2016

Small-scale biogeographic patterns of benthic bacterial and ciliate communities in the saline ponds of Lake MacLeod, North-Western Australia

Christopher Kavazos
Edith Cowan University

Small-scale biogeographic patterns of benthic bacterial and ciliate communities in the saline ponds of Lake MacLeod, North-Western Australia

Christopher Ronald James Kavazos

Thesis submitted for the degree of Doctor of Philosophy

in the School of Science

Edith Cowan University

2016

Edith Cowan University

Copyright Warning

ABSTRACT

This biogeographical thesis tests for the presence of taxa-area and distance-decay relationships, which are common among macrobionts, in prokaryotic (bacterial) and micro-eukaryotic (ciliate) communities. Microbial biogeographical patterns may be distinct because of the high abundances, diversity and dispersal capabilities of microbes, in comparison to macrobionts. The Northern Ponds of Lake MacLeod, north-western Australia, provide an ideal location to address this topic, because the ponds are effectively hydrogeomorphologically identical, other than in surface area, and biotic histories can be assumed to differ only according to distance of separation. This means that hypotheses concerning species-sorting and neutral processes on microbial assemblages can be tested in a natural setting.

Characteristics of the physical environment were determined using bathymetric and hydrodynamic surveys in eight ponds. For each pond, e vaporative outflow was determined using pan evaporation rates, and the hydrodynamic characteristics of each pond were described by measuring water flowing out of the ponds. Four pond morphotypes were distinguished on the basis of physical characteristics (surface area, volume and mean depth) and hydrodynamic properties (water residence time and percentage of evaporative loss).

For ionic and nutrient variation within and between the ponds, concentrations were expected to vary based on residence time of the brine within the ponds, evapo-concentration and subsequent precipitation of mineral phases. The water chemistry was found to be similar to seawater, with major ionic ratios remaining rather constant throughout each pond. C ygnet Pond differed from the other ponds in that it was enriched in Mg and Ca and depleted in K. Sediment characteristics were also investigated by microscopy. Six sediment types were described based on the particles found in each sample. There were no clear relationships between sedimentology types and water chemistry, and between each of these and the pond morphotypes.

The bacterial and ciliate biofilm communities were analysed using DNA community fingerprinting methods, and constrained to the above environmental parameters using redundancy analyses. Distance-decay relationships were found for the bacterial

communities within the ponds, and occurred at relatively short distances (<100m). There were no such relationships for the ciliate communities. Taxa-area relationships were not found in either community. Spatial redundancy analyses suggest that β-diversity across the pond complex manifests itself mainly because of the differentiation of taxa occurrences among the ponds, and could not be explained by the environmental variables. Species co-occurrence models suggest significant segregation in community composition (i.e. not randomly assembled) while none of the communities appear to conform to predictions based on neutral theory.

The results therefore provide evidence that microbial bacteria and ciliate biofilm communities can conform to observed biogeographic patterns for macrobionts, although neither community displayed taxa-area relationships. The communities differed in that a distance-decay relationship was only found in the bacterial community, where ciliate taxa are distributed 'patchily', and not as a function of distance. An alternative model is proposed for the bacteria and ciliate communities of Lake MacLeod; each pond, because of their isolation from one another, is influenced by stochastic events which differentiate the ponds via ecological drift. This thesis demonstrates that these microbial communities are capable of having complex biogeographies, and that processes such as ecological drift may be important determinants of their structure.

COPYRIGHT AND ACCESS DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

(i) incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;

(ii) contain any material previously published or written by another person except where due reference is made in the text; or

(iii) contain any defamatory material.

Signed (signature not included in this version of the thesis)

Date…….

ACKNOWLEDGMENTS

I acknowledge the people of the Yamatji Nations as the traditional custodians of Lake MacLeod. I also acknowledge the Whadjuk people of the Noongar Nations as the traditional custodians of the land on which I have worked and lived during my studies.

I would like to thank my three supervisors, Prof. Pierre Horwitz, Dr. Megan Huggett and Assoc. Prof Ute Mueller. Without their prompt replies to my e-mails, frequent meetings and readings of drafts, this thesis would never have been accomplished. Ute's statistical knowledge has been invaluable in helping me design and analyse my study, and her suggestion for me to learn how to use R has changed the way I will work with data forever. Megan's knowledge and passion for the microbial world, and her constantly open door has helped me immensely. Pierre has been a great mentor and a good friend. It has been a pleasure to work with you and I have learnt a lot just being in your presence.

Dampier Salt Ltd. provided the funds for this project, for which I am grateful. The Dampier Salt Ltd staff provided amazing practical, logistical and administrative support. I would particularly like to thank Steve Rusbridge, Matt Inman, Brent Tobin, Starr Gianatti and Brett Renton. Ultimately, Dampier Salt Ltd provided me the opportunity to access and study the unique wetlands of Lake MacLeod, which I feel extremely fortunate. I must also thank the owners of Quobba and Gnaraloo Stations for providing access to the lake on their tracks. Dave Bauer also provided much insight and knowledge on Lake MacLeod, particularly with advice on how to access areas of the lake by foot, four-wheel drive and boat.

The molecular methods used in this study were done by the Australian Genome Research Facility Ltd (AGRF). Without their help, and in particular, the help and advice from David Chandler, it would have been impossible to analyse the microbial communities.

I have relied on t he help of countless people throughout my PhD. The other team members involved with other Lake MacLeod projects have become great friends, and going through the PhD process with them has undoubtedly enriched my experience. In particular, I want to acknowledge Natasha Dunham, Hannah Cameron-Caluori,

Demetris Bertzeletos, Ben Jones, Michelle Newport and Dr. Sora Estrella. Natasha and Hannah both helped me immensely in the field, and were great companions. Alison Sampey provided invaluable assistance in the lab. Rob Czarnik was with me during almost every field trip, and assisted with just about every logistical and practical component of this study. This study could not have been completed without his tireless efforts in the field.

I am thankful to all my friends who have shared this journey with me, including: Maryam Abdolahpour, Pierre Bouvais, Udhi Hernawan, Flavia Tarquinio, Tim Doherty, Audrey Cartraud, Kwasi Frimpong, Alba Esteban Pacheco, Aldo Turoc and Federico Vitelli. Above all, I would like to thank Ana Luiza Muler and Bruno Buzatto for being great house mates and helping me finish my thesis. My life would have been so much more difficult without your kindness.

I will be eternally indebted to my girlfriend, Ashlee Peeters, because of the sacrifices she has made to continue to be with me in Perth. Thanks Ash for your moral support and constant source of sanity. I could not have done it without you.

This thesis stands as a testament to the support and encouragement I have received from my family throughout my life. My parents, Judy and Michael Kavazos, have always wanted the best for me, and have encouraged me to follow my dreams. I'm also thankful to my late Grandfather, Pop Ron. Pop was always proud of my achievements at university, and I am sorry he cannot be here to see me at the end.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This thesis is dedicated to the memory of my Grandfather.

Ronald Toohey
(1926-2014)

CHAPTER 1. INTRODUCTION AND STUDY DESIGN

1.1 Introduction

The study of species distributions and changes in ecological communities along geographic gradients and habitat patches is known as biogeography, and was first described in MacArthur and Wilson's seminal thesis, 'The theory of island biogeography' (1967). Today, one of the key goals of ecology is to understand how taxonomic groups contribute to the formation of communities and how communities are structured spatially. Understanding patterns across the spatial scale gives important clues regarding the underlying mechanisms that regulate community structure and biodiversity, ultimately playing a pivotal role in the development of theories that explain the nature of biological diversity (MacArthur and Wilson 1967, Hubbell 2001). While investigated in hundreds of studies using plants and animals, studies of scaling patterns in the microbial world are only just emerging, and often with conflicting conclusions. Because microbes play pivotal roles in the Earth's biogeochemical cycles (Falkowski *et al.* 2008) and functioning of ecosystems (Bell *et al.* 2005b, Langenheder *et al.* 2005, Reed and Martiny 2007) in addition to their overwhelming abundance and diversity (Whitman *et al.* 1998, Torsvik *et al.* 2002), there is sufficient grounds to increase knowledge on how these communities are spatially structured, and the mechanisms that create these patterns.

The discipline of microbial ecology involves the study of both prokaryotic and eukaryotic organisms that are not usually visible to the naked eye. The division between prokaryotes and eukaryotes is one of the greatest evolutionary discontinuities of life (Boon *et al.* 2014). Familiar life on earth, such as plants, animals and fungi are all eukaryotes, however, the less obvious, and microscopic prokaryotes consists mostly of bacteria and archaea. The main difference between prokaryotes and eukaryotes is that nucleic material is not membrane bound in prokaryotes, while in eukaryotes nucleic material is found within a membrane bound nucleus. Prokaryotes are some of the smallest life forms on earth and range in size from 1 – 10 μm, while microbial eukaryotes, although larger than bacteria, are also tiny (10 – 100 μm). Margulis and

Schwartz (1988) and Boon *et al.* (2014) provide great introductions on prokaryotic and eukaryotic life.

Prokaryotic life has dominated the Earth's biodiversity since life first evolved 3.5 – 4.0 billion years ago (DeLong and Pace 2001). It represents the most diverse group of organisms on the planet (Figure 1), and they inhabit almost every niche on Earth (Dykhuizen 1998). The Bacteria domain dominates the majority of known prokaryotes, but a second domain, Archaea are also prokaryotic. These two domains differ from one another by a number of biochemical characteristics, such as ester-links in their lipid membranes (Boon *et al.* 2014). Bacteria, such as cyanobacteria, can photosynthesise, while others are chemotrophic and heterotrophic. See Madigan *et al.* (2011) for a review of bacterial diversity.

There are three main groups of eukaryotes which are important to microbial ecologists; algae, protists and fungi. Algae are ubiquitous across Earth, and estimated to be responsible for >40% of global photosynthesis (Boon *et al.* 2014). Microalgae are important primary producers in almost all aquatic ecosystems, where they contribute up

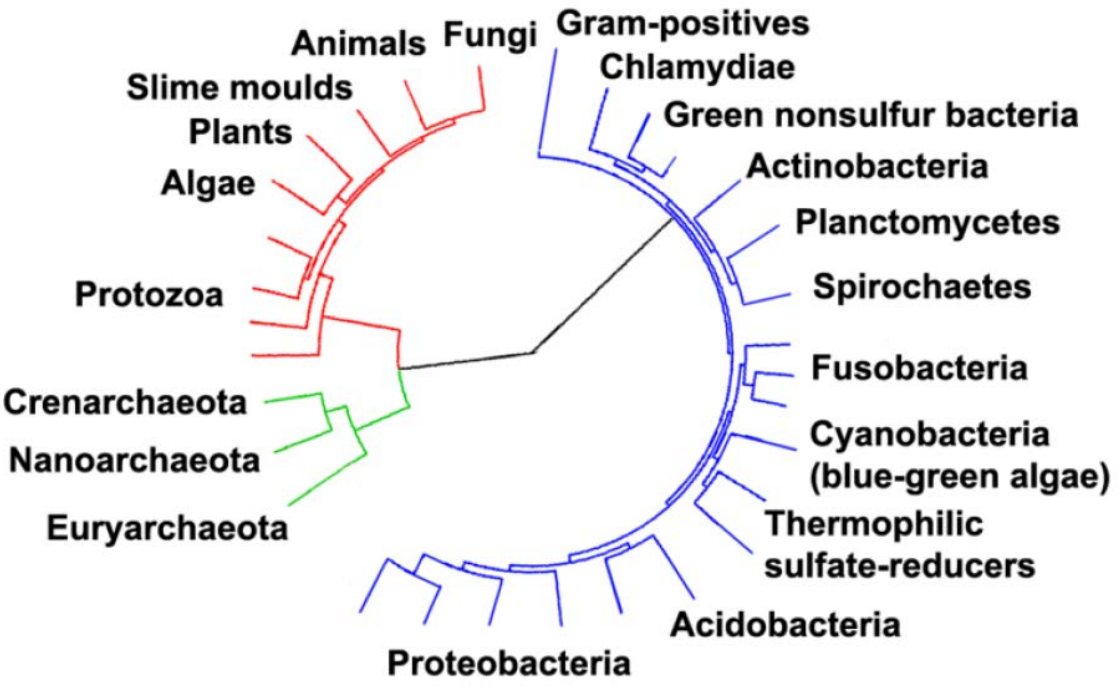


Figure 1: The phylogenetic tree of life displaying the three domains of life; Bacteria (Blue), Archaea (Green) and Eukaryotes (Red). Each branch represents phyla for the bacteria and a kingdom for the eukaryotes. The closer a branch, the more related they are. Adapted from Lee (2014).

to 30% of total production (Vymazal 1995). Fungi are found in aquatic systems as filamentous fungi, which dominate the littoral zone, and unicellular fungi, which dominate the pelagic zones (Boon *et al.* 2014). Fungi consume coarse particulate matter, such as leaves and wood, as well as other complex materials, however, their biogeochemical functions are not well understood but thought to have a smaller impact on biogeochemistry than bacteria (Wurzbacher *et al.* 2010).

Molecular approaches have recently revealed a large 'hidden world' of microbial eukaryotic diversity, mostly consisting of protists (Weisse 2008, Caron *et al.* 2009), bringing this important group of microbes into the forefront of modern ecological studies. Protists provide many ecological roles, including primary production and parasitism (Caron 2009) and it is thought that protists in aquatic systems are often the major source of much of the organic matter used by prokaryotes and other, multi-cellular eukaryotes. Many protists are believed to be ubiquitously distributed, an idea which is presently a topic of rich debate (Fenchel and Finlay 2004, Foissner 2006, Weisse 2008), arguments which are analogous to those regarding the distribution of bacteria (Horner-Devine *et al.* 2004, Martiny *et al.* 2005).

There are two main reasons why there is little understanding of scaling in microbial communities. First, conceptually it has long been assumed that microbial taxa have cosmopolitan distributions, as outlined by the Baas-Becking hypothesis, "*Everything is everywhere; but the environment selects*" (Baas-Becking 1934). Due to the small size and high abundance of microbes it has long been thought that dispersal capacities (distance and rate) are so great that they are effectively not subjected to any dispersal limitation at all (De Wit and Bouvier 2006). With such continuous, large-scale dispersal, microbes would have cosmopolitan distributions, and this is fundamentally different to the biogeographic patterns found in plants and animals (Fenchel and Finlay 2004). Second, it has been, and still is, technically very difficult to quantify microbial diversity. Prokaryotes, and many eukaryotic organisms, cannot be identified morphologically, but instead must be characterised and quantified using culture-based biochemical tests, or via molecular methods. Before the development of DNA-based methods, culture-based techniques were primarily used, which lead to severe underestimations of both microbial abundance and diversity in early studies (Hugenholtz *et al.* 1998). This is because it has been found that up to 99.99% of

bacterial taxa cannot be cultured under laboratory conditions (Torsvik *et al.* 1990). When DNA techniques were eventually applied to soil samples to look at bacterial richness, Tiedje *et al.* (1999) immediately found that microbial richness was orders of magnitude greater than what had been previously estimated using culture based techniques.

The '*everything is everywhere*' hypothesis is based on the premise that microbes have unique biologies because of their large population sizes, high dispersal capacity and low extinction rates (De Wit and Bouvier 2006). These characteristics have widely been assumed to be features of all microbes. Having a high capacity for dispersal is perhaps the most important factor in allowing microbes to have cosmopolitan distributions, and it has indeed been demonstrated that some bacterial taxa (Lighthart 1997, Leff *et al.* 1998, Gage *et al.* 1999), as well as protists (Finlay *et al.* 1999, Fenchel and Finlay 2004, Finlay and Fenchel 2004), have the potential for very high rates of passive dispersal. However, recent molecular techniques are beginning to reveal patterns in the microbial world that contradict the traditional cosmopolitan view of microbial diversity (Green *et al.* 2004, Horner-Devine *et al.* 2004, Bell *et al.* 2005a, Bell 2010, King *et al.* 2010, Caruso *et al.* 2011, Lear *et al.* 2014).

This introductory chapter outlines some fundamental ecological theories, such as taxa-area and distance-decay relationships, which have enhanced our understanding of biodiversity and distributions of higher organisms, and how they have recently been applied to microbial biogeography using molecular techniques. Furthermore, as it is becoming evident that similar spatial patterns exist in the microbial world as for macrobionts[1], theories based on neutrality using the meta-community framework will be discussed in order to understand the structuring mechanisms driving these patterns.

The second half of this chapter will identify the aims and hypotheses of this thesis, and provide a rationale for the methodology used throughout the study. This rationale will include a discussion on which groups of prokaryotes and eukaryotes should be investigated to test these hypotheses, as well as the type of study system that would be needed to test the generated hypotheses. Subsequently, Lake MacLeod will be

[1] In this thesis the term *macrobiont* will be used to denote multicellular macro-organisms.

introduced as a suitable system in which to conduct the study and an overview of its ecology will be given. Finally, the experimental design and statistical methodology will be introduced.

1.1.1 Taxa-area relationships

Perhaps the relationship between ecosystem size and species richness is one of the few universal relationships in ecology (Lawton 1999, Ricklefs and Lovette 1999, Lomolino 2001). MacArthur and Wilson's (1967) 'theory of island biogeography' provides the foundation of such a relationship, allowing ecologists to make predictions about species diversity within heterogeneous habitats, or 'islands'. Their theory specifically allows ecologists to make predictions on species richness with respect to an island's area by assuming "*smaller islands have extinction rates higher than large islands, for the same number of species present*" (Schoener 2010), implying that large islands will have more species than small islands.

When the number of species is plotted against island size, a species-area curve is constructed. This relationship is usually log-linear, where the log of island area is proportional to richness (MacArthur and Wilson 1967), and has more empirical evidentiary support than perhaps any other diversity relationship in ecology (Rosenzweig 1995). The slope of the species-area relationship (the z-value) for islands is typically between 0.2 and 0.35, however it is also taxon specific (Rosenzweig 1995). In general, a z-value less than 0.20 is indicative of organisms with high dispersal and low extinction rates and/or terrestrial islands such as lakes, high mountains or fragmented habitats (Connor and McCoy 1979).

There are at least three mechanisms which explain the species-area effect. First, large islands contain more individuals and thus have lower extinction rates (MacArthur and Wilson 1967, Reche *et al.* 2005). Secondly, larger islands have greater habitat heterogeneity, thus allowing them to support a greater variety of species (Ricklefs and Lovette 1999). And thirdly, the '*sampling effect*' could produce the species-area effect without the occurrence of a physical mechanistic process (Wardle 1999, Cam *et al.* 2002). This artefact can arise when larger islands, with greater species richness, is more likely to be dominated by a single, highly productive species, than a less species rich small island. Therefore, sampling more individuals from larger islands will result in the

larger islands appearing less rich (Connor and McCoy 1979, Rosenzweig 1995, Wardle 1999, Cam *et al.* 2002, Schoener 2010).

The theory of island biogeography is not exclusive to islands in a sea. In fact, it can be used to describe richness within any insular area. Rosenzweig (1995) defined an island as "*a self-contained region whose species originate entirely by immigration from outside the region*". Therefore the theory of island biogeography provides an ideal framework for investigating diversity related questions within lakes and ponds. Water bodies have been viewed as the islands in a terrestrial 'sea' by other ecologists (Dodson 1992), where the higher diversity found in larger lake systems (see Dodson 1992), has been explained using island biogeography mechanisms (Rosenzweig 1995, Brose *et al.* 2004, Schoener 2010).

Studies on North American lakes show that lake area alone accounts for more than 50% of among-lake variability in crustacean zooplankton diversity, with the effect of lake area being so strong that larger lakes contain more zooplankton species regardless of any other factors, including productivity or resource availability (Dodson *et al.* 2000). Cladocerans, fish, macrophytes and rotifers (Dodson *et al.* 2000), as well as benthic ciliates (Finlay *et al.* 1998), benthic microalgae (Hillebrand *et al.* 2001) and aquatic insects (Gaston 1992) have also been found to increase in species diversity with lake area. Dodson *et al.* (2000) found that copepods and phytoplankton showed no such increase in diversity with lake size, although microcosm studies by Smith *et al.* (2005) suggest that phytoplankton do display species-area curves.

Few studies have investigated species-area curves for microbes and other unicellular organisms. As stated above, the lack of studies in this area is probably attributable to the difficulty in measuring the richness of a microbial community. Finlay *et al.* (1998) provide perhaps the earliest account of the species-area relationship for a microbial taxon. They found unicellular protists to have extremely low z-values (0.043) and attributed this to their high dispersal ability. However, their study was limited to only detecting protist richness morphologically, and therefore probably underestimated community richness (Dopheide *et al.* 2009). Recent advances in molecular techniques now allow ecologists to measure genetic richness as a function of species richness, and

thus detect a large amount of previously undiscovered microbial richness (Fisher and Triplett 1999, Dopheide *et al.* 2009).

Horner-Devine *et al.* (2004) investigated the species-area curves for benthic bacteria in salt marshes using molecular techniques. The reported z-values are among the lowest recorded for any organism (Woodcock *et al.* 2006). The authors suggest that as bacteria are unlikely to be dispersal limited, the low z-values may be attributed to difficulties in resolving equivalent taxonomic units for bacteria as metazoans, the low habitat specificity of bacteria, and horizontal transfer of genes between phylotypes and ecotypes. However, the studies by Finlay *et al.* (1998) and Horner-Devine *et al.* (2004) suggest that bacteria and protists may obey similar ecological scaling relationships to macrobionts.

Bell *et al.* (2005a) investigated bacterioplankton diversity in tree holes ranging in volume from 0.04 to 24 litres, and similarly to the previously mentioned studies, found bacterial richness to increase according to the species-area law. They, however, found the z-value to be much higher than that found by previous studies, and suggest that microbes can show species turnover rates similar to those of plants and animals. Similarly high z-values were found by van der Gast *et al.* (2005) for bacteria inhabiting oil sump tanks. Continuing this trend, Reche *et al.* (2005) found a strong positive relationship between lake area and bacterioplankton richness. They suggested that because of the huge population sizes and high dispersal abilities of bacteria, immigration and extinction processes are probably insufficient to explain their findings. Instead, they suggested that the greater number of niches available within larger lakes is a better determinant of bacterioplankton richness. Other studies, however, have failed to detect a positive species-area relationship for bacteria (Humbert *et al.* 2009), even with the use of high throughput molecular methods (Logue *et al.* 2012). Further, the studies by Bell *et al.* (2005a) and Reche *et al.* (2005) have been contested by several researchers (Fenchel *et al.* 2005, Lindström *et al.* 2007) who argue that the respective sampling designs did not adequately address island biogeography hypotheses.

1.1.2 Distance-decay relationships

The similarity between two observations often decreases as the distance between them increases, a p attern which has been described by ecologists as the distance-decay

relationship. This negative relationship between distance and similarity allows analyses of spatial autocorrelation and is implicit in our understanding of species turnover along environmental gradients (Whittaker 1975, Cody 1985). There are two general mechanisms that can generate distance-decay patterns. Firstly, increased distances among samples will likely increase the degree of environmental difference (for example changes in geology and climate), and species-sorting processes will thus decrease the similarity of the communities (Nekola and White 1999). The second mechanism suggests that dispersal limitation of organisms produces the distance-decay pattern as the further the distance, the more effort required by an organism to move, or the more likely that a barrier will impede it. Therefore, if taxa are limited in their dispersal capabilities, samples that are closer will tend to have higher community similarity, and can generate a distance-decay pattern in the absence of environmental heterogeneity and species sorting processes.

The size and isolation (spatial configuration), the nature of surrounding habitats (spatial context) and time will all influence how genes move around landscapes (Nekola and White 1999). However, different species may perceive the spatial configuration and context differently, and this will lead to variation in the distance-decay rates among organisms. Further, if there is high resistance to movement across a landscape, i.e. dispersal or time are limiting, the nature of the spatial configuration of habitats will affect how long the effects of historic processes persist in structuring a community (Nekola and White 1999). Theories based on neutrality can be used to explain the generation of these patterns by taking into account the rates of immigration between samples (Hubbell 2001, Chave and Leigh Jr. 2002).

Distance-decay relationships have been reported in a number of microbial studies (Langenheder and Ragnarsson 2007, Soininen *et al.* 2007, Fuhrman *et al.* 2008, Bell 2010, Martiny *et al.* 2011, Astorga *et al.* 2012, Jones *et al.* 2012, Lear *et al.* 2013, Lepère *et al.* 2013, Barreto *et al.* 2014b, Lear *et al.* 2014). Nonetheless, the occurrence of the distance-decay relationship in microbes is still debatable and has not been detected in some studies, even with the use of molecular methods (Yannarell and Triplett 2004). Whether distance-decay patterns are driven by environmental heterogeneity or dispersal limitation remains unknown. For example, Lear *et al.* (2014) provide a case where distance-decay was found to occur at the extremely fine scales of

less than 20m within well mixed water bodies where dispersal limitation is unlikely to be the determinant, while, at a larger global scale, Sul *et al.* (2013) found dispersal limitation to be the primary determinant of the marine bacterial communities found in polar regions.

The distance-decay and taxa-area relationships have both been pivotal to the development of macro-ecology since MacArthur and Wilson (1967) described them. Recently, these concepts have been applied to microbial communities, allowing microbial ecologists to begin to investigate if the determinants of these patterns in microbes are similar to those of plants and animals. The prospect that microbes display taxa-area and distance-decay patterns similar to other organisms is exciting as it can allow for the discovery of mechanisms that may explain the ecology for all organisms. Furthermore, elucidating the mechanisms that create these biogeographic patterns may provide evidence for the hypothesis that all life on Earth is governed by universal ecological processes and allow a fuller understanding of the processes that regulate Earth's biodiversity. However, these two relationships are mostly explained by species sorting processes (environmentally determined) or neutral processes (determined by dispersal and/or extinction), and the 'everything is everywhere, but, the environment selects' argument remains central in contemporary microbial ecology papers (De Wit and Bouvier 2006). Given that inferring structuring processes from biogeographic processes is inadequate to increase our knowledge of microbial ecology, studies need to begin testing species-sorting and neutral processes in microbial communities (Hanson *et al.* 2012).

1.1.3 Species sorting

Consider a community of ecologically similar organisms (neglect predator-prey, parasite-host and mutualistic relationships), that are required to interact with one another, so that they have a negative effect on each other's success. Such relationships are competitive, and competition is expected to select for the single best-adapted organisms and exclude the others. Under such interactions, the spatial patterns of the community will begin to emerge where each species will tend to be found where it survives best (Bell 2001). Species-sorting, and other niche based process, predict community structure by assuming that competition for resources dictates the

assemblage of organisms able to coexist in an equilibrium state (Tilman 1982, Leibold 1995).

Species sorting processes have been central to the 'everything is everywhere, but, the environment selects' hypothesis because they suggest that dispersal is not limiting for microbes and that the environment is the primary determinant of whether a particular microbe is present or active in a particular habitat (Figure 2). There is strong evidence that species-sorting processes are important for structuring bacterial (Van der Gucht *et al.* 2007, Jones and McMahon 2009, Logue and Lindström 2010) as well as micro-eukaryote (Hambright *et al.* 2015) populations.

Recent studies, using the fine taxonomic resolutions that modern genetic methods allow, have shown that some microbes seem to exhibit biogeographic patterns that are unrelated to environmental conditions (Hanson *et al.* 2012). Gibbons *et al.* (2013) used a powerful deep sequencing method to investigate a marine bacterial community and

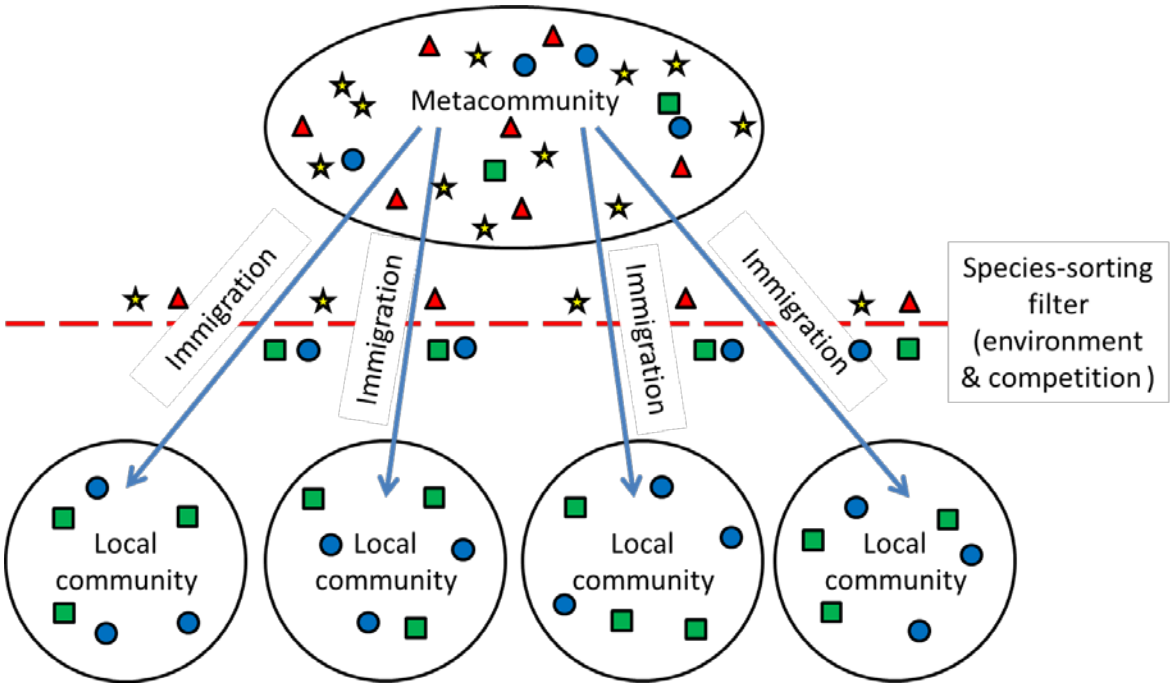


Figure 2: A simple conceptual diagram showing the role that species-sorting processes can play by providing a species-sorting filter (biotic and abiotic) which determines which taxa can inhabit the local communities. Different taxa are represented by different coloured shapes. This model shows that despite the green-squares and blue-circle taxa being uncommon in the metacommunity source, they are the only members of the metacommunity to be selected by the species-sorting filter and inhabit the local communities. In natural communities, it would be expected that each local community is subjected to unique species-sorting filters and thus have local communities which are different to one another. Adapted from Lee (2014).

found support for the Baas-Becking hypothesis. They found that global endemism does not exist, and that all community members were present at each location, but the relative abundances varied in orders of magnitude as the environmental conditions changed (Gibbons *et al.* 2013). Importantly though, Gibbons *et al.* (2013) suggest that studies that claim 'everything is everywhere' may only reflect technical limitations in detecting community members (Hambright *et al.* 2015), but the 'environment selects' part of the hypothesis determines which members are active components of the community.

1.1.4 Neutral processes

The unified neutral theory of biogeography and biodiversity (Hubbell 2001) is a recent theoretical addition to ecology that predicts community structure based solely on demographic stochasticity, immigration and speciation. Essentially, the concept of neutral models is that a larger community acts as a source of immigrants to a local community and that local community structure is the result of a balance between immigrants and local extinctions (Figure 3). The neutral models of Hubbell (2001) and Bell (2000) suggest that the abundance and distribution of organisms is predictable using the three parameters: immigration rate, local community size and a fundamental biodiversity index, where immigration rate can be estimated as the proportion of individuals that come from elsewhere in the community (Latimer *et al.* 2005, Etienne *et al.* 2006). Despite being built on s uch few assumptions, the neutral theory has successfully reproduced observed communities of macrobionts (Bell 2000, Hubbell 2001).

The neutral theory presented by Hubbell (2001) suggests that local communities within meta-communities are linked to one another through immigration and emigration. The rate of immigration for a species is determined by its ability to disperse. Local and regional richness and diversity are strongly affected by the level of dispersal (Mouguet and Loreau 2003), with diversity maximised at intermediate dispersal rates (Cadotte 2006). At low dispersal rates, stochastic extinctions and negative species interactions may maintain high beta and regional diversity (Figure 4), but low alpha diversity. Conversely, high dispersal rates allow dominant taxa to be introduced into all local communities, reducing alpha, beta and region diversity because of competition.

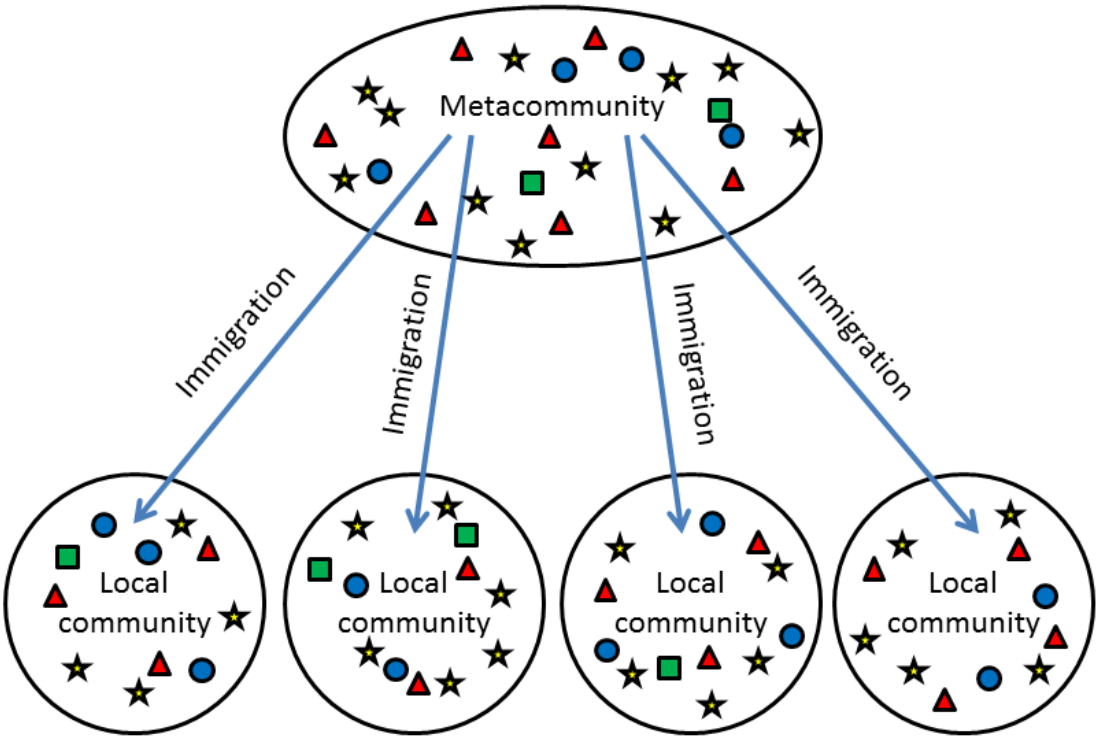


Figure 3: A simple conceptual diagram showing how neutral processes can determine the taxa that inhabit a local community. Following neutral processes, note that the relative abundances of the taxa in the metacommunity remains nearly the same in the local communties. Deviations in the relative abundances are due to extinction and speciation processes instead of environmental differences. Adapted from Lee (2014).

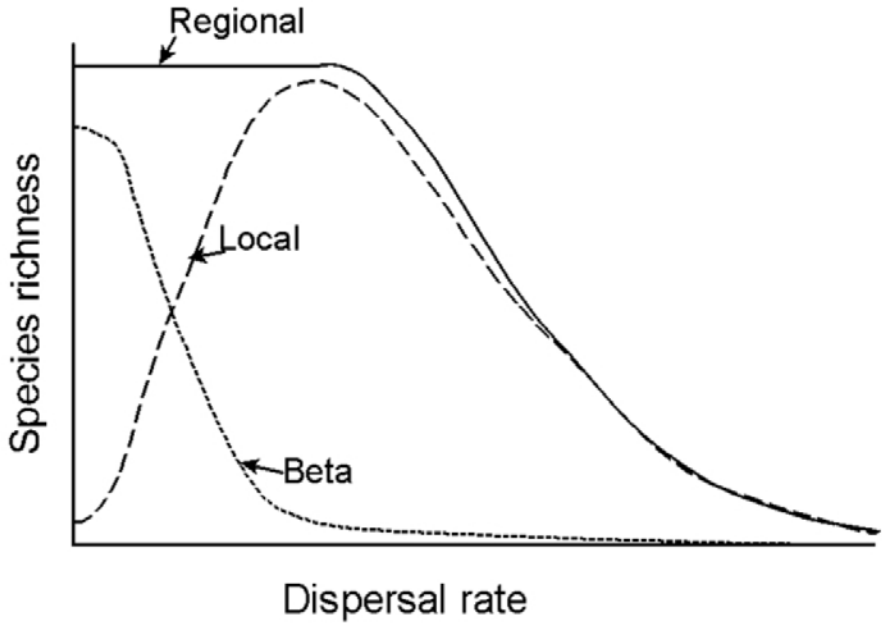


Figure 4: Hypothesised interaction between alpha (Local), beta and gamma (Regional) diversity on communities at different dispersal and spatial scales (adapted from Mouguet and Loreau 2003).

Neutral models have also been able to accurately predict the complex community structures of bacterial communities (Sloan *et al.* 2006, Sloan *et al.* 2007, Woodcock *et al.* 2007, Lee 2014), despite being dependent on only three parameters. Neutral theory predicts the relative and rank abundances of taxa within a local community to be the same as the source community (Langenheder and Székely 2011), and indeed, Sloan *et al.* (2006) have shown this to be the case for numerous bacterial communities. Sloan *et al.* (2006, 2007) provide some simple neutral models that predict the frequency by which members of highly abundant communities, such as microbial communities, are observed. However, studies testing the role of neutral processes in microbial communities have been limited as the models these studies used apply to single, dispersal limited local communities which are embedded in a meta-community (Alonso *et al.* 2006). Newer models, such as those developed by Etienne (2009), allow for testing of neutral predictions using discontinuous habitats, over larger spatial scales where dispersal limitation is likely to play a larger role. To date, these models have only been applied in a single study on global bacterial diversity (Caruso *et al.* 2011), where they found that different functional components of bacterial communities (photo- and hetero-trophic) were subject to different assembly processes depending upon the continent. Their results suggest that multi-trophic microbial systems may not be fully described by a single set of neutral assembly rules and that stochasticity is likely a major determinant of such communities.

1.1.5 Difference between microbial groups

This brief review on the patterns of microbial community composition and assembly has been written in the context of microbes, which is a term that includes both prokaryotic and eukaryotic unicellular organisms. Prokaryotic communities include bacteria and archaea, which are unicellular organisms that do not have a membrane bound nucleus (Walsh and Doolittle 2005). Microbial eukaryotes are also unicellular, but have a membrane bound nucleus. Microbial eukaryotes are mostly composed of organisms that belong to the protist kingdom (Walsh and Doolittle 2005). There are however, important differences in the biology of prokaryotes and eukaryotes. Bacteria, for example, are capable of reproducing asexually, whilst many protists, including ciliates, are capable of sexual reproduction. Horizontal gene transfer between bacteria is

thought to have a similar effect on genetic diversity as that of sexual reproduction in micro-eukaryotes (Ochman *et al.* 2005).

Clearly, recent studies suggest that bacterial communities have complex biogeographies, akin to those found in plants and animals (Horner-Devine *et al.* 2004, Bell *et al.* 2005, Martiny *et al.* 2005, Lear *et al.* 2014). However, in the review by Finlay (2002), it is suggested that, despite microbial eukaryotes being larger, and less abundant than prokaryotes, they are even less likely to display biogeographic patterns found in other domains of life because of their high dispersal rates. Recently though, small scale dispersal limitation was found in lake protist communities, suggesting that protists may not be ubiquitous (Rengefors *et al.* 2012). Studies are required to investigate whether microbial eukaryotes display similar patterns in species richness and distribution, as well as assembly processes, as other groups of microbes.

The unique biology of prokaryotes and eukaryotes could be expected to produce different biogeographic patterns. Conversely, similarities between microbial taxa, such as their small size and ability to disperse across large distances, should produce similar biogeographic patterns. Although there are studies investigating biogeographic processes on prokaryotes (Caruso *et al.* 2011, Lear *et al.* 2013, Barreto *et al.* 2014a, Livermore and Jones 2015) and eukaryotes, such as protists (Rengefors *et al.* 2012, Lepère *et al.* 2013) and fungi (Green *et al.* 2004), there are none directly comparing patterns between prokaryotic and eukaryotic microbes. By considering (vastly) different taxa within the same analysis, differences in landscape and environmental conditions can be used to directly compare the role of different factors in structuring the communities (Beisner *et al.* 2006).

1.1.6 Theoretical synthesis

It is clear that microbial communities display taxa-area and decay-distance relationships at a range of scales[2]. However, these patterns need to be studied using hypotheses generated from well-established ecological theories in order to understand the mechanisms driving these patterns. Species-sorting and neutral models provide a

[2] In this thesis, small-scales refer to distances in the centimetre to metre range. Large-scales refer to distances in the kilometre and beyond range.

framework where hypotheses can be generated and tested (Leibold 1995). Even though many studies have investigated the environmental determinants of microbial community structure, few studies have also incorporated neutral processes into their analyses (Caruso *et al.* 2011, Langenheder and Székely 2011, Lee *et al.* 2013), even though there they are likely to play an important role (Woodcock *et al.* 2007, Langenheder and Székely 2011, Lee *et al.* 2013). Future studies need to disentangle the roles of species-sorting and neutral processes if our knowledge of microbial ecology is to be on par with that of the macro world.

Furthermore, microbial ecology has tended to focus on prokaryotic organisms to test these ideas. Micro-eukaryotes, such as protists and fungi, are also part of the microbial component of ecosystems. Micro-eukaryotes are fundamentally different to bacteria because they have different metabolic requirements and usually represent a different trophic level. Nonetheless, they are still extremely small, unicellular organisms that are found in high abundances and diversity, albeit less than bacteria.

1.2 Aims of this study and hypotheses

This thesis has two main aims, under which some specific hypotheses have been developed and will be tested in the following chapters. The main aims of this thesis are to test if:

1. the biogeography of microbes is similar to that of macrobionts; and
2. Biogeographic patterns, such as TAR and DDR, and relationships with environmental and spatial drivers will be similar in prokaryotes and eukaryotes.

If microbes have biogeographies similar to macrobionts, it is expected that a distance-decay relationship will be observed. Distance-decay relationships provide evidence for changes occurring in the composition of a community that are caused by environmental gradients (species-sorting) or dispersal limitation (neutral theory), which are known to be determinants of community structure in macrobionts. Similarly, if prokaryotes and eukaryotes have different biogeographical patterns, or have different mechanisms structuring them (species sorting vs. neutral), there will be detectable differences in the distance-decay and taxa-area relationships. By addressing these two broad aims, knowledge on whether there is variation in the patterns and mechanisms influencing the

biogeographic processes for two abundant, diverse, unicellular microbial groups, which represent two domains of life, with very different cell biology and reproductive characteristics will be directly compared. Furthermore, this study will present an example of the biogeography of microbes, and the role of species-sorting and neutral processes on them by testing the following null hypotheses:

- prokaryotic and eukaryotic communities do not display distance-decay and taxa-area relationships;
- Isolated habitat patches do not contain microbial communities which are composed of different taxa or dominated by similar taxa with similar relative abundances;
- environmental characteristics and geographic location are not important determinants of microbial community composition;
- microbial communities are not randomly assembled; and
- immigration is not an important determinant of community composition using.

In order to test these hypotheses certain premises need to be met. Firstly, a suitable study system needs to be identified. The system needs to have suitable habitat patches incorporated in its landscape, which are replicated, and found at a range of sizes. This enables one to investigate the role of habitat/patch size on the biogeographic patterns of the organisms. Similarly, these habitats need to be arranged throughout the landscape in which there are no barriers for dispersal, thus ensuring that geographic distance become the primary determinant of dispersal limitation between patches. As historical events are likely to have affected the composition of contemporary microbial communities, it is also important to choose a system where the habitats are formed from the same geological event and are defined by the same geomorphological features, including hydrology. In essence, the system needs to be consanguineous (have the same origin, mineral and chemical composition) and consist of spatially isolated analogous habitat patches.

The second criterion that needs to be met is that suitable prokaryotic and eukaryotic microbial communities need to be selected. These organisms need to be abundant in the chosen habitats, and easily detectable. Measuring community composition of microbial communities mostly relies on the use of DNA analyses, and the chosen communities

will need to be able to be measured using comparable techniques to ensure that reasonable comparisons can be drawn. Further, by using DNA techniques, the two communities can be measured using a single sample, of which suitable environmental parameters can also be measured.

Selection of a suitable system and suitable microbial communities will assume that the experimental design can be rigorous enough to address the main aims of this thesis. In doing so, the study will test the role that species-sorting and neutral processes play in shaping the microbial world. These processes are central to understanding the biogeography and assembly of ecological communities.

1.3 Overall design to test hypotheses

Field studies in aquatic environments have played a large role in the development of microbial ecology in aquatic systems. This is partly because microbes are vital in aquatic communities as they drive the transformation and cycling of most biologically important elements in these systems (Newton *et al.* 2011), but also because aquatic habitats, such as lakes and ponds, provide ideal study systems for ecologists studying biogeographic patterns (Dodson 1992). As described earlier, lakes and ponds can be thought to represent a group of environmentally similar, insular, regional ecosystems that are embedded within a larger, landscape scale, meta-ecosystem. Ponds and lakes usually represent habitats of various sizes, and thus taxa-area and distance-decay patterns can be tested. Furthermore, ecologists can test hypotheses based on species-sorting and neutrality that might be driving microbial taxa-area and distance-decay patterns based on the spatial arrangement of the water bodies within the landscape (Declerck *et al.* 2011, Lepère *et al.* 2013, Lear *et al.* 2014).

Saline systems are unique habitats in that they usually occur in arid areas and are either a permanent or ephemeral aquatic system in an otherwise dry landscape. Often saline lakes are poor in plant and animal richness, but harbour abundant, diverse and unique microbial communities (Oren 2002). Saline systems are usually caused by subsurface hydrogeological processes that result in the creation of lakes scattered throughout the landscape. The lakes within a given saline systems usually have similar hydrochemistries, hydrologies and geomorphologies, and thus represent

environmentally similar habitats (Radke *et al.* 2002, Timms 2009, Castañeda *et al.* 2013). They also represent habitats which have shared geological and evolutionary histories, are exposed to the same disturbances (e.g. rainfall/drought and flood events) and are therefore ideal systems for testing the hypotheses of this thesis. The microbial meta-community is likely to be delivered into the ponds via the underground seepage. Preliminary results (Huggett *et al.* Unpublished) suggest the communities are primarily composed of marine taxa found in coastal, estuarine and deep sea environments.

A salt lake system, such as Lake MacLeod in north-western Australia, provides a suitable system in which to study the biogeography of microbes. The salt ponds in Lake MacLeod represent habitats formed at the same time by a single geological event during the early Holocene and have similar mineralogy and chemistry. They also occur at a range of sizes and are largely isolated from one another. Such a system allows for the testing of the hypotheses in this thesis.

Prokaryotic organisms have been well documented in salt lakes, and are an important component of aquatic ecosystems (Oren 2002, Casamayor *et al.* 2013). However, the diversity of other microbial taxa which inhabit these systems, particularly unicellular eukaryotes, is less well understood, thus detailed investigations into these environments are justified (Barberán and Casamayor 2011). Although extreme environments, such as salt lakes should have lower taxonomic richness based on general ecological principles (Frontier 1985), it has been found that the high salinities found in salt lakes may not reduce microbial richness (Casamayor *et al.* 2013). Furthermore, bacteria are not the only group of organisms thought to have high richness in saline systems. Some microbial eukaryotes, particularly ciliates, have recently been identified to have high levels of richness in salt lakes (Casamayor *et al.* 2013). Ciliates are also predators of bacteria, and thus represent a higher trophic level than bacteria (Finlay and Esteban 1998). Studying bacteria and ciliates in conjunction will allow for the study to test the hypotheses on two groups of microbes, which represent two taxonomically rich, trophically separated domains of life.

Because of the spatial arrangement and the consanguineous nature of the ponds, Lake MacLeod will be used to empirically investigate biogeographic theories in a natural system, and test the hypotheses generated above. Bacteria and ciliates are also abundant,

taxonomically rich and are important components of saline systems such as Lake MacLeod, and will therefore be used to analyse the differences between prokaryotic and eukaryotic microbial ecology. The benthic microbial communities form the same ecological functions across all the ponds and therefore allows for direct comparisons of communities between the ponds. Another important aspect of ecological studies concerned with spatial processes is that a rigorous and well developed sampling design needs to be employed. Although specific experimental designs and statistical analyses relevant to individual components of the thesis will be presented in the relevant chapters, the choice of ponds and the spatially explicit sampling design, which are important to the thesis as a whole, will be introduced in detail here.

1.3.1 Setting - Lake MacLeod

Lake MacLeod is a large inland saline lake located 20 km inland of the Carnarvon-Ningaloo coastline in Western Australia (Figure 5). The MacLeod lakebed is predominantly dry and covers an area of 2000 km^2, 60 km^2 of which is covered by numerous permanent bodies of brine[3]. The brine bodies persist due to a marine-water seepage-face that sustains brine inflow along the lake's north-western edge despite the huge evaporative outflow. Due to the continual seepage of marine water into the lake, Lake MacLeod contains areas of permanent saline wetlands and unique mangrove swamps (Ellison and Simmonds 2003, Dunham 2014), and it provides a major feeding habitat for migratory birds (Phillips *et al.* 2005, Bertzeletos *et al.* 2012).

Brine seepage occurs along seepage faces, of which there are five at Lake MacLeod, with four being active (Logan 1987). The Cygnet seepage face is the only one to discharge brine (originating from the Indian Ocean) freely to the surface to form the extensive ponds and brine sheets that are characteristic of the area (Logan 1987). The active part of the Cygnet seepage face covers 113 km^2, runs approximately 30 km along the north-western barrier base and extends up to 5 km eastward towards the basin interior (Shepherd 1990). The lake can be divided into two distinct systems, the northern Cygnet seepage face, or the Chirrida System, and the southern Cygnet seepage

[3] In this thesis, the term 'brine' is used to describe the water flowing through the ponds. At the vents, where brine enters the pond systems, the brine has a similar ionic composition to seawater.

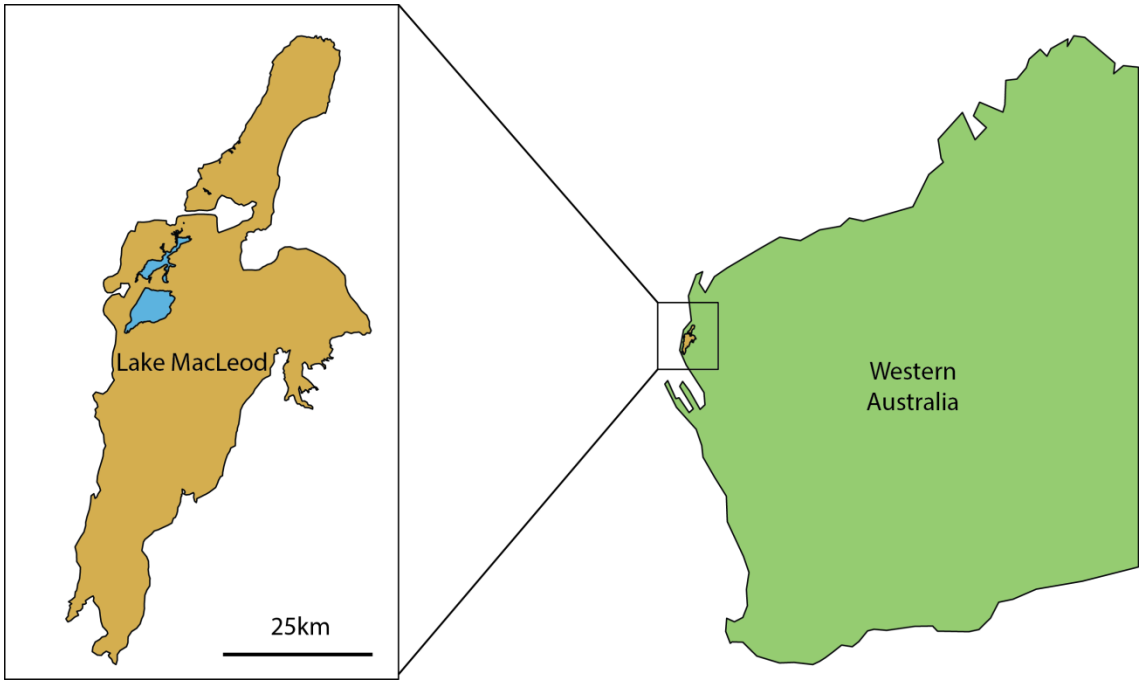


Figure 5: Map showing location of Lake MacLeod in Western Australia. Blue areas represent regions with permanent surface water. The brown area represents the dry lake basin. The two large blue areas in Lake MacLeod are Cygnet Pond and Ibis Pond, which is located south of Cygnet Pond.

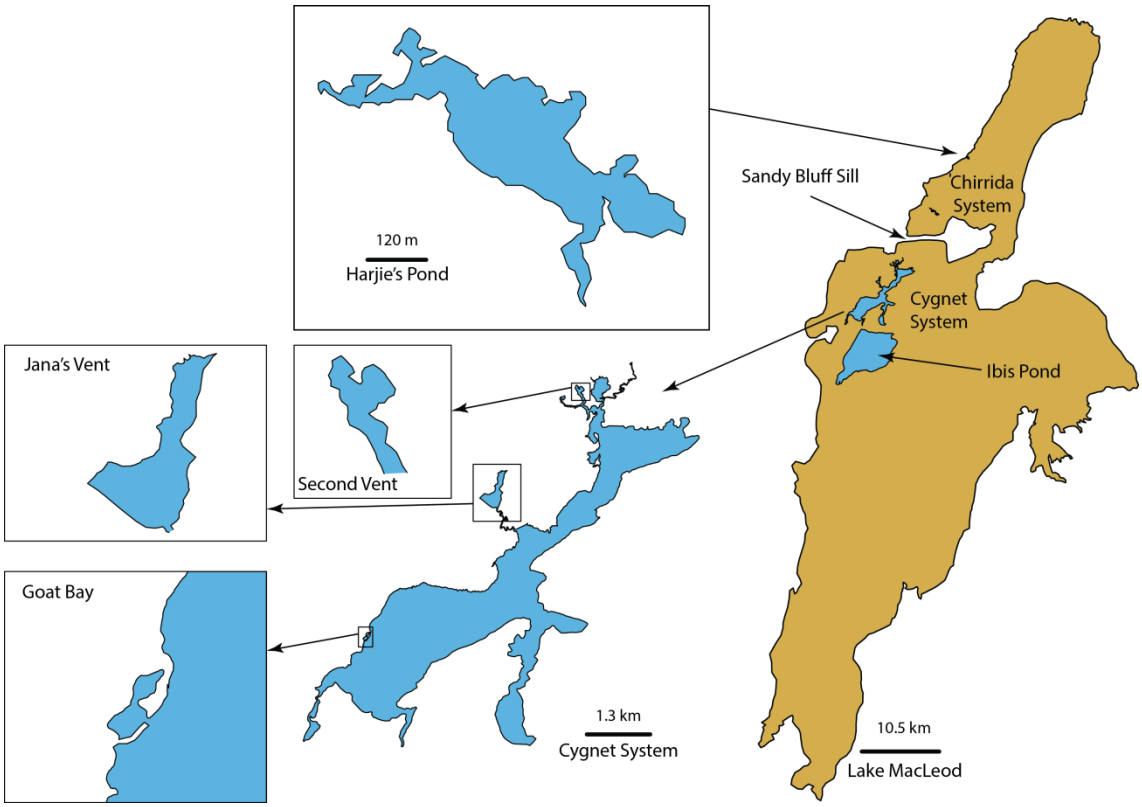


Figure 6: Map showing the locations of some ponds and landmarks of Lake MacLeod.

face, or the Cygnet System (Figure 6). The two systems are separated by an extension of the impermeable coastal barrier that extends eastward into the basin, known as Sandy Bluff Sill. Seepage is maintained by a hydrostatic head caused by the seepage points being up to 3.5m below sea level, and very high evaporation rates.

In the Cygnet System, seawater is discharged at numerous points along the western edge, flowing directly into Cygnet Pond and continuing southwards into Ibis Pond (Figure 6). Some of the larger vents, like Second Vent and Jana's Vent, consist of 1 to 2m deep ponds that flow through channels into Cygnet Pond. Other vents, like those at Goat Bay, open directly into the pond. Cygnet Pond is approximately 12 km^2 in area and has an average depth of 0.4m (Logan 1987). Brine flows south-east and over the south-eastern pond margin into Ibis Pond. Ibis Pond is the largest body of water in the basin, with possible depths of 1.5m and a surface area between 32 and 260km^2, depending on evapotranspiration.

The Chirrida System contains eight main discrete ponds, with each pond having a large vent and numerous smaller millimetre sized vents (Shepherd 1990). The largest of these ponds, Harjie's Pond, is over 300m long, whilst the smaller ponds are less than 10m in diameter (Figure 6). The ponds are mostly confined by mangrove roots and pneumatophores and accumulated sediment, preventing water from spilling over the banks. The ponds are shallow, rarely deeper than 1 m, and usually spill out over the eastern edges to form spill sheets. These spill sheets are extremely shallow and cover extensive cyanobacterial mats, which tend to flow south-east, although they can be moved throughout the basin by the wind.

Since the severe floods of 2000, the exotic fish, *Tilapia mossambicus*, has been recorded extensively throughout the ponds (Phillips *et al.* 2005). Invasive species are known to change important community characteristics of communities, such as, population sizes, community structure, disturbance regimes and biogeochemical processes (Vitousek 1990). *Tilapia mossambicus* has been found to clear large areas of aquatic macrophyte growth and form large breeding leks (Doupe and Burrows 2008, Cameron-Caluori 2014). Within the leks, males establish their territories by building large depressions in the sediment. These depressions have been found to be as large as

50cm in diameter. Cameron-Caluori (2014) provides a study of the effect of the *T. mossambicus* on the fish communities of Lake MacLeod. Although *T. mossambicus* were found to occupy the same habitats and resources as the endemic fish, *Amniataba caudavittata*, it is likely that the two species are able to co-exist because of their different feeding strategies.

A study commissioned by Dampier Salt Limited on the food web structure of Cygnet Pond (Streamtec 2003) found that mangroves did not contribute significantly to the aquatic food web, even though they are the largest source of organic matter at the lake. Instead, particulate organic matter in the ponds was found to be of aquatic origin, which was consumed by the amphipods which were consumed by the fish *A. caudavittata*. Similarly, Cameron-Caluori (2014) found the benthic biofilm to be a major component of the *T. mossambicus* and *C. pauciradiarus* diets. These studies highlight the important role that algae, diatoms and bacteria may be playing in the system.

The ponds are the most extensive aquatic habitat at Lake MacLeod, and benthic habitats are one of the most extensive habitats on the planet (Raffaelli *et al.* 2003). Benthic communities are important as they control biogeochemical processes, such as mineralisation and sedimentation, and have a profound effect on pelagic processes. The role of benthic processes at Lake MacLeod is evident by the apparent dependence of particulate organic matter on the Cygnet Pond food web (Streamtec 2003) and its contribution to the fish diets (Cameron-Caluori 2014). Benthic habitats are heterogeneous (Raffaelli *et al.* 2003), and this pattern is indicated in the Streamtec (2003) report by the great variability found in food web structure within Cygnet Pond. Microbes play a pivotal role in benthic processes, especially in tropical systems where they are more productive and efficient in biogeochemical processes than any other part of the world (Alongi 1994). Currently, no information is available on the diversity, functioning and spatial patterns of benthic biofilm microbial communities in Lake MacLeod. The importance of benthic microbes, and how they contribute to Lake MacLeod food webs is also unknown. Maintaining the health and the biogeochemical processes of benthic microbes is an important component of conservation and therefore, in Lake MacLeod, requires investigation to enhance management strategies.

1.3.2 Study design

Many previous studies investigating biogeographic patterns in microbial taxa have relied on collecting either a single, or very few samples to represent an entire aquatic habitat, such as a pond or lake (Reche *et al.* 2005, Yannarell and Triplett 2005, Pagaling *et al.* 2009, Romina Schiaffino *et al.* 2011). The few samples used in these studies are inadequate as no information is gained about the within habitat variation in community composition. In order to account for the effect of environmental heterogeneity, an adequate number of samples need to be collected which captures the variation within the habitat/pond. Similarly, if spatial processes are to be investigated, samples need to be collected in a way which accounts for the spatial extent of the habitats being investigated. This section describes the sampling design used in the following chapters, from the choice of ponds and sampling design, to the general statistical methods used to answer the questions, although specific protocols and statistical methodology will be presented in their respective chapters.

Measuring habitat, environmental and community variation

The ponds have different physical attributes, hydrodynamics and, although they are all seawater derived saline wetlands, are also likely to vary on a theme by having slightly different hydrochemistries and nutrient levels. These differences in habitat structure are likely to influence the microbial communities found in them (Radke *et al.* 2002, Radke *et al.* 2003, Boggs *et al.* 2006, Boggs *et al.* 2007, Long *et al.* 2009). It is therefore necessary to determine the morphometric characteristics of the salt lake systems, primarily the sizes, depth and the hydrodynamics of each pond. The morphometric and hydrodynamic characteristics could also be expected to cause variation in the water chemistry, primarily because of differences in residency time and evaporation. Water chemistry parameters that are likely to differ between ponds include ionic concentrations and nutrient concentration, both of which have been shown to alter the microbial communities of salt lakes (Radke *et al.* 2003, Casamayor *et al.* 2013). This thesis will provide an analysis of the morphometric and hydrodynamic characteristics of each pond, before an investigation of the ionic and nutrient concentrations is presented. These two sets of data will be fundamental to characterising the environmental variation

that occurs across the study ponds and is likely to be important in structuring the microbial communities.

Community variation will be measured using fragment analyses targeting gene regions which are commonly used in studies to fingerprint communities. Automated ribosomal intergenic spacer analysis (ARISA) is a popular method used for fingerprinting bacterial communities, and has a relatively high taxonomic resolution compared to other DNA methods (Ranjard *et al.* 2001). Unfortunately, ARISA cannot be used to profile ciliate communities, therefore terminal restriction fragment length polymorphism (T-RFLP) will be used (Dopheide *et al.* 2009). These two methods are methodologically very similar, except different gene regions will be targeted for the two groups which have been shown to provide the maximum taxonomic resolution.

Pond selection

In order to account for the range of variation in pond characteristics, it is important that this study is done in ponds from a range of spatial scales. The different sized ponds represent different ecosystem sizes, and will allow for the taxa-area relationship to be calculated. For this reason, eight ponds were chosen that represent the change in spatial scales of the ponds found at Lake MacLeod. Ponds were also chosen for their accessibility. The ponds chosen for this study were (from smallest to largest): Pete's Vent (95m^2), Donut Pond (6,000m^2), Annie's Pond (6,700m^2), Pete's Pond (14,900m^2), Whistler's Pond (46,500m^2), Harjie's Pond (65,100m^2), Jana's Vent (187,500m^2) and Cygnet Pond (7,147,700m^2; Figure 7), representing ecosystems ranging in size from tens of square metres to square kilometres.

Field sampling design

In order to account for within pond va riability of the environmental conditions and microbial communities, multiple samples need to be collected within each pond. These samples need to be collected in a manner which is spatially explicit. This requires that the samples are collected in an arrangement where the separation distances between samples ranges from fine (centimetres) to broad scales (dependent on pond size). To

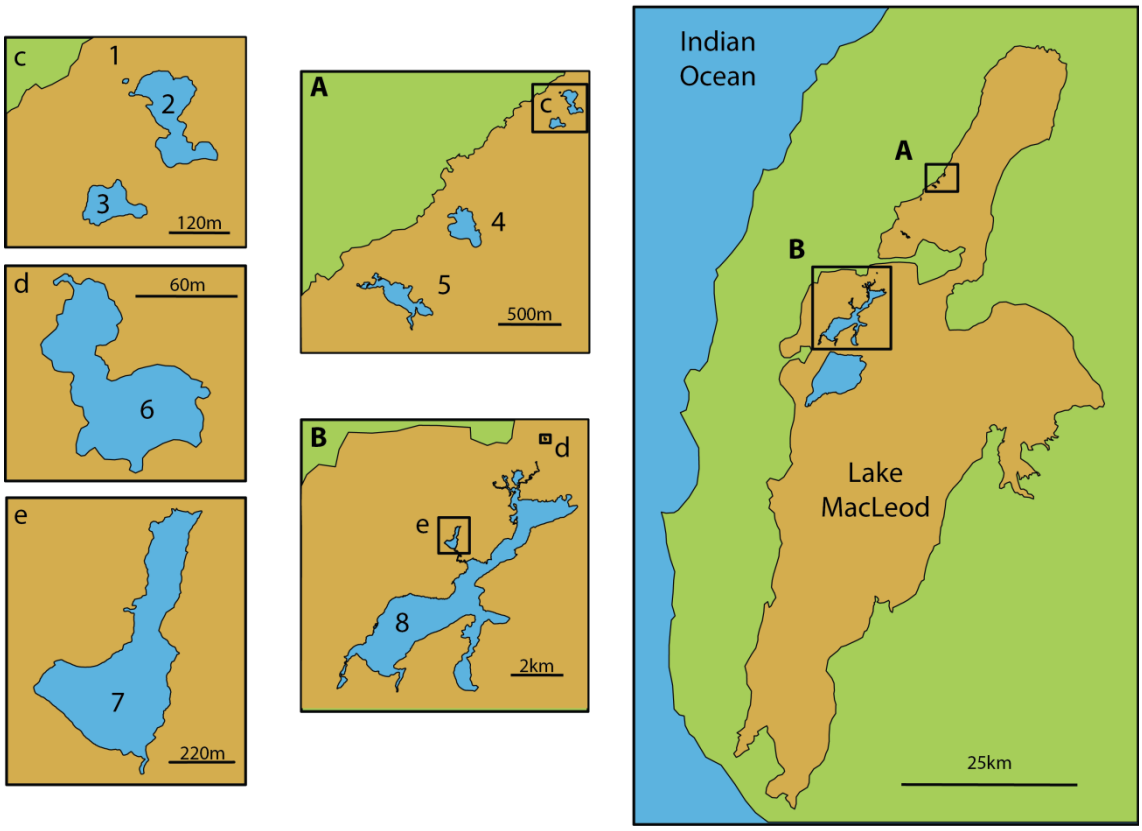


Figure 7: Map of the MacLeod basin showing the location of the ponds. Box A shows the location of the northernmost five ponds which include Pete's Vent (1), Pete's Pond (2) and Annie's Pond (3; box *c*) as well as Whistler's Pond (4) and Harjie's Pond (5). Box B shows the locations of Donut Pond (6), Jana's Vent (7) and Cygnet Pond (8).

achieve this, the samples must be collected in a design which allows for these separation distances, but also enables the location of each sample, with respect to the other samples, to be determined. Such a sampling design will account for small scale variation in the environment and the communities.

Using transects to define where samples are taken allows for the easy collection of geolocated samples, particularly at spatial scales where GPS does not provide adequate resolution. An adaptation of the sampling design used by Horner-Devine *et al.* (2004) was ultimately used in this study. Samples were collected from each pond along transects. In each pond four transects were established, with each transect originating from randomly selected point and orientated in a random direction. The length of the transects ranged from 3m in Pete's Vent to 2000m in Cygnet Pond (Table 1). Samples were collected at predefined distance along each transect, ranging from separation distances of 0.1m to the length of each transect. Care will be taken to only sample the microphytobenthic communities, also known as biofilm communities, by sampling the

top 5mm of sediment only. Hence, this spatially explicit sampling design can account for biotic and abiotic structures at scales ranging from 0.1m to thousands of metres. Further, this method allows for the creation of a data set where the location of each sample point in space is known, and the distances between samples can be determined. This information is fundamental when investigating the effect of spatial structures on ecological communities as it allows for differences in the distance where communities cease to be spatially autocorrelated to be measured. These distances will provide evidence of the spatial scale at which structuring processes are operating within a given pond. Comparing autocorrelation distances between the ponds allows for the effect of pond size on community structure to be determined.

1.3.3 Statistical approach

This section describes the statistical methods used in the following chapters to test for differences in the environmental and community structure of each pond. Details on the spatial analysis are also given to provide an understanding of how the data collected using the field sampling design were analysed. These analyses, unless otherwise stated, were all conducted using Euclidean distances. In order to use linear methods, such as principal components analysis (PCA) and redundancy analysis (RDA), community data were Hellinger transformed (Rao 1995). For each sample, Hellinger transformation is computed as the square root of the quotient of the abundance value for each species and the total abundance (Borcard *et al.* 2011). This transformation is recommended for clustering and ordination of community datasets (Legendre and Gallagher 2001). Specific details of how each analysis is used will be given in the relevant chapters.

Table 1: Summary of design used to collect samples in the ponds. Each pond had four transects established in random directions, and a variable number of samples collected at set intervals along each transect. The length and number of samples collected per transect increased as the size of the ponds increased.

Pond	Samples per transect	Length of transect (m)	Pond size (m^2)
Pete's Vent	12	3	95
Donut Pond	12	15	6000
Annie's Pond	12	15	6700
Pete's Pond	12	15	14,900
Whistler's Pond	16	30	46,500
Harjie's Pond	16	30	65,100
Jana's Vent	16	30	187,500
Cygnet Pond	22	2000	7,147,700

Non-metric Multidimensional Scaling

Non-metric multidimensional scaling (nMDS) is a method that can produce an ordination using any distance matrix. Because it can handle data with missing values, or zeros, it is ideal for producing ordinations of community data. Exact distances, however, are not preserved in nMDS ordinations, but rather represent the relationship in rank similarity between samples in two dimensions (Borcard *et al.* 2011). The stress value indicated the degree of uncertainty within the ordination, with values greater than 0.20 indicating an ordination which does not accurately portray the similarities/dissimilarities between points.

Principal Components Analysis

Unconstrained ordination analyses use a single data matrix and reveal its major structure as a reduced set of orthogonal axes which can be used in ordination. These methods decompose the total variation in the data into a set of linear components which can be used as axes for ordination. Principal component analysis (PCA) is an unconstrained ordination method of a data matrix consisting of normally distributed, quantitative variables. It is therefore not suitable for analysis of raw community data, unless used with an appropriate pre-transformation, such as the Hellinger transformation. PCA derives a maximum of p axes from a data set containing p number of orthogonal variables, which are commonly ordered representing decreasing contribution to the total variance (Borcard *et al.* 2011). The principal components give the positions of the objects (samples when using community data) in the new set of coordinates.

Redundancy Analysis

Redundancy analysis (RDA) is a method that allows the variation in a set of response variables (community data) that can be explained by a set of explanatory variables to be extracted and summarised. RDA can be considered as a constrained version of PCA, where constrained (or canonical) analyses associate two or more data matrices into the ordination processes. This allows for structures from a data set to be associated with structures in other data sets, and formally test hypotheses about the significance of these relationships. RDA is essentially a multivariate multiple linear regression analysis of the response data (e.g. community matrix) with the explanatory data (e.g. environmental or

spatial data), followed by a PCA of the fitted values. RDA therefore produces, in successive order, a series of linear combinations of the explanatory variables that best explain the community data. The axes define the space of the explanatory variables, and are orthogonal to one another. This allows for testing the hypothesis that the relationship between the explanatory variables and the community matrix is linear. The canonical axes which are produced by RDA, are linear combinations (multiple regression model) of all the standardised explanatory variables. Variation which cannot be constrained (i.e. the residuals of the regressions) is expressed as unconstrained PCA eigenvalues. Permutation tests are implemented to test the null hypothesis that no linear relationship exists between the community data and the explanatory variables (Borcard *et al.* 2011, Legendre *et al.* 2011).

Permutational Multivariate Analysis of Variance

Permutational multivariate analysis of variance (PERMANOVA) is used for testing the simultaneous response of one or more variables to a factor in an analysis of variance using permutational methods. PERMANOVA is directly analogous to a multivariate analysis of variance (MANOVA), in that it partitions the sums of squares of a multivariate data set (Anderson 2001). PERMANOVA can, however, use distance matrices as the response data, which is important when ecologists wish to test for significance between groups of community data, thus also making it analogous to distance-based RDA (Legendre and Anderson 1999). PERMANOVA assumes that the groups have the same multivariate spread among the factors being tested. This requires the data to be tested for homogeneity of variance. In this thesis, PERMANOVA is used to test for the effects of factors, such as pond identity, on microbial community structure.

Analysis of Similarities

Because the assumptions of PERMANOVA can often be difficult to meet, analysis of similarities (ANOSIM) was also used to test for the effect of factors on the microbial communities. This method is non-parametric, in that dissimilarities are ranked in order, similar to nMDS. The method produces an R-statistic which can be tested for significance using permutational tests (Clarke 1993). The R-statistic ranges from -1 to +1, with a value of 0 indicating the factors are random. Typically, ecologically

meaningful values lie between 0 and +1. Values close to +1 indicate that within group similarity is greater than between group similarity.

Variation Partitioning – Spatial Analysis

The structure and spatial patterns of the microbial communities detected in this study will be the result of environmental, spatial and unaccounted sources of variation. Disentangling the relative contributions of environmental and spatial mechanisms has been a focus of recent studies (Borcard *et al.* 1992, Beisner *et al.* 2006, Langenheder and Ragnarsson 2007, Caruso *et al.* 2011, Lear *et al.* 2014) and is important when trying to understand the role of deterministic and stochastic processes. If environmental variables explain much of the variation among communities, species-sorting processes can be assumed to be the dominant determinant of community structure. Similarly, if spatial factors explain significant proportions of variation in community structure, then neutral processes related to dispersal limitation and/or immigration can be assumed to be important determinants of community structure. This thesis will investigate the role of the environment and space by using variation partitioning methods to explain the amount of variation each source explains independently of one another. Specifically, this analysis can test the following:

- if there are gradients of change in microbial communities (indicative of processes operating at a larger scale than the sampling design);
- if there are spatial structures of both biotic and abiotic variables (identifiable patterns occurring at scales finer than the sampling area); or
- if there is a random component of the variation that indicates effects operating independently at each sample.

The first two data chapters of this these are primarily concerned with explaining and understanding the variation in the physical, hydrological and chemical components of the ponds. The outcome of these chapters is to produce a set of geolocated environmental determinants that can be used to explain the microbial community patterns. This set of determinants, combined with the spatial data (location of each sample in space) allows for an analysis into the relative contributions that each source of variation contributes to the community data.

Early methods for integrating space as a predictor variable into ecological models were based on Mantel regression (Mantel 1967) and partial Mantel tests (Smouse *et al.* 1986). However, Mantel tests only detect significant relationships between two or more similarity or distance matrices and do not provide information on the fraction of total variation explained by environmental and spatial matrices independently or the amount of variation related to both sets of explanatory variables (Borcard *et al.* 1992). However, ordination techniques, in particular RDA, do allow for the testing of two or more sets of explanatory variables (e.g. environmental and spatial data) on a response set of data (usually a community matrix when dealing with ecological data) in combination and independently of one another (Figure 8).

Using constrained and partial ordination techniques to partition variation in community data instead of partial Mantel tests allows one to account for the spatial component of the community and environmental data. This is particularly important for those concerned with spatial patterns in ecological data because it allows for the spatial component of communities to be accounted for, and the independent roles of environmental and spatial processes to be determined. Ordination methods also provide a measure of how much variation is not explainable by the datasets, which is either real stochasticity and/or unexplained variation from unknown determinants.

The use of variation partitioning has become a common method to investigate the roles of deterministic and spatial processes on microbial communities (Beisner *et al.* 2006, Langenheder and Ragnarsson 2007, Caruso *et al.* 2011, Lear *et al.* 2014). However, the amounts of unexplained variation in these studies are usually high (more than 50%) and these analyses cannot discern potentially explainable variation from unexplainable

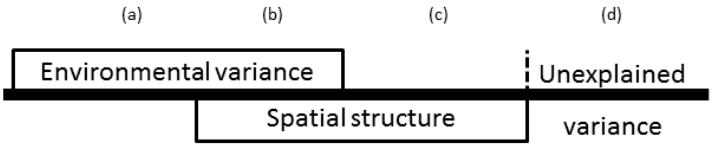


Figure 8: Variation partitioning of an ecological community matrix showing the fraction due to only environmental sources (a), both environmental and spatial (b), spatial sources (c) and residual variation (d). Sourced from Borcard *et al.* (1992).

variation due to stochastic processes. This can be expected as it is not possible to measure all the environmental factors that are possible determinants of community structure. However, by considering more than two spatial variables (e.g. northings and easting), one can ensure that as much of the spatial variation of the community data as possible is explained (Borcard *et al.* 1992).

Because this thesis is concerned with processes which are occurring at a range of scales, it is important to account for these different spatial scales. This means that the spatial variables need to model structures at the broadest scales (the entire sampling area), down to the finest scales, which is the 0.1m sampling interval. To achieve this, spatial variables that represent the structures of all relevant scales must be constructed. This is what Moran's eigenvector maps (MEM), and principal coordinates of neighbour matrices (PCNM) do (Borcard and Legendre 2002, Borcard *et al.* 2004, Dray *et al.* 2006).

MEM analysis allows for the modelling of spatial patterns by using a matrix of Euclidean distances among samples to build a truncated matrix which retains only the distances of close neighbours (defined by the smallest distance where all points remain connected by links smaller than or equal to the truncation distance). Sample pairs connected at distances greater than the threshold receive an arbitrary 'large' distance that corresponds with a distance four times the threshold. Principal coordinates analysis (PCoA) is used to compute the eigenvectors from the truncated distance matrix, with only eigenvectors corresponding to positive eigenvalues retained. These eigenvectors, or MEM variables, are retained and are interpreted as representing positive spatial correlation (Dray *et al.* 2006).

1.4 Thesis organisation

Although not an exhaustive summary of the analytical and statistical methodology that will be used in this study, this introduction provides the rationale and the foundation for the sampling approach used in the following chapters. There are four main data chapters presented in this thesis that address questions on the environmental variation found within the ponds of Lake MacLeod to a detailed biogeographic analysis of the bacterial and ciliate communities. **Chapter Two** is an investigation of the morphometrics and

hydrological regimes of each pond, **Chapter Three** provides information on the habitat heterogeneity found within the ponds, focusing primarily on variation in water chemistry and sedimentology characteristics for the benthic biofilm habitat. **Chapter Four** and **Chapter Five** link the variation in habitat characteristics found in Chapter Two and Three with the benthic bacterial and ciliate communities respectively. These two chapters will investigate, specifically, the biogeography of the two microbial communities using the theoretical frameworks and statistical methodology discussed in this chapter. The final chapter, **Chapter Six**, will provide a synthesis of the results, particularly concentrating on the differences and similarities of the bacterial and ciliate communities.

CHAPTER 2. POND MORPHOMETRY AND HYDRODYNAMICS

2.1 Introduction

Saline lakes are important features of the landscape in arid countries such as Spain (Mees *et al.* 2011, Casamayor *et al.* 2013, Castañeda *et al.* 2013) and Australia (Timms 2005). Salt lakes have unique ecologies because of their extreme conditions, such as variable and often very high salinity, high solar radiation, high temperatures and alternating drought and flooding periods. These extreme conditions result in the formation of unique habitats, which in turn enable distinctive assemblages of plants, animals and microbes to inhabit them (Timms 2009). Like other wetland systems, geomorphology, climate and hydrology are fundamental drivers of these systems, and understanding the way they operate and are expressed in a particular system is important when determining ecological processes (Mitsch and Gosselink 2000).

De Deckker (1983) describes a categorisation of salt lakes based on geomorphological grounds: i) large closed basins with (often) extensive internal drainage areas; ii) small closed basins with small internal drainage areas; and iii) crater lakes. A fourth category, coastal dune barrier lakes, is excluded because they do not conform to the notion of athalassohaline lakes. The term salt lake, or evaporite lake (also known as a salina), refers to these landlocked water bodies which have concentrations of salts and other dissolved minerals that are considerably higher than most lakes, and often higher than seawater (De Deckker 1983). A rarer class of salt water bodies contravene most of these general categorisations of salt lakes, under the following circumstances seen in the Northern Ponds at Lake MacLeod. In arid (hot and dry) subtropical coastal regions, where marine water travels through karst (Tertiary and Pleistocene age barriers; Wyrwoll *et al.* 2000), and upwells into a basin that is slightly below sea level, inland ponds can exist in a s tate with permanent seawater, and with relatively constant salinities.

Nutrient dynamics, planktonic biomass and biological community structure of lakes and ponds, are known to be affected by the hydrological characteristics of the system, such

as water depth, evaporation rates, temperature and water residence time (Olding *et al.* 2000, Baranyi *et al.* 2002, Rennella and Quiros 2006). These hydrologic characteristics are, in turn, influenced by the morphology and morphometric properties of the water bodies, of which little is known for the ponds of Lake MacLeod. Only crude estimates of surface area and volume are given by Shepherd (1990), and there is no understanding about the rate of water loss, or residence time of these systems. It is important to document these morphometric properties in order to understand how the physical differences in pond structure, such as size and residence time, determine the environmental characteristics, such as salinity and nutrients, which in turn might determine the biotic assemblages.

Lake morphology can have consequences for water temperature and dissolved oxygen, as well as other abiotic factors. For example, larger ponds may have increased fetch which contributes to increased wave action and mixing (Jackson *et al.* 2001, Wetzel 2001). Furthermore, as the surface area-to-volume ratio (SA:V) increases, so does the effect of evaporation on the concentration of ions. The mean depth of lakes has been shown to be a good determinant of fish and invertebrate community composition, as well as a good correlate for numerous limnological characters relevant for biological communities (Mehner *et al.* 2007, Jyväsjärvi *et al.* 2009).

Some studies have found a relationship between lake 'morphotypes', aquatic communities (Jackson and Harvey 1993, Mehner *et al.* 2007) and environmental parameters (Jackson *et al.* 2001, Jyväsjärvi *et al.* 2009). However, the morphometric characteristics of the Northern Ponds of Lake MacLeod, as well as the habitat structure and hydrochemistry found in these environments is not well understood, and it is unknown if they can be grouped into 'morphotypes' based on the physical attributes.

The ponds represent the areas of greatest seawater inflow into the MacLeod basin, with the vents marking the entry point for brine and the ponds providing a topographically enclosed area where evaporation acts, altering the water chemistry as water moves away from the vent. This hydrodynamic feature forms an integral part in the formation of aquatic habitats within the Northern Ponds. However, each pond has different physical features, such as shape, size and depth. Some ponds appear to have large vents where inflow seems to be high, whilst others seem to have no distinct vents. Similarly, some

ponds have relatively large outlet channels where water flows at a high rate out of the systems to form spill sheets, whilst in other ponds there is no clear out flow channels. The most comprehensive work attempting to classify the environmental units found in the varied landscape of the Northern Ponds of Lake MacLeod has been done by Shepherd (1990) in his thesis about the hydrological environments in Lake MacLeod. The work done in this thesis broadly follows the environmental unit classification scheme that Shepherd constructed.

The main distinction between the environmental units drawn by Shepherd (1990) was whether they belong to the brine sheet environment, where brine inflow is greater than evaporative outflow, or the majana environment, where evaporative outflow is great enough to supress brine to below the surface. The brine sheet environment was divided into four sub-environments, the marginal sheet, pond, deep sheet and spill sheet sub-environments. In Lake MacLeod, the brine sheet environments can be permanent or temporary, often encompassing discrete areas isolated from each other. The water bodies range in size from the square centimetres to the square kilometres and reach depths from centimetres to hundreds of centimetres. Within the discharge domain, brine is supplied directly from seepage, often through vents. It is in this domain that permanent bodies of water (ponds) are often found in deep depressions. These pond areas have a long-term positive water balance because seepage inflow is directly from a vent and the water body is in equilibrium with spill over and evaporation. Shepherd (1990) recognised that each of these environmental units consisted of unique biotic communities.

This chapter will investigate the morphology and morphometric properties of the eight ponds used in this study and investigate the similarities and differences between the ponds. Various properties of the ponds will be described, such as area, volume, perimeter length and water depth, as well as hydrodynamic properties, such as residence time and water temperature. The ponds will also be classified into ‘morphotypes’ in order to understand the variability in pond structure found at Lake MacLeod, and also the context for the hypotheses tested in Chapter 3.

2.2 Methods

2.2.1 Climate Data

Climatic data for the study period, January 2012 to December 2014, are presented. This three year period covers all months in which work was conducted in Lake MacLeod. These dates do not reflect individual sampling times. Time of sampling for each specific set of data is detailed in the relevant methodological sections. The variables include mean minimum and maximum temperatures, mean daily precipitation, mean daily evaporation and mean daily wind run for the Carnarvon Airport weather station. The weather station is located approximately 100kms south of the Northern Ponds. All data were retrieved from the Australian Government Bureau of Meteorology (2015) website. Mean daily wind speed (km/hr.) was calculated by dividing the mean daily wind run for each month by 24 hours.

2.2.2 Bathymetric and temperature profiles

Field

The use of a sonar measurement device has become a well-established method for many marine applications where an accurate water depth measurement is required. Many modern sonar devices also have the advantage of being able to continuously measure water depth at high resolution while traveling at speed. These advantages have led sonar technology to become the most popular method for reliably measuring water depth for marine applications like pond floor mapping (Morgan 2010, Coggins *et al.* 2014).

Fish-finding instruments are typically utilised in the fishing industry to measure water depth, as well as location. These units are relatively cheap and combine GPS technology with sonar to provide real-time depth-location data. The Lowrance HDS-5 Fish Finding Sonar and GPS unit was chosen for this study due to its cost, the sensitivity of the sonar device and the data-logging function. For each location the GPS co-ordinates, a depth and temperature measurement are recorded.

The fish-finding unit was calibrated for shallow water operation, as recommended by the manufacturer. This setup included setting the ping (pulse) speed to maximum and selecting the 200kHz transducer. These settings give the unit a resolution of ±0.01m

with a minimum depth reading of 0.30m. Data was logged using the unit's maximum settings of 3200 Bytes per ping and recorded in LS2 format.

Bathymetric measurements were made for each of the eight ponds in July 2013. A floating device to be towed as close as possible to the water surface was built using a surfing body board. The sonar transducer was attached beneath the body board and was towed approximately 1m behind the fish finding unit. The unit's computer, where the GPS is located, was kept at the rear of a boat, in order to ensure the GPS co-ordinates represented the depth measurements. The boat was a small inflatable which was manoeuvred with a small electric motor (with the exception of Cygnet Pond, where an aluminium dinghy was used with a small petrol motor). Travel speed was kept to a minimum in order to reduce the amount of displaced water at the rear of the boat.

Statistical interpretation

Depth and temperature data were downloaded from the Lowrance HDS-5 unit and run on the Lowrance software SonarViewer (Version 2.1.2, dow nloadable from http://www.lowrance.com/en-AU/Support/Sonar-Log-Viewer). Basic variography was conducted using ISATIS (Geovariances and Paris 2012), a geostatistical program that allows for calculation, inspection and manipulation of semi-variograms. Omnidirectional variograms were calculated and the lag distance and number of lags altered in order to achieve a smooth semi-variogram that could be modelled. Modelling was done automatically, with the program being able to select a nugget and multiple Spherical and Exponential structures. Variogram maps based on 36 directions were then inspected using the parameters defined in the omnidirectional variogram. The variogram maps are a means to check for anisotropy in the data, i.e., if the calculated semi-variograms are dependent on direction. Based on the variogram maps, two directions, one along the direction of greatest continuity, and one perpendicular to it, were chosen, and the directional semivariograms calculated and models fitted.

Using the parameters calculated in ISATIS (lag distance, lag number, direction of anisotropy and model structures), variogram modelling, kriging and simulation of the data were run in R using the RGeostats package (Renard *et al.* 2014). A conditional simulation approach was done using turning bands because it allows for assessment of uncertainty. Kriging was confined to a grid that fitted into the pond outline, and

consisted of 150 to 250 nodes in each direction. A moving neighbourhood was used where the search distances corresponded to the range in each variogram model, and a minimum and maximum of 10 and 50 samples were chosen, respectively. The simulation was run 100 times with 100 turning bands as additional simulation runs and turning bands were found to not alter the results. A normal-score data transformation was done on the depth and temperature data before computing and modelling the variograms. Data were back-transformed after simulations to obtain the simulated values for depth and temperature at each node. Pond water volume was calculated by summing the mean depth values after 100 simulations, and surface area calculated by summing the area of each grid node.

2.2.3 Water flow

Field

Water outflow rates from the ponds was measured using a portable flowmeter (Mash-McBirney, Inc. FLO-MATE^{TM} Model 2000). Flow measurements were taken during November 2013. The sensor was zero adjusted daily by placing it in a bucket of still water from the ponds. Areas expected to be outflow regions of each pond were found using satellite imagery, then inspected on foot before the measurements were taken.

A transect was established across the outflow area, and a flowmeter was used to measure the water flow at the bottom, middle and top of the water column at different distances along the transect. The sensor was mounted onto a wooden pole and was directed towards the expected direction of flow, and the user stood sidewards of the sensors so as to limit flow disturbance.

Calculation

The discharge rates were calculated following the Mash-McBirney FLO-MATE manufacturer's instructions. Using the distances along each transect, trapezoid areas were calculated, along with the mean bottom, mean middle and mean top flow (Figure 9). Water flow was calculated by:

$$Seg_{top}^{1} = \frac{top_1 + top_2 + top_3}{3}$$

$$Seg_{mid}^{1} = \frac{mid_1 + mid_2 + mid_3}{3}$$

$$Seg_{bot}^{1} = \frac{bot_1 + bot_2 + bot_3}{3}$$

$$A_1 = \frac{Depth_1 + Depth_2}{2} \times (d_3 - d_1)$$

$$T_{flow} = \sum \left(\frac{Seg_{top}^{1} + 2Seg_{mid}^{1} + Seg_{bot}^{1}}{4} \right) \times A_1$$

where variables are shown in Figure 9 and T_{flow} denotes the total water flow.

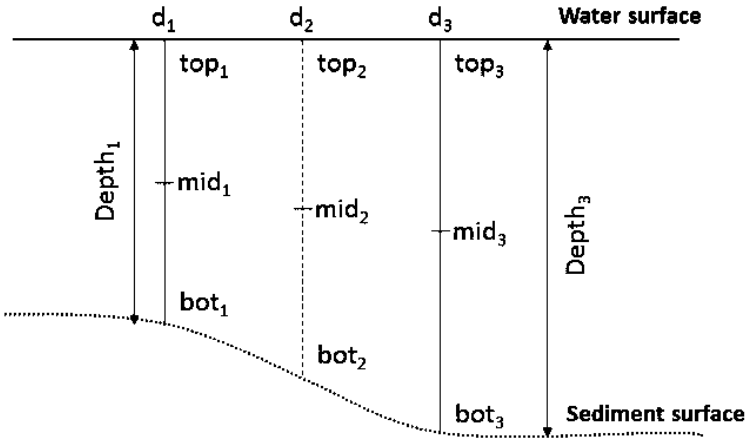


Figure 9: Diagram showing the variables required to calculate water flow accros a channel of variable depth.

Evaporative processes are inherently difficult to model due to the complex interaction between many climatic and hydrodynamic variables. The easiest and the most popular methods make use of weather data available at nearby weather stations to predict evaporative loss in a water body. Evaporative loss in each pond w as calculated following Abdelrady (2013):

$$E = 0.7 \times SF \times \left(E_p - R\right) \times A$$

$$SF = 1 - (S \times 0.00086)$$

where E_p, R, A and S denote the pan evaporation (m), rainfall (m) , lake area (m^2) salinity (‰), respectively. Salinity was assumed to be 35‰. Pan evaporation was averaged during the period that the outflow measurements were taken using data from the Bureau of Meteorology, where daily records from class A evaporation pan in a Carnarvon Airport were used.

Total outflows were calculated by summing water flow through the outlet channel with the evaporation rate at each pond. Residence time was measured as the time required for total outflow to equal pond volume. Water outflow was also expected to occur via groundwater loss, but was assumed to be negligible when compared to loss via channel flow and evaporation.

2.2.4 Classifying morphotypes

Pond volume, perimeter, mean water depth, maximum water depth, daily outflow volume, daily evaporation outflow volume and mean water temperature was used to classify the ponds into morphotypes. This was done using k-means clustering in the stats package for R, but because the ponds chosen in this study range in size, it is unlikely that these morphotypes are representative of all the Northern Ponds unless size is the primary determinant of morphotypic structure. The k-means clustering method uses the local structure of the data to delineate clusters based on high-density regions of within-group sums-of-squares (Borcard *et al.* 2011). A sum-of-squared-error scree plot was used to find a suitable number of clusters for the k-means cluster analysis. This method typically produces a bend in the plot when the sum-of-squared-error is plotted against the number of clusters. The number of clusters was thought to be appropriate at the bend of the scree plot because additional clusters beyond do not have a substantial impact on the total sum of squared error.

2.3 Results

2.3.1 Climate

There were negligible differences in mean monthly minimum and maximum daily air temperatures throughout the three year study period (Table 2, Table 3 and Table 4). The

wettest year was 2012, and the wettest month was March 2012. Similarly, evaporation was greatest during the hotter summer months, with February 2013 having the highest average daily evaporation of 11.3 m m/day. The summer months had the strongest average wind speeds, generally surpassing 25 km/hr in summer and being less than 17 km/hr in winter. The winds were generally dominated by southerlies in the morning and south-westerlies in the afternoons from November to April. Between May and October, the winds tend easterly to south-easterly in the mornings and swing to the south-west in the afternoons (Appendix 1). There were no cyclones in the area during the three years, however, there was 56.2mm of precipitation between the 24th and 25th March 2012.

2.3.1 Morphometric characteristics

The bottom geometry of each pond was determined from the results of the simulation output (see Appendix 2 for the variogram parameters). The ponds have been described by the total water volume within each pond, the surface area of the pond, mean depth and maximum depth (Table 5). Volume was directly proportional to surface area (Figure 10). Where water outflow through the channels could be measured using the flow meter, the rate was as low as 13 L/s in the smallest pond and as great as 578 L/s in the largest pond. The average pan evaporation rate in Carnarvon weather station was 10.0 mm/day during the period when water flow was being measured. Evaporative outflow increased with surface area. Similarly, the proportion of outflow estimated to be due to evaporation increased with surface area. Flow rate out of the ponds increases with pond size, as well as the magnitude of evaporative outflow. In the smallest pond, Pete's Vent, evaporative outflow is negligible (less than 1 L/s), whilst in the largest pond, Cygnet Pond, the proportion of the total outflow explained by evaporation is almost 50%.

Table 2: Summary of climatic conditions during 2012. All data are presented as monthly means. Data were retrieved from Bureau of Meteorology station at Carnarvon Airport, which is located approximately 100 km south of the study site.

	Minimum Temp	Maximum Temp	Mean Daily Precip. (mm).	Mean Daily Evap. (mm).	Mean Daily Wind Speed
January	24.2	34.3	21.0	9.9	22.5
February	23.7	32.8	15.2	9.8	26.8
March	22.1	34.3	56.2	9.8	21.4
April	18.8	30.8	6.6	7.1	19.0
May	14.2	29.0	0.6	5.7	17.2
June	12.3	24.7	32.2	3.8	15.5
July	8.8	24.7	10.4	4.6	16.3
August	11.3	24.5	6.8	5.1	18.6
September	14.8	25.9	4.6	7.1	20.8
October	17.0	28.1	0.0	8.8	24.3
November	18.9	27.7	5.0	9.8	27.5
December	21.8	30.6	14.6	9.5	25.0

Table 3: Summary of climatic conditions during 2013. All data are presented as monthly means. Data were retrieved from Bureau of Meteorology station at Carnarvon Airport, which is located approximately 100 km south of the study site.

	Minimum Temp	Maximum Temp	Mean Daily Precip. (mm).	Mean Daily Evap. (mm).	Mean Daily Wind Speed
January	24.3	35.0	25.8	10.9	27.1
February	24.8	35.8	0.2	11.3	24.0
March	22.4	32.2	2.2	9.4	24.0
April	20.6	31.4	0.0	7.6	19.5
May	14.1	26.3	11.0	5.1	16.5
June	11.2	24.0	29.0	4.5	18.2
July	9.9	23.2	0.6	4.5	16.2
August	12.2	25.4	0.2	5.7	18.4
September	14.9	24.1	7.0	6.7	21.8
October	17.1	27.9	2.8	8.3	26.0
November	20.5	29.2	0.0	9.2	26.3
December	22.2	31.9	0.4	10.5	27.3

Table 4: Summary of climatic conditions during 2014. All data are presented as monthly means. Data were retrieved from Bureau of Meteorology station at Carnarvon Airport, which is located approximately 100 km south of the study site.

	Minimum Temp	Maximum Temp	Mean Daily Precip. (mm).	Mean Daily Evap. (mm).	Mean Daily Wind Speed
January	23.3	32.9	9.0	10.9	28.4
February	24.5	35.1	0.0	10.3	23.3
March	23.3	33.1	0.4	9.7	24.3
April	19.1	30.1	7.8	7.4	20.2
May	15.4	26.0	23.6	5.0	17.7
June	10.9	25.3	14.0	4.7	17.0
July	10.2	23.8	2.4	4.4	16.0
August	12.5	28.2	0.2	6.0	16.1
September	15.8	27.1	31.4	6.7	21.3
October	17.6	26.6	1.0	8.3	25.4
November	18.8	27.9	0.0	9.7	26.7
December	20.7	30.4	0.0	10.7	26.0

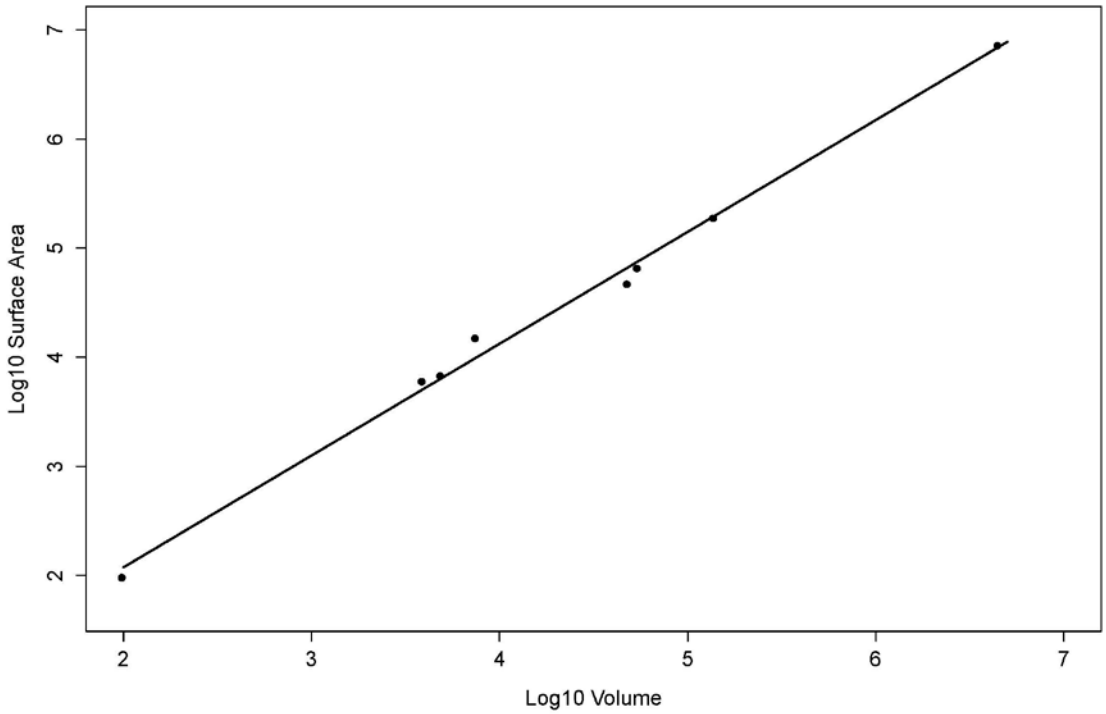


Figure 10: Linear relationship between pond surface area (m^2) and pond water volume (m^3). $R^2 = 0.99$, slope = 1.02

Table 5: Summary of morphometric and hydrodynamic characteristics of each pond. All temperature measurements were taken from approximately 10 cm below the water surface.

Pond	Volume (m^3)	Surface Area (m^2)	Perimeter (m)	Mean depth (m)	Maximum depth (m)	Outflow Rate (m^3/day)	Evaporation (m^3/day)	Total Water loss (m^3/day)	Residence Time (days)	Mean Temp. (°C)	Min Temp. (°C)	Max Temp. (°C)
Pete's Vent	98	95	40	1.03	1.63	17	1	18	5	27.3	27.2	27.4
Donut Pond	3,848	5,978	500	0.64	1.67	0*	41	41	95	23.6	23.4	24.0
Annie's Pond	4,826	6,712	400	0.72	2.40	1,094	46	1,140	4	22.2	22.0	23.3
Pete's Pond	7,390	14,880	810	0.50	0.82	0*	101	101	73	22.3	21.1	25.4
Whistler's Pond	47,470	46,544	1,250	1.02	2.11	11,214	316	11,530	4	22.1	21.5	25.2
Harjie's Pond	53,655	65,069	2,570	0.82	2.82	0*	442	442	1,198	22.5	21.5	25.7
Jana's Vent	136,528	187,509	2,950	0.73	2.57	27,173	1,273	28,446	5	19.5	17.5	24.3
Cygnet Pond	4,436,004	7,147,702	22,650	0.62	2.18	49,926	48,528	98,454	45	15.7	12.3	24.0

*These systems had no defined channel for outflow and it is likely that wind direction has a huge influence on water outflow OR water inflow = evaporative loss.

2.3.2 Pond details

Pete's Vent

Pete's Vent is the smallest pond in this study, being only 95m^2 in area and 98m^3 in volume (Figure 11). Despite its small area, it has an average depth of just over 1m, with a maximum depth of 1.6m where the vent occurs. There is a large channel where the water is discharged at a rate of 17m^3/day. Evaporative loss from this pond is negligible. Due to its small volume and the large discharge rate, the residence time in this pond is only 3 hours. Discharged water leaves the pond and travels down the channel where it is connected with the main water body in this system, Pete's Pond. Water temperature was relatively constant at 27.3°C.

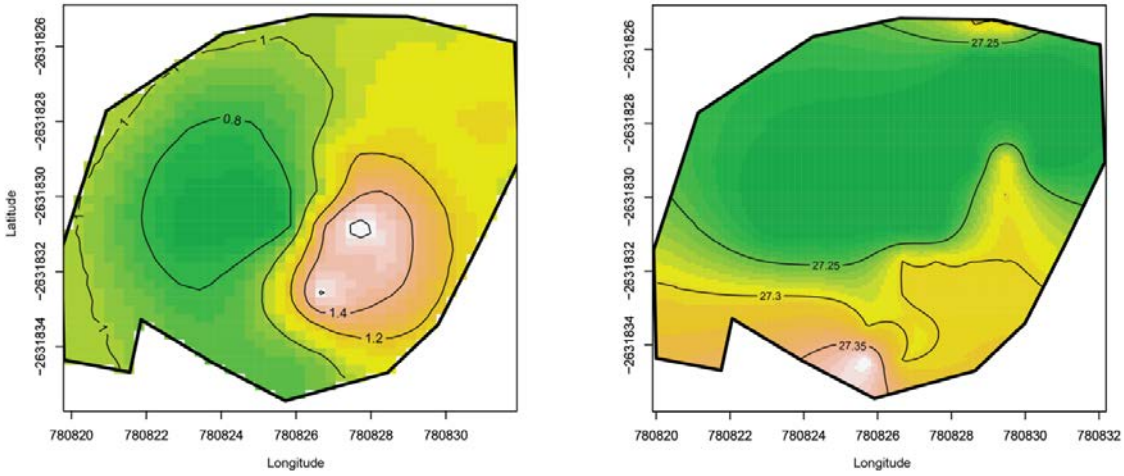


Figure 11: Bathymetric map of Pete's Vent with depth contours in meters (left) and water temperature (°C; right). All measurements were taken in July 2013.

Donut Pond

Donut Pond is found in the northern area of the Cygnet System. It is one of the larger of the ponds located between Sandy Bluff Sill and Cygnet Pond, covering an area of 3,800m^2 and containing 6,000m^3 of water (Figure 12). The vent is found in a small channel in the north-west of the pond, and reaches a depth of 1.7m. The average depth of the pond is slightly more than 0.6m. No definitive outlet channel could be found for this pond suggesting outflow is relatively slow. Water temperature was highest in the northern channel where the vent is located, where it was 24°C, and decreased to 23.4°C in the southern section.

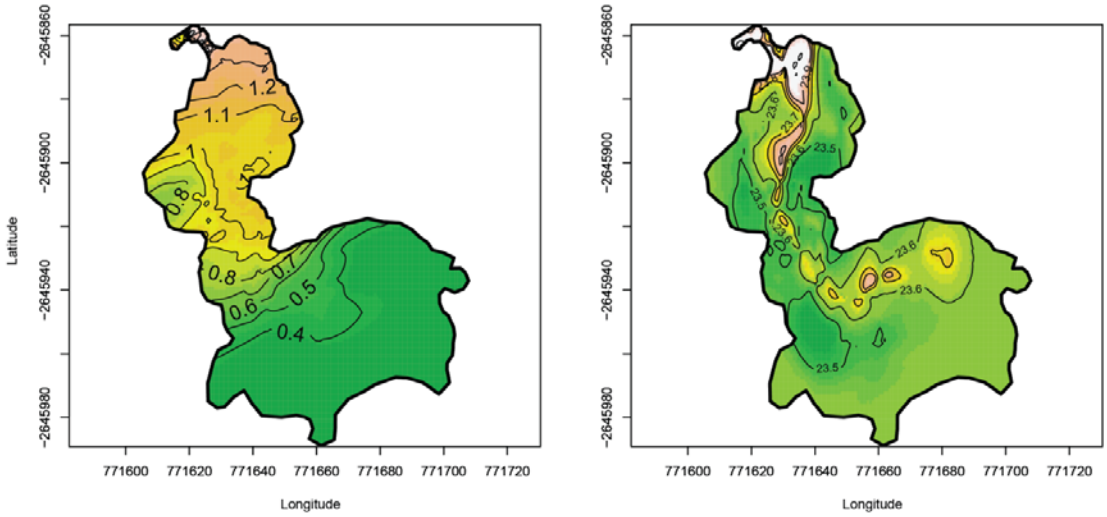


Figure 12: Bathymetric map of Donut Pond with contours in meters (left) and water temperature (°C; right). All measurements were taken in July 2013.

Annie's Pond

Annie's Pond is a medium sized water body found in the northern part of the Chirida System. It has a maximum width of just over 140m, a surface area of 6700m^2 and a water volume of 4,800m^3 (Figure 13). Annie's Pond reaches a maximum depth of 2.40m in the north-western corner, which is where the inlet vent is found. The pond becomes shallower in all directions away from the vent, with an average depth of 0.72m. The major outflow region was found at a very small outlet channel along the southern bank, where a flow rate of 1,094m^3/day was measured. Residence time 4 days once evaporation has been accounted for. Water temperature ranged from 23.3°C to 22.2°C.

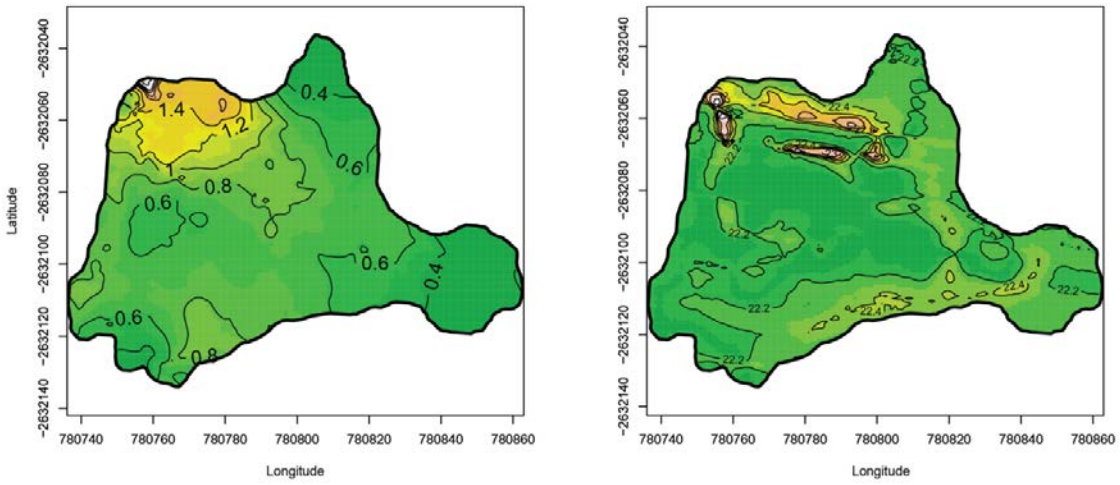


Figure 13: Bathymetric map of Annie's Pond. Contour lines are in meters (left) and water temperature (°C); right. All measurements were taken in July 2013.

Pete's Pond

Pete's Pond is the northern most water body in the Chirida system, and represents the northern limit of the ponds examined in this study. It has a surface area of 7,400m^2, larger than the neighbouring pond, Annie's Pond. It also has a water volume which is more than double that of Annie's Pond, (14,900m^3; Figure 14). This pond has no visible large vent area, and therefore has a relatively shallow maximum depth slightly deeper than 0.8m and an average depth of 0.5m. There were no signs of an outflow channel so water flow was measured across the narrowest point of the pond where it was believed a flow rate could be measured. However, no flow was measured, probably due to the rate being less than the sensitivity of the sensor (1cm/s). Residence time in this pond is longer than that of other ponds. The maximum water temperature was 25.4°C and minimum 21.1°C. The higher temperatures are found in the north-west section of the pond.

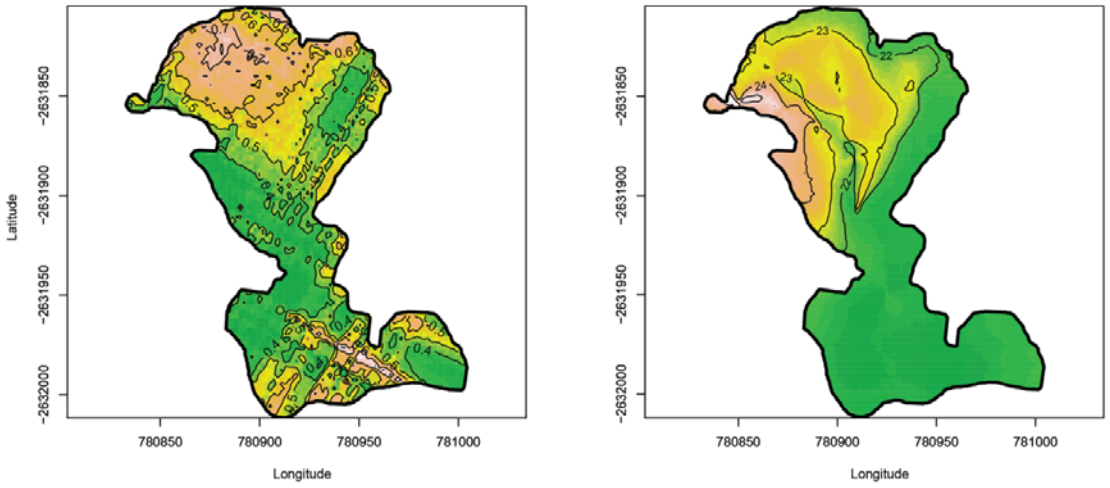


Figure 14: Bathymetric map of Pete's Pond with contours in meters (left) and water temperature (°C; right). All measurements were taken in July 2013.

Whistler's Pond

Whistler's Pond is a medium sized pond located in the Chirida System, north of Harjie's Pond. The pond covers an area of 46,500m^2 and contains 47,500m^3 of water (Figure 15). There are three distinct vent areas found in the northern sector of the pond, with the deepest having a maximum depth of 2.1m. The pond is quite deep compared to the other ponds, with an average depth of just over 1m. There is a well-defined outlet for this pond along its eastern margin, where water flows out at 11,214m^3/day into a large shallow spill sheet. Evaporation loss is estimated at 316m^3/day, which is relatively low in comparison to water discharged through the channel (3% of total outflow). The residence time in this pond is 4.1 days. Water temperature in the pond varied between 25.2°C near the vents to 21.5°C.

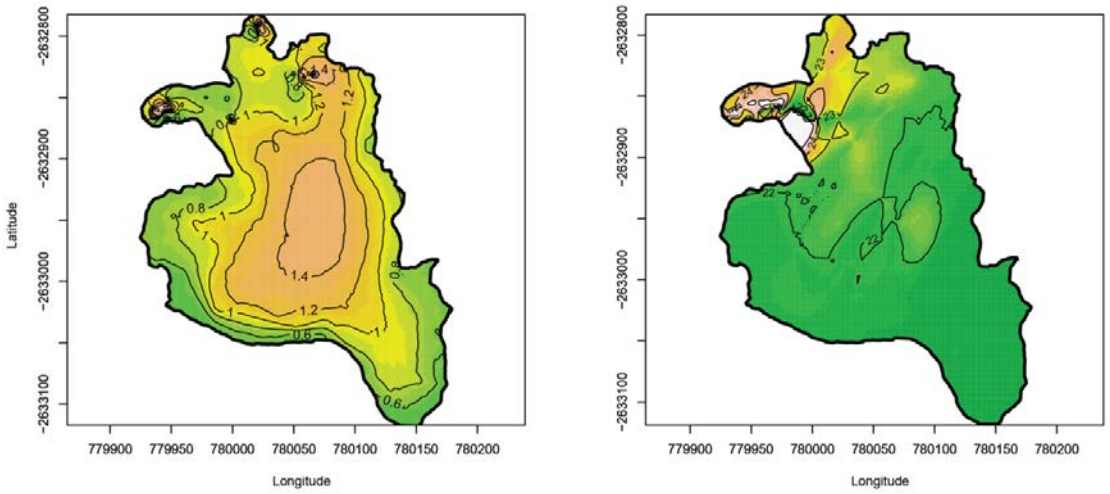


Figure 15: Bathymetric map of Whistler's Pond. Contours are in meters (left) and water temperature (°C; right) All measurements were taken in July 2013.

Harjie's Pond

Harjie's Pond, located immediately south of Whistler's Pond, is larger than its neighbour. It is a narrow pond that extends in a south-east direction with a surface area of 53,700m^2 and a water volume of 65,000m^3 (Figure 16). It is shallower than Whistler's Pond, with a mean depth only slightly greater than 0.8m, however, it has a deep vent area in its western reaches with a depth over 2.8m. Although water flowing out of the vent was difficult to observe visually, the warmer water temperatures in this area suggest that this is a significant entry point for water into the system, as increased water temperatures were found to correspond to vent areas in the other ponds. Another shallower vent is found in the northern limits of the pond, and there is probably a third vent that was not measured in the extreme western reach of this large pond. No outlet point could be found for this pond and therefore water flow was measured across a narrow point where water was thought to be flowing eastwards. No flow was detected, and this is probably attributable to the flow being less than 1 cm/s and therefore not detectable by our sensors. However, the residence time is longer than that of the other ponds where channel outflow was detected. Water temperature varied from 25.7°C near the vents to 21.5°C in the western regions.

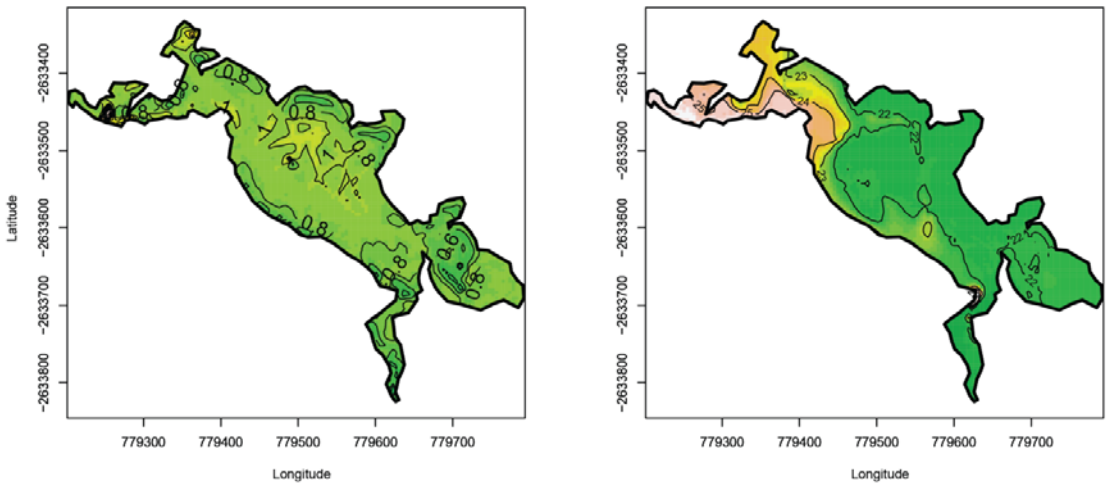


Figure 16: Bathymetric map of Harjie's Pond with contours in meters (left) and water temperature in Harjie's Pond (°C; right). All measurements were taken in July 2013.

Jana's Vent

Jana's Vent is found in the Cygnet system, south of Sandy Bluff Sill and west of Cygnet Pond. It is the second largest pond in the study, being 187,500m^2 in surface area and 136,500m^3 in volume (Figure 17). The deepest regions of the pond are in the north where the vent reaches depths of 2.6m. Water flows south from the vent and is discharged via an outlet channel found in the extreme south of the pond. Water is discharged at a rate of 27,173m^3/day and flows through a long series of channels before being released into Cygnet Pond. Discharge also occurs along the south-western margin of the pond, however, this was perceived to be negligible compared to that occurring from the channel region when visited during the study. Evaporation was 1,273m^3/day and the residence time of water in the pond is 4.8 days. Water temperature at the vent region was 24.3°C and decreased in to 17.5°C in the southern most parts of the lake where it is more strongly influenced by wind and ambient temperatures.

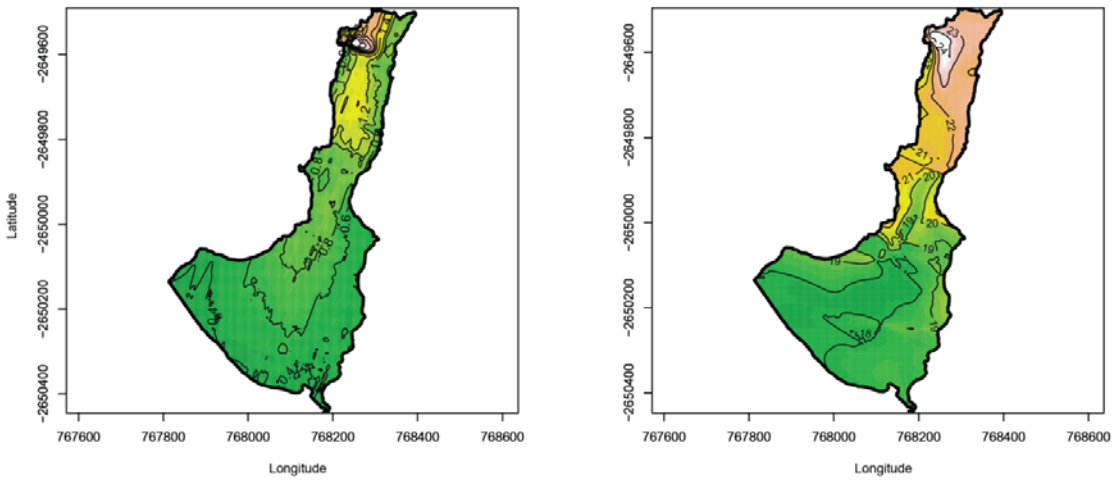


Figure 17: Jana's Vent bathymetric map with contours in meters (left) and water temperature in Jana's Vent (°C; right). All measurements were taken in July 2013.

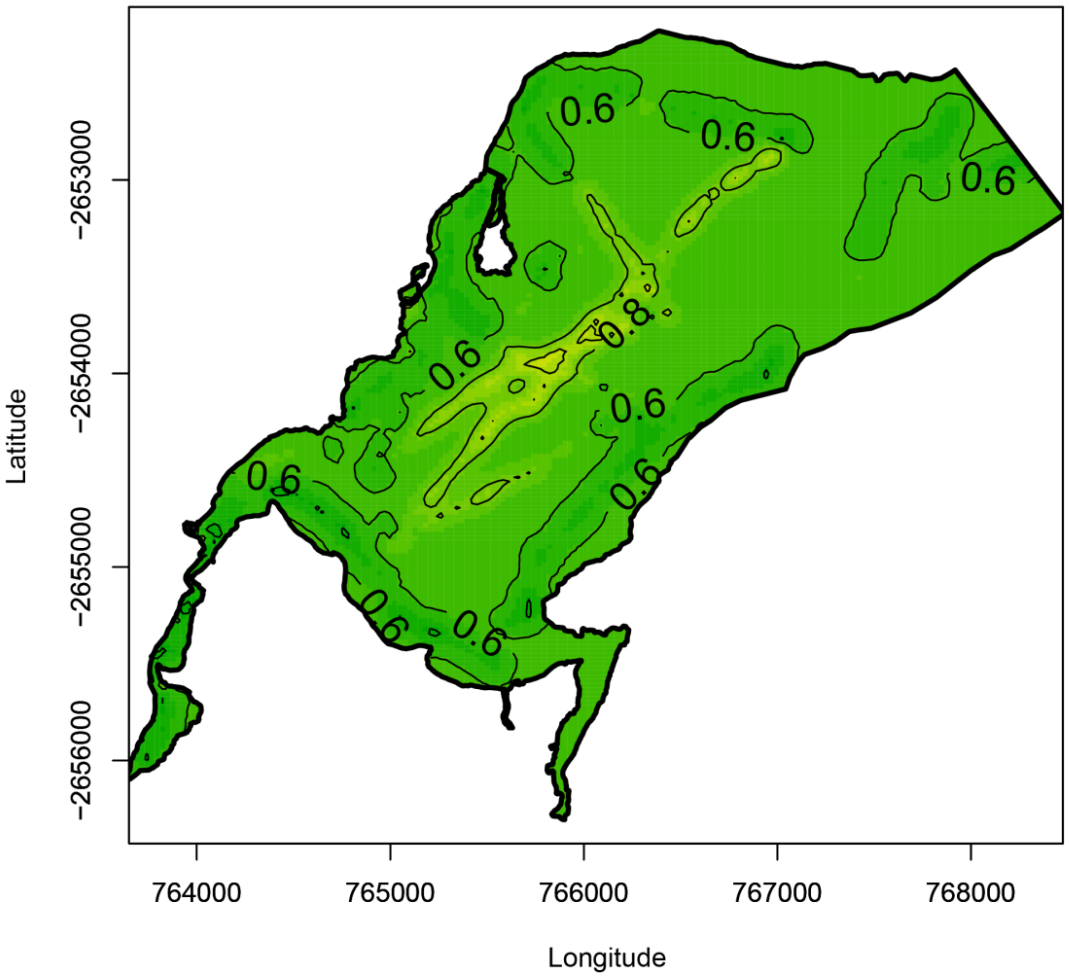


Figure 18: Bathymetric map of Cygnet Pond. Goat Island can be seen on near the western shore. Contours are in meters. All measurements were taken in July 2013.

Cygnet Pond

Cygnet Pond is the largest in this study and the second largest pond in Lake MacLeod, second to Ibis Pond, into which Cygnet Pond drains. Only the southern sector of this pond was examined in this study (this is the 'pond' part of the Cygnet system, not the 'channel like' part of it). The northern sector contains numerous vents and channels that are relatively deep and connected to the southern sector via a wide, shallow channel. Cygnet Pond is the largest pond used in this study, covering an area of 7.1km^2 and containing 4,436,000m^3 of water (Figure 18). In addition to the outlet channel for Jana's Vent that functions as an inflow, there are three vents found in Goat Bay, and at least

two more in the south-west of the pond. The maximum water depth is about 2.2m, but the pond is mostly shallow with a mean depth slightly greater than 0.6m. The primary outlet for this pond is in the north-east, however, there is probably substantial flow occurring across most of the south-western margin, as water flows over a shallow sill into Ibis Pond. A flow of 49,926m^3/day was measured flowing out of this pond across a 200m wide channel. Evaporative loss was approximately equal to that of channel outflow (48,528m^3/day), meaning that 49% of total outflow was caused by evaporative loss. This gives the pond a residence time of 45 days. Temperature maps could not be made for Cygnet Pond, because of the length of time required to cover the area in a boat, and the drastic change in water temperature caused by changes in wind. It is likely that because of the large, shallow nature of this pond, water temperatures are subjected to great change caused by interaction with ambient temperatures and wind.

2.3.3 Morphotypes

The eight ponds could be grouped into four clusters using k-means clustering and the morphotype variables for pond volume, perimeter length, mean water depth, maximum water depth, channel outflow, evaporative outflow and mean water temperature (Figure 19). The scree plot flattens after four clusters, therefore additional clusters do not greatly reduce the sum-of-square-error. The first group consists of Annie's Pond, Harjie's Pond and Jana's Vent. These ponds had the greatest mean and maximum depths (Table 6). The second group, which only consisted of Cygnet Pond, was different to the other groups because of its large volume, substantial channel and evaporative outflow, as well as having the coldest water temperature. The third group consisted of Pete's Vent and Whistler's Pond and had the highest mean water temperatures. The final group consisted of Pete's Pond and Donut Pond, which were characterised by the lowest mean water depth. Daily evaporative outflow was less than 1% of total water volume for all the groups. Group 3 had the lowest SA:V ratio (1m^{-1}), while group 1 had a SA:V ranging from 1.2 to 1.4m^{-1}. Group 2 and 4 had SA:V ratios ranging from 1.6 to 2.0 m^{-1}.

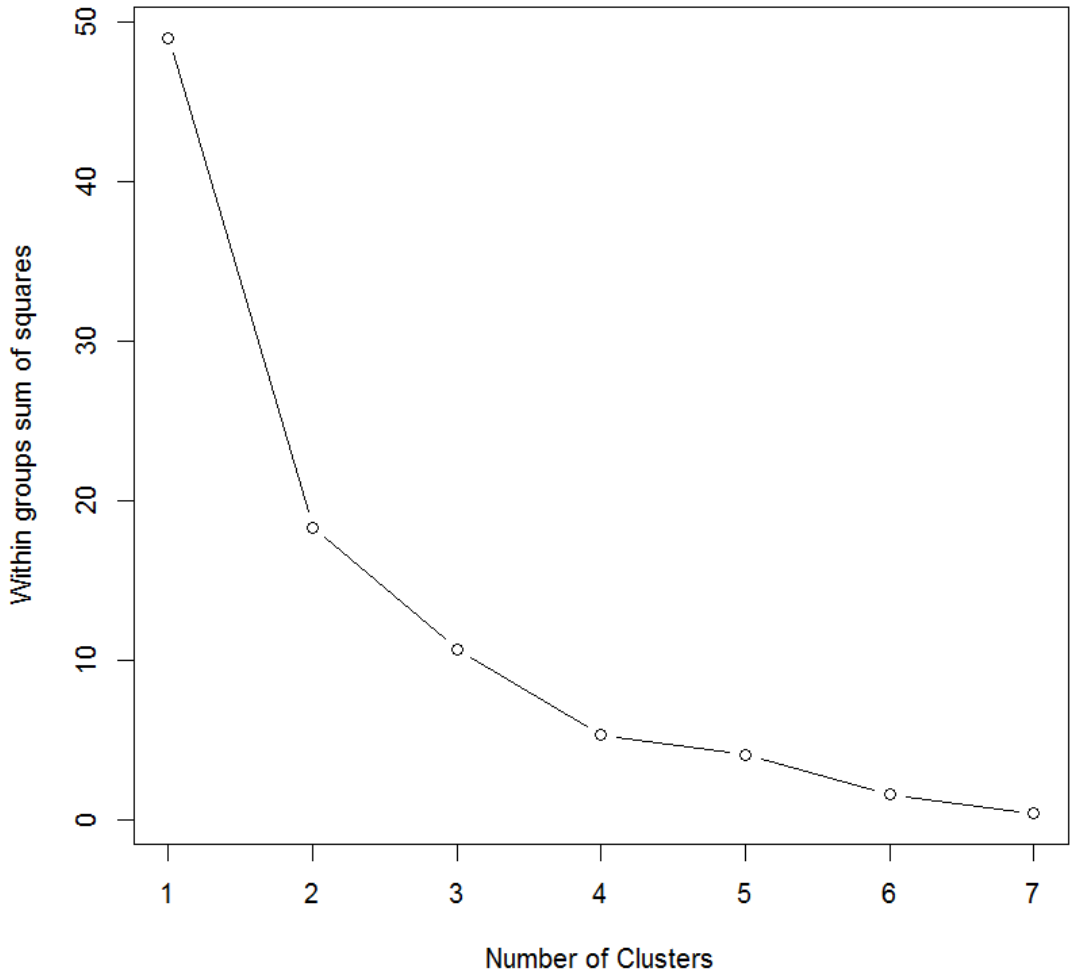


Figure 19: Scree plot of the change in sum of squares as k-means clusters are added using the morphometric variables volume, perimeter, mean depth, max depth, outflow, evaporative outflow and mean water temperature for each pond. Four clusters were chosen as the change in the within group sum of squares was minimal as additional clusters were added. The scree plot flattens after 4 clusters, therefore, the additional clusters do not reduce the sum-of-square-error greatly.

Table 6: Minimum and maximum values of morphometric properties used to define the four pond 'morphotypes'. Outflows are recorded as the proportion of pond volume loss per day. A: Annie's Pond; C: Cygnet Pond; D: Donut Pond; H: Harjie's Pond; J: Jana's Vent; PP: Pete's Pond; PV: Pete's Vent; W: Whistler's Pond.

Group	Ponds	Volume (m^3)	Perimeter (m)	Mean depth (m)	Max Depth (m)	Channel Outflow (%)	Temp	SA:V
1	A, H, J	4826-136,528	400-2650	0.7-0.8	2.4-2.8	0-22	19.5-22.5	1.2-1.4
2	C	4,436,004	22650	0.6	2.2	1.0	15.7	1.6
3	PV, W	98-47470	40-1250	1.0	1.6-2.1	17-24	22.1-27.3	1
4	PP, D	3848-7390	500-810	0.5-0.6	0.8-1.7	0	22.3-23.6	1.6-2.0

2.4 Discussion

Pond morphometry can be a useful proxy for the environmental parameters (Jackson *et al.* 2001, Jyväsjärvi *et al.* 2009) and ecologies (Jackson and Harvey 1993, Mehner *et al.* 2007) of aquatic systems. In this study, the ponds could be grouped into four 'morphotypes' using the morphometric and hydrodynamic characteristics determined for each pond. The four morphotype groups were mainly distinguished by mean water depth and proportional volume of channel outflow. This chapter shows that some ponds have similar morphometric characteristics. These morphotypes may be useful determinants of the environmental conditions and microbial communities found within each pond.

2.4.1 Update on the Lake MacLeod Ponds

Bathymetry

The integration of water temperature and bathymetric data yielded maps for eight of the main ponds in Lake MacLeod. The geological maps of Logan (1987) and sedimentology maps of Shepherd (1990) provide a geomorphological context for the ponds, while the maps produced in this study provide detailed information on water flow from the vents to the surrounding habitats. The morphologies of the ponds exhibited similar patterns, in that water temperature was highest nearest to the vents, and that all the ponds have simple topographies, where they tend to get shallower towards their eastern margins.

Surface area of the ponds studied varied from 95m^2 to 7km^2, and water volume from 98,000L to over 4.4GL. As was expected, water volume in the ponds was proportional to pond surface area. Either of these measurements would be a useful proxy for defining the size of the pond. However, as all the ponds are relatively shallow with similar mean depths, it is more useful to define pond size as a function of surface area. Aerosols are an important transport vector for microbes (Hörtnagl *et al.* 2010), and the surface of the lakes are an important entry point for these immigrants into the ponds. It is therefore likely that ponds with larger surface areas are subjected to higher levels of immigration (Comte *et al.* 2014). For these reasons, surface area will be used in the remainder of this thesis as a proxy for ecosystem size.

The ponds in Lake MacLeod have many similar morphometric characteristics. For example, all of the ponds, with the exception of Pete's Pond and Harjie's Pond, have a deep region that corresponds to a vent. The vents are almost exclusively found in the north-western region of each of the ponds, and the water bodies generally extend in a south or south-east direction and become shallower. Most of the ponds have a defined boundary around the perimeter, although the shore is generally steeper along the western and northern edges of the ponds (Shepherd 1990 and pers. obs.). These vent areas are easily identified when in the ponds by looking for depressions in the pond floor.

The vent regions were all found to be more than 1.5m deep, with the deepest one at Jana's Vent (2.6m); elsewhere in the Lake MacLeod system, the vents are up to 6m deep (Hidden Vent; pers. obs.). The shallower reaches of the ponds, where depth measurements were difficult to make as the sonar device can only measure greater than 0.30m, were generally found in the south to south-eastern margins of the ponds. For example, Annie's Pond, Pete's Pond and Donut Pond do not have definite shorelines along their south-eastern margins. In these ponds, the water bodies typically reach a depth of less than 0.30m and mangrove pneumatophores become dense and impede water flow along their eastern margins.

Hydrodynamics

The ponds either had a large outflow channel, where water flowed out of the pond into adjacent spill sheets, or were confined water bodies with little or no surface outflow. In the former scenario where there were outflow channels, water inflow through the vents must be sufficiently great enough, and evaporative loss low enough, to cause a net positive water balance. Alternatively, in the other ponds with no outlet channels, water inflow must be in equilibrium with evaporative loss, hence the constant volume of water despite the constant inflow from the vents. These ponds have a neutral water balance. It could be expected that these two different hydrological characteristics could influence the environmental conditions and/or the microbial communities found within them (Fisher *et al.* 2009). Additional outflow is likely to occur via subterranean seepage, as brine moves towards the basin centre through the sediment column. In this study, this effect was assumed to be negligible compared to other sources of outflow, but in the

ponds with a neutral water balance, it is assumed that subterranean seepage is not the most important method of brine loss.

Attention should, however, be paid to the flow rate measured for Cygnet Pond, which is probably strongly affected by wind strength and direction. The rate calculated was measured during a period of sustained northerly wind, and therefore is probably at the high end of the range that could be expected to flow out of the pond. When the measurements were being made, the wind direction changed to a southerly and a reduction in water flow was observed. Water levels in Cygnet Pond have been seen to increase by up to 5cm when there is a sustained southerly wind (per. obs.). It is therefore likely that wind plays an important role in determining water volume and flow out of Cygnet Pond. A southerly wind is likely to 'push' water upslope from Ibis Pond and thus limit the flow out of Cygnet Pond, resulting in raised water depths throughout the pond (probably limited to 5cm). Conversely, under northerly wind conditions, water is push downslope, and the water movement through the channels from Cygnet Pond and into Ibis Pond is facilitated, and the water level drops.

Some of the ponds received inflow via channels from adjacent ponds. This includes Pete's Pond, Harjie's Pond and Cygnet Pond. Pete's Pond receives water directly from the outflow channel of Pete's Vent, with the channel being approximately 30m long. No channel outlet or significant areas of spill over were found around Pete's Pond, so the system is likely to be in equilibrium, with total inflow/seepage equal to total evaporation loss. It can therefore be estimated that inflow from Pete's Vent is approximately 17% of total inflow into the system and that other sources of seepage are contributing water to the system. A similar scenario occurs with the outflow from Jana's Vent, which flows into Cygnet Pond. Cygnet Pond, however, has numerous other vents within the pond, and total inflow of brine originating from Jana's Vent should be expected to represent an small proportion of the total water body. Nonetheless, the data in this chapter suggests that inflow from Jana's Vent into Cygnet Pond is large, and represents 28% of the total daily water loss of Cygnet.

Morphotype groups

Mean water depth, channel outflow and water temperature were found to be good descriptors of the four pond morphotypes found for the eight ponds used in this study.

The shallowest groups, group 2 and 4, were characterised by very low, or no channel outflow, although Harjie's Pond (group 1) also had no channel outflow. Cygnet Pond has morphometric and hydrodynamic characteristics which are different to the other ponds, hence it was the only member of the group 2 morphotype. The most obvious feature that separates Cygnet Pond from the other ponds is its large water volume, which is an order of magnitude greater than that of the second largest pond, Jana's Vent. Cygnet Pond also has a very low mean water temperature, which because of the high surface area-to-volume ratio, is probably affected by prevailing winds, which may act to cool the water body through evapo-transpiration (Venäläinen *et al.* 1998, Van Cleave 2012). On the other hand, group 3, which includes Pete's Vent and Whistler's Pond, were characterised by the large proportion of water loss daily through an outlet channel, although Jana's Vent (group 1) has a comparable flow through an outlet channel.

2.4.2 Comparison with other saline wetlands

The closest analogues to the Northern Ponds in northwestern Australia are the "birradas" of Shark Bay, and Mandora Marsh which is inland of Eighty Mile Beach in the Pilbara. Birridas are close to the coast, evaporite pans occurring in interdunal depressions, especially around Shark Bay. Halse *et al.* (2000) state "Most of the birridas contain gypsum and, although they may dry intermittently, anecdotal information suggests their water levels show subdued response to oceanic tides." Indeed they liken Lake MacLeod, as a whole to be an example of a very large birrida. The Northern Ponds themselves, some 200 km north of Shark Bay, are clearly part of this analogous regional system as permanent water bodies fed by sea water in which the tidal signal appears to be completely, or almost so, obliterated by the long passage through the karst. Mandora Marsh is also a permanent arid zone wetland in north-western Australia, contains waters that range from fresh to saline, however they are not of marine origin. Although each wetland is not particularly diverse, because of the high levels of environmental heterogeneity, when considering the entire suite of wetlands, the marsh supports a relatively rich fauna assemblage (Storey *et al.* 2011).

The geological, hydrological, edaphic and biological characteristics of the saline wetlands located throughout Monegros, north-east Spain, has been well documented (Mees *et al.* 2011, Casamayor *et al.* 2013, Castañeda *et al.* 2013), and is perhaps the

best example of where the morphometric properties of saline wetlands has been investigated (Castañeda *et al.* 2013), although some morphological characteristics have been documented for the salt lakes in South Australia (Timms 2009) and Western Australia (Boggs *et al.* 2006). Similar to the ponds investigated here, the Monegros wetlands are typically shallow, with about 80% of the lakes ranging in depth from 0.3m to 5.0m. Similarly, the salt lakes of the Eyre Peninsula of South Australia are mostly less than 50cm deep, with few lakes being deeper than 1m (Timms 2009).

An interaction with subsurface geomorphology has been described for the geographic location of the Monegros salinas because these systems are usually the result of subsurface hydrogeological features. Similarly, the morphology of some playas of the Yarra Yarra salt lake system of Western Australia were found to be determined by the underlying geology (Boggs *et al.* 2006). The ponds in Lake MacLeod are also located along the seepage face, a subsurface hydrological feature where seawater enters the basin.

The ponds of Lake MacLeod displayed elongation along a south-east orientation, which not only corresponds to the direction of the basin slope, but also with the direction of the prevailing winds. Similarly, the Monegros salinas are also predominately elongated along the direction of the prevailing winds (Castañeda *et al.* 2013). Not only do the prevailing winds appear to be elongating or orientating the ponds in Lake MacLeod, evidence of wind driven fluctuations in water level, and steepening and deepening of the northern shorelines was observed. Similar interactions with local climate are also thought to be determining the orientation and elongation of the morphology of the Yarra Yarra playas (Boggs *et al.* 2006). Together these studies show that prevailing winds can be important in defining the morphology of saline wetlands/lakes

Most inland saline lakes are not permanently inundated like the ponds of Lake MacLeod. Despite the similar morphological characteristics the ponds have with other salt lake systems, it is likely that the environmental conditions and the ecology of permanently wet systems are different to the typical ephemeral systems as there is no alteration between wet and dry phases. An analogue might exist for the four saline lakes located on the Eyre Peninsula of South Australia, where seawater is delivered via marine springs, and are also permanently inundated (Timms 2009, 2010). These

systems were found to be distinct from the surrounding ephemeral salt lakes in that they have significantly different environmental conditions, zooplankton communities and littoral invertebrate communities (Timms 2009). The permanently inundated ponds of Lake MacLeod are therefore likely to have distinct environmental conditions and ecologies to those of other salt lakes, despite their similar morphometric characteristics.

2.4.3 Role of geomorphology and climate

The geomorphology, hydrology and climatic features of wetlands are known determinants of the ecological functions and processes that occur in wetlands (Mitsch and Gosselink 2000). This chapter describes those determinants by investigating the morphometry and hydro-dynamical characteristics each of the eight ponds studied. The seepage face, which runs along the western margin of the lake basin, is the dominant geomorphological feature which determines where the ponds are located within the basin. It is along this seepage face where water seepage can occur through the vents (Logan 1987). The geomorphology therefore plays an important role in determining the size of the vents and the volume of water which is discharged, which in turn affects the residency time of the discharged water within the pond.

Climatic drivers, and in particular wind driven processes affect the morphology of each pond. Although the ponds are mostly elongated along the basinal slope, their shape has been influenced by wind induced currents which cause the ponds to have steeper banks and deeper depths along the northern and north-western parts of the ponds. Wind also plays an important role in the hydrologies of the ponds as it is causes fluctuations in water depth and rate of discharge through channels (particularly in Cygnet Pond). Other climatic variables, such as ambient temperatures and precipitation affect the rate of water loss via evapo-transpiration, although wind also plays a role in this process too. The differences in these factors, which are attributed to the geomorphological and climatic drivers operating at each pond, are likely to be reflected in their hydrochemistry and ecologies.

2.4.4 Limitations

The Northern Ponds are a dynamic environment with the wetlands expanding and contracting with seasonal changes in evaporation rates, although the volume and

distribution of pond habitats remain relatively consistent (Logan 1987, Shepherd 1990). As discussed above, as well as by Shepherd (1990), variation in water depth and extent has been noted in Cygnet Pond due to wind forcing water 'upslope' and the subsequent reduction of outflow. Other sources of outflow, particularly subterranean seepage, are also assumed to be negligible compared to outflow via evaporation and channel outflow in this study. Seepage loss is probably a significant source of water loss in the ponds where water balances were found to be neutral (Donut Pond, Pete's Pond and Harjie's Pond), although the relative amount of seepage loss compared to evaporative loss remains unknown. Nonetheless, the contribution of seepage loss and wind induce variation need to be acknowledged as unaccounted sources of brine loss, which may have significant effects on the water volumes and residence times of the ponds.

All the measurements in this section were made at a single time point, and therefore ignore the effects of daily, seasonal and annual variation on the morphometry and hydrology of the ponds. The climatic data presented in this chapter highlights the strong seasonal variation in evaporation (ranging from 10 mm/day in summer to 4.5 mm/day in winter), and to a lesser extent, small inter-annual variation in evaporation, rainfall and temperature. Because the ponds are topographically confined, and don't appear to change size with season (personal observation), it may be assumed that outflow through the discharge channels increases during the winter months. However, under the assumption that total water loss from the ponds is always in equilibrium with evaporative loss and channel outflow, the residence time of water within the ponds must also remain relatively constant.

Residence time may have also been underestimated because of the detection limit of the flow meter (1 cm/sec). If flow was below detection limit in Donut Pond, Pete's Pond and Harjie's Pond, discharge volumes could have been underestimated by up to 850 m^3/day for every square meter of discharge area. This large source of error needs to be considered when in view of the high water residence times calculated for some of the ponds. This undetected flow can exceed the estimated evaporation rates for these ponds and may have important consequences on the water chemistry and biology of these systems. Furthermore, some areas of the ponds were not measured for water depth because of the 0.3m detection limit of the fish finding software. These areas, however, were relatively small and usually consisted of regions of dense pneumatophore beds.

Despite this limitation, water depth was measured throughout all the other regions of the pond, and surface areas estimated using satellite imagery. It is therefore unlikely that by not including these shallow areas in the depth surveys that pond volumes were underestimated, especially in the larger ponds.

2.4.5 Conclusion

The ponds are similar in their morphology, despite their differences in size. They are all shallow, and have seepage vents located along the western and north-western shores. All the ponds tend to become shallower in an easterly direction, where an outlet channel is sometimes located. The volume of seepage inflow through the vents determines whether the pond has a net positive water balance or a neutral water balance. When a pond has a positive water balance, surface water can leave the pond and flow into adjacent spill sheet areas. When the pond has a neutral water balance, seepage is in equilibrium with evaporative loss and there is no surface outflow into adjacent spill sheet areas. Ponds with a positive water balance include Pete's Vent, Whistler's Pond, Jana's Vent and Cygnet Pond, while ponds with a neutral water balance include Donut Pond, Annie's Pond, Pete's Pond and Harjie's Pond.

Pete's Vent is the smallest pond studied in this thesis and was characterised by having a relatively deep mean depth and low water residence time. Donut Pond and Annie's Pond are similar sizes, but Annie's Pond has a positive water balance, and thus a lower residency time than Donut Pond, which has a neutral water balance. Pete's Pond receives water from Pete's Vent, although only about this inflow represents less than a fifth of its water budget. The remaining water enters the shallow Pete's Pond via other small vents. Whistler's Pond is a medium sized pond which is characterised by the large outflow channel. The large volume of water that leaves this pond is supplied via the numerous large vents in its north-western edge. On the other hand, although of a similar size to Whistler's Pond, Harjie's Pond has a neutral water balance, and is probably supplied water through vents which discharge small volumes of water. Jana's Vent has a large vent in its northern sector, and large volumes of water are discharged from the pond via a channel in its south-eastern edge into Cygnet Pond. Cygnet Pond receives the majority of its water from other vents located along its long western shoreline though. Cygnet Pond has a large channel region where water is discharged into Ibis Pond,

although water probably leaves the pond at many points along its south-eastern shore. Cygnet Pond also loses a large proportion of its volume through evaporation because of its large surface area.

Although the ponds at Lake MacLeod share many morphological characteristics with other salt lake systems, it is likely that the environmental and ecological characteristics of the ponds are unique because they are permanently inundated, and represent a rather rare system. Nonetheless, this chapter shows that the geomorphological and climatic characteristics at Lake MacLeod are important drivers of the hydrology of each pond. The influence of the geomorphological, climatic and hydrologic drivers is likely to be reflected in the pond w ater chemistry and sedimentology, which in turn affects the composition of the microbial communities found there.

CHAPTER 3. SEDIMENTOLOGY AND HYDROCHEMISTRY

3.1 Introduction

Many studies have shown that chemical characteristics of the environment, such as salinity and nutrient levels, are good determinants of microbial community structure in inland waters (Domènech *et al.* 2006, Andrushchyshyn *et al.* 2009, Buosi *et al.* 2011, Lei *et al.* 2014). Physical characteristics, such as water temperature and sediment composition, have also been shown to be important determinants of microbial community structure (Patterson *et al.* 1989, Finlay and Esteban 1998, Gücker and Fischer 2003, Andrushchyshyn *et al.* 2009). It is therefore important to understand the chemical and physical properties of the ponds at Lake MacLeod in order to understand the structure and variability of the habitats. If the species-sorting process is an important mechanism that structures the microbial communities of Lake MacLeod, changes in chemical and physical conditions of the habitat would be expected to be an important component of the species-sorting 'filter' that structures the microbial assemblages.

Chapter 2 showed that there are four pond morphotypes found at Lake MacLeod, which are characterised by different water temperatures, mean depths, water outflow, and surface area-to-volume ratios (Chapter 2). Similar lake morphological properties, such as mean depth and surface area, have been used as surrogates for environmental variables in some studies because they have been found to be good predictors of nutrients and oxygen concentrations (Jackson *et al.* 2001, Jyväsjärvi *et al.* 2009), as well as determinants of the community structuring for fish (Jackson *et al.* 2001, Mehner *et al.* 2007) and microbes (Olding *et al.* 2000). For the Northern Ponds of Lake MacLeod, seawater enters the ponds chemically similar, at least with a similar ionic composition, to the seawater found at the intake zone/ocean feedstock (Shepherd 1990). The seawater is subsequently modified by evaporation, infiltration and runoff as it moves throughout the ponds, and it is therefore likely that the pond morphotypes have different physical and chemical conditions and thus represent different habitats.

Both Shepherd (1990) and Logan (1987) observed that the evapo-concentration of brines within the ponds can often exceed the precipitation thresholds of many mineral phases, including aragonite, gypsite, huntite, dolomite and halite. Since these precipitates accumulate in the sediment column, it is likely that the sediment composition of the ponds will reflect, at least to a degree, the overlying water chemistry. These processes become complicated by geographic (such as bathymetric) and climatic variation (Shepherd 1990). For example, water in faster flowing areas, such as channels, can be delivered further from the source before precipitation occurs, and seasonal variation in evaporation rates may result in precipitation occurring in summer months, and dissolution occurring during winter months (Shepherd 1990).

Few studies have considered relationships with physical and chemical aspects of bottom sediments in shallow saline lakes and ponds (Crosbie and Chow-Fraser 1999, Rowan *et al.* 2012, Kissoon *et al.* 2015), where water-sediment interactions play greater roles on the water chemistry than in deeper lakes (Scheffer 2004, Kissoon *et al.* 2015). In well mixed, shallow ponds which are free of macrophytes, it can be expected that sediments are prone to re-suspension and subsequent nutrient release (Faafeng and Mjelde 1998, Horppila and Nurminen 2003). However, in ponds where macrophytes are present, the physical stabilisation of sediments and the additional link between water and sediments through plant mediated elemental cycling may affect water chemistry (Barko and James 1998, Nurminen and Horppila 2009). Element cycling is also influenced by changes in oxidation and reduction reactions as well as photosynthesis and metabolism of surface microbial biofilms (Jackson and Harvey 1993, Wetzel 2001). For example, Wong and Yang (1997) showed that PO_4 and NH_4 concentrations increased in water ate the sediment-water interface as the redox potential and pH of underlying sediments decreased. These studies show the clear link between the physical and chemical structure of habitats at the sediment-water interface.

In Lake MacLeod, the ions and minerals are thought to become increasingly concentrated as water moves further from the entry point through evaporation driven concentration (Logan 1987, Shepherd 1990). However, the detailed effects of the passage of the seawater through the karst on the water chemistry are not known, and there is little information regarding the effects of evaporation of the brines as they travel throughout the ponds. Evapo-concentration of the ions would be expected to induce

changes in the microbial communities, even if the precipitation thresholds for the above mentioned minerals are not reached. Changes in microbial community structure have previously been shown to be related to changes in the ionic composition of the waters (Langenheder and Ragnarsson 2007, Dupont 2014). Understanding the changes in salinity in a system such as Lake MacLeod, where ionic concentrations can be expected to vary over small scales, is important when trying to elucidate the roles that environmental conditions, and species-sorting processes, have on microbial community structure.

Sediment structure has also been found to play an important role in structuring benthic microbial communities (Andrushchyshyn *et al.* 2009), making the characteristics and distributions of sediments another important habitat component. Shepherd (1990) provides the only analysis of the sedimentology within the pond environments at Lake MacLeod. Within the ponds, the sediments are mostly composed of pelletal-skeletal sands, although organic and aragonite mud also forms important components. Pelletal-skeletal sands consist primarily of biogenic material, which is incorporated into the sediment column from deceased organisms.

The most abundant biogenic sedimentary products are pellets, which are usually white to light grey, and well rounded. Some of these pellets reach 0.5mm in length and are produced by faecal matter and cyanobacterial boring of carbonate grains such as *Acetabularia* sp. rods and foraminifera tests (Shepherd 1990). *Acetabularia* sp. rods are also white, but have a cylindrical shaped and reach 1.5 mm in length and 0.3 mm in diameter. These rods are derived from the breakdown of the aragonite stem fragments that the green alga grows. Foraminifera tests, which belong to the genus *Spirolina* (Shepherd 1990), can reach 3 mm in diameter and are usually circular with a coiling pattern. Like *Acetabularia*, *Spirolina* organisms are likely to occur most abundantly throughout the littoral margins of the ponds amongst the epiphytic growth on the mangrove pneumatophores and roots, however, they probably also inhabit the biofilm matrix on the sediment surface. Other biogenic contributors to the sediment column include cyanobacterial fibres, seagrass remains, mangrove peat, and skeletal remains from invertebrates, including gastropods (*Marginella* sp. and *Hydrococcus* sp.), ostracods, bivalves and polychaetes, all of which have been found in the ponds of Lake MacLeod (McLure 2011).

The chemical and physical compositions of salt lake systems define the habitat and determine the nature of the biological communities (De Deckker 1983, Boggs *et al.* 2006, Timms 2009). In turn, the biological communities determine the physical structure of the sediments (Schnurrenberger *et al.* 2003) and possibly even the water chemistry at the sediment-water interface (Woodruff *et al.* 1999, Gainswin 2004). The aims of this chapter are therefore to investigate and describe the physical (sedimentology) and chemical (salinity and nutrients) habitat conditions at the sediment-water interface in the ponds that are likely to be important determinants of microbial community structure. These results will provide insight into how pond morphology and hydrology can influence the physical and chemical habitats found within salt lakes, as well as the relationship between sediment and ionic compositions in an evaporative systems. In this chapter it is hypothesised that there is a relationship between the physical and chemical habitat characteristics of a pond with its morphometric and hydrological features. The results of this chapter will be used in the following chapters to investigate the biological effects of these environmental changes.

3.2 Methods

3.2.1 Sampling design

Samples were collected along the transects described in Chapter 1 (section 1.3.2). Samples were collected using 70 mL sterile jars. The jars were first opened underwater, in proximity to where the sediment scoop was to be taken. The scoops were taken to include only the top layer of the sediment (approximately the top 5mm), which included as much as possible, the biofilm and the water immediately above it. Samples were not meant to remain as intact sediment profiles. In some of the deeper ponds, SCUBA was required for sampling, however, most samples were collected by snorkelling when the sediment was shallower than arms reach. Care was taken to not disturb the sediment and to avoid cross-contamination of nearby samples. This was done by the sampler floating above the sediment surface in a manner which minimised disturbance and suspension of the delicate biofilms. Care was taken to gently scoop the top layer of sediment so that resuspended material was minimised. All benthic disturbance in the pond was minimised by having only a single person sampling in the water body at a time, except when SCUBA was employed in the deeper regions and a swimmer was required to

transport samples and establish transects. Extreme care was taken to ensure samplers did not disturb the sediments, and if there was a disturbance, the transects were orientated in a location where there were no disturbances. The samples were transported on ice to a field laboratory on the day of collection, then kept cold in a fridge for 24 hours to allow particles in the water to settle. After 24 hours, water overlying the samples was decanted and remaining sediment was frozen in the collection jar. The decanted liquid was stored at -20°C and used for water chemistry measurements (this chapter) and the sediment sample used for sediment characterisation (this chapter) and subsequent DNA analysis of bacterial (Chapter Four) and ciliate (Chapter Five) diversity.

3.2.2 Microscopy

A small sub sample of sediment, approximately 5g, was placed on a petri dish and spread into a thin layer. A random dot sheet was placed beneath the petri dish and an optical microscope used to count 135 particles that were above the dots. The particles were visually inspected to determine if they were pellets, skeletal remains, organic detritus, aragonite mud or diatoms. These particles were defined as per Shepherd (1990). Pellets are white to light grey, rounded aggregates of aragonite, and were the most common material found in the samples. *Acetabularia* remains were included in this category. Skeletal remains included the remains of organisms such as foraminifera, arthropods, gastropods and bivalves, while organic detritus included remains of seagrasses, algae and mangroves. Aragonite mud consisted of aggregations of small crystals bound in a matrix of cyanobacteria mucilage. Using accumulation curves, it was found that counting more than 100 points and replicating within samples yielded no additional particle species.

3.2.3 Water chemistry analysis

Decanted water samples were removed from -20°C and defrosted at 5°C overnight. Using an unused 20mL syringe, approximately 20mL of sample water was removed and filtered through a 0.45µm syringe filter. 15mL of the filtrate was used for soluble reactive phosphorus (SRP) and nitrate/nitrite (NO_x) analysis. 100µL was diluted with 10mL of milliQ water for SO_4^{2-} and Cl^- analysis. A further 250µL of sample was added to 10mL of 2% nitric acid solution prepared by adding 200µL of concentrated HNO_3

with 10mL of millQ water for cation analysis (Na^+, Mg^+ Ca^+, K^+). All samples were stored at 5°C until analysis was done using an Inductively Cooled Plasma Mass Spectrometer (ICP-MS).

3.2.4 Statistical analyses

Because the sedimentary data did not conform to normality, non-parametric methods were used. Using the *vegan* package for R (Oksanen *et al.* 2013), Analysis of Similarities (ANOSIM) and Permutation Multivariate Analysis of Variances (PERMANOVA) were applied to test for differences between the sedimentology of the ponds (Clarke 1993). In order to classify different groups of sediments, k-means clustering was also done using the *vegan* package. The Simple Structure Index (SSI) was used to determine the quality of the clusters. SSI combines the three elements, the maximum difference of each variable between clusters, the size of the most contrasting clusters and the deviation of the variable in the cluster centres, in order to assess cluster interpretability (Borcard *et al.* 2011). The names of each sediment group were assigned using the two most abundant components of the sediment group.

Salinity measurements are expressed as milliequivalents (meq). As the sum of the ionic concentrations equals a constant, the data were treated as a composition (van den Boogaart and Tolosana-Delgado 2013), and log-ratio transformations used to conduct Linear Discriminant Analysis (LDA). Using log-ratios allows for tests between groups while accounting for the effect of correlated variables. Compositional data analysis, including transformations, were conducted using the compositions package (van den Boogaart *et al.* 2014) and LDA using MASS package (Ripley *et al.* 2014) in R. Inverse Distance Weighted Interpolation (IDW) and graphical techniques were performed using Euclidean distances and an inverse distance weighting power of 2 in the *gstat* (Pebesma and Graeler 2014) and *sp* packages in R (Pebesma *et al.* 2015). Descriptive statistics was used to relate the environmental variables measured in this chapter with the pond morphotypes described in the previous chapter.

3.3 Results

3.3.1 Sedimentology

In total, 456 samples were collected from Pete's Vent (n=47), Donut Pond (n=46), Annie's Pond (n=48), Pete's Pond (n=47), Whistler's Pond (n=63), Harjie's Pond (n=62), Jana's Vent (n=59) and Cygnet Pond (n=84). These samples were used for sediment and water chemistry analysis. These samples represent the physical and chemical characteristics found at the sediment-water interface, which are thought to be important in determining microbial biofilm communities. The mean number of pelletal particles observed per sample was 97, and was much higher than the mineral, skeletal, diatomaceous and detrital components (mean numbers per sample of 15, 11, 10 and 3, respectively). The frequency distribution of pelletal particles per sample is negatively skewed (Figure 20), whilst all other components have positively skewed frequency distributions (Figure 21).The only component, with the exception of pellets, to be unimodal in frequency greater than 0 was the component skeletal fragments.

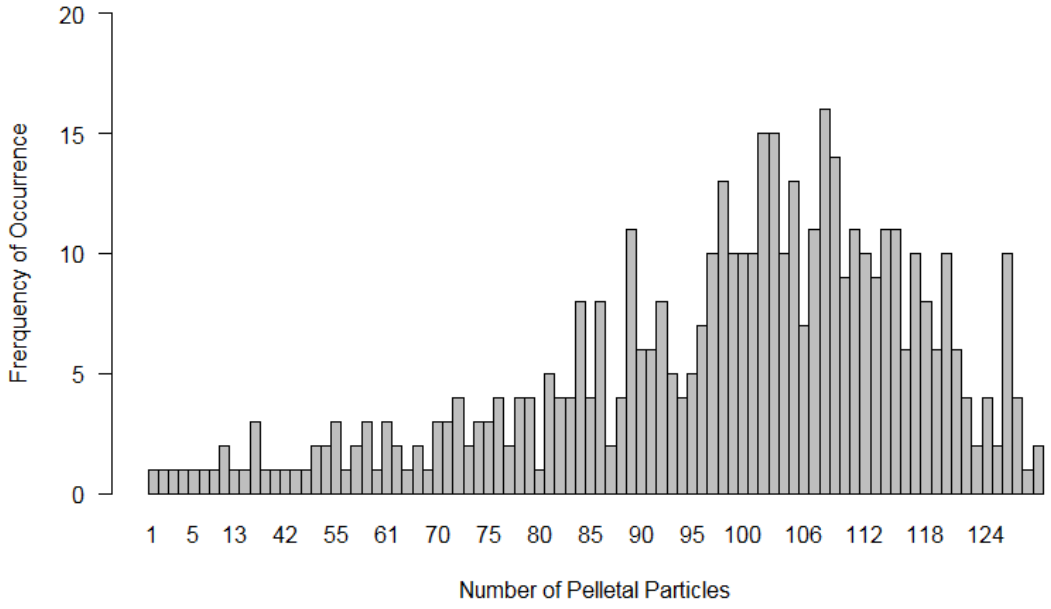


Figure 20: Barplot showing the frequency of observations for the pelletal component of the 457 sediment samples taken from the eight study ponds. The mean number is 97 ± 23 (mean ± SD; n=457).

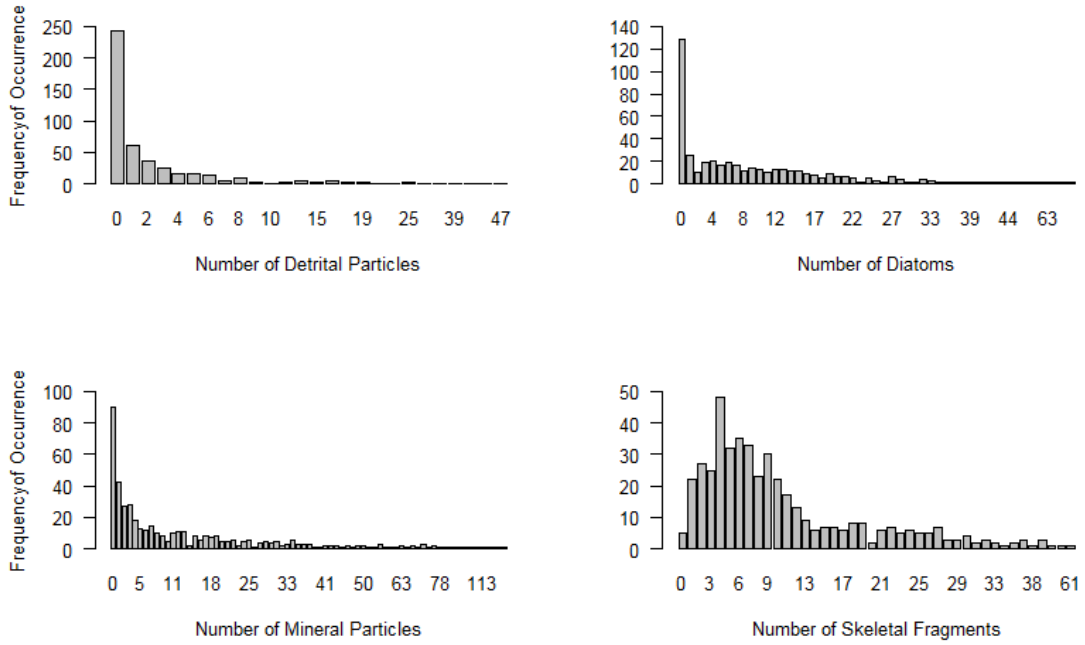


Figure 21: Barplots showing frequency of observations of detrital, diatomaceous, mineral and skeletal fractions of the sediments from the 457 sediment samples taken from the eight study ponds. The mean number of detrital particles was 3 ± 6. The mean number of diatoms was 10 ± 12. The mean number of mineral particles was 15 ± 22. The mean num,ber of skeletal fragments was 11 ± 9. All values are mean ± SD, n=457.

Pond differences

The pelletal composition of each pond was relatively constant. In all the ponds, it was the major component of the sediment, representing approximately 80% of the counted particles. The skeletal composition in the sediments was greatest for the samples belonging to Harjie's Pond, then Donut Pond. The other six ponds had relatively low counts of skeletal fragments. Detrital matter was highest in Donut Pond, followed by Cygnet Pond and moderately high in Pete's Vent and Jana's Vent. Annie's Pond, Harjie's Pond, Pete's Pond and Whistler's Pond all had very low amounts of detrital matter. Diatom frustules were very uncommon in Cygnet Pond and Jana's Vent sediments, but relatively common in the other ponds. Mineral components were particularly low in Annie's Pond and Whistler's Ponds, and very high in Cygnet Pond.

The similarity of the sediments in ponds was greater within rather than between ponds (ANOSIM; R=0.36, P=0.001), suggesting that the ponds have different sediment compositions (Figure 22). This result was reinforced by the PERMANOVA results, which indicated that sediment compositions differed significantly by pond (P=0.001), however, the assumption of homogeneity of variances was violated as there was a

significant difference in multivariate spread between ponds ($F_{7,448}$=20.2, P<0.001). Care needs to be taken when considering the PERMANOVA results as the differences in variance between the ponds may contribute to the significant differences found. In particular, the spread within samples belonging to Cygnet and Donut Ponds was high (the mean distances to centroids were 0.17 and 0.20, respectively), compared to the other ponds (between 0.04 and 0.09).

Sediment groups

Using k-means clustering, and the Simple Structure Index (SSI), it was found that there were six groups of sediments that showed good cluster quality (Figure 23). Although eight groups gave a greater SSI value, it was decided that some of these groups contained too few samples (<10 samples) and were of little use. Three groups gave the best SSI, however, this was inadequate as it only gave a single large group and two small groups. Using six groups meant that the group sizes were large enough to be meaningful, whilst maintaining good cluster quality, as determined by the high SSI value.

The groups are described based on the composition of the sediments within each group, with the main components used to define the sediment type. The most common sediment components, in order of abundances, were pellets, aragonite mud, skeletal fragments, diatomaceous material and organic detritus. These components are mostly autochthonous, being derived from biological and/or chemical processes operating within the lake itself. Few allochthonous particles were found in the biofilm and sediments immediately beneath; these particles consisted of mostly sand grains, probably transported into the seepage face from the surrounding Quobba sands formation.

The different sediment groups, identified by k-means clustering, were different in their relative composition of pellets, aragonite mud, skeletal fragments, diatomaceous material and organic detritus (Figure 24). The most common sediment types were Pelletal Sand (48%), Pelletal Sand and Aragonite Mud (20%) and Pelletal-Skeletal Sand (13%; Figure 25). The remaining sediment types consisted of Pelletal-Diatomaceous Sand (9%), Pelletal-Aragonite Mud (7%) and Aragonite Mud (3%). There were no clear patterns or trends of sediment distributions within the ponds.

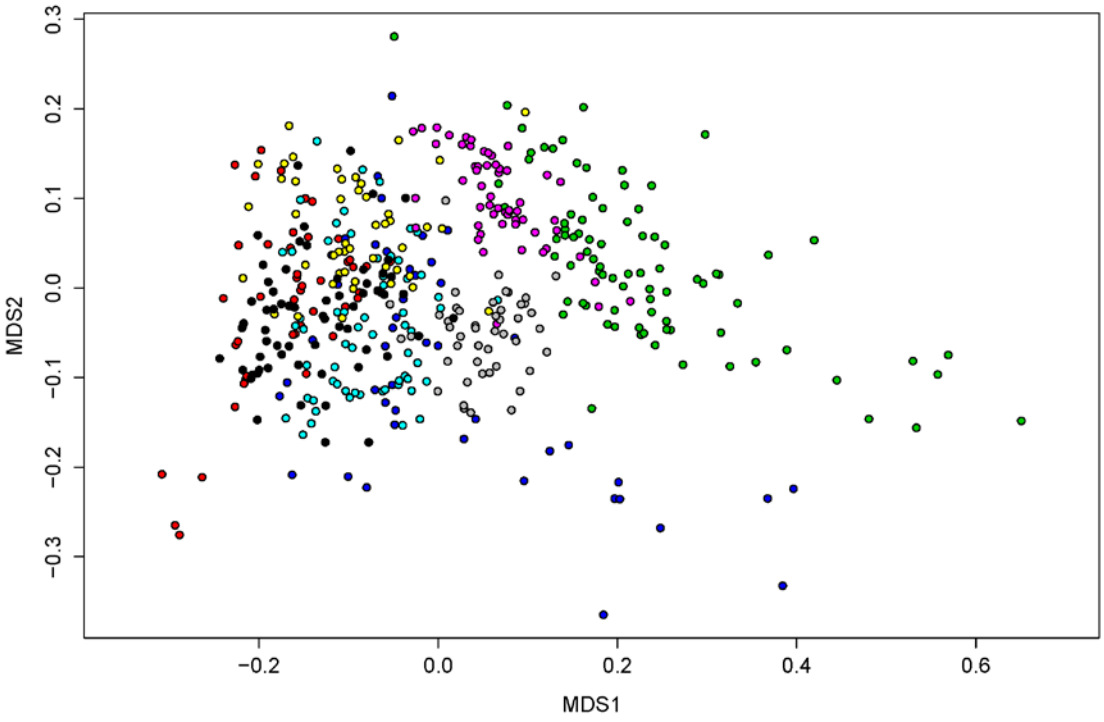


Figure 22: Non-metric Multi-Dimensional Scaling (NMDS) plot showing the similarity between sediment samples between Ponds determined by Euclidean distances. Red: Annie's Pond; Green: Cygnet Pond; Blue: Donut Pond; Light Blue: Harjie's Pond; Purple: Jana's Vent; Yellow: Pete's Pond; Grey: Pete's Vent; Black: Whistler's Pond. Stress=0.12.

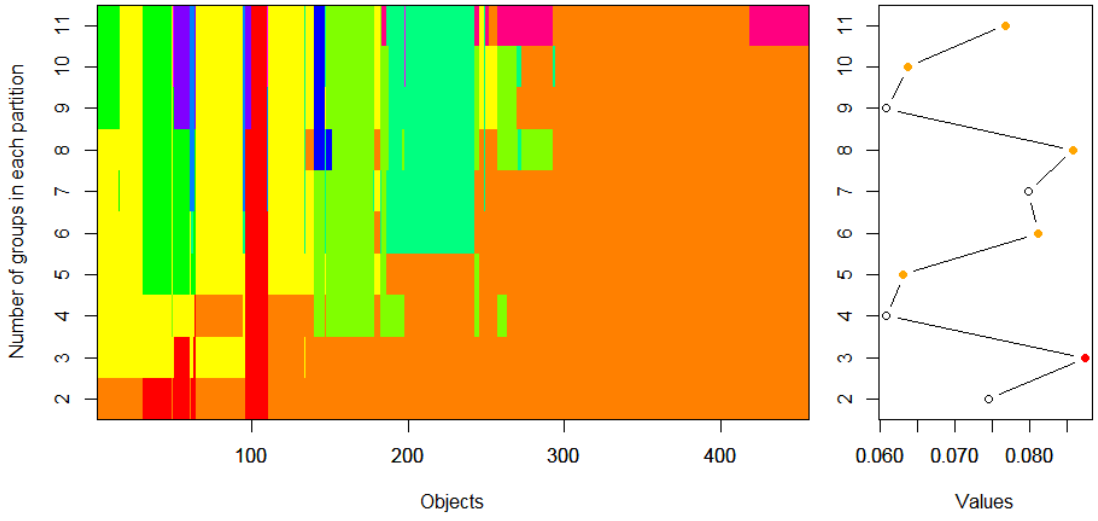


Figure 23: k-means cascade plot showing the group attributed to each object (sample) for each cluster partition. The graph on the right shows the values of the Simple Structure Index (SSI) for determining the best number of partitions. The highest SSI value is marked in red. Points in orange indicate partions that showed an increase in SSI as the number of groups increased. The best number of clusters was 3, followed by 8 and 6. Because 8 clusters has some small groups, and 3 clusters is dominated by a single large cluster, 6 clusters was chosen to represent the variability in the sediment data.

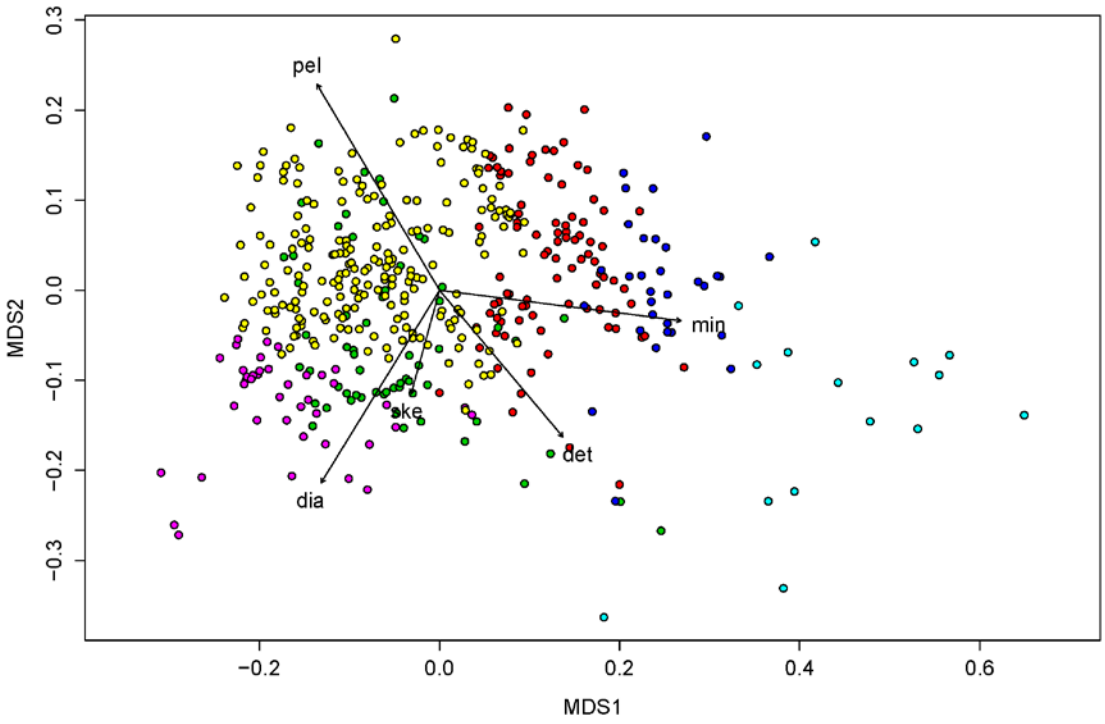


Figure 24: Non-metric Multi-Dimensional Scaling (NMDS) showing the relationship of each sediment group, determined using k-means partitioning, with the sediment components. Arrows represent the direction of high composition of the different sediment components, pelletal (pel), diatoms (dia), skeletal (ske), detrital (det) and mineral (min). Yellow: Pelletal Sand; Red: Pelletal Sand & Aragonite Mud; Green: Pelletal Skeletal Sand; Purple: Pelletal Diatomaceous Sand; Blue: Pelletal-Aragonite Mud; Light Blue: Aragonite Mud. Stress = 0.12.

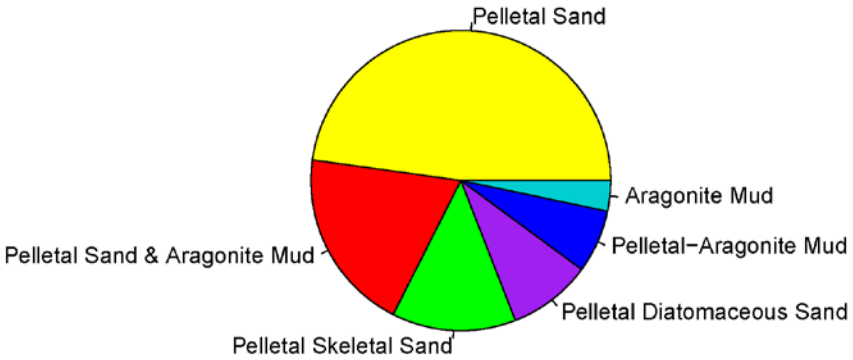


Figure 25: Pie chart showing the relative proportions of each sediment type across all the ponds studied.

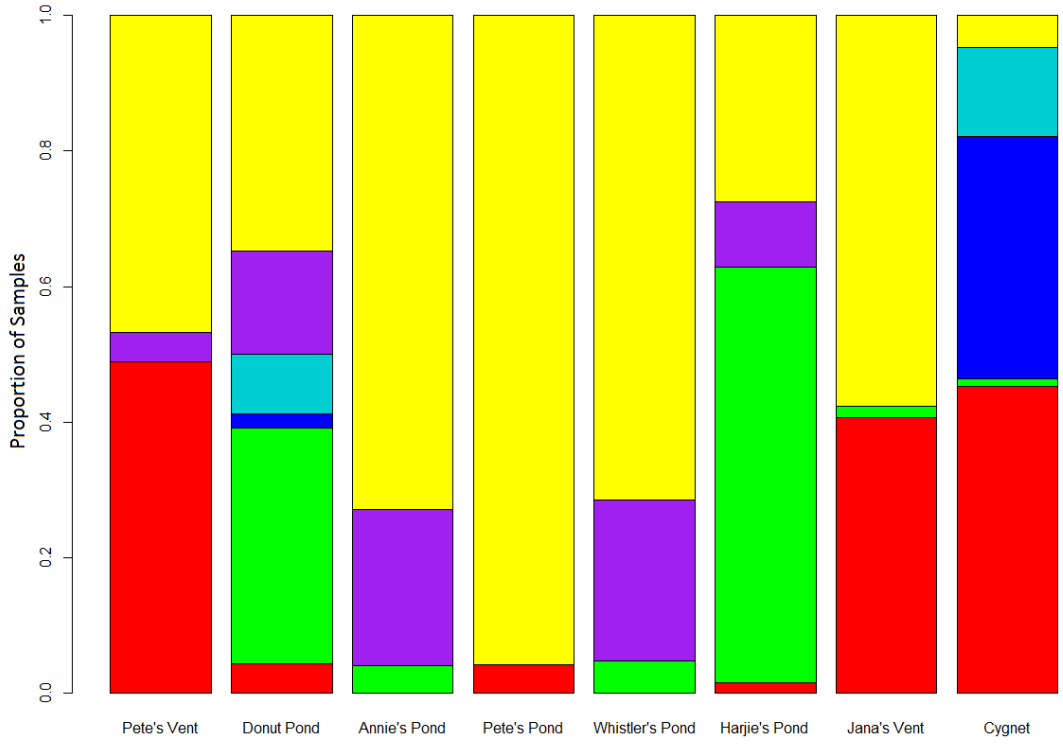


Figure 26: Distribution of the six sediment groups determined by k-means clustering. Red: Pelletal Sand & Aragonite Mud, Green: Pelletal Skeletal Sand, Blue: Pelletal-Aragonite Mud, Light Blue: Aragonite Mud, Purple: Pelletal Diatomaceous Sand, Yellow: Pelletal Sand.

Sediment composition and distribution

The proportion of each sediment type within each pond was different. Some ponds, such as Pete's Pond were relatively homogeneous, being dominated by pelletal sand, while other ponds, such as Donut Pond, contained up to six sediment types. Following is a short description of the composition and distribution of each sediment type defined in this study.

Pelletal-Aragonite Mud

Pelletal-Aragonite Mud was found only in the southern Cygnet seepage area, and almost exclusively within Cygnet Pond. It consists of 40-55% pellets, 30-50% Aragonite with 0-10% Skeletal fragments and plant detrital matter.

Aragonite Mud

This sediment type consists of 50-90 % Aragonite Mud, 1-20% pellets and 5-20% plant detrital matter. Similar to Pelletal-Aragonite Mud, this sediment type is exclusive to the southern Cygnet seepage face, and most common in Cygnet Pond, although it was found in Donut Pond.

Pelletal-Diatomaceous Sand

Sediments belonging to this group consist of 50-70% pellets, 20-40% diatomaceous matter and 3-15% skeletal fragments. Pelletal-Diatomaceous Sand is found most abundantly in Whistler's Pond and Annie's Pond, although it was also present in Donut Pond and Harjie's Pond.

Pelletal Skeletal Sand

Pelletal-Skeletal Sand was mostly found in Harjie's Pond, although it was also a major component of Donut Pond. It was a minor component of Whistler's Pond, Annie's Pond, Cygnet Pond and Jana's Vent. Pellets were the major component (55-80%) followed by skeletal fragments (15-25%) and diatomaceous matter (1-10%).

Pelletal Sand

Pelletal Sand is mostly composed of pellets (75-90%), diatomaceous matter (2-15%) and skeletal fragments (2-10%). Pelletal Sand was the only sediment type to be found in all ponds, however it was present in different abundances. It was mostly found in Pete's Pond, and also in Annie's Pond, Whistler's Pond and Jana's Vent in high abundance.

Pelletal Sand and Aragonite Mud

Pelletal Sand and Aragonite Mud differ from Pelletal-Aragonite Mud by the much larger proportion of pellets (60-80%) and lower proportion of Aragonite Mud (15-25%). Skeletal fragments are also present (2-8%). This sediment type was mostly found in Cygnet Pond, but also in Pete's Vent and Jana's Vent. It is absent from Annie's Pond and Whistler's Ponds.

3.3.2 Water chemistry

Chloride showed the largest variation in concentration, ranging from 9,033 mg L^{-1} to 34,295 mg L^{-1}, while sodium showed the least variation, ranging from 6,147 to 15,492 mg L^{-1}. The nutrients, ammonium, nitrate/nitrite and phosphate all showed larger levels of variation in concentration than the remaining ions. All three nutrient variables had positively skewed distributions (Table 7), and thus log-transformations of the data set were required to increase normality. The major ions, sodium, magnesium, calcium, potassium, chloride and sulphate displayed bimodal distributions. Because of their bimodal distributions, and the fact that the summation of ionic concentration should equal a constant, the water chemistry dataset were treated as compositions.

The mean concentration of sodium was 456 ± 83 meq, magnesium was 100 ± 28 meq, calcium was 21 ± 7 meq and potassium was 8 ± 2 meq. The mean concentrations of chloride and sulphate were 516 ± 136 meq and 53 ± 21 meq, respectively. All these ion species had large variances in respect to the means, suggesting that ionic concentrations vary between ponds (Table 8). The sum of equivalents was close to zero, ranging between -0.52 and 0.36 meq, with a mean of 0.017 ± 0.13 meq.

Table 7: Summary statistics for water chemistry data including: minimum and maximum values, 1st and 3rd quartiles, and mean and median values. All units are mg L^{-1}.

	Na^+	Mg^+	Ca^+	K^+	Cl^-	SO_4^{2-}	NH_4^-	NO_x^-	PO_4^{3-}
Minimum	6,147	629	215	76	9,033	743	0.004	0.003	0.006
1st Quartile	9,052	989	342	271	14,936	1,940	0.086	0.018	0.013
Median	10,292	1,132	391	323	17,625	2,337	0.221	0.034	0.017
Mean	10,474	1,216	429	308	18,304	2,529	0.259	0.037	0.023
3rd Quartile	11,804	1,341	456	372	20,343	2,732	0.360	0.051	0.024
Maximum	15,492	2,224	899	471	34,295	11,033	1.083	0.107	0.126

Table 8: Summary statistics of major ions across the eight ponds including: minimum and maximum values, 1st and 3rd quartiles, and mean and median values. All units are millequvalents. PV: Pete's Vent; D: Donut Pond; A: Annie's Pond; PP: Pete's Pond; W: Whistler's Pond; H: Harjie's Pond; J: Jana's Vent; C: Cygnet Pond. *n* represents the number of samples collected from each pond.

	PV	D	A	PP	W	H	J	C
	Sodium							
Min	32	27	35	31	26	27	37	45
1st Qua.	392	417	430	386	352	351	466	547
Median	417	480	451	410	392	393	501	585
Mean	422	471	453	415	400	403	495	585
3rd Qua.	446	528	488	434	446	455	526	624
Max	534	578	559	512	546	556	588	674
	Magnesium							
Min	69	56	72	64	52	55	82	125
1st Qua.	81	86	86	79	70	71	103	155
Median	84	101	89	83	84	81	108	160
Mean	85	98	92	84	84	83	109	160
3rd Qua.	89	112	101	88	95	97	118	169
Max	111	121	111	110	116	114	133	183
	Calcium							
Min	15	12	14	13	11	11	16	18
1st Qua.	17	17	17	16	17	16	20	34
Median	18	20	19	16	19	18	21	38
Mean	18	20	20	17	19	18	21	37
3rd Qua.	19	23	22	18	22	20	23	41
Max	24	25	26	22	25	26	26	45
	Potassium							
Min	7	5	6	6	5	5	8	2
1st Qua.	8	8	8	7	7	7	10	3
Median	8	10	9	7	8	9	10	4
Mean	8	9	9	7	8	8	10	4
3rd Qua.	9	10	10	7	9	9	11	4
Max	11	11	12	9	11	12	12	6
	Chloride							
Min	295	367	313	255	267	304	313	423
1st Qua.	432	437	479	357	379	388	469	647
Median	484	508	545	382	432	477	513	676
Mean	554	495	611	402	426	464	513	684
3rd Qua.	715	544	772	437	483	517	551	729
Max	883	608	967	689	547	716	673	879
	Sulphate							
Min	26	30	34	27	24	27	22	15
1st Qua.	40	45	50	44	35	36	46	52
Median	51	48	74	49	37	45	50	64
Mean	55	50	71	49	39	44	60	60
3rd Qua.	77	53	91	54	44	49	55	69
Max	93	128	101	73	54	89	230	85
n	47	46	48	47	63	62	59	84

Between pond variability

Chloride was the dominant ion in all the ponds, except for Pete's Pond, where sodium levels were slightly greater. Levels of potassium were much lower in Cygnet Pond than the other ponds, whilst magnesium and calcium levels were higher (Table 9). The first linear discriminant explained 96.5% of the variation and separated the samples into two, those belonging to Cygnet Pond, and those belonging to the rest of the ponds, a trend also found using principal components (Figure 27). This discriminant was largely influenced by the magnesium ratio with potassium and sodium (Table 10).

Table 9: Percentage composition coefficients of the major ions for each pond determined using LDA.

	Na	Mg	Ca	K	Cl	SO_4
Pete's Vent	37.7	7.6	1.6	0.7	47.7	4.7
Donut Pond	41.2	8.6	1.7	0.8	43.4	4.3
Annie's Pond	36.7	7.5	1.6	0.7	48.0	5.5
Pete's Pond	42.9	8.7	1.7	0.7	41.0	5.0
Whistler's Pond	41.0	8.6	1.9	0.8	43.7	4.0
Harjie's Pond	39.6	8.2	1.8	0.8	45.4	4.3
Jana's Vent	41.3	9.1	1.8	0.8	42.6	4.5
Cygnet Pond	38.4	10.5	2.4	0.2	44.7	3.8

Table 10: Coefficients of the major ions with the linear discriminants with the proportion of explained variance for each of the five discriminates. The proportion of trace represents the proportion of between-ion variation that is explained by each LD axis.

	LD1	LD2	LD3	LD4	LD5
Sodium	-5.4	-4.5	11.8	-12.3	-7.0
Magnesium	12.3	-9.0	-11.3	10.9	-1.3
Calcium	1.8	9.8	-0.4	-3.5	1.8
Potassium	-9.5	1.3	-2.9	1.5	1.4
Chloride	0.7	4.9	1.4	2.4	-4.8
Sulphate	0.04	-2.6	1.3	1.1	3.9
Proportion of trace	96.5	2.0	1.0	0.4	0.2

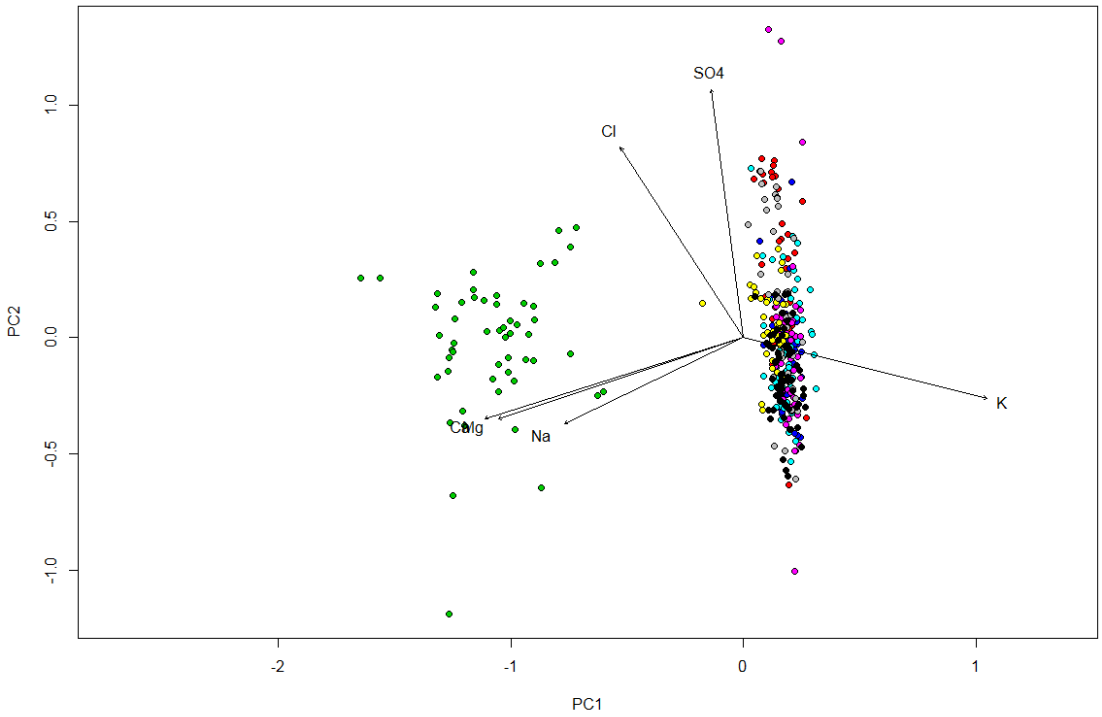


Figure 27: PCA biplot showing the differences between ponds (coloured points) with changes in water ionic composition. The first two axes explain 92% of the variation (62 and 30% for PC1 and PC2, respectively). Red: Annie's Pond; Green: Cygnet Pond; Blue: Donut Pond; Light Blue: Harjie's Pond; Purple: Jana's Vent; Yellow: Pete's Pond; Grey: Pete's Vent area; Black: Whistler's Pond.

Within pond variation

There was variation in cation concentrations within each pond (see individual pond descriptions below). Generally, the metals, sodium, magnesium, calcium and potassium decreased in concentration as water moved in an easterly direction. These changes in cation concentration relative to the inert chloride concentration suggest that the gradients are occurring because of chemical changes and/or biological processes. The changes found in ionic concentrations are described below for each pond.

Pete's Vent

The pond is dominated by a vent, and all water collected here is probably similar in composition to the water being transported through the karst barrier. However, there were areas of high concentration of the four cations in the southern part of the plot, near the vent (Figure 28). Magnesium concentration differed markedly between the north and the south of the plot, with the higher concentrations being found near the vent. The ratio

of major ions was 30:3:1:1 (Na:Mg:Ca:K). Fluctuations in ionic concentrations are likely to be from numerous small seepage points.

Donut Pond

The ratio of major ions was 30:3:1:1 (Na:Mg:Ca:K), with calcium and potassium deviating the greatest in relative concentration Generally, all ionic concentrations were lowest at the southern, shallow regions of the plot (Figure 29).

Annie's Pond

The ratio of major ions was 30:3:1:1 (Na:Mg:Ca:K), and they all showed similar rates of depletion in an easterly direction across the plot. All the cations were at highest concentration in the western section of the pond and decreased in an easterly direction, although they were at greatest concentration in the south-west section of the pond (Figure 30). The areas of highest sodium and magnesium concentration occurred where chloride and sulphate concentration was lowest. The concentration of the anions was mostly uniform throughout the pond, at 555-660 meq for chloride and 65-80 meq for sulphate, except in the south-western section of the plot, where concentrations were as low as 345 meq and 35 meq for chloride and sulphate, respectively.

Pete's Pond

Sodium and magnesium were slightly enriched compared the other ponds, with the ratio being 33:4:1:1 (Na:Mg:Ca:K). Sodium and calcium concentrations decreased in an easterly direction, whilst sulphate concentration was highest along the western margin (Figure 31). Similar to other ponds, there was slight variation in the other ions throughout the pond, but they did not show any trends.

Whistler's Pond

The ratio of major ions was similar to the other Northern Ponds, 29:3:1:1 (Na:Mg:Ca:K). Sodium, calcium and potassium displayed similar trends in concentration change, with an area of low concentration in the northern part of the plot, and higher concentration in the south (Figure 32). Magnesium concentration was the

greatest in the south-east part of the plot, whilst chloride concentration increased in a south-easterly direction.

Harjie's Pond

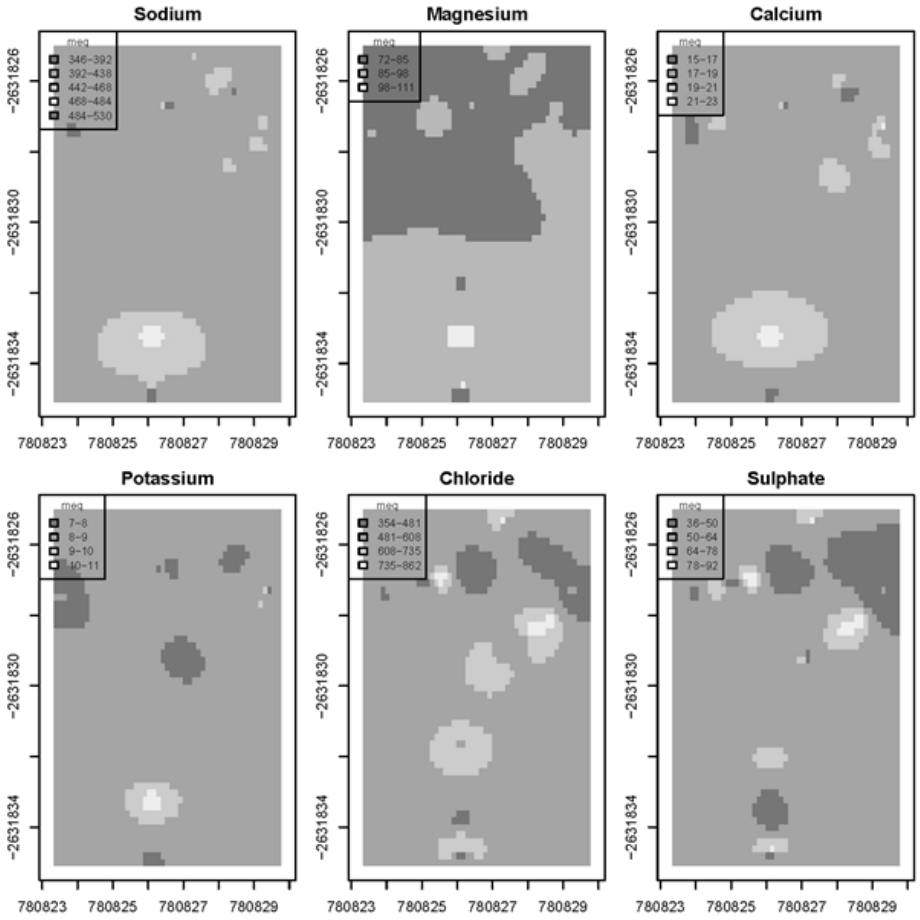
The ratio of major ions was 29:3:1:1 (Na:Mg:Ca:K). Sodium and magnesium concentrations did not differ markedly throughout the plot, but calcium and potassium did show some variation. There was a band of low calcium concentration through the centre of the plot (Figure 33). A similar pattern was seen for chloride concentration. Potassium concentration was also lowest in the centre of the plot. Sulphate levels were high in the south-western part of the plot, but were generally constant throughout the rest of the pond.

Jana's Vent

The ratio of major ions was 29:3:1:1 (Na:Mg:Ca:K). Sodium concentration was greatest in the south-eastern section of the plot, whilst potassium concentration was lowest in the north-eastern section (Figure 34). Sulphate concentration was greatest in the eastern side of the pond. Magnesium and calcium concentrations did not show any spatial trends, whilst chloride was highest in the north-western section of the plot.

Cygnet Pond

The ratio of major cations was 82:12:5:1 (Na:Mg:Ca:K), which differs markedly from the other ponds. Sodium showed a gradient of increasing concentration in a south-easterly direction, with low concentrations in the northern areas, and a point on the western margin, where an inflow vent is located (Figure 35). A further area of low concentration in the southern area may indicate another seepage area. Similar trends were seen for the other ions, with the increase in concentration of all the elements, a possible result of evapotranspiration. Potassium, however, was an exception, with changes in concentration showing the opposite pattern to the other cations. Chloride and sulphate showed very similar patterns in change in concentration, which was also similar to the cation changes. The low concentration in the southern part of the pond may indicate a seepage point, or precipitation of some ions.



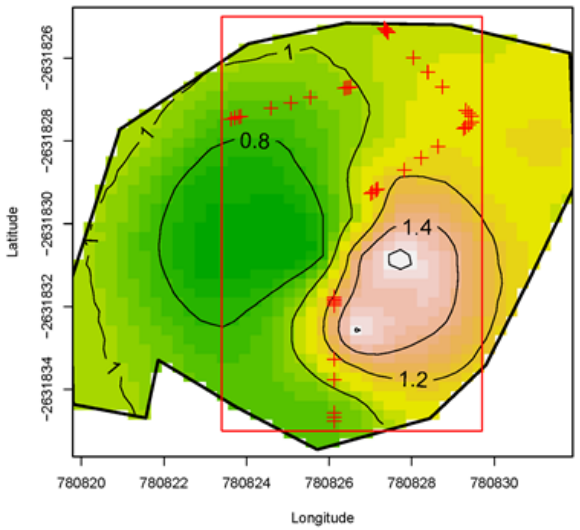


Figure 28: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Pete's Vent. n = 47.

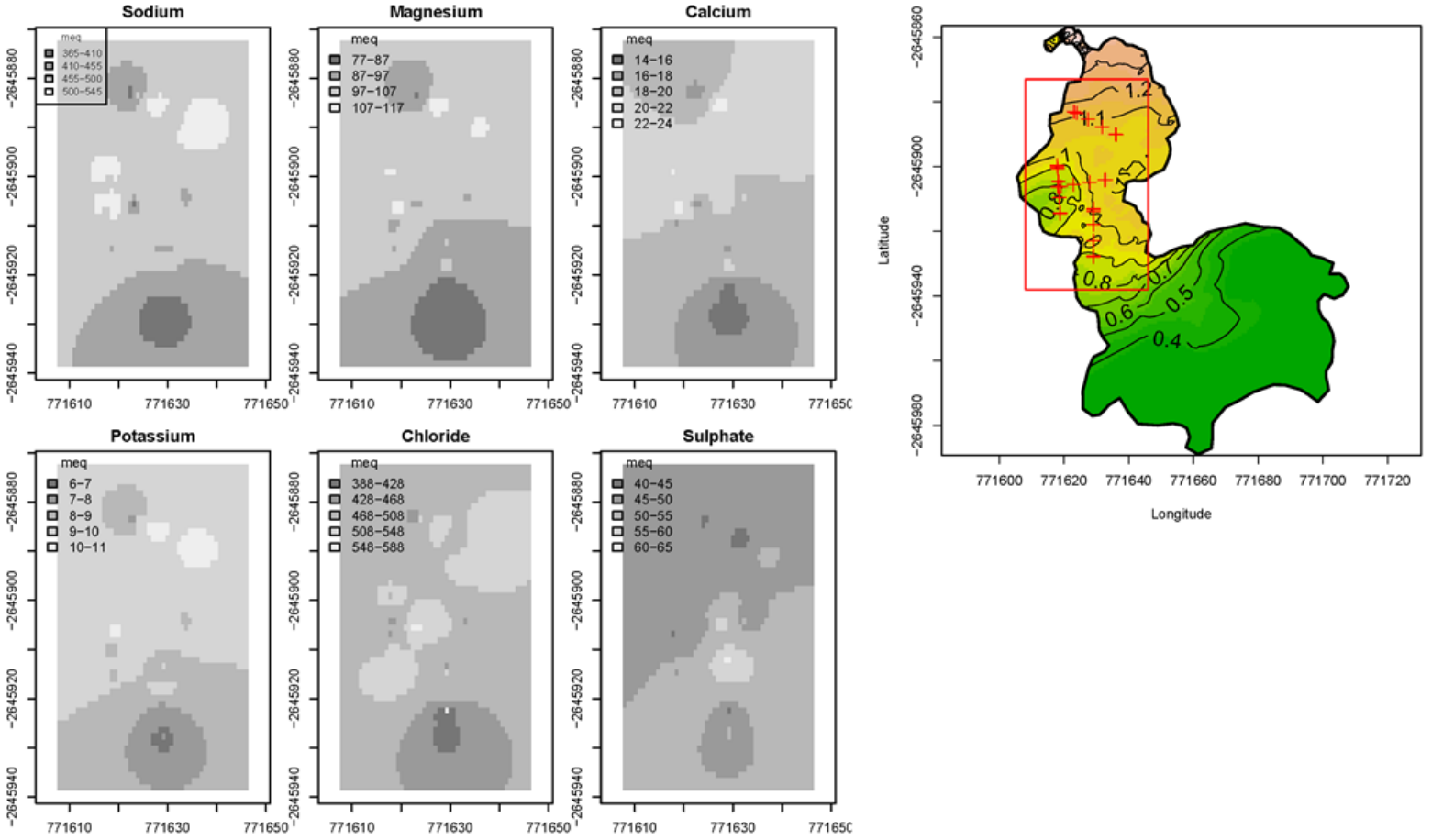


Figure 29: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Donut Pond. n = 46

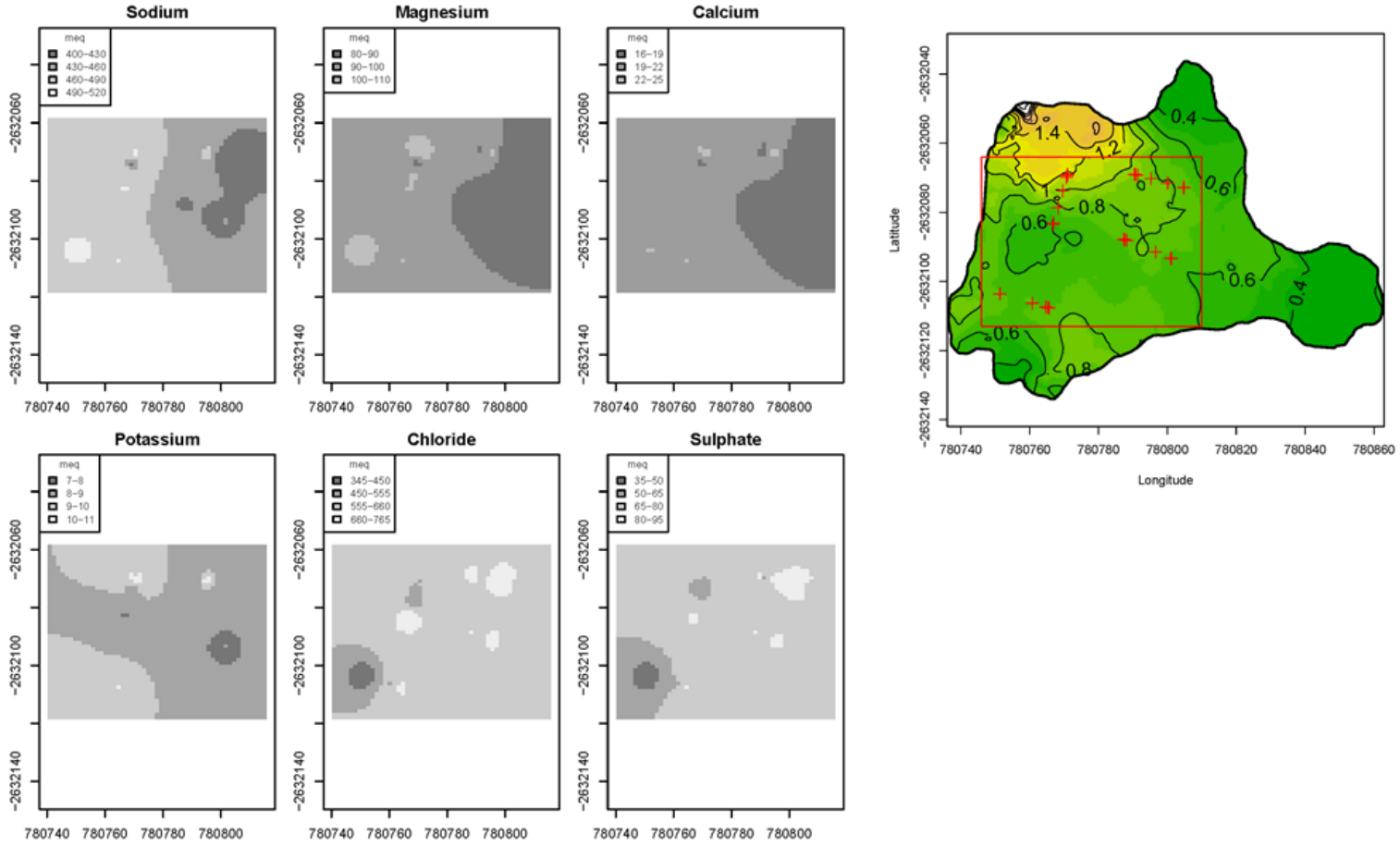


Figure 30: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Annie's Pond. n = 48

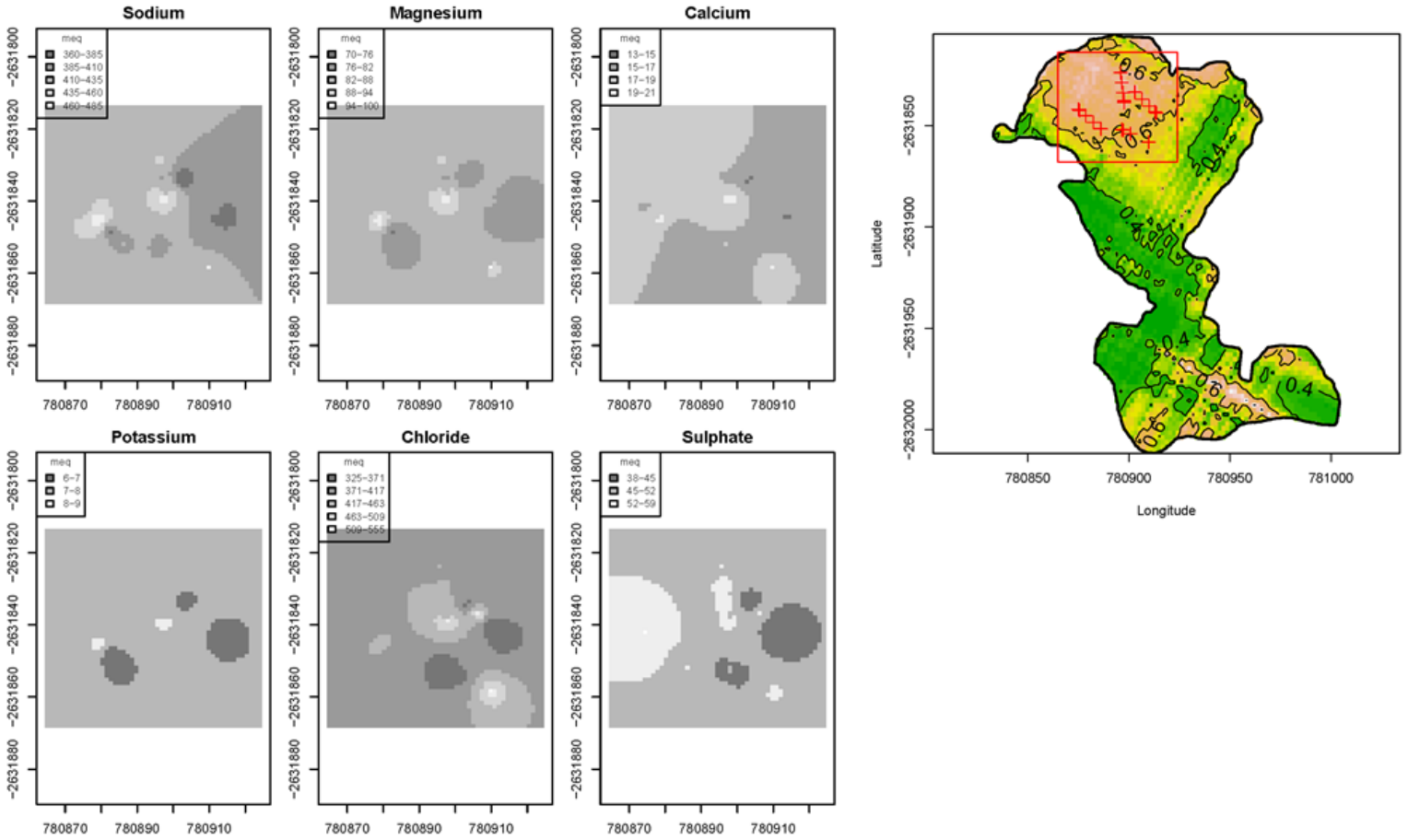


Figure 31:Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Pete's Pond. n = 47

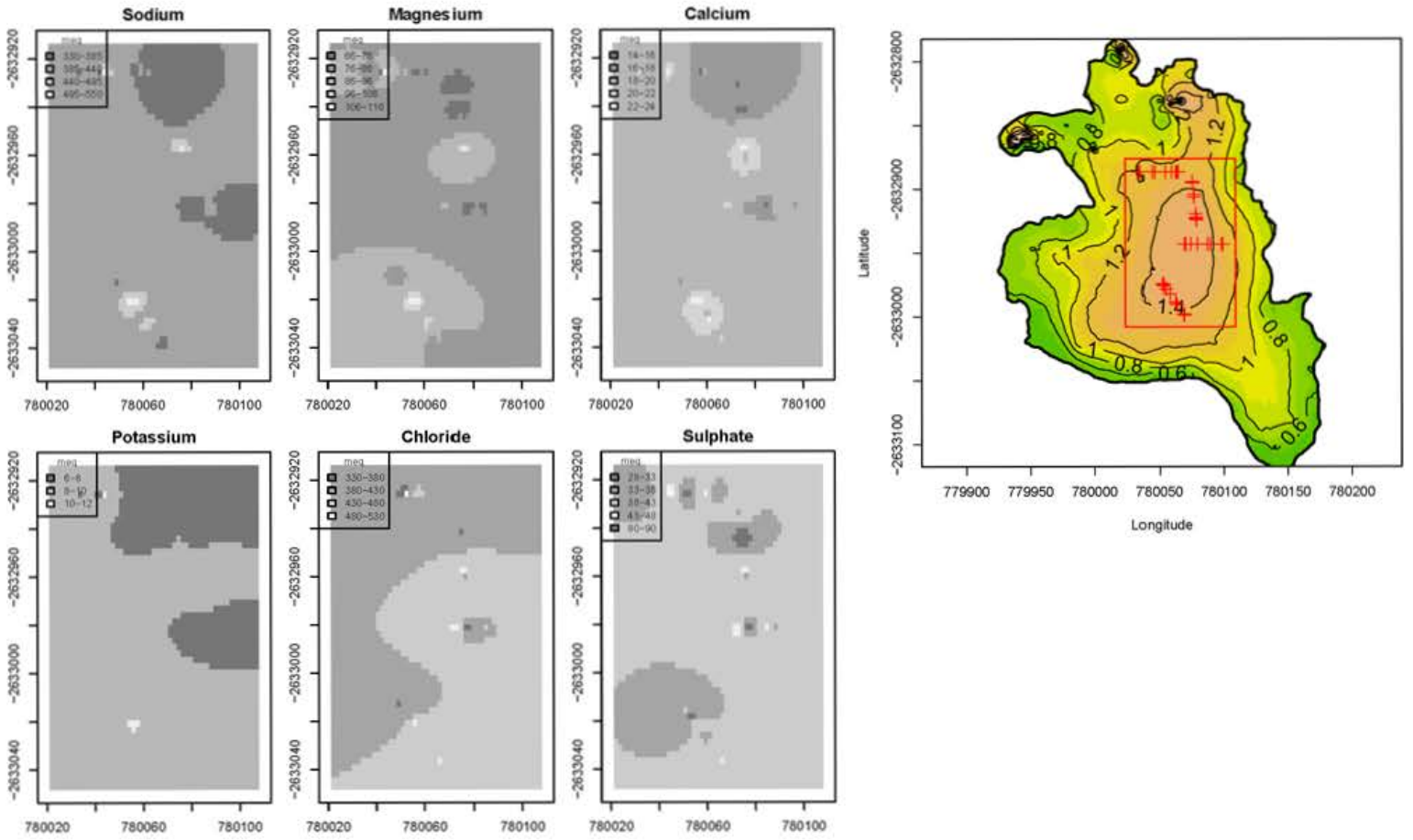


Figure 32: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Whistler's Pond. n = 63.

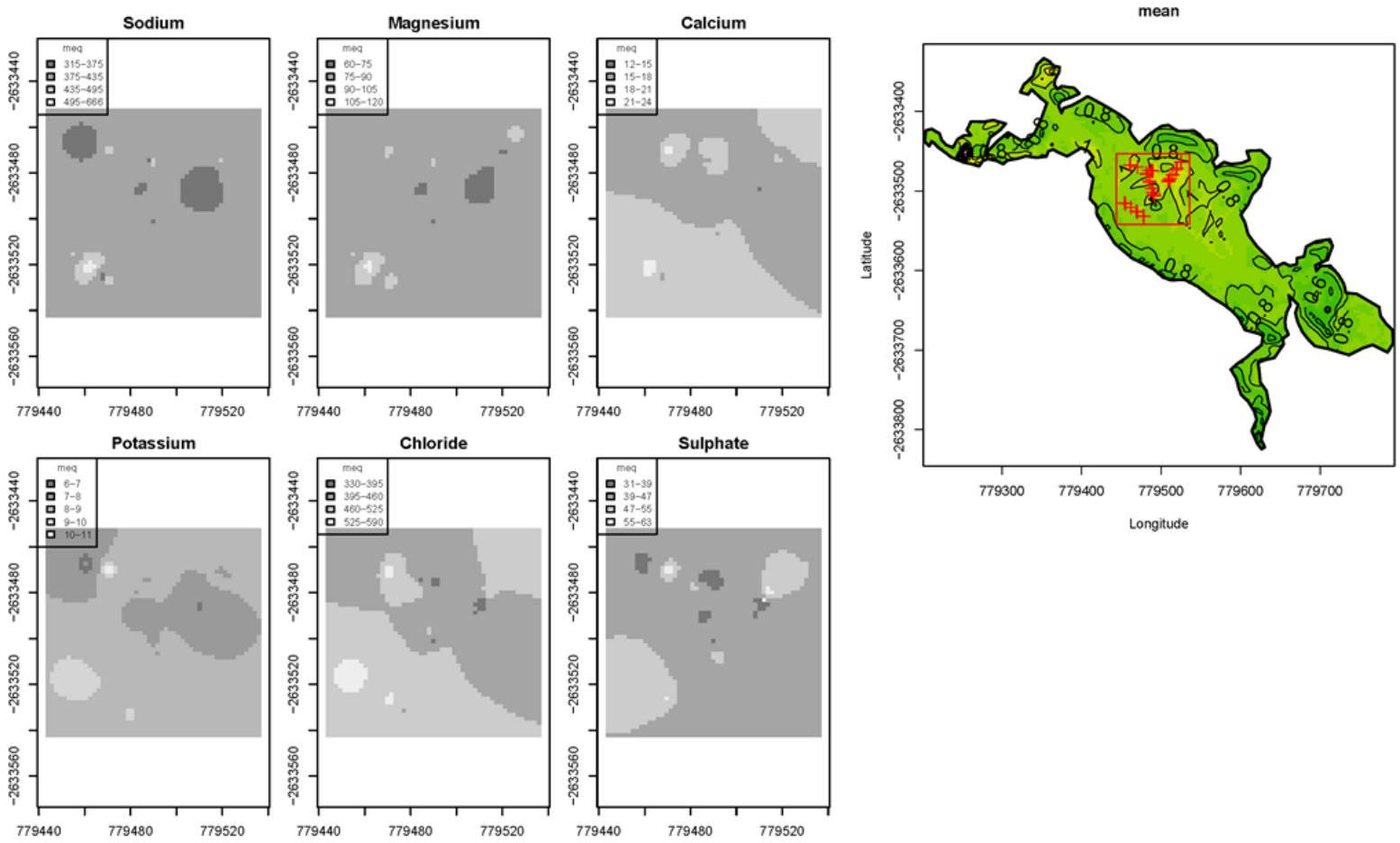


Figure 33: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Harjie's Pond. n = 62

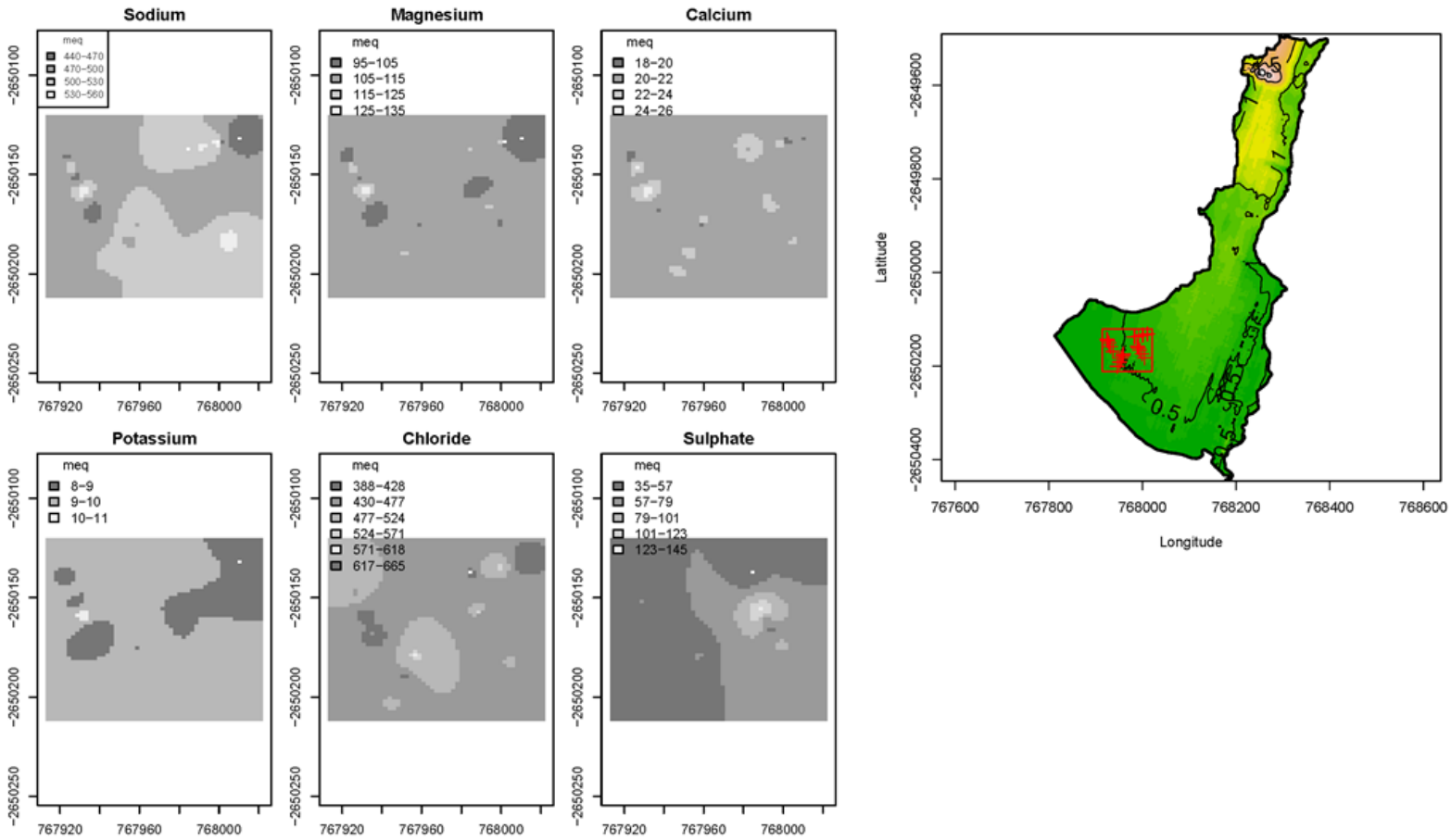


Figure 34: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Jana's Vent. n = 59.

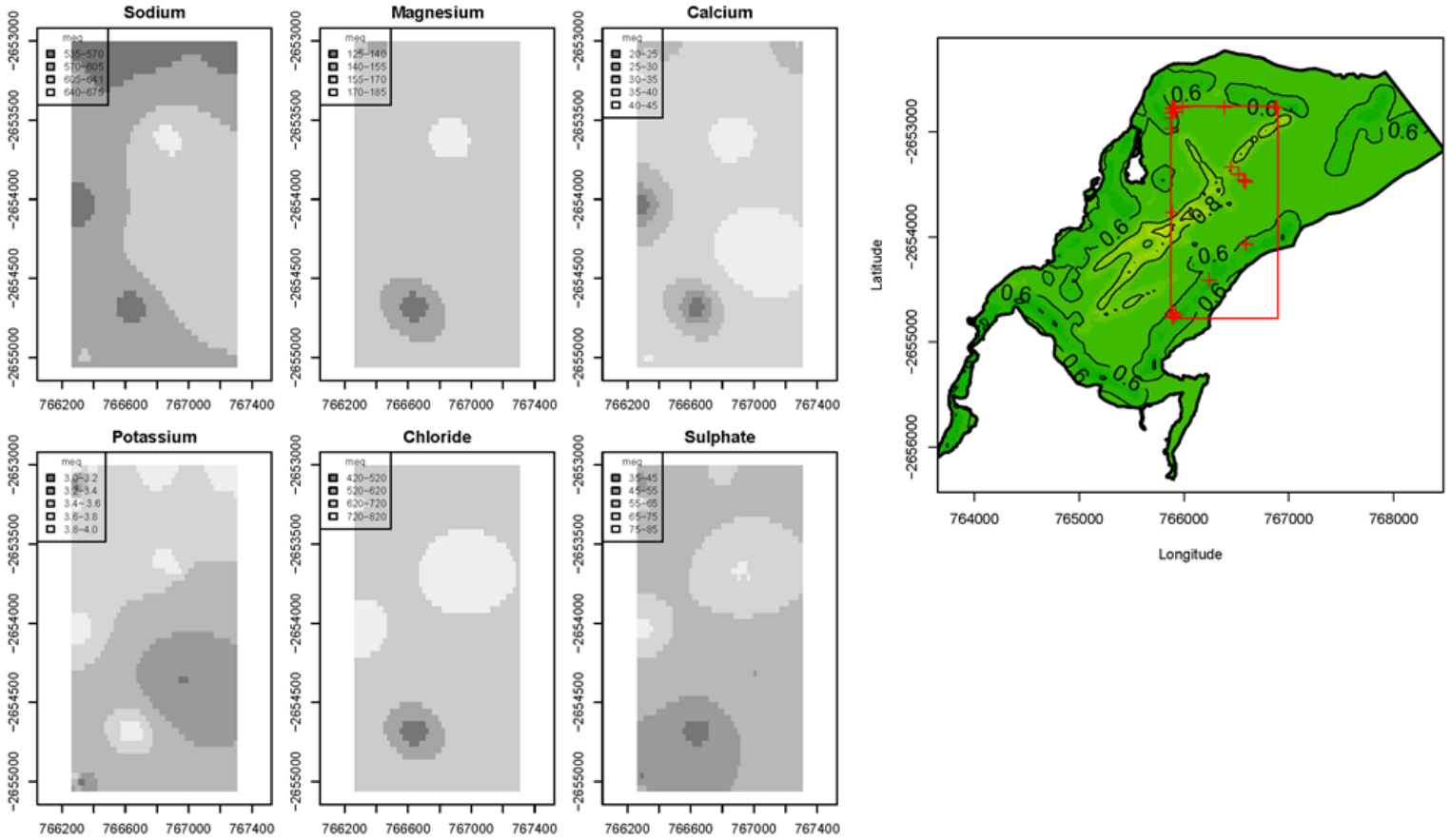


Figure 35: Bathymetric map showing samples locations (right) and inverse distance weighted map of ionic concentrations based on the concentration of each ion at each sample location (left) in Cygnet Pond. n = 84.

3.3.3 Nutrients

All the nutrient concentrations showed high variability. The mean ammonium concentration was 257 ± 211 µg/l, nitrate/nitrate concentration was 37 ± 23 µg/l and phosphate concentration was 23 ± 20 µg/l (Figure 36). There were significant differences in log transformed nutrient concentrations for ammonium ($F_{7,402}$=30.08, P<0.001), nitrite/nitrate ($F_{7,402}$=91.58, P<0.001) and phosphate ($F_{7,402}$=35.25, P<0.001) between ponds (Figure 37). Cygnet and Harjie's Ponds consistently had the highest and lowest concentrations of ammonium, nitrite/nitrate and phosphate, respectively.

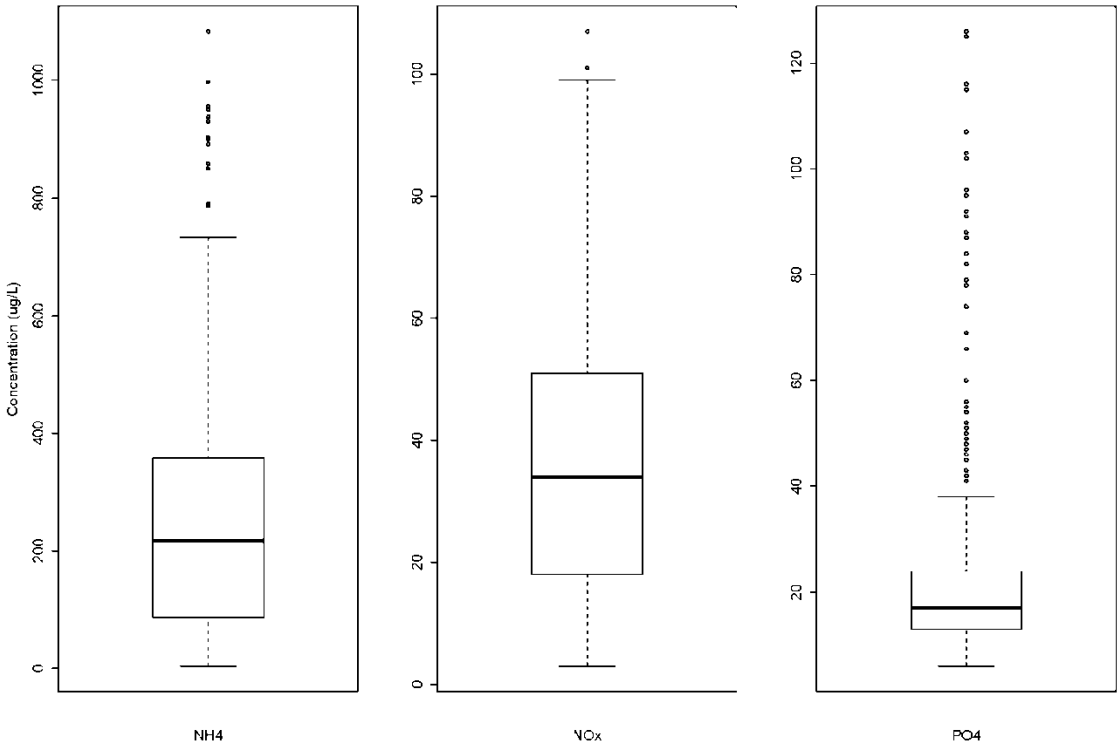


Figure 36: Boxplots showing concentrations of ammonium (NH_4), nitrite/nitrate (NO_x) and phosphate (PO_4) at the sediment-water interface. Dots represent outliers.

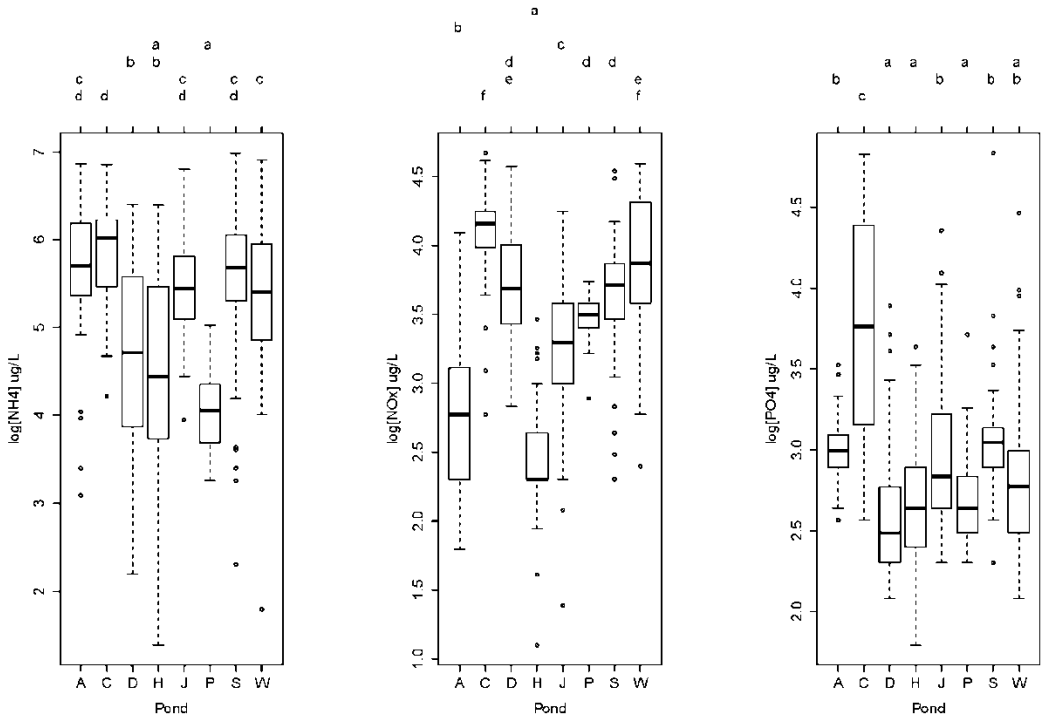


Figure 37: Boxplots showing significant differences between ponds and ammonium (left), nitrite/nitrate (middle) and phosphate (right) concentrations. Columns sharing letters at top are not significantly different to each other based on a Tukey's HSD test ($P<0.05$).

There were slight differences between ponds nutrient profiles (ANOSIM R=0.19, P=0.001). PCA of the nutrient data explained 86% of the variation along two axes (Figure 38 a). The first axis represents a gradient from low to high ammonium concentration, and the second axis represents a gradient from low to high nitrite/nitrate concentration. The pond groups were separated along both these axes (Figure 38 b). Donut Pond, Pete's Pond and Harjie's Pond were found to the left of the axis, whilst Pete's Vent, Annie's Pond, Whistler's Pond and Jana's Vent were found in the centre. Cygnet Pond, found on the right side of the axis, which means that it was different to the other ponds because of its high ammonium concentration. Harjie's and Annie's Ponds were found at the bottom of the second axis, which corresponds to low nitrate/nitrite concentration. Pete's Vent and Jana's Vent were found towards the centre of the axis whilst Donut Pond, Pete's Pond, Whistlers Pond and Cygnet Pond were found to have the highest levels of nitrite/nitrates.

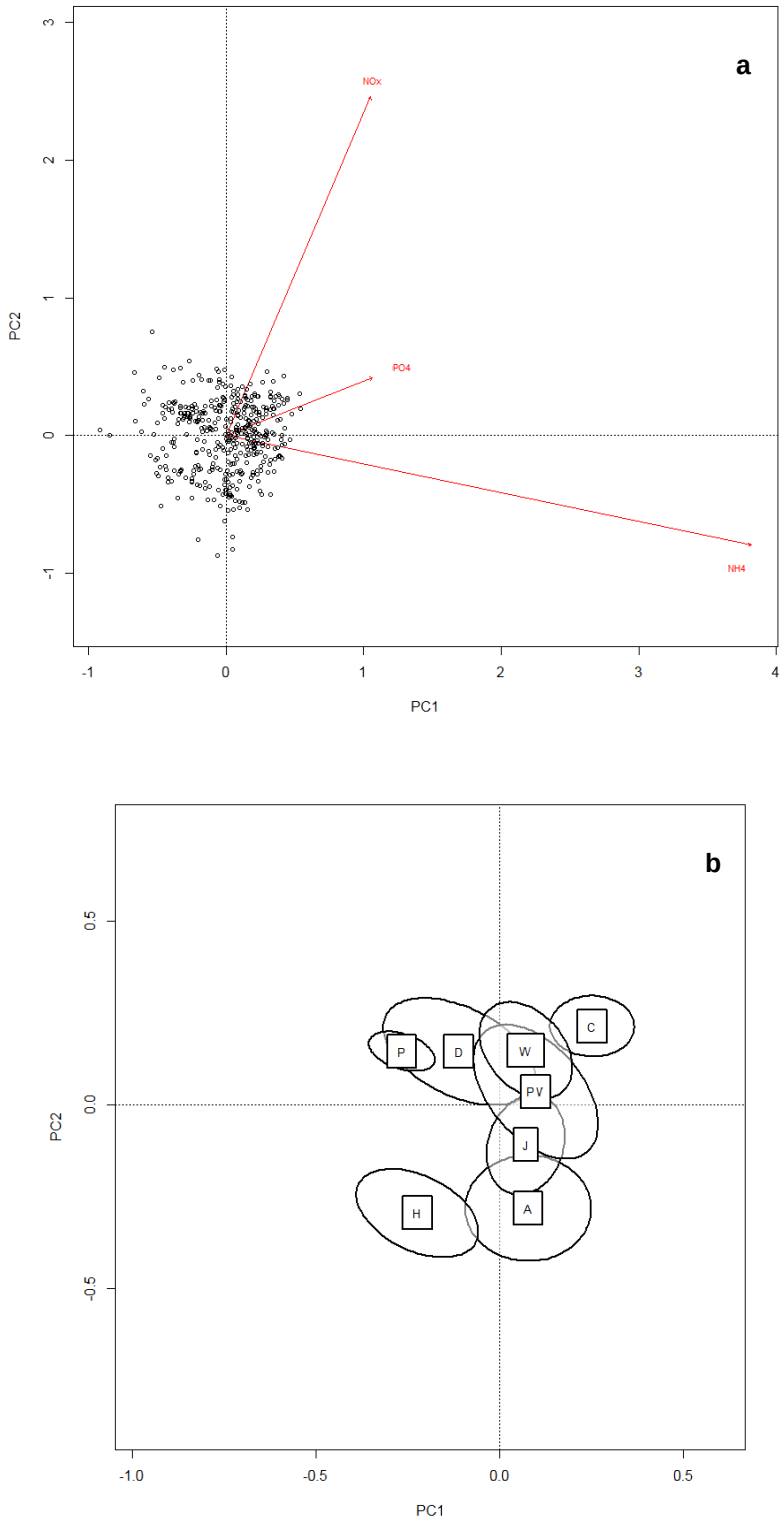


Figure 38: a) PCA biplot showing differences between the 410 samples and the log transformed nutrient variables; ammonium (NH_4), nitrite/nitrate (NO_x) and phosphate (PO_4) concentrations. The first and second PCA axes explain 61 and 25% of the variation, respectively. b) PCA plot showing the pond grouping, represented by 0.99 confidence level ellipses. Letters within squares represent the pond in which the samples originated; PV: Pete's Vent, D: Donut Pond, A: Annie's Pond, P: Pete's Pond, W: Whistler's Pond, H: Harjie's Pond, J: Jana's Vent; C: Cygnet Pond.

3.3.4 Sediment-water relationship

Linear discriminant analysis (LDA) of the salinity and nutrient data was able to discriminate between two groups of sediment types based on t he overlying water chemistry. The first linear discriminant explained 88.5% of the variation (Figure 39 a). The first group contained Pelletal-Aragonite Mud and Aragonite Mud, in addition to a few samples consisting of Pelletal Sand & Aragonite Mud and Pelletal Sand. The second, and much larger group, consisted of Pelletal Skeletal Sand and Pelletal Diatomaceous Sand, as well as most of the Pelletal Sand & Aragonite Mud and Pelletal Sand samples. All samples belonging to the small group were found in Cygnet Pond, while the other ponds were clustered within the larger group (Figure 39 b). Magnesium and potassium levels were the strongest determinants separating the sediment groups (Table 11 and Table 12).

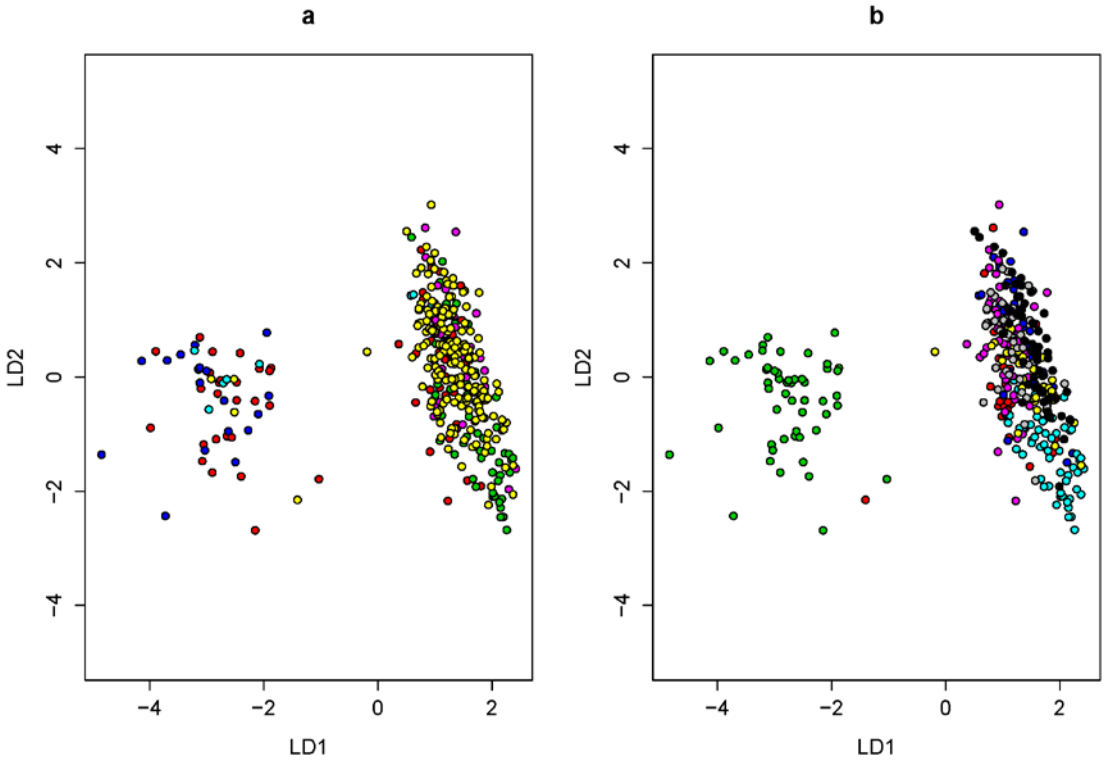


Figure 39: a) Linear discriminant analysis separating the different sediment types based on the overlying water chemistry (Yellow: Pelletal Sand; Red: Pelletal Sand & Aragonite Mud; Green: Pelletal Skeletal Sand; Purple: Pelletal Diatomaceous Sand; Blue: Pelletal-Aragonite Mud; Light Blue: Aragonite Mud); b) the same LDA showing the ponds which the samples belong to (Red: Annie's Pond; Green: Cygnet Pond; Blue: Donut Pond; Light Blue: Harjie's Pond; Purple: Jana's Vent; Yellow: Pete's Pond; Grey: Pete's Vent area; Black: Whistler's Pond).

Table 11: Coefficients of linear discriminants and the proportion of explained variation between groups for each of the 5 linear discriminants.

	LD1	LD2	LD3	LD4	LD5
Na	0.07	-0.81	7.40	-9.02	1.10
Mg	2.88	1.93	5.50	11.37	-0.95
Ca	-0.71	-0.74	-9.63	-6.38	0.9
K	-2.87	<-0.01	-1.70	1.60	0.16
Cl	0.63	-4.18	-1.44	2.47	-3.49
SO_4	-0.36	2.80	0.19	-1.04	1.78
NH4	0.19	-0.01	-0.14	0.48	0.78
NO_x	-0.12	1.16	-0.48	0.47	-0.72
PO_4	0.29	-0.15	0.31	0.06	0.40
Proportion of trace	88.5	6.6	2.8	2.1	<0.1

Table 12: Percentage composition coefficients of the major ions for each sediment type determined using LDA.

	Na	Mg	Ca	K	Cl	SO_4	NH_4	NOx	PO_4
Pelletal Sand & Aragonite Mud	31.2	3.8	1.3	0.7	56.0	7.0	<0.1	<0.1	<0.1
Pelletal Skeletal Sand	31.5	3.4	1.2	1.1	55.6	7.2	<0.1	<0.1	<0.1
Pelletal-Aragonite Mud	31.2	4.5	1.7	0.3	55.9	6.4	<0.1	<0.1	<0.1
Aragonite Mud	31.8	4.0	1.5	0.6	55.0	7.1	<0.1	<0.1	<0.1
Pelletal Diatomaceous Sand	31.4	3.4	1.2	1.1	55.4	7.5	<0.1	<0.1	<0.1
Pelletal Sand	32.5	3.6	1.2	1.1	54.0	7.7	<0.1	<0.1	<0.1

3.3.5 Relationship with pond morphotypes

No relationships between pond sedimentology and pond morphotypes (from Chapter 2) could be found. Pond morphotypes 1, 2 and 3 had similar ionic characteristics (Table 13). Morphotype group 2, which consists only of Cygnet Pond, had higher concentrations of sodium, magnesium and calcium, but lower concentrations of potassium. Nutrient characteristics were, however, different for the morphotypes. Morphotype groups 1 and 2 had similar concentrations of NH_4 (259 ± 17 and 301 ± 21 µg/l, respectively), while group 2 had the highest concentration (395 ± 26 µg/l) and group 4 had the lowest (119 ± 13 µg/l; Table 14). Group 1 also had the lowest concentration of NO_x (20 ± 1 µg/l), while the other groups had mean concentrations that varied between 39 and 62 µg/l). Morphotype 2 had the greatest levels of PO_4 (52 ± 4 µg/l).

Table 13: Range of major ionic concentrations in the four pond morphotypes. All units are millequivalents. Pond abbreviations are A: Annie's Pond, C: Cygnet Pond, D: Donut Pond, H: Harjie's Pond, J: Jana's Vent, PP: Pete's Pond, PV: Pete's Vent, W: Whistler's Pond.

Morphotype	Ponds	Na	Mg	Ca	K	Cl	SO_4
1	A, H, J	274-588	55-133	11-26	5-12	304-967	22-230
2	C	457-674	125-183	18-45	2-6	423-879	15-85
3	PV, W	267-546	52-116	11-25	5-11	267-883	24-93
4	PP, D	273-578	56-121	12-25	5-11	255-689	27-128

Table 14: Summary of nutrient data for each pond morphotype. Data are presented as means ± SE. All units are in µg/l. Pond abbreviations are A: Annie's Pond, C: Cygnet Pond, D: Donut Pond, H: Harjie's Pond, J: Jana's Vent, PP: Pete's Pond, PV: Pete's Vent, W: Whistler's Pond.

Morphotype	Ponds	NH_4	NO_x	PO_4
1	A, H, J	259 ± 17	20 ± 1	20 ± 1
2	C	395 ± 26	63 ± 2	52 ± 4
3	PV, W	301 ± 21	49 ± 2	21 ± 1
4	PP, D	119 ± 13	39 ± 2	15 ± 1

3.4 Discussion

There were six different sediments identified in this study, which were mostly of biogenic origin, and were mainly different types of skeletal and pelletal sands, with organic material and aragonite being minor constituents. Cygnet Pond had different sediments than the other ponds, and was characterised by sediments high in aragonite mud and low in pelletal sands. Water chemistry was also found to be different in the ponds, with ionic concentrations showing Cygnet Pond being very distinct from the other ponds. The ponds also had different nutrient profiles, which showed good agreement with the pond morphotype groups described in the previous chapter.

3.4.1 Relationship with morphometry

The morphometric properties of a water body can have profound effects on its biology. This suggests that there can be distinct limnological differences between water bodies that are difficult to explain using other variables, such as water chemistry. Indeed, morphometric variables, such as mean water depth and residence time, have been shown to be better predictors of biological community structure, including cyanobacterial and phytoplankton communities , as well as trophic status (Olding *et al.* 2000), than water chemistry variables themselves.

The highest nutrient levels were found in pond morphotype 2, which consists only of Cygnet Pond. This pond is shallow, with a high residence time and a high surface area-to-volume ratio. The ponds with high surface area-to-volume ratios are likely to be more

susceptible wind induced water column mixing and sediment resuspension, which is clearly the case for Cygnet Pond because of its large potential fetch. In non-stratified lakes, and particularly those which are shallow, nutrient release from sediments is likely to be greater than inputs from terrestrial sources because of sediment resuspension (Jensen and Anderson 1992, Jensen *et al.* 1994). Furthermore, the high residence time for this pond allows for the nutrients to be locally used, hence making for a more productive system (Tang and Xie 2000, Fisher *et al.* 2009). Morphotype 3, which included the lakes with the greatest discharge rate, highest mean depths and lowest surface area-to-volume ratios, on the other hand had intermediate levels of nutrients.

Unexpectedly, there was no clear link between morphotypes and ionic concentration. Odour and Schagerl (2007) found that lakes with low surface area-to-volume ratios tended to have stable ionic concentrations over time, whilst lakes with high surface area-to-volume ratios showed fluctuations in ionic concentration due to changes in evaporation. It is likely, however, that there are pond specific processes occurring, such as localised precipitation/dissolution, making it difficult to account for using morphometric variables. Similarly, no link between pond morphometry and sedimentology could be made suggesting that ionic concentrations and sedimentology are independent of pond morphometry.

3.4.2 Sedimentology

The sediments of Lake MacLeod were mostly dominated by pelletal fragments and, as a result, the most abundant sediment types were those composed of Pelletal Sands. There is a clear distinction between sediments dominated by aragonite mud and pelletal fragments: however, at the pond level, there were only two real groupings of sediments, those being high in aragonite mud and found in Cygnet Pond, and those high in pelletal sand. The latter sediment group dominated all the other ponds. Pelletal dominated sediments were also the most abundant sediment types found by Shepherd (1990) in the pond sub-environment. Pelletal sands are formed by the post-mortem accumulation of skeletal and faecal fragments (Shepherd 1990), and it is this biogenic feature of the sediments that make the sedimentology of the northern ponds of Lake MacLeod unique.

Shepherd (1990) also found pelletal-skeletal sands to be the dominant sediments in all the ponds, with the exception of the southern reaches of Cygnet Pond. In Cygnet Pond,

where the samples were taken, Shepherd (1990) describes 'continuous sheets of mud that are up to 30 cm thick". These are the sediments collected in this study which make the Cygnet Pond samples distinct from the other ponds and may be produced by the precipitation of carbonates (Shepherd 1990). Further, the large expanses of *Ruppia* sp. in Cygnet Pond may also make the Cygnet Pond sediments unique as seagrasses have been shown to affect sedimentation rates within meadows (Garcia *et al.* 2003). It is likely that such differences may be important when trying to understand the distribution of the microbial communities (Dupraz *et al.* 2004).

3.4.3 Ionic composition

The ponds at Lake MacLeod are dominated by sodium and chloride ions, similar to most other saline systems in Australia (Radke *et al.* 2002, Timms 2009). However, despite the strong sodium and chloride signatures typically found in Australian saline lakes, there still remains strong hydrochemical diversity among the lakes in terms of relative ionic concentrations (Radke *et al.* 2002), a pattern that was also found between and within the ponds studied here. The ionic maps generated in this study shown gradients in ionic concentration occurring within the ponds, and at small scales (<10m). The relative concentrations of major ions in salt lakes are believed to induce physiological responses in the aquatic organisms inhabiting them, thus causing organisms to have hydrochemical preferences (Radke *et al.* 2003).

Akin to the sedimentology differences observed between the ponds, Cygnet Pond was again different to all the other ponds in ionic concentrations. This difference was driven by the higher sodium, magnesium, calcium and chloride concentrations and the low potassium concentration of Cygnet Pond. Within the ponds, gradients in ionic concentrations were detected over small scales. Small scale environmental gradients in salinity have shown to be important determinants of microbial community structure (Dupraz *et al.* 2004, Langenheder and Ragnarsson 2007). The strength of these gradients would undoubtedly be much larger if the sampling areas included the entire pond areas, including the vent and outflow regions. However, because the plot areas could not cover the entire pond areas, particularly in Harjie's Pond and Jana's Vent, these results do not portray the entire ionic gradients which occur within the ponds.

In Annie's, Donut and Pete's Ponds, ionic concentrations typically decreased with distance from the source. A similar pattern was seen for magnesium in Pete's Vent over a small distance. This could be an indication of precipitates forming and dropping out of solution. Similar observations have been made by Shepherd (1990) at Lake MacLeod. There were areas of high magnesium and calcium concentrations found in Cygnet Pond which also coincided with high sodium, chloride and sulphate concentration. This indicates that dissolution of some precipitates, such as gypsum, huntite and dolomite may be occurring. Also of interest is the area of low ionic concentration in the southern area of Cygnet Pond. This could be due to a previously unknown seepage point introducing water from the ocean feedstock.

It is difficult to predict what processes are driving these patterns based on this data set. The reason for the decrease in potassium concentration when the other ions increased in concentration in Cygnet Pond is unknown, but could either be due to potassium precipitation, or biological uptake, in particular, by the extensive *Ruppia* sp. meadows found there. Potassium depletion has been recorded in some Australian saline lakes, where it is thought potassium can be taken up by reactive surfaces, such as an interaction with clay minerals in the sediment column (Herczeg and Berry Lyons 1991, Radke *et al.* 2002). The aragonite mud dominated sediments found in Cygnet Pond, which were not well established in any of the other ponds, may therefore be important in supressing potassium levels in the Cygnet Pond water column.

Generally, the changes in ionic concentrations found in Annie's, Cygnet, Donut and Pete's Pond are what could be expected to occur under an evapo-transpiration driven system, where concentration of brine has either resulted in an increase in ionic species concentration, or passed a threshold and precipitation of various salts has occurred (Logan 1987). Strong effects of evaporation would be expected in these ponds as they all have high surface area-to-volume ratios (1.4-2.0 m^{-1}). Pete's Vent and Whistler's Pond have low surface area-to-volume ratio of 1.0 m^{-1}, and therefore are influenced to a lesser degree by evaporative forces. However, the general trend seen in all the ponds of increasing chloride concentration in a south-easterly direction for these ponds is probably a result of evaporative concentration of the brine.

3.4.4 Sediment-water relationship

The water chemistry overlying the sediments were related to the type of sediment found there and could be caused by the precipitation of various mineral phases out of solution and being deposited within the sediment column, or vice versa. It is likely both precipitation and dissolution are occurring throughout the ponds, but in different areas, and is largely controlled by localised evapotranspiration, or seasonal shifts in salinity (Shepherd 1990). However, the relationship between water chemistry and sedimentology was not specific enough to distinguish between all sediment types and water chemistry. Water chemistry could only discriminate two broad groupings of sediments, those dominated by pelletal sands, and those dominated by aragonite mud, with the aragonite muds being found in waters with low potassium concentrations. Furthermore, these two broad groups contained samples that were either found in Cygnet Pond, or found in the other ponds, and thus provides further evidence of the unique environmental conditions and habitats found in Cygnet Pond, as well as the possible removal of potassium ions from the water body by the aragonite mud rich sediments in Cygnet Pond (see Herczeg and Berry Lyons 1991).

The relationships between sediment structure and water chemistry have also been detailed by Kissoon *et al.* (2015). Sediment redox potential is likely to have been important in determining the chemical characteristics of the water, at least at the pond scale. It has been shown that decreased redox potentials in sediments leads to the release of PO_4 and NH_4 into overlying waters (Wong and Yang 1997). The variation in nutrients found in this study could be related to changes in the redox chemistry in the underlying sediments, which may have been enhanced in Cygnet Pond because of the stabilisation of sediments from the large expanses of macrophytes. Understanding sediment depth changes in redox potential and element concentrations will enhance our knowledge on p rocesses affecting the water composition within the ponds. Unfortunately this was not accomplished in this study because of the focus on detailing the habitat structure directly important for microbial biofilm communities at the sediment-water interface of these ponds.

3.4.5 Conclusion

The ponds of Lake MacLeod represent unique habitats, particularly Cygnet Pond when compared to the other ponds. Within the ponds, there is considerable variation in ionic concentrations, and to a lesser degree, nutrient concentrations, and a patchy distribution of sediment types. Not only does this chapter highlight the distinctness in habitat structure of Cygnet Pond when compared to the other ponds, this chapter also presents a case where changes in hydrochemistry could be associated with changes in sedimentology, and how the morphometric properties of a water body can be related to the nutrient levels of the waters. It is expected that the habitats presented and described in this chapter will contain unique microbial communities, if species-sorting processes are the main structuring process operating for these communities at Lake MacLeod.

CHAPTER 4. BACTERIAL ECOLOGY OF THE PONDS

4.1 Introduction

The importance of bacteria in aquatic ecosystems was first recognised by Lindeman (1942) when he put 'bacterial ooze' at the centre of the trophic dynamics of a temperate lake. Since then, limnologists have regarded aquatic bacterial communities as fundamental components of aquatic ecosystems because of the critical role they serve by regenerating and mobilising nutrients (Newton *et al.* 2011). In fact, bacteria are responsible for driving the transformations and the cycling of most biologically active elements in these ecosystems (Cotner and Biddanda 2002, Newton *et al.* 2011). Furthermore, in aquatic systems, bacteria often represent the main group of primary producers, and, because of the subsequent grazing upon them by higher trophic levels (such as heterotrophic bacteria and ciliates), they form the base of the complex aquatic food webs (Pernthaler and Amann 2005).

Understanding the mechanisms that regulate bacterial distributions and diversity still remains an important topic of study in ecology. Early microbiologists assumed that bacterial biogeography is fundamentally different to macrobionts because they are not dispersal limited (O'Malley 2008). Recently though, studies have begun to discover biogeographic patterns similar to what has been seen in macrobionts (Horner-Devine *et al.* 2004, Martiny *et al.* 2011, Lear *et al.* 2013, Lear *et al.* 2014). Distance-decay relationships have been found in bacterial communities (Astorga *et al.* 2012), at both large geographical scales of hundreds of kilometres (Cho and Tiedje 2000, Whitaker *et al.* 2003, Fuhrman *et al.* 2006, Schauer *et al.* 2010, Lear *et al.* 2013) and also at much finer, within habitat scales (Franklin and Mills 2003, Horner-Devine *et al.* 2004, Bell 2010, Östman *et al.* 2012, Logares *et al.* 2013, Lear *et al.* 2014).

While it is now well established that bacteria are capable of producing distance-decay relationships, there is little understanding of the ecological processes involved in generating them (Martiny *et al.* 2005, Astorga *et al.* 2012). Even the idea that bacteria are dispersal limited remains controversial (Finlay *et al.* 1998, De Bie *et al.* 2012). If, however, dispersal limitation is the main driver of bacterial biogeographic patterns, and the neutral theory applies, then geographic distance should be the best predictor of the

distance-decay relationship. On the other hand, if taxa are assembled according to environmental parameters (species sorting and niche theory) then measures of environmental heterogeneity will be the best predictors for the distance-decay relationship.

It is possible that both types of mechanisms are mutually responsible for bacterial assemblages. There is now evidence that in combination with neutral processes, deterministic mechanisms, like those predicted by species-sorting and niche theories, are important drivers of bacterial community structure (Caruso *et al.* 2011, Lee *et al.* 2013). Furthermore, the recent study of Lear *et al.* (2014) shows that populations of bacteria separated by less than 20m in a highly continuous lentic environment show similar biogeographic patterns to those separated at larger, less connected, landscape scales. Understanding the ecological processes that drive the distance-decay relationship and community structure of bacteria will provide evidence for which processes play important roles in regulating the vast bacterial diversity.

Aquatic systems have been used by microbial ecologists to understand bacterial biogeography with water bodies representing relatively homogenous habitats, which have been treated as 'aquatic islands', within a 'terrestrial sea'. A major shortcoming in many studies is the assumed even distribution of bacterial taxa throughout the water body because of the environmental homogeneity and high levels of mixing thought to occur within these habitats. As a result, many studies have relied on taking only one or two water samples as a representation of an entire lake or pond bacterial community (Yannarell and Triplett 2004, Reche *et al.* 2005, Yannarell and Triplett 2005, Pagaling *et al.* 2009, Romina Schiaffino *et al.* 2011). Recently though, small-scale variation in bacterial distributions have been detected within lakes (Jones *et al.* 2012, Garcia *et al.* 2013) and fine scale variation occurring at distances less than 20m was recently found in the bacterioplankton communities in the alpine tarns of New Zealand (Lear *et al.* 2014, Lee 2014). It is therefore clear that within lake variation in bacterial community composition needs to be accounted for when attempting to elucidate the role of biogeographic processes on bacterial richness.

There are many technical challenges in studying the vast richness of the bacterial world. While up to one million bacterial cells per millilitre of water are thought to be present in

salt lakes (Whitman *et al.* 1998), it has been technically difficult to measure the number of taxonomic groups present. Molecular approaches are a well-established tool used by microbiologists to define taxonomic groups. These techniques have concentrated on defining bacterial taxa based on variation of either the lengths of (fragment analysis) or the sequences of the ribosomal RNA (rRNA) gene. A common definition of bacterial species is '*a group of strains that have some degree of phenotypic consistency, exhibit at least 70% DNA-DNA hybridisation, and greater than 97% 16S rRNA sequence similarity*' (page 734; Gevers *et al.* 2005). Although sequencing methods, such as 454 pyrosequencing, are very powerful and can generate datasets at fine taxonomic resolutions (usually 97 or 99% similarity), fragment analysis techniques are much cheaper, and have comparable results to sequencing when comparing the composition of bacterial communities between samples (Jones *et al.* 2012, Shade *et al.* 2012, Lee 2014).

This study utilises the DNA fragment analysis, automated ribosomal intergenic spacer analysis (ARISA) to investigate the spatial patterns, and the role of species-sorting and neutral processes as structuring mechanisms on the benthic bacterial communities found in the ponds of Lake MacLeod. The eight study ponds make an ideal study to examine the biogeographic patterns of benthic bacteria. Given the small-scale variation in planktonic bacterial communities, it can be expected that benthic communities display similar biogeographic patterns. The benthic habitat of the Northern Ponds, although appearing rather homogenous, has been shown in the previous chapter to display some degree of heterogeneity in physical and chemical heterogeneity which can be expected to alter the bacterial communities. Using the data from the previous chapters on the changes in water chemistry and sediment characteristics within the ponds, and the morphometric differences between the ponds, the contribution of the environment and geographic location will be examined in order to observe the role of deterministic processes on community assembly. Furthermore, because it is likely that community structure is in fact a result of both deterministic and stochastic processes, the effect of species co-occurrence patterns and neutrality and will be investigated using null and neutral models, respectively. By testing for the roles of stochasticity (co-occurrence models) and dispersal limitation (neutral models) on community composition, the role

of non-deterministic processes on c ommunity assembly can be disentangled from deterministic processes.

This chapter will test the following hypotheses central to understanding the environmental and demographic processes that structure the benthic bacterial communities in the ponds of Lake MacLeod:

1. larger ponds will have higher richness of taxonomic groups;
2. the bacterial assemblages among ponds differ significantly;
3. environmental heterogeneity and spatial variables explain significant amounts of variation in bacterial community composition;
4. OTU co-occurrence patterns are not random; and
5. immigration of taxa between sample locations affects the composition and diversity of bacterial communities.

This chapter will provide information on t he fine scale structure of bacterial communities, as well as whether species-sorting or neutral processes are the important determinants of bacterial community composition in Lake MacLeod.

4.2 Methods

4.2.1 Sampling

Sediment samples were collected in conjunction with the water and sediment samples described in Chapter 3 (section 3.2.1) using the spatial design described in Chapter 1 (section 1.3.2). As previously described, samples were collected in unused 70 mL sterilised jars and the jars were opened underwater in proximity to where the sediment scoop was to be taken. The scoops were taken to include only the top layer of the sediment. In some of the deeper ponds, SCUBA was required for sampling, however, most samples were collected by snorkelling when the sediment was shallower than arms reach. Care was taken to not disturb the sediment prior to collection. Samples were kept at 4°C for 24 hours to allow suspended particles in the water to settle. After 24 hours, the liquid was decanted and the sediment samples were frozen at -20°C. The decanted liquid was used for water chemistry measurements and the sediment sample used for sediment characterisation and subsequent DNA analyses.

4.2.2 Community fingerprinting

Automated ribosomal intergenic spacer analysis (ARISA) was used to describe the structure of the bacterial communities. This method has been widely used (Fisher and Triplett 1999, Ranjard *et al.* 2001, Ramette 2009, Lear *et al.* 2011, Lear *et al.* 2014), and has been found to be better for estimating community diversity of complex bacterial communities than other fingerprinting methods such as terminal restriction length polymorphism (T-RFLP) or denaturing gradient gel electrophoresis (DGGE) with greater power for detecting less abundant taxonomic groups (Danovaro *et al.* 2006). The 16S-23S intergenic spacer (IGS) region of bacterial rRNA was used to build a community profile, based on the length of the amplified fragment. This fragment has been shown to display strong heterogeneity between bacterial species (Fisher and Triplett 1999, Ranjard *et al.* 2001).

DNA was extracted from the frozen sediment samples using the PowerLyzer™ PowerSoil® DNA Isolation Kit (MO BIO Laboratories Inc., Carlsbad) following the manufacturer's instructions. PCR was used to amplify the IGS region of the extracted DNA using the bacterial primers LDBact (5'-CCG GGT TTC CCC ATT CGG-3') and SDBact (5'-TGC GGC TGG ATC CCC TCC TT-3') as detailed by Ranjard *et al.* (2001). The SDBact was labelled at the 5' end with 6-carboxyfluorescein (FAM) flourophore. Following Lear *et al.* (2008), the PCR conditions were: (i) 95 °C for 5 min; (ii) 35 cycles of 95 °C for 30s, 61.5 °C for 30 s, 72 °C for 90 s and then (iii) 72 °C for 10 min. Bovine serum albumin was added (0.1mg/mL) to overcome inhibitors that were preventing amplification from working. 1 µL of PCR product was combined with 10 µL of Hi Di formamide and 0.6 µL GeneScan™ 1200 LIZ® dye Size Standard (Applied Biosystems Ltd, Melbourne, Australia), before being heat treated at 95 °C for 5 minutes and cooled on ice. The samples were run on a 3730XL DNA Analyser (applied Biosystems Ltd) using a 50 cm capillary. All DNA extractions, PCR amplifications and capillary separations were done at the Australian Genome Research Facility.

The resulting data were binned according to the method described by Ramette (2009). This method reduces the effect of background noise generated during the automated analysis and aims to identify 'true' peaks. To compensate for uncertainty in size-calling between different samples, fragments were assigned to bins with the width of the bin

differing according to the length of the fragment (Brown *et al.* 2005). To include the maximum number of peaks, yet exclude background fluorescence, only peaks with a fluorescence value 150 U or greater, and between 130 and 1000 bp were analysed. Using the R-scripts provided by Ramette (2009), fragments up to 600 bp were assigned to bins of 3 bp (± 1 bp), bins of 5 bp for fragments between 600 and 700 bp (± 2 bp) and 7 bp for any fragments greater than 700 bp (± 3 bp). The total area under the curve for each sample was standardised to 100 to reduce the effect of differences in profiles due to different initial DNA template, and all peaks with a relative intensity less than 1% of the sample total were removed.

4.2.3 Richness and diversity

The relative fluorescence intensity (RFI) was used as a proxy for relative abundance for each OTU. Richness of a sample was defined as the number of different OTUs (different fragment lengths), and diversity for a sample was characterised using Shannon and Simpson indices calculated using the R-package *vegan* (Oksanen *et al.* 2013). Species accumulation models were calculated using Kindt's exact method (Ugland *et al.* 2003). Because species accumulation models usually indicate that not all OTUs will have been sampled, the total OTU richness for each pond was calculated using the *specpool* function in *vegan* (Oksanen *et al.* 2013). A measure of the taxa-area relationship occurring within the ponds was calculated using an adapted Arrhenius species-area model:

$$z = \frac{(\log(2) - \log(2a + b + c) + \log(a + b + c))}{\log(2)}$$

where a is the number of shared OTUs in two sites, and b and c are the numbers of OTUs unique to each site (Harte and Kinzig 1997, Koleff *et al.* 2003).

Hellinger transformations of the community matrix were used as it allows dissimilarity distances to be calculated in Euclidean space, thus allowing the use of linear methods of analysis, such as PCA and redundancy analysis (RDA; Legendre and Gallagher 2001, Borcard *et al.* 2011). Differences between ponds were tested for significance using permutational multivariate analysis of variance (PERMANOVA). Tukey's post hoc tests were done to test which ponds were different from one another. The significance of

the differences in structure of bacterial communities between ponds and pond morphotypes were also compared using an analysis of similarities (ANOSIM; Clarke 1993) and graphically represented via non-metric multidimensional scaling (nMDS) plots using Hellinger transformed bacterial community data and Euclidean distances.

Mantel correlograms were used to do pa irwise comparisons of bacterial community structure of each sample at different distances in the R package *vegan* (Oksanen *et al.* 2013). Using Euclidean distances derived from the Hellinger transformed community data, bacterial community resemblance can be compared between samples belonging to the same distance class. The Mantel statistic is then tested through a permutational Mantel test. When interpreting a Mantel correlogram, the sign of the correlation coefficient is important, and one must observe the pattern drawn by the significant correlation values (Borcard and Legendre 2012). A positive (and significant) mantel coefficient indicates that for a given distance class, the multivariate similarity among the samples is higher than expected with chance, and the reverse is true for significant negative coefficients. The expected value when there is no correlation is 0 (Borcard and Legendre 2012).

4.2.4 Effect of environmental and spatial determinants

RDA was used to constrain the community data to the environmental data at each sampling site. The environmental data included water chemistry (mean values for each pond are given in (Table 15), sedimentology (Table 16) and pond m orphotypes. A forward selection process was used to select the best environmental determinants of bacterial community composition using the *packfor* package in R (Dray *et al.* 2007). Variation partitioning, using the environmental data, pond identity and spatial variables, was done using partial RDAs. Moran's eigenvector maps (MEM) were created to represent the spatial variation across a range of spatial scales amongst the study sites.

Table 15: Summary of water chemistry variables (mean ± SE) used in environmental modelling for this chapter. All units are mg/l.

	Na	Mg	Ca	K	NH_4	NO_x	PO_4	Cl	SO_4
Pete's Vent	9770 (170)	1047 (19)	372 (7)	326 (6)	0.34 (0.04)	0.042 (0.003)	0.025 (0.003)	19954 (941)	2706 (151)
Donut Pond	10940 (210)	1206 (26)	396 (10)	360 (8)	0.18 (0.02)	0.045 (0.003)	0.016 (0.001)	17604 (340)	2394 (103)
Annie's Pond	10414 (170)	1123 (20)	395 (10)	351 (8)	0.38 (0.04)	0.018 (0.002)	0.021 (0.001)	21675 (952)	3420 (166)
Pete's Pond	9548 (160)	1021 (19)	339 (7)	285 (5)	0.06 (0.01)	0.033 (0.001)	0.015 (0.001)	14256 (447)	2346 (64)
Whistler's Pond	9197 (204)	1018 (24)	374 (8)	318 (7)	0.27 (0.02)	0.056 (0.003)	0.019 (0.001)	15091 (307)	1862 (45)
Harjie's Pond	9207 (207)	1008 (25)	359 (9)	322 (8)	0.15 (0.02)	0.012 (0.001)	0.016 (0.001)	16321 (399)	2116 (73)
Jana's Vent	11379 (154)	1321 (19)	430 (7)	394 (5)	0.29 (0.02)	0.028 (0.002)	0.023 (0.002)	18184 (329)	2835 (260)
Cygnet Pond	13428 (158)	1942 (21)	736 (17)	142 (3)	0.41 (0.03)	0.063 (0.002)	0.053 (0.005)	24231 (402)	2855 (101)

Table 16: Percentage contribution of each sediment type defined in Chapter 3 for each pond used in environmental modelling.

	Pelletal Sand & Aragonite Mud	Pelletal Skeletal Sand	Pelletal – Aragonite Mud	Aragonite Mud	Pelletal Diatomaceous Sand	Pelletal Sand
Pete's Vent	55	0	0	0	5	40
Donut Pond	5	31	0	10	17	38
Annie's Pond	0	5	0	0	23	73
Pete's Pond	5	0	0	0	0	93
Whistler's Pond	0	5	0	0	25	70
Harjie's Pond	2	64	0	0	7	27
Jana's Vent	35	2	0	0	0	63
Cygnet Pond	50	0	37	10	0	4

4.2.5 Co-occurrence patterns and neutral processes

Null model analysis of OTU co-occurrence was performed using EcoSim700 software (Gotelli and Entsminger 2004). EcoSim measures patterns in community matrices using co-occurrence indices, and creates a sample of random matrices and statistically compares the co-occurrence index in the observed and simulated data sets. The C-score index (Stone and Roberts 1990) was used to measure co-occurrence in this study. This method quantifies the average amount of co-occurrence among all unique pairs of OTUs in the assemblage. In a competively structured community, the observed C-scores should be larger than what is expected by chance (Gotelli and Entsminger 2004). For each pond community, 10,000 randomisations of presence/absence transformed community matrices were used to create null expectations for the C-score. The random matrices retained the same number of OTUs and samples as the original matrix. If the observed C-score was higher than that of 95% of the simulated datasets, the observed community was considered to display significant segregation of OTUs ($\text{C-score}_{\text{observed}} > \text{C-score}_{\text{simulated}}$). It was deemed to show significant aggregation if $\text{C-score}_{\text{observed}} < \text{C-score}_{\text{simulated}}$ for 95% of the simulated datasets. If there were no differences in C-scores ($\text{C-score}_{\text{observed}} = \text{C-score}_{\text{simulated}}$), then the OTUs are considered to be randomly assembled (no evidence of segregation or aggregation of OTUs).

The neutral diversity (θ) and immigration (I) parameters were estimated using a maximum likelihood approach and the sampling formula for multiple samples developed by Etienne (2009), which is an improved version of Etienne (2007). Simulated communities that correspond to the estimates of θ and I can then be predicted. In order to calculate θ and I, and simulate the neutral communities, the PARI/GP codes given by Etienne (2007, 2009) were used. Based on those parameters, 4999 simulated communities were predicted by the neutral model from the metacommunity that corresponds to the estimates of θ and I. The immigration parameter, m, as defined by Hubbell (2001), was not used as a simulation parameter. Instead the number of immigrants, I, was used to express the number of individuals that are immigrants to the local community. The relationship between m and I is:

$$m = \frac{I}{I + J - 1}$$

where *J* is the number of individuals in the community (Etienne 2007, Maaß *et al.* 2014).

The bacterial community matrix for each sample was used to compare observed β-diversity to what would be expected following a general neutral model (Hubbell 2001). Beta diversity of the real communities, and the 4999 simulated neutral data sets was calculated following Legendre and De Cáceres (2013) using the R script developed by Maaß *et al.* (2014). If the observed beta diversity was higher or lower than 95% of the simulated datasets, the observed community was then considered to have a different beta diversity than would be expected under neutral assembly processes (Maaß *et al.* 2014).

4.3 Results

4.3.1 Richness and diversity

Bacterial community fingerprinting was successfully completed for 463 samples across the eight ponds. In total, 171 OTUs were detected, although most OTUs were only represented in 0-10% of the samples (Figure 40). Very few OTUs were found in more than 80% of the samples. The least number of OTUs were found in Whistler's Pond (56 OTUs) and the greatest in Cygnet Pond (127 OTUs). Richness was estimated using the species accumulation plots (Figure 41) and a boot-strapping method, and show good agreement with the number of OTUs detected using ARISA (Table 17). Larger ponds did not have more OTUs ($F_{1,6}$=0.813, P=0.402; Figure 42), although the largest pond was the richest; Whistler's Pond and Harjie's Pond had low richness values given the size of the ponds. Similarly, there was no trend of decreasing species-area exponent (*z*) with pond size. Diversity was greatest in Cygnet Pond, based on Shannon and Simpson Diversity Indices (Table 17). Similarly, Cygnet Pond, as well as Pete's Vent and Donut Ponds, had high levels of evenness, based again on Shannon and Simpson Evenness indices. Whistler's Pond showed the lowest levels of diversity and evenness.

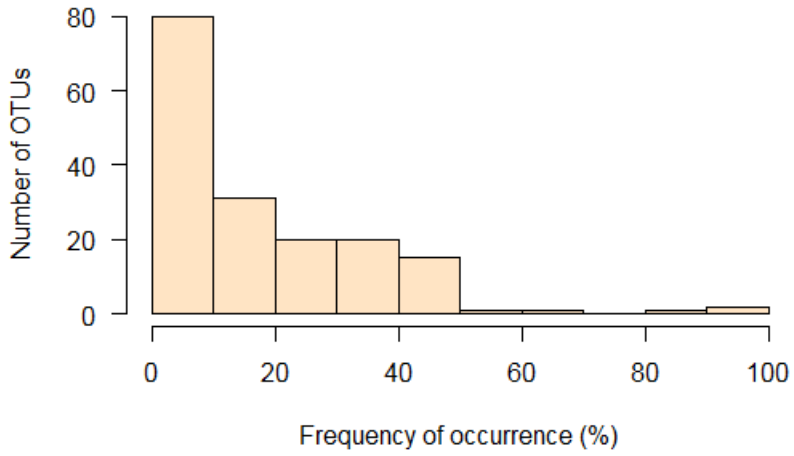


Figure 40: Frequency histogram of the frequency of occurrence of the 171 OTUs from the eight sampled ponds.

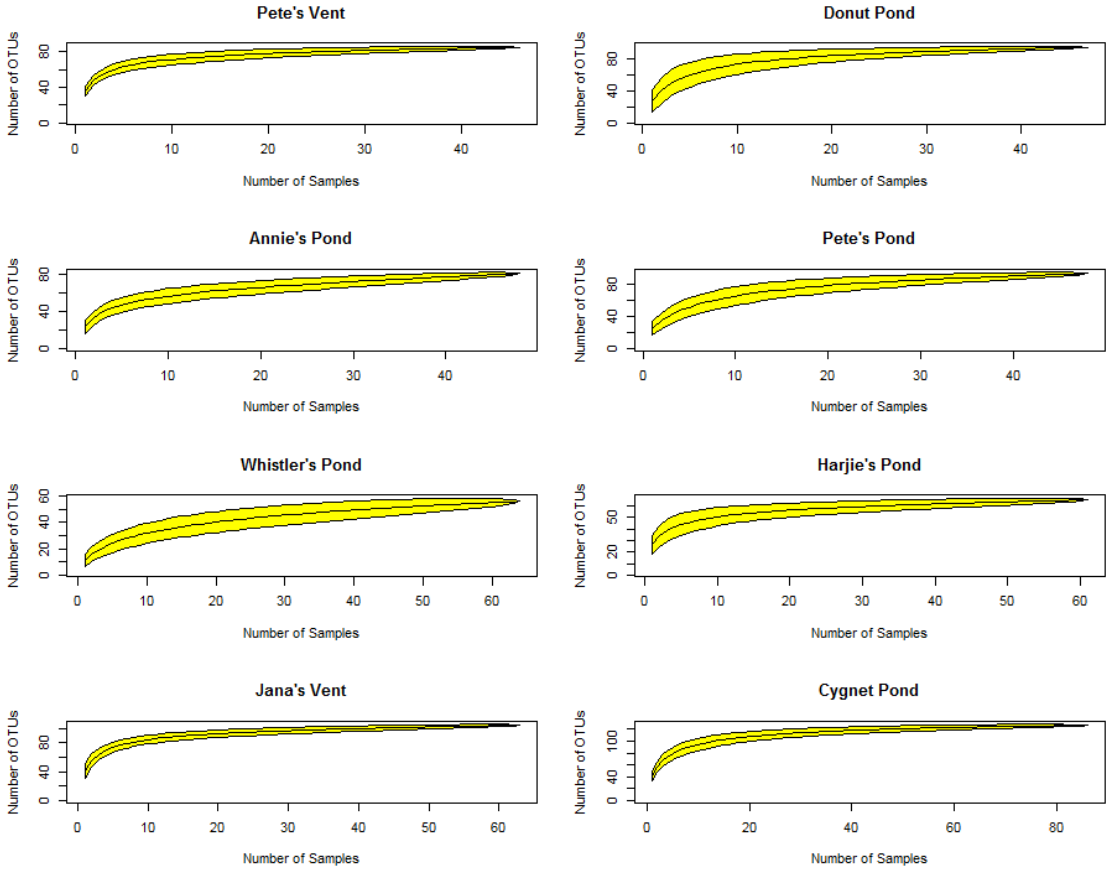


Figure 41: Species accumulation plots for each pond with yellow shaded area representing standard deviation.

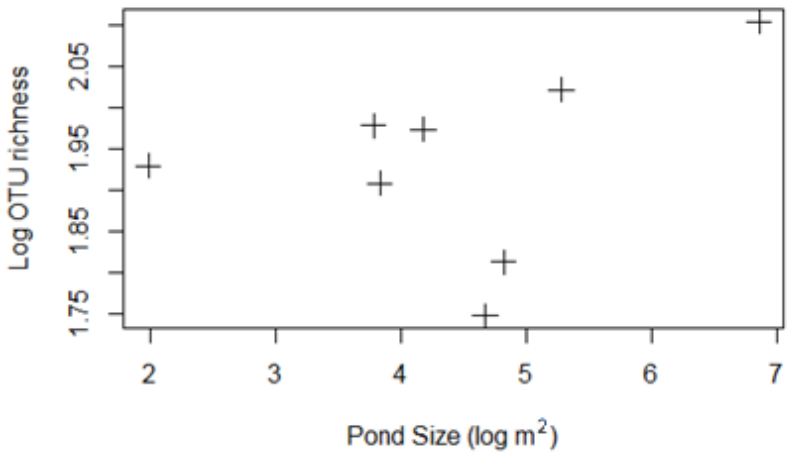


Figure 42: Scatter plot of OTU richness against log Pond Size (m^2). There was no significant relationship between OTU richness and pond size ($F_{1,6}$=0.813, P=0.402).

Table 17: Diversity indices for each pond. Total OTU richness and estimated richness represent the total for the entire pond, whilst the diversity and evenness values are mean for all samples within in pond.

	Total OTU Richness	Shannon Diversity	Shannon Evenness	Number of Samples	Estimated Richness	*z*
Pete's Vent	85	37.5	0.44	46	89 ± 2	0.31
Donut Pond	95	31.1	0.33	47	100 ± 3	0.39
Annie's Pond	81	19.0	0.23	48	89 ± 3	0.34
Pete's Pond	94	16.6	0.18	48	102 ± 3	0.33
Whistler's Pond	56	6.9	0.12	64	62 ± 3	0.34
Harjie's Pond	65	20.1	0.31	61	69 ± 2	0.26
Jana's Vent	105	43.9	0.42	63	110 ± 2	0.31
Cygnet Pond	127	48.3	0.38	86	132 ± 2	0.33
Entire Data Set	171	44.1	0.26	463	175 ± 2	

The composition of the bacterial communities varied among the ponds (Global PERMANOVA Pseudo $F_{7,462}$=74.65, P=0.001; Figure 43). Pairwise comparisons indicate that Cygnet Pond and Jana's Vent have similar bacterial communities, and so do Annie's Pond, Pete's Pond, Whistler's Pond and Harjie's Pond (pairwise PERMANOVAs, all P_{adj}<0.05; Table 18). Interestingly, Pete's Vent was different to Pete's Pond, into which it flow. Overall there are difference in community composition for the different pond morphotypes that could be attributed to morphotype 1 (which consists of Annie's Pond, Harjie's Pond and Jana's Vent) being different from morphotypes 2, 3 and 4 (Global PERMANOVA Pseudo$F_{3,383}$=36.46, P=0.001; Figure 44). ANOSIM results show that within group similarity was better explained using the pond identities (R=0.839), rather than morphotypes (R=0.356).

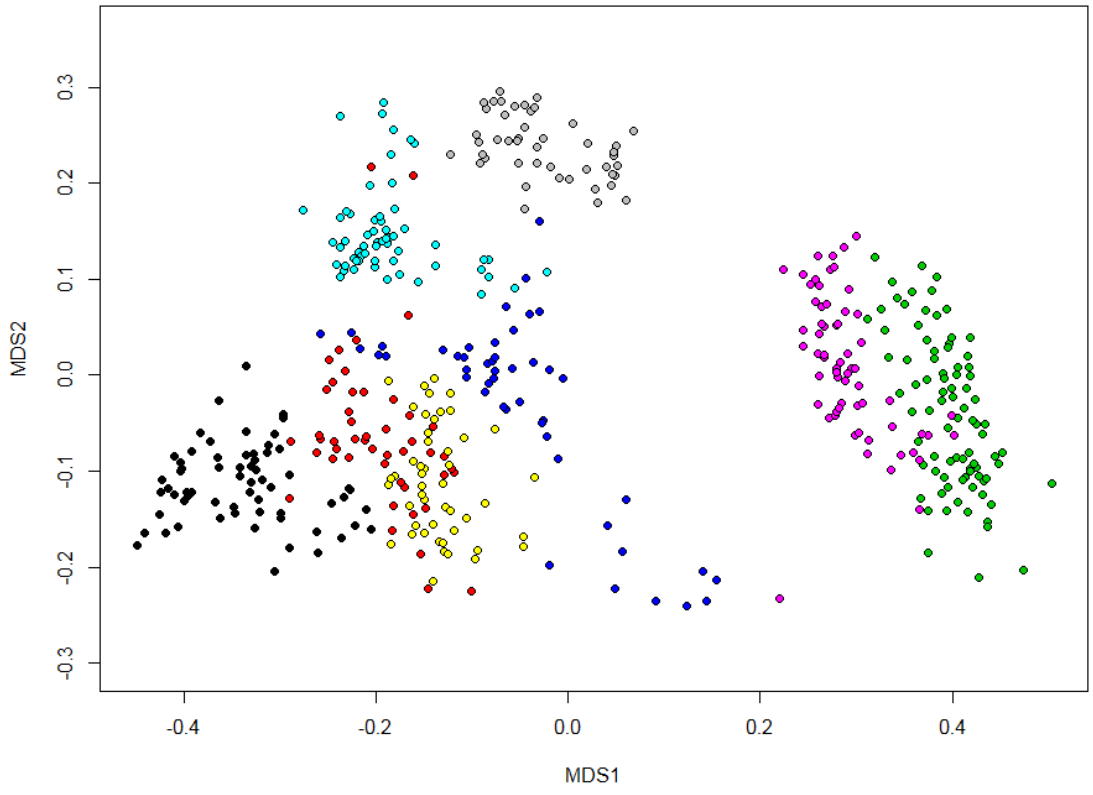


Figure 43: MDS ordination showing where each sample is located in relation to other samples in species space. Samples closer to each other are more similar than those separated by larger distances on the plot. A Hellinger transformed community matrix was used, and MDS was conducted using Euclidean distances. Grey: Pete's Vent; Dark Blue: Donut Pond; Red: Annie's Pond; Yellow: Pete's Pond; Black: Whistler's Pond; Light Blue: Harjie's Pond; Purple: Jana's Vent; Green: Cygnet Pond. Stress=0.15.

Table 18: Distance matrix showing the distances between the centroids of each pond group in principal coordinate space (PCoA) using Hellinger transformed bacterial community data. Bold values indicate significant differences in the bacterial community composition of the ponds (P<0.05).

	Pete's Vent	Donut Pond	Annie's Pond	Pete's Pond	Whistler's Pond	Harjie's Pond	Jana's Vent
Donut Pond	0.044						
Annie's Pond	0.035	**0.079**					
Pete's Pond	**0.103**	**0.147**	0.068				
Whistler's Pond	**0.125**	**0.169**	**0.090**	0.022			
Harjie's Pond	**0.069**	**0.114**	0.035	0.033	**0.056**		
Jana's Vent	0.008	0.052	0.027	**0.095**	**0.118**	**0.062**	
Cygnet Pond	0.037	0.007	**0.072**	**0.140**	**0.162**	**0.107**	0.045

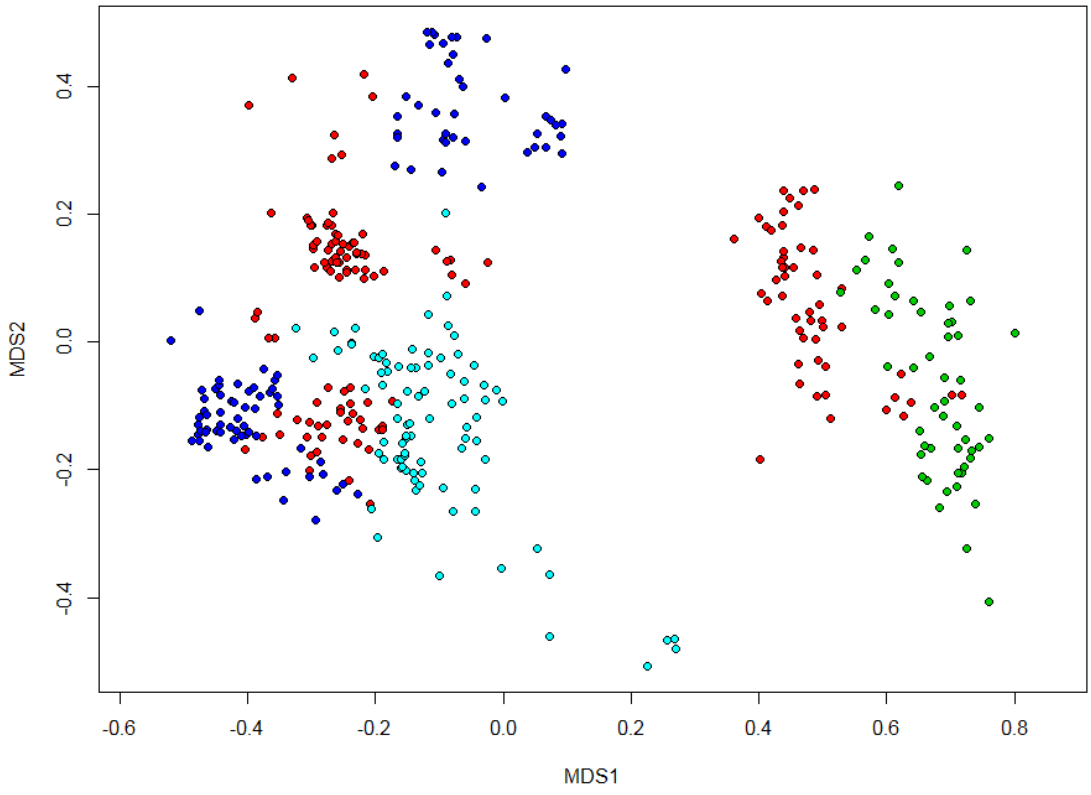


Figure 44: MDS ordination showing where each sample is located in relation to other samples in species space. Samples closer to each other are more similar than those separated by larger distances on the plot. A Hellinger transformed community matrix was used and MDS was conducted using Euclidean distances. Red: Morphotype 1; Green: Morphotype 2; Dark Blue: Morphotype 3; Light Blue: Morphotype 4. Stress=0.14.

Multivariate Mantel correlograms showed that for all eight ponds, similarity in bacterial communities tends to decrease with geographic distance (Figure 45). The distance where the similarity of bacterial community composition becomes non-significant, or the correlation coefficient becomes negative, represents a measure of distance where samples separated by greater distances are likely to differ significantly in composition. This relationship was dependent upon pond size, with larger ponds showing similarity occurring between samples separated by larger distances than at the smaller ponds. In Pete's Vent, samples were negatively correlated after 2.6m, but significant positive correlation was also detected at 4.7m. Similarity in bacterial communities was also detected between samples separated by 5m in Donut Pond. Whistler's Pond, Harjie's Pond and Jana's Vent all showed significant similarity in bacterial communities up to about 20m separation. Although samples were initially positively correlated at fine scales in Cygnet Pond, there was little evidence of any correlation occurring at scales up to 1000m.

4.3.2 Environmental control

The environmental variables (water chemistry and sediments) were used to constrain the bacterial community data using RDA. The RDA was able to significantly constrain 27.5% (R^2_{adj}) of the variation in bacterial communities to the environmental data, while 72.5% remained as residual variation ($F_{13,370}$=12.612, P=0.005), although the first eight RDA axes explained a significant proportion of the total constrained analysis (Table 19). The first RDA axis was the most important in explaining the environmental relationship of the communities as it explained 63.9% of the constrained variation.

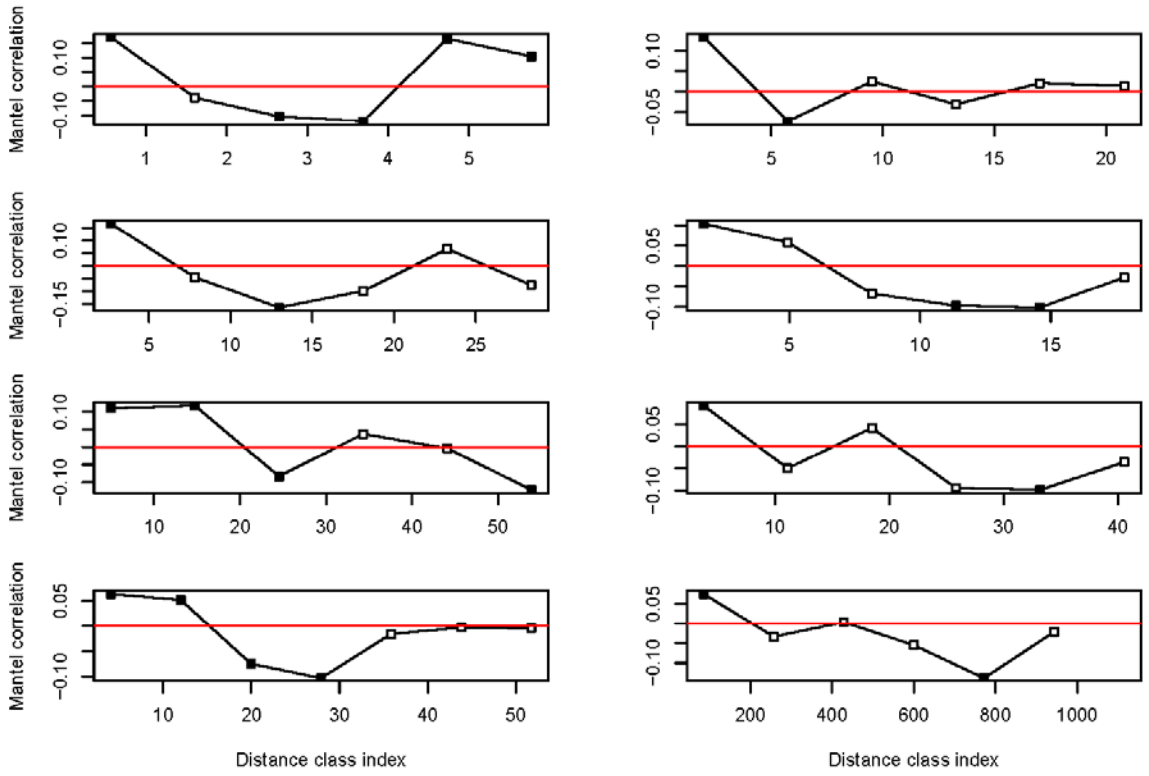


Figure 45: Multivariate correlograms showing the significance of spatial correlation in bacterial composition from transects in the eight study ponds. Black points represent scales with statistically significant similarity (positive Mantel correlation values) or dissimilarity (negative Mantel correlation values). White points represent non-significant values. a) Pete's Vent; b) Donut Pond; c) Annie's Pond; d) Pete's Pond; e) Whistler's Pond; f) Harjie's Pond; g) Jana's Vent; h) Cygnet Pond.

Table 19: Permutation ANOVA results (permutations=999) for each RDA axis using Hellinger transformed bacterial community and environmental variables. RDA model was significant ($F_{13,370}$=12.612, P=0.005).

	Df	Eigenvalue	Cum. Var. (%)	F	P>F	Sig
RDA1	1	0.09926	63.9	104.6229	0.001	***
RDA2	1	0.01906	76.1	20.0880	0.001	***
RDA3	1	0.01219	84.0	12.8467	0.001	***
RDA4	1	0.00790	89.1	8.3311	0.001	***
RDA5	1	0.00471	92.1	4.9670	0.001	***
RDA6	1	0.00434	94.9	4.5700	0.001	***
RDA7	1	0.00309	96.9	3.2540	0.001	***
RDA8	1	0.00177	98.0	1.8695	0.019	*
RDA9	1	0.00124	98.8	1.3117	0.153	
RDA10	1	0.00077	99.3	0.8125	0.647	
RDA11	1	0.00055	99.7	0.5843	0.934	
RDA12	1	0.00041	99.9	0.4344	0.996	
RDA13	1	0.00025	100.0	0.2658	1.000	
Residual	370	0.35105				

Table 20: Loadings of environmental variables for the first six RDA axes, which represent 94.9% of the constrained variation, done using the Hellinger transformed bacterial community matrix.

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
Na	0.01	0.20	-0.05	0.01	-0.09	0.30
Mg	-0.56	0.16	-0.27	-0.05	-0.18	0.21
Ca	-0.49	0.05	-0.43	-0.05	-0.16	-0.07
K	0.65	-0.10	0.18	0.13	-0.13	0.57
Cl	0.02	-0.28	-0.04	-0.15	0.04	-0.31
SO_4	0.08	0.13	0.29	0.31	0.19	-0.08
NH_4	-0.08	-0.14	-0.29	0.00	0.06	0.03
NO_x	0.04	0.27	-0.71	-0.14	0.04	0.32
Pelletal Skeletal Sand	0.27	-0.32	0.35	-0.37	-0.56	-0.11
Pelletal – Aragonite Mud	-0.46	0.04	-0.20	-0.14	-0.01	-0.33
Aragonite Mud	-0.20	0.15	-0.14	-0.48	-0.16	-0.10
Pelletal Diatom Sand	0.27	0.12	-0.27	0.13	-0.21	-0.03
Pelletal Sand	0.32	0.35	-0.06	0.53	0.23	0.08
Morphotype 2	-0.80	0.08	-0.34	-0.16	-0.05	-0.40
Morphotype 3	0.46	-0.35	-0.64	-0.13	0.46	0.14
Morphotype 4	0.21	0.81	0.36	-0.33	0.07	0.15

The different environmental variables contributed to the different RDA axes differently (Table 20). Potassium concentrations in the water were strongly associated with RDA1, and weakly associated with RDA2, making potassium concentration an excellent determinant of RDA1. Magnesium, calcium and Pelletal – Aragonite Mud were also associated with RDA1. The first RDA axis also separated ponds belonging to morphotype 2 (Cygnet), from the other morphotype groups. NO_x concentration and Pelletal Skeletal Sand were strongly associated with RDA2. The communities could therefore be described along axis RDA1, which represents samples with high magnesium and calcium concentration and low potassium concentration, to samples with low magnesium and calcium concentrations and high potassium concentration. Similarly, RDA2 represents a gradient from low to high NO_x concentration.

Three bacterial communities could be distinguished, although the samples were still grouped into pond groups (Figure 46). The first group, defined by low potassium levels, represents those communities found at Jana's Vent and Cygnet Pond. The second community, which had low concentrations of NO_x, consisted of samples collected from Pete's Vent and Harjie's Pond, whilst the third group, defined by low magnesium and calcium concentration and high NO_x consists of samples collected at Annie's Pond, Pete's Pond, Whistler's Pond and Harjie's Pond.

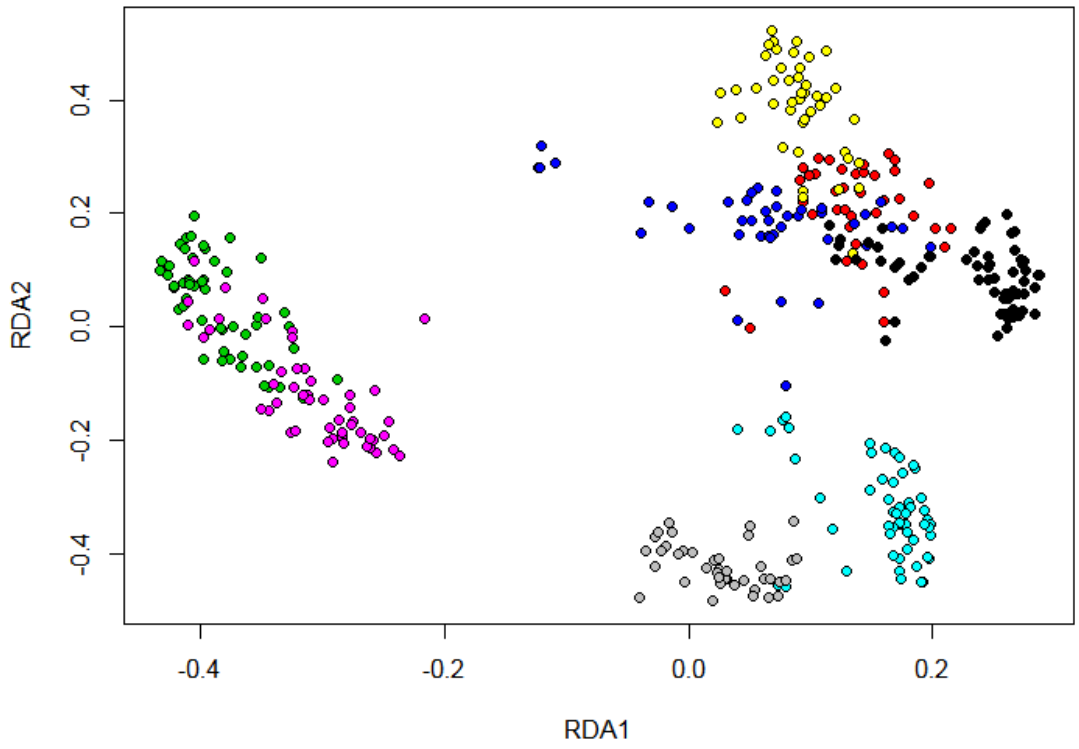


Figure 46: Ordination plot showing sample locations in environment space using RDA. The first axis explains 63.9% and the second axis 12.2% of the constrained variation. Red: Annie's Pond; Green: Cygnet Pond; Blue: Donut Pond; Light Blue: Harjie's Pond; Purple: Jana's Vent; Yellow: Pete's Pond; Grey: Pete's Vent; Black: Whistler's Pond.

4.3.3 Environmental vs. spatial determinants

Geographic location was accounted for using MEM variables, calculated using the coordinates of each sample and the Hellinger transformed community data. In total, 45 MEM variables were computed, but they did not explain a significant amount of variation in the bacterial communities (1.3%; Table 21). Forward selection of the environmental variables reduced the environmental matrix from 11 to 9 variables, whilst still explaining the same amount of variation. The selected variables included pond morphotype, sediment type, NO_x, Mg, Ca, SO_4^{-2}, Na, K and NH_4. The entire set of variables (Env + MEM + Pond) explained 61% of the variation found in the bacterial community. The environmental model (Env) explained 32.9 % of variation among bacterial community structure ($F_{14,369}$=14.39, P=0.005; Table 21). A third matrix consisting of the pond identity for each sample (Pond) was also constructed and

explained a significant proportion of the variation in the bacterial assemblages (52.1%; $F_{7,376}$=60.4, P=0.005; Table 21).

Variation partitioning of the bacterial community matrix with the environmental (Env), geographic location (MEM) and pond identity (Pond) variables revealed which factors were the best determinants driving the changes in bacterial community composition. The conditional effects (variation explained when the effect of the other two variables is removed) of all three variables was significant, although they explained different proportions of community variation (Table 21). Pond identity was the best determinant of bacterial community composition when the variation contributed by the other variables was accounted for (Pond|(MEM+Env)), explaining 22.1% of the total variation. The conditional effects of geographic location (MEM|(Env+Pond)) explained 7.7% of the variation while the environmental variables explained only 0.6% of the variation.

Table 21: Variation partitioning summary table done using partial RDAs. Env represents the 9 forward selected variables. Pond represents pond identity, and MEM represents the spatial location of the samples.

	R^2_{adj} (%)	DF_{model}	$DF_{residual}$	F	P-Value
Marginal Effects					
Env	32.9	14	369	14.39	0.005
MEM	1.3	45	338	1.11	0.074
Pond	52.1	7	376	60.40	0.005
Conditional Effects					
Env|Pond	1.2	11	365	1.91	0.005
Env|MEM	37.7	14	324	15.92	0.005
MEM|Pond	8.4	45	331	2.77	0.005
MEM|Env	6.1	45	324	1.82	0.005
Pond|Env	20.4	4	365	41.37	0.005
Pond|MEM	59.2	7	331	73.17	0.005
Env|(MEM+Pond)	0.6	11	320	1.47	0.005
MEM|(Env+Pond)	7.7	45	320	2.61	0.005
Pond|(MEM+Env)	22.1	4	320	46.86	0.005
Residual	39.0				

4.3.4 Demographic processes

Patterns of OTU co-occurrence were not consistent with a random pattern that would be expected using null models. There was highly significant segregation in bacterial OTUs in all eight ponds, with the observed C-scores always higher than the expected C-score (Table 22). These results suggest that bacterial OTUs co-occurred less frequently than would be predicted under a random process. The greatest amount of segregation in OTUs (largest effect size) was observed in Pete's Vent, the smallest study pond, whilst the least amount of segregation was found at Whistler's Pond. At the between pond scale ('*All ponds*' in Table 22), there was also highly significant segregation of OTUs in communities.

The neutral diversity parameter (θ) was greatest in Cygnet Pond and lowest in Whistler's Pond (Table 23). The estimated values of immigration (m) were low, however, with the highest value found in Pete's Pond of 0.96 ± 0.02, and the lowest in Donut Pond with 0.62 ± 0.06. The observed levels of β-diversity were all lower than the expected values (Table 24) estimated using simulated datasets generated by the neutral model of Etienne (2007, 2009). The effect sizes were all negative and significant, meaning more than 5% of simulated β-diversities were larger than the observed β-diversity.

Table 22: Results of testing if OTU co-occurrence in bacteria communities can be approximated by random distributions generated by null model analysis with 10,000 permutations.

Pond	C-score Obs	C-score Exp	C-score Var	Effect Size	P-value
Pete's Vent	52.49	45.33	0.01	99.03	<0.0001
Donut Pond	41.67	36.87	0.01	39.47	<0.0001
Annie's Pond	32.38	30.92	0.01	14.29	<0.0001
Pete's Pond	25.46	24.30	0.01	11.63	<0.0001
Whistler's Pond	26.42	25.19	0.04	6.11	<0.0001
Harjie's Pond	52.28	46.60	0.03	31.04	<0.0001
Jana's Vent	77.18	72.83	0.02	29.80	<0.0001
Cygnet Pond	119.44	113.61	0.02	44.46	<0.0001
All ponds	2.37	2.09	0.0002	19.10	<0.0001

Table 23: Estimates of neutral diversity (θ) and immigration (m ± standard error) for each pond using the multiple sampling formula given by Etienne (2009).

Pond	θ	m (±SE)	n
Pete's Vent	23.18	0.77 ± 0.03	46
Donut Pond	17.68	0.62 ± 0.06	47
Annie's Pond	10.69	0.89 ± 0.03	48
Pete's Pond	10.33	0.96 ± 0.02	48
Whistler's Pond	4.43	0.74 ± 0.05	64
Harjie's Pond	12.98	0.77 ± 0.03	61
Jana's Vent	26.35	0.93 ± 0.02	63
Cygnet Pond	29.51	0.83 ± 0.02	86
All Ponds	32.54	0.72 ± 0.13	8

Table 24: The observed β-diversity compared to expected β-diversity values obtained by simulating neutral communities. Communities were simulated using estimated parameters of neutral diversity (θ), immigration (*m*) and OTU abundance (*J*) from Etienne (2009) and 4999 simulations. Observed and expected β-diversity were calculated using Hellinger transformed OTU by site abundance matrices following Maaß *et al.* (2014).

Pond	Observed β-diversity	Simulated β-diversity	Effect Size	P-value
Pete's Vent	0.28	0.90	-21.32	<0.0001
Donut	0.32	0.89	-17.31	<0.0001
Annie's	0.25	0.80	-9.51	<0.0001
Pete's	0.19	0.78	-9.79	<0.0001
Whistler's	0.17	0.68	-5.01	0.0006
Harjie's	0.21	0.84	-13.04	<0.0001
Jana's	0.27	0.90	-24.37	<0.0001
Cygnet	0.31	0.91	-28.44	<0.0001
All Ponds	0.50	0.64	-3.98	<0.001

4.4 Discussion

The bacterial communities in the ponds of Lake MacLeod exhibit clear biogeographic patterns in spatial distribution. This work complements a growing body of literature that has found significant taxa-area and distance-decay relationships in bacterial communities (Langenheder and Ragnarsson 2007, Soininen *et al.* 2007, Fuhrman *et al.* 2008, Bell 2010, Martiny *et al.* 2011, Astorga *et al.* 2012, Lear *et al.* 2013, Lear *et al.* 2014). This study is however, unique in that it is one of the few investigations (but see Yannarell and Triplett 2004, Jones *et al.* 2012, Lear *et al.* 2014) of a) a salt lake system, b) conducted at fine spatial scales (<1km), and c) which is not limited by the collection of a small number of samples for each water body. Concentrating on t he benthic communities allowed for the sampling of communities using a spatially explicit sampling design that minimised the effect of sampling disturbances and enabled

comparisons across small distances, something difficult to achieve when investigating planktonic communities.

This study found that significant variation in bacterial communities occurs at spatial scales similar to those observed by Lear *et al.* (2014), but largely underestimated by previous studies (for example, Horner-Devine *et al.* 2004, Martiny *et al.* 2011, Lear *et al.* 2013). However, this study also found that the amount of variation in bacterial communities was different for each pond, and that the distance where the bacterial communities were not autocorrelated generally increased with an increase in pond size. This distance varied from meters in the smaller ponds to hundreds of meters in the largest pond. Reasons for these changes may be due to the characteristics of smaller ecosystems, where 'edge-effects' induce greater spatial variability in environmental conditions and higher levels of immigration (Lear *et al.* 2014). Indeed, there was a large amount of within pond variability in bacterial community composition, which contradicts previous studies where bacterial communities were thought to be largely homogeneous within aquatic systems (Reche *et al.* 2005, Dorigo *et al.* 2006, Humbert *et al.* 2009).

Samples within ponds were largely dissimilar to samples from other ponds, despite some of the ponds being similar environmentally. This could be expected because of the high connectivity between samples located in the same pond. When samples (or communities) are highly connected, it is difficult for segregation to occur (Shade *et al.* 2010). Highly connected aquatic habitats are also likely to show less environmental heterogeneity (Shade *et al.* 2008). Nonetheless, three main bacterial community groups were distinguishable: those belonging to (i) Donut, Annie's, Pete's and Whistler's Pond, (ii) those belonging to Pete's Vent and Harjie's Pond, and (iii) those belonging to Cygnet Pond and Jana's Vent.

The similarity between the Jana's Vent and Cygnet Pond communities could be expected due to both these water bodies being located in the southern reach of the Northern Ponds, and by the fact that water is discharged from Jana's Vent, via a channel, into the southern area of Cygnet Pond. On the other hand, Pete's Vent discharges water into Pete's Pond, but these two ponds had very distinct communities. It is possible that these communities are different because Pete's Vent is exposed to

greater ‘edge-effects’ from the surrounding terrestrial environment when compared to the community from the larger Pete’s Pond system. Similar conclusions have been made before regarding bacterioplankton in small alpine tarns (Lear *et al.* 2014).

Some environmental variables were found to be more important than others as determinants of bacterial community composition. For example, the water chemistry, sedimentology and pond morphotype groups explained approximately 30% of the variation in bacterial community composition. However, most of this explained variation (64%) was driven by the unique bacterial community composition of Cygnet Pond its unique environmental and morphometry characteristics as described in the previous two chapters. Whether the environmental variables accounted for in this study are the actual drivers of change in bacterial community composition, or if they represent intrinsic pond-level environmental variation is, however, difficult to elucidate.

The high levels of unexplained variation found when constraining bacterial communities to a set of environmental parameters is common (Beisner *et al.* 2006, Van der Gucht *et al.* 2007, Lindström *et al.* 2010, Caruso *et al.* 2011, Martiny *et al.* 2011). For example, 22% of the variation in hypolithic bacterial communities was explained using spatial and environmental parameters (Caruso *et al.* 2011). Similarly, Martiny *et al.* (2011) could only account for 27% of the variation in salt marsh bacterial communities using environmental parameters. Lear *et al.* (2014) suggest that when accounting for variation across sites that belong to different isolated systems, such as ponds, the composition of dissimilar communities is more likely to be due to stochastic events, such as immigration and disturbances than environmental variation.

Generally though, ecological studies, including non-bacterial studies, explain a greater amount of community variation when environmental parameters are accounted for in combination with spatial factors (as reviewed by Cottenie 2005, Hanson *et al.* 2012). Although initial RDA analyses revealed a large amount of variation could be explained using the environmental variables alone, when the effect of geographic location and pond identity were accounted for, using partial RDAs, the explained variation contributed by the environmental variables was found to decrease from 38% to less than 1%. This suggests that there is no effect of the within pond environmental variation, such as salinity gradients, on bacterial community composition despite the analysis

suggesting that the environmental differences detected between ponds (previous chapter) may be important in describing the changes at the pond scale.

Similar levels of variation were accounted for using partial RDAs by Caruso *et al.* (2011). They found that similarly low levels of variation in bacterial community composition could be attributed to the marginal effects of geographic location and environment. However, they could explain a significant amount of variation when they used a factor to group their samples into ecologically meaningful categories, analogous to the pond morphotype and pond identity categories used in this study. Pond identity was the best determinant of bacterial community composition in this study, with the conditional effect explaining 22% of the variation. There are two possible explanations for this. Either an important environmental parameter, which is driving the variation in bacterial community composition among ponds, was not measured, or immigration rates are not sufficiently great to prevent the bacterial communities from differentiating via ecological drift.

The reasons as to why the environmental parameters failed to explain a large proportion of the variability in community structure could be due (i) the lack of data of potentially important environmental variables, such as trace metal concentrations (Mason 2013), (ii) ignoring the effect of biotic variables, including the top-down effects of higher order organisms (Caruso *et al.* 2011, Hanson *et al.* 2012) and (iii) the limitations of bacterial DNA fingerprinting in describing the bacterial communities (Blackwood *et al.* 2007). But it could also indicate a possible mechanistic role, whereby environmental determinants and species sorting play a minor role in structuring the communities, and other demographic processes, perhaps interacting with stochastic events, play a pivotal role.

This study was done at a single point in time, and even though every attempt was made to capture the spatial variation in environmental conditions, it is possible that important, unaccounted, temporal environmental variation is driving the observed variation in bacterial community composition (Jones *et al.* 2012). Temporal variation of the environmental conditions could be expected to affect each pond differently, particularly because of the different morphometric properties of each pond. For example,

fluctuations in temperature, dissolved oxygen and pH can be expected to vary differently in water bodies with different surface areas and volumes (Fee *et al.* 1996).

Patterns of community composition showed evidence of niche partitioning and non-neutrality when null and neutral models were applied to the data. Using null model analysis (Gotelli and Entsminger 2004), it was found that in all eight ponds, OTU co-occurrence was non-random, and the bacterial communities within the ponds displayed high levels of divergence (OTUs co-occurred less frequently than randomly predicted). The β-diversity was also lower in every pond than what was expected using the models assuming using neutrality (Etienne 2009). Non-neutrality of bacterial communities has been found before (Caruso *et al.* 2011, Livermore and Jones 2015), and further supports the idea of species-sorting (and niche partitioning) processes being enhanced in harsh environments, such as deserts and salt lakes (Chase 2007, Ofiteru *et al.* 2010, Caruso *et al.* 2011). It is also possible that the neutral models proposed by Etienne (2009) underestimate immigration rates.

Given that the pond identity explained the greatest amount of variation in bacterial community composition, and this was independent of the environmental and spatial variables, it is possible that a large amount of the variation in these communities can be explained due to demographic stochasticity, such as species interactions and/or immigration and dispersal effects (Tilman 1982, Hubbell 2001). If it can be assumed that beta diversity depends on the dispersal and stochastic demographic fluctuations upon which neutral theories are based (Maaß *et al.* 2014), it is unlikely that dispersal limitations between the samples within ponds are drivers of the bacterial community structure and diversity. Indeed, immigration rates need to be relatively high to induce changes in bacterial community structure. It has been shown that dispersal rates need to be greater than 10% of the standing stock per day to have any significant effect on the composition of bacterial communities (Lindström and Östman 2011).

Non-random species co-occurrence and non-neutrality were features of some of the hypolithic desert bacterial communities that were sampled from deserts around the world (Caruso *et al.* 2011). In such a type of ecosystem, the influence of null models (species co-occurrence) and neutrality on the structure of bacterial communities was different for two trophic guilds, phototrophs and heterotrophs (Caruso *et al.* 2011).

Similarly, different responses to environmental determinants and the role of species sorting processes have been observed between generalists and specialists functional groups of bacteria in a community (Székely and Langenheder 2013). These studies are important as they highlight that assembly processes may differ between different guilds within a bacterial community.

This highlights another limitation of the current study is the broad taxonomic resolution that ARISA, as well as other fingerprinting methods (Blackwood *et al.* 2007). Rare species within bacterial communities have been found to occur at very low relative abundances (probably less than 0.1%) within microbial communities. These rare taxa probably make up the majority of the taxonomic richness found in microbial communities (Reid and Buckley 2011). It is therefore likely that richness was underestimated by the broad OTU categories used with community fingerprinting techniques like ARISA. Nonetheless, the majority of OTUs detected in this study were uncommon, with nearly 50% of the OTUs being present in less than 10% of the samples. Furthermore, the taxonomic resolution of ARISA has been found to be sufficient when testing distribution based hypotheses of bacterial communities (Lee 2014). Besides, the money saved from using a fingerprinting technique, such as ARISA, allows for a greater number of samples to be processed than could be achieved using sequencing methods.

The findings of this study add to a growing body of literature showing that bacterial communities have complex biogeographic patterns similar to those observed in other groups of organisms. The structure of these communities was found to not only be driven by environmental determinants (the traditional view of microbial ecology) but also by demographic processes. Species-sorting processes, as well as other deterministic mechanisms, permit some groups of bacteria to inhabit a community, whilst the species co-occurrence models suggest that OTUs are largely segregated and either do not co-occur, or exist at abundances below the detection limit of ARISA.

CHAPTER 5. CILIATE ECOLOGY OF THE PONDS

5.1 Introduction

Ciliates are an important component of all aquatic microbial communities, and are represented by free-living, mostly heterotrophic, species as well as host-bound species (Adl *et al.* 2005). Current estimates suggest that there are up to 40,000 free living ciliate species globally (Foissner *et al.* 2008). They are a diverse group of unicellular eukaryotic organisms that range in size from 0.2 - 200μm and exist in complex communities, which are diverse in terms of physiology, life cycles, and reproduction and dispersal abilities. Aspects of this diversity, in particular, the ability to sexually reproduce, make them fundamentally different to bacteria, which have largely been at the focus of microbial ecology (Caruso *et al.* 2011, Lear *et al.* 2014). The biogeography of ciliates, however, remains a controversial topic (Caron 2009).

Early studies suggest that ciliates have cosmopolitan distributions, leading some researchers to hypothesise that the entire global freshwater ciliate community could be found in a single pond (Finlay *et al.* 1998). Ciliates were thought to have cosmopolitan distributions, due to characteristics of their tiny size and ability to disperse easily, high numbers in populations and low likelihood of extinction (Finlay and Clarke 1999, Finlay 2002, Fenchel and Finlay 2003, Fenchel and Finlay 2004). Yet, such high global diversity of ciliates could not be expected if the species are cosmopolites, and, as traditional hypotheses predict, are '*found everywhere, where the environment permits*' (Baas-Becking 1934).

Evidence that ciliates are not '*everywhere*' has recently been found by looking for endemic taxa. For example, Foissner *et al.* (2008) found that past geological events, including the split of Pangaea, have deeply influenced ciliate distributions, and that there is evidence of continental, regional and local endemism of ciliates that is not dissimilar to patterns found in other eukaryotes (such as plants and animals), although the proportion of cosmopolites is much higher. Foissner *et al.* (2008) provide evidence for restricted continental distributions in some ciliate species. For example, *Apofrontonia dohrni* and *A. lemetschwandtneri* both inhabit muddy coastal puddles, but have only ever been found on di fferent continents (Europe and South America,

respectively; Foissner 2007). By studying conspicuous taxa, which are easily detectable if they are in a sample, these studies have been able to show that ciliates have distinct distributions, and are thus not '*everywhere the environment permits*'.

Speciation by isolation has been demonstrated by Miao *et al.* (2004) across geographically separated ciliate (*Carachesium polypinum*) populations in the Yangtze and Pear Rivers of China. Interestingly though, the genetic polymorphism and phylogeographic relationships between *C. polypinum* populations was remarkably similar to that of the freshwater fish populations inhabiting the same habitats, providing more evidence that ciliate populations are structured similarly to other eukaryotes. Although this study provides evidence of speciation occurring between two, geographically distinct populations, Miao *et al.* (2004) did not test for the presence of distance-decay and taxa-area relationships, which are central to understanding the biogeography of ciliates.

Recent evidence for small-scale biogeographic patterns in ciliate distributions has been found in lake protist communities. Significant distance-decay and taxa-area patterns were found in lake protist communities, even when the effect of the environmental variation was accounted for (Lepère *et al.* 2013). Similar patterns have been found in the genetic structure of protist communities in Antarctic saline lakes, where geographic variation in the populations was detected at scales less than 9km (Rengefors *et al.* 2012). These two studies show that inland aquatic systems, including saline lakes, act as ecological islands, and despite the close proximity of some lakes, can contain distinct protist communities. Although these two above studies considered the biogeographic patterns of protists, they both targeted dinoflagelates. It is not known if distance-decay and taxa-area relationships, as well as the determinant of these patterns, can also be found in aquatic ciliate communities in general.

Most studies of ciliate diversity and ecology use microscopy-based methods for identification of species (Andrushchyshyn *et al.* 2007, Reiss and Schmid-Araya 2008, Dopheide *et al.* 2009). Although microscopy techniques are inexpensive, using morphological markers to characterise taxa is difficult because of their small sizes. It is also impossible to distinguish between many ciliate species morphologically (Dopheide *et al.* 2009), which has led to the development of DNA-based methods to detect species.

Few studies have used molecular techniques to investigate micro-eukaryotic biogeography (Dopheide *et al.* 2009, Monchy *et al.* 2011, Rengefors *et al.* 2012, Lepère *et al.* 2013). Molecular techniques provide a method which can detect taxonomic richness of a large number of samples, relatively quickly. These methods are also not restricted to detecting 'morphotypes', and can thus have the potential to capture a greater number of taxa (Dopheide *et al.* 2008). Capturing as much of the taxonomic richness as possible is important when trying to understand small-scale variation in community composition.

Even though the literature on protist diversity in ponds is fairly rich (Finlay and Esteban 1998, Foissner 2006, Foissner *et al.* 2008), few studies have investigated patterns of alpha and beta-diversity among lakes at fine scales (Rengefors *et al.* 2012, Lepère *et al.* 2013), and only Dopheide *et al.* (2009) and Reiss and Schmid-Araya (2008) have specifically targeted ciliates. Previous studies suggest that spatial processes and environmental variation are strong determinants of ciliate community composition in lakes (Rengefors *et al.* 2012, Lepère *et al.* 2013).

This study used molecular techniques to investigate small-scale biogeographic patterns in the ciliate communities of Lake MacLeod. Because each pond has a unique set of morphotypic and environmental variables (Chapter 2 and Chapter 3), it is expected that the composition of each pond community will be different. Because each pond represents an 'island' it is expected that demographic processes based on immigration and dispersal limitation between samples will be able to explain large scale (between pond) diversity patterns based on the assumptions of neutrality (Etienne 2007, 2009). It is predicted that distance-decay and taxa-area relationships will be found for the ciliate communities in Lake MacLeod. Using partial redundancy analysis, the effect of environmental and spatial variation on the composition of the ciliate communities will be investigated. Similarly, species co-occurrence and neutral models will be used to predict if communities are randomly assembled or if the community composition can be predicted using neutral models. Specifically, this chapter will test the following hypotheses central to understanding the environmental and demographic processes that structure the benthic ciliates communities in aquatic systems:

1. larger ponds have higher richness of taxonomic groups;
2. the ciliate assemblages among ponds differ significantly;
3. environmental heterogeneity and spatial variables explain significant amounts of variation in bacterial community composition;
4. OTU co-occurrence patterns are not random; and
5. immigration of taxa between sample locations affects the composition and diversity of bacterial communities.

5.2 Methods

Sediment samples were collected in conjunction with those collected for water chemistry (Chapter 3) and used for bacterial community analyses (Chapter 4) and ciliate community analyses (this Chapter). Similarly, the genomic DNA that was extracted for the bacterial analysis described in Chapter 4 were also used to profile the ciliate community. As the methodology of this chapter largely overlaps that of Chapter 4, only steps specific to this chapter will be described here. The reader should refer to section 4.2 of this thesis for methodological details. Because ciliate DNA failed to amplify in some samples, a slightly different set of variables to those used in Chapter 4 were used to represent the environmental conditions. A summary of the water chemistry and sediment variables for each pond used in this chapter is provided in Table 25 and Table 26.

Table 25: Summary of water chemistry variables (mean ± SE) used in environmental modelling for this chapter. All units at mg l^{-1}.

	Na	Mg	Ca	K	NH_4	NO_x	PO_4	Cl	SO_4
Pete's	9685	1033	370	323	0.36	0.042	0.022	19033	2525
Vent	(195)	(21)	(8)	(7)	(0.05)	(0.003)	(0.001)	(1069)	(172)
Donut	10481	1147	376	344	0.21	0.048	0.016	17001	2368
Pond	(290)	(36)	(12)	(11)	(0.03)	(0.003)	(0.002)	(446)	(147)
Annie's	10194	1100	383	341	0.37	0.017	0.021	21590	3470
Pond	(264)	(30)	(13	(10)	(0.07)	(0.002)	(0.001)	(1714)	(272)
Pete's	9583	1032	342	288	0.06	0.033	0.015	14130	2288
Pond	(185)	(22)	(8)	(5)	(0.01)	(0.001)	(0.001)	(505)	(65)
Whistler's	9329	1031	379	323	0.26	0.056	0.019	14969	1852
Pond	(239)	(28)	(9)	(8)	(0.02)	(0.003)	(0.002)	(348)	(48)
Harjie's	9377	1027	364	326	0.15	0.013	0.016	16207	2095
Pond	(227)	(27)	(10)	(8)	(0.02)	(0.001)	(0.001)	(456)	(80)
Jana's	11412	1327	432	395	0.29	0.028	0.023	17881	2884
Vent	(171))	(22)	(8)	(6)	(0.03)	(0.002)	(0.002)	(358)	(308)
Cygnet	13461	1962	742	139	0.43	0.065	0.059	24637	2917
Pond	(202	(23)	(22)	(3)	(0.03)	(0.003)	(0.006)	(461)	(131)

Table 26: Proportional contribution (as a percentage) of each sediment type defined in Chapter 4 for each pond used in environmental modelling.

	Pelletal Sand & Aragonite Mud	Pelletal Skeletal Sand	Pelletal – Aragonite Mud	Aragonite Mud	Pelletal Diatomaceous Sand	Pelletal Sand
Pete's Vent	61	0	0	0	4	36
Donut Pond	3	41	0	6	17	31
Annie's Pond	0	6	0	0	19	75
Pete's Pond	3	0	0	0	0	97
Whistler's Pond	0	4	0	0	24	71
Harjie's Pond	2	63	0	0	2	30
Jana's Vent	35	2	0	0	0	63
Cygnet Pond	52	0	36	6	0	6

5.2.1 Community fingerprinting

Terminal restriction fragment length polymorphism (T-RFLP) is an efficient and inexpensive means to compare ciliate diversity between samples (Dopheide *et al.* 2009, Lear *et al.* 2011). DNA was extracted from the frozen sediment samples using the PowerLyzer™ PowerSoil® DNA Isolation Kit (MO BIO Laboratories Inc., Carlsbad) following the manufacturer's instructions. PCR was used to amplify a ~700 BP fragment of the ciliate 18S rRNA gene using the primers 384F (5'-YTB GAT GGT AGT GTA TTG GA-3') and 1147R (5'-GAC GGT ATC TRA TCG TCT TT-3') designed by Dopheide *et al.* (2008). The 5' termini of 384F and 1147R were labelled with 6-carboxyhexachlorofluorescein (HEX) and 6-carboxyfluorescooein (FAM) fluorophores, respectively. PCR was done following Dopheide *et al.* (2009): (i) incubation at 94°C for 5 mins; (ii) 35 cycles of 94°C for 45s, 55°C for 60s and 72°C for 90s before a final step of 72°C for 7 mins. PCR products were digested for 4 hours at 37°C with the restriction endonucleases HaeIII and RsaI in 10 µl reaction mixtures. PCR products were purified using Sera-mag Speedbeads and AMPure™ PCR cleanup kits. Each digestion mixture contained 0.5 µl of each enzyme, 1.5 µl of reaction buffer and about 100ng of purified amplicon. Purified PCR digested fragments were combined with 10µl of HiDi formamide and 0.6 µl GeneScan™ 1200 LIZ® dye size standard (Applied Biosystems Ltd., Melbourne, Australia), before being heat treated at 95°C for 5 mins and cooled on ice. The samples were run on a 3730XL DNA Analyser (applied

Biosystems Ltd) using a 50cm capillary. All DNA extractions, PCR amplifications, digestions and separation were done at the Australian Genome Research Facility.

5.2.2 Statistical analyses

As in the previous chapter, the peak profiles generated by the genetic analyser were binned using the R-scripts of Ramette (2009). To include the maximum number of peaks, and exclude background fluorescence, only peaks with a height greater than 150 relative fluorescence units, and with between 10 and 650 BP were used (Dopheide *et al.* 2009). Bins of 3 BP (±1 BP) were used. This yielded two T-RFLP profiles, one blue (HEX- labelled) and one green (FAM- labelled). Because each fragment is expected to produce a HEX-labelled and a FAM-labelled peak, richness and diversity estimates were made using only HEX-labelled peaks. However, for multivariate analyses, both HEX-labelled and FAM-labelled peaks were used as a single data set for each sample (Dopheide *et al.* 2009). Multivariate analyses, as well as null and neutral models, were performed as described in section 4.2 of this thesis.

5.3 Results

5.3.1 Richness and diversity

Community fingerprinting was successfully completed for 330 samples across the eight ponds. Similar amounts of OTU richness were measured for the blue (5' label) and green fragments (3' label; Table 27), however, since the blue fragments consistently accounted for slightly higher richness, only these fragments were considered in further measurements. In total, 30 OTUs were detected, although most OTUs were only represented in 1-10% of samples (Figure 47). Four dominant OTUs were found in more than 20% of the samples, with one OTU in particular, with a blue fragment length of 293 BP, being found among 80% of samples and accounting for 46.6% of the total abundance. The second and third most frequently occurring peaks occurred at 641 and 608 BP, respectively. Furthermore, 75.9% of the OTU's accounted for less than 1% of the total abundance. Five OTUs, occurring at 56, 77, 248, 431 and 446 BP, were found to occur only in a single pond, whilst 11 OTUs were found in 6 or more of the ponds; only one of these OTUs, that occurred at 293 BP, was found in all eight ponds.

Richness was estimated using the species accumulation plots (Figure 48) and the boot-strapping method of Smith and van Belle (1984), and show good agreement with the number of OTUs detected using T-RFLP (Table 27). There was no significant linear relationship between ciliate OTU richness and pond size ($F_{1,6}$=0.504, P=0.505), although the largest pond was the richest. There was no trend of decreasing species-area exponent (*z*) with pond size, with the two smallest ponds Pete's Vent and Donut Pond having the same *z*-value as Cygnet Pond, the largest pond. The least number of OTUs were found in Annie's Pond (6 OTUs) and the greatest richness found in Cygnet Pond (22 OTUs). Diversity was greatest in Pete's Vent, which had 18 OTUs and lowest at Annie's Pond, based on both Shannon and Simpson Diversity Indices of the blue T-RFLP fragments. Similarly, Pete's Vent had the highest levels of community evenness, whilst Pete's Pond had the lowest.

Table 27: Diversity indices for each pond and both blue and green T-RFLP fragment. All indices are calculated using the total community found within each pond. Values for the blue fragments are recorded above those for the green fragments.

	Total OTU Richness	Shannon Diversity	Shannon Evenness	Number of Samples	Estimated Richness	*z*
Pete's Vent	**18** **16**	**8.37** **8.01**	**0.47** **0.50**	32	19	0.47
Donut Pond	**15** **14**	**5.36** **4.89**	**0.36** **0.35**	32	17	0.47
Annie's Pond	**6** **7**	**1.67** **3.17**	**0.28** **0.45**	19	7	0.34
Pete's Pond	**14** **12**	**2.70** **2.87**	**0.19** **0.24**	39	15	0.37
Whistler's Pond	**11** **11**	**3.62** **3.08**	**0.33** **0.28**	52	12	0.32
Harjie's Pond	**14** **14**	**4.64** **4.64**	**0.31** **0.33**	52	14	0.26
Jana's Vent	**17** **14**	**3.32** **3.45**	**0.20** **0.25**	54	17	0.28
Cygnet Pond	**22** **21**	**4.94** **4.50**	**0.22** **0.21**	50	24	0.47
Entire Data Set	**30** **29**	**6.33** **6.31**	**0.21** **0.22**	330	32	

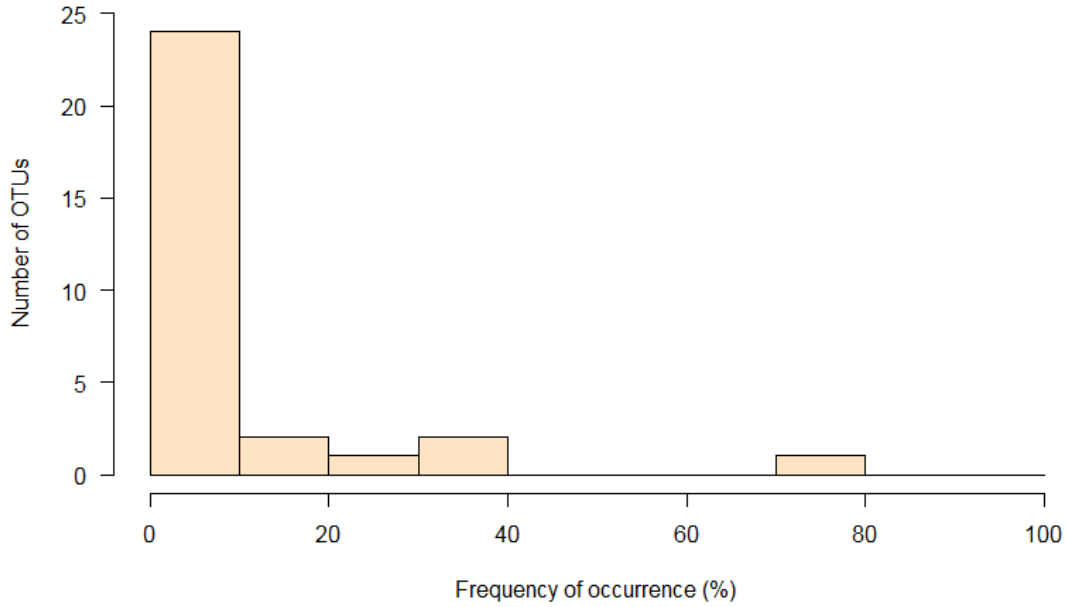


Figure 47: Frequency histogram of species occurrences of all 30 blue-labelled T-RFLP fragments (OTUs).

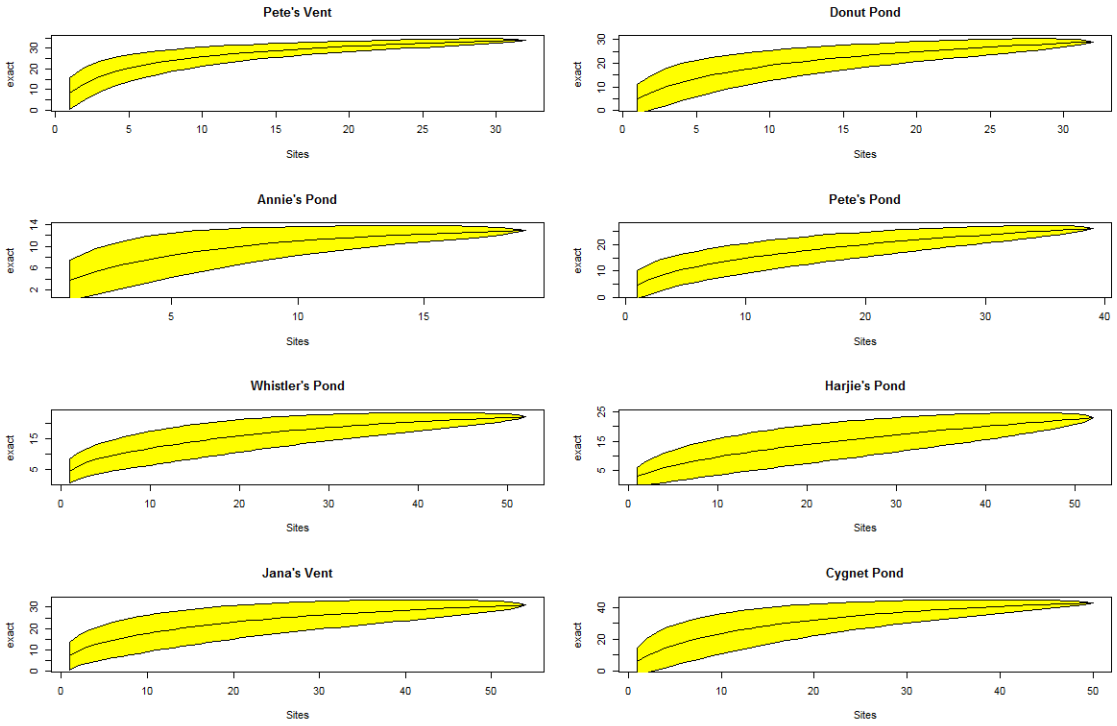


Figure 48: Species accumulation plots for each pond with yellow shaded area representing standard deviation.

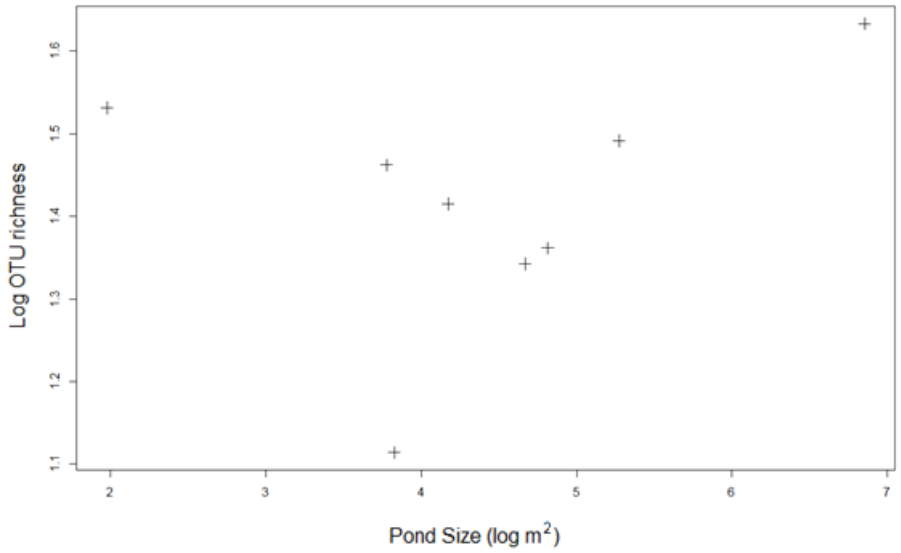


Figure 49: Plot of OTU richness against log Pond Size (m^2). There was no significant linear relationship between OTU richness and pond size ($F_{1,6}$=0.504, P=0.505).

The composition of the ciliate communities was different for each pond (Global PERMANOVA Pseudo $F_{7,322}$=29.52, P=0.001; Figure 50). Pairwise comparisons indicate that Cygnet Pond and Jana's Vent have different ciliate communities, as well as Donut Pond with Harjie's Pond and Jana's Vent (pairwise PERMANOVAs, all P_{adj}<0.05; Table 28). Unlike to what was found for the bacterial communities in section 4.3.1, Pete's Vent had a similar composition to Pete's Pond, which into flows. There were three main OTUs that explained the variation in community structure in these two axes. These OTUs had blue fragment lengths of 293, 608 and 641 BP. The OTU at 293 BP was associated with Annie's Ponds, Donut Pond, Harjie's Pond, Pete's Pond and Whistler's Pond. The OTU with a fragment length of 608 BP was strongly associated with Pete's Vent. Fragment length 641 BP was associated with Jana's Vent, while Cygnet Pond showed associations with both 608 and 641 BP fragments. ANOSIM results using Bray-Curtis dissimilarity matrix show that between pond dissimilarity was greater than within pond dissimilarity (R=0.460, P=0.001; Figure 50). Likewise, ANOSIM showed that between morphotype dissimilarity was greater than within morphotype dissimilarity (R=0.219, P=0.001). Similarly, the pond morphometry (morphotypes) had different ciliate communities (Global PERMANOVA Pseudo$F_{3,279}$=16.46, P=0.001).

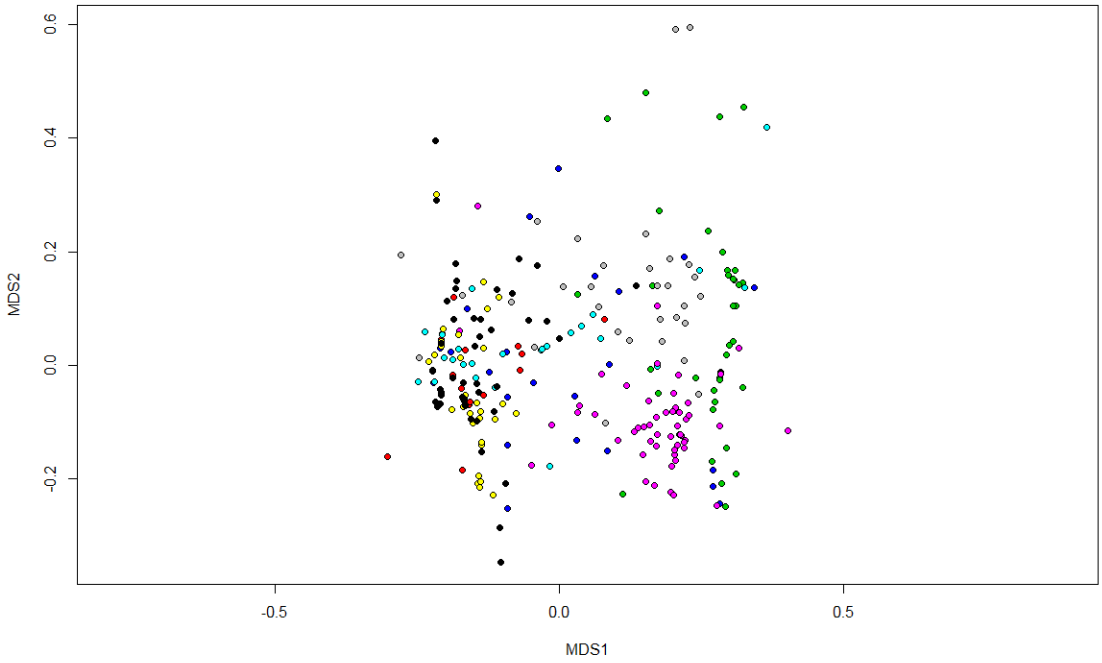


Figure 50: MDS ordination showing where each sample is located in relation to other samples in species space. Samples closer to each other are more similar than those separated by larger distances on the plot. A Hellinger transformed community matrix was used, and MDS was conducted using Euclidean distances. Grey: Pete's Vent; Dark Blue: Donut Pond; Red: Annie's Pond; Yellow: Pete's Pond; Black: Whistler's Pond; Light Blue: Harjie's Pond; Purple: Jana's Vent; Green: Cygnet Pond. Stress=0.15.

Table 28: Distance matrix showing the distances between the centroids of each pond group in principal coordinate space (PCoA) using Hellinger transformed ciliate community data. Bold values indicate significant differences in the ciliate community composition of the ponds (P<0.05).

	Pete's Vent	Donut Pond	Annie's Pond	Pete's Pond	Whistler's Pond	Harjie's Pond	Jana's Vent
Donut Pond	16.98						
Annie's Pond	20.00	36.99					
Pete's Pond	12.09	29.07	7.92				
Whistler's Pond	9.23	26.21	10.78	2.86			
Harjie's Pond	23.72	**40.70**	3.72	11.63	14.49		
Jana's Vent	21.41	**38.39**	1.41	9.33	12.19	2.31	
Cygnet Pond	11.91	5.07	31.92	24.00	21.14	**35.63**	**33.32**

Multivariate Mantel correlograms of the Hellinger transformed community data using Euclidean distances show that some of the samples within the ponds were similar, although somewhat weakly, at distances of less than 15m (Figure 51). However, the small ponds, including Pete's Vent, Donut Pond, Annie's Pond and Pete's Pond were not significantly similar at any distance. Whistler's Pond and Harjie's Pond both showed significant similarity at small scales (between 10 and 15 m) and Jana's Vent had significant similarities between samples at small scales. There was no significant auto-correlation at any scale in Cygnet Pond.

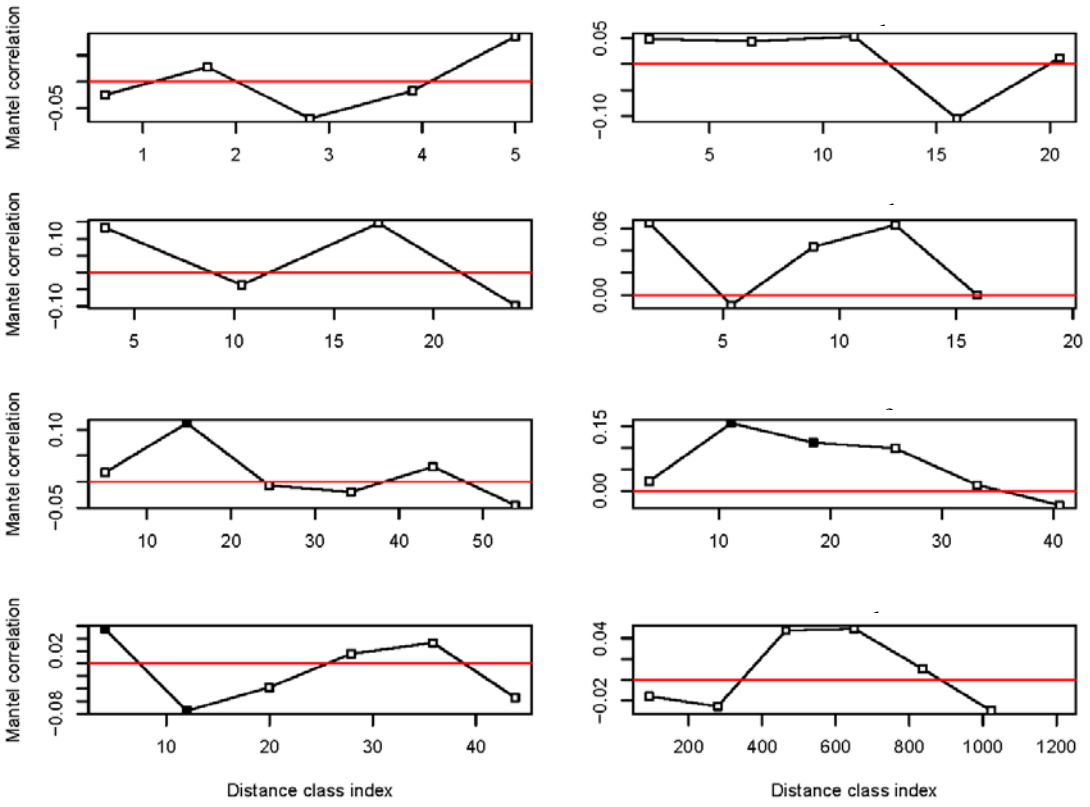


Figure 51: Multivariate correlograms showing the significance of spatial correlation in ciliate community composition from the transects in the eight study ponds. Black points represent scales with significant spatial autocorrelation (positive Mantel correlation values) or spatial clustering (negative Mantel correlation values), while white points represent non-significant values. a) Pete's Vent; b) Donut Pond; c) Annie's Pond; d) Pete's Pond; e) Whistler's Pond; f) Harjie's Pond; g) Jana's Vent; h) Cygnet Pond.

5.3.2 Environmental control

The environmental variables (water chemistry and sediments) were used to constrain the ciliate community data using RDA, of which, the first four axes were significant (Table 29). The RDA was able to significantly constrain 21.1% (R^2_{adj}) of the variation in ciliate communities with the environmental data, whilst 79.9% remained as residual variation ($F_{16,263}$=5.67, P=0.005; Figure 52). The first RDA axis was the most important in explaining the environmental relationship of the communities and it explained only 68.7% of the constrained variation.

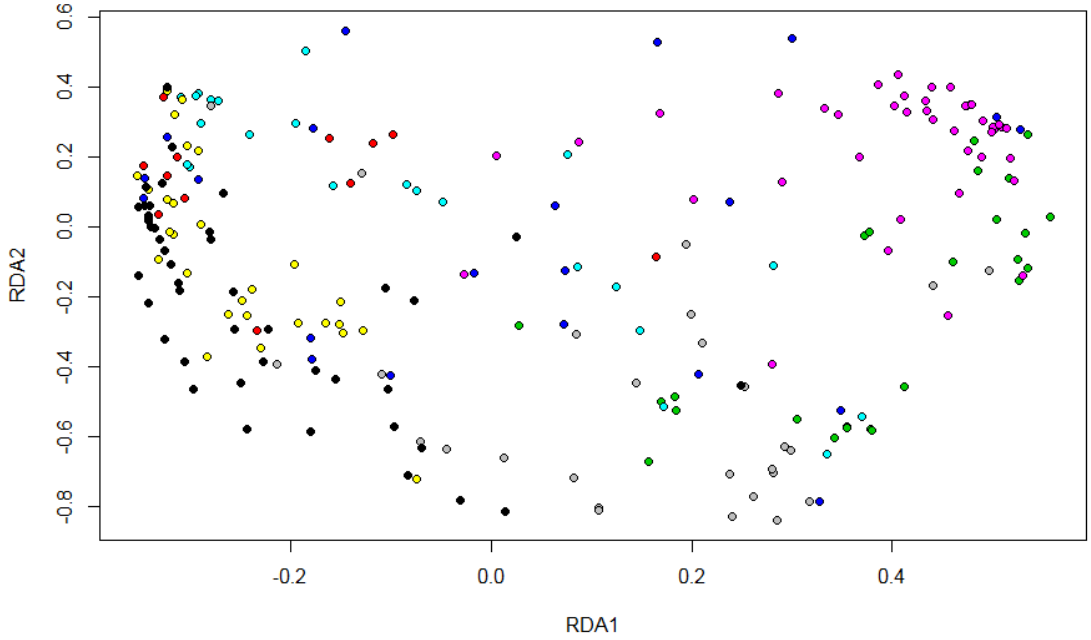


Figure 52: RDA plot showing sample locations relative to one another after the Hellinger transformed ciliate community data was constrained to the environmental data. The first axis explains 69% and the second axis 15% of the constrained variation. Different colours represent different ponds. Red: Annie's Ponds; Green: Cygnet Pond; Dark Blue: Donut Pond; Light Blue: Harjie's Pond; Yellow: Pete's Pond; Grey: Pete's Vent; Black: Whistlers Pond.

Table 29: Permutation ANOVA results (permutations=999) for each RDA axes using Hellinger transformed ciliate community and environmental variables. RDA model was significant ($F_{13,266}$=5.725, P=0.005).

	Df	Eigenvalue	Cum. Var	F	P>F	Sig
RDA1	1	0.108	0.69	62.35	0.001	***
RDA2	1	0.023	0.83	13.16	0.001	***
RDA3	1	0.011	0.90	6.51	0.001	***
RDA4	1	0.007	0.95	4.08	0.001	***
RDA5	1	0.003	0.97	1.63	0.111	
RDA6	1	0.002	0.98	1.10	0.308	
RDA7	1	0.001	0.99	0.57	0.798	
RDA8	1	0.001	0.99	0.43	0.898	
RDA9	1	0	0.99	0.29	0.981	
RDA10	1	0	0.99	0.17	1.000	
RDA11	1	0	1.00	0.11	1.000	
RDA12	1	0	1.00	0.10	1.000	
RDA13	1	0	1.00	0.09	1.000	
RDA14	1	0	1.00	0.07	1.000	
RDA15	1	0	1.00	0.05	1.000	
RDA16	1	0	1.00	0.04	1.000	
Residual	263	0.457				

Table 30: Loadings of environmental variables for the first four RDA axes, which represent 95% of the constrained variation, done using the Hellinger transformed bacterial community matrix.

	RDA1	RDA2	RDA3	RDA4
Na	-0.09	0.07	-0.22	-0.13
Mg	0.48	-0.02	-0.33	0.17
Ca	0.41	-0.12	-0.32	0.35
K	-0.54	0.15	0.07	-0.42
Cl	0.06	-0.17	0.37	0.24
SO_4	-0.04	0.21	-0.20	-0.25
NH_4	0.14	-0.18	0.01	-0.23
NO_x	-0.09	-0.48	-0.42	0.38
Pelletal Skeletal Sand	-0.21	0.35	0.63	0.33
Pelletal – Aragonite Mud	0.42	-0.01	-0.24	0.26
Aragonite Mud	0.07	0.01	0.04	-0.10
Pelletal Diatomaceous Sand	-0.22	-0.18	-0.21	0.11
Pelletal Sand	-0.43	0.12	-0.46	-0.26
Morphotype 2	0.69	-0.17	-0.12	0.51
Morphotype 3	-0.29	-0.81	0.03	-0.09
Morphotype 4	-0.41	0.10	-0.27	-0.02

The first four RDA axes explained significant proportions of variation, with the first two axes representing 83% of the variation (Table 30). Potassium and magnesium concentration, as well as the presence of pelletal-aragonite mud were both strongly associated with RDA1, and weakly associated with RDA2, making them excellent determinants of RDA1. The first RDA axis was also strongly separated pond belonging to morphotype 2 (Cygnet), with the other ponds. NO_x concentration was strongly associated with RDA2. Like the bacterial communities in the previous chapter, the ciliate communities could be described along RDA1, which represents samples with high magnesium and calcium concentration and low potassium concentration, to samples with low magnesium and calcium concentrations and high potassium concentration. RDA2 represents a gradient from low to high NO_x concentration.

5.3.3 Environmental vs. spatial determinants

MEM variables were calculated using the coordinates of each sample and Hellinger transformed community data. 41 MEM variables were calculated which did not explain any of variation in the ciliate communities (Table 31). The environmental variables were used in a forward selection procedure, reducing the environmental component from 11 to 5 variables, while still explaining the same amount of variation. The selected variables included pond morphotype, sediment type, Mg, Ca, SO_4^{-2}. This reduced model

explained 21.3% of the variation in ciliate community structure ($F_{9,270}$=9.37, P=0.005). The effect of pond identity was the best determinant, explaining 37.1% of the variation in community composition ($F_{7,272}$=24.55, P=0.005).

Variation partitioning of the ciliate community matrix with the reduced environmental data, MEM variables and pond identity revealed which factors were the best determinants driving the changes in community composition (Table 31). The entire set of environmental and spatial variables (MEM + pond identity) explained 42% of the variation found in the ciliate communities. The environmental differences between ponds explained an insignificant amount of variation (0.3%), however, when the effect of spatial patterns (MEM variables) were accounted for, the effect of the environmental variables increased to 25.6%. The greatest amount of variation was, however, explained by pond identity, when the MEM variables were accounted for (43.7%). Pond identity was also the best discriminant for ciliate community composition when the variation contributed by the other variables was accounted for, with 18.2% of variation in ciliate composition being explained by pond identity alone.

Table 31: Variation partitioning summary table done using partial RDAs. Env represents the 5 forward selected variables (pond morphotype, sediment type, as well as Mg, Ca and SO_4^{-2} concentrations). Pond represents pond identity, and MEM represents the spatial location of the samples.

	R^2_{adj} (%)	DF_{model}	$DF_{residual}$	F	P-Value
Marginal Effects					
Env	21.3	9	270	9.37	0.005
MEM	-2.0	41	238	0.86	0.89
Pond	37.1	7	272	24.55	0.005
Conditional Effects					
Env|Pond	0.3	6	266	1.22	0.16
Env|MEM	25.6	9	229	9.86	0.005
MEM|Pond	4.5	41	231	1.51	0.005
MEM|Env	2.3	41	229	1.20	0.025
Pond|Env	16.2	4	266	18.47	0.005
Pond|MEM	43.7	7	231	26.44	0.005
Env|(MEM+Pond)	0.2	6	225	1.11	0.32
MEM|(Env+Pond)	4.3	41	225	1.48	0.005
Pond|(MEM+Env)	18.2	4	225	18.93	0.005
Residual	58.2				

5.3.4 Co-occurrence patterns and neutral processes

The three ponds, Whistler's Pond, Jana's Vent and Cygnet Pond, showed patterns in species co-occurrence that suggest the composition of the communities are non-random as they all had observed C-scores greater than the expected C-scores, and therefore have greater segregation than a random model (Table 32). In other words, there was a trend for OTUs not to coexist with each other in these ponds. The other ponds however, were not significantly different to what could be expected to occur randomly and therefore, there is no evidence of co-occurrence and/or segregation in the ciliate communities of those ponds.

The neutral diversity parameter (θ) was greatest in Pete's Vent, whilst Donut Pond and Cygnet Pond were also high (Table 33). The lowest levels of neutral diversity were observed in Annie's Pond, Pete's Pond and Harjie's Pond. The immigration parameters (m) within the ponds were very low, with Jana's Vent having the highest value. Cygnet Pond and Pete's Pond also had relatively high values for the immigration parameter, whilst Harjie's and Donut Pond had the lowest values. The observed levels of diversity were all lower than the expected values (Table 34) estimated using simulated datasets generated by the neutral model of Maaß *et al.* (2014).

Table 32: Results of species co-occurrence models using 10,000 permutations. The C-score is an index which summarises the co-occurrence pattern in a community matrix. Effect size is a measure of the difference between the observed and expected C-scores as a function of standard deviation.

	C-score Obs	C-score Exp	C-score Var	Effect Size	P-value
Pete's Vent	15.54	15.19	0.08	1.24	0.12
Donut Pond	7.78	7.83	0.18	-0.01	0.55
Annie's Pond	1.53	1.66	0.03	-0.73	0.29
Pete's Pond	4.91	5.30	0.11	-1.17	0.09
Whistler's Pond	14.64	13.12	0.66	1.87	0.05
Harjie's Pond	6.80	7.38	0.39	-0.93	0.18
Jana's Vent	8.43	7.30	0.22	2.39	0.03
Cygnet Pond	13.52	10.93	0.19	5.91	<0.0001
All ponds	1.44	1.427	0.002	0.35	0.31

Table 33:Estimates of neutral diversity (θ) and immigration (m ± standard error) for each pond using the multiple sampling formula given by Etienne (2007, 2009). The number of samples used to calculate each parameter is given (n).

	θ	m (±SE)	*n*
Pete's Vent	9.91	0.023 ± 0.003	32
Donut Pond	6.15	0.011 ± 0.003	32
Annie's Pond	2.19	0.024 ± 0.014	19
Pete's Pond	2.46	0.075 ± 0.029	39
Whistler's Pond	2.93	0.019 ± 0.003	52
Harjie's Pond	2.49	0.007 ± 0.002	52
Jana's Vent	3.23	0.096 ± 0.028	54
Cygnet Pond	4.82	0.077 ± 0.030	49
All ponds	2.84	0.045 ± 0.051	8

Table 34: The observed β-diversity compared to expected β-diversity values obtained by simulating neutral communities. Communities were simulated using estimated parameters of neutral diversity (θ), immigration (*m*) and ciliate OTU abundance (*J*) from 4999 simulations. Observed and expected β–diversity was calculated using Hellinger transformed OTU by site abundance matrices following Maaß *et al.* (2014). The effect size indicates the direction and strength of the difference between observed and simulated β-diversities.

	Observed β-diversity	Simulated β-diversity	Effect Size	P-value
Pete's Vent	0.58	0.91	-6.78	0.001
Donut	0.59	0.87	-3.87	0.008
Annie's	0.27	0.72	-2.48	0.028
Pete's	0.29	0.72	-2.84	0.020
Whistler's	0.34	0.91	-14.16	<0.0001
Harjie's	0.30	0.74	-2.99	0.017
Jana's	0.26	0.77	-4.09	0.005
Cygnet	0.53	0.83	-3.54	0.010
All ponds	0.36	0.75	-2.13	0.045

5.4 Discussion

This is the first study to simultaneously investigate the changes in ciliate alpha and beta diversity at two levels of scale using both deterministic and demographic processes. The traditional cosmopolitan view of protist biogeography is that one might expect to find similar communities in similar habitats (Finlay *et al.* 1998, Green and Bohannan 2006, Lepère *et al.* 2013). In this study, RDA analyses demonstrated strong differences in ciliate community structure in biofilm among ponds, even when the effects of spatial patterns and environmental variation were accounted for and thus challenges the traditional cosmopolitan view. Although there was an association between the OTU community and some environmental variables (particularly cation concentration and sediment type), this relationship was not important when the other spatial factors were accounted for. It was also found that null and neutral models were not good determinants of community structure. Unlike these results, other studies using bacteria (Caruso *et al.* 2011) and oribatid mites (Maaß *et al.* 2014) have shown that community

structure can be determined by dispersal and/or immigration limitation and large-scale environmental gradients.

5.4.1 Environmental determinants

The ciliate communities were mostly related to changes in cation concentrations, although there were also significant effects of nitrite/nitrate concentration and sediment type. However, once the effect of pond identity was accounted for, they were no longer significant determinants of community structure. Furthermore, the Mantel correlograms showed little evidence of distance-decay relationships in the ciliate communities occurring with-in the ponds, a pattern one would expect if communities were responding to environmental gradients within the ponds. RDA analyses suggested that the large amount of unconstrained variation was still attributable to pond groups, which could indicate an effect of unaccounted environmental factors operating at the between pond scale.

The cation concentrations matched the patterns in ciliate community structure more closely than any of the other variables. Salinity has also been found to be the best determinant of ciliate community structure in the previous study by Lei *et al.* (2014), who found distinct assemblages in freshwater, brackish and marine/brackish waters. Similar changes in diversity and structure of ciliate communities was found along a 237 km longitudinal gradient in the Gulf of Gabes (Elloumi *et al.* 2015). This study was different in that it was confined to salinities matching waters ranging from seawater concentrations to one and a half times the concentration of seawater. There were differences in cation concentrations between the ponds, with Cygnet Pond having the highest concentrations of sodium and Whistler's Pond and Harjie's Pond having the lowest concentration of sodium ions.

Nitrate/nitrite concentration was also a significant environmental determinant of ciliate community structure. However, nitrite/nitrate was not correlated with ammonium concentration and only weakly correlated with phosphate concentration, therefore samples with high nitrite/nitrate concentration do not necessarily mean high nutrient concentration. Nonetheless, the ponds were different in their nitrite/nitrate concentration (Chapter 3). Buosi *et al.* (2011) and Domènech *et al.* (2006) also observed an increase in ciliate richness and changes in ciliate composition of communities fertilised with

nitrate and phosphate, although Hingsamer *et al.* (2013) found species specific responses to nutrient enrichment and no overall changes in species richness. Andrushchyshyn *et al.* (2009) have shown ciliate communities to be associated with other physiochemical factors, such as water temperature, dissolved oxygen levels, as well as concentrations of ammonia and nitrates. It is likely though, that the effect of nutrients is more complex than experimental studies suggest because of the additional response of the grazers in the macrozoobenthic community, and subsequent alteration of the 'top-down' effects (Wickham *et al.* 2004).

Ciliate communities were also found to be related to the presence of the sandy sediments, pelletal sand, pelletal-skeletal sand and pelletal-diatomaceous sand. The other sediment types, which consisted of aragonite mud, were found not to be important in describing changes in the ciliate communities. Studies on ciliates and their association with different sediment types show that many factors, including temperature and salinity, but also silt and clay concentration and particle grain size can affect the distribution of ciliates (Finlay *et al.* 1997, Burgess 2001, Gücker and Fischer 2003, Reiss and Schmid-Araya 2008, Lei *et al.* 2014, Plebani *et al.* 2015). Sediment structure effects pore volume and particle surface area, which are the main features of the interstitial habitat for microbes (Gücker and Fischer 2003). Generally, the highest abundances of ciliates have been found in sandy sediments that contained small amounts of fine sediment, which could represent a trade-off between mean pore volume and particle surface area (Gücker and Fischer 2003). For example, in an intertidal area, ciliates were generally more abundant in medium and fine sand areas than in silty and muddy areas (Patterson *et al.* 1989).

One source of unaccounted variation is the hydrological regimes of the ponds. Data presented in Chapter 2 shows the differences in the hydrodynamic characteristics of the ponds. In particular, it was found that the different morphometric properties of the systems produced varying water residency times. Residency times were estimated to range from 4 days to over a year, assuming a balance between water inflow from the vents and outflow via channels and/or evapotranspiration. Some of the ponds were dominated by direct channel outflow, whilst outflow in others was dominated by evapotranspiration. For example, evaporative loss from Pete's Vent, Annie's Pond, Whistler's Pond and Jana's Vent was less than 5% of total water loss, whilst it

represented 50% of water loss in Cygnet and 100% water loss in Donut, Harjie's and Pete's Ponds. It is likely that the physical characteristics of each pond is an important structuring force on both abiotic variables, such as water mixing, turbidity, light penetration, salinity and nutrients, and biotic variables, such as immigration and top-down effects.

The few abundant OTUs detected in the ciliate meta-community are probably generalist species which are able to function given the available habitat characteristics found within the ponds. The single dominant OTU (found at 293 BP) accounted for nearly 50% of the total fragments detected and would have reduced the statistical ability to detect differences in pond communities. Although 75% of taxa represent less than 1% of the detected fragments, it was these rare taxa which were important in characterising the pond communities. Ciliate communities in saline lagoons and ponds have also been found to be dominated by generalist species (Elloumi *et al.* 2006, Lei *et al.* 2014), but are more abundant and diverse than other extreme environments, such as deep hypersaline anoxic basins (Orsi *et al.* 2012, Stock *et al.* 2013). A rare taxonomic component is frequently detected in ciliate communities (Weisse 2014) and may play an important role as a 'seed bank' in communities if physical, chemical and/or biological variables in their habitats change. Temporal and spatial changes in the rare component of microbial communities remains a challenge for contemporary studies in microbial ecology, but disentangling the role of neutral and deterministic processes on the rare component of these communities, especially using a meta-community perspective, will enhance our understanding on microbial community dynamics and functioning.

5.4.2 Demographic determinants

Similar to the previous chapter on the bacterial communities, neutral and null models were tested at two levels of spatial scale, the pond scale and the with-in pond scale. At the between-pond level, ciliate communities seemed to co-occur randomly, with neither evidence for significant segregation nor aggregation. However, at the within pond scale, some ciliate communities showed significant segregation (OTUs co-occurred less frequently than randomly predicted) in Whistler's Pond, Jana's Vent and Cygnet Pond. Furthermore, this study shows that the ciliate communities in all the ponds have diverged relative to a reference point provided by the general neutral model of Etienne

(2009), with the observed beta diversities lower than what is expected under neutral dynamics. These results, however, assume that expected beta diversity depends solely on the basic processes postulated by neutral theory (dispersal and stochastic fluctuations). Furthermore, there is also evidence suggesting that functional groups within microbial communities are structured differently by deterministic drivers and stochastic processes (Caruso *et al.* 2011). In real communities, it is therefore likely a complex interaction between deterministic processes and stochastic drivers, such as historical events, act on populations at different spatial, temporal and trophic scales.

Studies investigating ciliate biogeography at fine, local scales are rare. Lepère *et al.* (2013) have demonstrated a distance-decay relationship in French lake protist populations over scales ranging from 133 to 400 km. Similarly, Rengefors *et al.* (2012) found limited gene flow between Antarctic lakes separated by distances ranging from 1 and 9 km, and was probably the first to show that dispersal limitation, at scales less than 10 km can be important in structuring ciliate communities. For a ciliate, the ponds at Lake MacLeod are essentially 'islands' isolated by a 'sea' of uninhabitable habitat of dry lake bed inundated only by extremes of flooding or temporary wind-driven surface flows. It is therefore likely that dispersal between ponds is somewhat limited and probably an important factor in shaping the ciliate communities. This conclusion is in agreement with that of Wey *et al.* (2009), who have shown that mature ciliate communities are controlled independently of immigration processes, and influenced more by local resources and historical events.

The floods of 2010 may represent a historical event that caused homogenisation of the pond communities through enhanced turbulence in ponds and spatial connectivity for a short period. Microbial communities could be expected to rapidly adapt to the post-disturbance effects of changes in abiotic (nutrients and salinity) and biotic (new predators) stressors. However, re-establishment of the community is thought to be enhanced by a resting propagule bank found within each pond for aquatic organisms (De Meester *et al.* 2002). Such a process would maintain each pond on s eparate structuring 'trajectories' due to the propagule bank providing a powerful buffer against new immigrants (De Meester *et al.* 2002). Indeed, many protists, including ciliates, will pass through a benthic resting stage during their life cycles (Kremp 2001, Müller *et al.* 2002), allowing them to overcome periods of unfavourable conditions (Lei *et al.* 2014).

This hypothesis, called the Monopolization Hypothesis (De Meester *et al.* 2002), has been postulated by Lepère *et al.* (2013) and Rengefors *et al.* (2012) to be a significant force that shapes the structure of lake protist communities, but has never been tested previously.

5.4.3 Conclusions

This study presents a case where environmental and demographic processes seem to be weak structuring determinants of the ciliate communities. There are relatively few studies looking at diversity changes and community structure of protists, and fewer targeting ciliates specifically (Reiss and Schmid-Araya 2008, Dopheide *et al.* 2009). This is the first study to partition the spatial and environmental effects separately, and use demographic models to explain the structure of these important communities. Ciliates provide an important trophic link in aquatic ecosystems, and understanding community diversity is thus important in understanding ecosystem processes. Even though salinity, nitrate and sediment type all influenced the communities, it was the conditional effect of pond identity itself that explained the most amount of variation. These results suggest that nearby habitats can have different communities, and these differences in community structure occur largely independently of environmental and demographic processes.

CHAPTER 6. GENERAL DISCUSSION

The aim of this thesis was to compare the role of species-sorting and neutral processes that influence the distributions of microbial prokaryotes and eukaryotes. Two main questions addressed by this study were to test: if similar biogeographic processes important in determining the distribution of plants and animals are important to microbial communities, and if there are differences between bacterial (prokaryotes) and ciliate (eukaryotes) distributional and biogeographic patterns. This research contributes towards our understanding of species-sorting and neutral processes in microbial ecology, as well as taxon-area and distance-decay relationships. The data herein demonstrates that distance-decay and taxa-area relationships can exist for microbial communities, however the relationships may be different for prokaryotes (bacteria) and micro-eukaryotes (ciliates; Table 35). It was also demonstrated that environmental heterogeneity was a poor determinant of community composition, and that each pond habitat had unique communities that were not determined by the environment and spatial processes. The work in this thesis contributes significantly to the field of microbial ecology, by understanding the roles of environmental and demographic processes involved in structuring these ecologically important communities.

Table 35: Summary of hypothesis testing in Chapters 4 (bacteria) and Chapter 5 (ciliates).

Hypothesis	Bacteria	Ciliates
Larger ponds have greater richness	False	False
Communities differ between ponds	True	True
Community variation is explained by environmental and spatial variation	<10% can be explained by spatial variation. <1% explained by environment	<5% can be explained by spatial variation. Environment not significant
Co-occurrence patterns are random	True	Pond dependent
Immigration is a good determinant of beta diversity	False	False

The ponds of Lake MacLeod were chosen to empirically investigate these questions because of the ideal spatial arrangement of the ponds and their consanguinity. Moreover, using a natural system allows for the testing of processes which occur at scales much larger than can be achieved using laboratory experiments. However, when using natural systems, one must understand the variation within and between the habitats when attempting to understand the role of possible environmental determinants. Chapter 2 does this, by providing information about the physical differences of the ponds in terms of morphometry and hydrology, whilst Chapter 3 is an investigation into the fine-scale variation in water chemistry parameters likely to either be determinants, or proxies for other sources of variation, that alter the composition of microbial communities. Chapter 4 and 5 provide data that contribute to our knowledge of microbial ecology, by showing the biogeographic processes that are important in structuring these microbial communities using the data collected in Chapters 2 and 3 as explanatory variables. In this final chapter, a synthesis of the main findings of this thesis is presented, concentrating in particular on the roles of species-sorting and neutral processes, as well as the differences and similarities found in the distributions of the bacteria and ciliate communities.

6.1 Biogeographic patterns

There were 171 bacterial and 30 ciliate OTUs detected in this study. In both the bacterial and ciliate communities, Cygnet Pond (the largest pond in this study) was the most OTU rich pond. Similarly, Pete's Vent showed the highest levels of diversity among the ponds for both the bacterial and ciliate communities. Patterns of similarities between pond communities seemed to be different for bacteria and ciliates, although there was a very weak positive relationship between bacterial richness and ciliate richness (R^2_{adj}=0.10; $F_{1,324}$=36.76, P<0.0001; Figure 53). The bacterial communities in Cygnet Pond were different to most of the other ponds, although they were not significantly different to those found in Pete's Vent, Donut Pond and Jana's Vent. For the ciliate communities, Jana's Vent and Harjie's Pond were different in composition to the other ponds. Nonetheless, pond identity was the most important determinant of community composition for both communities, while environmental and spatial variation seemed to be unimportant.

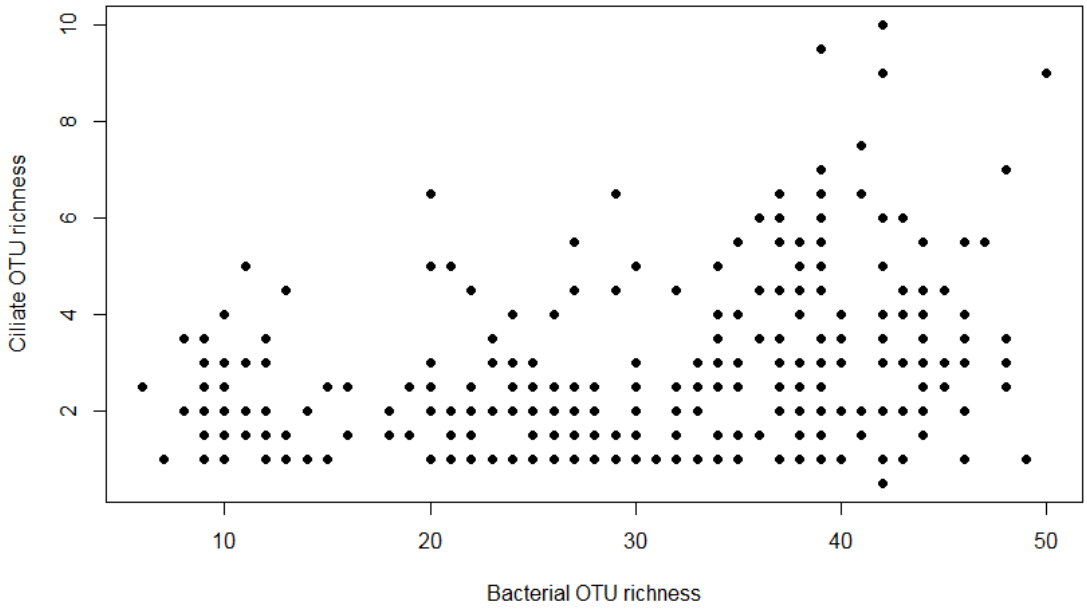


Figure 53: Scatterplot of bacterial and ciliate OTU richness. There was a significant linear relationship ($F_{1,324}$=36.76, P<0.0001). R^2_{adj}=0.10.

Distance-decay and taxa-area relationships are both predicted to occur if species-sorting and neutral processes are operating, but for different reasons. Although no t axa-area relationships were detected, this is the first study to describe distance-decay relationships for bacteria and ciliates in a single study using a contiguous environment. There was no evidence for the taxa-area relationship for either bacteria nor ciliates as there was no increase in OTU richness with increased pond size. This result is particularly interesting because taxa-area relationships are thought to be one of the few general patterns in ecology (Lawton 1999), and a link between the ecology of macro- and micro-organisms (Horner-Devine *et al.* 2004, Bell *et al.* 2005a, Reche *et al.* 2005, Woodcock *et al.* 2006). This thesis provides a case where ecosystem size does not have an effect on bacterial and ciliate richness, and therefore disagrees with the results of previous studies where evidence of taxa-area relationships has been found (Bell *et al.* 2005a, Reche *et al.* 2005, van der Gast *et al.* 2005).

Lee (2014) and Logue *et al.* (2012) also did not observe taxa-area relationships in their aquatic bacterial communities using high-throughput sequencing. Both studies found that environmental variables were more important than habitat size in determining

richness. The absence of a taxa-area relationship may be caused by the vast population sizes and high dispersal rates that are typical of microbes, as well as their low extinction rates (Fenchel 1993, Fenchel and Finlay 2003), or simply because niche diversity was low within the ponds (Angermeier and Schlosser 1989). In this study, niche diversity may have been reduced because the ponds are all shallow, non-stratified and exposed to high amounts of sun light, despite being different sizes. Because T-RFLP has a rather coarse taxonomic resolution, the ability to detect taxa-area and distance-decay relationships may have been compromised. Furthermore, the uneven communities, being dominated by relatively few taxa, would also reduce the ability to detect either relationship. Even though the pond communities are dominated by few abundant taxa, using sequencing methods to measure community structure would enhance the ability to resolve many of the rare taxa, and may provide fruitful in establishing whether taxa-area and distance-decay relationships are found for ciliate communities. Additionally, because microbes can be expected to function at rather small scales, the failure to detect a taxa-area relationship could be due to the habitats being large enough that the scales are not relevant to the microbial processes and functioning. (Logue *et al.* 2012). Even though the smallest pond was 6 orders of magnitude smaller than the largest, these scales may still be too large to detect taxa-area relationships in microbes. This study, as well as the results of Lee (2014) and Logue *et al.* (2012), cast doubt on the presence of microbial taxa-area relationships.

Bacteria and ciliates differed in their distance-decay relationships. This suggests that fundamentally different distributional patterns between bacteria and ciliates occur, where the bacterial communities show gradual changes in community composition, perhaps in response to environmental gradients, and the ciliate taxa have 'patchy' distributions, where the patterns in composition do not change predictably with distance. The differences in the distance-decay patterns described for bacteria and ciliates should, using the theories of biogeography, be explained by either differences dispersal and/or environmental specificity (Barreto *et al.* 2014a). The patterns observed in this study could therefore either be due to ciliates having greater dispersal rates than bacteria and/or bacterial communities being governed by species-sorting processes more than the ciliate communities. Because the distances involved in this study are unlikely to be great enough to significantly limit the dispersal of bacteria, it is more likely that

deterministic processes play a stronger role in structuring the communities than in ciliate communities. High redundancy is a feature of bacterial communities (Dopheide *et al.* 2015) which may permit the establishment, and subsequent exclusion, of different taxa within a niche. The high levels of OTU segregation found in the communities using the co-occurrence models are evidence of high redundancy in the bacterial communities. The establishment of different taxa in a given niche within the ponds may be caused either by stochastic events, or because of their slightly different environmental conditions. Nonetheless, the ciliate communities are represented by more abundant, generalist taxa, which, because of the lower amount of redundancy in community composition, and their ability to exist in a wider range of environmental conditions, showed lower levels of community segregation.

Although the ubiquity of ciliates has been suggested to be caused by their vast capacity to disperse across large distances (Finlay and Clarke 1999), it is unlikely that high dispersal rates are different, or in fact greater than, those of bacterial taxa (Green and Bohannan 2006), especially within a set of contiguous habitats and at the scales relevant to this study. Species-sorting processes have been shown to cause distance-decay relationships in microbes as the diversity of habitats tend to increase with increasing area (Hanson *et al.* 2012, Nemergut *et al.* 2013).

Microbial studies using rock pools (Langenheder and Ragnarsson 2007) and alpine tarns (Lear *et al.* 2014, Lee 2014) have reported variation which can be accounted for by environmental variation and geographic location on bacterial communities. For example, bacterial communities in rock pools had up to 12% of variation explained by environmental variables and up to 10% of variation explained by spatial variables (Langenheder and Ragnarsson 2007). In the alpine tarns, up to 38% variation in bacterial communities could be explained by spatial variation and up to 32% by spatially structured environmental variation (Lear *et al.* 2014). These results suggest that contemporary environmental conditions are important in determining bacterial community composition (Horner-Devine *et al.* 2004, Beisner *et al.* 2006), as well as spatial processes. Conversely, the data presented in this thesis suggested that the measured environmental variables do not explain any of the bacterial and ciliate community composition, and spatial variation only accounts for a small amount of the variation.

It is difficult to explain the distance-decay patterns found in this study using species-sorting processes because there were neither clear environmental gradients within the ponds, nor was there a significant effect of environmental variation when geographic distance and pond identity were accounted for. Spatial variation did however explain some of the community variation for both bacteria and ciliate communities (approx. 8% and 4%, respectively; Figure 54). In this study though, pond identity, rather than geographic location and environmental variables, was the strongest determinant for both bacterial and ciliate communities. This suggests that either an unaccounted determinant, which is spatially independent but varies at the pond level, and/or a neutral process which operates independently of immigration processes, is driving the differences in these communities. It therefore remains possible that species-sorting mechanisms are important for these communities (Kent *et al.* 2007, Mosher and Findlay 2011, Lindström and Langenheder 2012, Lee *et al.* 2013). For example, intrinsic lake variables important in defining bacterial and ciliate community composition may not have been included in this study. Kent *et al.* (2007) suggest that mechanisms such as competition and nutrient cycling are important for describing among lake variation in aquatic communities. Nonetheless, this study found that species-sorting processes are important in structuring the bacterial and ciliate communities in these ponds when considering ionic and nutrient dynamics, variables which would otherwise be expected to explain microbial community variation in salt lakes (Casamayor *et al.* 2013).

The ponds have unique bacterial and ciliate communities inhabiting them, and perhaps the distinction between the communities is more pronounced with the bacterial communities than the ciliates. For example, the within pond similarity was greater for bacteria (R=0.839) than it was for ciliates (R=0.460).These results show that the ponds represent habitats in which immigrants do not contribute significantly to the community composition, because the proximity of the ponds to each other appeared to have little influence on the microbial taxa found there. Furthermore, the inability of the neutral models to predict β-diversity of the ponds, suggests that immigration and connectivity are also unimportant structuring mechanisms of these microbial communities. Instead, it is likely that the community structure of the microbes diverge because of a neutral process that is driven by stochasticity instead of immigration.

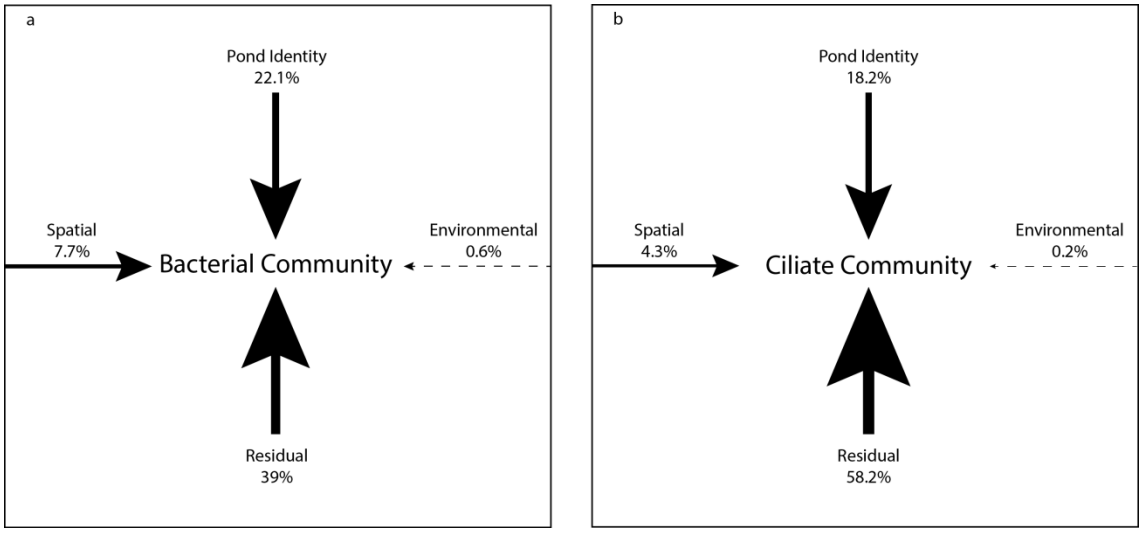


Figure 54: Conceptual models comparing the amount of variation explained by the spatial variables, pond identity and the environmental variables for the bacterial community (a) and the ciliate communities (b). The percentage of variation explained by each of these factors were calculated from the variation partitioning method in the previous two chapters. The size of the arrows are proportional to the square root of the variation which they explain, therefore the larger the arrow, the more important the factor. Dashed lines represent non-significant relationships. Residual variation explains variation which could not be explained by the models.

Ecological drift in microbial communities has been shown to cause divergence in community composition as taxa are lost by chance, or become common from exposure to novel conditions (Hubbell 2001, Lankau *et al.* 2012, Nemergut *et al.* 2013). This is a neutral process that explains community structure based on changes in the relative abundances of organisms due to stochastic events. It is analogous to theories of genetic drift (Masel 2011), which predicts the abundances and distribution of genes among populations. Specifically, the theory predicts that species in low abundance in a community are more likely to proceed towards local extinction. This is important when considering microbial communities because often most of the taxa are found at low relative abundances (Pedrós-Alió 2006, Nemergut *et al.* 2013).

The effect of ecological drift can be the dominant process governing composition of ecological communities (Clark and McLachlan 2003), and can produce community patterns which are difficult to interpret (Volkov *et al.* 2007). Ecological drift is driven by three processes that simultaneously operate in metacommunities in the absence of environmental variation (Ruokolainen *et al.* 2009). Firstly, population densities change (drift) due to demographic stochasticity and interspecific competition. Secondly, species composition in local communities drift due to local extinctions and dispersal limitation. And thirdly, the metacommunity species pool can experience drift as individual species

become extinct across all local communities (global extinction). These three mechanisms operate at different strengths based on the importance of species-sorting and neutral processes (Ruokolainen *et al.* 2009).

Testing for ecological drift is inherently difficult. Firstly, pure ecological drift is unlikely to occur in real communities because there will always be some effect of deterministic processes. Secondly, unexplained variance cannot automatically be attributed to drift because much of that variance may be due to the effect of unaccounted factors. For these reasons, the plausibility of ecological drift occurring in many ecological studies is often neglected (Vellend 2010). How to interpret stochasticity in ecological models, and methods to model elements that are unknown, present future challenges for neutral theory and ecology as a whole (Rosendell *et al.* 2012).

Ecological drift is introduced here only as a possible explanation for the patterns found in the bacterial and ciliate communities in this study. The fact that there remained a large amount of unexplained variation in the models after environmental and spatial patterns were accounted for, as well as the significant role that pond identity has on community composition, a stochastic process, such as ecological drift, could be an important factor driving the patterns found here. However, it remains impossible to detect the unknown in this case, and the role that ecological drift has on the microbial communities in Lake MacLeod will remain as speculation.

6.2 Where has time gone?

Temporal variation in communities is an important factor in the assembly of microbial communities which has largely been ignored in this thesis (Newton *et al.* 2011). In this study, temporal variation was controlled for by collecting all the samples at the same time of day, during the same week and season, thus allowing for a 'snap shot' of the microbial communities and environmental conditions of the Northern Ponds at a single time point. By removing sources of temporal variation, and focusing on assembly processes within a single period, patterns of dissimilarity and relationships with environmental and spatial factors can be determined (Horner-Devine *et al.* 2004, Bell *et al.* 2005, Martiny *et al.* 2011, Lear *et al.* 2014). Seasonal and inter-annual temporal

variation undoubtedly affects microbial communities (Jones *et al.* 2012, Lee 2014), particularly the Northern Ponds communities (Huggett *et al.* unpublished). In their analysis of the spatial and temporal variation of bacterioplankton in the ponds of Lake MacLeod (the same ponds used in this study), Huggett *et al.* (unpublished) provide evidence of strong seasonal shifts in the microbial communities, but importantly show that the spatial patterns and separation of pond communities remain. Those results reinforce the data presented in this thesis by suggesting that the spatial patterns found for the bacterial and ciliate communities are likely to persist throughout time despite seasonal changes. Including a temporal component to this study would have reduced the number of samples that could have been used for the spatial analysis, because of budget, time and logistical reasons, and thus limited the power of the variation partitioning and spatial analyses. Nonetheless, temporal changes in microbial communities need to be considered when interpreting structuring processes of these communities. This section will present a brief overview of research on the temporal variation of aquatic microbial communities.

Typically, bacterial communities become increasingly dissimilar with time (Wells *et al.* 2011, Cabrol *et al.* 2012, Portillo *et al.* 2012). For example, in his thesis, Lee (2014) demonstrates changes in microcosm communities over 82 weeks. These microcosm communities displayed strong temporal patterns where communities sampled in similar weeks were not only more similar in composition, but also diverged from one another over time. Jones *et al.* (2012) suggest that spatial variation and temporal variation in microbial communities are quite comparable, with the average community similarity across multiple temporal scales falling between average intra- and inter- lake community similarity. When spatial and temporal beta-diversities were calibrated to each other, it was found that communities separated by a single meter were comparably similar to those separated by a day. It can therefore be concluded that similar ecological processes drive community assembly over these spatial and temporal scales (Soininen 2010, Shade *et al.* 2013), and further iterates the importance of temporal variation.

It is likely that intrinsic factors, such as body size and dispersal rate, and extrinsic factors, such as ecosystem size and isolation, are drivers of bacterial turnover in both space and time (Soininen 2010, Jones *et al.* 2012). This may be because changes in environmental conditions at the metre scale are similar to the environmental changes

during the course of a day. Alternatively, the fast generation times (on the order of days) closely corresponds to the dispersal distances. Incorporating temporal sampling into ecological studies is difficult (Redford and Fierer 2009, Lauber *et al.* 2013), especially when considering fine-scale spatial variation has only recently been detected in microbial communities (Lear *et al.* 2014).

6.3 Limitations and future directions of study

One of the limitations of this project was the use of ARISA and T-RFLP, and the coarse taxonomic resolutions these methods achieve. Therefore, the rare taxa, which contribute greatly to microbial diversity (Reid and Buckley 2011, Dunthorn *et al.* 2014, Weisse 2014), were not accounted for. Also, the identities of the taxa detected remain unknown when using ARISA and T-RFLP. These problems can largely be overcome by using 454-pyrosequencing techniques (Lee *et al.* 2012). Although more expensive than fragment analyses, such as ARISA and T-RFLP, pyrosequencing will improve future estimations of the distance-decay and taxa-area relationships and the understanding of the assembly processes occurring in bacteria and ciliate communities. Future studies should also target other gene regions which discriminate functional and trophic guilds in conjunction with the 16S and 18S rRNA genes (Burke *et al.* 2011), which are good for taxonomic discriminations. This would allow for investigations into the biogeographic patterns of organisms representing different functional and trophic guilds (Caruso *et al.* 2011, Lear *et al.* 2014). Such data would enable for biogeographic patterns to be distinguished between functional and trophic groups within these diverse communities. It would also allow for a thorough investigation into the role of species-sorting processes as the composition of functional guilds can be compared across sites (Burke *et al.* 2011, Dopheide *et al.* 2015).

By using molecular methods with higher OTU resolution, and having the ability to separate function guilds within microbial communities, the high amount of unaccounted variation, which is typical of these studies, can be reduced and thus allow for the building of more representative models. This would also give greater power to discern the role of neutrality and species-sorting mechanisms, allowing ecologists to begin identifying the role of ecological drift in structuring microbial communities.

6.4 Conclusion

The objective of this thesis was to understand the role that environmental and demographic determinants have on the bacteria and ciliate communities found in Lake MacLeod. By understanding the variation in the physical and chemical environments found among the ponds, and the variation in bacteria and ciliate communities, it was established that the communities are determined primarily by the pond in which they originate from, despite the environmental similarities and the geographic locations of the ponds. These results suggest that ecological drift processes, where isolation has enabled each bacterial and ciliate pond community to diverge in composition, is an important mechanism in regulating the diversity of the ponds and the lake basin itself.

New and exciting molecular methods are allowing for ecologists to affordably generate vast datasets to test concepts which are central to our understanding of the diversity on Earth. Previously, the microbial world has been treated as an '*ecological black box*' without any discernible biogeographic patterns. Today this is not the case, and this study adds to a growing body of literature that suggests that microbial communities are structured within our environment in a similar manner to plants and animals. It is not only important to pursue an understanding of the biogeography and ecology of the most diverse, rich, abundant and functionally important organisms on our planet to test theory, but it is also important simply for the pursuit of intellectual inquiry.

REFERENCES

Abdelrady, A. R. 2013. Evaporation over fresh and saline water using SEBS. University of Twente.

Adl, S. M., A. G. B. Simpson, M. A. Farmer, R. A. Andersen, O. R. Anderson, J. R. Barta, S. S. Bowser, G. Brugerolle, R. A. Fensome, S. Fredericq, T. Y. James, S. Karpov, P. Kugrens, J. Krug, C. E. Lane, L. A. Lewis, J. Lodge, D. H. Lynn, D. G. Mann, R. M. McCourt, L. Mendoza, Ø. Moestrup, S. E. Mozley-Strandridge, T. A. Nerad, C. A. Shearer, A. Y. Smirnov, F. W. Spiegel, and M. F. J. R. Taylor. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. Journal of Eukaryotic Microbiology **52**:399-451.

Alongi, D. 1994. The role of bacteria in nutrient recycling in tropical mangrove and other coastal benthic ecosystems. Hydrobiologia **285**:19-32.

Alonso, D., R. S. Etienne, and A. J. McKane. 2006. The merits of neutral theory. Trends in Ecology and Evolution **21**:451-457.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology **26**:32-46.

Andrushchyshyn, O. P., K. P. Wilson, and D. D. Williams. 2007. Ciliate communities in shallow groundwater: Seasonal and spatial characteristics. Freshwater Biology **52**:1745-1761.

Andrushchyshyn, O. P., K. P. Wilson, and D. D. Williams. 2009. Climate change-predicted shifts in the temperature regime of shallow groundwater produce rapid responses in ciliate communities. Global Change Biology **15**:2518-2538.

Angermeier, P. L. and I. J. Schlosser. 1989. Species-area relationships for stream fishes. Ecology **70**:1450-1462.

Astorga, A., J. Oksanen, M. Luoto, J. Soininen, R. Virtanen, and T. Muotka. 2012. Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules? Global Ecology and Biogeography **21**:365-375.

Australian Bureau of Meteorology. 2015. Climate statistics for Australian locations. http://www.bom.gov.au/climate/averages/tables/cw_006011_All.shtml.

Baas-Becking, L. G. M. 1934. Geobiologie of inleiding tot de milieukunde. van Stockum, W. P.

Zoon, N. V., Den Haag.

Baranyi, C., T. Hein, C. Holarek, S. Keckeis, and F. Schiemer. 2002. Zooplankton biomass and community structure in a Danube River floodplain system: Effects of hydrology. Freshwater Biology **47**:473-482.

Barberán, A. and E. O. Casamayor. 2011. Euxinic freshwater hypolimnia promote bacterial endemicity in continental areas. Microbial Ecology **61**:465-472.

Barko, J. W. and W. F. James. 1998. Effects of submerged aquatic macrophytes on nutrient dynamics, sedimentation, and resuspension.*in* E. Jeppesen, M. Søndergaard, M. Søndergaard, and K. Christoffersen, editors. The structuring role of submerged macrophytes in lakes. Springer Science+Business Media, New York.

Barreto, D. P., R. Conrad, M. Klose, P. Clause, and A. Enrich-Prast. 2014a. Distance-decay and taxa-area relationships for bacteria, archaea and methanogenic archaea in a tropical lake sediment. PLoS One **9**:e110128.

Barreto, D. P., R. Conrad, M. Klose, P. Clause, and A. Enrich-Prast. 2014b. Distance-decay and taxa-area relationships for bacteria, archea and methanogenic archaea in a tropical lake sediment. PLoS One **9**:e110128.

Beisner, B. E., P. R. Peres-Neto, E. S. Lindström, A. Barnett, and M. L. Longhi. 2006. The role of environmental and spatial processes in structuring communities from bacteria to fish. Ecology **87**:2985-2991.

Bell, G. 2000. The distribution of abundance in neutral communities. The American Naturalist **155**:606-617.

Bell, G. 2001. Neutral macroecology. Science **293**:2413-2418.

Bell, T. 2010. Experimental tests of the bacterial distance-decay relationship. The ISME Journal **4**:1357-1365.

Bell, T., D. Ager, J. I. Song, J. A. Newman, I. P. Thompson, A. K. Lilley, and C. J. van der Gast. 2005a. Larger islands house more bacterial taxa. Science **308**:1884.

Bell, T., J. A. Newman, B. W. Silverman, S. L. Turner, and A. K. Lilley. 2005b. The contribution of species richness and composition to bacterial services. Nature **1157-1160**.

Bertzeletos, D., R. A. Davis, and P. Horwitz. 2012. Importance of Lake MacLeod, northwestern Australia, to shorebirds: a review and update. Journal of the Royal Society of Western Australia **95**:115-124.

Blackwood, C. B., D. Hudleston, D. R. Zak, and J. S. Buyer. 2007. Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. Applied and Environmental Microbiology **73**:5276-5283.

Boggs, D. A., G. S. Boggs, I. Eliot, and B. Knott. 2006. Regional patterns of salt lake morphology in the lower Yarra Yarra drainage system of Western Australia. Journal of Arid Environments **64**:97-115.

Boggs, D. A., G. S. Boggs, B. Knott, and I. Eliot. 2007. The hydrology and hydrochemistry of six small playas in the Yarra Yarra drainage system of Western Australia. Journal of the Royal Society of Western Australia **90**:15-32.

Boon, P. I., P. C. Pollard, and D. Ryder. 2014. Wetland Microbial Ecology and Biogeochemistry.*in* D. P. Batzer and R. R. Sharitz, editors. Ecology of Freshwater and Estuarine Wetlands. University of California Press, Oakland, Californai.

Borcard, D., F. Gillet, and P. Legendre. 2011. Numerical Ecology with R. Springer Science + Buisness Media, New York.

Borcard, D. and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling **153**:51-68.

Borcard, D. and P. Legendre. 2012. Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. Ecology **93**:1473-1481.

Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto. 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology **85**:1826-1832.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology **73**:1045-1055.

Brose, U., A. Ostling, K. Harrison, and N. D. Martinez. 2004. Unified spatial scaling of species and their trophic interactions. Nature **428**:167-171.

Brown, M. V., M. S. Schwalbach, I. Hewson, and J. A. Fuhrman. 2005. Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. Environmental Microbiology **7**:1466-1479.

Buosi, P. R. B., G. M. Pauleto, F. A. Lansac-Tôha, and L. F. M. Velho. 2011. Ciliate community associated with aquatic macrophyte roots: Effects of nutrient enrichment on the community composition and species richness. European Journal of Protistology **47**:86-102.

Burgess, R. 2001. An improved protocol for separating meiofauna from sediments using colloidal silica sols. Marine Ecology Progress Series **214**:161-165.

Burke, C., P. Steinberg, D. Rusch, S. Kjelleberg, and T. Thomas. 2011. Bacterial community assembly based on functional genes rather than species. Proceedings of the National Academy of Sciences **108**:14288-14293.

Cabrol, L., L. Malhautier, F. Poly, A.-S. Lepeuple, and J.-L. Fanlo. 2012. Bacterial dynamics in steady-state biofilters: beyond functional stability. FEMS Microbiology Ecology **79**:260-271.

Cadotte, M. W. 2006. Dispersal and species diversity: a meta-analysis. The American Naturalist **167**:913-924.

Cam, E., J. D. Nichols, J. E. Hines, J. R. Sauer, R. Alpizar-Jara, and C. H. Flather. 2002. Disentangling sampling and ecological explanations underlying species-area relationships. Ecology **83**:1118-1130.

Cameron-Caluori, H. 2014. A preliminary investigation of the potential effects of the invasive Mozambique tilapia Oreochromis mossambicus on the native fish assemblages of Lake MacLeod, Western Australia. Edith Cowan University.

Caron, D. A. 2009. Past President's Address: Protistan Biogeography: Why All The Fuss? Journal of Eukaryotic Microbiology **56**:105-112.

Caron, D. A., A. Z. Worden, P. D. Countway, E. Demir, and K. B. Heidelberg. 2009. Protists are microbes too: a perspective. The ISME Journal **3**:4-12.

Caruso, T., Y. Chan, D. C. Lacap, M. C. Y. Lau, C. P. McKay, and S. B. Pointing. 2011. Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. ISME J **5**:1406-1413.

Casamayor, E. O., X. Triadó-Margarit, and C. Castañeda. 2013. Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. FEMS Microbiology Ecology **85**:503-518.

Castañeda, C., J. Herrero, and J. A. Conesa. 2013. Distribution, morphology and habitats of saline wetlands: a case study from Monegros, Spain. Geologica Actta **11**:371-388.

Chase, J. M. 2007. Drought mediates the importance of stochastic community assembly. Proceedings of the National Academy of Sciences **104**:17430-17434.

Chave, J. and E. G. Leigh Jr. 2002. A spatially explicit neutral model of beta-diversity in tropical forests. Theoretical Population Biology **62**:153-168.

Cho, J.-C. and J. M. Tiedje. 2000. Biogeography and Degree of Endemicity of Fluorescent Pseudomonas Strains in Soil. Applied and Environmental Microbiology **66**:5448-5456.

Clark, J. and J. S. McLachlan. 2003. Stability of forest biodiversity. Nature **423**:635-638.

Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology **18**:117-143.

Cody, M. L. 1985. Towards a theory of continental species diversities. Pages 214-257 *in* M. L. Cody and J. M. Diamond, editors. Ecology and evolution of communties. Harvard University, Cambridge, Massachusetts.

Coggins, L., A. Ghadouani, and M. Ghisalberti. 2014. Bathymetry mapping using a GPS-sonar equipped remote control boat: Application in waste stabilisation ponds. European Geosciences Union General Assembly, Vienna, Austria.

Comte, J., E. S. Lindström, A. Eiler, and S. Langenheder. 2014. Can marine bacteria be recruited from freshwater sources and the air? The ISME Journal **8**:2423-2430.

Connor, E. F. and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. The American Naturalist **113**:791-833.

Cotner, J. B. and B. A. Biddanda. 2002. Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. Ecosystems **5**:105-121.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecological Letters **8**:1175-1182.

Crosbie, B. and P. Chow-Fraser. 1999. Percentage land use in the watershed determines the water and sediment quality of 22 marshes in the Great Lakes basin. Canadian Journal of Fisheries and Aquatic Sciences **56**:1781-1791.

Danovaro, R., G. M. Luna, A. Dell'Anno, and B. Pietrangeli. 2006. Comparison of two fingerprinting techniques, terminal restriction fragment length polymorphism and automated ribosomal intergenic spacer analysis, for determination of bacterial diversity in aquatic environments. Applied and Environmental Microbiology **72**:5982-5989.

De Bie, T., L. De Meester, L. Brendonck, K. Martens, B. Goddeeris, D. Ercken, H. Hampel, L. Denys, L. Vanhecke, K. Van der Gucht, J. Van Wichelen, W. Vyverman, and S. A. J. Declerck. 2012. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. Ecology Letters **15**:740-747.

De Deckker, P. 1983. Australian salt lakes: their history, chemistry, and biota - a review. Hydrobiologia **105**:231-244.

DeLong, J. R. and N. R. Pace. 2001. Environmental diversity of bacteria and archaea. Systematic Biology **50**:470-478.

De Meester, L., A. Gómez, B. Okamura, and K. Schwenk. 2002. The Monopolization Hypothesis and the dispersal–gene flow paradox in aquatic organisms. Acta Oecologica **23**:121-135.

De Wit, R. and T. Bouvier. 2006. '*Everything is everywhere*, but, *the environment selects*': what did Baas Becking and Beijerinck really say? Environmental Microbiology **8**:755-758.

Declerck, S. A. J., J. S. Coronel, P. Legendre, and L. Brendonck. 2011. Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High-Andes wetlands. Ecography **34**:296-305.

Dodson, S. I. 1992. Predicting crustacean zooplankton species richness. Limnology and Oceanography **37**:848-856.

Dodson, S. I., S. E. Arnott, and K. L. Cottingham. 2000. The relationship in lake communities between primary productivity and specie richness. Ecology **81**:2662-2679.

Domènech, R., A. Gaudes, López-doval, J. C., H. Salvadó, and I. Muñoz. 2006. Effects of short-term nutrient addition on microfauna density in a mediterranean stream. Hydrobiologia **568**:207-215.

Dopheide, A., G. Lear, Z. Z. He, J., and G. Lewis. 2015. Functional gene composition, diversity and redundancy in microbial stream biofilm communities. PLoS One **10**:e0123179.

Dopheide, A., G. Lear, R. Stott, and G. Lewis. 2008. Molecular characterization of ciliate diversity in stream biofilms. Applied and Environmental Microbiology **74**:1740-1747.

Dopheide, A., G. Lear, R. Stott, and G. Lewis. 2009. Relative diversity and community structure of ciliates in stream biofilms according to molecular and microscopy methods. Applied and Environmental Microbiology **75**:5261-5272.

Dorigo, U., D. Fontvieille, and J.-F. Humbert. 2006. Spatial variability in the abundance and composition of the free-living bacterioplankton community in the pelagic zone of Lake Bourget (France). FEMS Microbiology Ecology **58**:109-119.

Doupe, R. G. and D. Burrows. 2008. Thirty years later, should we be more concerned for the ongoing invasion of Mozambique Tilapia in Australia? Pacific Conservation Biology **14**:235-238.

Dray, S., P. Legendre, and F. G. Blanchet. 2007. Packfor: forward selection with permutation - R package ver. 0.0-7.

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling **196**:483-493.

Dunham, N. R. 2014. Influence of hydrological and environmental conditions on mangrove vegetation at coastal and inland semi-arid areas of the Gascoyne region. Edith Cowan University, Joondalup, WA.

Dunthorn, M., T. Stoek, J. C. Clamp, A. Warren, and F. Mahé. 2014. Ciliates and the rare biosphere: a review. Journal of Eukaryotic Microbiology **61**:404-409.

Dupont, C. 2014. Functional tradeoffs underpin salinity-driven divergence in microbial community composition. PLoS One **9**:e89549.

Dupraz, C., P. T. Visscher, L. K. Baumgartner, and R. P. Reid. 2004. Microbe-mineral interactions: early carbonate precipitation in a hypersaline lake (Eleuthera Island, Bahamas). Sedimentology **51**:745-765.

Dykhuizen, D. 1998. Santa Rosalia revisited: Why are there so many species of bacteria? Antonie van Leewenhoek **73**:25-33.

Ellison, J. C. and S. Simmonds. 2003. Structure and productivity of inland mangrove stands at Lake MacLeod, Wetern Australia. Journal of the Royal Society of Western Australia **86**:25-30.

Elloumi, J., J.-F. Carrias, H. Ayadi, T. Sime-Ngando, M. Boukhris, and A. Bouaïn. 2006. Composition and distribution of planktonic ciliates from ponds of different salinity in the solar saltwork of Sfax, Tunisia. Estuarine, Coastal and Shelf Science **67**:21-29.

Elloumi, J., Z. Drira, W. Guermazi, A. Hamza, and H. Ayadi. 2015. Space-time variation of ciliates related to environmental factors in 15 nearshore stations of the Gulf of Gabes (Tunisia, Eastern Mediterranean Sea). Mediterranean Marine Science **16**:162-179.

Etienne, R. S. 2007. A neutral sampling formula for multiple samples and an 'exact' test of neutrality. Ecology Letters **10**:608-618.

Etienne, R. S. 2009. Improved estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. Ecology **90**:847-852.

Etienne, R. S., A. M. Latimer, J. A. Silander, and R. M. Cowling. 2006. Comment on "Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot". Science **311**:610.

Faafeng, B. A. and M. Mjelde. 1998. Clear and turbid water in shallow Norwegian lakes related to submerged vegetation.*in* E. Jeppesen, M. Søndergaard, M. Søndergaard, and K. Christoffersen, editors. The Structuring Role of Submerged Macrophytes in Lakes. Springer Science+Business Media, New York.

Falkowski, P. G., T. Fenchel, and E. F. Delong. 2008. The microbial engines that drive Earth's biogeochemical cycles. Science **320**:1034-1038.

Fee, E., R. Hecky, S. Kasian, and D. Cruikshank. 1996. Effects of lake size, water clarity, and climatic variability on mixing depths in Canadian Shield lakes. Limnology and Oceanography **41**:912-920.

Fenchel, T. 1993. Are there more small than large species? Oikos **68**:375-378.

Fenchel, T. and B. Finlay. 2003. Is microbial diversity fundamentally different from biodiversity of larger animals and plants? European Journal of Protistology **39**:486-490.

Fenchel, T. and B. J. Finlay. 2004. The ubiquity of small species: Patterns of local and global diversity. BioScience **54**:777-784.

Fenchel, T., B. J. Finlay, E. A. D. Mitchell, T. Bell, J. A. Newman, I. P. Thompson, A. K. Lilley, and C. J. van der Gast. 2005. Bacteria and island biogeography. Science **309**:1997-1999.

Finlay, B. and K. J. Clarke. 1999. Ubiquitous dispersal of microbial species. Nature **400**:828.

Finlay, B., G. Esteban, and T. Fenchel. 1998. Protozoan diversity: Converging estimates of the global number of free-living ciliate species. Protist **149**:29-37.

Finlay, B. and T. Fenchel. 2004. Cosmopolitan metapopulations of free-living microbial eukaryotes. Protist **155**:237-244.

Finlay, B. J. 2002. Global dispersal of free-living microbial eukaryote species. Science **296**:1061-1063.

Finlay, B. J. and G. F. Esteban. 1998. Freshwater protozoa: biodiversity and ecological function. Biodiversity and Conservation **7**:1163-1186.

Finlay, B. J., G. F. Esteban, J. L. Olmo, and P. A. Tyler. 1999. Global distribution of free-living microbial species. Ecography **22**:138-144.

Finlay, B. J., S. C. Maberly, and J. I. Cooper. 1997. Microbial diversity and ecosystem function. Oikos **80**:209-213.

Fisher, J., C. J. Stratford, and S. Buckton. 2009. Variation in nutrient removal in three wetland blocks in relation to vegetation composition, inflow nutrient concentration and hydraulic loading. Ecological Engineering **35**:1387-1394.

Fisher, M. M. and E. W. Triplett. 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Applied and Environmental Microbiology **65**:4630-4636.

Foissner, W. 2006. Biogeography and dispersal of micro-organisms: A review emphasizing protists. Acta Protozzologica **45**:111-136.

Foissner, W. 2007. Dispersal and biogeography of protists: Recent advances. Japanese Journal of Protozoology **40**:1-16.

Foissner, W., A. Chao, and L. A. Katz. 2008. Diversity and geographic distribution of ciliates (Protista: Caliophora). Biodiversity and Conservation **17**:345-363.

Franklin, R. B. and A. L. Mills. 2003. Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. FEMS Microbiology Ecology **44**:335-346.

Frontier, S. 1985. Diversity and structure in aquatic ecosystems. Oceanography and Marine Biology Annual Review **23**:253-312.

Fuhrman, J. A., I. Hewson, M. S. Schwalbach, J. A. Steele, M. V. Brown, and S. Naeem. 2006. Annually reoccurring bacterial communities are predictable from ocean conditions. Proceedings of the National Academy of Sciences of the United States of America **103**:13104-13109.

Fuhrman, J. A., J. A. Steele, I. Hewson, M. S. Schwalbach, M. V. Brown, J. L. Green, and J. H. Brown. 2008. A latitudinal diversity gradient in planktonic marine bacteria. Proceedings of the National Academy of Sciences **105**:7774-7778.

Gage, S. H., S. A. Isard, and G. M. Colunga. 1999. Ecological scaling of aerobiological dispersal processes. Agricultural and Forestry Meteorology **30**:249-261.

Gainswin, B. 2004. Biotic influences on chemical fluxes and sediment-water exchanges in sediment deposits. University of Birmingham.

Garcia, E., C. M. Duarte, N. Marba, J. Terrados, H. Kennedy, M. D. Fortes, and N. H. Tri. 2003. Sediment deposition and production in SE-Asia seagrass meadows. Estuarine, Coastal and Shelf Science **56**:37-47.

Garcia, S. L., I. Salka, and H.-P. Grossart. 2013. Depth-discrete profiles of bacterial communities reveal pronounced spatio-temporal dynamics related to lake stratification. Environmental Microbiology Reports **5**:549-555.

Gaston, K. J. 1992. Regional numbers of insect and plant species. Functional Ecology **6**:243-247.

Geovariances and G. d. l. E. d. M. d. Paris. 2012. Isatis. France.

Gevers, D., F. M. Cohan, J. G. Lawrence, B. G. Spratt, T. Coenye, E. J. Feil, E. Stackebrandt, T. Van de Peer, P. Vandamme, F. L. Thompson, and J. Swings. 2005. Re-evaluating prokaryotic species. Nature Reviews. Microbiology **3**:733-739.

Gibbons, S. M., J. G. Caporaso, M. Pirrung, D. Field, R. Knight, and J. A. Gilbert. 2013. Evidence for a persistent microbial seed bank throughout the global ocean. Proceedings of the National Academy of Sciences **110**:4651-4655.

Gotelli, N. J. and G. L. Entsminger. 2004. EcoSim: Null models software for ecology. Acquired Intelligence Inc. & Kesey-Bear, Jericho, VT 05465.

Green, J. L. and B. J. M. Bohannan. 2006. Spatial scaling of microbial biodiversity. Trends in Ecology and Evolution **21**:501-507.

Green, J. L., A. J. Holmes, M. Westoby, I. Oliver, D. Briscoe, M. Dangerfield, M. Gillings, and A. J. Beattle. 2004. Spatial scaling of microbial eukaryote diversity. Nature **432**:747-750.

Gücker, B. and H. Fischer. 2003. Flagellate and ciliate distribution in sediments of a lowland river: relationships with environmental gradients and bacteria. Aquatic Microbial Ecology **31**:67-76.

Halse, S. A., R. J. Shiel, A. W. Storey, D. H. D. Edward, E. I. Lansbury, D. J. Cale, and M. S. Harvey. 2000. Aquatic invertebrates and waterbirds of wetlands and rivers of the southern Carnarvon basin, Wetern Australia. Records of the Western Australian Museum **61**:217-267.

Hambright, K. D., J. E. Beyer, J. D. Easton, R. M. Zamor, A. C. Easton, and T. C. Hallidayschult. 2015. The niche of an invasive marine microbe in a subtropical freshwater impoundment. The ISME Journal **9**:356-364.

Hanson, C. A., J. A. Fuhrman, M. C. Horner-devine, and J. B. H. Martiny. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature Reviews. Microbiology **10**:497-506.

Harte, J. and A. Kinzig, P. 1997. On the implications of species–area relationships for endemism, spatial turnover, and food web patterns. Oikos **80**:417-427.

Herczeg, A. I. and W. Berry Lyons. 1991. A chemical model for the evolution of Australian sodium chloride lake brines. Palaeogeography, Palaeoclimatology, Palaeoecology **84**:43-53.

Hillebrand, H., F. Watermann, R. Karez, and U. G. Berninger. 2001. Differences in species richness patterns between unicellular and multicellular organisms. Oecologia **126**:114-124.

Hingsamer, P., M. Striebel, M. Stockenreiter, R. Ptacnikova, R. Ptacnik, and S. A. Wickham. 2013. Interaction of productivity and disturbance in a marine ciliate community: a mesocosm study. Aquatic Microbial Ecology **70**:141-155.

Horner-Devine, C. C., M. Lage, J. B. Hughes, and B. J. M. Bohannan. 2004. A taxa-area relationship for bacteria. Nature **432**:750-753.

Horppila, J. and L. Nurminen. 2003. Effects of submerged macrophytes on sediment resuspension and internal phosphorus loading in Lake Hiidenvesi (southern Finland). Water Research **37**:4468-4474.

Hörtnagl, P., M. T. Pérez, M. Zeder, and R. Sommaruga. 2010. The bacterial community composition of the surface microlayer in a high mountain lake. FEMS Microbiology Ecology **73**:458-467.

Hubbell, S. P. 2001. The unified neutral theory of biodiversityand biogeography. Princeton University Press, Princeton, NJ.

Hugenholtz, P., B. M. Goebel, and N. R. Pace. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. Journal of Bacteriology **180**:4765-4774.

Humbert, J.-F., U. Dorigo, P. Cecchi, B. Le Berre, D. Debroas, and M. Bouvy. 2009. Comparison of the structure and composition of bacterial communities from temperate and tropical freshwater ecosystems. Environmental Microbiology **11**:2339-2350.

Jackson, D. A. and H. H. Harvey. 1993. Fish and benthic invertebrates: community concordance and community-environment relationships. Canadian Journal of Fisheries and Aquatic Sciences **50**:2641-2651.

Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities - the roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences **58**:157-170.

Jensen, H. S. and Ø. F. Anderson. 1992. Importance of temperature, nitrate, and pH for phosphate release from aerobic sediments of four shallow, eutrophic lakes. Limnology and Oceanography **37**:577-589.

Jensen, J. P., E. Jeppesen, K. Olrik, and P. Kristensen. 1994. Impact of nutrients and physical factors on the shift from cyanobacterial to chlorophyte dominance in shallow Danish lakes. Canadian Journal of Fisheries and Aquatic Sciences **51**:1692-1699.

Jones, S. E., T. A. Cadkin, R. J. Newton, and K. D. McMahon. 2012. Spatial and temporal scales of aquatic bacterial beta diversity. Frontiers in Microbiology **3**:e318.

Jones, S. E. and K. D. McMahon. 2009. Species-sorting may explain an apparent minimal effect of immigration on freshwater bacterial community dynamics. Environmental Microbiology **11**:905-913.

Jyväsjärvi, J., K. T. Tolonen, and H. Hämäläinen. 2009. Natural variation of profundal macroinvertebrate communities in boreal lakes is related to lake morphometry: implications for bioassessment. Canadian Journal of Fisheries and Aquatic Sciences **66**:589-601.

Kent, A. D., A. C. Yannarell, J. A. Rusak, E. W. Triplett, and K. D. McMahon. 2007. Synchrony in aquatic microbial community dynamics. The ISME Journal **1**:38-47.

King, A. J., K. R. Freeman, K. F. McCormick, R. C. Lynch, C. Lozupone, R. Knight, and S. K. Schmidt. 2010. Biogeography and habitat modelling of high-alpine bacteria. Nature Communications **1**:53.

Kissoon, L. T. T., D. L. Jaob, M. A. Hanson, B. R. Herwig, S. E. Bowe, and M. L. Otte. 2015. Multi-elements in waters and sediments of shallow lakes: relationships with water, sediment, and watershed characteristics. Wetlands **35**:443-457.

Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence-absence data. Journal of Animal Ecology **72**:367-382.

Kremp, A. 2001. Effects of cyst resuspension on germination and seeding of two bloom-forming dinoflagellates in the Baltic Sea. Marine Ecology Progress Series **216**:57-66.

Langenheder, S., E. S. Lindström, and L. J. Tranvik. 2005. Weak coupling between community composition and functioning of aquatic bacteria. Limnology and Oceanography **50**:957-967.

Langenheder, S. and H. Ragnarsson. 2007. The role of environmental and spatial factors for the composition of aquatic bacterial communities. Ecology **89**:2154-2161.

Langenheder, S. and A. Székely. 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. The ISME Journal **5**:1086-1094.

Lankau, E. W., P.-Y. Hong, and R. I. Mackie. 2012. Ecological drift and local exposures drive enteric bacterial community differences within species of Galápagos iguanas. Molecular Ecology **21**:1779-1788.

Latimer, A. M., J. A. Silander, and R. M. Cowling. 2005. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. Science **309**:1722-1725.

Lauber, C. L., K. S. Ramirez, Z. Aanderud, J. Lennon, and N. Fierer. 2013. Temporal variability in soil microbial communities across land-use types. The ISME Journal **7**:1641-1650.

Lawton, J. H. 1999. Are there general laws in ecology? Oikos **84**:177-192.

Lear, G., M. J. Anderson, J. P. Smith, K. Boxen, and G. D. Lewis. 2008. Spatial and temporal heterogeneity of the bacterial communities in stream epilithic biofilms. FEMS Microbiology Ecology **65**:463-473.

Lear, G., J. Bellamy, B. Case, S., J. Lee, E., and H. Buckly, L. 2014. Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. The ISME Journal **8**:1-12.

Lear, G., A. Dopheide, P. Ancion, and G. D. Lewis. 2011. A comparison of bacterial, ciliate and macroinvertebrate indicators of stream ecological health. Aquatic Ecology **45**:517-527.

Lear, G., V. Washington, M. Neale, B. Case, S., H. Buckley, and G. Lewis. 2013. The biogeography of stream bacteria. Global Ecology and Biogeography **22**:544-554.

Lee, C. K., C. W. Herbold, S. W. Polson, K. E. Wommack, S. J. Williamson, I. R. McDonald, and S. C. Cary. 2012. Groundtruthing next-gen sequencing for microbial ecology-biases and errors in community structure estimates from PCR amplicon pyrosequencing. PLoS One **7**:44224.

Lee, J. E. 2014. Niche and neutral processes in aquatic bacterial communities: Are all things equal? Lincoln University.

Lee, J. E., H. Buckley, R. S. Etienne, and G. Lear. 2013. Both species sorting and neutral processes drive assembly of bacterial communities in aquatic microcosms. FEMS Microbiology Ecology **86**:288-302.

Leff, L. G., J. V. McArthur, and L. J. Shimkets. 1998. Persistence and dissemination of introduced bacteria in freshwater microcosms. Microbial Ecology **36**:202-211.

Legendre, P. and M. J. Anderson. 1999. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. Ecological Monographs **69**:1-24.

Legendre, P. and M. De Cáceres. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Ecology Letters **16**:951-963.

Legendre, P. and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia **129**:271-280.

Legendre, P., J. Oksanen, and C. J. F. ter Braak. 2011. Testing the significance of canonical axes in redundancy analysis. Methods in Ecology and Evolution **2**:269-277.

Lei, Y., K. Stumm, S. Wickham, and U. G. Berninger. 2014. Distributions and biomass of benthic ciliates, foraminifera and amoeboid protists in marine, brackish, and freshwater sediments. Journal of Eukaryotic Microbiology **61**:493-508.

Leibold, M. A. 1995. The niche concept revisited: Mechanistic models and community context. Ecology **76**:1371.

Lepère, C., I. Domaizon, N. Taïb, J.-F. Mangot, G. Bronner, D. Boucher, and D. Debroas. 2013. Geographic distance and ecosystem size determine the distribution of smallest protists in lacustrine ecosystems. FEMS Microbiology Ecology **85**:85-94.

Lighthart, B. 1997. The ecology of bacteria in the alfresco atmosphere. FEMS Microbiology Ecology **23**:263-274.

Lindeman, R. L. 1942. The trophic-dynamic aspect of ecology. Ecology **23**:399-417.

Lindström, E. S., A. Eiler, S. Langenheder, S. Bertilsson, S. Drakare, H. Ragnarsson, and L. J. Tranvik. 2007. Does ecosystem size determine aquatic bacterial richness? Comment. Ecology **88**:252-253.

Lindström, E. S., X. M. Feng, W. Granéli, and E. S. Kritzberg. 2010. The interplay between bacterial community composition and the environment determining function of inland water bacteria. Limnology and Oceanography **55**:2052-2060.

Lindström, E. S. and S. Langenheder. 2012. Local and regional factors influencing bacterial community assembly. Environmental Microbiology Reports **4**:1-9.

Lindström, E. S. and Ö. Östman. 2011. The importance of dispersal for bacterial community composition and functioning. PLoS One **6**.

Livermore, J. A. and S. E. Jones. 2015. Local-global overlap in diversity informs mechanisms of bacterial biogeography. The ISME Journal **9**:2413-2422.

Logan, B. 1987. The MacLeod evaporite basin, Western Australia. The American Associaction of Petroleum Geologists, Tuksa, Oklahoma.

Logares, R., E. S. Lindström, S. Langenheder, J. B. Logue, H. Paterson, J. Laybourn-Parry, K. Rengefors, L. Tranvik, and S. Bertilsson. 2013. Biogeography of bacterial communities exposed to progressive long-term environmental change. The ISME Journal **7**:937-948.

Logue, J. B., S. Langenheder, A. F. Anderson, S. Bertilsson, S. Drakare, A. Lanzen, and E. S. Lindström. 2012. Freshwater bacterioplankton richness in oligotrophic lakes depends on nutrient availability rather than on species-area relationships. The ISME Journal **6**:1127-1136.

Logue, J. B. and E. S. Lindström. 2010. Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. The ISME Journal **4**:729-738.

Lomolino, M. V. 2001. Ecology's most general, yet protean pattern: the species-area relationship. Journal of Biogeography **27**:17-26.

Long, D. T., W. Berry Lyons, and M. E. Hines. 2009. Influence of hydrogeology, microbiology and landscape history on the geochemistry of acid hypersaline waters, N.W. Victoria. Applied Geochemistry **24**:285-296.

Maaß, S., M. Migliorini, M. C. Rillig, and T. Caruso. 2014. Disturbance, neutral theory, and patterns of beta diversity in soil communities. Ecology and Evolution **4**:4766-4774.

MacArthur, R. and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, N. J.

Madigan, M. T., J. M. Martinko, D. A. Stahl, and D. P. Clark. 2011. Brock Biology of Microorganisms. 13th edition. Benjamin Cummings, San Francisco.

Mantel, N. 1967. The detection of disease clustering and a gerneralized regression approach. Cancer Research **27**:209-220.

Margulis, L. and V. Schwartz. 1988. Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth. W. H. Freeman & Company, New York.

Martiny, J. B. H., B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, C. C. Horner-Devine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Ovreas, A. L. Reysenbach, V. H. Smith, and J. T. Staley. 2005. Microbial biogeography: putting microorganisms on the map. Nature Reviews. Microbiology **4**:102-112.

Martiny, J. B. H., J. A. Eisen, K. Penn, S. D. Allison, and C. C. Horner-Devine. 2011. Drivers of bacterial β-diversity depend on spatial scale. Proceedings of the National Academy of Sciences **108**:7850-7854.

Masel, J. 2011. Genetic drift. Current Biology **21**:R837-R838.

Mason, R. P. 2013. Trace Metals in Aquatic Systems. John Wiley & Sons, Somerset, NJ, USA.

McLure, N. 2011. Temporal and spatial variation in aquatic invertebrate communities at Lake MacLeod, Northwestern Australia. Edith Cowan University, Joondalup, WA.

Mees, F., C. Castañeda, and E. Van Ranst. 2011. Sedimentary and diagenetic features in saline lake deposits of the Monegros region, northern Spain. CATENA **85**:245-252.

Mehner, T., K. Holmgren, T. L. Lauridsen, E. J. Jeppesen, and M. Diekmann. 2007. Lake depth and geographical posiiton modify lake fish assemblages of the European 'Central Plains' ecoregion. Freshwater Biology **52**:2285-2297.

Miao, W., Y. Y. Shen, and X. Zhang. 2004. Intraspecific phylogeography of *Carchesium polypinum* (Peritrichia, Ciliophora) from China, inferred from 18S-ITS1-5.8S ribosomal DNA. Science in China Series C: Life Sciences **47**:11-17.

Mitsch, W. J. and J. G. Gosselink. 2000. Wetlands. 3rd edition. John Wiley & Sons, New York.

Monchy, S., G. Sanciu, M. Jobard, S. Rasconi, M. Gerphagnon, M. Chabé, A. Cian, D. Meloni, N. Niquil, U. Christaki, E. Viscogliosi, and T. Sime-Ngando. 2011. Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. Environmental Microbiology **13**:1433-1453.

Morgan, D. 2010. Appliation of sonar for the measurment of sludge heights in wastewater stabilisation ponds. University of Western Australia, Perth, Australia.

Mosher, J. J. and R. H. Findlay. 2011. Direct and indirect influence of parental bedrock on streambed microbial community structure in forested streams. Applied and Environmental Microbiology **77**:7681-7688.

Mouguet, N. and M. Loreau. 2003. Community patterns in source-sink metacommunities. The American Naturalist **162**:544-557.

Müller, M., P. Stadler, and T. Weisse. 2002. Seasonal dynamics of cyst formation of strombidiid ciliates in alpine Lake Mondsee, Austria. Aquatic Microbial Ecology **29**:181-188.

Nekola, J. C. and P. S. White. 1999. The distance decay of similarity in biogeography and ecology. Journal of Biogeography **26**`:867-878.

Nemergut, D. R., S. K. Schmidt, T. Fukami, S. P. O'Neill, T. M. Bilinski, L. F. Stanish, J. E. Knelman, J. L. Darcy, R. C. Lynch, P. Wicky, and S. Ferrenberg. 2013. Patterns and processes of microbial community assembly. Microbiology and Molecular Biology Reviews **77**:342-356.

Newton, J. A., S. E. Jones, A. Eiler, K. D. McMahon, and S. Bertilsson. 2011. A guide to the natural history of freshwater lake bacteria. Microbiology and Molecular Biology Reviews **75**:14-49.

Nurminen, L. and J. Horppila. 2009. Life form dependent impacts of macrophyte vegetation on the ratio of resuspended nutrients. Water Research **43**:3217-3226.

O'Malley, M. A. 2008. 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences **39**:314-325.

Ochman, H., E. Lerat, and V. Daubin. 2005. Examining bacterial species under the specter of genne transfer and exchange. Proceedings of the National Academy of Sciences **102**:6595-6599.

Odour, S. O. and M. Schagerl. 2007. Temporal trends of ion contents and nutrients in three Kenyan Rift Valley saline–alkaline lakes and their influence on phytoplankton biomass. Hydrobiologia **584**:59-68.

Ofiteru, I. D., M. Lunn, T. P. Curtis, G. F. Wells, C. S. Criddle, C. A. Francis, and W. T. Sloan. 2010. Combined niche and neutral effects in a microbial wastewater treatment community. Proceedings of the National Academy of Sciences **107**:15345-15350.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, and H. Wagner. 2013. vegan: Community Ecology Package. r package version 2.0-10. http://CRAN.R-project.org/package=vegan.

Olding, D., D., J. Hellebust, A., and M. Douglas, S., V. 2000. Phytoplankton community composition in relation to water quality and water-body morphometry in urban lakes, reservoirs, and ponds. Canadian Journal of Fisheries and Aquatic Sciences **57**:2163-2174.

Oren, A. 2002. D iversity of halophilic microorganisms: Environments phylogeny, physiology, and application. Journal of Industrial Microbiology and biotechnology **28**:56-63.

Orsi, W., S. Charvet, P. Vd'acný, J. M. Bernhard, and V. P. Edgcomb. 2012. Prevalence of partnerships between bacteria and ciliates in oxygen-depleted marine water columns. Frontiers in Microbiology **3**:341.

Östman, Ö., S. Drakare, E. S. Kritzberg, S. Langenheder, J. Logue, and E. S. Lindström. 2012. Importance of space and the local environment for linking local and regional abundances of microbes. Aquatic Microbial Ecology **67**:35-45.

Pagaling, E., H. Wang, M. Venables, A. Wallace, W. D. Grant, D. A. Cowan, B. E. Jones, Y. Ma, A. Ventosa, and S. Heaphu. 2009. Microbial biogeography of six salt lakes in inner Mongolia, Chin and a salt lake in Argentina. Applied and Environmental Microbiology **75**:5750-5760.

Patterson, D. J., J. Larsen, and J. O. Corliss. 1989. T he ecology of heterotrophic flagellates and ciliates living in marine sediments. Progress in Protositology **3**:185-277.

Pebesma, E., R. Bivand, B. Rowlingson, V. Gomez-Rubio, and R. Hijmans. 2015. sp: classes and methods for spatial data. r package versio 1.0-17. http://cran.r-project.org/package=sp.

Pebesma, E. and B. Graeler. 2014. gstat: Spatial and spatio-temporal geostatistical modelling, prediction and simulation. r package version 1.0-22. http://cran.r-project.org/package=gstat.

Pedrós-Alió, C. 2006. M arine microbial diversity: can it be determined. Trends in Microbiology **14**:257-263.

Pernthaler, J. and R. Amann. 2005. F ate of heterotrophic microbes in pelagic habitats: Focus on popul ations. Microbiology and Molecular Biology Reviews **69**:440-461.

Phillips, B., R. Butcher, J. Hale, and M. Coote. 2005. Ecological character of the Lake MacLeod wetland of international importance. Department of Conservation and Land Managment, Western Australia.

Plebani, M., K. E. Fussmann, D. Hansen, E. J. O'Gorman, R. I. A. Stewart, G. Woodward, and O. L. Petchey. 2015. Substratum-dependent responses of ciliate

assemblages to temperature: a n atural experiment in Icelandic streams. Freshwater Biology **60**.

Portillo, M. C., S. P. Anderson, and N. Fierer. 2012. T emporal variability in the diversity and composition of stream bacterioplankton communities. Environmental Microbiology **14**:2417-2428.

Radke, L. C., K. W. F. Howard, and P. A. Gell. 2002. C hemical diversity in south-eastern Australian saline lakes I: geochemical causes. Marine and Freshwater Research **53**:941-959.

Radke, L. C., S. Juggins, S. A. Halse, P. De Deckker, and T. Finston. 2003. C hemical diversity in south-eastern Australian saline lakes II: biotic implications. Marine and Freshwater Research **54**:895-912.

Raffaelli, D., E. Bell, G. Weithoff, A. Matsumoto, J. J. Cruz-Motta, P. Kershaw, R. Parker, D. Parry, and M. Jones. 2003. T he ups and downs of benthic ecology: considerations of scale, heterogeneity and surveillance for the benthic-pelagic coupling. Journal of Experimental Marine Biology and Ecology **285-286**:191-203.

Ramette, A. 2009. Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. Applied and Environmental Microbiology **75**:2495-2505.

Ranjard, L., F. Poly, J.-C. Lata, C. Mougel, J. Thioulouse, and S. Nazaret. 2001. Characterization of bacterial and fungal soil communities by automated ribosomal intergenic spacer analysis fingerprints: Biological and methodological variability. Applied and Environmental Microbiology **67**:4479-4487.

Rao, C. R. 1995. A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. Qüestiió **19**:23-63.

Reche, I., E. Pulido-Villena, R. Morales-Baquero, and E. O. Casamayor. 2005. D oes ecosystem size determine aquatic bacterial richness? Ecology **86**:1715-1722.

Redford, A. J. and N. Fierer. 2009. B acterial succession on t he leaf surface: a novel system for studying successional dynamics. Microbial Ecology **58**:189-198.

Reed, H. E. and J. B. H. Martiny. 2007. Testing the functional significance of microbial composition in natural communities. FEMS Microbiology Ecology **62**:161-170.

Reid, A. and M. Buckley. 2011. The Rare Biosphere. Washington DC.

Reiss, J. and J. M. Schmid-Araya. 2008. E xisting in plenty: abundance, biomass and diversity of ciliates and meiofauna in small streams. Freshwater Biology **53**:652-668.

Renard, D., N. Bez, N. Desassis, H. Beucher, F. Ors, and F. Laporte. 2014. RGeostats: The Geostatistical package [10.0.3]. MINES ParisTech.

Rengefors, K., R. Logares, and J. Laybourn-Parry. 2012. P olar lakes may act as ecological islands to aquatic protists. Molecular Ecology **21**:3200-3209.

Rennella, A., M. and R. Quiros. 2006. The effects of hydrology on plankton biomass in shallow lakes of the Pampa Plain. Hydrobiologia **556**:181-191.

Ricklefs, R. E. and I. J. Lovette. 1999. T he roles of island area per se and habitat diversity in the species-area relationships of four Lesser Antillean faunal groups. Journal of Animal Ecology **68**:1142-1160.

Ripley, B., B. Venables, D. M. M. Bates, K. Hornik, A. Gebhardt, and D. Firth. 2014. MASS: Support Functions and Datasets for Venables and Ripley's MASS.*in* r. p. v. 7.3-39, editor. http://cran.r-project.org/package=MASS.

Romina Schiaffino, M., F. Unrein, J. M. Gasol, R. Massana, V. Balagué, and I. Izaguirre. 2011. Bacterial community structure in a latitudinal gradient of lakes: the roles of spatial versus environmental factors. Freshwater Biology **56**.

Rosendell, J., S. P. Hubbell, F. He, L. J. Harmon, and R. S. Etienne. 2012. The case for ecological neutral theory. Trends in Ecology and Evolution **27**:203-208.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, UK.

Rowan, J. S., S. Black, and S. W. Franks. 2012. S ediment fingerprinting as an environmental forensics tool explaining cyanobacteria blooms in lakes. Applied Geography **32**:832-843.

Ruokolainen, L., E. Ranta, V. Kaitala, and M. S. Fowler. 2009. When can we distinguish between neutral and non-neutral processes in community dynamics under ecological drift? Ecology Letters **12**:909-919.

Schauer, R., C. Bienhold, A. Ramette, and J. Harder. 2010. B acterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. The ISME Journal **4**:159-170.

Scheffer, M. 2004. Ecology of shallow lakes.*in* M. B. Usher, D. L. De Angelis, and B. F. J. Manly, editors. Population and Community Biology Series 202. Kluwer Academic Publishers, Dordrecht.

Schnurrenberger, D., J. Russell, and K. Kelts. 2003. C lassification of lacustrine sediments based on s edimentary components. Journal of Paleolimnology **29**:141-154.

Schoener, T. W. 2010. The MacArthur-Wilson equilibrium model.*in* J. B. Losos and R. E. Ricklefs, editors. The Theory of Island Biogeography Revissted. Princeton University Press, Princeton, NJ.

Shade, A., C.-Y. Chiu, and K. D. McMahon. 2010. Seasonal and episodic lake mixing simulate differential planktonic bacterial dynamics. Microbial Ecology **59**:546-554.

Shade, A., S. E. Jones, and K. D. McMahon. 2008. T he influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. Environmental Microbiology **10**:1057-1067.

Shade, A., J. G. Caporaso, J. Handelsman, R. Knight, and N. Fierer. 2013. A meta-analysis of changes in bacterial and archaeal communities with time. The ISME Journal **7**:1493-1506.

Shade, A., J. S. Read, N. D. Youngblut, N. Fierer, R. Knight, T. K. Kratz, N. R. Lottig, E. E. Roden, E. H. Stanley, J. Stombaugh, R. J. Whitaker, C. H. Wu, and K. D. McMahon. 2012. Lake microbial communities are resilient after a whole-ecosystem disturbance. The ISME Journal **6**:2153-2167.

Shepherd, M. J. 1990. Hydrologic environments and sedimentation, Cygnet Seepage Face, Lake MacLeod, Western Australia. University of Western Australia.

Sloan, W. T., M. Lunn, S. Woodcock, I. M. Head, S. Nee, and T. P. Curtis. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environmental Microbiology **8**:732-740.

Sloan, W. T., S. Woodcock, M. Lunn, I. M. Head, and T. P. Curtis. 2007. Modeling taxa-abundance distributions in microbial communities using environmental sequence data. Microbial Ecology **53**:443-455.

Smith, E. P. and G. van Belle. 1984. Estimation of species richness. Biometrics **40**:119-129.

Smith, V. H., B. L. Foster, J. P. Grover, R. D. Holt, M. A. Leibold, and J. R. Denoyelles. 2005. Phytoplankton species richness scales consistently from laboratory microcosms to the world's oceans. Proceedings of the National Academy of Sciences **102**:4393-4396.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology **35**:627-632.

Soininen, J. 2010. Species turnover along abiotic and biotic gradients: patterns in space equal patterns in time? BioScience **60**:433-439.

Soininen, J., R. McDonald, and H. Hillebrand. 2007. The distance decay of similarity in ecological communities. Ecography **30**:3-12.

Stock, A., V. Edgcomb, W. Orsi, S. Filker, H.-W. Breiner, M. M. Yakimov, and T. Stoeck. 2013. Evidence for isolated evolution of deep-sea ciliate communities through geological separation and environmental selection. BMC Microbiology **13**:150.

Stone, L. and A. Roberts. 1990. The checkerboard score and species distributions. Oecologia **85**:74-79.

Storey, A. W., S. A. Halse, and S. Creagh. 2011. Aquatic fauna and water chemitry of the mound springs and wetlands of Mandora Marsh, north-western Australia. Journal of the Royal Society of Western Australia **94**:419-437.

Streamtec. 2003. Lake MacLeod northern ponds food web structure. Report ST 05/03 to Dampier Salt Pty Ltd. Streamtec Pty Ltd.

Sul, W. J., T. A. Oliver, H. W. Ducklow, L. A. Amaral-Zettler, and M. L. Sogin. 2013. Marine bacteria exhibit a bipolar distribution. Proceedings of the National Academy of Sciences **110**:2342-2347.

Székely, A. and S. Langenheder. 2013. The importance of species sorting differs between habitat generalists and specialists in bacterial communities. FEMS Microbiology Ecology **87**:102-112.

Tang, H. and P. Xie. 2000. Budgets and dynamics of nitrogen and phosphorus in a shallow, hypereutrophic lake in China. Journal of Freshwater Ecology **15**:505-514.

Tiedje, J. M., S. Asuming-Brempong, K. Nüsslein, T. L. Marsh, and S. J. Flynn. 1999. Opening the black box of soil microbial diversity. Applied Soil Ecology **13**:109-122.

Tilman, D. 1982. Resource competition and community structure. Princeton University Press, Princeton, NJ.

Timms, B. V. 2005. Salt lakes in Australia: present problems and prognosis for the future. Hydrobiologia **552**:1-15.

Timms, B. V. 2009. A study of the salt lakes and salt springs of Eyre Peninsula, South Australia. Hydrobiologia **626**:41-51.

Timms, B. V. 2010. Blue Lagoon, South Australia: A closed marine lake harbouring potential invaders of continental saline lakes? Verhandlungen des Internationalen Verein Limnologie **30**:1425-14228.

Torsvik, V., J. Goksøyr, and F. L. Daae. 1990. High diversity in DNA of soil bacteria. Applied and Environmental Microbiology **56**:782-787.

Torsvik, V., L. Øvreås, and T. F. Thingstad. 2002. Prokaryotic diversity--magnitude, dynmacis and controlling factors. Science **296**:1064.

Ugland, K. I., J. S. Gray, and K. E. Ellingsen. 2003. The species-accumualtion curve and estimation of species richness. Journal of Animal Ecology **72**:888-897.

Van Cleave, K. 2012. Interaction among evaporation, ice cover, and water temperature on Lake Superior: decadal, interannual, and seasonal variability. University of Nebraska, Lincoln, Nebraska.

van den Boogaart, K. G. and R. Tolosana-Delgado. 2013. Analyzing Compositional Data with R. 1 edition. Springer-Verlag, Berlin Heidelberg.

van den Boogaart, K. G., R. Tolosana-Delgado, and M. Bren. 2014. compositions: Compositional Data Analysis. r package version 1.40-1. http://CRAN.R-project.org/package=compositions.

van der Gast, C. J., A. K. Lilley, D. Ager, and I. P. Thompson. 2005. Island size and bacterial diversity in an archipelago of engineering machines. Environmental Microbiology **7**:1220-1226.

Van der Gucht, K., K. Cottenie, K. Muylaert, N. Vloemans, S. Cousin, S. Declerck, E. Jeppesen, J.-M. Conde-Porcuna, K. Schwenk, G. Zwart, H. Degans, W. Vyverman, and L. De Meester. 2007. The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. Proceedings of the National Academy of Sciences of the United States of America **104**:20404-20409.

Vellend, M. 2010. Conceptual synthesis in community ecology. The Quarterly Review of Biology **85**:183-206.

Venäläinen, A., M. Heikinheimo, and T. Tourula. 1998. Latent heat flux from small sheltered lakes. Boundary-Layer Meterology **86**:355-377.

Vitousek, P. M. 1990. Biological invasions and ecosystem processes: towards an integration of population biology and ecosystem studies. Oikos **57**:7-13.

Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2007. Patterns of relative species abundance in rainforest and coral reefs. Nature **450**:45-49.

Vymazal, J. 1995. Algae and element cycling in wetlands. Lewis Publishers, University of California.

Walsh, D. A. and W. F. Doolittle. 2005. The real ‘‘domains’’ of life. Current Biology **15**:R237-R240.

Wardle, D. A. 1999. Is "sampling effect" a problem for experiments investigatin biodiversity-ecosystem function relationships? Oikos **87**:403-407.

Weisse, T. 2008. Distribution and diveristy of aquatic protists: an evolutionary and ecological perspective. Biodiversity and Conservation **17**:243-259.

Weisse, T. 2014. Ciliates and the Rare Biosphere—Community Ecology and Population Dynamics. Journal of Eukaryotic Microbiology **61**:419-433.

Wells, G. F., H.-D. Park, B. Eggleston, C. A. Francis, and C. S. Criddle. 2011. Fine-scale bacterial community dynamics and the taxa–time relationship within a full-scale activated sludge bioreactor. Water Research **45**:5476-5488.

Wetzel, R. G. 2001. Limnology. Lake and river ecosystems. Academic Press, London, U.K.

Wey, J. K., H. Norf, H. Arndt, and M. Weitere. 2009. Role of dispersal in shaping communities of ciliates and heterotrophic flagellates within riverine biofilms. Limnology and Oceanography **54**:1615-1626.

Whitaker, R. J., D. W. Grogan, and J. W. Taylor. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science **301**:976+.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: The unseen majority. Proceedings of the National Academy of Sciences of the United States of America **95**:6578-6583.

Whittaker, R. H. 1975. Communities and ecosystems. MacMillan Publishing, New York.

Wickham, S. A., S. Nagel, and H. Hillebrand. 2004. Control of epibenthic ciliate communities by grazers and nutrients. Aquatic Microbial Ecology **35**:153-162.

Wong, J. W. C. and C. L. Yang. 1997. The effect of pH and redox potential on the release of nutrients and heavy metals from a contaminated marine sediment. Toxicological and Environmental Chemistry **62**:1-10.

Woodcock, S., T. P. Curtis, I. M. Head, M. Lunn, and W. T. Sloan. 2006. Taxa–area relationships for microbes: the unsampled and the unseen. Ecology Letters **9**:805-812.

Woodcock, S., C. J. van der Gast, T. Bell, M. Lunn, T. P. Curtis, I. M. Head, and W. T. Sloan. 2007. Neutral assembly of bacterial communities. FEMS Microbiology Ecology **62**:171-180.

Woodruff, S. L., W. A. House, M. E. Callow, and B. S. C. Leadbeater. 1999. The effect of a developing biofilm on chemical changes across the sediment-water interface in a freshwater environment. International Review of Hydrobiology **84**:509-532.

Wurzbacher, C. M., F. Bärlocher, and H.-P. Grossart. 2010. Fungi in lake ecosystems. Aquatic Microbial Ecology **59**:125-149.

Wyrwoll, K., T. Stoneman, G. Elliot, and P. Sandercock. 2000. Geoecological setting of the Carnarvon Basin, Western Australia: geology, geomorphology and soils of selected sites. Records of the Western Australian Museum **61**:29-75.

Yannarell, A. C. and E. W. Triplett. 2004. Within- and between-lake variability in the composition of bacterioplankton communities: Investigations using multiple spatial scales. Applied and Environmental Microbiology **70**:214-223.

Yannarell, A. C. and E. W. Triplett. 2005. Geographic and environmental sources of variation in lake bacterial community composition. Applied and Environmental Microbiology **71**:227-239.

APPENDIX 1 – ROSE OF WIND DIRECTION VS SPEED

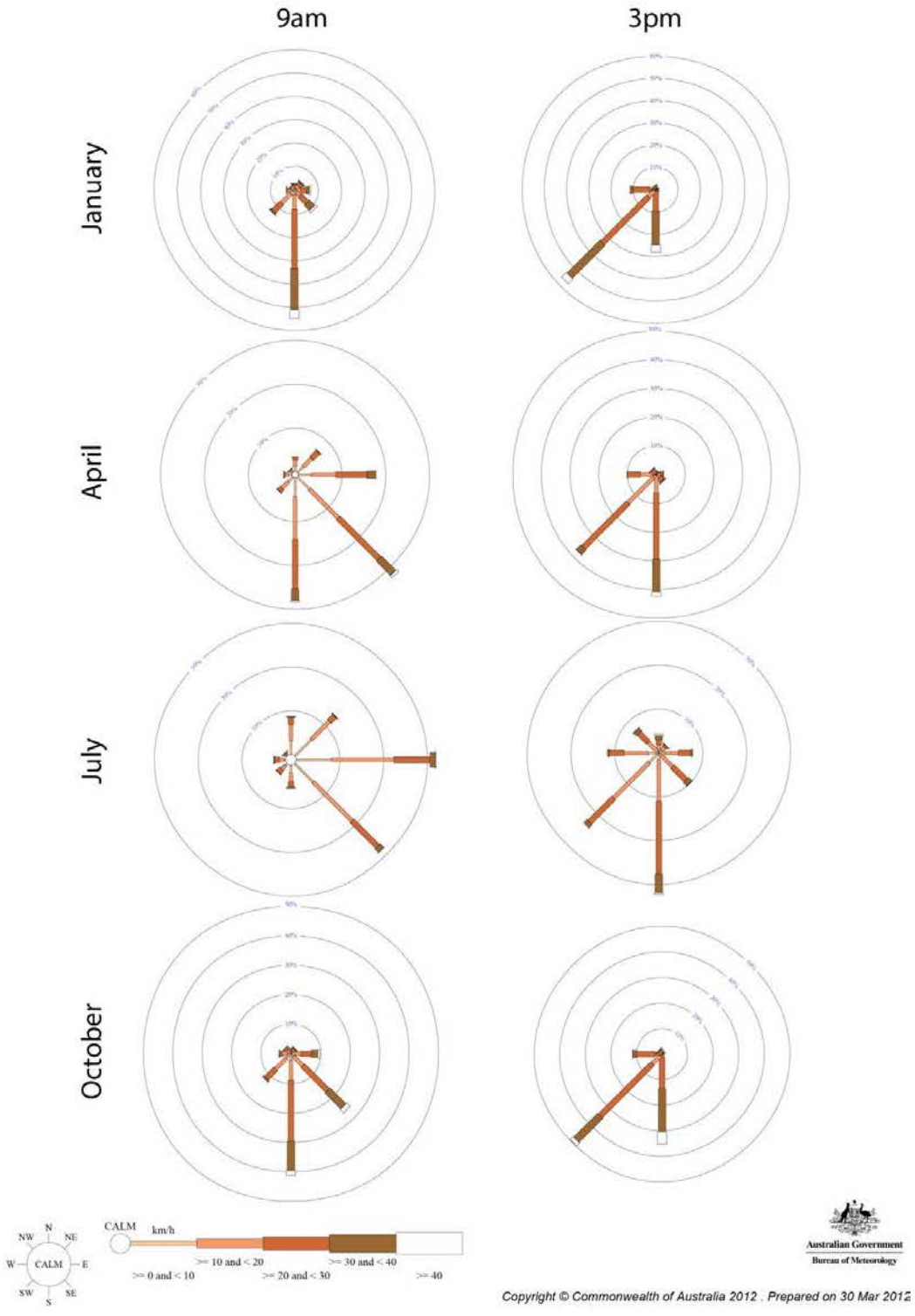


Figure 55: Rose of wind direction versus wind direction in km/h for the months January, April, July and October. All data collected at Carnarvon Airport (BOM Site: 006011) for the period between 1945 to 2010. Graphs reproduced from Bureau of Meteorology.

APPENDIX 2 – VARIOGRAM MODELS

Table 36: Summary of parameters used to derive the variogram models used for the kriging processes for determining the bathymetry of each pond.

	Lag distance	Number of lags	Direction of anistropy	Range (m)	Total sill	Model structure
Pete's Vent	0.5	12	0	11.7	0.096	Spherical
Donut Pond	3	15	30, 120	178.0	7.506	Spherical
Annie's Pond	5	10	60, 150	60.8	0.088	Spherical, Cubic
Pete's Pond	10	10	60, 150	114.7	0.037	Spherical
Whistler's Pond	10	30	70, 160	193.8	0.121	Spherical
Harjie's Pond	2	20	55, 90, 145, 180	13.1	0.053	Spherical
Jana's Vent	20	30	70, 160	118.9	1.102	Spherical
Cygnet Pond	10	100	0, 90	201.6	0.071	Spherical

APPENDIX 3 – R-SCRIPTS

Bathymetric determination

This R-Script was written to build the bathymetric maps of the ponds and to determine some of the morphometric characteristics of the ponds. The raw data file (*data*) should contain northing [,1] and easting [,2] values for each recorded depth [,3] and temperature [,4] measurement and imported in .csv format. The outline of the water body can be imported directly in .kml format, although the file needs to be in the working directory.

```
#Load required packages
library(sp)
library(rgdal)
library(RGeostats)
library(maptools)

#Load depth and temperature data file with coordinated in UTM
data <- read.csv("[….].csv", header=TRUE)
dt.z <- data
dt.z$Temp <- NULL
dt.t <- data
dt.t$z <- NULL

#load outline of pond as .kml file into UTM coordinates
spa <- data.frame(getKMLcoordinates(kmlfile="[….].kml", ignoreAltitude=TRUE))
colnames(spa) <- c("x", "y")
coordinates(spa) <- c("x", "y")
proj4string(spa) <- CRS("+proj=longlat +datum=WGS84")
spa <- data.frame(spTransform(spa, CRS("+proj=utm +zone=49 ellps=WGS84")))

#Build polygon shape of pond outline
polygon <- polygon.create(x=spa$x, y=spa$y)
plot(polygon)
axis.values <- par("usr")
n.nodes.x1 <- 200
n.nodes.x2 <- 200
n.dim <- c(((axis.values[2]-axis.values[1])/n.nodes.x1),
      ((axis.values[4]-axis.values[3])/n.nodes.x2), autoname=F)

#Create grid for kriging
grid.db <- db.create(flag.grid=TRUE,
            dx=c(n.dim[1], n.dim[2]),
            nx=c(n.nodes.x1, n.nodes.x2), x0=c(axis.values[1], axis.values[3]))
grid.poly <- db.polygon(grid.db, polygon)

#Build variograms and variogram models
z.db <- db.create(dt.z, flag.grid=FALSE, ndim=2)
z.db <- db.locate(z.db, c("x","y"), "x")
```

```
lag.distance.z <- [INSERT LAG DISTANCE]
lag.number.z <- [INSERT NUMBER OF LAGS]
dir1 <- [INSERT DIRECTION OF ANISTROPY]
dir2 <- dir1+90
z.dir.vario <- vario.calc(z.db, lag=lag.distance.z, nlag=lag.number.z)
plot(z.dir.vario,npairdw=TRUE,npairpt=TRUE, title="Omnidirectional variogram")
z.dir.model <- model.auto(z.dir.vario,
        struct=c("Nugget Effect", "Spherical", "Spherical", "Exponential", "Exponential"),
        title = "1 direction variogram model")

#Build neighbourhood matrix
unique.neigh <- neigh.init(type=0, ndim=2)
moving.neigh <- neigh.input()

#Kriging & simulation
depth.db <- db.polygon(z.db, polygon)
depth.db <- xvalid(depth.db, model=z.dir.model, neigh=moving.neigh)
depth.db <- db.locate(depth.db, seq(6,7))
depth.db <- db.locate(depth.db, "z", "z")
grid.poly.z <- kriging(depth.db, grid.poly, z.dir.model, unique.neigh, radix="KU.Part")
grid.poly.z <- kriging(depth.db, grid.poly.z, z.dir.model, moving.neigh, radix="KM.All")
grid.poly.z <- neigh.test(depth.db, grid.poly.z, z.dir.model, moving.neigh, radix="Moving")
depth.anam <- anam.fit(depth.db, "z")
depth.db <- anam.z2y(depth.db, "z", anam=depth.anam)
depth.g.vario <- vario.calc(z.db, nlag=lag.number.z, lag=lag.distance.z)
plot(depth.g.vario, npairdw=T, npairpt=T)
depth.g.model <- model.auto(depth.g.vario, struct=c("Nugget Effect", "Spherical", "Spherical",
        "Exponential", "Exponential"))
grid.poly.z <- simtub(depth.db, grid.poly.z, depth.g.model, unique.neigh, nbsimu= 100, nbtuba=100)
grid.poly.z <- anam.y2z(grid.poly.z, ngrep="Simu.Gaussian.z", anam=depth.anam)
grid.poly.z.mean <- db.compare(grid.poly.z, ngrep="Raw.Simu.Gaussian.z", fun="mean")
grid.poly.z.stdv <- db.compare(grid.poly.z, ngrep="Raw.Simu.Gaussian.z", fun="stdv")

#Lake characteristics
depth.points <- data.frame(grid.poly.z.mean@items$mean)
depth.points <- na.omit(depth.points)
names(depth.points) [1] <- "Depth"
Lake.character <- data.frame((nrow(depth.points)*(grid.poly@dx[1])*(grid.poly@dx[2])),
                  (sum(depth.points$Depth*(grid.poly@dx[1])*(grid.poly@dx[2]))),
                  mean(depth.points$Depth), max(depth.points$Depth))
names(Lake.character) <- c("S.A.", "V", "Mean", "Maximum")
Lake.character <- round(Lake.character, digits=2)

#Output Map
plot(grid.poly.z.mean, col=terrain.colors(21), asp=1, xlab= "Latitude", ylab="Longitude")
plot(grid.poly.z.mean, name.contour="mean", nlevels=10, col="black", labcex=1, add=TRUE)
plot(polygon, lwd=3, add=TRUE)
```

Sediment analysis

R script used to conduct the sedimentology analyses. This script relies on two sets of data: the sediment data (*sed*) which contains the counts of each particle type for each sample, and the coordinates for each sample (*spa*).

```
#Load required packages and codes
library(vegan)
source("sr.value.R") # http://www.davidzeleny.net/anadat-r/doku.php/cs:numecolr:sr.value
source("scores.cca.R")# https://searchcode.com/codesearch/view/15734587/

#Load data
sed  <- read.csv("[….].csv", row.names=1, header=TRUE)
spa <- read.csv("[….].csv", row.names=1, header=TRUE)

#Remove un-used coordinates
spa <- spa[Reduce(intersect, lapply(list(sed, spa), rownames)), ]

sed.prop <- data.frame(mean(sed $pel), mean(sed $ske), mean(sed $det),
mean(sed $dia),  mean(sed $min))

sed.norm <- decostand(sed, "normalize")
sed.mds <- (metaMDS(sed.norm))
sed.anosim <- anosim(sed.norm, grouping=groups$Pond, distance="bray")
sed.PERMANOVA <- adonis(sed.norm~Pond, groups)
sed.PERMANOVA
sed.cn.dist <- vegdist(sed.norm, distance="bray")
anova(betadisper(sed.cn.dist, groups$Pond))
betadisper(sed.cn.dist, groups$Pond)
plot(sed.mds, type="n",
   main="Sediment composition of each location")
points(sed.mds, select=which(groups$Pond=="A"), col="blue")
points(sed.mds, select=which(groups$Pond=="C"), col="olivedrab")
points(sed.mds, select=which(groups$Pond=="D"), col="green4")
points(sed.mds, select=which(groups$Pond=="H"), col="purple")
points(sed.mds, select=which(groups$Pond=="J"), col="red")
points(sed.mds, select=which(groups$Pond=="P"), col="darkorange1")
points(sed.mds, select=which(groups$Pond=="S"), col="darkred")
points(sed.mds, select=which(groups$Pond=="W"), col="violetred")
ordihull(sed.mds, group=groups$Pond)

#K-means clustering to identify sediment groups
sed.KM.cascade <- cascadeKM(sed.norm, inf.gr=2, sup.gr=11,
                        iter=9999, criterion="ssi")
plot(sed.KM.cascade, sortg=TRUE)
sed.kmeans <- kmeans(sed.norm, centers=6, nstart=9999)
library(labdsv)
iva <- indval(sed.norm, sed.kmeans$cluster)
iva <- indval(sed.norm, sed.kmeans$cluster)
#Table of significant indicators
gr <- iva$maxcls[iva$pval<=0.05]
iv <- iva$indcls[iva$pval<=0.05]
```

```
pv <- iva$pval[iva$pval<=0.05]
fr <- apply(sed.norm > 0, 2, sum) [iva$pval<=0.05]
fidg <- data.frame(group=gr, indval=iv, pvalue=pv, freq=fr)
(fidg <- fidg[order(fidg$group, -fidg$indval),])
clus$cluster <- as.factor(clus$cluster)
summary(clus)

#Extract information on the compositions of the sediment types.
Sed.nominclature <- cbind(sed, clus$cluster)
names(Sed.nominclature)[6] <- "cluster"
one.sed <- subset(Sed.nominclature, Sed.nominclature$cluster == "1")
two.sed <- subset(Sed.nominclature, Sed.nominclature$cluster == "2")
three.sed <- subset(Sed.nominclature, Sed.nominclature$cluster == "3")
four.sed <- subset(Sed.nominclature, Sed.nominclature$cluster == "4")
five.sed <- subset(Sed.nominclature, Sed.nominclature$cluster == "5")
six.sed <- subset(Sed.nominclature, Sed.nominclature$cluster == "6")

one.sed.prop <- data.frame(mean(one.sed$pel), mean(one.sed$ske), mean(one.sed$det),
mean(one.sed$dia),
                  mean(one.sed$min))
two.sed.prop <- data.frame(mean(two.sed$pel), mean(two.sed$ske), mean(two.sed$det),
mean(two.sed$dia),
                  mean(two.sed$min))
three.sed.prop <- data.frame(mean(three.sed$pel), mean(three.sed$ske), mean(three.sed$det),
mean(three.sed$dia),
                   mean(three.sed$min))
four.sed.prop <- data.frame(mean(four.sed$pel), mean(four.sed$ske), mean(four.sed$det),
mean(four.sed$dia),
                   mean(four.sed$min))
five.sed.prop <- data.frame(mean(five.sed$pel), mean(five.sed$ske), mean(five.sed$det),
mean(five.sed$dia),
                   mean(five.sed$min))
six.sed.prop <- data.frame(mean(six.sed$pel), mean(six.sed$ske), mean(six.sed$det), mean(six.sed$dia),
                  mean(six.sed$min))
names(one.sed.prop)[1:5] <- c("pel", "ske", "det", "dia", "min")
names(two.sed.prop)[1:5] <- c("pel", "ske", "det", "dia", "min")
names(three.sed.prop)[1:5] <- c("pel", "ske", "det", "dia", "min")
names(four.sed.prop)[1:5] <- c("pel", "ske", "det", "dia", "min")
names(five.sed.prop)[1:5] <- c("pel", "ske", "det", "dia", "min")
names(six.sed.prop)[1:5] <- c("pel", "ske", "det", "dia", "min")

Sed.groups.prop <- rbind(one.sed.prop, two.sed.prop, three.sed.prop, four.sed.prop, five.sed.prop,
six.sed.prop)
Sed.groups.prop$Group <- c("One", "Two", "Three", "Four", "Five", "Six")
Sed.groups.prop

clus$Pond <- row.names(clus)
clus$Pond <- substring(clus$Pond, 1,1)
clus$Pond <- as.factor(clus$Pond)
counts <- table(clus$cluster, clus$Pond)

barplot(counts,  xlab="Pond", ylab="Number of Samples", col=c("red", "blue", "green", ”yellow",
"purple", "orange"))
legend("topright", legend=rownames(counts), fill=c("red", "blue", "green", "yellow",
"purple", "orange"), horiz=TRUE)
```

```
# Pie Chart with Percentages
slices <- c(31, 15, 41, 61, 218, 90)
lbls <- c("Pelletal-Aragonite Mud", "Aragonite Mud", "Pelletal Diatomaceous Sand",
        "Pelletal Skeletal Sand", "Pelletal Sand", "Pelletal Sand & Aragonite Mud")
pie(slices,labels = lbls, col= c("red", "blue", "green",
                          "yellow", "purple", "orange")) #rainbow(length(lbls)))
```

Water chemistry analysis

Raw data file (*data*) consists of the concentrations of sodium, magnesium, calcium, potassium, chloride, sulphate, ammonia, nitrite/nitrate and phosphate, in that order, in mg/L. The *data* data.frame has two additional columns for sediment cluster and pond identity. The *spa* data.frame contains the UTM coordinates for each sample. The data will need to be subset into pond groups for the IDW analysis.

```
library(compositions)
library(vegan)
library(MASS)
library(gstat)
library(sp)
library(lattice)

data <- read.csv("[….].csv", header=TRUE, row.names=1)

#Transform ionic concentrations into millequivalents
sal <- data[,c(1,2,3,4,5,6)]
sal[,1] <- (sal[,1]/(22.99/1))*1000
sal[,2] <- (sal[,2]/(24.31/2))*1000
sal[,3] <- (sal[,3]/(40.08/2))*1000
sal[,4] <- (sal[,4]/(39.10/1))*1000
sal[,5] <- (sal[,5]/(35.45/1))*1000
sal[,6] <- (sal[,6]/(96.06/2))*1000
sal <- round(sal, 0)

#Transform into composition data matrix
x <- acomp(sal, parts=c("Na", "Mg", "Ca", "K", "Cl", "SO4"))

#PCA_ Exploring Codependence in Salinity measures
pcx <- princomp(x) #Conduct PCA on Salinity Composition
sum(pcx$sdev[1:1]^2/mvar(x)) #Proportion of explained variance biplot captures

#Draw PCA biplot
opar <- par(mar=c(1,1,1,1))
dots <- rep(".", times=nrow(x))
biplot(pcx, xlabs=dots)
par(opar)

#Three groups in biplot. Measure proportionality of each group
sal.grp1 <- mvar(acomp(x[,c("Ca", "Mg", "Na")]))
sal.grp2 <- mvar(acomp(x[,c("Cl", "SO4")]))
sal.grp3 <- mvar(acomp(x[,c("K")]))
mvar(x)
(prop.grp1 <- sal.grp1/mvar(x))
(prop.grp2 <- sal.grp2/mvar(x))
(prop.grp3 <- sal.grp3/mvar(x))

#Scree and form plots and variation explained by PCA
plot(pcx, type="screeplot")
```

```
plot(pcx, type="variance")
sum(pcx$sdev[1]^2/sum(pcx$sdev^2)) #First two variances against total variation

#Loadings of ions against PC axes
loadings(pcx)[,1:2]
colSums(loadings(pcx))

comprel2d <- function(data, fixedvar){
  diag(1/data[,fixedvar]) %*% unclass(data)
}
comprel1d <- function(data, fixedvar){
  unclass(data)/data[fixedvar]
}

princ.comp <- 1
element <- "Cl"
fk <- pcx$scores[,princ.comp]
vd <- pcx$Loadings[princ.comp,]*pcx$sdev[1]
vd <- comprel1d(vd, element)
mn <- pcx$Center
mn <- comprel1d(mn, element)
matplot(fk, log(comprel2d(x, element)), pch=19, col=rainbow(10))
for(i in 1:6){
  abline(a=log(mn[i]), b=log(vd[i]), col=rainbow(10)[i], lwd=2)
}
fkdens <- seq(from=min(fk)*1.1, to=max(fk)*1.1, length.out=200)
compdens <- clrInv(outer(fkdens, clr(vd))) + pcx$Center
compdens <- comprel2d(compdens,element)
etqy <- compdens[length(fkdens),]
par(mfrow=c(1,2), mar=c(3,3,1,1))
for(logscale in c("","y")){
  matplot(fk, comprel2d(x,element),
          pch=19, col=rainbow(10), log=logscale, cex=0.75)
  matlines(fkdens, compdens, lty=1, col=rainbow(10))
}
text(x=fkdens[length(fkdens)], y=etqy,
     labels=colnames(x), pos=2)

#Cluster Analysis: Detecting Natural Groups
xc = acomp(x)
dd = dist(xc)
hc = hclust(dd, method="ward")
plot(hc)
h = locator(1)$y
rect.hclust(hc, h=h)
gr = cutree(hc, h=h)
plot(x,col=gr)
dd = as.dist(variation(x))
hc = hclust(dd, method="ward")
plot(hc)

#Linear Discriminant Analysis
res <- lda(x=data.frame(ilr(x)), grouping=pond)
ilrInv(res$means, orig=x)
V <- ilrBase(x)
```

```
rownames(V) <- colnames(x)
t(ilr2clr(t(res$scaling),V=V))
grint <- as.integer(pond)
pairs(res, abbr=1, col=1:8[grint], cex=1.2)

#Inverse distance weighting maps
meq.sp <- cbind(sal, spa)
coordinates(meq.sp) <- c("x", "y")
x.min <- round(min(spa$x), 0)
x.max <- round(max(spa$x), 0)
y.min <- round(min(spa$y), 0)
y.max <- round(max(spa$y), 0)
nodes <- 50
spa.grid <- (cbind(seq(x.min, x.max, by=((x.max-x.min)/nodes)), seq(y.min, y.max, by=((y.max-
y.min)/nodes))))
spa.grid <- expand.grid(spa.grid[,1], spa.grid[,2])
spa.grid$x <- spa.grid$Var1
spa.grid$y <- spa.grid$Var2
spa.grid$Var1 <- NULL
spa.grid$Var2 <- NULL
spa.grid$fict <- rnorm(nrow(spa.grid), 34000,88541)
coordinates(spa.grid) <- c("x", "y")
spa.grid <- as(spa.grid, "SpatialPixelsDataFrame")

Na.idw <- idw(Na ~ 1, meq.sp, spa.grid)
Mg.idw <- idw(Mg ~ 1, meq.sp, spa.grid)
Ca.idw <- idw(Ca ~ 1, meq.sp, spa.grid)
K.idw <- idw(K ~ 1, meq.sp, spa.grid)
Cl.idw <- idw(Cl ~ 1, meq.sp, spa.grid)
SO4.idw <- idw(SO4 ~ 1, meq.sp, spa.grid)

par(mfrow=c(2,3))
greys = grey.colors(4, 0.55, 0.95)
image(Na.idw, main="Sodium", col=greys, breaks=seq(400,520, 30), axes=TRUE)
legend("topleft", legend = c("400-430", "430-460", "460-490", "490-520"),
    fill = greys, bty = "n", title = "meq")

greys = c("#8C8C8C", "#B0B0B0", "#CDCDCD")
image(Mg.idw, main="Magnesium", col=greys, breaks=seq(80,110, 10), axes=TRUE)
legend("topleft", legend = c("80-90", "90-100", "100-110"),
    fill = greys, bty = "n", title = "meq")

greys = c("#8C8C8C", "#B0B0B0", "#CDCDCD")
image(Ca.idw, main="Calcium", col=greys, breaks=seq(16,25, 3), axes=TRUE)
legend("topleft", legend = c("16-19", "19-22", "22-25"),
    fill = greys, bty = "n", title = "meq")

greys = grey.colors(4, 0.55, 0.95)
image(K.idw, main="Potassium", col=greys, breaks=seq(7,11, 1), axes=TRUE)
legend("topleft", legend = c("7-8", "8-9", "9-10", "10-11"),
    fill = greys, bty = "n", title = "meq")

greys = grey.colors(4, 0.55, 0.95)
image(Cl.idw, main="Chloride", col=greys, breaks=seq(345,765, 105), axes=TRUE)
legend("topleft", legend = c("345-450", "450-555", "555-660", "660-765"),
```

```
    fill = greys, bty = "n", title = "meq")

greys = grey.colors(4, 0.55, 0.95)
image(SO4.idw, main="Sulphate", col=greys, breaks=seq(35,95, 15), axes=TRUE)
legend("topleft", legend = c("35-50", "50-65", "65-80", "80-95"),
    fill = greys, bty = "n", title = "meq")

SpatialPolygonsRescale(layout.scale.bar(), offset = locator(1), scale = 20, fill = c("transparent", "black"),
plot.grid = FALSE)
text(locator(1), "0")
text(locator(1), "20 m")
par(mfrow=c(1,1))

#LDA on water chemistry and nutrients using sediment type as group
clus <- as.factor(data$clus)
x2 <- acomp(data, c(1:9))
table(clus)

x2.ilr <- data.frame(ilr(x2))

x2.dist <- dist(x2.ilr)
(x2.MHV <- betadisper(x2.dist, clus))
anova(x2.MHV)
permutest(x2.MHV) #Not homogenous but close enough

res2 <- lda(x2.ilr, grouping=clus)

round(ilrInv(res2$means, orig=x2), 10)
V2 <- ilrBase(x2)
rownames(V2) <- colnames(x2)
t(ilr2clr(t(res2$scaling), V=V2))

grint2 <- as.integer(clus)
pairs(res2, abbr=1, col=(1:6)[grint2], cex=1.2)

grw <- as.numeric(clus)
plot(res2$scores , asp=1, xlab="PC1", ylab="PC2", add=TRUE)
k <- length(levels(factor(grw)))
for (i in 1:k) {
  points(pcx$scores[grw==i,1], pcx$scores[grw==i,2], pch=21, cex=1,
      col="black", bg=i+1)
}
legend("bottomright", paste("Pond", 1:k), pch=21,
          col="black", pt.bg=2:(k+1), pt.cex=2, bty="n")
plot(envfit(pcx$scores, sal, Permu=999), col="black")
```

Biological analysis

This script is a summary of the analysis used for understanding the structure of the bacteria and ciliate chapters. It uses three sets of data: bacterial or ciliate community matrices (*com*), UTM coordinates for each sample (*com.spa*) and a set of environmental determinants (*expl*) which consists of water chemistry, nutrient, sediment and morphotype data.

```
library(vegan)
library(compositions)
library(PCNM)
library(packfor)
source("create.MEM.model.R") # www.davidzeleny.net/anadat-
        r/doku.php/en:numecolr:create.mem.model

#Import data and define env, pond and morph data.frames
com <- read.csv("[….].csv", row.names=1, header=TRUE)
com.spa <- read.csv("[….].csv", row.names=1, header=TRUE)
expl <- read.csv("[….].csv", row.names=1, header=TRUE)
pond <- data.frame(substr(row.names(com), 1,1))
morph <- data.frame(expl[,12])
rownames(morph) <- rownames(expl)
rownames(pond) <- rownames(expl)
sed <- data.frame(as.factor(expl$clus))
wat <- data.frame(acomp(expl, parts=c("Na", "Mg", "Ca", "K", "Cl", "SO4", "NH4", "NOx", "PO4")))
env <- cbind(wat, sed)
names(env)[10] <- "Sed"
com <- round(com, 0)
com <- com[,!apply(com==0,2,all)]
com.pond <- data.frame(substr(row.names(com), 1,1))

# Basic Functions
ncol(com) #Number of OTUs
colnames(com) #OTU names
rownames(com) #Site names
summary(com) #All OTUs non-symmetrical in abundance due to large number of zeros in dataset

# Overall distributions of abundances (dominance codes)
#Minimum and maximum of abundance values in the whole dataset
range(com)
ab <- table(unlist(com))
ab
# Barplot of the distributions, all species confounded
barplot(ab, las=1, xlab="Abundance class", ylab="Frequency", col=gray(5:0/5)) #Highly skewed towards
0
# Number of absences
sum(com==0)
# Proportion of zeros in the data set
sum(com==0)/(nrow(com)*ncol(com))

# Compare species: number of occurrences
```

```
com.pres <- apply(com>0, 2, sum)
# Sort in increasing order
sort(com.pres)
# Compare percentage frequencies
com.relf <- 100*com.pres/nrow(com)
round(sort(com.relf), 1)
#Plot histogram
par(mfrow=c(1,2))
hist(com.pres, main="OTU Occurrences", right= FALSE, las=1, xlab="Number of occurrences",
ylab="Number of OTUs",
   breaks=seq(0,480,by=20), col="bisque")
hist(com.relf, main="Species Relative Frequencies", right=FALSE, las=1, xlab="Frequency of occurrence
(%)", ylab=
     "Number of OTUs", breaks=seq(0,100, by=10), col="bisque")

#Diversity Summary
com.tot <- data.frame(t(colSums(com)))
N0 <- rowSums(com.tot>0)  #Species richness
H <- diversity(com.tot)   #Shannon entropy
N1 <- exp(H)      #Shannon diveristy number
N2 <- diversity(com.tot, "inv") #Simpson diveristy number
J <- H/log(N0)  #Pielon evenness
E1 <- N1/N0 #Shannon evenness
E2 <- N2/N0 #Simpson evenness
(div.tot.com <- data.frame ((N0), (H), (N1), (N2), (E1), (E2), (J)))
sac <- specaccum(com) #RAREFACTION/ Species ACCUMULATION
#Betadiversity 'z' values
com.z <- betadiver(com, "z")
quantile(com.z)
mean(com.z)
z.mod <- betadisper(com.z, pond[,1])
z.mod
boxplot(z.mod)

#Linear Regression of TAR
richness <- rbind(A.tot.div.com[,1], C.tot.div.com[,1], D.tot.div.com[,1], H.tot.div.com[,1],
J.tot.div.com[,1],
            P.tot.div.com[,1], S.tot.div.com[,1], W.tot.div.com[,1])
log.rich <- log10(richness)
size <- rbind(6715, 7147702, 5978, 65069, 187509, 14880, 95, 46544) #Areas of each pond
size.richness <- data.frame(cbind(log.rich, size))
size.richness$LogSize <- log10(size.richness[,2])
size.richness.lm <- lm(log.rich ~ LogSize, size.richness)
summary(size.richness.lm)
plot(size.richness$X1~size.richness$LogSize, ylab="OTU richness", xlab="Pond Size (log m2)", type="n")
points(size.richness$X1~size.richness$LogSize, pch=3, cex=1.5, lwd=0.5)
axis.values <- par("usr")
coef <- data.frame(size.richness.lm$coefficients)
temp <- data.frame(x=seq(axis.values[1]+1.5, axis.values[2]-.1, length=20))
temp$y <- temp$x*coef[2,1]+coef[1,1]
lines(y~x, data=temp, col="black", lwd=2)

# Mantel correlogram
com.hel <- decostand(com, "hellinger") # Perform Hellinger transformation
# Data is first detrended
```

```
anova(rda(com.hel, com.spa))
com.hel.det <- resid(lm(as.matrix(com.hel)~., data=com.spa))

com.h.D1 <- dist(com.hel.det)
(com.correlog <- mantel.correlog(com.h.D1, XY=com.spa, cutoff=FALSE, nperm=999))
plot(com.correlog)

#MDS
com.nmds <- metaMDS(vegdist(decostand(com, "hellinger")), "euclidean")
plot(com.nmds, type="t", main=paste("NMDS/BRay Lp1 - Stress =", round(com.nmds$stress, 3)))

par(mfrow=c(1,1))
com.Pond <- data.frame(substr(row.names(com), 1,1))
names(com.Pond)[1] <- "Pond"
grw <- as.numeric(com.Pond$Pond)
plot(com.nmds$points , asp=1, xlab="MDS1", ylab="MDS2", add=TRUE)
k <- length(levels(factor(grw)))
for (i in 1:k) {
  points(com.nmds$points[grw==i,1], com.nmds$points[grw==i,2], pch=21, cex=1,
       col="black", bg=i+1)
}
ordispider(com.nmds, com.Pond$Pond, label=TRUE)
legend("bottomright", paste("Pond", 1:k), pch=21,
      col="black", pt.bg=2:(k+1), pt.cex=2, bty="n")

morph.nmds <- metaMDS(com.hel, "euclidean")
par(mfrow=c(1,1))
grw <- as.numeric(morph2[,1])
plot(morph.nmds$points , asp=1, xlab="MDS1", ylab="MDS2", add=TRUE)
k <- length(levels(factor(grw)))
for (i in 1:k) {
  points(morph.nmds$points[grw==i,1], morph.nmds$points[grw==i,2], pch=21, cex=1,
       col="black", bg=i+1)
}
ordispider(morph.nmds, morph2$morph2, label=TRUE)

#PERMANOVA
com.perm <- vegdist(decostand(com, "hellinger"), "euclidean")
perm.res <- adonis(com.perm~Pond, data=com.Pond)
perm.beta <- with(com.Pond, betadisper(com.perm, Pond))
plot(perm.beta)
TukeyHSD(perm.beta)

#ANOSIM
anosim(vegdist(decostand(com, "hellinger"), "euclidean"), com.Pond$Pond)
adonis(vegdist(com.hel, "euclidean")~morph2, data=morph2)
with(morph2, betadisper(vegdist(com.hel, "euclidean"), morph2))
plot(with(morph2, betadisper(vegdist(com.hel, "euclidean"), morph2)))
TukeyHSD(with(morph2, betadisper(vegdist(com.hel, "euclidean"), morph2)))
anosim(vegdist(com.hel, "euclidean"), morph2$morph2)

# RDA of Hellinger-transformed OTU data by environmental data
com.hel <- decostand(com2, "hellinger")

com.rda <- rda(com.hel ~., cbind(com.env, morph2)) # Sediment Dummy variable
```

```
summary(com.rda)
coef(com.rda)

# Retrieval of adjsuted R2
(R2 <- RsquareAdj(com.rda)$r.squared) #Unadjusted R@ from rda result
(R2adj <- RsquareAdj(com.rda)$adj.r.squared) # Adjusted R2 retrieved from rda object

# Global test of RDA result
anova.cca(com.rda, stop= 1000)
# Test of all canonical axes
anova.cca(com.rda, by="axis", step=1000)

com.rda.summary <- summary(com.rda)
com.rda.summary[15]

#Forward selection of environmental variables using ordiR2step
com.env.com <- cbind(com.env, morph2)
rda.com.0 <- rda(com.hel ~ 1, data=com.env.com)
rda.com.all <- rda(com.hel ~ ., data= com.env.com)
ordiR2step(rda.com.0, scope=formula(rda.com.all), direction="forward")
temp <- rda(com.hel~., com.env.com)
RsquareAdj(temp)
forward.sel(com.hel, model.matrix(~.,com.env.com)[,-1], adjR2thresh=RsquareAdj(temp))
plot(com.rda2)
com.env.com2 <- model.matrix(~.,com.env.com)[,-1][,c(15,16,17,10,8,13,14,12,2,3,6,1,4,7)]

#RDA DONE USING NEW VARIABLES
com.rda2 <- rda(com.hel~com.env.com2)
summary(com.rda2)
com.rda2
(R2 <- RsquareAdj(com.rda2)$r.squared) #Unadjusted R2 from rda result
(R2adj <- RsquareAdj(com.rda2)$adj.r.squared) # Adjusted R2 retrieved from rda object

#MEM Analysis
ngr=8 # Number of groups in data
nsites.per.group= c(40,52,42,55,51,44,40,60) #Insert number of observation of for each pond
MEM <- create.MEM.model(com.spa, ngroups=ngr, nsites=nsites.per.group)
MEM <- data.frame(MEM)
summary(MEM)
str(MEM)

# Variation partitioning
pond.group <- model.matrix(~., pond2)[,-1]
varp <- varpart(com.hel, com.env.com2, pond.group, MEM)
plot(varp, digits=2)

#Marginal Effects
env.rda <- rda(com.hel ~ com.env.com2)
anova.cca(env.rda)
RsquareAdj(env.rda)
MEM.rda <- rda(com.hel ~., MEM)
anova.cca(MEM.rda)
RsquareAdj(MEM.rda)
pond.rda <- rda(com.hel ~ pond.group)
anova.cca(pond.rda)
```

```
RsquareAdj(pond.rda)

#2 way Conditional Effects
wipond.varpart.ad <- rda(com.hel, com.env.com2, pond.group)
anova.cca(wipond.varpart.ad)
RsquareAdj(wipond.varpart.ad)
wipond.varpart.af <- rda(com.hel, com.env.com2, MEM)
anova(wipond.varpart.af)
RsquareAdj(wipond.varpart.af)
wipond.varpart.bd <- rda(com.hel, MEM, pond.group)
anova(wipond.varpart.bd)
RsquareAdj(wipond.varpart.bd)
wipond.varpart.be <- rda(com.hel, MEM, com.env.com2)
anova.cca(wipond.varpart.be)
RsquareAdj(wipond.varpart.be)
wipond.varpart.ce <- rda(com.hel, pond.group, com.env.com2)
anova(wipond.varpart.ce)
RsquareAdj(wipond.varpart.ce)
wipond.varpart.cf <- rda(com.hel, pond.group, MEM)
anova(wipond.varpart.cf)
RsquareAdj(wipond.varpart.cf)

#3 way Conditional Effects
wipond.varpart.a <- rda(com.hel, com.env.com2, cbind(MEM,pond.group))
anova.cca(wipond.varpart.a)
RsquareAdj(wipond.varpart.a)
wipond.varpart.b <- rda(com.hel, MEM, cbind(com.env.com2,pond.group))
anova(wipond.varpart.b)
RsquareAdj(wipond.varpart.b)
wipond.varpart.c <- rda(com.hel, pond.group, cbind(com.env.com2,MEM))
anova(wipond.varpart.c)
RsquareAdj(wipond.varpart.c)
```