
#occ.:	17	
N2:	7.7	
Sites:	C, I, VI	(pl. 6: f. 12, 13, 14)

Pseudoraphid

Genus: **Eunotia**

Eunotia bilunaris (Ehrenberg) Mills 1934

Floras:	Krammer & Lange-Bertalot, 1991a (pl. 137, f. 1-7)	
Copper:	opt. ± tol.: 30.3 ± 47.53 mg/L	
pH:	opt. ± tol.: 4.85 ± 1.58	
#occ.:	18	
N2:	6.6	
Sites:	C, I, VI	(pl. 6: f. 19)

Eunotia camelus v. ***didymodon (***Grunow) Frenguelli 1933

Floras:	Metzeltin & Lange-Bertalot, 1998 (pl. 29, f. 3-5)	
Copper:	opt. ± tol.: 38.36 ± 50.54 mg/L	
pH:	opt. ± tol.: 4.9 ± 1.6	
#occ.:	14	
N2:	5.7	
Sites:	C, I, VI	(pl. 6: f. 20)

Eunotia naegelii Migula in Thomas 1907

Floras:	Krammer & Lange-Bertalot, 1991a (pl. 140, f. 1-6) Lange-Bertalot & Mezeltin, 1996 (pl. 9, f. 8-13)	
Copper:	opt. ± tol.: 19.16 ± 35.23 mg/L	
pH:	opt. ± tol.: 4.8 ± 1.4	
#occ.:	16	
N2:	7.9	
Sites:	C, I, VI	(pl. 6: f. 25, 26)

Eunotia cf. ***rabenhorstiana*** (Grunow) Hustedt 1949

Floras:	Lange-Bertalot, 1998 (pl. 10, f. 1-13)	
Copper:	opt. ± tol.: 2.29 ± 4.6 mg/L	
pH:	opt. ± tol.: 5.7 ± 1.2	
#occ.:	10	
N2:	7.3	
Sites:	I, VI	(pl. 6: f. 21, 22)

Eunotia veneris (Kützing) De Toni 1892

Floras:	Krammer & Lange-Bertalot, 2004. (pl 163, f. 14-19)	
Copper:	opt. ± tol.: 7.06 ± 25.46 mg/L	
pH:	opt. ± tol.: 6.15 ± 1.44	
#occ.:	14	
N2:	6.3	
Sites:	C, I, VI	(pl. 6: f. 23)

Eunotia spec.9

Copper:	opt. ± tol.: 6.31 ± 22.21 mg/L	
pH:	opt. ± tol.: 6.1 ± 1.1	
#occ.:	11	
N2:	8.6	
Sites:	C, I, VI	(pl. 6: f. 24)

Biraphid

Genus: ***Brachysira***

Brachysira neoexilis Lange-Bertalot 1994

Floras:	Lange-Bertalot & Metzeltin, 1996 (pl. 36, f. 24-28)	
Copper:	opt. ± tol.: 7.49 ± 27.05 mg/L	
pH:	opt. ± tol.: 6.4 ± 1.3	
#occ.:	31	
N2:	8.8	
Sites:	C, I, VI	(pl. 6: f. 15, 16, 17, 18)

Genus: ***Chamaepinnularia***

Chamaepinnularia mediocris (Krasske) Lange-Bertalot 1932

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 78, f. 14-16)	
Copper:	opt. ± tol.: 7.29 ± 21.77 mg/L	
pH:	opt. ± tol.: 4.3 ± 1.2	
#occ.:	11	
N2:	4.7	
Sites:	C, I, VI	(pl. 1: f. 8)

Genus: ***Craticula***

Craticula halophilioides nov. comb. (Hustedt) Lange-Bertalot

Floras:	Lange-Bertalot, 2001 (pl. 91, f. 1-7)	
Copper:	opt. ± tol.: 0.55 ± 2.45 mg/L	
pH:	opt. ± tol.: 6.9 ± 1.1	
#occ.:	15	
N2:	5.1	
Sites:	C, I, VI	(pl. 1: f. 9, 10)

Genus: ***Encyonema***

Encyonema minutum (Hilse ex Rabenhorst) Mann 1990

Floras:	Krammer & Lange-Bertalot 1986 (pl. 305, f. 119: 1-13)	
Copper:	opt. ± tol.: 2.44 ± 10.17 mg/L	
pH:	opt. ± tol.: 6.15 ± 1.30	
#occ.:	27	
N2:	14.3	
Sites:	C, I, VI	(pl. 1: f. 15, 16)

Encyonema silesiacum (Bleisch in Rabenhorst) Mann 1990

Floras:	Krammer & Lange-Bertalot, 1991a (pl. 64, f. 1-8)	
Copper:	opt. ± tol.: 2.48 ± 5.68 mg/L	
pH:	opt. ± tol.: 6.25 ± 1.29	
#occ.:	22	
N2:	6.5	
Sites:	C, I, VI	(pl. 1: f. 11, 12, 13, 14)

Genus: ***Encyonopsis***

Encyonopsis perborealis nov. spec. Krammer

Floras:	Krammer, 1997 (pl. 154, f. 1-12)	
Copper:	opt. ± tol.: 3.39 ± 15.31 mg/L	
pH:	opt. ± tol.: 6.7 ± 1.69	
#occ.:	12	
N2:	6.1	
Sites:	C, I, VI	(pl. 1: f. 17, 18, 19)

Genus: ***Eolimna***

Eolimna minima (Grunow) Lange-Bertalot 1997

Floras:	Krammer & Lange-Bertalot, 1986 (pl. 229, f. 76: 39-47)	
Copper:	opt. ± tol.: 1.79 ± 5.38 mg/L	
pH:	opt. ± tol.: 7.2 ± 1.2	
#occ.:	6	
N2:	5.3	
Sites:	C, I	(pl. 5: f. 13, 14)

Genus: ***Epithemia***

Epithemia cf. ***cistula*** (Ehrenberg) Ralfs in Pritchard 1861

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 105, f. 7-11)	
Copper:	opt. ± tol.: 0.17 ± 1.68 mg/L	
pH:	opt. ± tol.: 8.3 ± 0.5	
#occ.:	7	
N2:	3.8	
Sites:	C, I	(pl. 1: f. 20, 21)

Genus: ***Gomphonema***

Gomphonema cf. ***exilissimum*** nov. stat. (Grunow) Lange-Bertalot & Reichardt

Floras:	Lange-Bertalot 1996. (T. 62, f. 23-27)	
Copper:	opt. ± tol.: 1.1 ± 3.64 mg/L	
pH:	opt. ± tol.: 6.6 ± 1	
#occ.:	16	
N2:	7.9	
Sites:	C, I, VI	(pl. 2: f. 1, 2, 3)

Gomphonema cf. ***exilissimum*** group 2

Copper:	opt. ± tol.: 3.59 ± 4.84 mg/L	
pH:	opt. ± tol.: 5.4 ± 1.7	
#occ.:	12	
N2:	6.3	
Sites:	C, I, VI	(pl. 2: f. 4, 5, 6)

Gomphonema gracile Ehrenberg 1838

Floras:	Krammer & Lange-Bertalot, 1986 (pl. 361, f. 154, 26,27)	
Copper:	opt. ± tol.: 0.21 ± 0.97 mg/L	
pH:	opt. ± tol.: 6.6 ± 0.9	
#occ.:	17	
N2:	7.4	
Sites:	C, I, VI	(pl. 2: f. 13, 14)

Gomphonema lagenula Kützing 1844

Floras:	Krammer & Lange-Bertalot, 1986 (pl. 358, f. 154)	
Copper:	opt. ± tol.: 2.05 ± 4.69 mg/L	
pH:	opt. ± tol.: 6.6 ± 1.7	
#occ.:	17	
N2:	8.3	
Sites:	C, I, VI	(pl. 2: f. 7, 8, 9)

Gomphonema cf. ***vibrioides*** Reichardt & Lange-Bertalot 1991

Floras:	Lange-Bertalot, 2000 (T.155, f. 1-4)	
Copper:	opt. ± tol.: 0.84 ± 3.23 mg/L	
pH:	opt. ± tol.: 6.8 ± 1	
#occ.:	11	

N2: 6.2
Sites: C, I, VI (pl. 2: f. 15, 16, 17)

Gomphonema spec. 15

Copper: opt. ± tol.: 0.60 ± 0.76 mg/L
pH: opt. ± tol.: 5.8 ± 1.2
#occ.: 10
N2: 5.2
Sites: C, I, VI (pl. 2: f. 10, 11, 12)

Genus: ***Luticola***

Luticola mutica (Kützing 1844) Mann in Round *et al.* 1990

Floras: Krammer & Lange-Bertalot, 1997 (pl. 61, f. 1-11)
Copper: opt. ± tol.: 15.41 ± 33.75 mg/L
pH: opt. ± tol.: 5.7 ± 1.8
#occ.: 15
N2: 9.3
Sites: C, I (pl. 2: f. 18, 19)

Genus: ***Navicula***

Navicula cincta f. ***minuta*** Grunow in van Heurck 1885

Floras: Krammer & Lange-Bertalot, 1997 (pl. 28 f. 16)
Copper: opt. ± tol.: 2.49 ± 6.12 mg/L
pH: opt. ± tol.: 7.1 ± 1.4
#occ.: 7
N2: 5.6
Sites: C, I (pl. 2: f. 20)

Navicula gerloffii Schimanski 1978

Floras: Krammer & Lange-Bertalot, 1986 (pl. 80, f. 18-21)
Copper: opt. ± tol.: 1.26 ± 3.91 mg/L
pH: opt. ± tol.: 6.5 ± 1.5
#occ.: 19
N2: 7.6
Sites: C, I, VI (pl. 3: f. 1)

Navicula heimansioides Lange-Bertalot 1993

Floras: Krammer & Lange-Bertalot, 1991 (pl. 388, f. 70. 1-8)
Copper: opt. ± tol.: 0.74 ± 2.56
pH: opt. ± tol.: 6.4 ± 1.1
#occ.: 20
N2: 7.6
Sites: C, I, VI (pl. 3: f. 2, 3, 4)

Navicula notha Wallace 1960

Floras: Krammer & Lange-Bertalot, 1991b (pl. 70, f. 19)
Copper: opt. ± tol.: 0.96 ± 6.37 mg/L

pH:	opt. ± tol.: 6.9 ± 1.1	
#occ.:	28	
N2:	12	
Sites:	C, I, VI	(pl. 3: f. 5, 6, 7)

Navicula radiosa Kützing 1844

Floras:	Krammer & Lange-Bertalot, 1986 (pl. 29, f. 1,2,4)	
Copper:	opt. ± tol.: 1.37 ± 4.65 mg/L	
pH:	opt. ± tol.: 7 ± 1.1	
#occ.:	11	
N2:	5.9	
Sites:	C, I, VI	(pl. 3: f. 8)

Navicula rostellata Kützing 1844

Floras:	Lange-Bertalot, 2001 (pl. 35, f. 1-6)	
Copper:	opt. ± tol.: 0.06 ± 0.08 mg/L	
pH:	opt. ± tol.: 8 ± 1	
#occ.:	6	
N2:	5.2	
Sites:	C, I, VI	(pl. 3: f. 9, 10)

Navicula spec. 31

Copper:	opt. ± tol.: 0.59 ± 2.42 mg/L	
pH:	opt. ± tol.: 6.8 ± 1.5	
#occ.:	20	
N2:	2.8	
Sites:	C, I, VI	(pl. 3: f. 14, 15, 16)

Genus: ***Naviculadicta***

Naviculadicta difficillima Hustedt 1950

Floras:	Krammer & Lange-Bertalot, 1986 (pl. 80, f. 7,8)	
Copper:	opt. ± tol.: 0.69 ± 2.11 mg/L	
pH:	opt. ± tol.: 6.5 ± 1.3	
#occ.:	11	
N2:	8.6	
Sites:	C, I, VI	(pl. 2: f. 21)

Naviculadicta subtilissima Cleve 1891

Floras:	Lange-Bertalot & Metzeltin, 1996 (pl. 24, f. 4-6)	
	Krammer & Lange-Bertalot, 1997 (pl. 79, f. 22-26)	
Copper:	opt. ± tol.: 11.14 ± 24.58 mg/L	
pH:	opt. ± tol.: 4.6 ± 1.1	
#occ.:	10	
N2:	6.7	
Sites:	C, I, VI	(pl. 3: f. 11, 12)

Naviculadicta tridentulata Krasske 1923

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 80, f. 1-3)	

Copper: opt. ± tol.: 0.38 ± 1.87 mg/L
pH: opt. ± tol.: 6.7 ± 1
#occ.: 11
N2: 4.4
Sites: C, I, VI (pl. 3: f. 13)

Genus: ***Nitzschia***

Nitzschia amphibia Grunow 1862

Floras: Krammer & Lange-Bertalot, 1988 (pl. 78, f. 13-26)
Copper: opt. ± tol.: 1.51 ± 4.89 mg/L
pH: opt. ± tol.: 7.4 ± 1.3
#occ.: 9
N2: 6.1
Sites: C, I, VI (pl. 3: f. 17, 18)

Nitzschia archibaldii Lange-Bertalot 1980

Floras: Krammer & Lange-Bertalot, 1997 (pl. 81, f. 10-12)
Copper: opt. ± tol.: 1.89 ± 10.74 mg/L
pH: opt. ± tol.: 6.5 ± 1.7
#occ.: 19
N2: 9.7
Sites: C, I, VI (pl. 3: f. 19, 20)

Nitzschia gracilis Hantzsch 1860

Floras: Krammer & Lange-Bertalot, 1988 (pl. 66, f. 1-11)
Copper: opt. ± tol.: 5.49 ± 9.86 mg/L
pH: opt. ± tol.: 4.7 ± 1.6
#occ.: 13
N2: 5.4
Sites: C, I, VI (pl. 3: f. 24, 25)

Nitzschia cf. ***hantzschiana*** Rabenhorst 1860

Floras: Krammer & Lange-Bertalot, 1988 (pl. 73, f. 9-18)
Copper: opt. ± tol.: 0.04 ± 0.22 mg/L
pH: opt. ± tol.: 7.1 ± 0.7
#occ.: 15
N2: 6.2
Sites: C, I, VI (pl. 3: f. 21, 22, 23)

Nitzschia intermedia Hantzsch ex Cleve & Grunow 1880

Floras: Krammer & Lange-Bertalot, 1997 (pl. 61, f. 1-10)
Copper: opt. ± tol.: 0.09 ± 0.29 mg/L
pH: opt. ± tol.: 7.5 ± 1.2
#occ.: 10
N2: 3.3
Sites: C, I, VI (pl. 4: f. 1)

Nitzschia liebetruthii Rabenhorst 1864

Floras: Krammer & Lange-Bertalot, 1988 (pl. 96, f. 69, 14-32)

Copper:	opt. ± tol.: 0.69 ± 9.06 mg/L	
pH:	opt. ± tol.: 8.2 ± 0.8	
#occ.:	10	
N2:	3.7	
Sites:	C, I, VI	(pl. 4: f. 19, 20, 21)

Nitzschia nana Grunow in van Huerck 1881

Floras:	Lange-Bertalot, 2000 (pl. 206, f. 15-18)	
Copper:	opt. ± tol.: 0.57 ± 2.73 mg/L	
pH:	opt. ± tol.: 5 ± 1.4	
#occ.:	18	
N2:	2.9	
Sites:	C, I, VI	(pl. 4: f. 2, 3, 4)

Nitzschia palea v. ***palea*** (Kützing) W. Smith 1856

Floras:	Krammer & Lange-Bertalot, 1988 (pl. 85, f. 59, 1-24)	
Copper:	opt. ± tol.: 0.58 ± 2.52 mg/L	
pH:	opt. ± tol.: 7 ± 1.1	
#occ.:	38	
N2:	11.8	
Sites:	C, I, VI	(pl. 4: f. 14, 15, 16)

Nitzschia palea v. ***tenuirostris*** Grunow in Van Heurck 1881

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 59, f. 19-23)	
Copper:	opt. ± tol.: 1.1 ± 11.07 mg/L	
pH:	opt. ± tol.: 7.9 ± 1	
#occ.:	12	
N2:	4.4	
Sites:	C, I, VI	(pl. 4: f. 8, 9, 10)

Nitzschia paleaeformis Hustedt 1950

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 65, f. 3-8A)	
Copper:	opt. ± tol.: 0.94 ± 2.59 mg/L	
pH:	opt. ± tol.: 5.5 ± 1.3	
#occ.:	15	
N2:	6.7	
Sites:	I, VI	(pl. 4: f. 11, 12, 13)

Nitzschia cf. ***pseudofonticola*** Hustedt 1942

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 63, f. 11-13)	
Copper:	opt. ± tol.: 0.01 ± 0.01 mg/L	
pH:	opt. ± tol.: 7.2 ± 0.2	
#occ.:	11	
N2:	5.5	
Sites:	C, I, VI	(pl. 4: f. 5)

Nitzschia vasta Hustedt 1939

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 56, f. 8, 8a)	
Copper:	opt. ± tol.: 19.05 ± 34.71 mg/L	
pH:	opt. ± tol.: 3.9 ± 1.5	
#occ.:	15	

N2: 4
Sites: C, I, VI (pl. 4: f. 17, 18)

Nitzschia spec. 39

Copper: opt. ± tol.: 0.03 ± 0.06 mg/L
pH: opt. ± tol.: 7.8 ± 0.9
#occ.: 13
N2: 6.6
Sites: C, I, VI (pl. 4: f. 6, 7)

Genus: ***Pinnularia***

Pinnularia joculata (Manguin) Krammer 2000

Floras: Krammer, 2000 (pl. 13, f. 5-9)
Copper: opt. ± tol.: 0.24 ± 1.93 mg/L
pH: opt. ± tol.: 6.3 ± 1.5
#occ.: 5
N2: 1.2
Sites: I, VI (pl. 5: f. 4, 5, 6)

Pinnularia subcapitata Gregory 1856

Floras: Krammer, 1992 (pl. 107, fig 37: 17-31)
Copper: opt. ± tol.: 10.57 ± 28.45 mg/L
pH: opt. ± tol.: 5.4 ± 1.6
#occ.: 18
N2: 6.9
Sites: C, I, VI (pl. 5: f. 1, 2, 3)

Genus: ***Sellaphora***

Sellaphora pupula v. ***pupula*** Kützing 1844

Floras: Krammer & Lange-Bertalot, 1997 (pl. 68, f. 1-12)
Copper: opt. ± tol.: 0.31 ± 1.91 mg/L
pH: opt. ± tol.: 7 ± 0.9
#occ.: 27
N2: 8.8
Sites: C, I, VI (pl. 5: f. 7, 8, 9)

Genus: ***Stenopterobia***

Stenopterobia densestriata (Hustedt) Krammer 1987

Floras: Lange-Bertalot & Metzeltin, 1996 (pl.74, f. 9)
Copper: opt. ± tol.: 1.74 ± 12.69 mg/L
pH: opt. ± tol.: 6.4 ± 1.2
#occ.: 12
N2: 5.9
Sites: C, I, VI (pl. 5: f. 12)

Genus: ***Surirella***

Surirella roba Leclerq 1983

Floras:	Krammer & Lange-Bertalot, 1997 (pl. 148, f. 5-9)	
Copper:	opt. ± tol.: 0.21 ±1. 16 mg/L	
pH:	opt. ± tol.: 5.5 ± 1.3	
#occ.:	13	
N2:	2.3	
Sites:	C, I, VI	(pl. 5: f. 10, 11)

Plate 1 (x1500)

Figs. 1, 2, 3, 4: *Fragilaria tenera* (W. Smith) Lange-Bertalot 1980

Figs. 5, 6: *Fragilaria capucina* Desmazières 1925

Fig. 7: *Synedra ulna* Ehrenberg 1836

Fig. 8: *Chamaepinnularia mediocris* (Krasske) Lange-Bertalot 1932

Figs. 9, 10: *Craticula halophilioides* nov. comb. (Hustedt) Lange-Bertalot

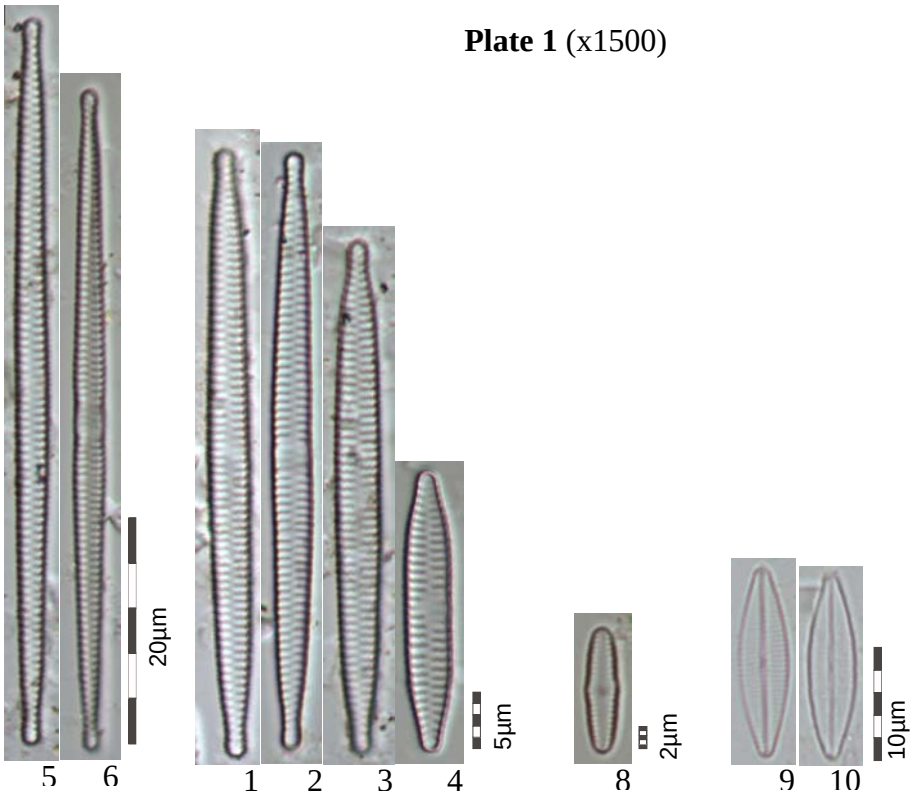
Figs. 11, 12, 13, 14: *Encyonema silesiacum* (Bleisch in Rabenhorst) Mann 1990

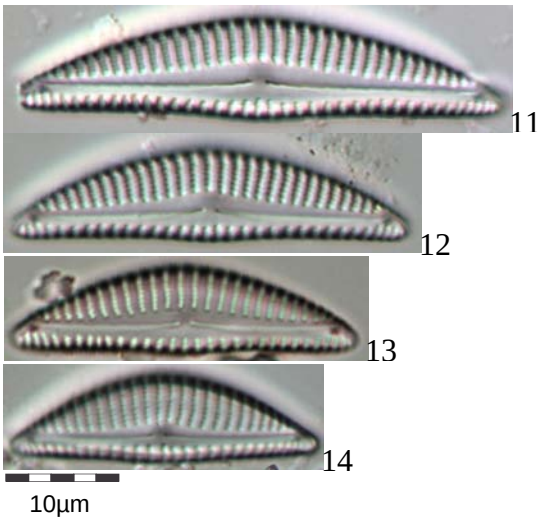
Figs. 15, 16: *Encyonema minutum* (Hilse ex Rabenhorst) Mann 1990

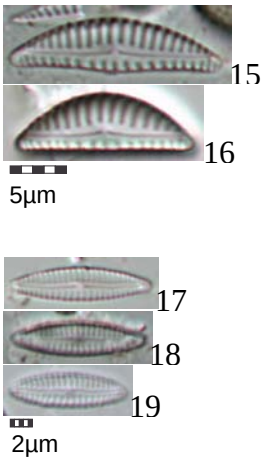
Figs. 17, 18, 19: *Encyonopsis perborealis* nov. spec. Krammer

Figs. 20, 21: *Epithemia* cf. *cistula* (Ehrenberg) Ralfs in Pritchard 1861

Plate 1 (x1500)







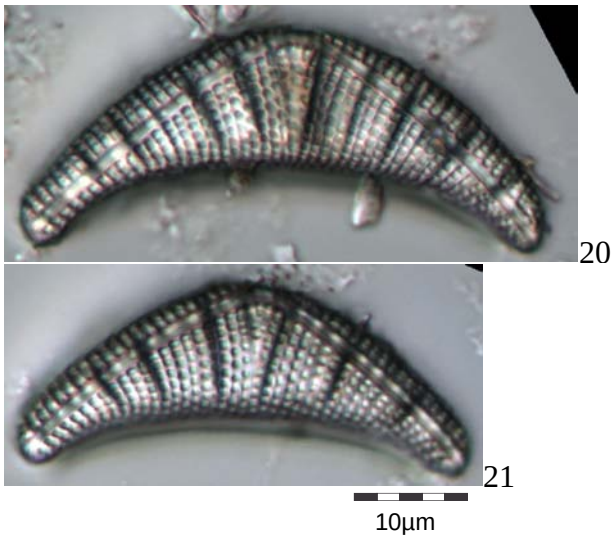


Plate 2 (x1500)

Figs.1, 2, 3: *Gomphonema* cf. *exilissimum* nov. stat. (Grunow) Lange-Bertalot & Reichardt

Figs. 4, 5, 6: *Gomphonema* cf. *exilissimum* group 2

Figs. 7, 8, 9: *Gomphonema lagenula* Kützing 1844

Figs. 10, 11, 12: *Gomphonema* spec. 15

Figs. 13, 14: *Gomphonema gracile* Ehrenberg 1838

Figs. 15, 16, 17: *Gomphonema* cf. *vibrioides* Reichardt & Lange-Bertalot 1991

Figs. 18, 19: *Luticola mutica (*Kützing 1844*)* Mann in Round *et al.* 1990

Fig. 20 : *Navicula cinta* f. *minuta* Grunow in van Heurck 1885

Fig. 21 : *Naviculadicta difficillima* Hustedt 1950

Plate 2 (x1500)

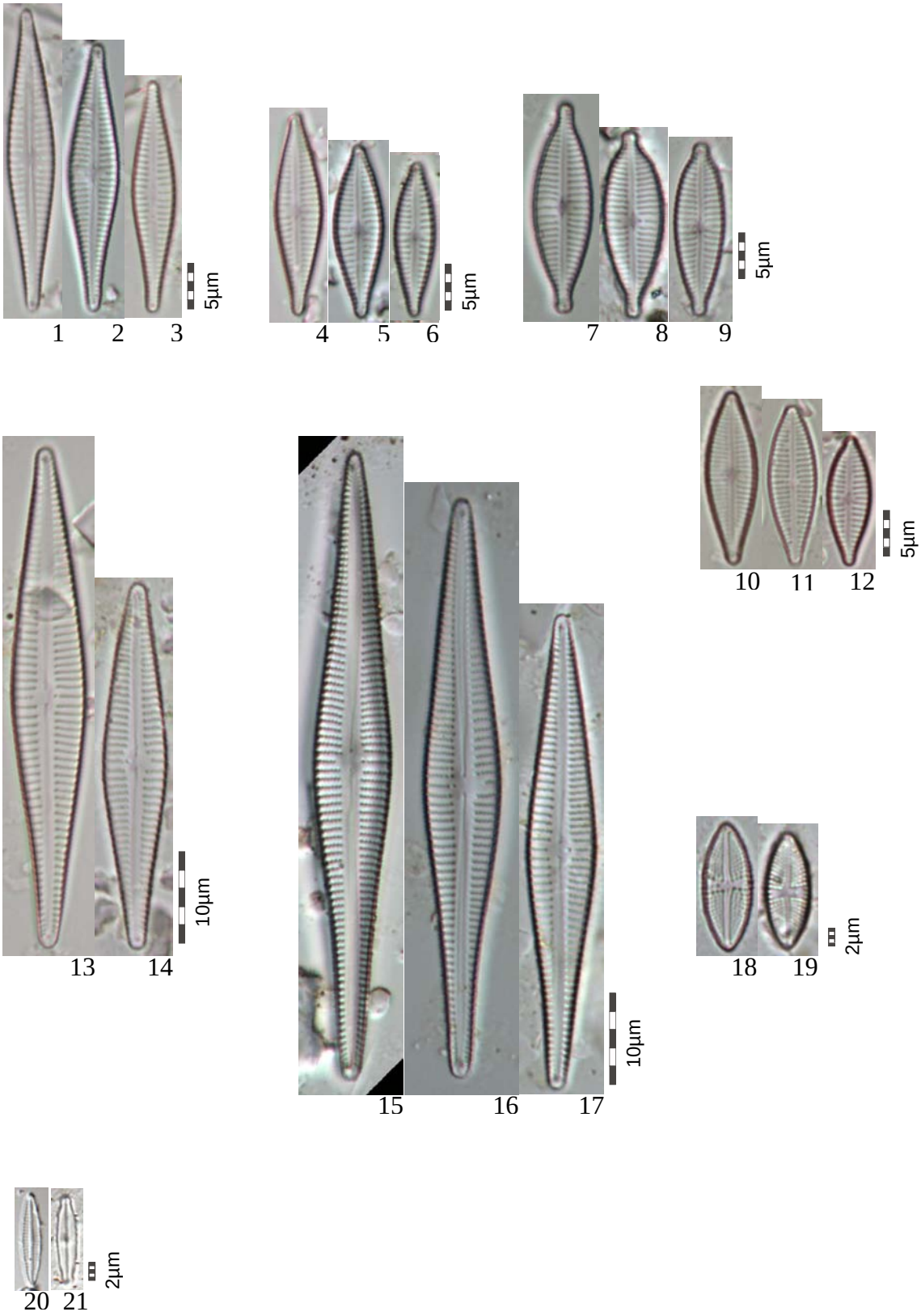


Plate 3 (x1500)

Fig. 1:	*Navicula gerloffii* Schimaski 1978
Figs. 2, 3, 4:	*Navicula heimansioides* Lange-Bertalot 1993
Figs. 5, 6, 7:	*Navicula notha* Wallace 1960
Fig. 8:	*Navicula radiosa* Kützing 1844
Figs. 9, 10:	*Navicula rostellata* Kützing 1844
Figs. 11, 12:	*Naviculadicta subtilissima* Cleve 1891
Fig. 13:	*Navicula tridentulata* Krasske 1923
Figs. 14, 15, 16:	*Navicula* spec. 31
Figs. 17, 18:	*Nitzschia amphibia* Grunow 1862
Figs. 19, 20:	*Nitzschia archibaldii* Lange-Bertalot 1980
Figs. 21, 22, 23:	*Nitzschia* cf. *hantzschiana* Rabenhorst 1860
Figs. 24, 25	*Nitzschia gracilis* Hantzsch 1860

Plate 3 (x1500)

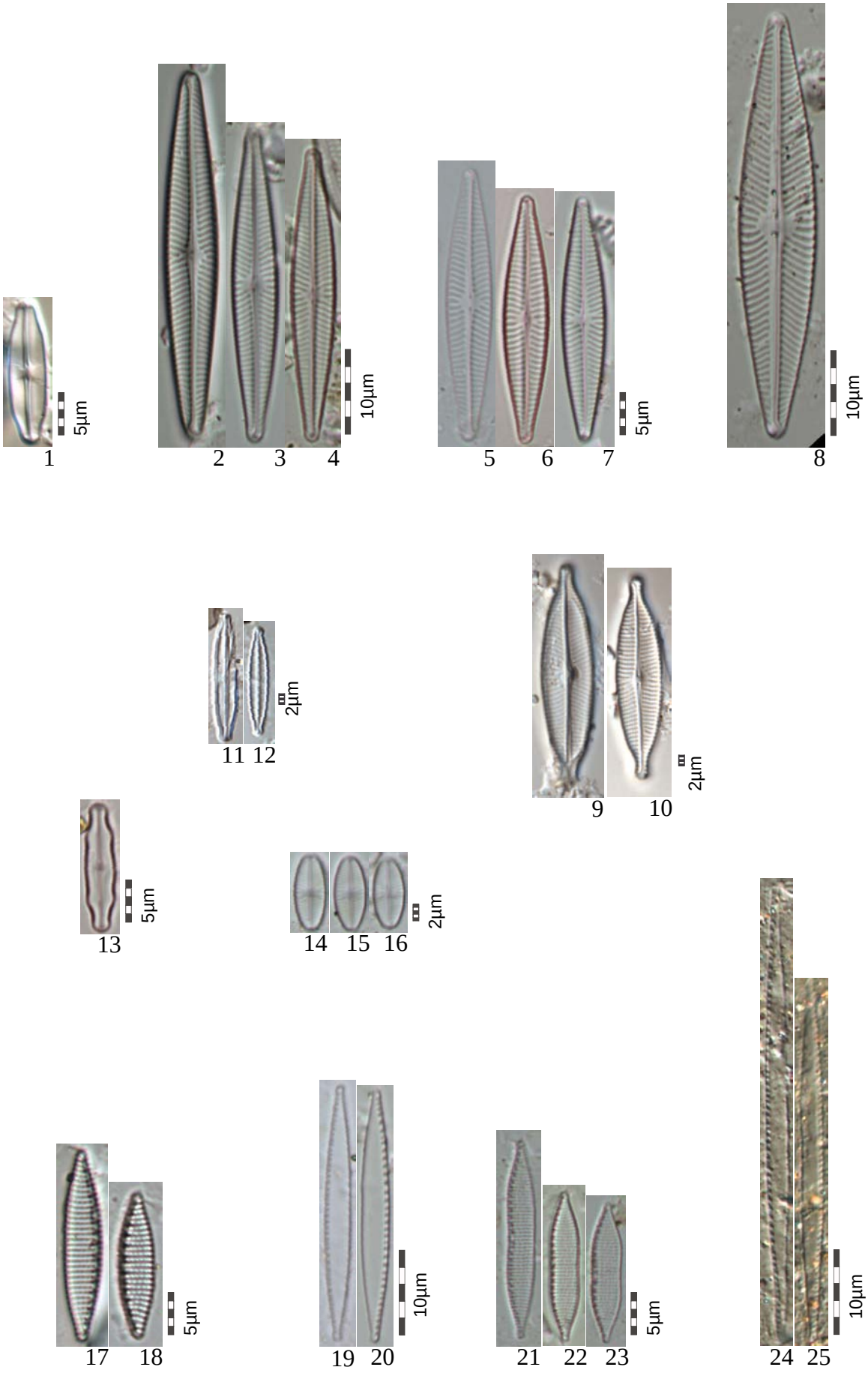


Plate 4 (x1500)

Fig. 1:	*Nitzschia intermedia* Hantzsch ex Cleve & Grunow 1880
Figs. 2, 3, 4:	*Nitzschia nana* Grunow in van Huerck 1881
Fig. 5:	*Nitzschia* cf. *pseudofonticola* Hustedt 1942
Figs. 6, 7:	*Nitzschia* spec. 39
Figs. 8, 9, 10:	*Nitzschia palea* v. *tenuirostris* Grunow in Van Heurck 1881
Figs. 11, 12, 13:	*Nitzschia paleaeformis* Hustedt 1950
Figs. 14, 15, 16:	*Nitzschia palea* v. *palea* (Kützing) W.Smith 1856
Figs. 17, 18:	*Nitzschia vasta* Hustedt 1939
Figs. 19, 20, 21:	*Nitzschia liebetruthii* Rabenhorst 1864

Plate 4 (x1500)

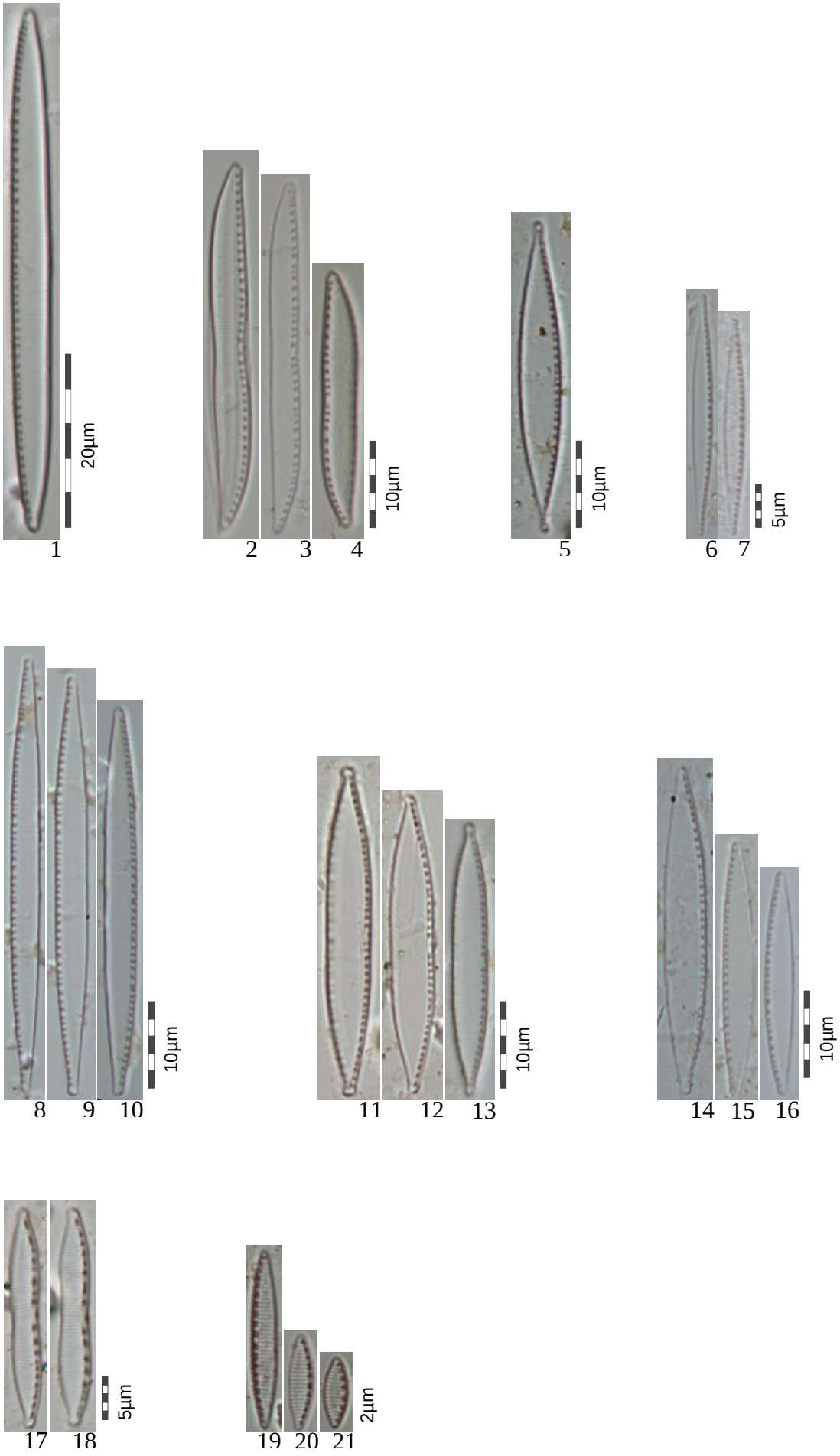


Plate 5 (x1500)

Figs. 1, 2, 3: *Pinnularia subcapitata* Gregory 1856

Figs. 4, 5, 6: *Pinnularia joculata* (Manguin) Krammer 2000

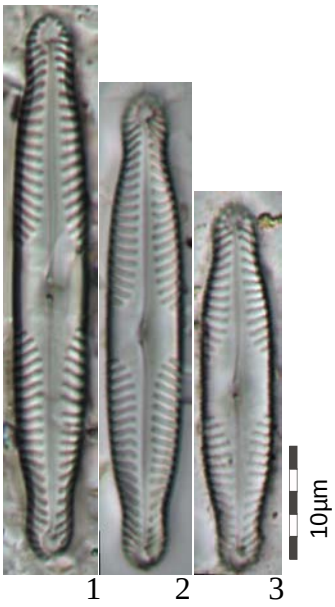
Figs. 7, 8, 9: *Sellaphora pupula* v. *pupula* Kützing 1844

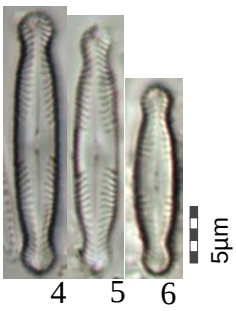
Figs. 10, 11: *Surirella roba* Leclerq 1983

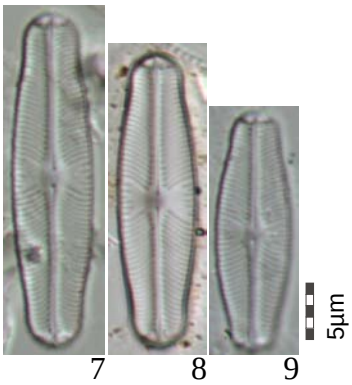
Fig. 12: *Stenopterobia densestriata* (Hustedt) Krammer 1987

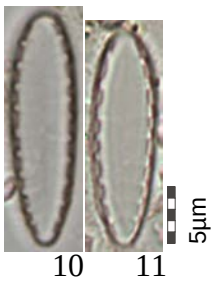
Figs. 13, 14 *Eolimna minima* (Grunow) Lange-Bertalot 1997

Plate 5 (x1500)









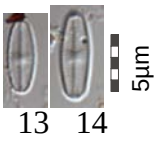


Plate 6 (x1500)

Figs. 1, 2, 3, 4: *Achnanthidium minutissimum* v. *minutissimum* Kützing 1833

Figs. 5, 6, 7, 8: *Achnanthidium minutissimum* v. *exilis* Kützing 1833

Figs. 9, 10, 11: *Achnanthidium exigua* Grunow in Cleve & Grunow 1880

Figs. 12, 13, 14: *Planothidium rostratum* (Østrup) Round & Bukhtiyarova 1996

Figs. 15, 16, 17, 18: *Brachysira neoexilis* Lange-Bertalot 1994

Fig. 19: *Eunotia bilunaris* (Ehrenberg) Mills 1934

Fig. 20: *Eunotia camelus* v. *didymodon* (Grunow) Frenguelli 1933

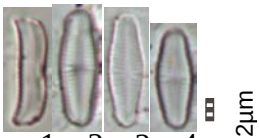
Figs. 21, 22: *Eunotia* cf. *rabenhorstiana* (Grunow) Hustedt 1949

Fig. 23: *Eunotia veneris* (Kützing) De Toni 1892

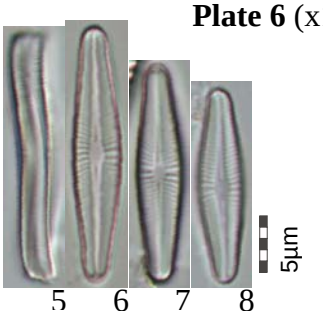
Fig. 24: *Eunotia* spec.9

Figs. 25, 26: *Eunotia naegelii* Migula in Thomas 1907

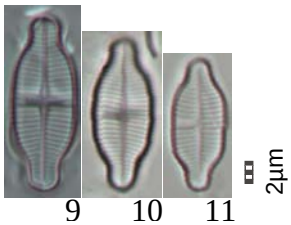
Plate 6 (x1500)



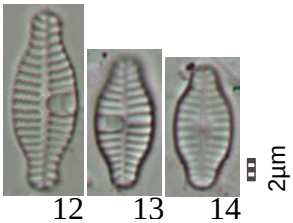
1 2 3 4



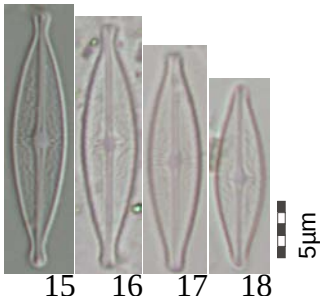
5 6 7 8



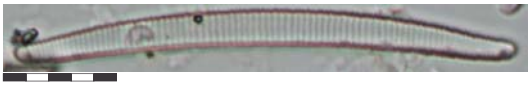
9 10 11



12 13 14



15 16 17 18



19

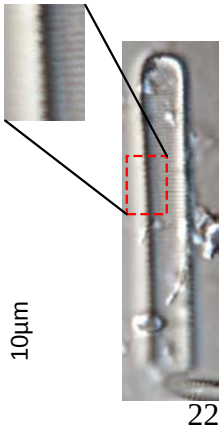
20

23

24

25 26

21



22

CHAPTER 5 - Genetics

5.1 Introduction

Diatoms are increasingly being used in ecological monitoring and paleoecological reconstructions (Stoermer & Smol 1999) in which accurate identification of species and knowledge of geographical distributions is essential. Despite this, diatom taxonomy and the species concept for phycologists is primarily based on the morphological characteristics rather than biological species concept {i.e. Mayr (1982); a species is a reproductive community of populations, reproductively isolated from others, that occupies a specific niche in nature}.

As the most species rich and productive group of eukaryotic algae, diatoms are recognised for their intricate bipartite, siliceous cell wall called the frustule (Round *et al.* 1990). Valve structure of the frustule has been studied more than any other aspect of the diatom cell (Round *et al.* 1990). Diatom classification depends to a great extent upon the intricacies of pore structure and the arrangement of the wall organelles (i.e. raphes, ocelli) as the valves carry most of the taxonomic features used in standard floras. While most key features of diatom valves can be recognised in the light microscope (LM), use of scanning electron microscopes (SEM) since the 1960's has revealed finer features and new structures of the cell wall (Battarbee *et al.* 2001) which have helped to define more species. However, the overall conclusion from purely morphometric studies, involving statistical treatment of morphological data and image analysis of taxa, is that traditional visual assessment considerably underestimates the complexity of species and infraspecific variation pattern (Mann 1999). For instance, the morphological characteristics, length, width, stria and fibula density, of species resembling *Nitzschia parvula* (Witkowski *et al.* 2004) was analysed by principle component analysis (PCA), resulting in the separation of four new species. Similarly, eleven new species of *Gomphonema* were described from Tasmania based on morphological characteristics (Kociolek *et al.* 2004).

In recent debates on species concepts in diatoms (Mann 1999, 2001) the need for culturing (breeding experiments) and molecular genetic data has been emphasised. Mann (2001) argued that identification based solely on morphological criteria may be misleading and that, to verify distribution, taxonomy should be based on a combination of morphological, molecular and breeding data. One reason why identification based solely on morphology may be misleading is that variation in valve morphology may be phenotypic or genetically caused. Morphological differences can be caused by physiological or chemical differences, for instance, salinity (Jahn 1986, Cox 1995, Trobajo *et al.* 2006) and temperature (Syversten 1977) affect diatom cell size, life cycle dynamics and morphology (i.e. Potapova & Snoeijs 1997; Snoeijs & Potapova 1998). Additionally, one problem is in defining the range of morphotypes to be included in a species and

distinguishing where the breaks occur between these and their nearest neighbours (Round *et al.* 1990). Consequently it is difficult to determine whether morphological variation within species indicates the existence of new species or simply the extreme ends of morphological ranges. In the *Gomphoneis herculeana* complex, some varieties have turned out to represent nothing more than larger or smaller ends of a particular size diminution series, while other specimens ascribed to the same complex are indeed different species (Kociolek & Stoermer 1988). Furthermore, only a tiny fraction of the total genotypic variation within a population or taxon is likely to manifest morphologically (Mann 1999). Population genetic studies have shown a high degree of invisible genetic differentiation between geographically or temporally separated populations of what is supposed to be the same species (Mann & Chepurnov 2004). Genetic studies on the species *Asterionella formosa* (Soudek & Robinson 1983), *Fragilaria capucina* (Lewis *et al.* 1997) and *Skeletonema costatum* (Medlin *et al.* 1988) have shown high levels of genetic diversity among populations of what is morphologically the same species. One instance where genetic variation has been coupled with morphological variation is in the morphologically variable complex *Sellaphora pupula* (Mann *et al.* 2004). Mann describes five new species within the *Sellaphora pupula* species complex based on morphometric analysis and mating data which is supported by molecular sequence data. During the last twenty years there has been an increase in the number of culture based studies and studies applying molecular data to examine phylogenetic and biogeographical relationships in diatoms (i.e. Medlin *et al.* 1991, 1993, 1996, 2000; Lundholm *et al.* 2003; Behnke *et al.* 2004). These studies have improved the taxonomy of various groups and added new knowledge on diatom phenotypic plasticity, mating systems, population genetics and molecular phylogeny (Mann & Chepurnov 2004).

The aim of this chapter is to examine morphologically variable *Nitzschia palea*-like cells collected from sites in the Northern Territory of Australia and to group the cultured cells according to their morphological characteristics. The cultures are then grouped by genetic techniques in order to ascertain whether the groups differ depending on the method used and to define whether optical or genetic methods are needed for the classification of *N. palea* cells. The cultures are compared genetically to sequenced *N. palea* collected from European sites to determine if they are the same species. Each genetically defined species (from the cultured diatoms) is described based on morphological characteristics of live and dead cells.

Nitzschia Hassall is a large and difficult genus of biraphid diatoms to identify. In the 19th Century, the genus was split into numerous sections (Witkowski *et al.* 2004). *Nitzschia palea* {(Kützing) W. Smith 1856} is accepted as being a cosmopolitan, ubiquitous diatom which is very abundant and has a wide tolerance range of pH. It is frequently found in high numbers in hypertrophic waters and has been noted in areas of extreme metal pollution (Dickman *et al.* 1990; Sabater 2000; Salomoni *et al.* 2006) and acidity (Jüttner *et al.* 2003) as well as in areas of high nitrate and phosphate

eutrophication (Maznah *et al.* 2002; Jüttner *et al.* 2003; Salomoni *et al.* 2006). *Nitzschia palea* has many varieties and can be often difficult to distinguish from one another under light microscope (i.e. var. *tenuirostris*, var. *debilis*, form *major*). Within the *Nitzschia palea* group the frustules are morphologically very variable (Krammer & Lange-Bertalot 1988) especially in length, width and stria density. Admiraal *et al.* (1999) noted that *N. palea* populations were morphologically different between upstream zinc polluted locations and downstream populations. Similarly, Tibby (unpublished thesis 2000) noted that *N. palea* exhibited a high degree of morphological variability. This inter-population morphological variation raises the question of whether these populations are different varieties or species.

5.2 Morphometric data

From the eight samples in which live *Nitzschia palea*-like cells were found, 48 cells were cultured. Of these cultures, SCUSCCD3, BKI2D6, TGI6G5 and SCUSCCD2 were excluded from morphometric clustering and discriminant analysis because under the light microscope, the frustules displayed features which were clearly not those of *N. palea*. Based on its frustule length (mean 113.5 µm), which exceeds Lange-Bertalot's 70 µm limit, and its regular fibula placement, the SCUSCCD3 culture was excluded. BKI2D6 was not included in analysis because of the distinct sigmoid shape of the frustule. TGI6G5 and SCUSCCD2 were very similar to *N. palea* however, the width of the frustule was slightly narrower (approximately 2.5 µm) and the transapical striae were clearly visible under light microscopy.

All other 44 cultures had the characteristic features of *Nitzschia palea* (i.e. Krammer & Lange-Bertalot 1988 , Sonneman *et al.* 2000) for instance, cells are linear-lanceolate to linear with capitate to subcapitate, rounded valve ends, 15-70 µm long, fibulae are short and irregularly spaced, striae are not visible with light microscopy. Chloroplast arrangement and number was not used as it was uniform in all cells. Two chloroplasts were seen in each cell, one at either end of the apex. Mann (1999) notes that chloroplast arrangement genus specific for almost all genera. Morphometric data of *N. palea* cultures are summarised in Table 5.1. The mean length of the cultures ranged from 20 µm (BKI2E5) to 60.8 µm (TGI6E5) while the mean frustule width ranged from 3.3 µm (RJI1D4) to 5.5 µm (SCUSCCG3). The breadth range given by Krammer and Lange-Bertalot (1988) is 2.5 to 5 µm while the fibulae count is 9-17 /10 µm. Culture RJI10G33 had the highest mean fibulae count (14.4/ 10 µm) while culture RJI10E44 had the lowest (10.0/10 µm). The mean length to width ratio for each culture has a large range from 5.38 (CHI5E5) to 14.08 (TGI1C1). There is low deviation of width from the cultures mean (0.1 to 0.5 µm). However, there is much variation of length within the cultures, for instance culture RJI5E7 has a standard deviation of 0.2 µm and RJI5F10 has 6.3 µm. The majority of the variation within the cultures is below 4 µm. This is not unexpected as

length varies over life cycle in most diatoms. For the fibula variable, the highest variation noted is 1.7 µm standard deviations in culture SCUSCCF3.

Table 5.1. Morphometric data on *Nitzschia palea* cultures (x ±stdn).

	Culture	n	L µm	W µm	F µm/10	L/W µm
1	**BKI2C6**	10	29.1±2.7	3.9±0.5	13.7±0.8	7.46
2	**BKI2D6**	10	34.39±0.43	4.2±0.1	10.2±0.6	8.26
3	**BKI2E5**	10	19.9±0.3	3.71±0.18	12.3±0.8	5.38
4	**CHI8B7**	8	31.9±4.3	4±0.3	12.7±0.9	7.94
5	**CHI8B8**	6	51.7±0.5	4.2±0.1	11±0.9	12.32
6	**CHI8B9**	6	37.7±4.6	4.1±0.5	14±1.6	7.74
7	**CHI8C7**	10	52.9±0.5	3.9±0.2	11.7±0.9	13.41
8	**CHI8C22**	4	50.5±0.5	4.5±0.5	11.5±0.6	11.35
9	**CHI8D7**	4	36.6±0.4	3.8±0.9	13.5±0.6	9.63
10	**CHI8D9**	10	23.5±1.7	3.6±0.4	11.1±1.4	6.6
11	**CHI8E9**	10	26.9±2.9	3.9±0.5	11.7±1.1	6.96
12	**CHI8G9**	10	44.5±4.6	4±0.4	11.7±1.2	11.14
13	**RJI1B4**	10	29.1±3.6	3.6±0.5	13.6±1.3	8.19
14	**RJI1B5**	10	27.9±1.9	3.6±0.4	12.3±1	7.82
15	**RJI1B6**	10	30.6±1.4	3.8±0.2	12.2±1	7.95
16	**RJI1D4**	10	23.8±2.6	3.3±0.4	13.7±1.2	7.28
17	**RJI1E3**	10	34.1±2.1	4.36±0.4	11.5±1.6	7.82
18	**RJI1F2**	4	31.8±0.3	4±0	13±0	7.89
19	**RJI5B11**	10	25.9±2.5	3.7±0.4	11.9±1	6.99
20	**RJI5C8**	4	37.1±0.6	3.7±0.3	12.7±0.6	9.77
21	**RJI5D8**	6	31.3±0.2	4.2±0.4	13±0	7.48
22	**RJI5E7**	4	37.4±0.2	3.7±0	12±0	10.18
23	**RJI5F10**	8	34.9±6.3	3.5±0.3	12.5±0.5	9.93
24	**RJI10C44**	10	28.1±3.2	3.6±0.4	13.7±1.1	7.89
25	**RJI10D33**	10	49.3±3.7	4.68±0.4	10.9±1	10.53
26	**RJI10D44**	10	27.7±2.6	3.5±0.3	14.1±1.1	7.89
27	**RJI10E22**	10	49.8±3.8	4.6±0.4	10.2±0.8	10.91
28	**RJI10E44**	10	50.8±3.5	4.6±0.3	10±0.8	11.02
29	**RJI10F33**	10	24.8±1.6	3.6±0.3	11.6±1	6.85
30	**RJI10F44**	10	40.5±2.8	4±0.3	10.7±0.8	10.04
31	**RJI10G33**	10	24.9±0.3	3.4±0.2	14.4±0.7	7.41
32	**SCUSCCB2**	10	55.2±3.7	4.8±0.4	10.6±0.5	11.48
33	**SCUSCCB3**	8	29.7±3.8	4.7±0.3	10.8±0.5	6.35
34	**SCUSCCD2**	10	35.8±0.4	2.6±0.1	14.1±0.6	13.66
35	**SCUSCCD3**	10	113.5±1	5±0.2	12.7±0.7	22.65
36	**SCUSCCF3**	10	20.1±2.0	3.2±0.5	13.5±1.7	6.21
37	**SCUSCCF4**	10	30.3±3.4	4.3±0.5	12.3±1.1	6.98
38	**SCUSCCG3**	6	50.4±1.2	5.5±0.3	11.3±0.5	9.11
39	**TGI6B4**	8	36.2±0.9	4.6±0.4	14±1.3	7.95
40	**TGI6D6**	10	23.5±1.7	3.6±0.4	11.1±1.4	6.6
41	**TGI6E5**	10	60.8±3.5	5.4±0.4	12±1	11.33
42	**TGI6E7**	6	31.8±0.6	3.4±0.2	11.3± 0.6	9.46

	Culture	n	L µm	W µm	F µm/10	L/W µm
43	**TGI6G4**	10	33.6±1.8	3.3±0.3	12.1±1	10.34
44	**TGI6G5**	10	24.0±0.2	2.5±0.1	14.3±0.5	9.58
45	**TGI1B2**	10	26.6±0.2	3.2±0.1	14±1.3	8.37
46	**TGI1C2**	10	46.4±3.2	3.3±0.3	14.1±1	14.08
47	**TGI1C3**	10	38.7±2.1	3.8±0.4	12.9±0.7	10.06
48	**TGI1F2**	10	36.1±3.1	3.4±0.3	12±1	10.55

L= valve length; W= valve width; F= number of fibulae in 10 µm; L/W= ratio of valve/length to width.

5.2.1 Cluster Analysis

Hierarchical cluster analysis (Figure 5.1), used to group cultured individuals based on their different morphological characteristics (length, width, fibula density and length to width ratio), produced five separate morphological groups.

Group 2 contains the highest number of cultures (n= 11) while group 4 contains only four cultures (Table 5.4). Each group contains cultures from every geographical area, except group 4, which is made up almost entirely of cultures from the Rum Jungle region. Group 1 contains cultures from sites representing the entire pH (3.4 to 8.4) and copper spectrum (5.9 to 0.002 mg/L) for sites RJI1 and SCUSCC respectively. Cultures from group 2 come from sites with pH values greater than 7.0 with the highest pH value being 8.4 (SCUSCC), while cultures from group 3 come from sites which have pH values higher than 6.2 (RJI5).

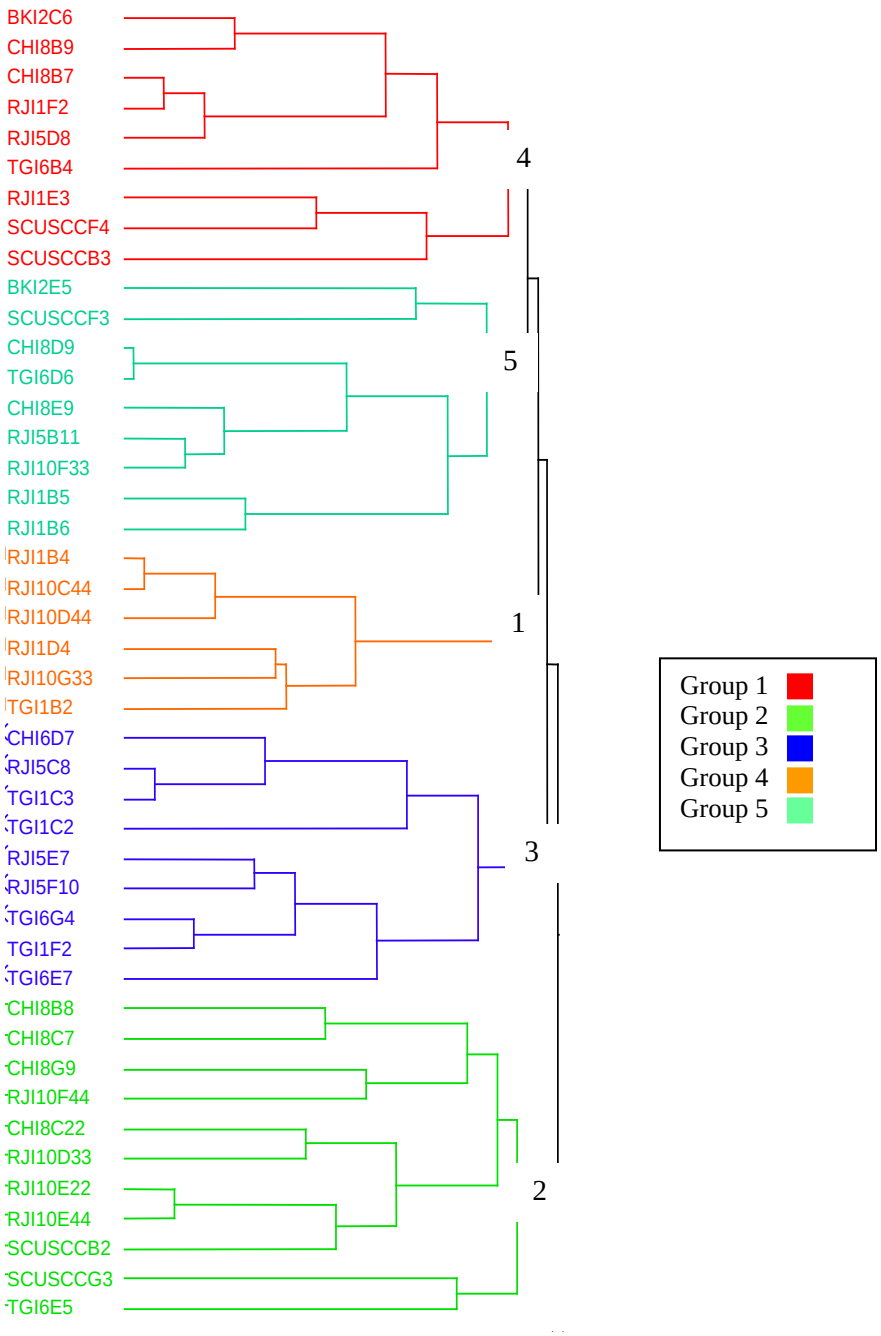


Figure 5.1. Dendrogram of 44 *Nitzschia palea*-like cultures.

5.2.2 Morphological correlations

Figure 5.2 shows a canonical plot of the first two principal components derived from the four-dimensional space created by the variables length (L), width of the central valve part (W), fibula density (F) and the length to width ratio (L/W). The two principal components (the first two axes) account, respectively, for 63.06 % and 23.05% (with a total of 86.11%) of the variation in the data (Table 5.2). The eigenvalue for axis 1 (2.52) is more than two times greater than axis 2 (0.92),

showing a high degree of reliability and suitability for principle component analysis (PCA). PC1 was formed by L, W, F and L/W which jointly explain 63.06% of the variation. PC1 is highly correlated with valve length; this variable has a considerable loading (0.61) on axis 1 and a much lower loading on axis 2. PC2 explains 23.05% of the total variation in the data. Both F and L/W have higher loadings on axis 2 (0.56 and 0.64) and are therefore more correlated to axis 2 than axis 1. The coefficient attributed to L (0.61) on the first axis is high and the coefficient of F has a negative value (-0.41), indicating that, with increased valve length, the number of fibulae per 10 µm decreases. The remaining axes PC3 and PC4 possessed low eigenvalues and variation percentages (13.89%) so they were not included in the results description.

Table 5.2. Summary of the principle component analysis of *Nitzschia palea* cultures.

Eigenvalues	2.52	0.92	0.55	0.01
% Variation	63.06	23.05	13.76	0.13
Cumulative % variation	63.06	86.11	99.87	100
Character/principal component	**PC1**	**PC2**	**PC3**	**PC4**
Length (µm)	0.61	0.24	0.15	-0.74
Width (µm)	0.47	-0.47	0.64	0.37
FIbulae/10µm	-0.41	0.56	0.72	-0.01
Length/Width ratio	0.49	0.64	-0.21	0.56

The correlations between each of the morphological variables are further illustrated in Table 5.3. Length and width are highly correlated (0.71) as are length and length/width ratio (0.87). This indicates that an increase in frustule length would be paired with an increase in frustule width. The lowest correlation exists between frustule length/width ratio and width (9.28) and then with fibulae (-0.36).

Table 5.3. Pairwise correlations of morphological variables.

Variable	by Variable	Correlation	Probability
Width	length	0.7110	0.0000
Fibulae	length	-0.5019	0.0005
Fibulae	width	-0.4674	0.0014
Length/width	length	0.8728	0.0000
Length/width	width	0.2878	0.0582
Length/width	fibulae	-0.3642	0.0151

As evident in the canonical plot (Figure 5.2) the cultures are most dispersed along the axis 1. The spread across axis 1 is not generated by single outliers but a clear continuum gradient of samples. The variables are most associated with axis 2 indicating that there is a lack of explanation for the horizontal variation of the spread of samples. The angle between the arrows representing width and length/width ratio is narrow, signifying a close correlation between the two variables. The variable

arrow for fibulae is the shortest with an F ratio of 15.6 indicating a small influence on the grouping of cultures. Length has the longest arrow and the highest F value of 85.9 followed by length/width (57.1) and width (20.8).

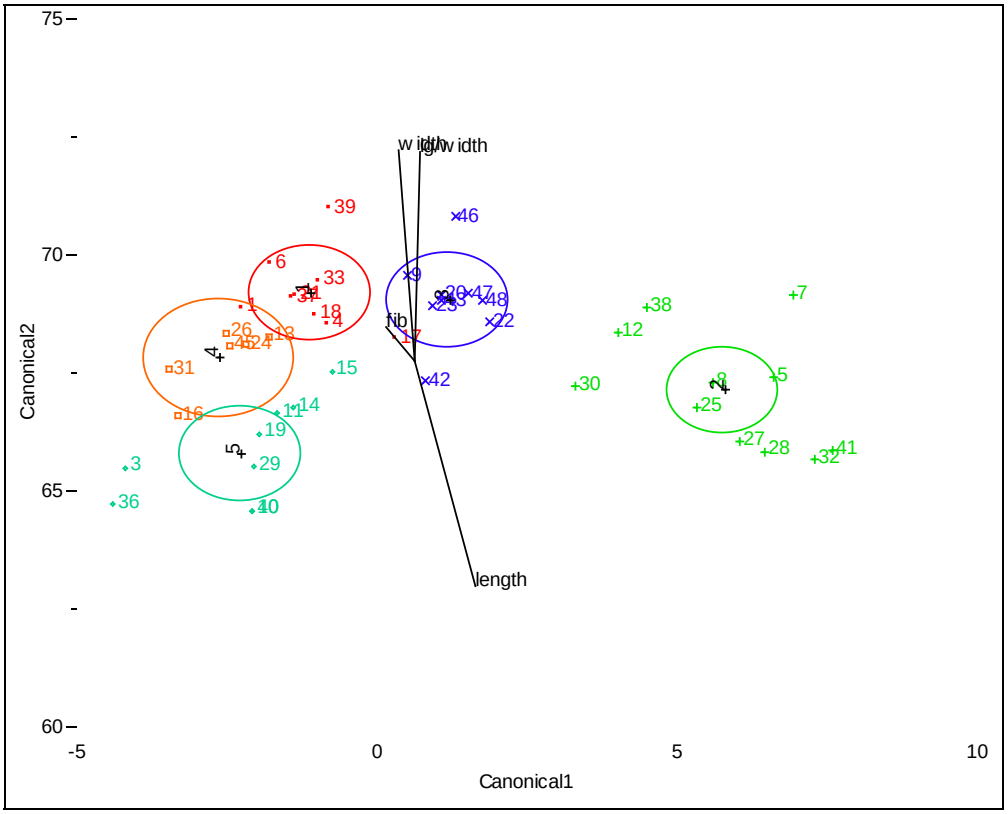


Figure 5.2. Canonical plot of clusters and morphological variables.

The scatter of the cultures suggests five different morphological groups. Groups 4 and 5 are characterised by low length measurements (< 27µm) whereas group 2 is characterised by higher length measurements (mean 50.6 µm) and broadest frustules (Table 5.4). Group 1, 4 and 5 are less significantly different from one another as they have intersecting circles, whereas the cultures in group 5 and 2 are most different. These groups have the lowest lengths and a length/width ratio between 6.8 and 7.8 µm. The group with the highest number of fibulae per 10 µm is group 4.

Table 5.4. Morphological means of clusters.

Cluster grouping	Count	Length	Width	Fibulae	Length/ Width
1	9	31.79	4.24	12.78	7.51
2	11	50.57	4.56	11.05	11.15
3	9	35.89	3.56	12.49	10.11
4	6	26.69	3.41	13.92	7.84
5	9	24.78	3.63	11.97	6.82
All	44	35.19	3.94	12.28	8.85

According to discriminant analysis, each morphological variable, length, width, fibula density and length/width ratio, are needed to separate the cultures into groups. All morphological variables are highly significant F <0.05 (0.00 for each) and, according to the Wilks' *Lambda* test, the probability of F is <0.0001 (indicating significance). The only culture misclassified by the hierarchical grouping analysis was culture 15 (RJIIB6). This culture was predicted to be a member of group 5 but, through discriminant analysis, was placed in group 1.

5.2.3 Discussion

From the *Nitzschia palea*-like cells cultured from mine polluted regions in the Northern Territory, 44 out of the 48 cultures have morphological characteristics which indicate that they are *Nitzschia palea* taxa. However, similar to findings from Krammer *et al.* (1988) and Tibby (unpublished thesis 2000), there is a large range in the length and breadth of the cells.

Despite their morphological similarity in LM and SEM images, cluster analysis of morphometric characteristics (length, breadth, fibulae and length/width) separates the cultures into five different groups, providing evidence for the existence of different morphological groups within the *N. palea* populations. Additionally, the ordination plot of the cultures produced five, well defined clusters of cultures. The major factors responsible for the scatter of cultures were valve length, length/width ratio, width and fibulae, in that order of importance. For instance, group 2 is made up of cultures which have long length and large width values whereas cultures in group 4 are shorter and narrower. However, there is a variable not yet identified which is important in explaining the variation in samples across axis 1. The only evidence of geography being a factor in differentiation the culture groups was that group 4 was made up almost entirely of cultures from sites at Rum Jungle. Similarly, indication of differentiation by pollution was not clear as most culture groups were sourced from environments with a pH from 3.4 to 8.4. Only groups 3 and 2 contained cultures only from sites with pH above 6.0. The application of multivariate techniques is able to define 5 separate morphological groups.

Application of multivariate techniques on previous studies, in some cases, has revealed polymorphism in diatoms (i.e. Theriot 1992; Mann 1999). Some authors observed that the number of morphotypes can range from two (e.g. *Biremis lucen* in Sabbe *et al.* 1995) to more than ten (e.g. *Diploneis smithii, D. fusca* complex in Droop 1994). Principal components analysis (PCA) has also been previously successfully applied by researchers, for instance Stoermer (*et al.* 1984), Theriot (1992), Droop (1994) Potapova *et al.* (2007) and Witkowski *et al.* (2004) to resolve differences within complexes, in the cases of *Gomphoneis herculeana, Stephanodiscus niagarae* , *Diploneis smithii, D. fusca, Achnanthidium minutissimum* and *Nitzschia tryblionella*. Witkowski *et al.* (2004) found that two clusters of specimens differing in geographical origins were identifiable and that the

major factor responsible for the scatter of specimens was valve width (PC1 0.53) and length (PC2 0.8), with specimens from the Antarctic having wider valves as well as lower stria and fibula densities, than named *Nitzschia buschbeckii*.

In order to strengthen results more morphological features could be used to define groups. This would require a more detailed SEM analysis of frustules to obtain stria densities and other ultrastructural features. Additionally, a method of simple shape descriptors, similar to studies by Droop *et al.* (2000), or more quantitative shape descriptors such as Legendre polynomials and contour segment analysis (Rhode *et al.* 2001; Pappas & Stoermer 2003; Mann *et al.* 2004), could be developed to measure frustule shape variation, especially to establish variation in shapes of *N. palea* frustule ends. These studies show that extraction of outline shape features combined with multivariate techniques such as PCA, is an immensely useful tool for quantifying subtle morphological variation in diatoms. The shape of the end of a frustule, (i.e. drawn out, bluntly rounded, or sharply rounded) can vary between *N. palea* varieties.

The uncertainty of whether these clusters represent phenotypic plasticity or genetic groups, or even natural variation within a species, reflects the recent debate on species concepts in diatoms (Mann 1999, 2001) highlighting the need for culturing and molecular genetic data. Taxa should be based on a combination of morphological, molecular and breeding data. It may be desirable to separate cluster groups 2 and 3 as a morphotype, but this could also be an expression of ecophenotypic variation. Environmental factors affect cell size and morphology in cultures and natural populations. For instance, water chemistry variables such as pH have been found by various researchers (i.e. Potapova & Snoeijs 1997; Trobajo *et al.* 2004; Cerino *et al.* 2005) to affect cell size and the morphology of the taxon. However, as with the *Gomphonema herculeana* complex, varieties may just be smaller ends of the species morphological range.

5.3 Phylogenetic tree

From the 48 cultures, 20 were chosen to be sequenced (Table 3.2). The sequenced cultures represented each geographical site and the large pH and heavy metal range. The cultures CHI5D6, TGI10G5 and SCUSCCD2 were sequenced even though they were not included in the morphometric analysis because they lacked *Nitzschia palea*-like characteristics. However, inclusion of these cultures in sequencing analysis was deemed important to assist with taxonomic classification and phylogenetic relationship between these and the remaining cultures.

The complete alignment of the sequenced cultures, including sequences from GenBank and *Navicula phyllepta* as the out-group, comprises 1808 nucelotide characters, but the first 191 sites and the last 80 were omitted from the analysis because this information was not available for all the

strains. Thus, of the 1536 remaining nucleotides, 244 positions were variable and contained 161 informative sites for parsimony analysis, which were used for the construction of a phylogenetic tree. The best phylogenetic tree obtained under Maximum-likelihood (ML) criteria is shown in Figure 5.3, rooted with *Navicula phyllepta*. *Navicula phyllepta* is an uncontroversial outgroup - close enough to allow inference with other species from sequence data, but far enough to be a clear outgroup

The relationships between the *Nitzschia* culture sequences are as follows. The sequenced SCUSCCD2 and TGI6G5 cultures group together on the phylogenetic tree. The clade has a bootstrap support value of 100%. A sequence taken from GenBank, *N.* cf. *supralitorea* (AJ867020), is basal to this clade (53% bootstrap support value). The genetic distance between the two cultures in clade 1 (SCUSCCD2 and TGI6G5) is 0.0007. The genetic difference between the clade 1 cultures and *Nitzschia* cf. *supralitorea* is 0.044. The genetic differences between thee cultures are given in Appendix 8.

Culture BKI2D6 is grouped with the *N. filiformis* sequence taken from GenBank (AJ866999). This clade (2) has a 100% bootstrap support value. The difference between the two sequences within the clade is 0.0046.

Aside from SCUSCCD2, TGI6G5 and BKI2D6, all the other *Nitzschia palea*-like cultures were grouped together with the *Nitzschia palea* sequences (taken from GenBank), within clade 3. The mean genetic distance (± standard deviation) for all pairwise comparisons within the culture sequences was small; 0.0015 (± 0.0014) with the maximum distance being 0.006. There was little genetic distance between the cultures and the *N. palea* sequences taken from GenBank 0.0014 (± 0.001). The next closest sequences to *N. palea* are *Nitzschia thermalis* (approx. 0.008), *Nitzschia pusilla* (approx. 0.012) and then *Nitzschia communis* (approx. 0.015).

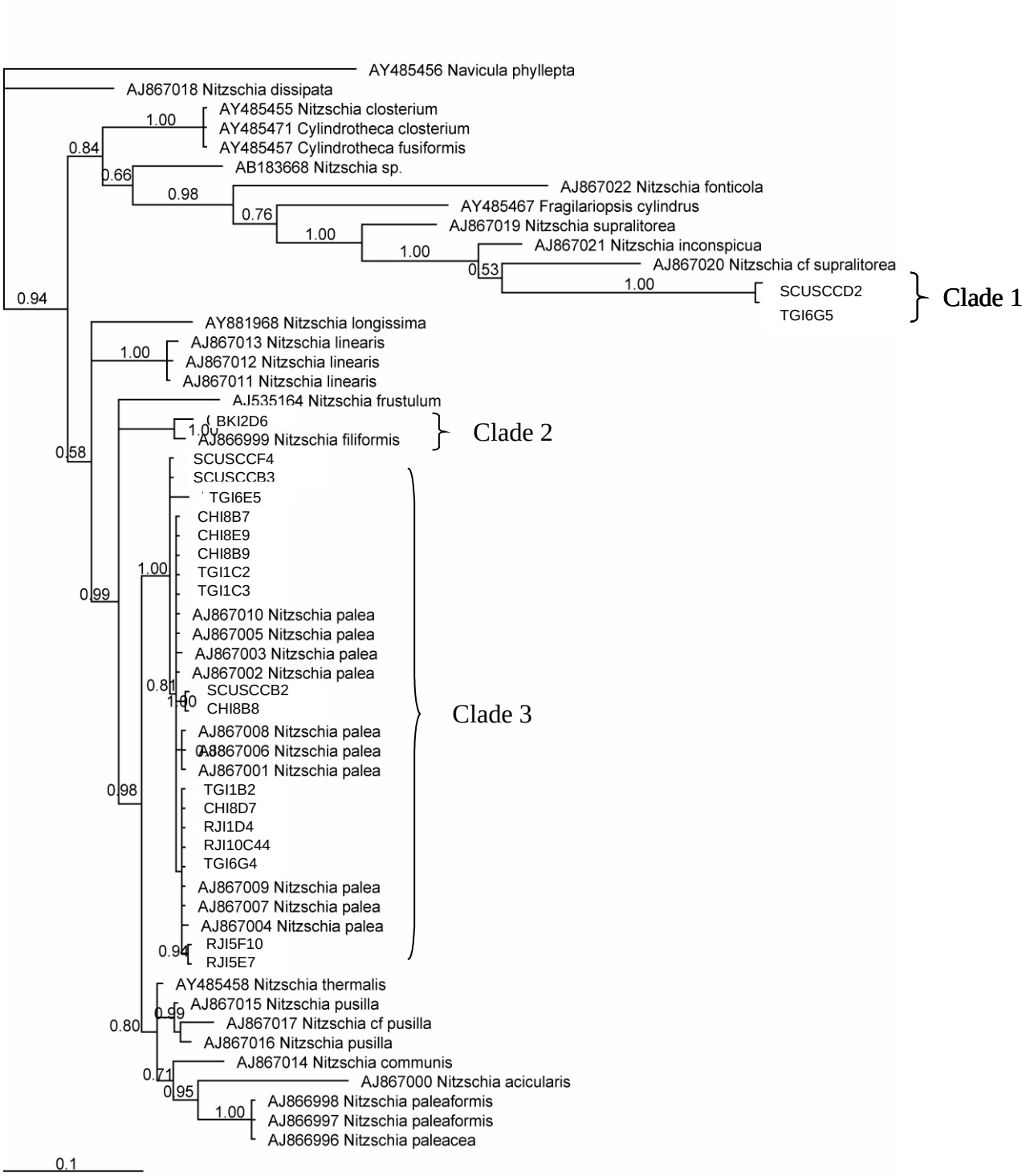


Figure 5.3. Maximum likelihood tree inferred for *Nitzschia* cultures based on SSU sequences.

5.3.1 Discussion

The sequences show little genetic divergence between the cultures. The small genetic divergence between the sequences suggests that there is no genetic difference between the *N. palea* cultures or the *N. palea* sequences taken from GenBank. However, cultures BKI2D6 and SCUSCCD2 and TGI6G5 have genetic sequences which are definitely not *N. palea*. BKI2D6 is shown to be *Nitzschia filiformis* and SCUSCCD2 and TGI6G5 are most genetically similar to *N.* cf.

supralitorea, but are a distinct species based on the genetic distances between these two species. These three cultures, which as defined by genetic sequencing are clearly not *N. palea* species, were also classed as separate species due to their morphological characteristics. However, while the *N. palea*-like cultures were separated into five separate groups based on their morphological characteristics, this separation was not reflected in the genetic study where each culture sequence was placed in one clade along with *N. palea* sequences taken from GenBank. Consequently, this study confirms that *Nitzschia palea* has a wide distribution and provides evidence of its tolerance to extreme pH and heavy metal environmental conditions. It also provides evidence that the noted morphological variability of the species is related to the upper and lower limits of the species size range.

Morphometric and molecular genetic analysis methods have provided evidence both supporting and disproving the theory that the presence of morphological variation indicates genetically distinct species. For instance, studies by researchers such as Lundholm *et al.* (2006) have found that grouping of morphological variable diatoms has been supported by genetic analysis. However, morphotypes within both *Cymbella cistula* (Pappas & Stoermer 2003) and *Gomphoneis herculeana* (Kociolek & Stoermer 1988) were not assigned any taxonomic status based on genetic analysis. In the case of *Gomphoneis herculeana,* morphotypes represent nothing more than larger or smaller representatives of a particular size diminution series. Perhaps before assigning variable populations within species, to new species, a process similar to Mann *et al.* (2004) should be followed. Mann only elevated the demes of *Sellaphora pupula* to species status after obtaining evidence of reproductive isolation inferred from morphometric analysis and mating tests.

However, one limitation of GenBank is that it does not define which sequences of *N. palea* are in fact *N. palea* varieties, for instance *N. palea* v. *debilis*. Consequently, it is indeterminable whether the small genetic variation within the clade represents varieties. However, varieties have been designed almost exclusively on morphological criteria (Round 1996), usually based on variations in overall form or striation. Although there has been an increased tendency to upgrade varieties to species this has been mainly based on morphometric evidence and not on genetic sequences of taxa.

While nuclear-encoded small subunits (SSU) can be used to assess genetic relationships, the internal transcriber spacer (ITS) may be a better tool for infraspecific analysis. Variations in ITS sequences have been used to separate isolates of diatoms from different geographic regions (Zechman *et al.* 1994; Godhe *et al.* 2006) and have revealed several lineages within *Pseudo-nitzschia delicatissima*-like species in the Mediterranean Sea (Orsini *et al.* 2002). In a recent study by Godhe *et al.* (2006) ITS 2 attributed the highest percentage to among-sample variability, and thus provided the best resolution for separating different geographic isolates of *Skeletonema marinoi*.

Taxonomic problems in diatom research are persistent within, and between, many species. This problem inhibits consistent and correct application of taxonomy which is essential to ecological and palaeoecological studies. Multivariate analysis of morphological characteristics may help in solving these problems but, ideally, a combination of morphological techniques and molecular genetic or mating test techniques seem to be the best way to solve such problems. Although *Nitzschia palea* is a geographically widespread taxon, which is adapted to varying environmental conditions and displays large phenotypic plasticity, this variation was not reflected genetically. Each of the cultures can be genetically classed as *Nitzschia palea*.

5.4 Species descriptions

The three species from this study, defined by molecular genetic analysis, are described from light microscope and electron microscope micrographs of live and dead cells. Key references for each species are given to allow for cross comparison between this study and others.

5.4.1 Nitzschia palea

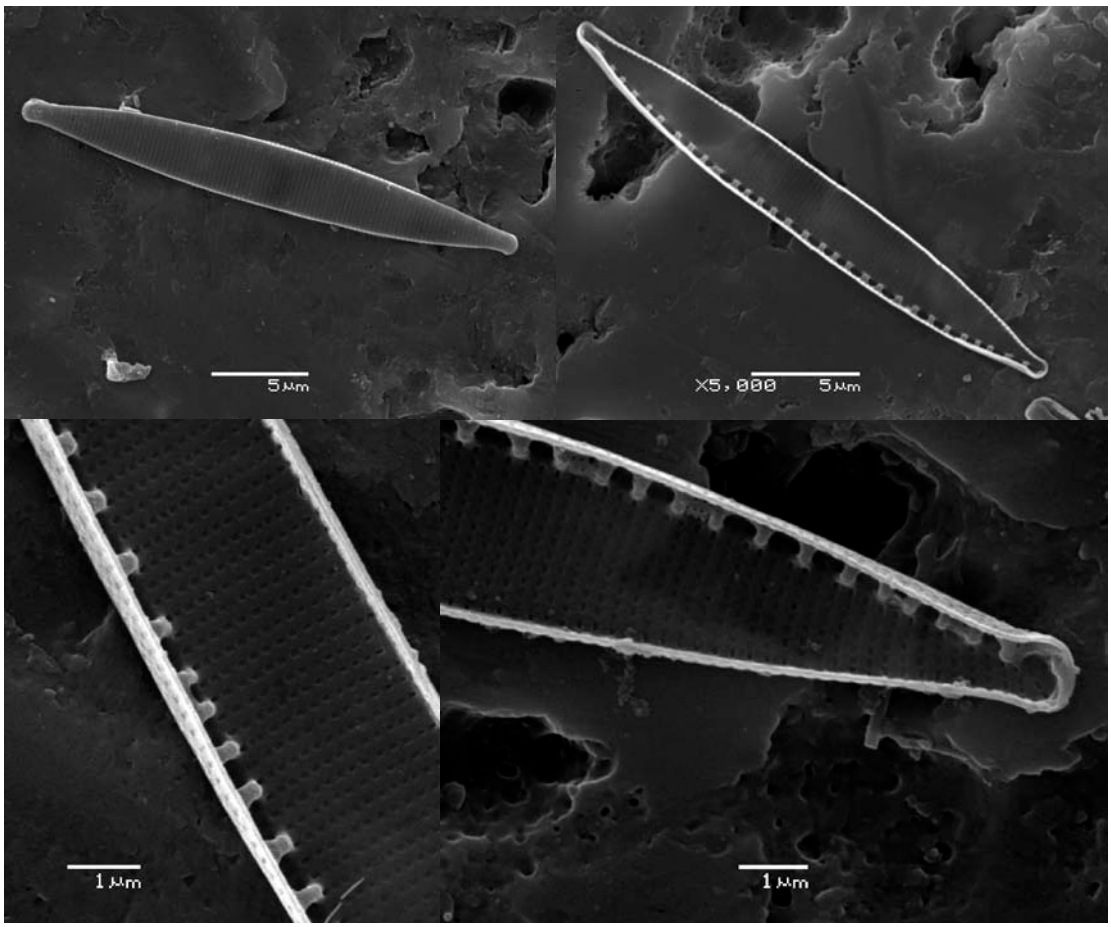
Nitzschia palea (Kützing) W. Smith 1856

Description:

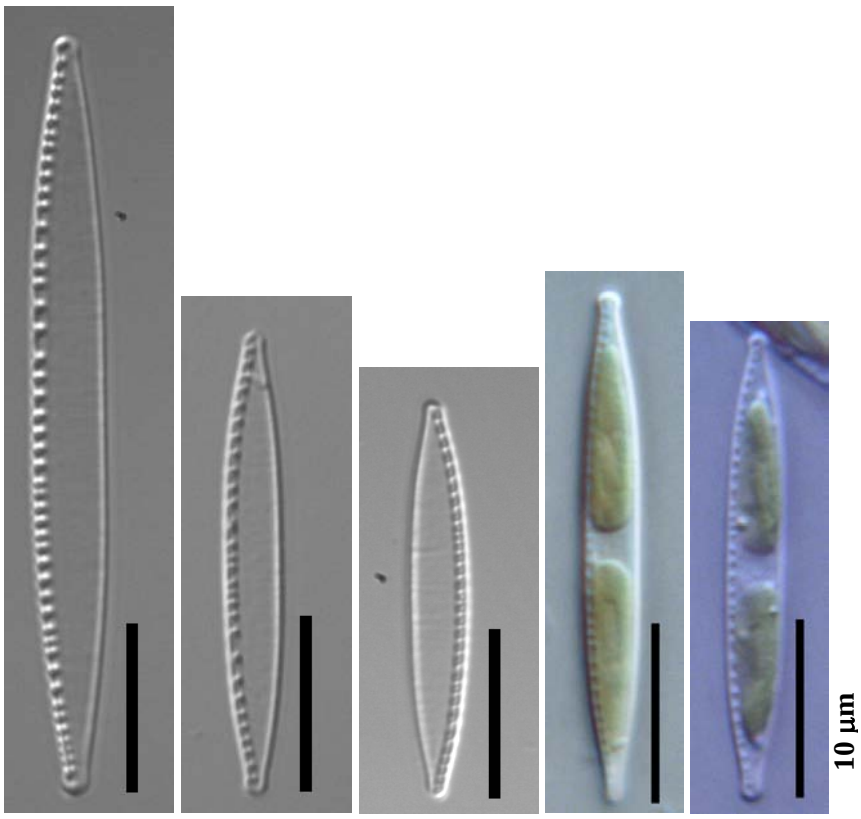
The valve is highly variable in shape, ranging from narrow/broadly linear-lanceolate to linear with narrowly elongated to pointedly rounded, subcapitate ends (Figure 5.5). The marginal raphe canal is supported by small, narrow and irregularly spaced fibulae. The transapical striae are moderately dense, 38-42/10 µm, and indistinguishable under the light microscope (Figure 5.8-5.12). Cell length ranges from 20.1 - 67.2 µm, while width ranges from 2.6 - 6 µm. The density of the fibulae is approximately 10 - 16/10 µm. In live cells, two plate-like chloroplasts sit towards each apex (Figure 5.11, 5.12). The areolae in the striae are slightly elongated and the number along the central striae is variable and unevenly spaced ranging from 12 to 15 in 10 µm (Figure 5.6).

Key references:

Foged 1978, p. 108, pl. 45, fig. 11, pl. 46, fig. 22.; Holland and Clark 1989, p. 22, pl. 18, figs. 7,8.; John 1983, p. 171, pl. 70, fig. 12; Krammer and Lange-Bertalot 1988, p. 85, figs. 59; 1-24, 60: 1-7; Thomas 1983, p. 60, pl.11, fig.15.



Figures 5.4-5.7. SEM micrographs of *Nitzschia palea* (TGI1B2). Figure 5.4 external valve view. Figures 5.5-7 internal valve view.



Figures 5.8-5.12. Light micrographs of *Nitzschia palea.* Figures 5.8-5.10 light micrographs of dead cells. Figures 5.11-12 chloroplast positioning in two live cultured cells.

5.4.2 Nitzschia nana

Nitzschia nana Grunow in Van Heurck 1881

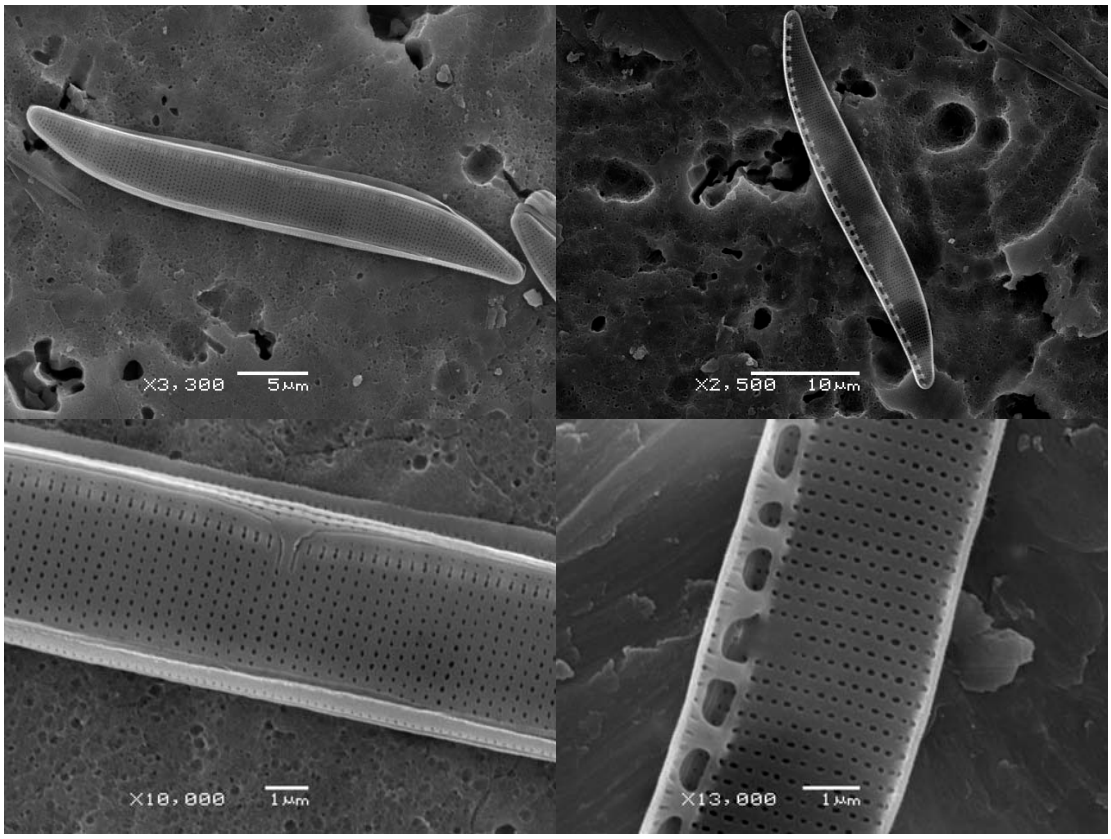
Description:

The frustules in valve view are small and linear with the ends distinctly curved in opposite directions giving the frustule a sigmoid shape (Figure 5.13). The ends of the frustule are bluntly rounded. The middle of the frustule is slightly constricted. The frustules taken from the Northern Territory have a length of 34 – 35 µm (mean 34.4 ± 0.4), breadth of 3.95 - 4.4 µm (mean 4.2 ± 0.1) and irregular, short, blunt fibulae, 9 – 11/10 µm (mean 10.2 ± 0.6). The middle two fibulae are distinctly distant. Additionally, with SEM imagery, the stria density was recorded to be 34/10 µm. On the outer frustule valve surface one can see the raphe ending in the middle and its internal deflection (Figure 5.15). On the internal frustule one can see that the striae are punctate and that, two areolae close to the raphe in the middle two stria rows are absent (Figure 5.16). The chloroplast positioning is typical of *Nitzschia*. The two chloroplasts are situated at each apex (Figure 5.19). This is a difficult species to classify especially with light microscopy. There has been a lot of

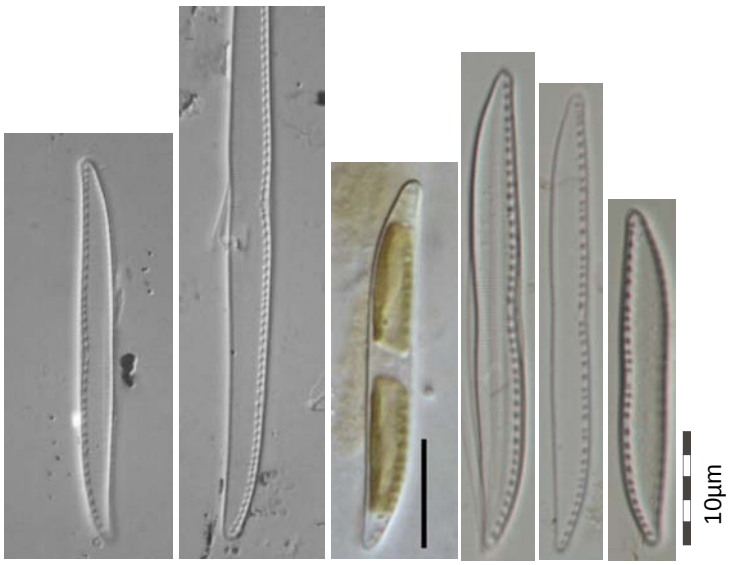
misclassification between *N. nana*, *N. filiformis* and *N. ignorata* as the differences between the species are poorly defined.

Key references:

Lange-Bertalot 2000, p. 206, f. 15-18.



Figures 5.13 – 5.16. SEM micrographs of *Nitzschia nana* (BKI2D6). Figures 5.13 and 5.15 external valvar view. Figures 5.14. and 5.16 internal valve view.



Figures 5.17 -5.22. Light micrographs of *Nitzschia nana.* Figures 5.17-18 type material (valves in Figure 5.17 measures 48.4 µm and in Figure 5.18 measures 83.2 µm in length). Figure 5.19 light micrograph of a live cultured cell (BKI2D6). Figures 5.20-5.22 light micrographs of dead cells.

Note: BKI2D6 is grouped genetically as having the same sequence as *Nitzschia filiformis*. However, by comparing the morphology of the BKI2D6 frustules with the *Nitzschia nana* type material (Synopsis type number 399, Haverfordwest, Wales, held in Zurich) (Figure 5.17-18), and with the micrographs from Lange-Bertalot (2000), I would classify this culture as *Nitzschia nana* rather than *N. filiformis*.

5.4.3 Nitzschia spec. 39

Nitzschia spec. 39

Description:

The frustules in valve view are linear-lanceolate with long, drawn out, sharply rounded ends (Figure 5.24, 5.25). The length of the species ranges from 24 – 36 µm (mean 29.9 ± 6.0 µm) while the width measures 2 – 3 µm (mean 2.6 ± 0.2). The density of the fibulae is 13 - 15/10 µm (mean is 14 ± 0.5µm). The transapical striae are easily visible under a light microscope (Figure 5.27). The fibulae are short, blunt and irregularly spaced. Two chloroplasts are evident at either end of the cell apex (Figure 5.28).

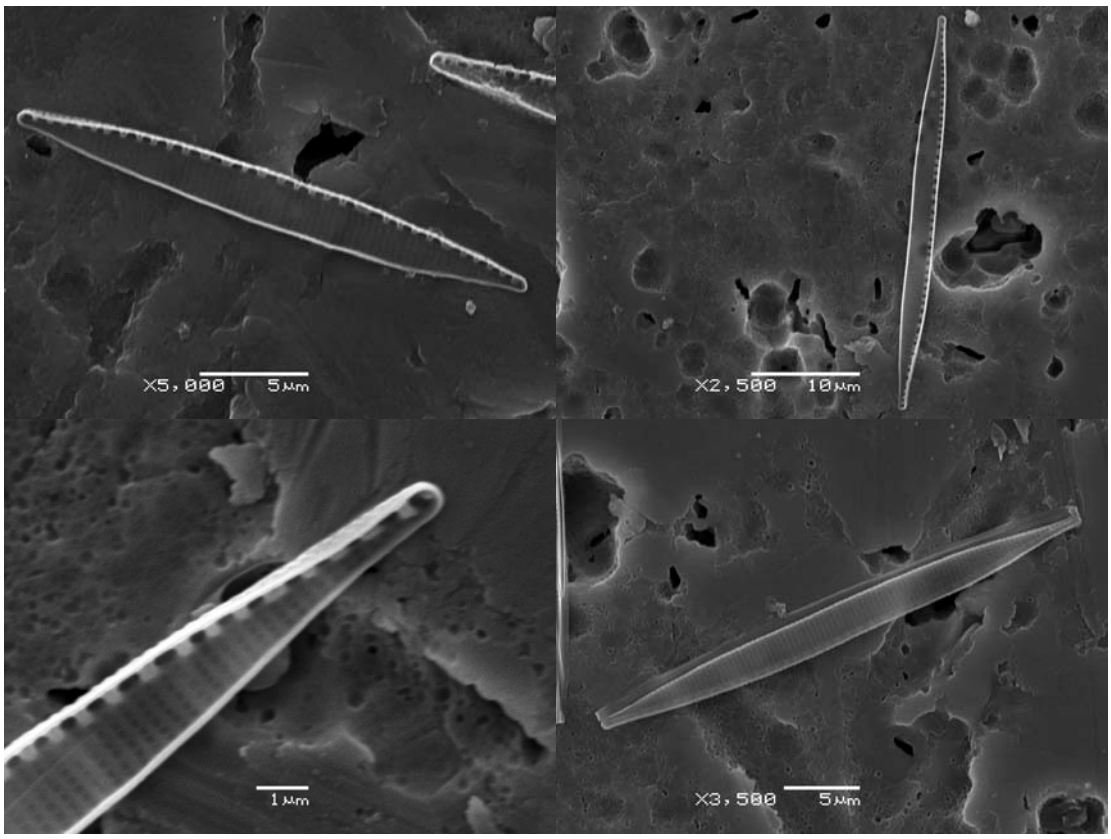
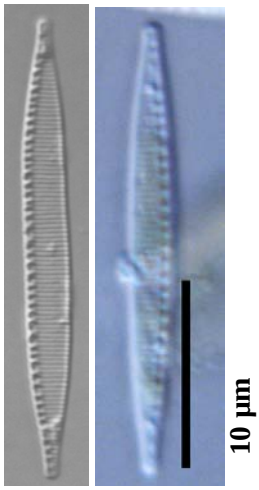


Figure 5.23-5.26. SEM micrographs of *Nitzschia* spec. 39 (TGI6G5 & SCUSCCD2). Figures 5.23-25 internal valve view. Figure 5.26 external valve view.



Figures 5.27-5.28. Light micrographs of *Nitzschia* spec. 39. Figure 5.27 light micrograph of a dead cell. Figure 5.28 chloroplast positioning in a live cultured cell.

Chapter 6 - Biological indicators

6.1 Introduction

The quality of water in numerous creeks and rivers in the Northern Territory has deteriorated as a result of mining. To manage effectively and restore aquatic resources managers must be able to identify which water bodies or portions of water are affected and to what degree the systems have degraded. Biological monitoring is used increasingly in the management of aquatic systems. However, it is pertinent that a method of monitoring which is effective and suitable for extreme environmental conditions, such as heavy metal pollution, is established. Diatom community assemblages have been successfully used as biological monitors to model and characterise current-day water quality conditions, and also to estimate the extent, magnitude and spatial pattern of past changes in water quality (Round 1991; Dixit *et al.* 1992; Dixit & Smol 1994; Smol & Cumming 2000). Most commonly, the multivariate technique of transfer functions have been utilised to relate changes in diatom communities to environmental variables. However, there are a range of biological monitoring methods suitable for diatoms. These include single species or a group of indicator species and biodiversity indices such as species richness, diversity and evenness measures. The identification of indicator species can be based on the autecology of diatom species or the use of presence and absence of a species. In order to address the issue of the management of tropical, metal-polluted, aquatic systems it is essential that an effective monitoring method is established.

6.2 Community structure transfer functions

Transfer functions are commonly used for diatom-based assessments of aquatic environments but were initially used primarily in lake studies, particularly for the reconstruction of historical and pre-historical environments (Cumming & Smol 1993; Dixit *et al.* 1993; Gell 1997; Smol *et al.* 1998; Pienitz *et al.* 1999; Battarbee *et al.* 2001). Recently however, transfer functions have been successfully used in streams to infer a variety of environmental variables, for instance, pH and conductivity (van Dam & Mertens 1995; Pan *et al.* 1996; Winter & Duthie 2000; Philibert *et al.* 2006). However studies by Philibert and Prairie (1999, 2002) and Pipp (2001) have suggested that diatom-based indices and transfer functions that were developed for one region are inaccurate in others. Consequently, it is important to develop site specific transfer functions for aquatic monitoring. Regional differences in water quality, force fitting of species to keys (Tyler 1998), endemism (Potapova & Charles 2002) and local history (Philibert & Prairie 1999, 2002) are some of the reasons why transfer functions may be inaccurate when applied to regions outside their scope. The aim of this section is to produce a regional transfer function for diatoms in tropical northern Australia, which inhabit waters impacted by acid mine drainage. An initial transfer function is

developed using the dataset of all species occurring more than once in the dataset. This is then compared to a transfer function produced utilising a reduced dataset of genus level identification and, then again, to a dataset utilising all species occurring more than once, together with the bioavailable fractions of heavy metals. The three datasets are compared to determine which level of diatom taxonomic identification and chemical components produce the strongest transfer function for inference models of selected environmental variables.

6.2.1 Species with occurrences greater than one

The first section in this chapter explores the relationship between diatom species assemblages and environmental variables in order to determine which environmental variables would best explain the diatom species distribution. Transfer functions are developed using the species dataset composed of the 193 species occurring in more than one sample (Appendix 6).

6.2.1.1 Similarities and dissimilarities in species assemblages

A detrended correspondence analysis (DCA) was used to examine the patterns in diatom variance from the dataset. DCA is an ordination technique derived from correspondence analysis (CA) but which detrends the data to counteract the arch affect which is common in a CA (ter Braak & Prentice 1988) and which was apparent in this dataset. Importantly, since a DCA creates a synthetic environmental gradient, to provide maximal site dispersal, the nature of the taxons unimodal response is improved.

A DCA was conducted on the 193 diatom species that occurred in more than one sample from the 50 site dataset. Table 6.1 indicates the variance of the diatom community explained by the four DCA axes while Figures 6.1 and 6.2 are plots of sample and taxon scores. Axis 1 explains only 9 % of species variance. This is not surprising given the large number of taxa included in the analysis. Large amounts of ecological information may be obtained from datasets with equal or less axis 1 variance than that observed in this study (e.g. Dufrene & Legendre 1997). Moreover, robust transfer functions have been derived from diatom calibration sets even with only relatively low percentages of diatom data variance explained by DCA analysis (e.g. Stevenson *et al.* 1991, DCA axis 1 = 9.7%). In the DCA of the species and sites, the percentage of cumulative variance captured by the first two axes reached 15.4% of the species data variance and 19.6% was captured on the third axis. The first four axes explained 23.1% of the cumulative variance in the diatom data. The first axis accounts for most of the explained variation in the species data. The first two axes of the DCA were significant according to Monte Carlo permutation tests (with 199 unrestricted permutations, $p < 0.05$).

Table 6.1. Summary statistics for a DCA of the entire dataset of 50 samples.

Axes	1	2	3	4	Total inertia
Eigenvalues:	0.739	0.527	0.348	0.289	8.222
Lengths of gradient:	5.032	3.594	3.678	3.018	
Cumulative percentage variance of species data:	9	15.4	19.6	23.1	

The DCA showed that there was a clear separation of low impacted sites and high impacted sites across axis 1 (Figure 6.1). Site codes are found in Appendix 1. The sites are spaced along axis 1 ranging from 0 to 5.0 deviation units. The species *Kobayasia* cf. nov. spec. and *Pinnularia joculata* were excluded from the DCA analysis as these are outlying species which greatly skewed the distribution of the DCA plot, disguising species and sample relationships. At Redbank site SC@BRX (30) *Kobayasia* cf. nov. spec. reaches 50% relative abundance. It is present at six other sites but at an abundance of less than 1%. Similarly, *Pinnularia joculata* has a high relative abundance of almost 95% at site RJI5. This species is present at four other sites but at a relative abundance less than 11%. However, the site SC@BRX is also dominated by *Nitzschia palea* v. *tenuirostris* (Ni.pvt) (48.5% relative abundance) and, therefore, the site plots high on axis 2 close to this taxon (Figure 6.2). Sites seem to cluster into two groups on the basis of their species composition. Sites belonging to each group have a distinct flora associated with them. The most highly impacted sites, according to their low pH and high metal concentrations (i.e. RJI1 = 1), are encompassed in a dashed circle. These 16 sites plot to the far right of axis 1. The sites encompassed by the dotted circle are sites which are either controls, or sites further downstream of the pollution source. This cluster of sites is more widely spread along axis 2 and halfway along axis 1.

The positioning of a number of sites high on axis 1 in Figure 6.1 appears to be strongly influenced by taxa such as *Nitzschia paleaeformis* (Ni.pa) and *Nitzschia vasta* (Ni.va). These species plot to the extreme right of axis 1 in Figure 6.2. These are key species of the impacted sites. The species associated with impacted sites are principally *Nitzschia vasta*, *Nitzschia nana* (Ni.na) and *Nitzschia paleaeformis*. Interestingly, the majority of the species noted at the impact sites belong to the genus *Nitzschia*. *Nitzschia paleaeformis* is most abundant at site RJI6 (6) at an abundance of 76%, followed by site BKI1 (41) at 74% abundance and site CHI2 (38) at 48%. It is not surprising that these sites are found to the extreme right of the plot. Impacted sites from Redbank mine site RJI1, RJI2, RJI3 and RJI4 (1, 2, 3, 4) plot lower on axis 1 and are associated with a large number of species with relative abundances lower than 10%. Site RJI3 (3) however, does not plot with any of the other impacted Rum Jungle sites as it is dominated by *Achnanthidium minutissimum* v. *minutissimum* (93% relative abundance) (Ac.mvm). This species is highly abundant in a number of

other sites, for instance RJI7 (7%) and RJI8 (8%) and other such low impacted sites, which all plot in the lower left quadrant of the DCA.

The presence of *Achnanthidium minutissimum* v. *minutissimum,* in sites of both low and high impact, may be partly related to its tolerance to many environmental conditions (Nakanishi *et al.* 2000). This species plots low on axis 1 and 2 of the DCA. Similarly, *Nitzschia palea* (Ni.pvp) is argued to be a pollution tolerant species (Dickman *et al.* 1990; Salomoni *et al.* 2006). However, it plots high on axis 2 close to sites of low impact (sites 28, 29 and 40). These two species are both present at sites with pH levels from pH 3.4 (RJI1) to pH 8.4 (SCUSCC) and copper concentrations ranging from 0.002 mg/L (CHI5) to 14.9 mg/L (ECUS12MC). Another species of interest is *Nitzschia vasta* which has been noted, in a previous study, as being tolerant to low pH (DeNicola 2004). The site CHI1 (37) is solely dominated (100%) by *Nitzschia vasta*. This site has a pH value of 2.95 and a copper concentration of 12.2 mg/L. Similarly, *Nitzschia paleaeformis* is abundant at the impacted sites for example, sites CHI2, BKI2 and TGI4. It appears that, from these observations, a number of taxa in the dataset constitute useful bioindicators of high levels of heavy metals while others are less useful as they have pollution tolerances which are very broad.

Although the Redbank site ECUS12MC (26) had a pH of 4.5 and copper concentration of 14.9 mg/L it does not plot with the other impacted sites. The site consists of many species of low abundance (> 10%), for instance, *Cymbella cymbiformis* (8.8%) (Cy.cy). The species of highest abundance at this site, *Nitzschia palea* v. *palea* and *Achnanthidium minutissimum* v. *minutissimum,* have wide copper tolerances. So, this site plots low on axis 1 and 2, away from the polluted sites, and closer to sites of lower pollution.

The sites which plot high on axis 2, of Figure 6.1, are sites of low impact with pH values above 7.0. These sites, with the exception of CHI4 (40), come from Redbank mine (28, 29, 30). Species influencing the positioning of these sites include *Nitzschia palea* v. *tenuirostris.* The remaining species in these sites have a relative abundance of less than 10%. The control sites of the Redbank mine plot low on axis 2. Species associated with Redbank sites 12MCUSEC (33), ECUSBRX (34), CCUSSC (35) and SCUSCC (36), which have relative abundances greater than 10%, comprise *Epithemia* cf. *cistula* (Ep.ci), *Nitzschia liebetruthii* (Ni.lie) and *Planothidium rostratum* (Pl.ro). These sites have similar physical and chemical characteristics. Two Redbank control sites, 12MCUSEC (31) and ECUSBRX (32), have a different selection of species leading them to plot away from the other Redbank control sites. The species most influencing these sites are the *Encyonema* species, *E. silesiacum* (En.si) and *E. minutum* (En.mi). The difference in species composition is reflected by the difference in the chemical composition of these waters compared to the other low impacted Redbank sites. These two sites have a low alkalinity level (< 15mg/L) which is more comparable with the impacted site 12MCDSEC than with the other less polluted

sites which have alkalinity levels above 200 mg/L. These sites also have low calcium but slightly higher iron levels than the other low impacted sites.

The low impact and control sites from the mines Cosmo Howley, Rum Jungle and Tom's Gully (i.e. sites 22, 23) plot in the central area of the DCA (Figure 6.1). Within this area cluster the majority of sites from the dataset cluster along with the majority of species. The taxa chiefly influencing this cluster of sites include *Achnanthidium minutissimum* v. *minutissimum,* which has a relative abundance of greater than 50% at more than eight low impacted sites (i.e. CHC2 and RJC2), and *Brachysira neoexilis* (Br.ne), which reaches over 10% at five sites, and over 38% at the sites CHI7 (45) and CHI8 (46). However, the majority of the other species that cluster with these two taxa occur in relatively low abundances. Thus the positioning of these sites in Figure 6.1 are influenced by no one species rather they are influenced by many. For instance, site RJC1 has 50 species which have a relative abundance of less than 10%. Only two species at this site have a relative abundance greater than 10%, namely *Achnanthidium minutissimum* v. *minutissimum* and *Navicula notha* (Na.no), but their relative abundances are less than 13%.

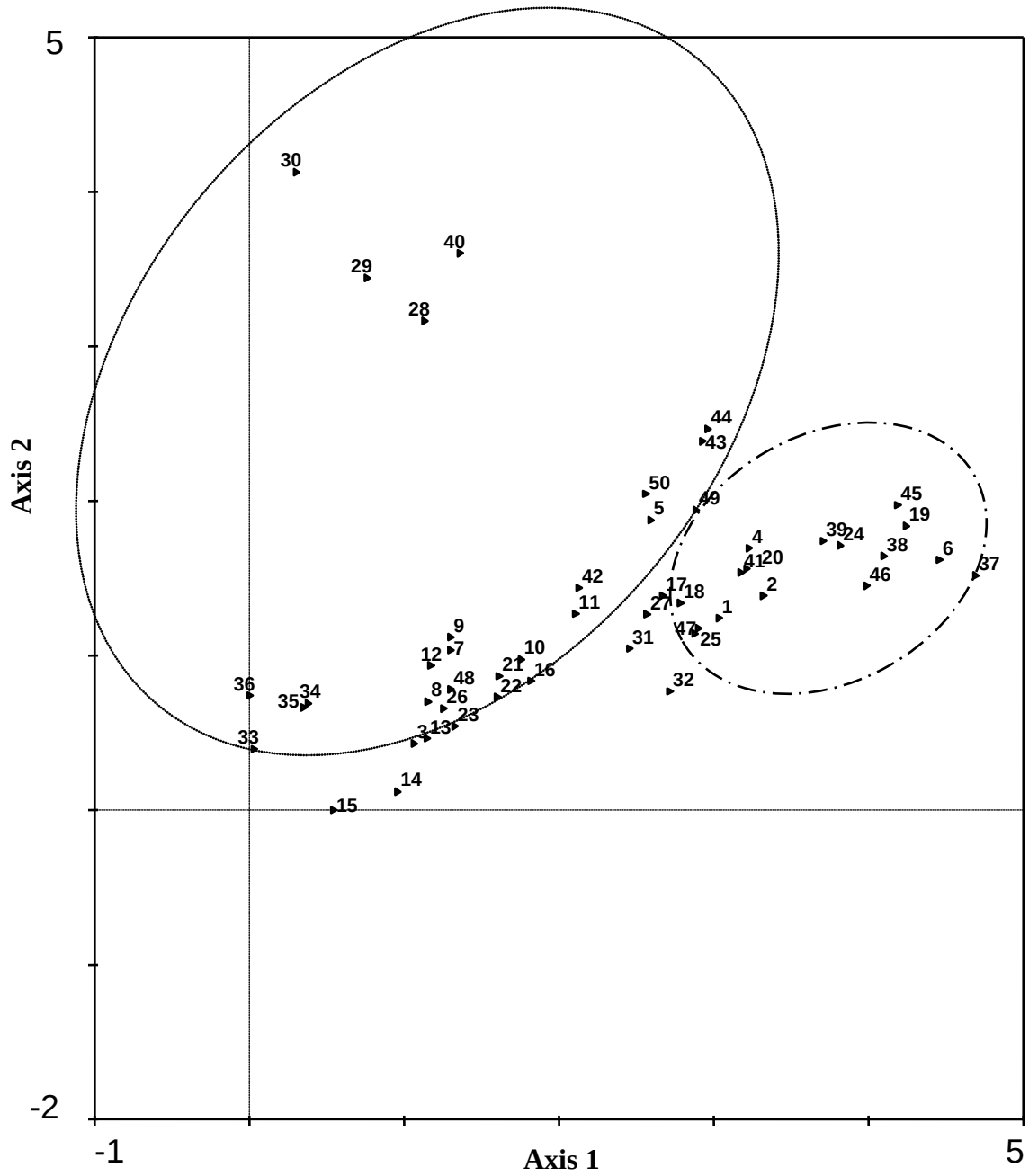


Figure 6.1. A plot of detrended correspondence analysis (DCA) of the 50 site dataset. Site numbers are listed in Appendix 1. The dotted circle encompasses low impact sites while the dashed circle encompasses highly impacted sites.

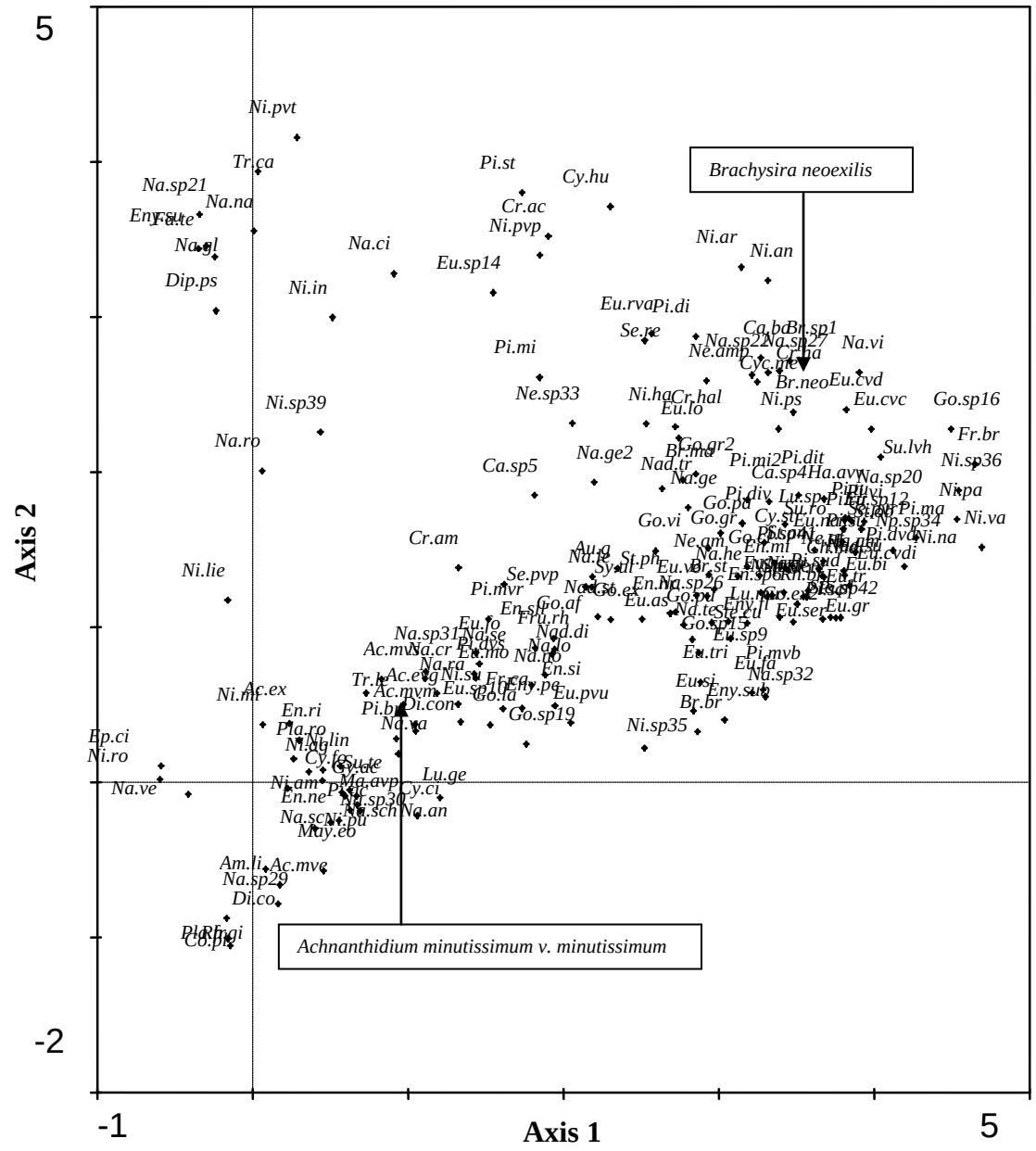


Figure 6.2. A biplot of detrended correspondence analysis of the species.

6.2.1.2 Analysis of diatom-environment relations

Canonical correspondence analysis (CCA) was undertaken to explore the relationship between diatom taxa and measured environmental variables from the dataset. A CCA is a multivariate direct gradient analysis technique which is based on constrained ordination. The method assumes that the abundance of a taxon is a symmetrical, unimodal function of position along environmental gradients (ter Braak 1987). The advantage of a CCA over unconstrained ordination techniques is that the position of a sample or taxon is not only determined by the taxa present in a sample but is also a function of a defined set of environmental variables. This then provides a visual and

mathematical expression of the degree of association between taxa and environmental variables (Charles & Smol 1994). Thus, one of the objectives of using a CCA is to identify which environmental variables are most suitable for weighted average regression and calibration. For the dataset a CCA was conducted with stepwise forward selection of environmental variables using the program CANOCO. All variables, except pH, were log10(x+1) transformed in order to reduce skewedness along the distributions. The ranges of the quantitative variables, retained after construction of a correlation matrix, are summarised in Table 4.1.

The relative influence and significance of each environmental variable in explaining diatom distribution was determined through the process of stepwise forward selection. The variables were tested for significance by Monte Carlo permutation tests, in which 99 unrestricted permutations were used. Variables were considered to be significant and included in the CCA at the $p \leq 0.05$ threshold point. Stepwise selection of variables aids in the elimination of variables which may co-vary with statistically significant variables but not exert independent explanatory power. The results of the present stepwise forward selection are presented shown in Table 6.2. Here, the difference between the percentage variance explained, before and after addition to the analysis, is due to the effects of co-variation between variables. For example, before aluminium (Al) was selected, the percentage variance attributed to aluminium also included a portion that was attributed to several other variables such as electrical conductivity (EC) and magnesium (Mg). Therefore, the variance explained by these variables decreased when aluminium was added to the dataset as some variance was already accounted for. Each of the variables indicated in the table below were selected based on their significance calculated by the Monte Carlo permutation test. The variables from the dataset not included in the table each had a p-value greater than 0.05 and were therefore not considered significant.

Table 6.2. Percentage of variance explained by the variables retained after stepwise forward selection.

Variable	% variance explained before addition	% variance explained after addition	p -value
Al	0.48	0.48	**0.046**
EC	0.43	0.90	**0.002**
SO_4	0.40	1.30	**0.002**
Mg	0.32	1.62	**0.008**
N	0.35	1.96	**0.002**
Na	0.31	2.28	**0.008**
pH	0.30	2.57	**0.014**

* Significant variables ($p \leq 0.05$) are emboldened.

Once the seven forward-selected variables (pH, Na, EC, Mg, SO_4, N and Al) were selected, a final CCA was executed on the variables, 50 samples and the 193 diatom taxa. A summary of the results are given in Table 6.3. The environmental variables used in the final CCA explained 18.7% (sum of all canonical eigenvalues = 2.521) of the total variation in the diatom data on the first four axes. This compares to 17% for the Surface Water Acidification Program (SWAP) dataset (Stevenson *et al.* 1991). As can been seen from examination of Table 6.2 the combined variance explained by axis 1 and 2 of the CCA is 16.1%. These two axes express over half of the constrained and approximately over one quarter of all taxon variance in the dataset (total = 49.9%). The relatively low portion of cumulative variance explained by the first 2 axes is typical of large noisy datasets which includes many zero values (Vyverman *et al.* 1995). The second and third ordination axes are less important than the first two, judged from their eigenvalues, relative to that of the first axis.

Table 6.3. Summary statistics for CCA of full dataset of 50 samples and 7 forward selected variables.

Axes	1	2	3	4	Total inertia
Eigenvalues:	0.563	0.492	0.461	0.361	10.036
Species-environment correlations:	0.935	0.853	0.891	0.86	
Cumulative percentage variance					
of species data:	5.6	10.5	15.1	18.7	
of species-environment relation:	22.4	41.9	60.2	74.5	
Sum of all canonical eigenvalues					2.521

Table 6.4 illustrates the environmental variable loadings for each forward selected variable in the CCA axes. The most influential variable identified, according to its loading on axis one, is the environmental variable pH. This variable is strongly, negatively correlated to axis 1 (-0.72). The second strongest variable is aluminium which is positively correlated to the axis (0.66), followed by sulphate (0.58) and nitrate (0.54). On axis 2 the strongest, correlated environmental variable is electrical conductivity which is positively correlated with the axis (0.44). The variables magnesium and sodium have the lowest loadings on the axes, 0.26 on axis 1 and 2 respectively.

Table 6.4. Environmental variable loadings.

	Axis 1	Axis 2	Axis 3	Axis 4
pH	**-0.72**	0.07	-0.13	-0.09
EC	0.10	0.44	0.57	-0.04
Al	**0.66**	-0.05	0.46	-0.29
Mg	0.26	0.25	0.64	-0.10
Na	-0.08	0.26	0.60	0.39
$S0_4$	**0.58**	0.33	0.44	0.17
N	**0.54**	0.07	-0.08	-0.40

Figure 6.3 shows a CCA biplot of the seven environmental variables, after stepwise forward selection, together with the sample sites. In this figure, the vector of an environmental variable and its length is proportional to the strength of correlation with changes in the diatom assemblage. Small angles between environmental variables and axis 1 suggest high correlations with changes in patterns in the biological data. As suggested by the statistical output pH is the primary variable constrained to axis 1, with 5.6% of taxon variance attributed to the first axis. Evident in the CCA biplot (Figure 6.3), pH and aluminium are highly negatively correlated. This relationship is not surprising as the solubility and speciation of aluminium is pH-dependant (Cummins 1994). Aluminium and pH are the variables which are most associated with axis 1. Magnesium, sulphate and nitrate are also correlated with axis 1 but the shorter lengths of their vectors and greater angle show that they are less important in explaining variation in the taxon data than the variables pH and aluminium. The second axis is represented, in order of influence, by electrical conductivity and then sodium. The two variables magnesium and sulphate are highly correlated, indicated by the narrow angle between the two arrows.

Figure 6.4 shows a biplot of taxa scores. A list of species codes is given in Appendix 4.When examining the diatom taxa spread along the pH, aluminium vectors, it can be seen that there are a number of taxa that are associated with either low or high pH and aluminium levels. For instance, the taxon *Nitzschia vasta* appears to favour low pH conditions and high aluminium concentrations, while *Epithemia cistula* (Ep.ci) prefers more alkaline waters with low concentrations of aluminium. Associated with high conductivity are *Pinnularia joculata* (Pi.jo) and *Nitzschia* spec. 36 (Ni.sp36). At the other end of the conductivity gradient are species, such as *Diadesmis confervacea* (Di.co) and *Cocconeis placentula* (Co.pl) which appear to be associated with more dilute waters. The third gradient is related to the environmental variables magnesium, sulphate and nitrate. Associated with high concentrations of these ions are *Pinnularia* cf. *schoenfelderi* (Pi.sc) and *P. subcapitata* (Pi.su). At the other end of this gradient are *Achnanthidium minutissimum* v. *exilis* (Ac.mve) and *Nitzschia* spec. 39 (Ni.sp39). The weakest gradient is sodium which seems to be associated with *Nitzschia archibaldii* (Ni.ar) and *Nitzschia hantzschiana* (Ni.ha). It appears then that there are a number of taxa in the dataset which may be good indicators of elevated aluminium concentrations and acidity.

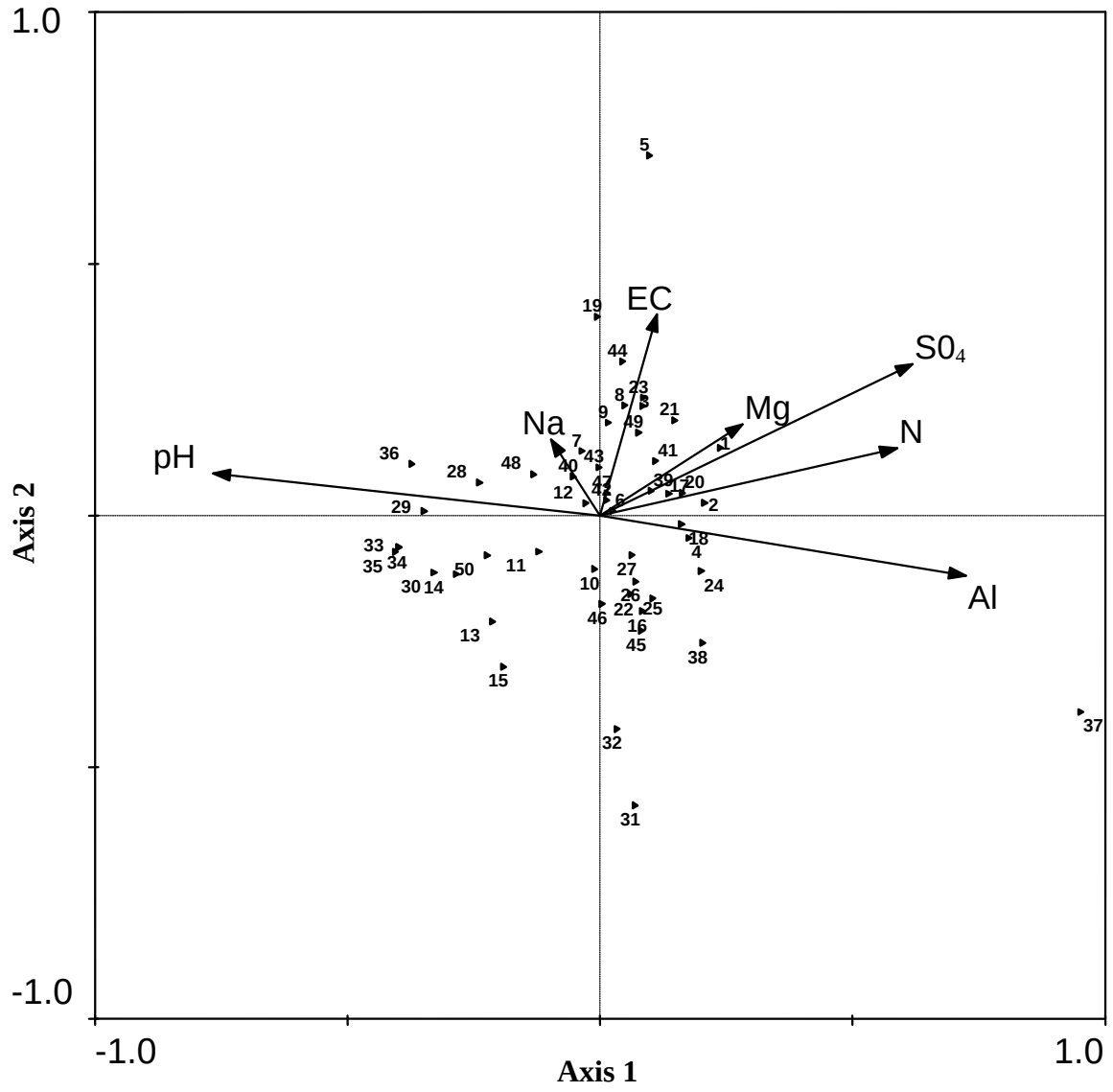


Figure 6.3. Canonical correspondence analysis (CCA) biplot showing the 7 forward-selected environmental variables and samples.

Figure 6.3 shows a CCA biplot of sample scores for the species dataset. Sites in the upper right quadrant of the biplot are associated with high sulphate, nitrate and magnesium. For instance, the concentration of sulphate at site RJI1 (1) was 4600 mg/L. The sites plotting high on axis 2 appear to be strongly influenced by *Pinnularia joculata* (Pi.jo) which plots to the extreme positive end of axis 2. *Pinnularia joculata* represents 94% of valves at site RJI5 (5). Therefore it is not surprising that this site plots high on axis 2. This site, along with TGI4 (19) and CHI8 (44), appear to be associated with higher conductivity, in all cases these sites measured above 770 µScm^{-1}. Site RJI1 (1) seems to be more influenced by magnesium than aluminium or pH. Sites which plot at the high end of the aluminium vector, and at the low end of the pH gradient, comprise CHI1 (37), CHI2 (38),

RJI2 (2), HCUSEC (24), BKI1 (45) and BKI2 (46). These sites were separated in the DCA as sites of high impact. Sites plotting at the other end of this gradient are the Redbank control sites CCUSSC (33), SCUSCC (34), SCDSCC (35), the Cosmo Howley control sites CHC2 (48), CHC4 (50) and the Rum Jungle control sites RJC3 (13), RJC4 (14) and RJC5 (15).

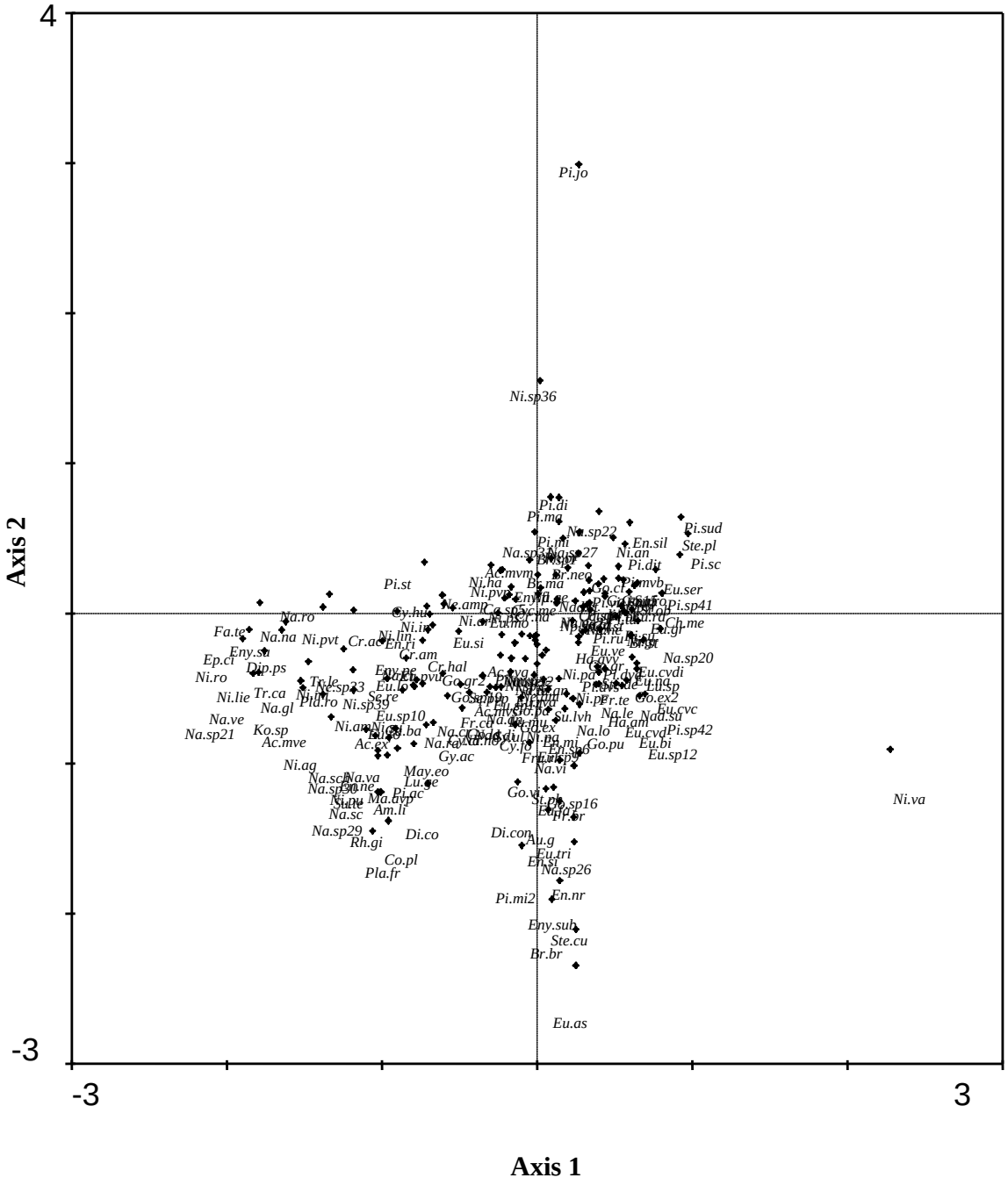


Figure 6.4. CCA biplot of all species from the dataset.

6.2.1.3 Gradient analysis of selected environmental variables

A series of detrended canonical correspondence analyses (DCCA), based solely on each selected variable, was undertaken to confirm their potential predictability. This was verified by testing the first axis of each constrained DCCA. The first axis of a DCCA, constrained to pH, had a gradient length of 2.44 standard deviations (S.D.), indicating unimodel-based techniques (weighted averaging) as most appropriate for pH inference models. The gradient lengths for the other selected variables, other than aluminium (1.12), were also high enough to justify the use of unimodel-based techniques (S.D. ≥ 2). Due to the short gradient length of aluminium, this variable cannot be used for the development of a predictive model. Although copper was not selected in the CCA forward selection process, it is highly correlated to pH (-0.71), which was selected in the CCA, and can therefore be potentially used for inference models. Variables eliminated in forward selection may well be ecologically important, but statistically unimportant. Its correlation to aluminium and other heavy metals may have contributed to insignificant p-values during forward selection. Copper was also ranked by AquaRisk analysis in section 4.4 as being one of the two top ranking bioavailable metals of potential concern which exceeds the Australian water quality guidelines. The gradient length of copper was also above two standard deviations. The potential of the dataset for the development of a pH model will be investigated in the next section using weighted averaging procedures.

6.2.1.4 Weighted Averaging Partial Least Squares

The inference ability of transfer functions can be examined by comparison of actual values recorded at each site against those predicted by Weighted Averaging (in the program CALIBRATE) species optima. According to the detrended canonical correspondence analysis (DCCA) gradient length results, unimodal inference models were used for the selected environmental variables. The performance of the dataset can be calculated as a measure of the inference power most commonly expressed as the root mean squared error (RMSE), which represents a reliable measure of the difference between measured and predicted values, and R^2, which is the correlation of determination between inferred and observed values. The RMSE unit depends on the variable predicted and, optimally, the RMSE should be lower than 10% of the range of the variable of interest (Philibert *et al.* 2006). Weighted average models with apparent R^2 values of less than 0.6, and R^2 jackknife less than 0.35, were rejected (as proposed in Philibert *et al.* 2006). Jackknifing was used in this study as a form of cross validation which involves the calculation of environmental estimates for each site based on a dataset which excludes that site.

The results of the weighted averaging partial least squares (WA-PLS) models are summarised in Table 6.5. The WA-PLS models of environmental variables with species data showed that

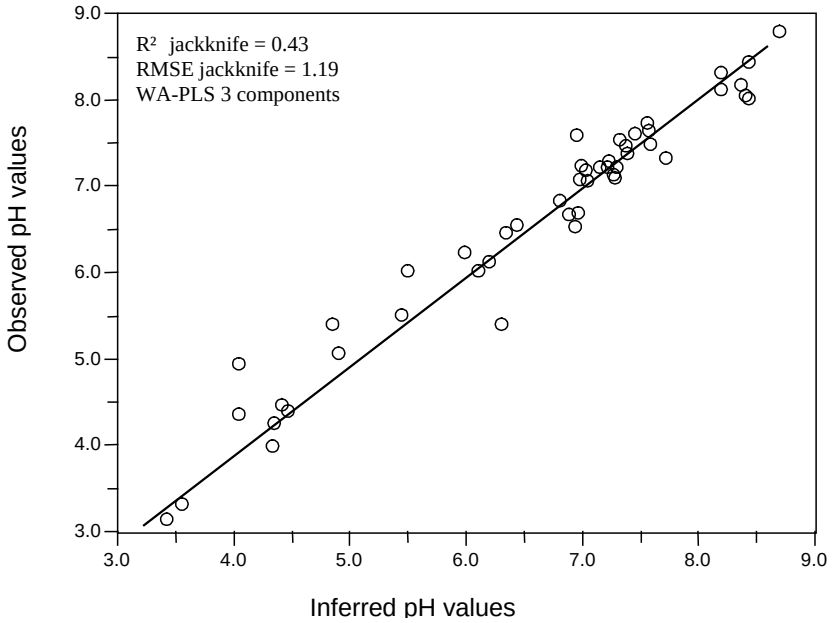
inferences from diatom species and each selected environmental variable (pH, EC, Mg, Cu and SO_4) were significant and strong showing high predictive capacity. The apparent R^2, for the environmental variables pH and sulphate, reached 0.96 ($R^2_{jackknife}$ = 0.43, RMSE = 0.29) and 0.97 ($R^2_{jackknife}$ = 0.41, RMSE = 0.16) respectively. This indicates that the relationship is strong and significant. In the case of conductivity the apparent R^2 reached 0.91 ($R^2_{jackknife}$ = 0.37). The WA-PLS model plots of inferred versus observed values for the five variables, pH, sulphate, electrical conductivity, magnesium and copper, are shown in Figure 6.5. The relation between the inferred and the observed values was also strong and significant for magnesium and copper. The apparent R^2 reached, 0.95 and 0.95 ($R^2_{jackknife}$ = 0.41 and 0.36) respectively.

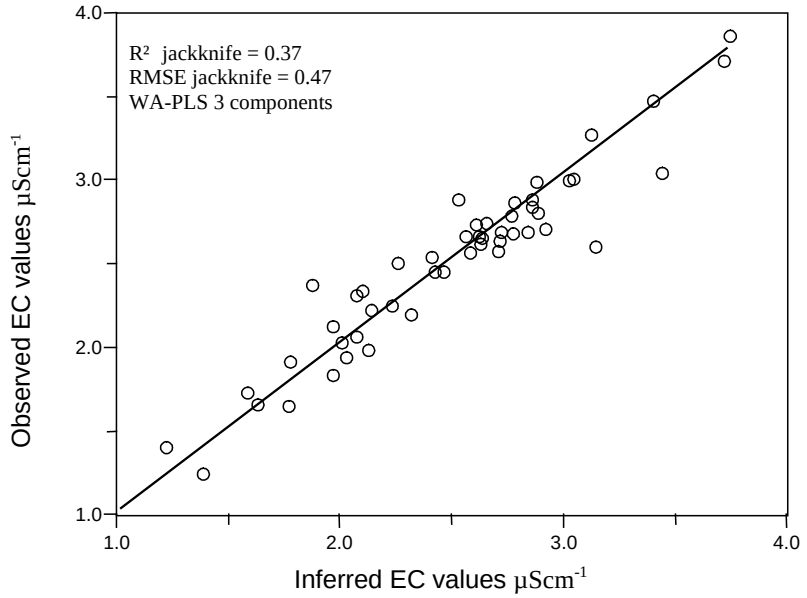
One sample was identified as an outlier that significantly decreased the performance of the model for the environmental variables sulphate and pH. The site RJI3 was removed from WA-PLS analysis of these two environmental variables. This site is unique as the diatom assemblage is almost completely dominated (93% relative abundance) by a single species, *Achnanthidium minutissimum* v. *minutissimum*. However this species is not rare or absent at other sites and has, in fact, high relative abundances at several other sites. The sample was only removed as the overall performance, in terms of RMSE and R^2, improved in its absence. In a number of other studies (Birks *et al.* 1990; Gasse *et al.* 1995; Bennion *et al.* 1996; Gell 1997) it has been shown that diatom water quality models may be improved by the removal of samples which reduce model predictive ability, but which add little in terms of ecological information. Often the presence of outlying samples may be due to truncated environmental gradients (Reid *et al.* 1995), which may improve by the addition, rather than the removal, of samples (Hall & Smol 1992). Each of the plots show good linear relationships (Figure 6.5) for the loged metal concentrations. However, the spread of the sites for the copper models is very clustered between the concentration values 0.0 and 0.2 (log) mg/L

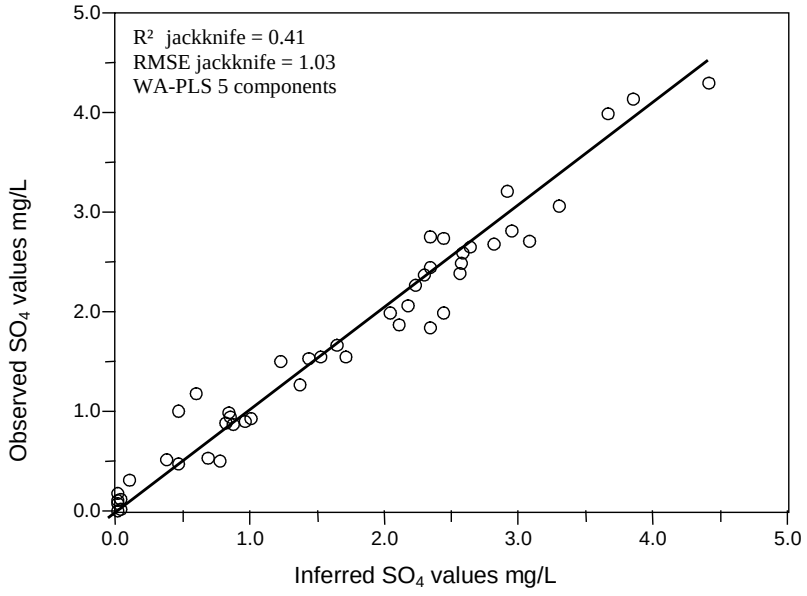
Table 6.5. Results of the WA-PLS for the calibration set (n=50).

	Mg	**Cu**	**SO_4***	**EC**	**pH***
WA-PLS components for apparent	4	3	5	3	3
Apparent R^2	0.95	0.95	0.97	0.91	0.96
R^2 jackknife	0.41	0.36	0.41	0.37	0.43
Apparent RMSE	0.15	0.09	0.16	0.17	0.29
RMSE jackknife	0.55	0.36	1.03	0.47	1.19
Number of samples used	50	50	49	50	49

*sample RJI3 was excluded


R² jackknife = 0.43
RMSE jackknife = 1.19
WA-PLS 3 components
Observed pH values
Inferred pH values
9.0
8.0
7.0
6.0
5.0
4.0
3.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0


R² jackknife = 0.37
RMSE jackknife = 0.47
WA-PLS 3 components
Observed EC values µScm-1
Inferred EC values µScm-1
4.0
3.0
2.0
1.0
1.0
2.0
3.0
4.0


R² jackknife = 0.41
RMSE jackknife = 1.03
WA-PLS 5 components
Observed SO4 values mg/L
Inferred SO4 values mg/L
5.0
4.0
3.0
2.0
1.0
0.0
0.0
1.0
2.0
3.0
4.0
5.0

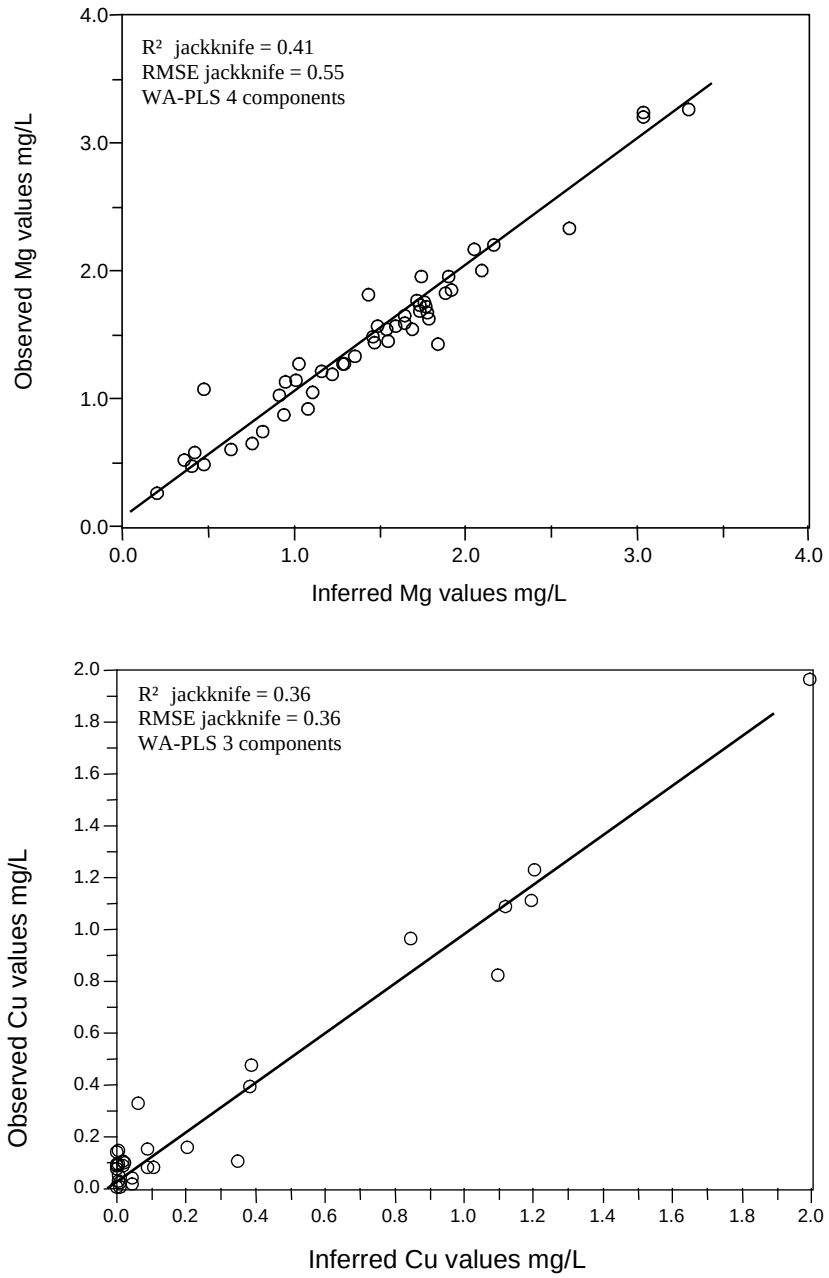


Figure 6.5 Plots of inferred versus observed values of pH, EC, $S0_4$, Mg and Cu concentrations (loged) based on WA-PLS models.

6.2.1.5 Discussion

A detrended correspondence analysis (DCA) was used to examine the patterns in diatom variance from the dataset. The four axes together explained 27.9% variance in the diatom data. The detrended canonical analysis of the species dataset was successfully able to separate very impacted sites from those which were moderately impacted or control sites. The sites which plotted to the far right of axis 1, and labelled as very impacted, include RJI1, RJI2, RJI3, RJI4, TGI4, TGI5, CHI1, CHI2, CHI3, BKI1, BKI2 and the Redbank sites HCUSEC, ECDSHC and ECUS12MC. The sites clustered into two groups on the bases of their species composition. The main species associated

with the impacted sites are *Nitzschia vasta, Nitzschia nana* and *Nitzschia paleaeformis*. *Nitzschia paleaeformis* is known to be characteristic of acidic waters (Denys & van Straaten 1992). Site RJI3, which is also highly impacted did not plot with the other species as it was dominated by *Achnanthidium minutissimum* v. *minutissimum.* This taxon, together with *Nitzschia palea*, was found by Sabater (2000) to withstand mine spillage in Spain. *Nitzschia palea* is argued to be a pollution tolerant species although it plots high on axis 2 close to sites of low impact (sites 28, 29 and 40). These taxa have been included among those most resistant to heavy metal pollution (Deniseger *et al.* 1986; Ivorra *et al.* 2000). It appears that, from these observations, a number of taxa in the dataset are good indicators of high heavy metal concentrations while others may have a pollution tolerance which is too broad. Taxa which are associated with the low impact sites are *Achnanthidium minutissimum* v. *minutissimum, Brachysira neoexilis* and *Nitzschia palea* v. *tenuirostris*.

This study followed the established procedure for the development of a WA-PLS transfer function. In summary, this involved, firstly, the use of CCA to determine which variables were responding most to the changes in diatom assemblages. Gradient analysis of a diatom dataset from the Northern Territory has shown that pH explains a relatively large, and significant, proportion of variation in the data. The identification of pH as a significant, independent, explanatory variable, which explains a relatively large portion of variation in the data, indicates the suitability of this dataset for the derivation of a model for pH. The next most influential variables in order of decreasing magnitude were aluminium, sulphate, conductivity, magnesium and sodium. It is not surprising that these variables are important as they are often associated with heavy metal pollution as with a decrease of pH levels to below 3.5, dissolution of aluminium and iron containing minerals occurs (Noller *et al.* 1997). Additionally, if conductivity is important, then so will other ions unless they are expressed as a proportion of cations, anions by milli equivalents. It is interesting however, that electrical conductivity is neither negatively nor positively correlated with pH on the CCA plot, as elevated levels of conductivity are often associated with acidity in heavy metal pollution. This suggests that, in the streams, diatoms were responding to changes in pH and heavy metals rather than to small changes in nutrient concentrations. The nutrient variable is shown to be important and in previous studies has been proven to be a factor affecting diatom composition. The amount of taxon variance explained by the first four axes was 18.7%. The six variables explained 49.9% of the total taxon variance. Possible indicator species associated with low pH and high aluminium concentration in the CCA are *Nitzschia vasta* and *Nitzschia paleaeformis* while *Nitzschia nana, Nitzschia palea* v. *tenuirostris* and *Epithemia cistula* seem to prefer low pollution levels. *Pinnularia joculata* and *Nitzschia* sp. 36 and, to a lesser extent, *Achnanthidium minutissimum* v. *minutissimum,* are associated with high levels of electrical conductivity. These findings are in contrast with those of John (1993) who found that *Achnanthidium minutissimum* and *Pinnularia subcapitata* can be used as markers for transition from acid to alkaline conditions. *Pinnularia*

subcapitata and *Pinnularia* cf. *schoenfelderi* in this study were actually associated with high concentrations of sulphate and magnesium. van Dam and Mertens (1990) found *Eunotia exigua*, *Navicula subtilissima* and *Pinnularia subcapitata* v. *hilseana* to be highly abundant in copper polluted lakes. The main species associated with low electrical conductivity, magnesium and sulphate are *Diadesmis confervacea, Cocconeis placentula, Nitzschia* sp. 39 and *Achnanthidium minutissimum* v. *exilis.*

Weighted averaging regression was carried out on the environmental variables pH, magnesium, sulphate, conductivity and copper. Copper was included as it is strongly correlated to pH, as seen in Table 4.1, and is an important variable of acid mine drainage (AquaRisk). Each inference model performed well ($R^2 > 0.6$) and the diatom inferred pH values are in agreement with the directly observed pH values. Compared to other published models, the goodness-of-fit (R^2) of the model was strongest for pH and was comparable to that found in Canada. For instance, the surface water acidification program pH calibration model (Dixit & Smol 1994) ($R^2 = 0.86$; RMSE = 0.25) is comparable to the pH calibration model developed for lakes in the Adirondacks ($R^2 = 0.91$; RMSE = 0.31; Dixit *et al.* 1993). However, the cluster of sites at the negative end of the copper gradient indicated that there was not a good spread of data. More sites need to be sampled at levels of high copper concentrations.

In conclusion, a transfer function for pH, magnesium and copper has been developed which is applicable to heavy metal impacted tropical water bodies. A number of taxa have been shown to be possible indicator species and this will be developed in a later section.

6.2.2 Genus Transfer function

There are a variety of methods in which diatoms are used to show the impacts of environmental disturbance. Most commonly, multivariate methods such as prediction models, ordination and transfer functions are used based on the composition of diatom assemblages. To date the use of diatoms in ecological disturbance monitoring using transfer functions has been mainly conducted with species level identification (e.g. Stevenson 1991; Whitton *et al.* 1991; Kwandrans 1993). However, it has been suggested that one obstacle to widespread use of diatoms for biological monitoring may be the large number of species and the need for taxonomic expertise (Kelly *et al.* 1995). A weakness of diatom analysis for routine monitoring in Canada is that the ecological preferences of many diatoms are not yet known. This pertains to species endemic to North America (Patrick & Reimer 1966), as most models are based on diatom observations in Europe. Moreover, because of the high number of species, continuously changing taxonomy, and incomplete or non-existent floras, diatom analysis can be very time consuming and, in contrast to species taxonomy, diatom genera can be well defined (Round 1996). When very similar species have contrasting ecological tolerances, expensive equipment, such as a scanning electron microscope, is mandatory for species identification. To counter this problem, a few researchers have used genus-level identification for the assessment of stream condition (e.g. Prygiel 1991; Prygiel & Coste 1993; Kelly *et al.* 1995; Chessman *et al.* 1999; Growns 1999; Wu 1999; Hill *et al.* 2000; Wunsam *et al.* 2002). From such studies, specific genera, for instance *Eunotia*, are known to be strong indicators of acid, oligotrophic waters (van Dam *et al.* 1994).

The main objective of this section was to explore the relationship between diatom genus assemblages along environmental pollution gradients from the four mine regions and to identify the environmental variables that would best explain the distribution of diatom genera and to reconstruct this variable/s to see if they can be used for inference models. Secondly, the species and genus transfer functions are compared to determine the level of taxonomic resolution which best infers impacts of heavy metals.

6.2.2.1 Similarities and dissimilarities in diatom genus assemblages

To explore the similarities or dissimilarities in diatom assemblages between impacted and non-impacted sites, a detrended correspondence analysis (DCA) was carried out grouping the 40 diatom genera which occurred at more than one of the 50 sites. Table 6.6 indicates the variance explained by the four DCA axes while the samples and taxon scores are plotted in Figure 6.6. The Rum Jungle site, site RJI5, was excluded from the final DCA analysis as, in the initial DCA, it was defined as an outlier. At this site the genus *Pinnularia* occurred at 96% relative abundance. This site skewed the DCA plot, disguising the relationships between species and sites. Once it was

removed from the dataset the percentage of species variance explained by four axes increased from 35.2% to 37.8%. In the DCA of the genera and sites the percentage of cumulative variance explained by the first axis was 14.3% increasing to 26% on the second axis and to 33.5% on the third axis. The four axes together explained 37.8% variance in the diatom genus data. The first two axes of the CA were significant according to Monte Carlo permutation tests (with 199 unrestricted permutations, $p < 0.05$).

Table 6.6. Summary statistics for DCA of the entire genus dataset with 49 samples.

Axes	1	2	3	4	Total inertia
Eigenvalues:	0.546	0.446	0.285	0.163	3.815
Lengths of gradient:	3.22	2.644	2.021	2.066	
Cumulative percentage variance of species data:	14.3	26	33.5	37.8	

The DCA plot (Figure 6.6) illustrates that the separation of low impacted and high impacted sites is less clear than the separation made with the species dataset. Sites which, in the previous section (section 6.2.1.1.) from the species dataset, were recognised as very impacted (RJI1, RJI2, RJI3, RJI4, TGI4, TGI5, HCUSEC, ECDSHC, CHI1, CHI2, CHI3, BKI1, BKI2) are highlighted on the genus DCA plot in red. The site numbers are listed in Appendix 1. The impacted sites, highlighted in red, plot in an almost linear fashion with site RJI3 (3) plotting at the lowest end of axis 1 and 2 and sites BKI1 (45) and CHI1 (37) plotting high on axis 1 and midway on axis 2. However, between the impacted sites ECDSHC (25) to CHI1 (37), circled, cluster an additional nine sites which were not previously indicated as very impacted in section 6.2.1.1. Of these nine sites, RJI6 (6), 12MCUSCC (28), SCDS12MC (29), CHI4 (40), CHI5 (41), CHI7 (43), CHI8 (44), CHC3 (49) and CHC4 (50), two are control sites and the remaining sites are downstream of the pollution source and not highly polluted in terms of their pH or metal concentrations. The pH of these nine sites ranges from 6.3 at site RJI6 (6) to pH 8.7 at SCDS12MC (29). Concentrations of aluminium (Al) are as low as 0.005 mg/L at site CHC4 (50) to 0.07 mg/L at site RJI6 (6). Site CHI4 (40) was excluded as this could potentially be included as an impacted site as its pH is 4.9 but, more importantly, its aluminium concentration is higher than 1 mg/L.

Several genera in diatom assemblages, those with high relative abundance, may be attributed to sites of high impact. The main genera associated with high impact sites are *Nitzschia* (Nit) and *Pinnularia* (Pin). The genus *Nitzschia* is present at each highly impacted site at at least 20% abundance and up to 99% abundance {CHI1 (1)}. One of the reasons why site RJI3 does not plot with the rest of the very impacted sites is that it has a relative abundance of *Nitzschia* of less than 3%. Sites RJI1 (1), RJI4 (4) and HCUSEC (24) have relative abundances of *Pinnularia* greater than 10%. The genus *Brachysira* (Bra) is present at 28% relative abundance at the impacted sites TGI4 (19) and 15% at site HCUSEC, however, it has higher relative abundances (>20%) at the low

impact sites CHI7 (43), CHI8 (44) and CHC3 (49). Consequently, *Nitzschia* and *Pinnularia* could be considered potential indicators of heavy metal pollution.

The positioning of a number of sites high on axis 2 appears to be strongly influenced by genera such as *Epithemia* (Epi) which plot to the extreme left of axis 2. The four sites CCUSSC (33), SCUSCC (34), SCDSCC (35) and SCUS12MC (36) are all control sites from the Redbank area. Each of these sites has more than 25% relative abundance of *Epithemia*. Site CCUSSC (33), which plots closest to *Epithemia*, have more than 60% relative abundance of this genus. These sites also contain low abundances (less than 10%) of *Diploneis* (Dip), *Encyonopsis* (Ency), *Cymbella* (Cym), *Navicula* (Nav) and *Planothidium* (Pla). At these sites *Nitzschia* relative abundances are greater than 10%. However, this genus has a relative abundance of up to 90% at more impacted sites, such as CHI1 (37), and therefore plots closer to CHI1. The control site RJC5 (15) plots nearby because of its greater than 20% abundance of each of the genera, *Cocconeis*, *Navicula* and *Planothidium*. Evidently, the genus *Epithemia* could be a potential indicator of clean sites.

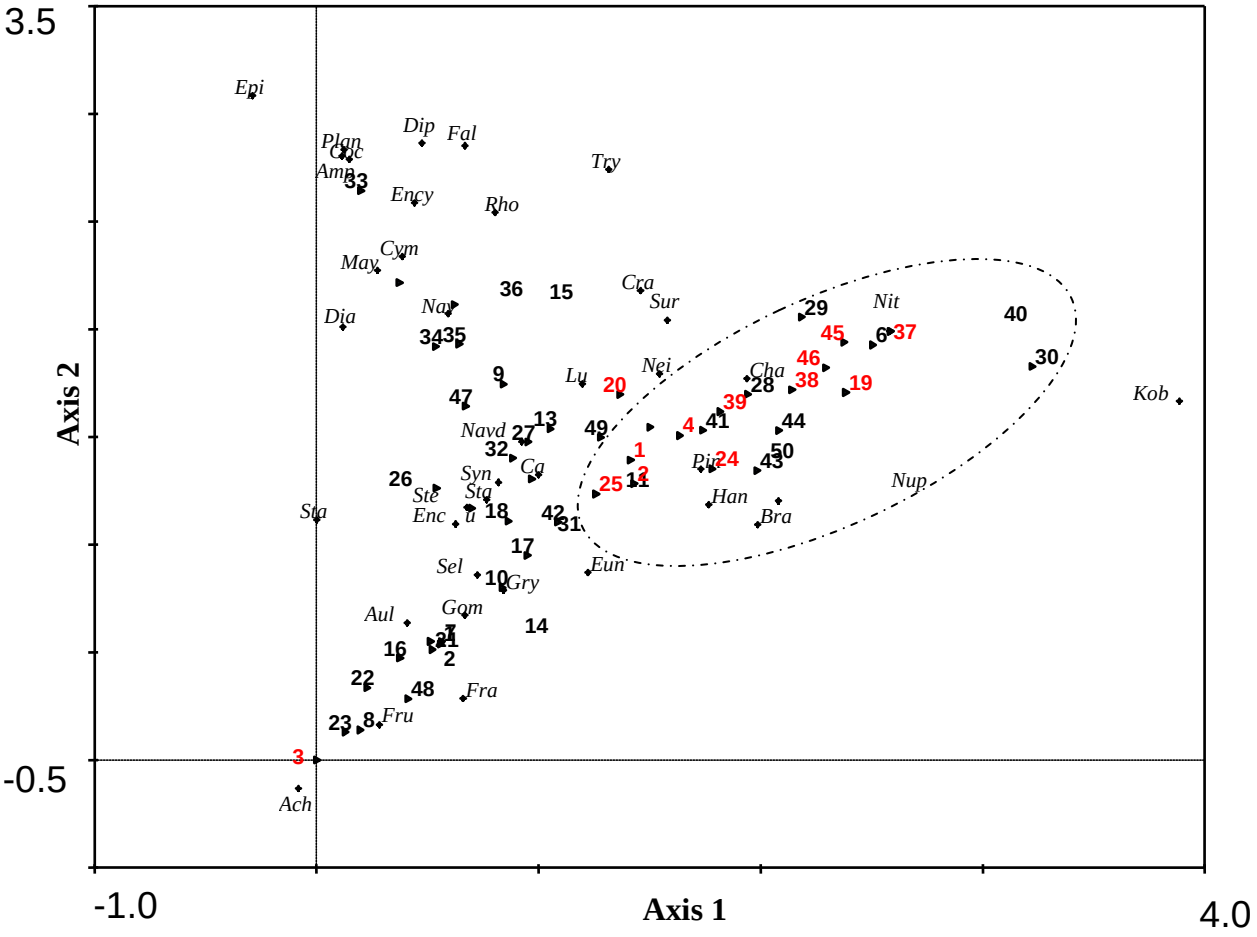


Figure 6.6. DCA biplot of genera and site scores. Sites classed as very impacted are highlighted in red.

The sites which plot lowest on axes 1 and 2 are most influenced by the genus *Achnanthidium* (Ach). This genus is present at greater than 50% abundance at any of the nearby sites, namely RJI3 (3), RJI7 (7) RJI8 (8), RJC2 (12), RJC4 (14), TGI1 (16), TGI6 (21), TGC1 (22), TGC2 (23) and site CHC2 (48). Other than site CHI3, these sites were each previously indicated as low impact by the species DCA in section 6.2.1.1. *Achnanthidium* is also present at a large number of other sites but at relative abundances of less than 50%. It seems that *Achnanthidium* could be an indicator genus for clean sites. Other genera related to these sites, such as *Aulacoseira*, *Frustulia* and *Fragilaria*, have relative abundances of less than 7%.

The Redbank site SC@BRX (30) is furthest downstream from the Redbank pollution source. However, it plots near to the cluster of very impacted sites on axis 1, to the right of axis 1 between *Nitzschia* and *Kobayasia*. The genus *Kobayasia* has a relative abundance of 50% and *Nitzschia* has a relative abundance of 48.5%. *Kobayasia* is present at other sites at abundances of less than 1%.

6.2.2.2 CCA of the genus dataset

Canonical correspondence analysis (CCA) of the 44 genera and 15 environmental variables was undertaken to explore the relationship between the diatom genera and measured environmental variables from the dataset. The relative influence and significance of each environmental variable in explaining diatom variation was determined through the process of stepwise forward selection. The variables were tested for significance by Monte Carlo permutation tests, in which 99 unrestricted permutations were used. Variables were considered to be significant, and included in the CCA, at the $p \leq 0.05$ threshold point. The results of the stepwise forward selection on the variables are shown in Table 6.7. Each of the seven variables indicated in the table below were selected based on their significance calculated by the Monte Carlo permutation test. The variables not included in the table each had a p-value greater than 0.05 level and were therefore not considered significant. Of the 15 variables, seven were chosen by forward selection, namely, turbidity (Turb), nitrate (N), aluminium (Al), sulphate (SO_4), conductivity (EC), magnesium (Mg) and sodium (Na).

Table 6.7. Percentage of variance explained by the variables retained after a stepwise forward selection.

Variable	% variance explained before addition	% variance explained after addition	p -value
Na	0.236	0.236	**0.002**
Turb	0.169	0.405	**0.002**
$S0_4$	0.158	0.562	**0.002**
EC	0.127	0.689	**0.01**
Mg	0.116	0.805	**0.022**
N	0.126	0.931	**0.016**
Al	0.116	1.047	**0.034**

A CCA was executed on the forward-selected variables (Na, EC, Mg, SO_4, N, Al and Turbidity), 50 samples and the 43 diatom genera. A summary of the results are given in Table 6.8. The Cosmo Howley site, CHI1, was excluded from analysis because of its extreme influence (>15x) detected by leverage diagnosis in the initial CCA. The environmental variables used in the CCA explained 25.9% (sum of all canonical eigenvalues = 3.4) of the total variation in the diatom data. As can been seen from examination of Table 6.8, the combined variance explained by axis 1 (9.7%) and 2 (16.3%) is 26%. However, these two axes express less than half of the variance in genera in the dataset (total = 73.4%). The second and third ordination axes are less important than the first two judged from their eigenvalues respective to that of the first axis.

Table 6.8. Summary of statistics for CCA of genus dataset of 49 samples and 7 stepwise forward selected variables.

Axes	1	2	3	4	Total inertia
Eigenvalues:	0.327	0.223	0.177	0.147	3.375
Species-environment correlations:	0.871	0.727	0.694	0.67	
Cumulative percentage variance					
of species data :	9.7	16.3	21.5	25.9	
of species-environment relation:	31.2	52.5	69.4	83.4	
Sum of all canonical eigenvalues					1.047

The most influential variables identified according to its loading on axis one is the environmental variable sodium. This variable is strongly, negatively correlated to axis 1 (0.61) (Table 6.9). The second strongest variable is conductivity, which is positively correlated to the axis (0.54), followed by magnesium (0.53). On axis 2 the strongest, correlated environmental variable is sulphate which is positively correlated with the axis (0.58). The variable turbidity has the highest loading on axis 4 (-0.45). The variables nitrate and aluminium have the lowest loadings on the axes, -0.40 and 0.35 on axis 1 and 2 respectively.

Table 6.9. Environmental variable loadings.

	Axis 1	Axis 2	Axis 3	Axis 4
EC	**0.54**	0.42	-0.17	0.00
Turb	-0.21	0.10	0.09	-0.45
Al	-0.01	0.35	0.11	0.17
Mg	**0.53**	0.39	-0.11	-0.08
Na	**0.61**	0.31	0.19	-0.15
$S0_4$	0.18	**0.58**	0.04	-0.13
N	-0.40	0.22	-0.28	-0.12

Figure 6.7 shows a CCA biplot of genus and environmental variables after stepwise forward selection. The seven environmental variables, turbidity, nitrate, aluminium, sulphate, electrical conductivity, magnesium and sodium are represented by arrows. In this figure, the vector of an environmental variable and its length is proportional to the strength of correlation with changes in the diatom assemblage. As suggested by the statistical output, sodium is the strongest variable constrained to axis 1, with 9.7% of taxon variance attributed to the first axis. However, as is evident in the CCA biplot, magnesium and conductivity are highly, positively correlated with sodium as indicated by the close angles between the variables vectors. Consequently, an increase in sodium is coupled with an increase in magnesium and conductivity. The environmental variables, turbidity and nitrate, are also correlated to axis 1 but the shorter lengths of their vectors show that they are less important than the variables sodium and conductivity in explaining variation in taxon data. The second axis is represented primarily by the variables sulphate and aluminium. Aluminium has a shorter vector length but both variables are highly correlated. The least correlated variables are turbidity and sodium, as indicated by the large angle between their two vectors.

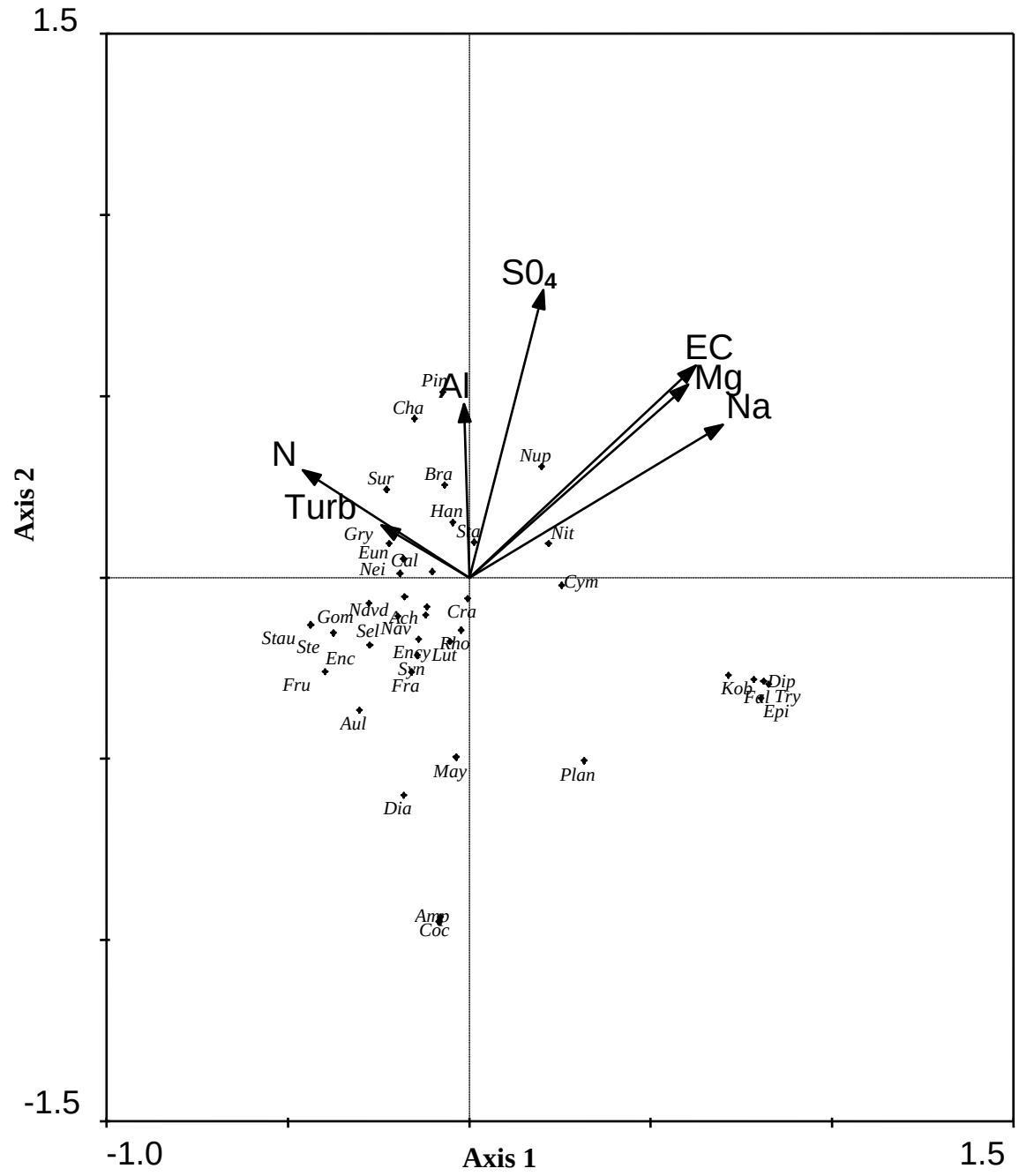


Figure 6.7. CCA biplot of genera scores and environmental variables.

Figure 6.7 shows a biplot of genus and environmental variable scores for the dataset. When examining the diatom taxa spread along the conductivity, magnesium and sodium vectors, it can be seen that there are a number of genera that are associated with either the low or high end of this gradient. For instance, *Nitzschia* (Nit) appears to favour high magnesium, sodium and conductivity conditions, while *Frustulia* (Fru) and *Aulacoseira* (Aul) prefer waters of lower metal concentration and conductivity. *Pinnularia* (Pin) is associated with high levels of aluminium and sulphate. *Amphora* (Amp) and *Cocconeis* (Coc) appear to be associated with low levels of aluminium and sulphate. However, *Amphora* has a relative abundance of less than 1% at any site and the genus

Cocconeis has only an abundance of greater than 1% at the site RJC5 (15.3%). This high abundance of this genus at site RJC5 (15) is indicated also by the close proximity of the genus and site on the biplot. The third gradient is comprised of the environmental variables, nitrate and turbidity. Associated with high measurements of these two variables are *Gyrosigma* (Gyr) and *Eunotia* (Eun) while at the other end of this gradient are *Kobayasia* (Kob), *Tryblionella* (Try), *Diploneis* (Dip), *Fallacia* (Fal) and *Epithemia* (Epi). It appears then that there are a number of genera in the dataset which may be good indicators of elevated aluminium and magnesium.

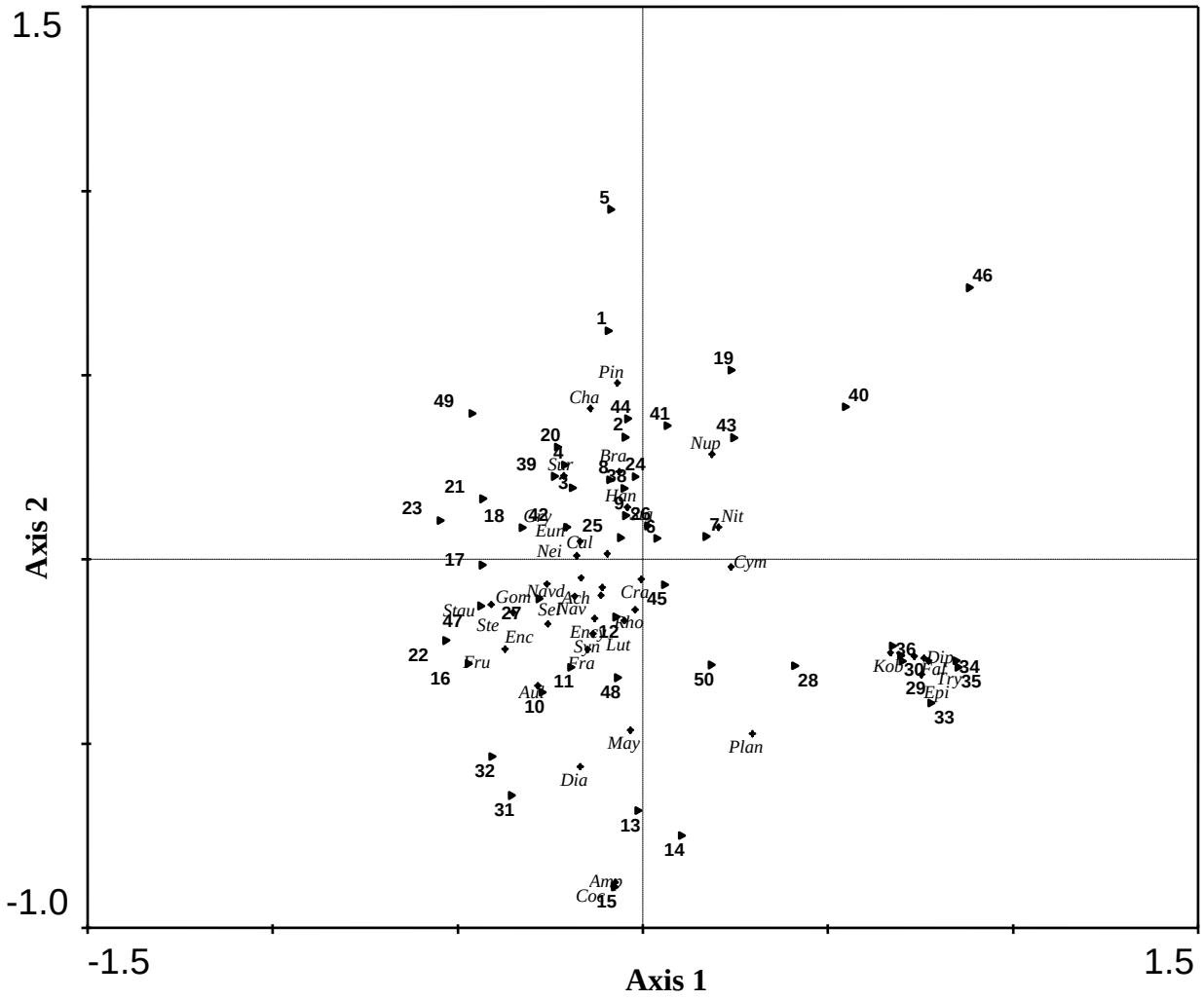


Figure 6.8. CCA biplot of genera and sample scores on axis 1 and 2.

Sites which are associated with high conductivity, magnesium and sodium plot in the upper right quadrant of Figure 6.8. These sites include RJI7 (7), CHI7 (43), CHI4 (40) and BKI2 (46). Site RJI7 plots closest to the central area of the biplot while site BKI2 (46) plots at the upper end of the conductivity, magnesium, sodium vectors and has the highest concentrations from these. Sites associated with high aluminium and sulphate plot in the upper central section of the CCA biplot. These sites include RJI1 (1), RJI2 (2), RJI5 (5), HCUSEC (24), CHI2 (38) and CHI8 (44). Sites which plot in the lower central and lower left quadrant of the biplot are associated with low metal concentrations of aluminium, sulphate, magnesium, sodium and electrical conductivity. Sites which

plot in the lower right quadrant are the Redbank sites SCDS12MC (29), SC@BRX (30), CCUSSC (33), SCUSCC (34), SCDSCC (35) and SCUS12MC (36). Each of these are low impact or control sites.

6.2.2.3 Gradient analysis of selected environmental variables

A series of detrended canonical correspondence analyses (DCCA), based solely on each variable, chosen in CCA stepwise selection, was undertaken to verify their potential inference ability. This was verified by testing whether the first axis of each constrained DCCA had a gradient length of greater than two standard deviations. The first axis of a DCCA constrained to sulphate had a gradient length of 1.48 standard deviations, indicating that unimodel-based techniques (weighted averaging) were not appropriate for sulphate reconstructions. The gradient lengths of the other six selected variables were each less than two standard deviations and were, therefore, also not associated with diatom assemblages in a unimodal fashion. The variable electrical conductivity had a gradient length of 1.92 standard deviations, sodium 1.9 standard deviations, turbidity 0.99 standard deviations, magnesium 1.8 standard deviations, aluminium 1.52 standard deviations and nitrate had a gradient length of 1.45 standard deviations.

As the gradient lengths for each of the seven environmental variables was lower than two standard deviations unimodal inference models were not able to be produced. Therefore, a partial least squares transfer function was run in the program CALIBRATE. The partial least squares (PLS) model of environmental variables with genus data showed that inferences between diatom genera and the seven environmental variables were not significant or strong, indicating no predictive capacity (R^2jack < 0.35 and R^2apparent < 0.6). For instance, the apparent R^2 for sulphate was 0.37 (R^2 jack 0.22).

6.2.2.4 Discussion

The initial detrended canonical analysis (DCA) of the genus dataset separated the impacted sites and the less impacted sites less well than the DCA analysis of the species dataset. An additional nine sites clustered with the 14 sites which were defined by the species dataset DCA as very impacted. However, when viewing the chemical composition of these nine sites, it is obvious that they are not impacted. Two of the sites included in the cluster are control sites from the Cosmo Howley area. The percentage of 'species' variance explained is higher (37.8%) for the genus dataset compared to (23.1%) the species dataset. There are certain genera which are clearly associated with the impacted and less impacted sites. *Nitzschia* and *Pinnularia* are associated with the impacted sites while *Epithemia* is associated with the Redbank control sites. However, *Achnanthidium* is associated with control sites of the other three mines. Similarly, in the species

DCA, three *Nitzschia* species (*N. vasta*, *N. nana* and *N. paleaeformis*) were identified as influencing impacted sites in the species dataset, while the species *Achnanthidium minutissimum* v. *minutissimum* and *Brachysira neoexilis* were associated with low impact and control sites.

The amount of taxon variance explained by the first four axes in the genus CCA is 25.9%. This is higher than the 18.7% explained by the first four axes in the species dataset. The seven variables (Na, Turb, SO_4, EC, Mg, N, Al) in the genus dataset explained 73.4% of the total generic variance which is higher than the 49.9% explained by the six variables in the species dataset. However, the amount of the total variation that can be explained by the overall inertia (or variance in species dispersion) in the genus dataset is 3.38. The inertia for the genus dataset is low compared to that of the species dataset which is 10.04. Similarly the environmental variables (canonical eigenvalues) explained only a small portion (1.05) of the variation in genus composition compared to that of the species dataset (2.52). However, how well the measured variables explain species composition can be ascertained by dividing the explained variance by the total variance. This is 31.1% for the genus dataset and 25.1% for the species dataset. While the genus CCA better explains diatom variation the species CCA better discriminates between non-impacted and impacted sites.

Sodium and sulphate were the most important environmental variables for explaining genus level variation in the dataset followed, in order, by conductivity, magnesium, nitrate, aluminium and turbidity. *Pinnularia* is associated with high concentrations of aluminium and sulphate. *Nitzschia* is associated with high concentrations of magnesium and sodium, while at the other end of the gradient, *Achnanthidium* is associated with low levels of these metals. Low levels of nutrients and turbidity were associated with *Epithemia*.

This compares to the species CCA where the variables, in order of importance in explaining species variation, are pH, aluminium, sulphate, nitrate, conductivity, magnesium and sodium. The difference between the two datasets is the importance of pH in explaining the species variation. *Nitzschia vasta* is associated with low pH and high aluminium and, at the other end of the gradient, *Epithemia cistula* is associated with alkaline conditions and low concentrations of aluminium. At high concentrations of magnesium *Pinnularia schoenfelderi* and *P. subcapitata* are associated. *Pinnularia joculata*, *Achnanthidium minutissimum* v. *minutissimum* and *Nitzschia palea* are associated with high levels of electrical conductivity. Typical to both datasets is that *Nitzschia* and *Pinnularia* genera are potential indicators of metals. These findings are similar to that of Hill *et al.* (2001) who classified *Nitzschia* and *Pinnularia* (as well as *Eunotia*) as pollution-tolerant genera. However, his conclusions were based on the environmental preferences of the genera determined as the mean environmental preference rating for species within a genus.

Evidently, different descriptors of diatom assemblages have different predictive potential. This was highlighted by Wunsam *et al.* (2002) who discovered that recognition of diatoms at the generic level explained a larger proportion of variance (47%) than recognition at the species level (31%). Additionally, in river monitoring studies in England (Kelly & Whitton 1995), and in France (Rumeau & Coste 1988), simplified diatom indices, based mostly on genera, responded well to nutrient gradients and were highly correlated to more complex, species-based indices. In Australia, in a comparison of the ability of different taxonomic levels of diatom assemblages to monitor the impacts of river regulation, Growns (1999) found similar performances at the generic and species level. Similarly, Lane (2007) concluded that indices of biotic integrity, when developed using autecological indices, provide similar qualitative, conditional information across taxonomic levels for isolated, herbaceous wetlands. However, in his study there were few species per genera, the highest being eight. Overall, 70% of the genera were represented by fewer than three species. Consequently, if these species had similar ecological requirements or responses to ecological disturbances, it would not matter which level was used. However, the Chessman *et al.* (1999) genus based index performs poorly (O_E/E) as most genera are predicted to occur everywhere. This has led to the production of a new species based diatom index for Australian rivers (DSIAR) (Chessman *et al.* 2007). However, this diatom index does not include diatoms from Northern Territory rivers.

Although, the generic level identification explained a larger proportion of variance than species level, it was not possible to construct weighted average models for the forward selected environmental variables to test their potential as inference models. A series of detrended canonical correspondence analyses (DCCA), based solely on each CCA stepwise selected variable, was done but was not successful because each variable had a standard deviation greater than two. Consequently, the transfer function using genus level identification was unsuccessful.

One disadvantage of not using species level identification is that several species within the same genus may have different ecological responses to a disturbance (Growns 1999). For example, in this study, there are some *Nitzschia* species which are acidophilic (associated with acidic conditions) while others are alkalophilic. For instance, *Nitzschia vasta* is associated with high concentrations of aluminium and acidic conditions while *Nitzschia nana* is associated with more alkaline waters. However, in the genus dataset, *Nitzschia* was associated with waters with high levels of electrical conductivity and magnesium. Both of these environmental variables are highly positively correlated with pH. This differing response of species within a genus may thus mask the response if only genus level information is used. Diatom species generally have much narrower environmental requirements and tolerances than whole genera (Round 1991).

6.2.3 Species transfer function with bioavailable metal concentrations

Diatoms can be powerful indicators of water quality because their species composition can be significantly related to water characteristics (Winter & Duthie 2000). One of the most severe instances of anthropogenic impact on waters is the acidification and heavy metal pollution of rivers and streams by mining activities. However, when relating heavy metals to species composition researchers have stressed that the total or dissolved concentration approach is inadequate as it does not reflect the true exposure to, and effect on, aquatic organisms (Payle *et al.* 1993; Janssen *et al.* 2000). This is true because the toxicity of a heavy metal is greatly dependant on how much metal is bioavailable to organisms. Increasingly, bioavailable fractions of metals are being included in Australian guidelines to protect aquatic environments (e.g. ANZECC 2000). It is important then, to consider the combined use of diatoms and the bioavailable fractions of heavy metals to produce stronger, and more realistic, results of species and environmental variable relationships. The aim of this section is to determine if the use of bioavailable metals in a diatom species transfer function produces stronger results than total metal concentrations.

6.2.3.1 Similarities and dissimilarities in species assemblages

A detrended correspondence analysis (DCA) was used to examine the patterns in diatom variance from the dataset. In the previous section all 50 sites were used in the dataset, however, with computation of the bioavailable fraction of each metal by AquaRisk, the program was not able to incorporate the two Redbank control sites CCUSSC (33) and SCUSCC (34) in the analysis. Consequently, the dataset is reduced in this section from 50 to 48 sites.

A DCA was conducted on 193 taxa which occurred in more than one sample from the 48 site dataset. Table 6.10 indicates the variance explained by four DCA axes while figures 6.9 and 6.10 are plots of sample and taxon scores. For the DCA analysis *Pinnularia joculata* and *Kobayasia* cf. nov. spec. were again identified as outliers and excluded from analysis. Axis 1 of the DCA explains 7.6 % of species variance. The fourth axis explains 21.1% of cumulative variance in the diatom data. Axis 1 accounts for the most of the explained variation in the dataset.

Table 6.10. Summary statistics for DCA of the entire dataset of 48 samples.

Axes	1	2	3	4	Total inertia
Eigenvalues:	0.735	0.601	0.391	0.328	9.73
Lengths of gradient:	4.775	4.282	3.254	2.911	
Cumulative percentage variance of species data:	7.6	13.7	17.7	21.1	

With the reduced dataset (48 sites), there is a greater spread of sites across axis 2 than with the previous species DCA. However, the positioning of the species and the relevant potential indicator species remains the same. For instance, *Nitzschia vasta* and *Nitzschia paleaeformis* plot on the far right of axis 1 while *Nitzschia palea* v. *tenuirostris* plots high on axis 2.

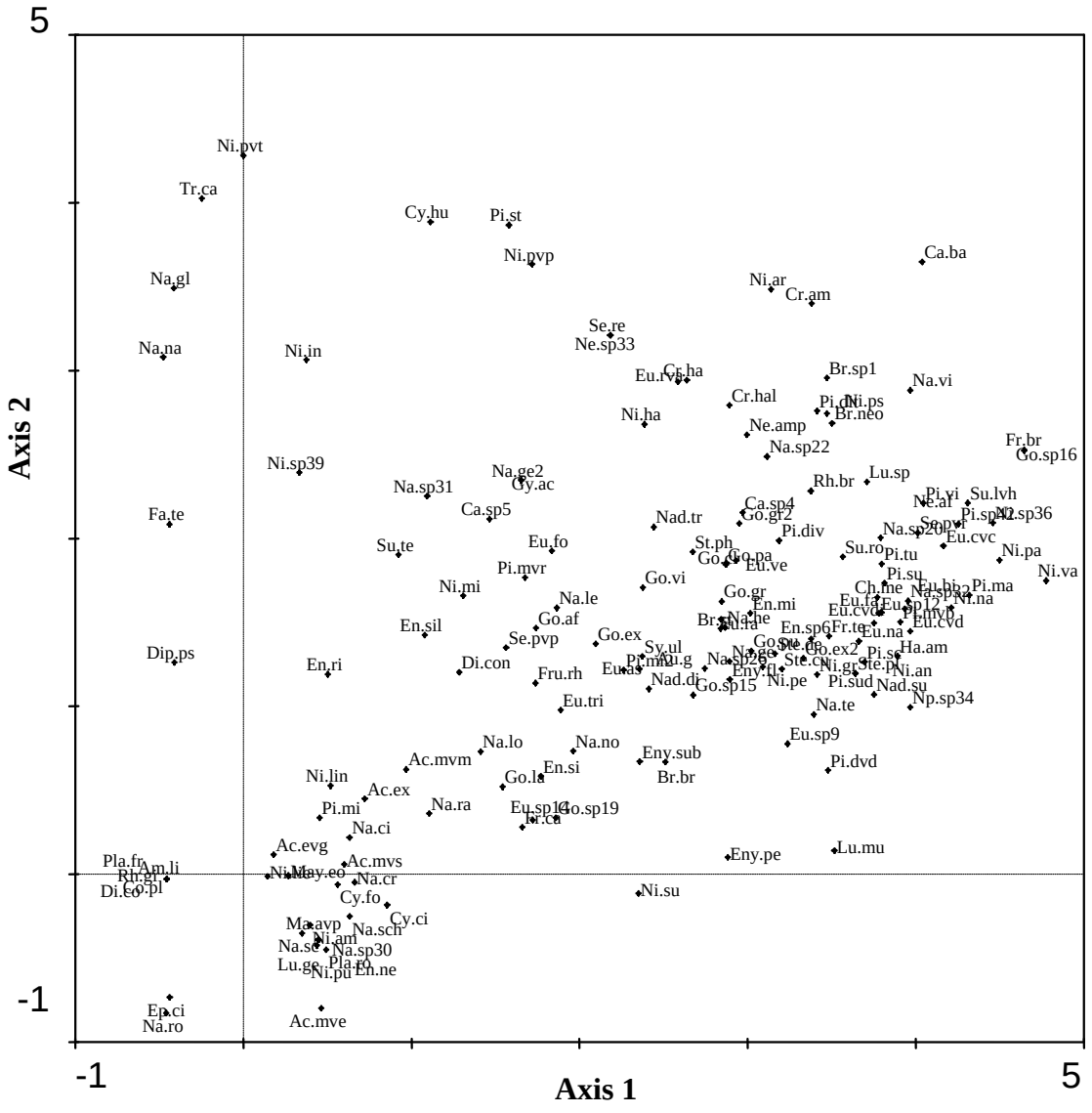


Figure 6.9. DCA scatterplot of taxon scores on axis 1 and axis 2 (48 sites).

When comparing the DCA of the reduced dataset to the initial species DCA in section 6.2.1.1., the sites exhibit the same clustering pattern. Only the position of the Rum Jungle sites RJI5 (5) and RJI9 (9) are slightly changed. These sites now plot higher on the axis 2 than previously.

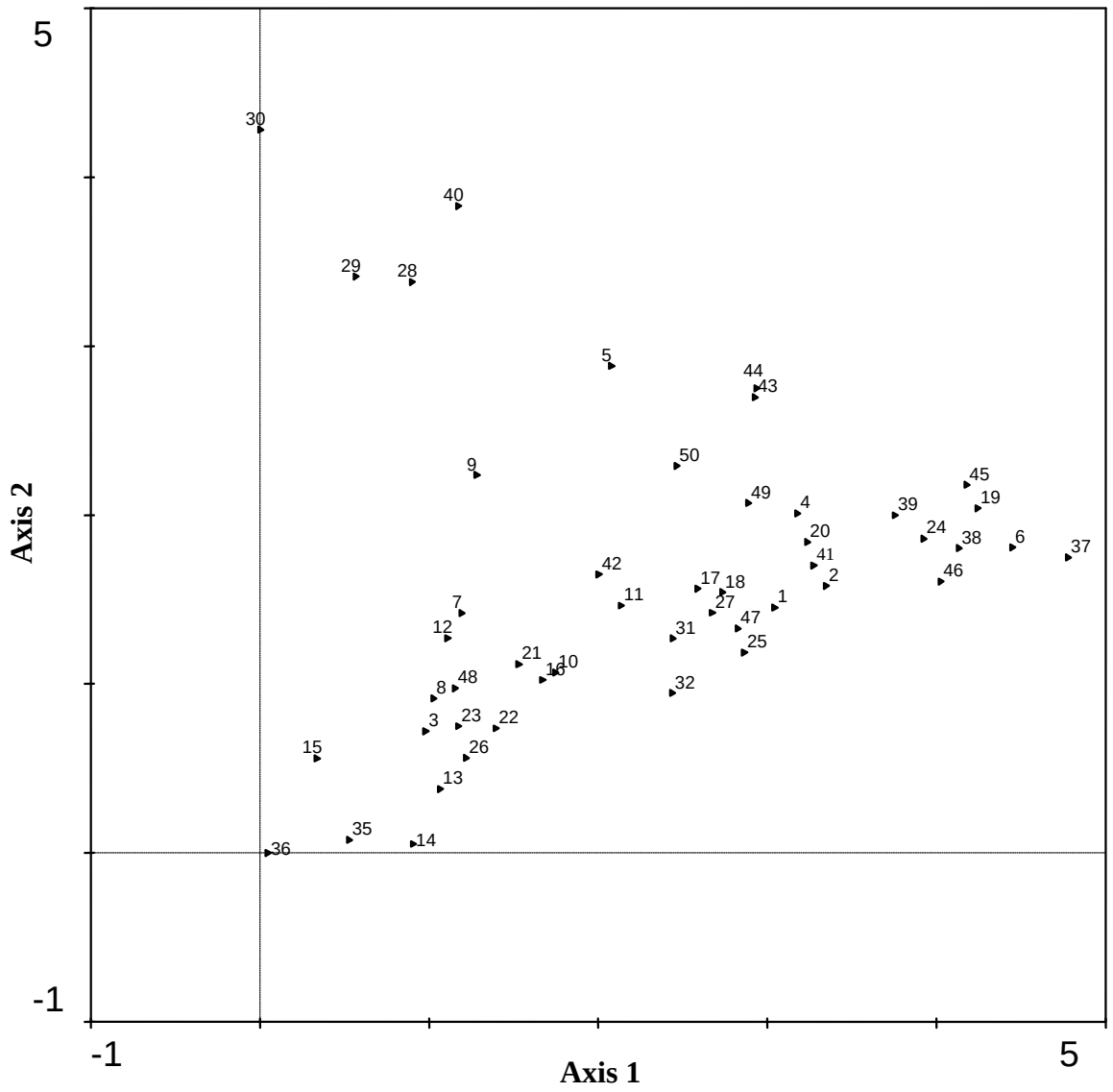


Figure 6.10. DCA scatterplot of site scores on axis 1 and axis 2 (48 sites).

6.2.3.2 CCA of the species dataset with bioavailable metal concentrations

Canonical correspondence analysis (CCA) was undertaken to explore the relationship between diatom taxa and the measured environmental variables in the dataset. The original dataset of 15 environmental variables is modified to include the bioavailable fractions of metals, estimated using the computer program AquaRisk, which contains the geochemical speciation code MODPHRQ. The bioavailable values of the metals are summarised in Appendix 5.

The relative significance and influence of each environmental variable in explaining diatom variation was determined through stepwise forward selection. Monte Carlo permutation tests were used to test each variable included in the CCA and to determine if the "p" threshold point was less than or equal to 0.05. The results of the stepwise forward selection on the dataset, which includes the difference between the percentage variance explained before and after addition, are shown in

Table 6.11. The significant forward selected variables include manganese (Mn), electrical conductivity (EC), copper (Cu) and zinc (Zn), noted in order of selection.

Table 6.11. Percentage of variance explained by variables retained after stepwise forward selection.

Variable	% variance explained before addition	% variance explained after addition	p -value
Mn	0.465	0.465	**0.002**
EC	0.389	0.854	**0.002**
Cu	0.351	1.205	**0.004**
Zn	0.333	1.538	**0.004**

A CCA was executed on the four, forward selected variables, 48 samples and the 193 diatom taxa. A summary of the results are given in Table 6.12. Site RJI5 was excluded from the CCA analysis of its influence (>15x) detected by leverage diagnosis. The first four axes of the CCA explained 15.2% of the total variation in the diatom data. The combination of the variance explained by axis 1 and 2 is 14.5%. These two axes express over half of the constrained and approximately one third of all taxon variance in the dataset (total = 42.8%). Judged by the eigenvalues, the third and fourth axes are least important as their values are lower than axis 1 and 2.

Table 6.12. Summary of statistics for CCA of genus dataset of 49 samples and 4 stepwise forward selected variables.

Axes	1	2	3	4	Total inertia
Eigenvalues:	0.493	0.477	0.355	0.213	10.096
Species-environment correlations:	0.888	0.871	0.891	0.743	
Cumulative percentage variance					
of species data:	4.9	9.6	13.1	15.2	
of species-environment relation:	32	63	86.1	100	
Sum of all canonical eigenvalues:					1.538

The most influential environmental variable identified according to its loading on axis one is zinc. As is evident in Table 6.12, zinc is negatively correlated to axis one (-0.60). The variable with the strongest loading is manganese. This variable is positively correlated to axis 2 (-0.74) as is conductivity (0.62). Copper, however, is most strongly correlated with axis three (0.71) and is negatively correlated with axis 1 (-0.37).

Table 6.13. Environmental variable loadings.

	Axis 1	Axis 2	Axis 3	Axis 4
Cu	-0.37	0.29	0.71	0.20
Mn	-0.36	**0.74**	0.28	0.07
Zn	**-0.60**	0.39	0.06	0.43
EC	0.37	0.62	0.08	0.42

Figure 6.11 shows a CCA biplot of environmental variables and the positioning of sample sites on axis 1 and 2. The vector of a variable and its length is proportional to the strength of correlation with changes in the diatom assemblage. As suggested by the statistical output manganese is the variable most strongly correlated to axis 2 and has the longest vector. This is followed by conductivity, zinc and copper. The small angle between the copper and zinc vectors indicates that these metals are covariant, but the longer zinc vector attests to its greater influence on diatom assemblages.

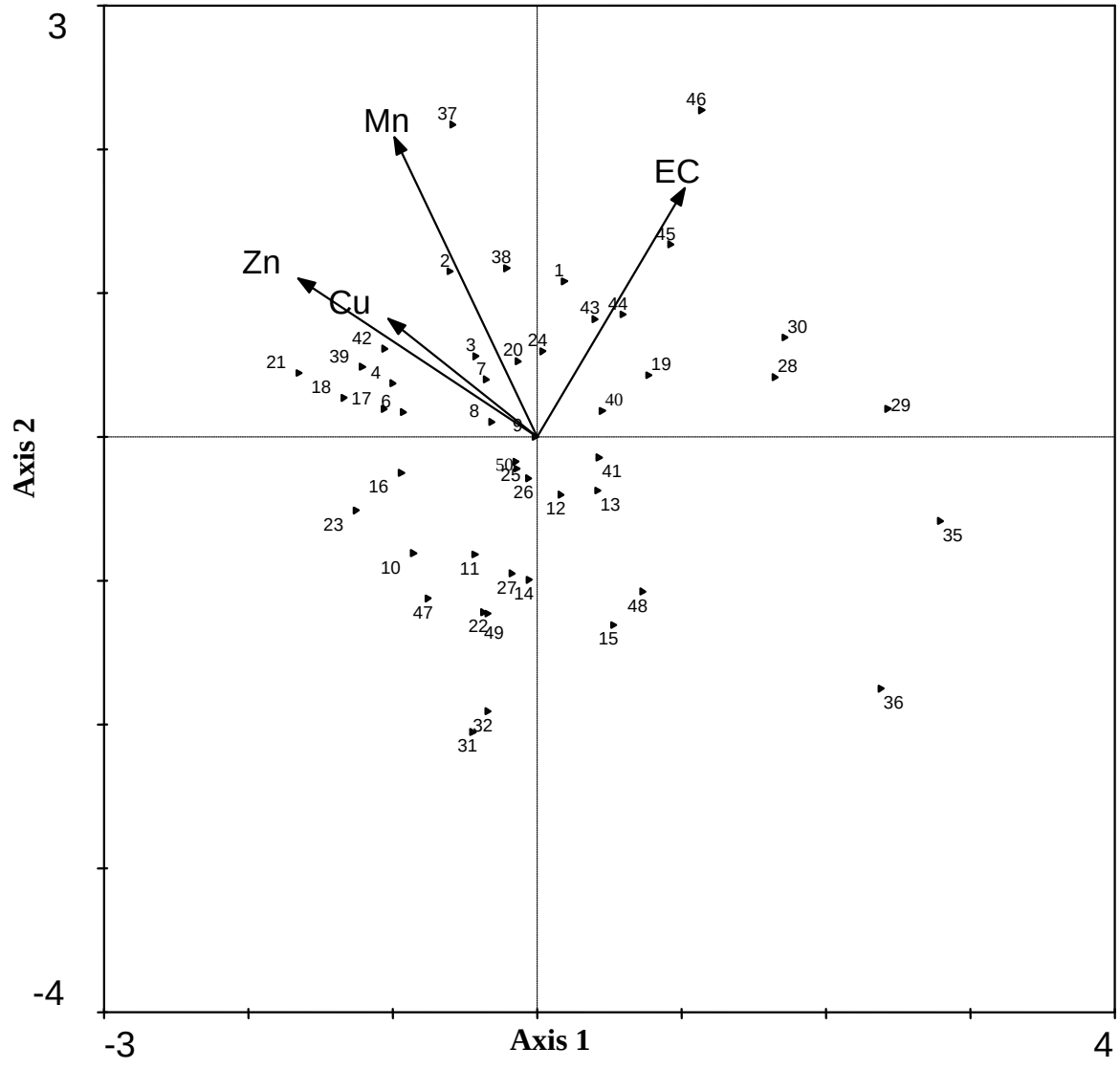


Figure 6.11 CCA biplot of environmental variables and sites on axis 1 and axis 2

When examining the spread of sample sites in the CCA biplot (Figure 6.11), the sites BKI1 (45) and BKI2 (47), plot at the positive end of the conductivity gradient. Sites which appear to be associated with the negative end of the conductivity gradient include the control sites from Redbank, 12MCUSEC (31) and ECUSBRX (32), as well as the Tom's Gully sites TGC1 (22) and TGC2 (23). Sites plotting in the upper left quadrant of the biplot are most influenced by high concentrations of manganese, copper and zinc. Sites CHI1 (37) and RJI2 (2) are most associated with high levels of manganese, and plot at the positive end of this gradient. The sites RJI4 (4) and CHI3 (39) plot at the positive end of the zinc and copper gradient. The position of the Rum Jungle site RJI1 (1) on the biplot seems to indicate that it is associated with both high levels of manganese and conductivity. Sites associated with low metal concentrations include the control sites SCDSCC (35), SCUS12MC (36) and CHC2 (48).

Without mentioning species with low relative abundances, *Pinnularia joculata* (Pi.jo) *Nitzschia palea* v. *palea* (Ni.pvp) and *Nitzschia nana* (Ni.na) are most associated with high measurements of conductivity. These species plot at the positive end of this gradient in the upper right quadrant of the CCA biplot (figure 6.12). At the negative end of this gradient plot *Achnanthidium minutissimum* v. *minutissimum* (Ac.mvm), *Gomphonema gracilis* (Go.gr) and *Encyonema silesiacum* (En.si). Associated with high levels of manganese is *Nitzschia vasta*. This taxon has a relative abundance of greater than 10% in sites RJI2 (2), RJI6 (6), HCUSEC (24), CHI1 (37) and CHI2 (38), all of which are situated in the upper left quadrant. The key species associated with low concentrations of manganese are *Achnanthidium minutissimum* v. *exilis* (Ac.mve), *Planothidium frequentissimum* (Pl.fr) and *Cocconeis placentula* (Co.pl). However, each of these is only present at one site at over 10% relative abundance. *Pinnularia subcapitata* (Pi.su) and *Eunotia naegelii* (Eu.na) are associated with waters of high copper and zinc concentrations while *Nitzschia liebetruthii* (Ni.lie) and *Planothidium rostratum* (Pl.ro) are associated with low concentrations of these metals. *Nitzschia paleaeformis* (Ni.pa) plots between the manganese and conductivity vectors, close to sites RJI1 (1) and CHI1 (38). *Nitzschia paleaeformis* reaches over 40% abundance at these sites.

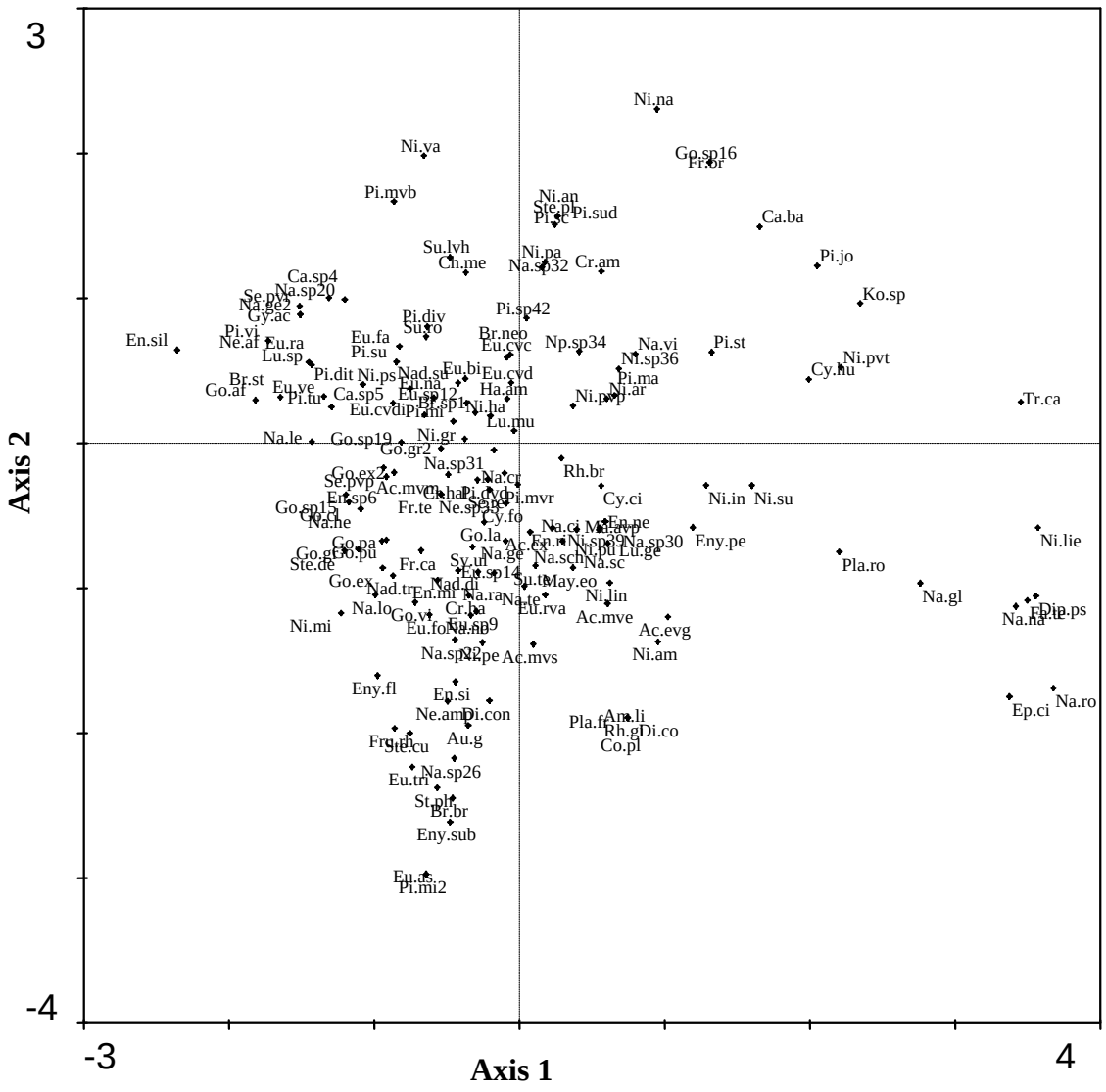


Figure 6.12. CCA scatter plot of species scores on axis 1 axis 2.

6.2.3.3 Gradient analysis of selected environmental variables

A series of detrended canonical correspondence analyses (DCCA), based solely on each CCA stepwise selected variable, was undertaken to verify their potential predictability. This was verified by testing the first axis of each constrained DCCA. The first axis of a DCCA constrained to copper had a gradient length of 2.5 standard deviations (S.D.), indicating unimodel-based techniques (weighted averaging) were appropriate to generate copper inferences. Each of the following three environmental variables had gradient lengths of greater the two standard deviations: manganese (3.5), conductivity (2.9) and zinc (2.7).

As the gradient lengths of each of the four variables were greater than two standard deviations, unimodal inference models were produced. The inference ability of transfer functions can be examined by comparison of actual values recorded at each site against those predicted by weighted

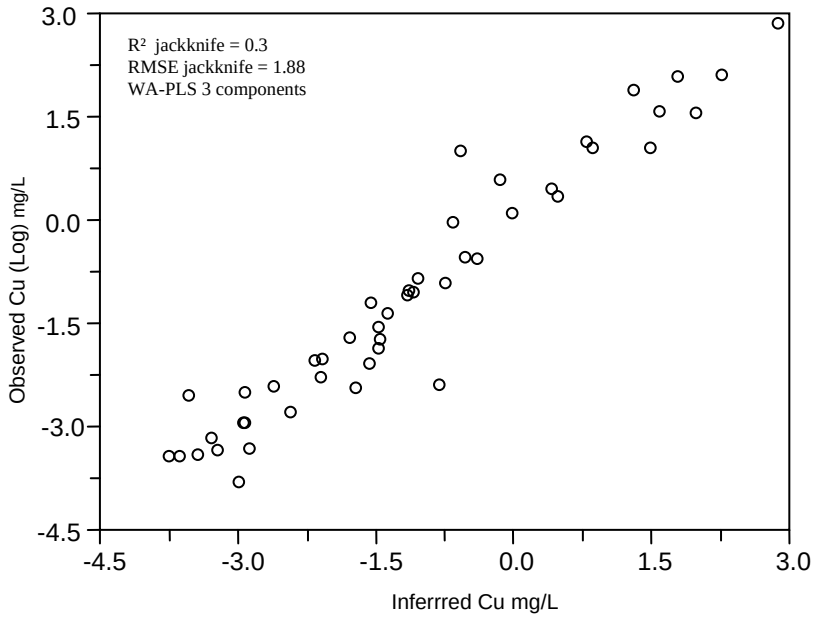
averaging (in the program CALIBRATE) of species optima. The sample CHC2 was identified as an outlier. This sample decreased the performance of the model for the environmental variable copper. Site CHC2 was removed from weighted average partial least squares (WA-PLS) analysis of the environmental variable as it was an outlier. This site is dominated by (>70% relative abundance) *Achnanthidium minutissimum* v. *minutissimum*. While this species is not rare or absent at other sites, site CHC2 was removed as it improved the overall performance in terms of RMSE and R^2 values.

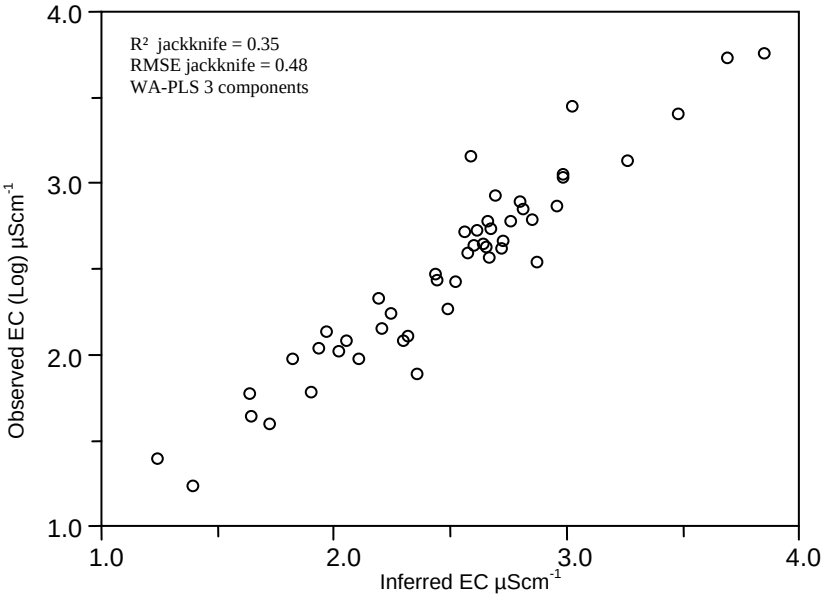
The weighted averaging partial least squares model of environmental variables with species data showed that inferences between diatoms and manganese were significant and strong, indicating high predictive capacity ($R^2_{jackknife}$ >0.35 and R^2apparent >0.6). The results of the models are summarised in Table 6.14. The apparent R^2 for manganese is 0.96 ($R^2_{jackknife}$ = 0.38, RMSE = 23). However, for the two variables copper and electrical conductivity, the models have weaker predictive power as the $R^2_{jackknife}$ values for each of the variables is equal to, or less than, 0.35. These two variables have apparent R^2 values above 0.90. Although the model for zinc has a high apparent R^2 (0.96), overall it has the least predictive ability as the $R^2_{jackknife}$ is just 0.1. The WA-PLS model plots of inferred versus observed values for each of the four variables are shown in Figure 6.13. Each of the relationships shows an exponential regression indicating that, with increases in observed values, there are corresponding increases in inferred values. Additionally, the strong relationship is further indicated by the tight cluster of the sites along the line of regression.

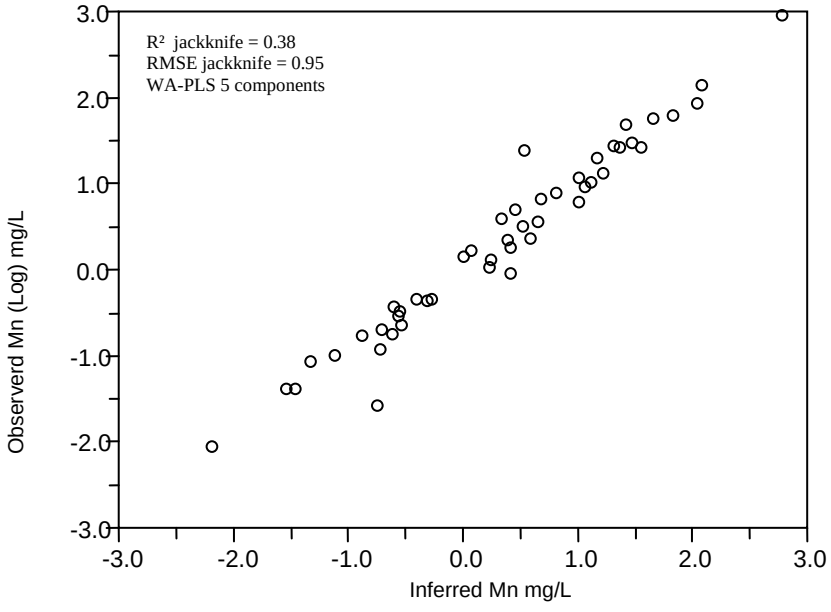
Table 6.14. Results of the WA-PLS for the bioavailable dataset (n=48).

	Cu	EC	Mn	Zn
WA-PLS components for apparent	3	3	5	5
Apparent R^2	0.93	0.91	0.96	0.96
R^2 jackknife	0.3	0.35	0.38	0.1
Apparent RMSE	0.48	0.17	0.23	0.24
Apparent jackknife	1.88	0.48	0.95	1.77
Number of samples	*47	48	48	48

*site CHC2 was excluded


R² jackknife = 0.3
RMSE jackknife = 1.88
WA-PLS 3 components
Observed Cu (Log) mg/L
Inferrred Cu mg/L
3.0
1.5
0.0
-1.5
-3.0
-4.5
-4.5
-3.0
-1.5
0.0
1.5
3.0


R² jackknife = 0.35
RMSE jackknife = 0.48
WA-PLS 3 components
Observed EC (Log) µScm-1
Inferred EC µScm-1
4.0
3.0
2.0
1.0
1.0
2.0
3.0
4.0


R² jackknife = 0.38
RMSE jackknife = 0.95
WA-PLS 5 components
Observerd Mn (Log) mg/L
Inferred Mn mg/L
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0

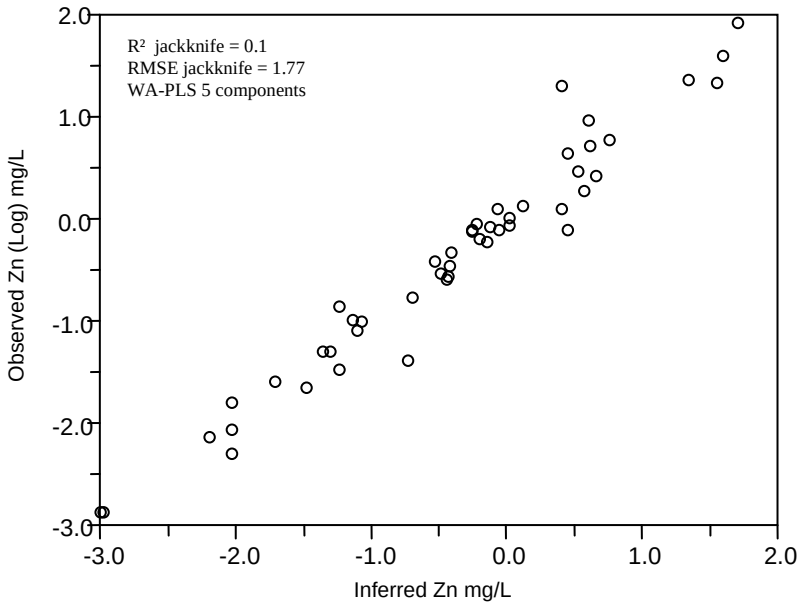


Figure 6.13. Plots of observed versus inferred concentrations of Cu, EC, Mn and Zn based on WA-PLS models.

6.2.3.4 Discussion

The DCA of the reduced dataset of 48 sites explained less variance on axis 4 (21.1%) compared to the full species dataset of 50 sites (23.1%). This reduction in variability may be due to the removal of the most variable sites. However, the general positioning and relationships of species and samples changed little. This finding is repeated for the CCA. The CCA of the dataset incorporating bioavailable metal concentrations explained less variance on axis 4 (15.2%) than the species dataset which incorporated total metal concentrations (18.7%). Additionally, the amount of total variation that can be explained by the overall inertia of the species dataset with total metals compared to that of the bioavailable metals dataset is very similar. The bioavailable environmental variables (canonical eigenvalues) explained only a small proportion of the variation in species composition (1.54) compared to that of the species dataset with total metal concentrations (2.52). The division of the explained variance by the total variance, results in the measured variables explaining 15.26% of variance compared to 25.12% of variance for the species dataset with total metal concentrations. Evidently, the species dataset utilising total metal concentration better explains species variation and the separation of sites.

For the bioavailable fraction of the metals dataset, the environmental variables manganese and electrical conductivity were the most important variables in explaining species variance. In comparison, CCA of the species dataset, with the total metal concentrations, determined pH to be the most important variable on axis 1 followed by aluminium. The most evident difference between the rankings of variables is that the metal variables, rather than pH, are most important in explaining diatom variance in the bioavailable dataset.

From the CCA plot it appears that associated with high concentrations of bioavailable magnesium is *Nitzschia vasta.* Associated with high concentrations of conductivity are *Nitzschia nana* and *Pinnularia joculata*. All three of these species are also associated with impacted sites of the species dataset utilising total metal concentrations. Similarly, the taxa *Achnanthidium minutissimum v. minutissimum* is associated with low measurements of electrical conductivity in both datasets. The use of these taxa as indicator species is described in a later section.

Weighted averaging partial least squares (WA-PLS) regression was carried out on the environmental variables, copper, conductivity, zinc and manganese. However, when utilising the bioavailable concentrations of metals, rather than the total metal concentrations, the WA-PLS models were less successful. For example, only the manganese inference model had a high predictive capacity. The inference models produced for copper, conductivity and zinc each displayed low apparent R^2 or $R^2_{\text{jackknife}}$ values indicating their low predictive capacity. This compares to the five models produced for the species dataset, utilising total metal concentrations, where the relationship between inferred and observed values were each strong and significant. The low performance of some models, in terms of R^2, may have been caused by the uneven spread of values across the ranges of the environmental variables. Additional sampling to fill gaps in the environmental data is likely to strengthen the predictive capacity of the models.

Although the use of bioavailable fractions of metal concentrations should theoretically provide more realistic results of species and environmental variable relationships there was no additional explanation of species variance and no significant inference models could be produced. Unfortunately, there are no similar studies which use bioavailable fractions with which to compare the outcomes.

6.3 Diversity Indices

Diversity indices are often used as monitoring tools as the concept of diversity is closely related to the nature of species-abundance distributions. Diversity indices have been found useful in characterising the biotic responses of aquatic communities to environmental disturbances (Khan 1991; Townsend & Riley 1999). However, there is still discussion of whether diversity indices can be used as indicators of water quality (Lobo & Kobayasi 1990), especially in waters impacted by heavy metals. A diversity index must take into account some statistical requisites such as independence of sample size and technique and emphasise the relative importance of each species rather than the presence or absence of species (Pielou 1975). Diversity is not only the number of species (richness), and the number of individuals, but also is the evenness of distribution (the property of a community that relates to the relative frequency of the species). Thus, the diversity is the result of the interaction between these basal indicators of the community structure. Two methods of determining species diversity are the Shannon-Wiener diversity index and Simpson's index of diversity. Margalef (1956) initiated the use of the Shannon function to measure species diversity, which integrates species richness and evenness into one index of diversity.

This section evaluates the manner in which acid mine drainage, represented by pH and copper, affects species diversity and richness of benthic diatom communities on natural substrata. The section particularly aimed at determining if changes in diversity and species richness can infer heavy metal pollution and, possibly, be used to monitor acid mine drainage.

6.3.1 Diversity indexes and species richness

The Shannon-Wiener (H) and Simpson's (D) indices, and species richness, for the selected mines, are presented in Table 6.15 along with their corresponding pH and copper measurements. According to the results of the detrended canonical analysis (DCA) of the dataset in section 6.2.1 there is a clear separation of 'low impact' and 'high impact' sites. Fourteen sites were determined to be very impacted. The Shannon-Wiener and Simpson's diversity values for the benthic diatom communities of very impacted sites range from 0.2 (CHI1) to 3.3 (RJI2) and 0.01(CHI1) to 0.95 (RJI1) respectively. The pH values for these sites are very acidic, ranging from 3.0 (CHI1) to pH 5.5 (CHI5). The copper measurements range from 97.8 mg/L at the Redbank site HCUSEC to 0.02 mg/L at the Cosmo Howley site BKI2. TGI4 is classified as very impacted although it had a pH of 7.6 and low copper (0.009 mg/L). Its classification as 'very impacted' is derived from its high magnesium, cadmium and potassium concentrations. The sites classified as 'impacted' have diversity ranges of 0.3 (RJI5) to 3.1 (12MCDSEC) for the Shannon-Wiener index and 0.1 (RJI5) to 1.0 for Simpson's index. The control sites have a Simpson's diversity range between 0.37 (TGC1)

and 0.93 (RJC1) and Shannon-Weiner diversity range between 1.0 (TGC1) to 3.2 (RJC3). These sites are characterised by pH values above pH 6.0 and copper concentrations of less than 0.3 mg/L.

Table 6.15. Diversity index and pH, copper values for each site.

Site	Degree of impact	Simpson's diversity index	Shannon-Wiener diversity index + (-1)	Species richness	pH	Cu (mg/L)
RJI1	v.impacted	0.95	3.24	45	3.42	5.99
RJI2	v.impacted	0.94	3.27	55	4.04	11.5
RJI3	v.impacted	0.13	0.33	7	3.65	1.22
RJI4	v.impacted	0.94	3.22	44	4.42	1.43
RJI5	Impacted	0.10	0.30	10	6.2	0.19
RJI6	Impacted	0.38	0.67	6	6.31	0.6
RJI7	Impacted	0.50	0.99	10	7.27	0.045
RJI8	Impacted	0.26	0.59	8	7.03	0.053
RJI9	Impacted	0.52	1.01	6	7.3	0.022
RJI10	Impacted	0.92	3.05	54	6.97	0.003
RJC1	Control	0.93	3.14	58	7.21	0.011
RJC2	Control	0.54	1.33	21	7.37	0.013
RJC3	Control	0.93	3.17	57	7.39	0.002
RJC4	Control	0.81	2.30	42	7.57	0.018
RJC5	Control	0.87	2.70	44	8.19	0.003
TGI1	Impacted	0.74	2.12	32	6.34	0.005
TGI2	Impacted	0.93	3.07	41	6.44	0.005
TGI3	Impacted	0.92	2.84	34	6.11	0.011
TGI4	v.impacted	0.67	1.43	14	7.58	0.009
TGI5	v.impacted	0.78	2.33	37	4.9	0.11
TGI6	Impacted	0.68	2.02	31	6.81	0.002
TGC2	Control	0.76	2.05	31	6.98	0.007
TGC1	Control	0.37	1.03	22	7.05	0.014
HCUSEC	v.impacted	0.87	2.61	33	3.56	97.8
ECDSHC	v.impacted	0.92	2.61	26	4.35	14.6
ECUS12MC	Impacted	0.92	3.04	50	4.47	14.9
12MCDSEC	Impacted	0.95	3.11	54	5.45	1.41

Site	Degree of impact	Simpson's diversity index	Shannon-Wiener diversity index + (-1)	Species richness	pH	Cu (mg/L)
12MCUSSC	Impacted	0.79	2.06	26	8.43	0.231
SCDS12MC	Impacted	0.76	2.05	26	8.7	0.109
SC@BRX	Impacted	0.52	0.78	6	7.56	0.0458
12MCUSEC	Control	0.92	2.84	33	6.89	0.0405
ECUSBRX	Control	0.84	2.19	26	5.99	0.281
CCUSSC	Control	0.63	1.73	26	8.2	0.0101
SCUSCC	Control	0.87	2.65	39	8.4	0.00193
SCDSCC	Control	0.87	2.44	20	8.37	0.106
SCUS12MC	Control	0.81	2.28	27	8.43	0.00084
CHI1	v.impacted	0.01	0.02	2	2.95	12.2
CHI2	v.impacted	0.73	2.01	27	4.05	0.157
CHI3	v.impacted	0.89	2.73	35	6.94	0.01
CHI4	Impacted	0.47	0.75	6	6.95	0.002
CHI5	v.impacted	0.87	2.65	45	5.5	0.231
CHI6	Impacted	0.91	2.84	36	7.15	0.003
CHI7	Impacted	0.73	1.78	18	7.28	0.005
CHI8	Impacted	0.79	2.09	25	6.99	0.004
BKI1	v.impacted	0.47	1.13	9	4.85	0.336
BKI2	v.impacted	0.66	1.49	18	4.33	0.016
CHC1	Control	0.88	2.58	34	7.32	0.002
CHC2	Control	0.49	1.38	27	7.72	0.0005
CHC3	Control	0.93	2.95	42	7.23	0.002
CHC4	Control	0.92	2.99	40	7.45	0.004

Species richness varied from 20 (SCDSCC) to 58 species (RJC1) in undisturbed stream sites (control sites) and from nine (BKI1) to 55 species (RJI2) in sites of very high impact. The sites RJI3, RJI5 and CHI1 have distinct diatom communities displaying dominance of >80% of one species and, consequently, have extremely low diversity indices and species richness values.

6.3.2 Relationships between indices and environmental variables

The accuracy of each of the diversity indexes and species richness as an indicator of water quality was evaluated by comparison with the pH and copper measurements obtained for each site. The variables pH and copper were chosen to represent heavy metal pollution at the sites based on their

AquaRisk ranking and importance in the CCA. The diversity indices for the two environmental variables were correlated with three datasets to determine which dataset provided the strongest relationship. The first dataset was composed of the entire 50 sample sites, while the second dataset excluded any outlier sites. The third dataset only included those sites which were determined to be very polluted by DCA analysis of the species dataset. These sites are indicated in Table 6.15 as "v.impacted". The significance of each correlation is assessed based on the R square (R^2), root mean square error (RMSE) and significance (p) values.

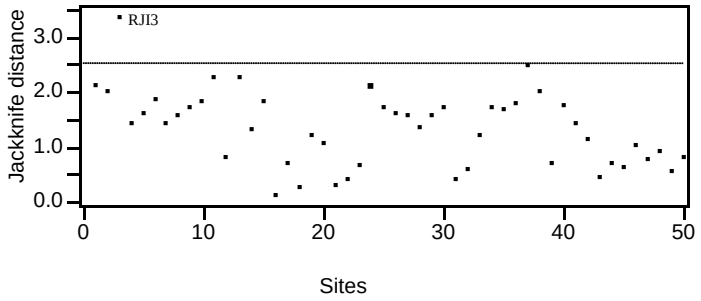


Figure 6.14. Sample outlier analysis of jackknife distances. Site numbers are shown in Appendix 1.

Outliers were determined by creating a distance plot using jackknifed distances in the computer program JMP (version 5.1) based on species richness. As an example, the bivariate analysis of species richness and pH (Figure 6.14) has one outlier, site RJI3 (2.98). The Figure 6.14 distance plot illustrates that site RJI3 is statistically an outlying sample as it plots above the cut-off distance of 2.5. Any sample with a jackknife distance of above 2.5 was excluded from diversity analysis of the second dataset. As already mentioned, site RJI3 is dominated by one particular species, *Achnanthidium minutissimum* v. *minutissmum,* which has a relative abundance of over 93%.

Table 6.16. Summary statistics for regressions of species richness and diversity indices for pH.

	Species richness			Simpson's index			Shannon-Wiener index		
	All sites	Excluding outliers	v.polluted sites	All sites	Excluding outliers	v.polluted sites	All sites	Excluding outliers	v.polluted sites
R^2	0.12	0.19	0.16	0.013	0.025	0.087	0.000	0.047	0.02
RMSE	0.106	13.84	13.79	0.245	0.207	0.302	0.920	0.827	1.09
p	0.12	0.002	0.17	0.423	0.283	0.306	0.959	0.140	0.62

The largest percentage variance that can be explained for pH and species richness, for any dataset, is displayed by the dataset which excludes outliers. This model explains 19.1% ($R^2 = 0.191$) of variance (Table 6.16). Additionally this is the only dataset which has a significant correlation ($p <$

0.05) between species richness and pH (0.002). The other models explain 16.2% (polluted) and 12.4% (all sites) of variations. This correlation is illustrated in Figure 6.15.

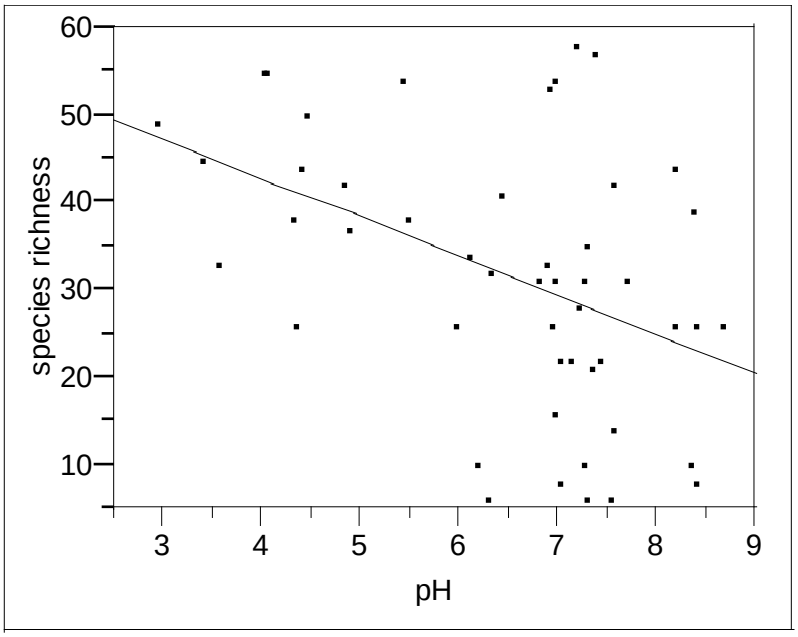


Figure 6.15. Bivariate fit of species richness and pH excluding outlier samples (RJI3).

There is a significant negative correlation between species richness and pH ($p = 0.002$). However, there is only one site that plots below pH 3.5 indicating that there is low representation of sites below this point. Additionally, the scatter of sites around the line of linear fit is very broad. For instance, between pH 6.0 and 7.0 there are over eight sites with species richness ranges from 2 to 6.

The relationship between Simpson's diversity index and pH is not significant ($p > 0.05$) for any of the datasets (Table 6.16). Two outlier samples were identified, CHI1 (10.4 jackknife distance) and RJI3 (3.4 jackknife distance), and were excluded from the outlier dataset. The relationship of the Simpson's diversity index and pH for this model had the highest significance value (p), 0.28, and the second largest percentage of variance explained (2.5%). However, the variance explained increased slightly (to 5.2%) when the dataset was reduced to include only sites classed as 'v.impacted'. Figure 6.16 shows the bivariate fit of Simpson's diversity index and pH using the polluted sites dataset. This indicates that there is an increase in diversity with increased acidity. This is however, not statistically significant.

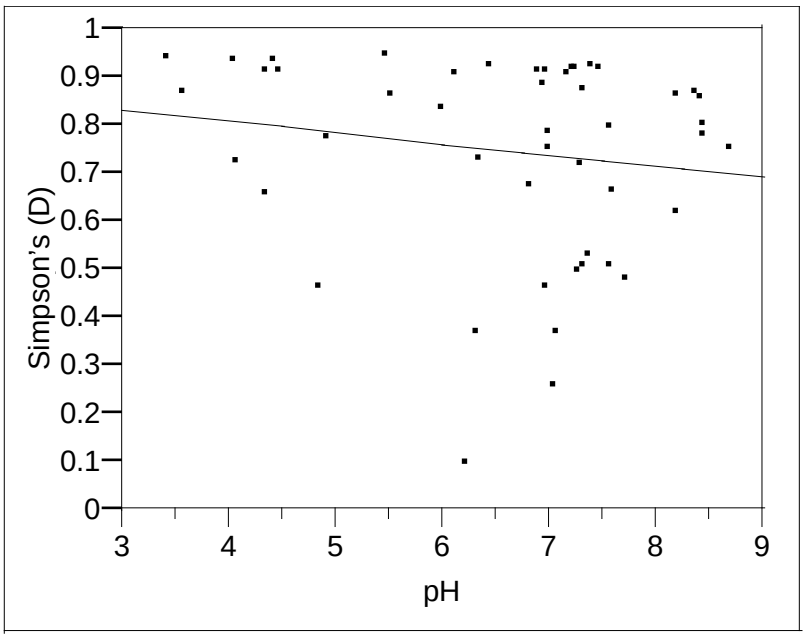


Figure 6.16. Bivariate fit of Simpson's index and pH (excluding outliers).

Table 6.16 details the statistical outcomes for the analysis of the relationship between Shannon-Wiener diversity index and pH for each of the datasets. CHI1 (8.3) and RJI1 (2.51) were both excluded from the bivariate analysis as they both have a jackknife distance above 2.5. The relationship for each of the datasets explains less than 5% of species variance and the probability factor is not significant ($p > 0.05$). The analysis with the dataset excluding outliers produces the strongest results with the highest variance ($R^2 = 0.047$) and highest significance ($p = 0.14$). The general relationship between the Shannon-Wiener diversity index and pH is illustrated in Figure 6.17. With an increase in pH there is a slight increase in diversity. This relationship is not significant ($p = 0.14$).

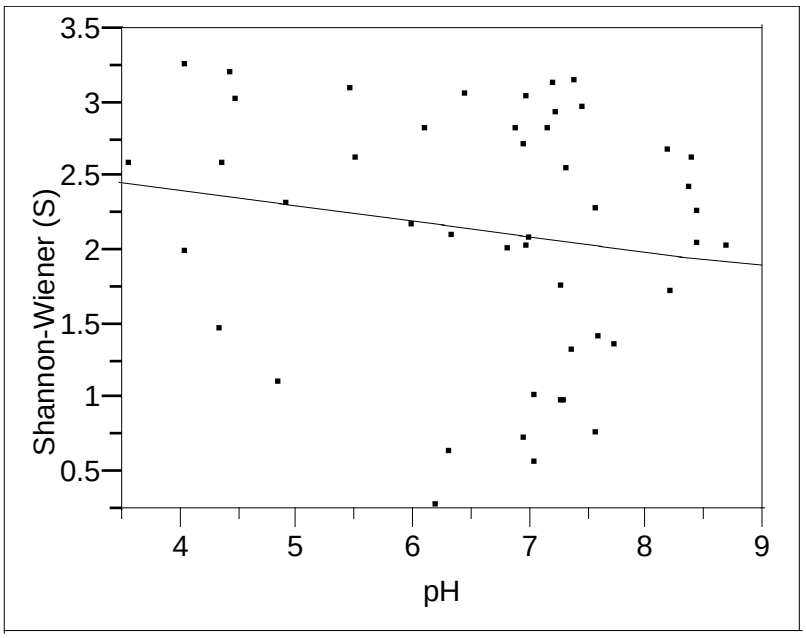


Figure 6.17. Bivariate fit of Shannon-Wiener index and pH (excluding outliers).

Table 6.17 shows the R^2, RMSE and probability values for each dataset for the relationship between the diversity indices and copper (mg/L). The strongest relationship for any of the diversity indices and copper is for the Shannon-Wiener diversity index and the dataset incorporating only the most polluted sites. The percentage of variance explained is 6.1% ($R^2 = 0.061$). However, this relationship is not significant. The most significant relationship is between species richness and copper using the full site dataset however, the relationship is 0.33 (p) indicating that it is not statistically significant. The regression for species richness and copper is illustrated in Figure 6.18. With an increase in copper there is an increase in species richness per site. For species richness, site HCUSEC was defined as an outlier by multivariate outlier analysis. It had a jackknife distance of 2.82 and was consequently excluded from bivariate analysis. For this dataset, there is no relationship between species richness and copper ($p = 3.24$) and only 2.1% ($R^2 = 0.021$) of variance is explained.

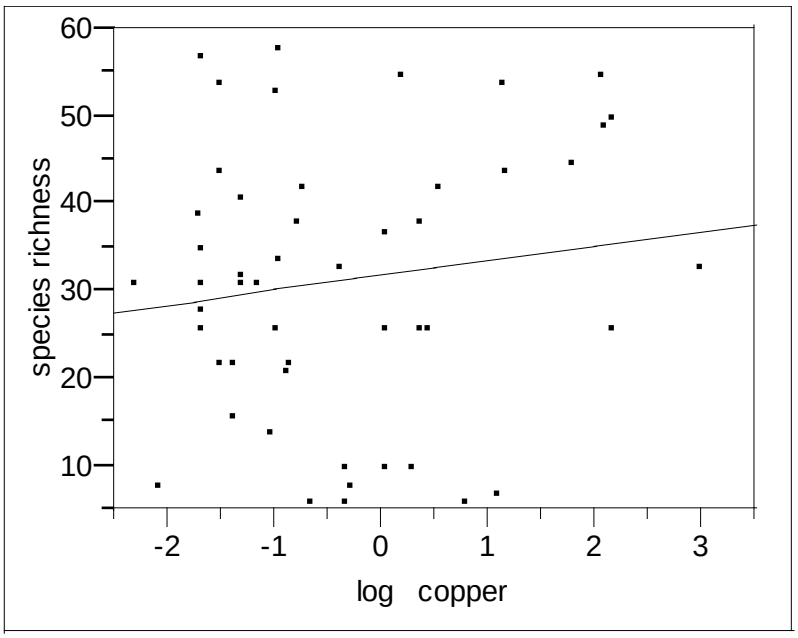


Figure 6.18. Bivariate fit of species richness and log copper (mg/L) (with all sites).

Table 6.17. Regressions of species richness and diversity indices for copper mg/L (log).

	Species richness			Simpson's index			Shannon-Wiener index		
	All sites	Excluding outliers	v.polluted sites	All sites	Excluding outliers	v.polluted sites	All sites	Excluding outliers	v.polluted sites
R^2	0.020	0.021	0.0974	0.009	0.000	0.002	0.001	0.002	0.012
RMSE	15.485	15.637	14.279	0.246	0.228	0.316	0.920	0.885	1.095
P	0.330	3.248	0.277	0.514	0.925	0.886	0.861	0.784	0.709

For the determination of the correlation between the Simpson's diversity index and copper, site CHI1 is again defined as an outlier (10.4 jackknife distance), as is site HCUSEC (3.12 jackknife distance). For each of the regressions involving the Simpson's diversity index, each of the datasets explain less than 1% of the variance ($R^2 < 0.01$) and have a significance of greater than 0.5 ($p > 0.5$). Consequently, none of the datasets produce a significant correlation between the diversity index and the environmental variable copper. The dataset incorporating all sites had the highest R^2 (0.009) and highest significance ($p = 0.51$) and, as illustrated in Figure 6.19, with an increase in copper there is a decrease in diversity.

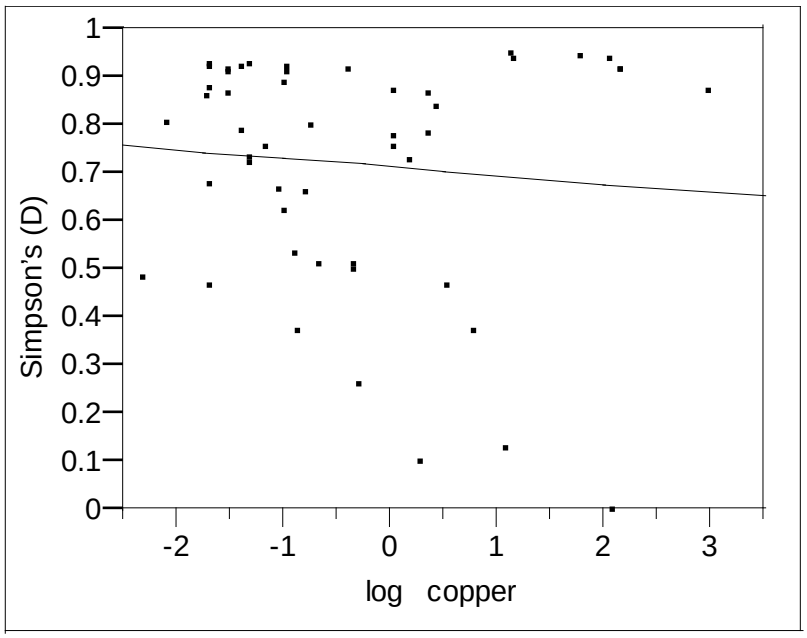


Figure 6.19. Bivariate fit of Simpson's index and copper (log mg/L) (with all sites).

The percentage variance explained by the relationship between the Shannon-Wiener diversity index and copper was highest for the dataset comprising the very polluted sites. For the outlier reduced dataset, sites CHI1 (jackknife distance = 8.2) and HCUSEC (jackknife distance = 3.1) were excluded from analysis. This dataset explains 6.1% of variance and has a probability factor of 0.42 which indicates that the relationship is not significant. However, depending on the dataset used, with an increase in copper concentration there is a slight increase (very polluted sites & outliers) or a decrease (all sites) in diversity.

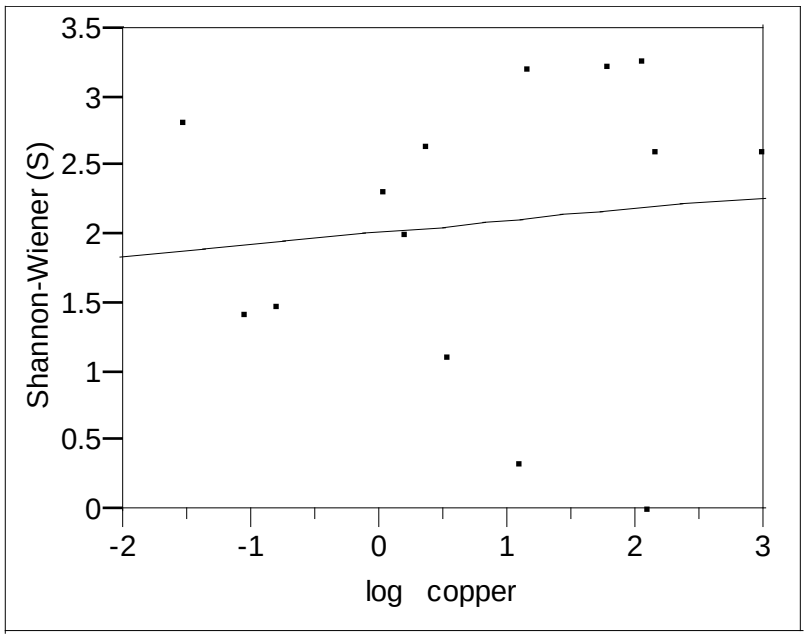


Figure 6.20. Bivariate fit of Shannon-Wiener index and copper (log mg/L) (v.polluted sites).

There is an intrinsic positive correlation between the Shannon-Wiener and Simpson's diversity indexes which is illustrated with bivariate analysis for each of the datasets. The relationship between the indices and the dataset utilising all sites has an R^2 value of 0.91 and a RMSE value of 0.26. The dataset, which includes only the 14 sites which are classed as very polluted, has a slightly higher R^2 (0.92) value and a lower RMSE (0.09) value indicating that it performs better. However, both datasets are significant ($p < 0.05$) with p values of less than 0.0001. The correlation between the two diversity indices is illustrated in Figure 6.21. There is a positive a linear fit of the two indices indicating that, as the Simpson's diversity measure of sites increases, so does the Shannon-Wiener diversity measure of sites.

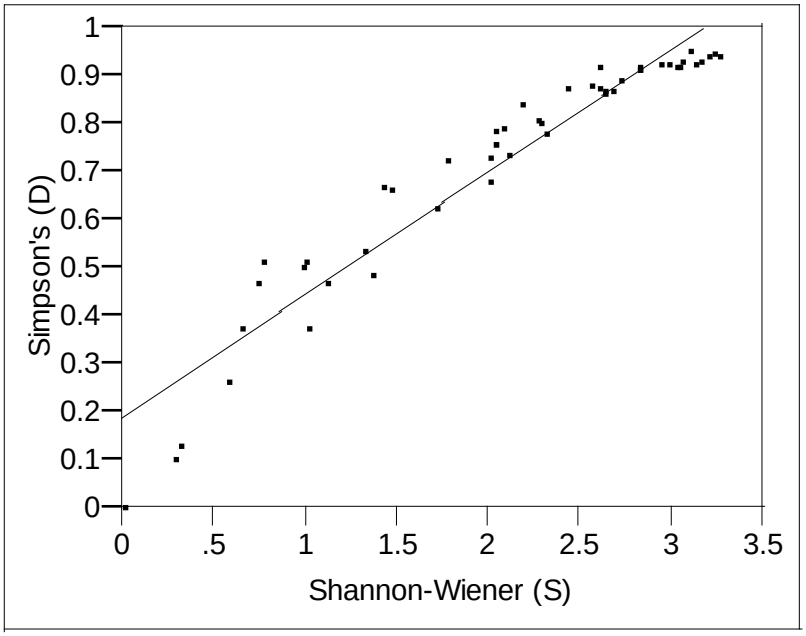


Figure 6.21. Bivariate fit of Shannon-Wiener and Simpson's diversity indexes.

6.3.3 Discussion

The sites classified as very impacted generally had high Simpson's (D) diversity values, close to 1, and high Shannon-Wiener (S) values. This indicates that these sites are characterised by high diversity. Maznah *et al.* (2002) similarly found the highest diversity and species richness at sites with highest pollution levels. However, at the most impacted sites, for instance, RJI3, RJI5, CHI1, and site BKI2, diversity values were low and the assemblages were dominated by one or two species. These species had relative abundances totalling over 80%. Consequently these sites have very low S and D values compared to the other sites within the very impacted group. The species abundance at these very impacted sites also meant that this group had a very large range of species

richness, from low to very high. Similarly the control sites also had generally high diversity values with only one site (TGC1) having a low diversity value as it was dominated by one species. The control sites however, had high species richness values. The site category with the highest variance in diversity values was the polluted group. This group had diversity values which ranged from the low to the high end of the diversity scale.

The only index which displayed a significant relationship with pH was species richness. This index was also able to explain the highest percentage of species variance. The relationship showed that, with an increase in acidity (pH), there is an increase in species richness. However, this relationship was only significant for the dataset which excluded outlier sites. This finding is emphasised again by the relationship displayed between species richness and copper. With increases in copper concentration there is an increase in species richness so that, in environments with high copper and high acidity, there is greater species richness. This result however, is not statistically significant for any of the datasets and they each explain little of the variance. Additionally, the fit of the line for the correlation tends towards horizontal. However, in contrast to these results, both Gomez (1998) and Fore *et al.* (2002) found that species richness declined at mining impacted sites.

Although low species diversity is often associated with stressful conditions (Sullivan 1986; Rott & Pfister 1988), such as heavy metal pollution (Lampkin & Sommerfeld 1982; Dickman *et al.* 1990), neither of the diversity indices used in this analysis could be significantly related to the environmental variables pH or copper. The fit of both these regressions is not strongly linear and the probability factors are not significant nor is the amount of variance explained. This finding corresponds with the conclusion of Archibald (1972) who found that no diversity index could relate diatoms to water quality in South African rivers. Other environmental diatomists have reported similar lack of association (van Dam 1982, Lobo & Kobayasi 1990). Nevertheless, John (1993) found that the Shannon-Wiener diversity index was a reasonable measure of acidity and that, with increased pH, there was increased diversity (0.28 H value for sites with pH 2-4). However, John did not test the significance of this correlation. Similarly, Jüttner *et al.* (2003) found that, in Kathmandu Valley streams, richness and diversity increased significantly with K, Cl, SO_4 and NO_3, but declined significantly with increasing metals such as Al, Fe, Mg and with acidity. In contrast, Stevenson (1984) stated that the Shannon-Wiener diversity index of diatom communities will often decrease with nutrient enrichment and increase with toxic impacts.

Although none of the relationships between the diversity indices and variables are significant or linear there is however, a difference in their statistical outcomes. This difference in statistical outcome is dependent on the dataset used. For instance, the relationship between pH and the Shannon-Wiener index was statistically different depending on the dataset used. The dataset excluding the outliers, although not significant, had the strongest relationship for each pH-diversity

index. However, for the relationships between diversity indices and copper there was no statistically significant result for any dataset.

The correlation between the two diversity indices, the Shannon-Wiener and Simpson's index, is strong and significant. This suggests that the relationship between Shannon-Wiener and Simpson's indexes can be equally described. According to Stevenson (1984), most measures of species diversity which incorporate evenness and richness components of diversity into one index of diversity are highly correlated. By incorporating evenness and richness into one index, species diversity could remain the same if both evenness and richness changed in opposite directions. It seems then that only one index is necessary to measure diversity.

Although there are certainly environmental factors affecting species diversity and richness, the results of this analysis are not conclusive, due to the lack of statistical significance, and also the lack of correlation with other studies. Fore *et al.* (2002), in a study of mine impacted Idaho rivers, found that, although the total number of diatom taxa declined at mining sites, they increased with other types of disturbances such as agriculture. Other studies have found that the number of taxa increased for moderate disturbance and declined with extreme urbanisation or metal contamination from mining (Jüttner *et al.* 1996; Medley & Clements, 1998; Stewar*t et al.* 1999; Genter & Lehman, 2000; Verb & Vis, 2000; Sonneman *et al.* 2001). However, the findings by Chessman *et al.* (1999) suggest that the total number of species only declines at intense levels of disturbance. Consequently, studies have concluded that the relationship between diversity and environmental quality is more complex than previously thought (Podani 1992). One reason for this is that there are too many environmental factors, other than water pollution (pH and copper), that affect species diversity of communities. Additionally, the incorporation of species richness and evenness into a single index of diversity, and the different types of responses that communities manifest with respect to different types of pollutants, is a problem. Ho and Peng (1997) stated that a higher diversity value (S) is not necessarily better than a lower one, nor does it follow that high diversity indices can be interpreted as being a reflection of high quality habitat. Archibald (1972) stated that, to accurately estimate water quality using species diversity, it is necessary to define precisely the species within the community and to have thorough knowledge of their autecology.

One problem with the analysis is that each of the regressions is calculated without some of the most extremely heavy metal polluted sites. Sites CHI1 from Cosmo Howley mine, HCUSEC from Redbank mine and RJI1 and RJI3 from Rum Jungle mine are repeatedly excluded from bivariate analysis and labelled as outliers due to high jackknife distances. These sites represent some of the most highly polluted sites along the pollution gradient in terms of pH (<4) and heavy metal concentrations (Cu >1 mg/L), but have low species richness and diversity values compared to other sites within this pollution category. This indicates that there is a problem with the spread of the data

and possibly some other form of analysis is required. For example, there were not enough samples representing low pH and high copper and the change in pH and copper concentration between the site closest to the pollution output to the second site is too great. The exclusion of outliers which represent sites of extreme impact may result in the loss of important ecological and community information. Specifically, species at these sites may be 'opportunistic species' which are best adapted to the specific conditions at the site which enhances their reproductive rates. This then reduces evenness and therefore diversity at the site. Often the presence of outlying samples may be due to truncated environmental gradients (Ried *et al.* 1995) which may be improved by the addition, rather than the subtraction, of samples (Hall & Smol 1992). RJI3 is found at the extreme end of the pH gradient and prediction errors for these sites are likely to have been derived from truncation of the gradient.

6.4 Indicator Species

Within the Northern Territory there are many disused mines which leak heavy metals into the streams and rivers causing aquatic degradation. Given the increasing pressure on streams from human development, effective assessment tools are needed for consistent and fast evaluation of the condition and stressors of water resources and to provide information for solving pollution problems. One form of water quality assessment is the utilisation of indicator species, defined by their optima and tolerances of specific environmental variables. Optima and tolerance of diatoms, or species indicator values, to environmental variables such as conductivity, pH, temperature and phosphorus have been developed for many hundred freshwater diatom taxa (i.e. ter Braak & van Dam 1989; van Dam *et al.* 1994; Pienitz *et al.* 1995; O'Connell *et al.* 1997; Potapova *et al.* 2004). Such species-specific indicator values of taxa, based on species optima, have been included in autecological indices such as the 'Trophic Diatom indices' of Kelly and Whitton (1995), the saprobic index of Sladeček (1973), and the 'Pollution Index' of Descy (1979) used to evaluate water quality in European and North American rivers (Whitton *et al.* 1991; Charles 1996; Prygiel *et al.* 1999). However, in the tropical north of Australia there exists almost no autecological information on diatom species found in the area. The optima and tolerances (responses) of species in the Northern Territory are not necessarily different from those found in other areas. However, the extreme environmental conditions existing in this region for instance, high temperatures and other physico-chemistry variables, provide the opportunity to determine more complete tolerance ranges. Additionally, there is, of course, the possibility of the existence of endemic species and this influences their importance as indicators. Defining the optima and tolerance of the local diatom species to environmental variables is important in order to establish potential indicator species for monitoring programs.

Recent advancements in statistical methods, such as Generalized Linear Models (GLM) and nonparametric (smoothed) regression techniques, have greatly improved the ability to investigate species responses to the environment and specific variables. These methods have also revealed that species often have complex response curves, including non linear or symmetrical bell-shaped responses (Austin 2002). The SWAP and PIRLA projects found that weighted average regression and calibration, implemented by the computer program WACALIB (Line *et al.* 1994), is the most effective statistical method to infer species responses to pH. Weighted averaging is effective because it does not assume a linear relationship between the inferred variable and diatoms, but assumes unimodel responses. Essentially, weighted averaging maximises the covariance between the diatom data and the measured environmental variable (Korsman & Birks 1996).

In this section, diatom indicators of acid mine drainage were identified by studying diatom species responses to acidity (pH) levels in streams of the Northern Territory. pH was chosen as a measure

of pollution because of its importance in acid mine drainage and its covariance with heavy metals such as copper, aluminium, zinc, iron and manganese and other environmental variables, which are responsible for species distributions. In section 6.2.1., pH was found to be an important environmental variable in explaining the variation in diatom communities.

6.4.1. Species occurrence and optima and tolerance values

The complete dataset included 270 diatom taxa but only the 75 species (Table 6.19) which occurred in at least seven sites of the 50 sample calibration dataset were retained for analysis. This number includes the eight species with percentages greater than 50% at one or more sites (Table 6.18). For instance, *Nitzschia palea* v. *palea* has an abundance of 65% at the Cosmo Howley site CHI7 but was also present in at least seven sites. There was, in fact, only one taxon, *Pinnularia joculata*, with abundances greater than 50% that occurred at less than seven sites. While this taxon had a relative abundance of 94.8% at RJI5, it was present at just five sites in the dataset, but was included in the analysis.

Table 6.18. Species with abundances greater than 50% at one or more sites.

Genus	Species	% Abundance	Site
Achnanthidium	*minutissimum v. minutissimum*	92.7	RJI3
Epithemia	*cistula*	59.8	CCUSSC
Navicula	spec 31	66.2	RJI9
Nitzschia	*palea v. palea*	65.0	CHI7
Nitzschia	*paleaeformis*	72.0	RJI6
Nitzschia	*vasta*	66.5	CHI1
Pinnularia	*joculata*	94.8	RJI5
Kobayasia	cf. nov. spec. Nr 941/6-9	50.0	SC@BRX

Table 6.19. WA-PLS optima and tolerance for pH, Hills N2 and occurrence values for diatoms which occur at 7 or more sites.

Species	Occur.	N2	pH Opt.	pH Tol.
Achnanthidium exilis	12	5.92	7.34	1.45
Achanthidium minutissimum v. *minutissimum*	38	14.09	6.53	1.45
Aulacosira granulata	7	3.74	6.48	1.06
Brachysira neoexilis	31	8.82	6.36	1.31
Brachysira spec. 1	8	4.35	7.09	0.30
Caloneis spec. 4	7	2.67	6.07	1.96
Chamaepinnularia mediocris	11	4.73	4.33	1.15
Craticula halophilioides nov. comb.	15	5.15	6.94	1.09

Species	Occ.	N2	pH Opt.	pH Tol.
Cymbella cistula	7	2.45	5.74	2.28
Cymbella hustedtii f. *stigmata*	8	3.25	7.37	1.80
Diploneis pseudovalis	7	4.93	8.26	0.92
Encyonema minutum	27	14.26	6.15	1.30
Encyonema silesiacum	22	6.54	6.25	1.29
Encyonema spec. 6	7	3.56	6.20	1.45
Encyonopsis perborealis nov. spec.	12	6.08	6.72	1.69
Encyonopsis subspicula nov. spec.	7	1.30	5.88	1.12
Epithemia cf. *cistula*	7	3.82	8.28	0.46
Eunotia bilunaris	18	6.58	4.85	1.58
Eunotia camelus v. *camelus*	7	1.82	4.39	2.30
Eunotia camelus v. *denticulata*	9	4.70	4.70	1.64
Eunotia camelus v. *didymodon*	14	5.70	4.89	1.60
Eunotia cf. *rabenhorstiana*	10	7.32	5.65	1.22
Eunotia naegelii	16	7.87	4.84	1.43
Eunotia spec. 9	11	8.65	6.11	1.11
Eunotia veneris	14	6.34	6.15	1.44
Fragilaria capucina	10	3.57	6.75	1.40
Fragilaria tenera	10	5.53	5.05	1.68
Frustulia rhomboides	7	4.67	6.61	0.59
Gomphonema cf. *exilissimum*	16	7.88	6.56	1.01
Gomphonema cf.*exilissimum* group 2	12	6.25	5.36	1.66
Gomphonema cf. *vibrioides*	11	6.17	6.78	0.95
Gomphonema gracile	17	7.39	6.57	0.89
Gomphonema gracile group 2	8	3.44	6.42	1.38
Gomphonema lagenula	17	8.32	6.62	1.69
Gomphonema spec. 15	10	5.18	5.81	1.22
Hantzschiana amphioxys	8	4.59	4.78	1.75
Kobayasia cf. nov. spec. Nr 941/6-9	7	1.13	7.45	1.86
Luticola mutica	15	9.25	5.70	1.78
Naviculadicta difficillima	11	8.63	6.45	1.34
Naviculadicta subtilissima	10	6.68	4.56	1.05
Naviculadicta tridentulata	11	4.40	6.69	1.02
Navicula cincta f. *minuta*	7	5.59	7.07	1.39
Navicula cryptotenella	9	4.88	6.65	1.55
Navicula gerloffii	19	7.64	6.49	1.52
Navicula heimansioides	20	7.55	6.35	1.07
Navicula notha	28	12.04	6.91	1.09
Navicula radiosa	11	5.94	6.96	1.12
Navicula spec. 31	20	2.83	6.84	1.50
Navicula vitabunda	9	4.22	5.54	1.68
Nitzschia amphibia	9	6.11	7.40	1.26
Nitzschia archibaldii	19	9.72	6.53	1.73
Nitzschia cf. *hantzschiana*	15	6.23	7.09	0.71
Nitzschia cf. *pseudofonticola*	11	5.49	7.15	0.23
Nitzschia gracilis	13	5.36	4.72	1.56
Nitzschia intermedia	10	3.28	7.49	1.23
Nitzschia liebetruthii	10	3.73	8.22	0.81

Species	Occ.	N2	pH Opt.	pH Tol.
Nitzschia linearis	8	4.59	7.13	1.72
Nitzschia nana	18	2.85	4.99	1.41
Nitzschia palea v. *palea*	38	11.84	6.97	1.07
Nitzschia palea v. *tenuirostris*	12	4.44	7.85	0.98
Nitzschia paleaeformis	15	6.69	5.47	1.32
Nitzschia spec. 39	13	6.64	7.78	0.86
Nitzschia subacicularis	8	4.75	7.20	1.56
Nitzschia vasta	15	4.01	3.93	1.46
Pinnularia cf. *schoenfelderi*	8	1.47	3.70	1.50
Pinnularia divergens	8	4.07	6.11	1.25
Pinnularia joculata nov. comb.	5	1.23	6.30	1.51
Pinnularia subcapitata	18	6.89	5.35	1.59
Pinnularia tumuscens nov. spec.	7	3.46	6.24	1.51
Planothidium rostratum	17	7.72	7.29	1.78
Sellaphora pupula v. *pupula*	27	8.81	7.03	0.93
Stenopterobia densestriata	12	5.88	6.35	1.20
Staurosira phoenicenteron f. *gracilis*	8	4.90	6.88	0.87
Surirella roba	13	2.27	5.49	1.26
Synedra ulna	31	14.52	6.65	1.21

Species N2 values range from 1.1 for *Kobayasia* to 14.5 for *Synedra ulna*. Species indicator (optima) values for pH, calculated as weighted averages (WA), ranged from 3.7 for *Pinnularia schoenfelderi* to 8.5 for *Fallacia tenera,* within the measured pH range (2.95 – 8.7) from the site dataset. The pH optima for the two morphological groups of *Gomphonema* cf. *exilissimum* are 6.56 (group 1) and 5.36 (group 2). There is more than 1 pH unit difference between the two morphological groups. Similarly, the pH optima for the varieties within *Nitzschia palea* are 6.97 for variety *palea* and 7.85 for variety *tenuirostris*. The range of pH optima within *Nitzschia* ranges from 3.93 (*Nitzschia vasta*) to 8.22 (*Nitzschia liebetruthii*), over 4 pH units. However, there is a smaller pH range for *Gomphonema,* from 5.36 (*Gomphonema* cf. *exilissimum* group 2) to 6.78 (*Gomphonema* cf. *vibrioides*).

The mean pH for the 14 sites identified as impacted by the DCA in the section 6.1 is 4.5. The standard deviation between the sites is 1.3. Consequently, possible indicator species for pH will have optima and tolerance ranges below pH 5.8. Based on this criterion four species are potential indicators. These include *Chamaepinnularia mediocris,* which has a pH optimum of 4.3 and a tolerance of 1.15 standard deviations. The combined pH optimum and tolerance, which represents the upper limit of its standard deviation range, is 5.5 (N2 4.7). This is 0.3 units below the pH cut-off point of 5.8. *Naviculadicta subtilissima* has a combined optimum and tolerance of pH 5.61 (pH opt. 4.56, tol. 1.05) and an N2 value of 6.7. *Nitzschia vasta* and *Pinnularia schoenfelderi* have combined pH optima and tolerances of 5.39 (pH opt. 3.93, tol. 1.46) and 5.2 (pH opt. 3.7, tol. 1.5) respectively. The N2 values of these taxa are 4.0 for *N.vasta* and 1.5 for *P. schoenfelderi*. Based on

the optimum occurrence values of each of the taxa the species can be classified as acidobiontic according to the classification system derived from Hustedt (1938-39). Acidobiontic species have an optimal occurrence below pH 5.5. There are an additional ten species which, according to their optima, could be classified as acidobiontic, however the range of their tolerances excludes them from further analysis. For instance, the three varieties of *Eunotia camelus* have pH optima below 5, but tolerances greater than 1.6 standard deviations, creating an upper standard deviation pH range above 6.0. Although *Eunotia* species are known to be tolerant to acidic waters (van Dam *et al.* 1994), their large tolerance range precludes them from further analysis.

Table 6.20. pH range of the 14 very impacted sites.

	pH
RJI1	3.4
RJI2	4.0
RJI3	3.7
RJI4	4.4
TGI4	7.6
TGI5	4.9
HCUSEC	3.6
ECDSHC	4.4
ECUS12MC	4.5
CHI1	3.0
CHI2	4.1
CHI3	6.9
BKI1	4.9
BKI2	4.3

In contrast, 28 of the 92 species have pH optima above 7.0. According to Hustedt's acidity classification, species with optima above 7.0 are classified as alkaliphilous. Of the *Nitzschia* species, ten (from 16) have pH optima greater than 7.0. Using the 17 control sites to represent natural aquatic conditions, with the exclusion of the Redbank site ECUSBRX which has a pH of 6.0 and may be impacted by an historic disused mine, the mean pH value for these sites is 7.61. The standard deviation of pH between the sites is small, 0.54. Consequently, possible indicator species for pH will have optima and tolerance ranges above 7.07. Such species would also need to occur in at least eight sites. Based on this criterion there is just one diatom which could be considered a potential indicator of this pH condition. *Nitzschia liebetruthii* occurs at ten sites (N2 3.7) and has a pH optimum of 8.22 and a tolerance range of 0.81 standard deviations. The lower range limit then, for this species, is 7.41. For the next three potential indicator species to be considered the occurrence value of eight needs to be lowered to seven or six. *Epithemia cistula* and *Diploneis pseudovalis* both occur at seven sites and have pH optima of 8.28 (tol. 0.46) and 8.26 (tol. 0.92), and N2 values of 3.8 and 4.9, respectively. The lower limit of the tolerances ranges for these two species is pH 7.82 and 7.28. The fourth species is *Fallacia tenera* (not noted in Table 6.19). This taxon has a pH optimum of 8.48 and a tolerance of 0.26 standard deviations which places it

within the defined range. However, it only occurs at six sites and has an N2 value of 3.6. Other than the Redbank control sites, it only occurs at one other control site (RJC3). Consequently, it may be a good indicator taxa for clean conditions at the Redbank region, but not for general conditions as it may not inhabit all regions of the Northern Territory.

Table 6.21 pH range of the 17 control sites

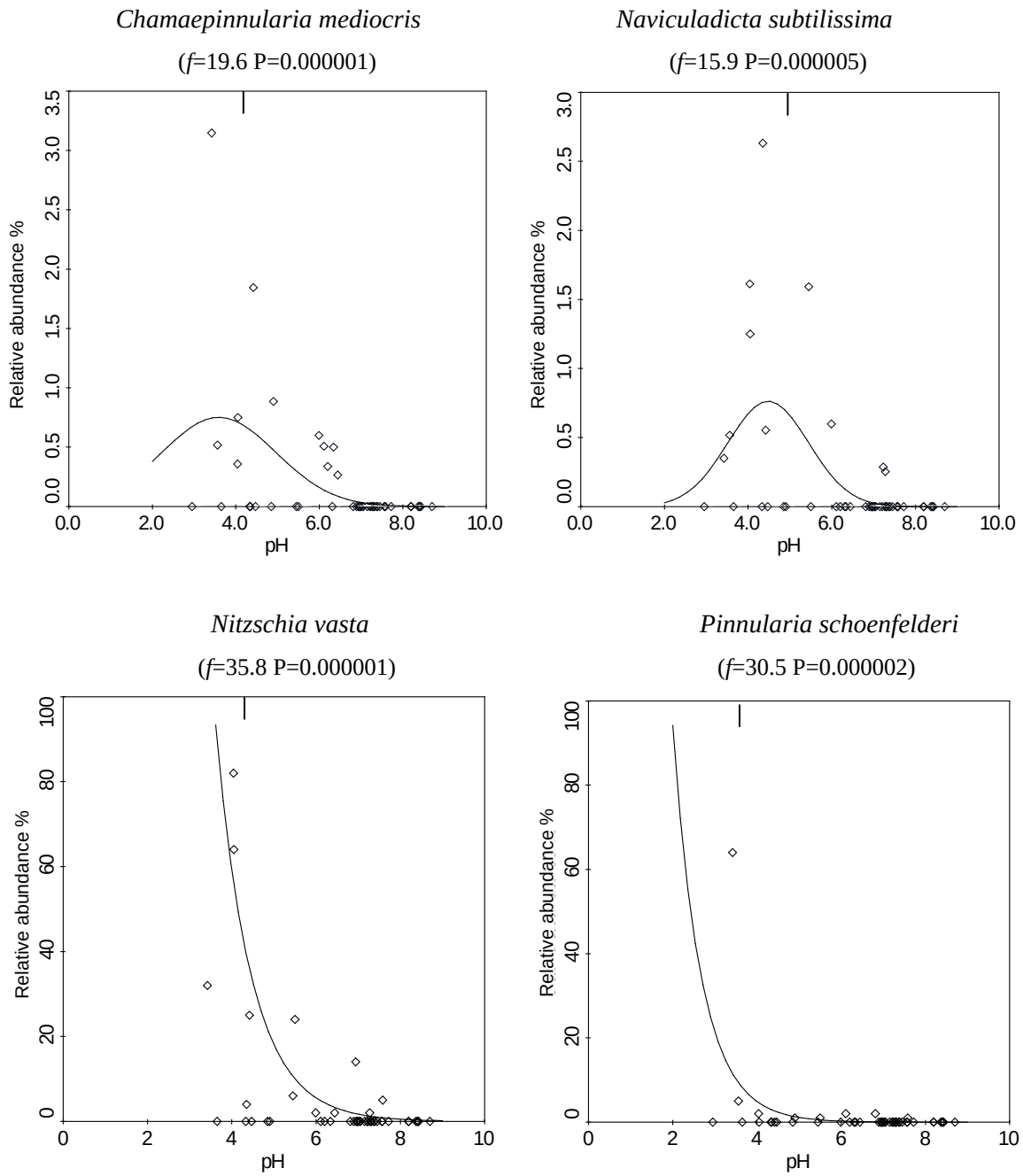
	pH
RJC1	7.2
RJC2	7.4
RJC3	7.4
RJC4	7.6
RJC5	8.2
TGC1	7.1
TGC2	7.0
12MCUSEC	6.9
ECUSBRX	6.0
CCUSSC	8.2
SCUSCC	8.4
SCDSCC	8.4
SCUS12MC	8.4
CHC1	7.3
CHC2	7.7
CHC3	7.2
CHC4	7.5

A Monte-Carlo permutation test in canonical correspondence analysis (CCA), with pH as the only constraining variable, proves that the variable was significant ($p < 0.05$) in explaining variability among diatom assemblages. The third CCA axis explained 18.8 % of the species variance in the dataset. The pH variable explains more than 10% variance of the taxon and can consequently be used in the selection of indicator species. This indicates that it is possible to develop pH inference models from the dataset. The Weighted Averaging-Partial Least Squares (WA-PLS) inference models were used as they performed better than simple WA models, and WA models with tolerance downweighting. The abundance values of species, with a tolerance less than the mean, were used in a CCA and WA regression, and are plotted against pH. The estimated WA optimum of the taxon is shown as a vertical line at the top of the plot. Species responses to pH were studied by fitting parametric regression models using a Generalised Linear Modelling (GLM) approach. The GLM was fitted with a Gaussian distribution using the program CANOCO. Parametric regression models were fitted with Poisson error distributions and tested to see if quadratic (symmetric bell shaped) or linear (monotonic) models were more significant ($p < 0.05$).

Of the taxa tested, 19 had a symmetrical unimodel response to pH, displaying significant ($p < 0.05$) fits with the quadratic regression. Eight diatom species had a monotonic response to pH based on their significant fit to the linear GLM models. When graphed, these responses appear as

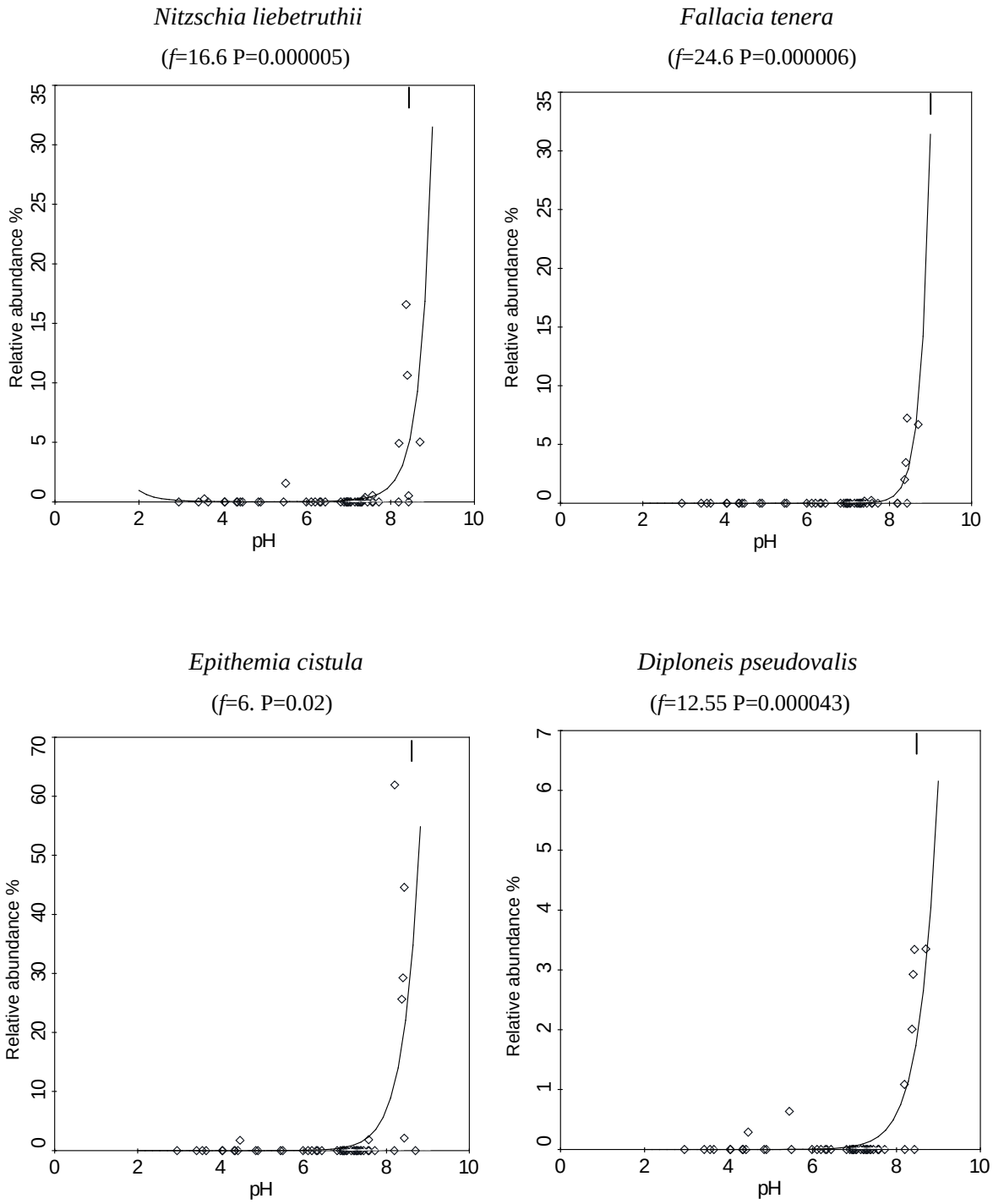
exponential curves because of the use of a log-link function (Figure 6.23). Three of these species also had a significant fit to a quadratic model but the residual deviance was not high, so the preferred parametric model was still linear. Additionally, *Nitzschia liebetruthii* and *Diploneis pseudovalis* appear to have linear response curves but a higher significant fit to a quadratic model than a linear one. All together, only 23 species showed significant responses.

The response curves of the four species which are potential indicators of aquatic environments with pH lower than 5.5 are shown in Figure 6.22. *Chamaepinnularia mediocris, Naviculadicta subtilissima* and *Pinnularia schoenfelderii* have a significant fit with quadratic regressions, while *Nitzschia vasta* has a monotonic response to pH, based on its significant linear fit. Each model was significant ($p < 0.05$). There seems to be only one site along the pH gradient, with a relative abundance of greater than 10%, which determines the low pH optimum of *P. schoenfelderii*. Additionally, only *Nitzschia vasta* has more than six sites with relative abundances of greater than 10%.



Figures 6.22 Relative abundance and modelled unimodal and monotonic responses of species to low pH. Arrows indicate species indicator values calculated as weighted averages.

Of the four species considered as possible indicators of alkaline, aquatic conditions, *Nitzschia liebetruthii* had a quadratic response as did *Diploneis pseudovalis*. *Fallacia tenera* and *Epithemia cistula* had linear responses. However, each model was significant ($p < 0.05$).



Figures 6.23 Relative abundance and modelled unimodal and monotonic responses of indicator species for high pH. Arrows show species indicator values calculated as weighted averages.

Results from this and similar studies are compared to investigate individual diatom species responses to specific environmental variables with the aim of gaining as much ecological information as possible about particular diatom taxa and to assess whether some species have different ecological ranges in other studies. In comparing the eight pH indicator species to the diatom indicator values from the Netherlands (van Dam *et al.* 1994), just two species can be

compared. These include *Navicula subtilissima* and *Navicula mediocris*, sampled from the Netherlands, and *Naviculadicta subtillissima* and *Chamaepinnularia mediocris* from this study. van Dam classed *Naviculadicta subtilissima* as acidobiotic and *Navicula mediocris* as acidophilous. However, *C. mediocris* has a pH optimum of 4.33 (1.15 standard deviations) while van Dam *et al.* (1994) found this taxon to occur optimally in waters of pH >7.0. The results of *N. subtilisima* compare closely to this study, having a pH optima of 4.56 (1.05 standard deviations). In a study of Malaysian diatoms (Maznah & Mashhor 2002) this species was found to occur in waters with a pH range of 5.7 to 8.7. Unfortunately not many comparisons can be made between the Northern Territory data and those obtained from other investigations because of the lack of similar quantitative estimates of pH optima and tolerances of diatom taxa from similar ecological conditions. In all of the above studies, there are no occurrences of species such as *Nitzschia vasta*. One of the only references to the ecology of *Nitzschia vasta* was by DeNicola (2000) who observed that it is very abundant in water of pH 2.8 to 3.4. Its pH optimum from my analysis 3.93 (1.46 standard deviations), outside the pH range noted by DeNicola.

6.4.2. Discussion

Of the 270 diatom taxa observed in the 50 samples, 92 diatoms met criteria one and four. Of these, just four species met criterion three. Four species had a combined weighted average optima and tolerance (which represents the upper limit of its standard deviation range) within the mean, plus the upper limit standard deviation range, of the very polluted sites. The levels for pH ranged from highly acidic (2.95) to alkaline (8.7). The potential indicators of acidic environments (pH <5.8) are *Chamaepinnularia mediocris* (opt. 4.33) and *Naviculadicta subtilissima* (opt. 4.56). Both have pH tolerances of more than one pH unit. The species which were defined as potential indicators of alkaline environments (pH > 7.1) are *Epithemia cistula*, *Nitzschia liebetruthii*, *Diploneis pseudovalis* and *Fallacia tenera*.

Not all taxa had clearly defined optima for pH. Rather, many taxa had a wide pH range over which they are abundant. For instance, *Achnanthidium minutissimum v. minutissimum* has pH a optimum of 6.5 (tol. 1.45 standard deviations) yet it occurs between pH 3.4 and 8.4. Within this range the taxon reaches over 20% abundance at each pH value between 3.0 and 9.0. However, despite this taxon having a relatively even distribution across this gradient, WA regression is used to generate a pH optimum. Essentially this is an arithmetic rather than an ecological optimum, and consequently, if the response is skewed, the optimum may grossly overestimate or underestimate pH conditions. Ideally what is needed for accurate derivation of environmental optima is broad gradients with minimal data gaps. This point is constantly emphasised by ecological statisticians (see Birks 1995, 1998) but is often difficult to employ. Additionally, if most of the samples are concentrated in the acidic range then the pH optimum is always going to be much lower than other studies with similar

ranges. Ideally again what is needed is an even spread of samples across the gradient range. This is however, difficult when the chemical conditions of stream change dramatically within a small geographical area.

Using, again, the example of *Achnanthidium minutissimum* v. *minutissimum*, this is just one of many species which has a broad ecological tolerance for pH. Many taxa occur over a broad ecological pH gradient and, therefore, convey little information on specific physico-chemical conditions. For instance, a taxon may have a pH optimum of 6.0 but occur at up to 40% of the total assemblage at both pH 4.0 and 8.0. If this optimum of 6.0 is applied to systems that were actually alkaline then the inferred pH will be overstated. Known taxa with broad ecological tolerances are small *Staurosira* taxa (Gell 1995; Tibby 2000).

Biological species are commonly assumed to have predominantly symmetrical, unimodal distributions patterns along environmental gradients (Whittaker 1967 in Potapova *et al*. 2004). The majority of the taxa examined here followed the Gaussian unimodel response model, but several followed a linear, logit fashion with their abundances increasing or decreasing with the increasing environmental variable. This is likely to be because only the upper or lower limit of their response range was sampled. Of the eight indicator species for acidic and alkaline aquatic conditions, each species, other than *Nitzschia vasta*, had a quadratic unimodal response to pH. Similarly, ter Braak & van Dam (1989) found that 16 of 26 common diatoms in the Western European lake dataset had significant fit to a symmetrical unimodal pH model. Additionally, the majority of the diatom taxa (58 out of 92) studied by Juggins (1992) had unimodal distribution curves along the salinity gradient of the Thames estuary. However, Potapova *et al.* (2004) found that exponential curves, fitted by linear regression in GLM, are common for the best TP indicators in the dataset. Linear regression patterns are difficult to assess as the actual optimum may not be within the measured range, indicating that the pH and copper pollution gradient lengths in the study are too short. There is a need for more sites, especially of low pH, so that there is less probability of artificial truncation of species ranges and more reliable estimates of pH optima, especially for uncommon taxa, can be gained. The lower performance of some models in terms of R^2 and associated errors (RMSE), also may have been caused by the uneven spread of values across the ranges of environmental variables. Additional sampling to fill the gaps in the environmental data is likely to strengthen the predictive capacity of the models. Because ecological species' response is assumed partly to be the result of competition, the shapes of response curves will depend, to some extent on the species composition of the calibration dataset, and consequently on the particular environmental and geographical settings of the study sites. Further investigation is needed to study how the shape of response curves for individual species might vary among different datasets, and one also needs to determine which factors are most important in determining the competitive ability of diatoms.

One of the most obvious problems with the eight diatom indicator species were low relative abundances. Small numbers of taxa may be found in locations where conditions are very different from optimal (Harper 1999). These low abundances may be due to the valves being dead at the time of sampling (washed downstream from areas of different aquatic conditions) or the lack of interspecific competition. Only *Nitzschia vasta* has a relative abundance of more than 10% at at least seven sites which is half the total, 'very impacted' sites. The second best indicator, in terms of relative abundance, is *Epithemia cistula* which has a relative abundance greater than 10% at four sites. *Pinnularia schoenfelderi* and *Nitzschia liebetruthii* exceed 10% of valves at only one or two sites. At no site was *Diploneis pseudovalis, Fallacia tenera, Naviculadicta subtilissima* and *Chamaepinnularia mediocris* above 10%. Consequently, taxa such as *Kobayasia* cf. *nov. spec.*, which has relative abundances of 0.7%, 0.2%, 0.7%, 0.5%, 50% 0.3% and 0.3%, will have its optimum heavily biased by the sample with the 50% abundance. Therefore the N2 for this taxon would be slightly greater than 1.0, as Hills N2 is an indication of the number of sites where a taxon occurs at a significant relative abundance. A taxon with an N2 of 50 (when there are 50 sites) is a taxon which occurs equally abundantly at all sites. In view of this, perhaps potential indicator species should have a high N2 value rather than an occurrence of more than seven. Taxa with high N2 values and narrow tolerance could be considered strong indicators whereas uncommon taxon with broad tolerances would be less effective. Consequently, because of the low effective number of occurrences of many taxa, the optimum and tolerance lists generated here should be considered preliminary.

Finally, difficulties were encountered when attempting to compare the ecological optima and tolerance of my eight potential indicator species to those from other studies. This problem was mainly due to the very different species communities present in the other studies and the unique chemical and bio-climatic conditions of my sites.

6.5 Teratological forms of diatoms

The presence of abnormal forms of organisms is often associated with extreme environments. In diatom research aberrant forms of diatoms have been termed 'teratological' (Miquel 1890). While cell shape and size can change during cell division, in natural diatom populations cell shape and ornamentation normally varies only slightly within a species, and deformities in diatom morphology from unpolluted waters are rare. Adverse environmental conditions induce a formation of teratological forms of many diatom species. One of the most commonly observed effects of heavy metal poisoning is a change is cell size or morphology. This has been observed in a wide variety of organisms, including Bacillariophyceae (Sunda & Guillard 1976). In diatoms variations include deformities of cell outline, interrupted patterns of striation and raphes, and bifurcated or ecentric raphes in *Cymbella*, *Diatoma*, *Ceratoneis*, *Fragilaria* and *Synedra* species (Antoine & Benson-Evans 1986). Fisher *et al.* (1981) treated diatom cultures with metals and showed reduced cell division rate and marked increase in cell size and volume. For deformities to occur, heavy metal concentrations must be high enough to cause the formation of abnormal frustules, but low enough to allow for at least one division (Pickett-Heaps *et al.* 1990). This range of concentration is often quite narrow (ter Braek & Jensen 1976). Dokulil *et al.* (1997) are among the few researchers who have found that teratological forms of diatoms can be used as bioindicators of polluted systems.

The aim of this section is to determine which diatom species display abnormal cell morphology and to describe the characteristics of the teratologies. A second aim is to determine if the presence of teratological diatom cells in samples can be used as an indicator of heavy metal pollution, and to reliably infer copper or acidity.

6.5.1 Occurrence of normal and abnormal cells

From the 270 species in the dataset, the only species found to display abnormal cell morphology at heavy metal impacted sites was *Nitzschia vasta*. This taxon occurred in eleven of the 50 sampled sites (Table 6.22), each of these having different levels of impact. No *N. vasta* frustules were observed at sites from Tom's Gully mine. From these eleven sites, aberrant forms of *N. vasta* were present at five sites with percentages ranging from 3% (CHI8) to 80% at site CHI1. *Nitzschia vasta* is present at low occurrences at ten of the eleven sites (<12%), with site CHI1 dominated by *N. vasta* (100%). At five of the eleven sites, aberrant and non-aberrant forms of *N. vasta* co-occurred.

Table 6.22. Occurrence of *Nitzschia vasta* frustules.

Site	% of *N. vasta* (n=200)	% of teratogenic *N. vasta*	pH	Cu mg/L
RJI1	1.5	0	3.42	5.99
RJI2	7	0	4.04	11.5
RJI4	4	0	4.42	1.43
RJI6	10	0	6.31	0.6
CHI1	100	80	2.95	12.2
CHI2	11.5	8	4.05	0.16
CHI3	7	5	6.94	0.01
CHI5	5	3	5.5	0.23
HCUSEC	10	10	3.56	97.8
12MCDSEC	3	0	5.45	1.41
ECDSHC	2	0	4.35	14.6

Nitzschia vasta occurs at sites with a pH range of 2.95 to 6.94 and a copper range of 0.01 to 97.8 mg/L. Aberrant frustules of the taxon are also present throughout this range. However, while aberrant *N. vasta* frustules are present at the most copper-polluted (HCUSEC) and acidic sites (CHI1), site RJI1, which is highly acidic (pH 3.4), contains no aberrant forms of *N. vasta*. This indicates that the aberrations are not restricted to polluted sites, as defined by their pH levels and copper concentrations. Additionally, a factor other than pH or copper may influence the formation of aberrant frustules.

6.5.2 Description of Nitzschia vasta

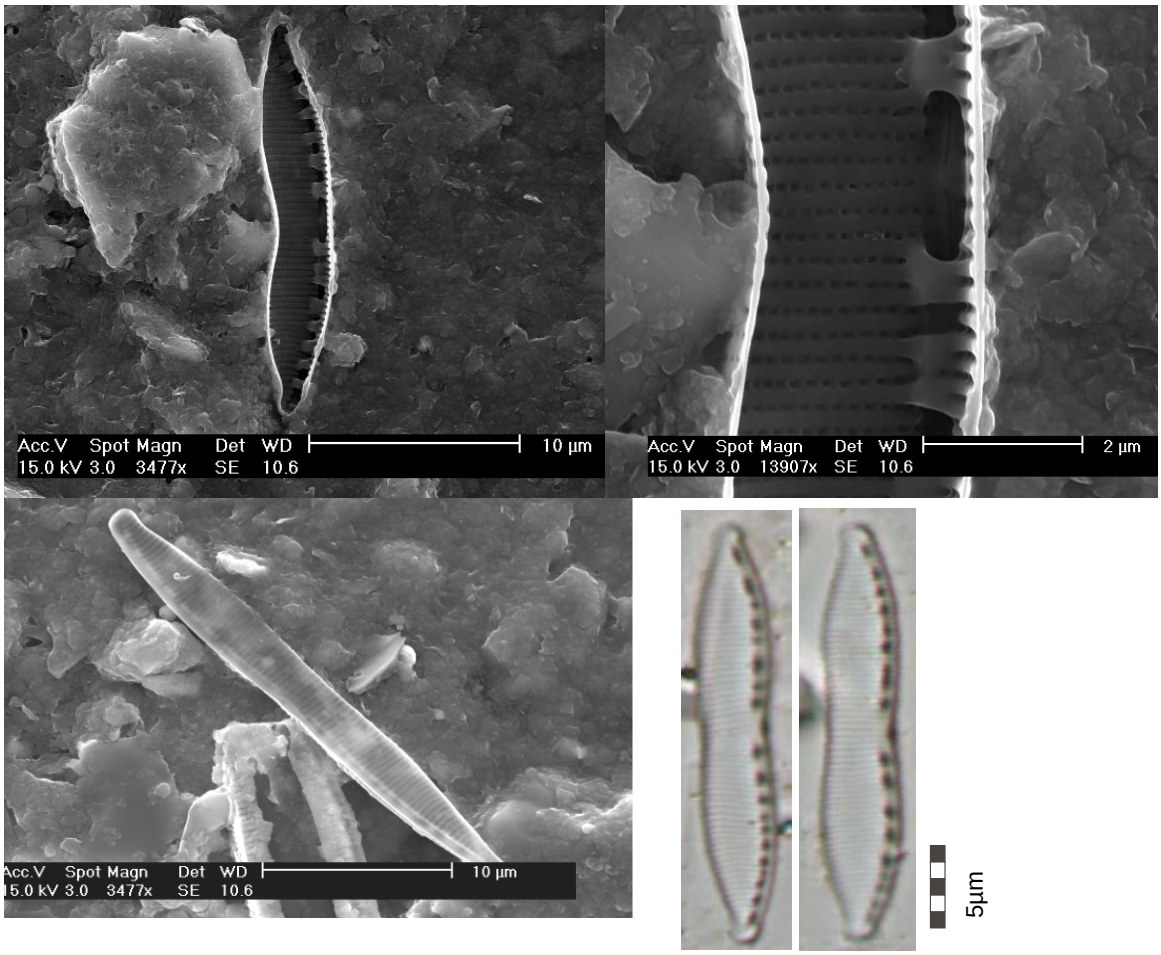
Lange-Bertalot's (2000) description of *Nitzschia vasta* is based on just one individual from Hustedt's type material. This individual had a length of 30 µm and a breadth of 5 µm. He described the fibulae as exaggeratedly large and irregularly spaced with the two middle fibulae spaced more widely apart. The fibula density noted was 6-8/10 µm and striae 27/10 µm. The distribution of the taxa has been noted in the Ems River mouth and the North Sea (Lange-Bertalot 2000), and in lignite springs in Texas, U.S.A (Winsborough, B., unpublished data in DeNicola 2000). There is essentially no ecological information for *Nitzschia vasta* except that it was identified by DeNicola (2000) as a true inhabitant of highly acidic waters (pH < 3.5).

Nitzschia vasta Hustedt 1939

The valves of *Nitzschia vasta* are linear and slightly constricted in the middle. The valve apices are obtusely rounded. The raphe is strongly eccentric and the fibulae are coarse and unevenly spaced. The two central fibulae are distinctly distant. The length of the frustule from the Northern Territory

samples ranges from 18-31 µm and its breadth, from 3-4 µm. The lengths of these valves are longer and narrower than noted by Lange-Bertalot. The fibulae were found to have a density of 7-8/10 µm. The striae are coarse and clearly visible with light microscopy (30-35/10 µm). The pattern of the striae, as seen in Figure 6.24, is parallel to slightly radial towards the apices. The striae are formed of single rows of areolae. As described in the previous section of this chapter, *Nitzschia vasta* has a copper optimum of 19.5 mg/L (± 34.7 standard deviations) and a pH optimum of 3.9 (± 1.5 standard deviations), indicating that it is highly pollution tolerant.

References: Krammer & Lange-Bertalot 1997 (pl. 56, f. 8, 8a)

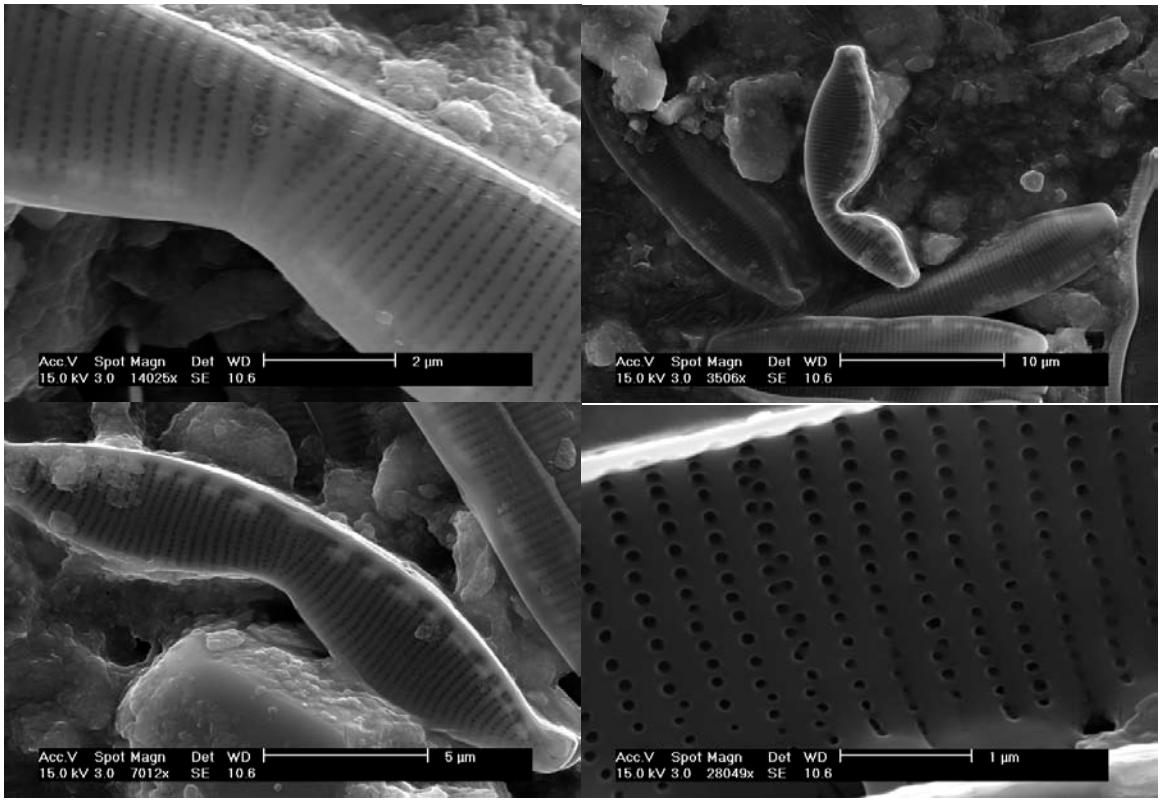


Figures 6.24-28. SEM and light micrographs of *Nitzschia vasta* frustules from site BKI2.

6.5.3 Description of teratological forms of Nitzschia vasta

The types of mutations observed are not consistently displayed in each frustule. The types of teratology include a change in the cell shape. As is evident in Figure 6.30, the valve is often severely constricted at the centre and one end is often disproportionately larger than the other. Figure 6.31 illustrates the larger node in this teratological frustule, with the two central fibulae extremely widely spaced. The stria pattern is interrupted, with rows of striae merging especially in

the central region of the frustule (Figure 6.29 and 6.31). The areolae are in some instances, unevenly spaced (Figure 6.32) and often in double rows. There is an extreme radial arrangement of the striae in the central region of the frustule.



Figures 6.29-32. SEM micrographs of teratological *Nitzschia vasta* valves from site CHI1.

6.5.4 Multivariate regressional analysis of relationships between Nitzschia vasta and pH and copper.

The relationship between the percentage of teratological *Nitzschia vasta* and the environmental variables pH and copper (Cu) was explored using bivarate regression analysis in JMP 5.1 (SAS Inc. 2003). Bivariate analysis explains how the distribution of one continuous variable is related to another variable. Linear regression fits a straight line through the data points using a partial least squares regression. As illustrated in Figure 6.33 a strong regression is not evident between the percentage of *N. vasta* teratologies and pH ($R^2 = 0.21$, $p = 0.16$). The same lack of association was observed for *N. vasta* and loged copper concentrations ($R^2 = 0.07$, $p = 0.43$) (Figure 6.34). Additionally, the sites do not cluster close to the line of regression. However, the fit of the linear line indicates that, with a decrease in pH, there is an increase in the percentage of teratological *N. vasta* valves. Similarly, with an increase in copper concentration, there is an increase in the percentage of teratological valves.

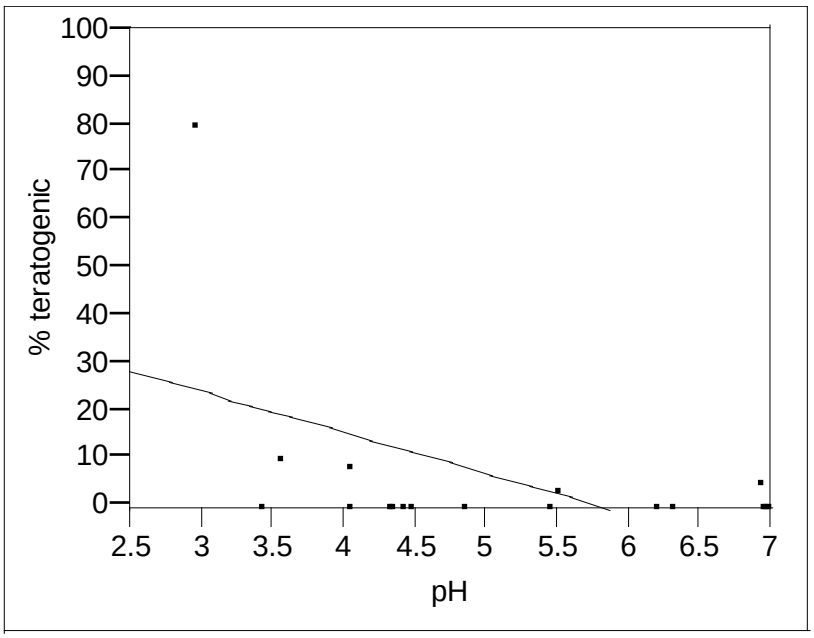


Figure 6.33. Bivariate fit of percentage of teratological *N. vasta* frustules and pH.

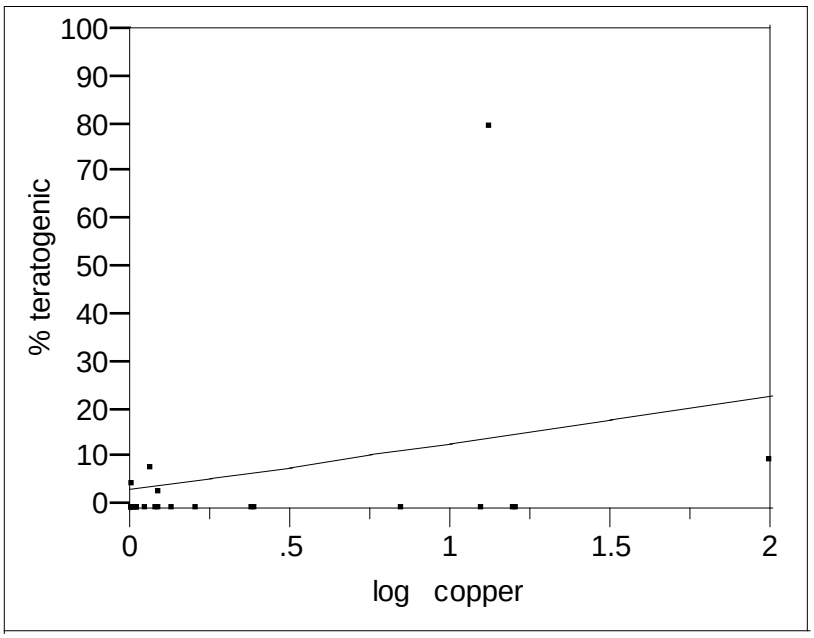


Figure 6.34. Bivariate fit of percentage of teratological *N. vasta* frustules and copper (mg/L log).

6.5.5 Non-parametric linear regressions

As neither copper nor pH showed linear relationships to the percentage of teratological *Nitzschia vasta* frustules, non-linear regressions between the variables were derived. This was accomplished with the computer program SigmaPlot (version 6.0) and it was found that a rational curve with two parameters was most significant (Figure 6.35). The fit of the rational curve has a 98% explanation

strength ($R^2 = 0.98$) indicating that there is a good relationship between the two variables. Additionally, the standard error, which measures the uncertainties in the estimates of the regression coefficients, is low, being 0.07. The analysis of variance is significant with p, the probability that there is an association between the two variables, being less than 0.05 ($p = 0.001$). However, the measured pH gradient did not extend below 3.0 so it is unknown what the response, or presence, of *N. vasta* would be below this level. What is evident is that, below pH 3.5, the number of teratological valves dramatically increases.

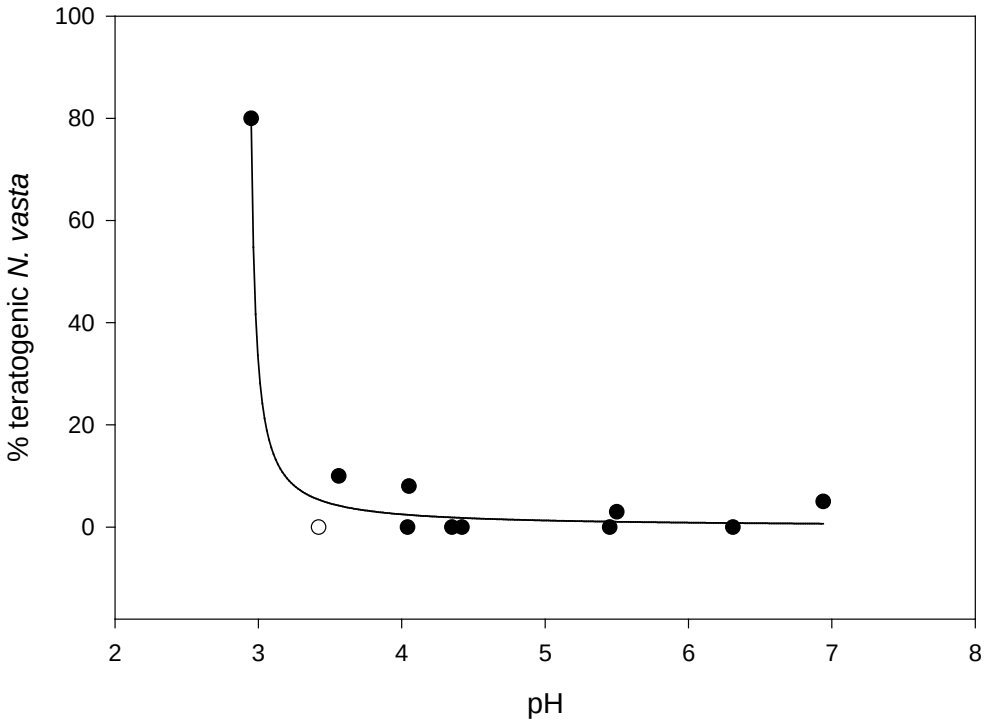


Figure 6.35. Rational curve fit of pH and percentage *N. vasta* valves with two parameters

The regression with the variables copper and percentage of teratological valves with a fitted rational curve was not significant. The R^2 value is 0.07. The p value for the analysis of variance is 0.44 and the p value 0.66.

6.5.6 Discussion

Of the 270 diatom species identified, only *Nitzschia vasta* had deformed valves at any of the sample sites. *Nitzschia* is one genus which has previously been identified, along with *Eunotia* and *Fragilaria*, as being more easily affected by pollution than other genera (Barber & Carter 1981). However, *Nitzschia vasta* has not, to date, been noted in any published studies as forming teratological valves. The teratological structures of these valves were expressed morphologically in a large variety of ways, differing in each valve. The variations occurred in the valve shape, striation pattern and density, as well as areola pattern and fibula spacing. The teratological valves were easily distinguishable from “normal” valves.

Of the 50 sampled sites *Nitzschia vasta* valves were present at eleven sites, and teratological forms were present at five sites. Four of these five sites were distinguished as very polluted by the DCA analysis in section 6.2.1.1. The relative abundance of teratological valves ranged from low (3%) to almost complete dominance at one site (80%). Given suitable conditions, these teratological forms may compete with other diatom species and will become common, if not dominant, constituents of the diatom community.

To assess the relationship between teratological *Nitzschia vasta* valves and environmental variables contributing to acid mine drainage at the sites, the percentage of teratologies was related to pH and copper. Both models showed that no statistically significant linear regression could be fitted for the percentage of teratologies and pH or copper. However, the relationship between the variables was then assessed using a non-linear regression fitted with a rational curve with two parameters. According to the high strength ($R^2 = 0.98$) of the fit, and the low probability error (< 0.05), it is clear that a relationship exists between the percentage of teratological *Nitzschia vasta* valves and pH. This indicates that, with a decrease in pH, there is an increase in the number of teratological valves. However, this association is based largely on site CHI1 which is almost totally dominated by teratological valves. This same positive association was not evident between the variable copper and the percentage of teratologies. This indicates that the teratologies are a response to increasing acidic conditions at sites, rather than increasing copper concentrations. However, teratological cell morphology observed in samples from nature must be used with caution as an indicator of heavy metal pollution because the abnormalities can be caused by a variety of other factors including nutrient limitation (William *et al.* 1980; Barber & Carter 1981; Dickman 1998).

An additional problem of using teratological diatoms collected from waters polluted by acid mine drainage is separating the effects of elevated metals and low pH, both of which are known to cause abnormalities in diatom cell walls (Carter 1990). To obtain greater statistically significant results, the relatively low percentage of teratological valves, at all but one site, needs to increase greatly. The lower performance of some models, for instance the relationship between copper and teratologies, in terms of R^2, may have been caused by the uneven spread of values across the ranges of the environmental variable. Additional sampling to fill gaps in the environmental data is likely to strengthen the predictive capacity of the models. However, in reality, there may be not enough sites with teratological frustules to improve the capacity of using teratogenic forms to monitor acid mine drainage impacts.

Chapter 7 - Study conclusions and future research possibilities

7.1 Taxonomic conclusions

Taxonomic precision is crucial for consistency in monitoring studies utilising biological organisms. In this study more than 260 diatom species and 45 genera were identified. However, within this dataset, from the tropical region of the Northern Territory, there is a clear indication that a proportion of the species identified are endemic to the region or at least have not been previously identified in any taxonomic texts. The most pertinent indication of this is that there were over 40 species which were unidentifiable using the current available taxonomic books.

Although taxonomic precision is crucial, the systematics of diatoms has been almost exclusively based upon frustule characteristics. This method of classification is problematic when phenotypic variation within and between species does not reflect genetic variation. By statistically grouping *Nitzschia palea*-like cultures, based on variations of particular morphological characteristics, the cultures could be separated into five distinct groups. However, although the cultures displayed statistically significant morphological variation the cultures were not found to be genetically variable. Each of the cultures could be identified as *Nitzschia palea.* Evidently, in the case of *Nitzschia palea*, phenotypic variation does not reflect genetic variation. A broad morphological classification approach, including the large length and width ranges, which relies purely on light microscopy, is appropriate for this taxon. The research also supports the cosmopolitan distribution of this taxon.

The type of sequencing applied to determine genetic variation of the *Nitzschia palea* cultures was not able to detect varieties within the *N.palea* complex. The definition of intraspecific variation is important even within the *N.palea* complex since, as demonstrated by weighted averaging analysis, there was over one pH unit of difference between the optimum of *Nitzschia palea* v. *tenuirostris* and *N. palea* v. *palea*. A one pH unit of difference, although seemingly small, is significant enough to indicated different environmental conditions. This suggests that there is a need for greater genetic resolution to define varieties and determine whether they are in fact genetic variants or mophotypes.

While northern hemisphere floristic texts can be applied to identify *Nitzschia palea* it would be hazardous to use this one taxon as an example for the region and assume that taxonomic texts produced for northern hemisphere regions can be applied to tropical regions of Australia. In addition to the 40 unidentifiable species, the genetic analysis of selected Northern Territory diatom taxa successfully defined one new species, *Nitzschia* spec. 39. Although this taxon is morphologically very similar to *Nitzschia subacicularis* it is genetically distinct. Consequently,

when applying a taxonomic text produced for the northern hemisphere to the Northern Territory one is at risk of taxonomic force fitting. As taxonomic uncertainties and force fitting of interregional taxonomic texts inhibits consistent and correct application of taxonomy, which is essential to ecological studies especially when identifying geographical distributions, this can mask endemic and indicator species. Multivariate analysis of morphological characteristics may help in solving these problems but, ideally, a combination of morphological techniques and molecular genetics or breeding tests is ultimately most desirable.

7.2 Conclusions from the monitoring study

Monitoring is an essential element in efforts directed at maintaining and improving water quality. It enables the detection of changes in aquatic biotic communities and water chemistry that have resulted from pollution. It therefore aids in the evaluation of the impact of mining activities. Biological monitoring is a valued extension of monitoring by chemical means as organisms better reflect changes in ecosystems as they continuously respond to changes within their environments. Though diatoms possess a great number of features which make them suitable biological indicators there are a variety of methods in which they can be applied. The tropical nature of the Northern Territory, coupled with the extent of acid mine drainage, constitutes a unique environment with extreme conditions suited to assessing the transferability of methods of monitoring developed overseas.

This study supports several established principles in diatom ecology and physico-chemical water quality monitoring. These include observations of high levels of responsiveness of diatoms to changes in water chemistry, particularly pH and heavy metals. As with other analyses of mine impacted waters, canonical correspondence analysis identified pH and aluminium as the principal environmental factors structuring the diatom communities. The calibration set models generated to infer pH and heavy metals had high predicative capabilities. Additionally, detrended canonical analysis of the species dataset was successfully able to separate very impacted sites from those which were moderately impacted or control sites. Overall, the species dataset, rather than the datasets utilising genus level identification or bioavailable metal fractions, provided the statistically strongest results. Although monitoring using genus level identification is preferred as identification is rapid, an advantage of diatoms for bioassessment of streams is that, even with species level identification, costs are low. The time which was taken to sample stream diatoms required less than 20 minutes per site. However, this will of course be dependent on the number of physical environmental variables sampled concurrently. By contrast, the AUSRIVAS protocols for sampling of macroinvertebrates require up to an hour of field or laboratory sub-sampling per sample, in addition to the time required to collect the bulk sample (Chessman *et al.* 2007). However, although identification of diatoms at species level does require considerable training and experience,

regional inconographs such as the one produced in this study will greatly facilitate the adoption and development of diatom monitoring studies.

Diversity indices were found to be less successful than transfer functions in indicating pollution. Neither the Shannon-Wiener nor the Simpson's diversity index could be significantly or strongly related to the variables pH or copper. Although species richness could be significantly related to pH it could not be closely related to copper. Contrary to findings from other studies, species richness tended to be higher in polluted waters than control sites.

Unlike similar studies of acidic environments, C*hamaepinnularia mediocris, Naviculadicta subtilissima*, *Nitzschia vasta* and *Pinnularia schoenfelderi* were found to be the best indicators of acidic environments (pH<5.8). Conversely, species which have been identified as indicators of acidic environments in other studies were not present in this research. Clearly, it is important to identify indicator species specific to particular regions and environments.

Nitzschia vasta was the only teratological diatom present within the dataset. The teratological structures of these valves were expressed morphologically in a variety of ways. However, as with other analysis, the relationship between teratological valves and mine impact was not significant. The lower performance of some models, for instance the relationship between copper and teratological, in terms of R^2, may have been caused by the uneven spread of values across the ranges of the environmental variable. Additional sampling to fill gaps in the environmental data is likely to strengthen the predictive capacity of the models.

7.3 Extension and improvement of diatom monitoring heavy metal pollution

Benchmark studies such as this are pertinent for understanding how the natural environment has been affected by mining and the subsequent impact of acid mine drainage. In this study, diatoms responded strongly to heavy metal pollution and associated characteristics of low pH. However, there are a number of factors in using diatoms as monitors which are unknown or which can be improved upon. For example, the sampling procedure used in this study does not make the distinction between live and dead cells. This seems particularly important in this study, as there were many diatoms in the highly polluted sites which were rare or had low relative abundance. Although the distinction between live and dead frustules has not been necessary to show relationships between taxon and environmental variables, the effect of including only live cells is unknown. Theoretically, the inclusion of dead cells may provide misleading data if they are washed to a site from upstream reaches or tributaries with different environmental conditions, or they may reflect past conditions too far removed from the time of sampling to be relevant to present biological integrity. Artificial substrata have been used to reduce incidence of dead cells but in

environments such as the tropical Northern Territory, flooding can result in the loss of substratum and dangers in setting substrata because of crocodiles, reduces the positive use of this method. Identification of live material (Cox, 1996) may be used, but requires considerable expertise and time, as morphological characteristics are hidden by staining methods. Additional microscopical examination of live material to check for the presence of large numbers of empty frustules (with no chloroplasts) may also be useful. This could be an interesting avenue of research, especially as monitoring of extreme environments may include a selection for more tolerant species when conditions change abruptly along a stream from natural to polluted.

As with this study, species tolerances ranges of ecological variables have been established almost completely empirically (Ivorra *et al.* 2002). The occurrence of certain species in relation to one or more metal is often difficult to interpret. One of the problems encountered using diatoms as bioindicators of acid mine drainage has been separating the effects and tolerances of diatoms to elevated metals and low pH. Ecotoxicology studies are essential in order to determine the precise tolerance range and responses to co-varying pollutants in aquatic environments. The autecology of taxa has rarely been verified experimentally (Cox 1991, 1993; van Dam *et al.* 1994). Cox (1993) concluded from culture experiments on four freshwater benthic diatom taxa, grown under contrasting light, temperature and pH conditions, that the physiological ranges were not consistent with the field distributional data of the species which indicated different capacities of the species. Additionally, the merging of existing diatom datasets from a variety of geographical locations reflecting different environmental conditions could provide the entire environmental gradient and, therefore, generate more robust optima and tolerance ranges.

One issue which was consistently problematic for the majority of the monitoring methods utilised was the nature of decreasing heavy metal concentration gradients downstream of the pollution source. There is a sudden change of metal concentrations when the metals drop out of the water column in response to increased buffering and pH levels. Consequently, it is almost impossible to sample a continuous concentration gradient. In statistical analyse this meant that the most polluted sites were considered as outlier samples. For instance, the copper concentration at site CHI1 was 12.2 mg/L, but at the following site concentrations dramatically decreased to 0.2 mg/L. Thus in many of the statistical models site CHI1 was considered an outlier and excluded from analysis even though it is an important representative of the most impacted sites. This problem may be lessened by closer positioning of site and inclusion of all outliers, even if the end result is less statistically significant.

7.4 Future diatom monitoring of mine impacted freshwater streams

Australia faces a daunting problem regarding its water quality and resources. With the knowledge that Australia is the driest continent after Antarctica and the country with the most variable precipitation, water quality issues are at the forefront of environmental management. However, as a nation which relies upon the exploitation of its natural resources, in this case mining, monitoring is a critical means directed at identifying the water quality issues that arise from the conflict between water management and development. Consequently, monitoring is now an integral part of environmental management. The findings of this study are extremely valuable and provide strong evidence to suggest that diatoms are an appropriate tool to indicate the water quality of tropical creeks and streams affected by acid mine drainage.

This study demonstrates the continued need for research in the tropics in order to be able to better understand biological responses and enable comparisons between studies. Additionally, the benefit of diatoms as biological monitors varies depending on the method utilised. In these highly impacted systems, transfer functions produced the strongest results. Further work culturing diatoms, and combining this technique with ecotoxicological research, will help verify autecologies of taxa and their responses to co-varying pollutants. This will strengthen the use of diatom taxa as indicator species. Continued monitoring of these sites can add much to our ecological understanding of these highly impacted systems which, in turn, will lead to better management of the systems for both sustainable resource development and conservation.