Climate change since the Last Interglacial in northern New Zealand inferred from pollen and chironomid records of the Auckland Maars

by

Valerie van den Bos

A thesis submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Victoria University of Wellington
2019

Abstract

In light of contemporary climate change it is more important than ever to understand past shifts in climate, especially past warm phases, and their effects on ecosystems and societies. From compilations of global climate reconstructions, several periods have been identified that might have been warmer than today, the two most recent of which are the Holocene Thermal Maximum (~11 – 5 kyr BP) and the Last Interglacial (~129 – 116 kyr BP). However, spatio-temporal complexities are typically smoothed out in global climate reconstructions and we do not have a good understanding of the regional differences in past climate. The southern mid-latitudes especially are underrepresented in palaeoclimate research.

For this thesis I analyse the sediments from two maars within the Auckland Volcanic Field: Orakei Basin, which erupted ~126.0 kyr BP and accumulated sediments until ~9 – 8.5 kyr BP; and Lake Pupuke, which still contains a lake today and therefore covers the Holocene. Quantitative climate reconstructions are necessary to put the Orakei Basin and Lake Pupuke records in a broad context and to enable comparisons of past and future climates. For this study I focus on biological proxies preserved by lake sediments, namely pollen, which primarily responds to mean annual air temperatures (MAAT), and chironomids, a surrogate for summer air temperatures (SmT). Together, MAAT and SmT reconstructions from the same site can provide insight into changing seasonality over time, an underexplored dimension of proxy-based reconstructions. The chironomid record covers just the last ~14 cal kyr BP however, because of low head capsule abundances in older sediment sections.

The Orakei Basin pollen record and associated MAAT reconstruction cover ~85 to 9 cal kyr BP and show five distinct phases comparable to Marine Isotope Stages (MIS) 5 to 1. This association is confirmed by the preliminary tephrochronology of the core. The broad similarity of the Orakei MAAT trend to the MIS and other records from New Zealand implies all were driven by northern high-latitude summer insolation, consistent with the Milankovitch orbital forcing hypothesis. Several patterns superimposed on the general trend stand out: first, MIS 4 is a brief cool period, which is inconsistent with the observation that glacier advances equivalent to those of the late last glacial maximum occurred ~65 kyr BP in the Southern Alps, possibly due to the seasonal distribution of energy from solar insolation. Second, MIS 3 displays an earlier warm phase followed by a progressive cooling trend which

might be correlated to decreasing local summer insolation intensity. Third, glacial conditions of MIS 2 appear consistent with the early onset of the last glacial maximum in the southern mid latitudes, which was likely driven by regional insolation intensity.

The Lake Pupuke pollen and chironomid records, covering the last ~14 cal kyr BP, show no evidence of a past warm period equivalent to the Holocene Thermal Maximum. MAAT is stable throughout the Holocene, whereas SmT increases between 10 and 3 cal kyr BP. The latter shows a strong relationship with integrated local summer insolation. The temperature reconstructions lead to the conclusion, first, that seasonality was low during the Early Holocene (12 to 9.3 cal kyr BP), and second, that during mid-to-late Holocene (after ~7 cal kyr BP) summers were hot and dry, allowing the tall conifer kauri to expand throughout northern New Zealand.

The Lake Pupuke chironomid-SmT reconstruction highlighted an issue with the transfer function model, namely, that it was not able to reconstruct values close to modern day (18.9°C). Therefore, I explore an extended training set which encompasses a longer temperature gradient. New models are fitted using both traditional techniques and modern machine learning methods. The new model improves the SmT reconstruction from Lake Pupuke, in the sense that reconstructed temperatures now reach modern day values. However, the SmT trend is the same as the original trend, substantiating the previously drawn conclusions.

During the course of this research, I discovered that density separation during pollen preparation can lead to varying relative abundances, depending on the specific gravity used. After some experimentation I found that using a low specific gravity (2.0; recommended value in the literature) can result in the overrepresentation of buoyant pollen grains, leading to erroneous interpretations.

Together, these results point out the importance of considering regional-to-local drivers of climate changes superimposed on global reconstructions. Multi-proxy records can help disentangle the different aspects of the climate system, where especially chironomids can be helpful to elucidate the role of SmT and local summer insolation. Finally, this thesis shows the importance of questioning the appropriateness of conventional methodologies and where possible, addressing their limitations.

Acknowledgements

First, I would like to thank my supervisors, Rewi Newnham and Andrew Rees for their unlimited support over these last three years. You were a great team: Rewi with his vast research experience teaching me the New Zealand palynoflora, and Andrew instructing me in the arts of chironomid analysis and transfer function development. I am especially grateful for the extra effort you both put in over the last few months and dragging me over the finish line.

Second, I want to acknowledge all the people that contributed to this thesis and the Auckland Maars Marsden project: Leonie Peti for providing the preliminary lithology of Orakei Basin and for sharing her insights and ideas; Jenni Hopkins for analysing the tephra layers in the Orakei core and her patience in explaining the results to me; Janet Wilmshurst for reconstructing mean annual air temperatures from the Pupuke pollen data; Marcus Vandergoes and Lizette Reyes for providing much of the chironomid training set data; Paul Augustinus for leading the greater project; Jenni, Elaine Smid and Natasha Piatrunia for their help drilling Orakei Basin. I also want to thank the staff at SGEES for their support and assistance, including, but not limited to: Kosta Tashkoff, Dez Tessler, Miranda Voke, Monika Hanson and Emma Fisher. A special thanks to Jane Chewings for her help in the lab and calm responses to my chemical spills…

I am thankful for the funding provided by the Royal Society of New Zealand that made this research possible, for the financial support of the Victoria Doctoral Submission Scholarship, and for the Faculty Strategic Research Grant that allowed me to attend the European Geosciences Association general assembly in 2017 and Australasian Quaternary Society's biennial meeting in 2018.

I also want to thank the friends I made over the last three years at Vic that made this experience enjoyable: Aidan, Ben, Ben, Ignacio, Jürgen, Lisa, Lou, Luisa, Lynda, Marcel, Matt, Sandra, Shaun, Tom; Aline and Margaret for their companionship in the microscope lab; Wokje for sharing her home with me when I'd just arrived. Thanks to my office mates, old and new, for a great working environment and always making time for a chat: Chris Conway, Sam Webber, Céline Mandon, Leo Pure, Willy Henriquez Gonzalez, Kate Mauriohooho, Hannah Elms, Katie Jones and Elliot Swallow. Elliot, I was so grateful the first

time you invited me to the tea room for a coffee break and introduced me to everyone important. You were truly a good friend throughout this experience.

I am eternally grateful to my family, who have supported me always, even when I said I was moving to the other side of the world.

Last, but definitely not least, Joshua. Thank you for your patience, your support. Thank you for dealing with my moods. Thank you for picking me up from the city at all hours. Thank you for feeding me. Thank you for introducing me to the Jurys, my family away from family.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction to thesis

1.1 Background

In light of contemporary climate change it is more important than ever to understand past shifts in climate and their effects on ecosystems and societies. Along with historical data, which is comparatively limited in the New Zealand context, we have three main tools at our disposal to infer past climates: instrumental data, proxy records and model simulations. Instrumental data, however, do not reach far beyond the past century and are insufficient to sample the full range of climatic behaviour of our planet (Alley et al., 2003). Proxy records can provide valuable information on past climate and, in combination with model simulations, can improve our understanding of the mechanisms that drove climate changes (Renssen et al., 2005). Yet, model-data comparisons do not always agree (e.g. Liu et al., 2014).

Part of the dissensus between proxy records and model simulations may be due to the poor spatial resolution of continuous proxy records around the globe. Any spatio-temporal complexities in past climate change are typically smoothed out in global climate reconstructions, although we do not have a good understanding of these complexities. In addition, some proxies may be more sensitive to seasonal climate parameters even though they are assumed to represent annually averaged climate in model-data comparisons.

In the context of contemporary and future climate change, the inability of models to simulate climate reconstructions from proxy data for warm climate intervals is especially problematic. This problem highlights the importance of exploring more proxy reconstructions for recent warm intervals, in particular the Holocene Thermal Maximum (HTM; ~11 – 5 kyr BP) and the Last Interglacial (LIG; ~129 – 116 kyr BP), that may be partly analogous to climate scenarios projected for the coming decades (Masson-Delmotte et al., 2013; Fischer et al., 2018).

Continuous records that reach back to the LIG or beyond are sparse, as sites where we may expect continuous deposition over thousands of years are rare wordwide. At high latitudes and in high alpine regions, long-term continuous ice deposition has taken place from which we are able to produce very detailed climate reconstructions (e.g. Petit et al., 1999; Kindler et al., 2014). However, their obvious limitation is spatial resolution. Long marine cores can fill in the gaps in the oceans but are often of low temporal resolution due to low sedimentation rates, and signals are frequently smeared by bioturbation. Terrestrial records such as lake

sediments can be of higher resolution than marine cores, but continuous deposition since the LIG is rare, as terrestrial basins can be overtopped by ice during glaciations, lakes can disappear during dry episodes, or basins can fill up with sediments. The southern mid-latitudes especially are underrepresented in global climate change compilations. New Zealand is one of the few landmasses in the southern mid-latitudes and covers both subtropical and subantarctic climatic regimes. There is one continuous record from southern New Zealand that covers the entire period from present to the LIG and beyond, i.e. Okarito Pakihi Bog (Vandergoes et al., 2005; Newnham et al., 2007a). The scarcity of information from terrestrial New Zealand provides an initial motivation for this thesis, which sets out to advance understanding of past climate change in northern New Zealand in the context of global events since the LIG.

In this study I analyse the sediments from two maars within the Auckland Volcanic Field (AVF; Fig. 1.1). The AVF comprises 53 Quaternary volcanic centres (Fig. 1.1b; Hopkins et al., 2015), each of which has resulted from a single eruption or short sequence of eruptions. The lakes that developed inside the volcanic craters form an excellent depositional environment for the accumulation of lacustrine sediments. Here we focus on Orakei Basin (Fig. 1.1c), which erupted ~126.0 kyr BP (Hopkins et al., 2017) and accumulated sediments until ~9 – 8.5 kyr BP (Hayward et al., 2008), and Lake Pupuke (Fig. 1.1d), which has an estimate age of ~193.2 kyr BP (Leonard et al., 2017) and is the only crater in the AVF that contains a lake today. The sediment core retrieved from Lake Pupuke spans the last ~ 48.2 cal kyr (Stephens et al., 2012a) but only the last 16 cal kyr BP are analysed for this study. The starting premise for this research was that the sediment cores from Orakei and Pupuke could provide a climate record from the LIG to present.

Quantitative climate reconstructions are necessary to put the Orakei Basin and Lake Pupuke records in a broader context and to enable more direct comparison between past and future climate. For this study I focussed on pollen and chironomids. From pollen we can reconstruct mean annual air temperatures (MAAT) using an existing transfer function for New Zealand (Wilmshurst et al., 2007). From chironomids we can reconstruct summer air temperatures (SmT). A published chironomid-SmT transfer function already exists for the South Island of New Zealand (Dieffenbacher-Krall et al., 2007) but, as part of this research, has been extended to be more appropriate for reconstructions in the North Island. Together, MAAT and SmT reconstructions from the same site can provide important insight into changing

seasonality over time, an often-ignored though important aspect of climate change discussions.
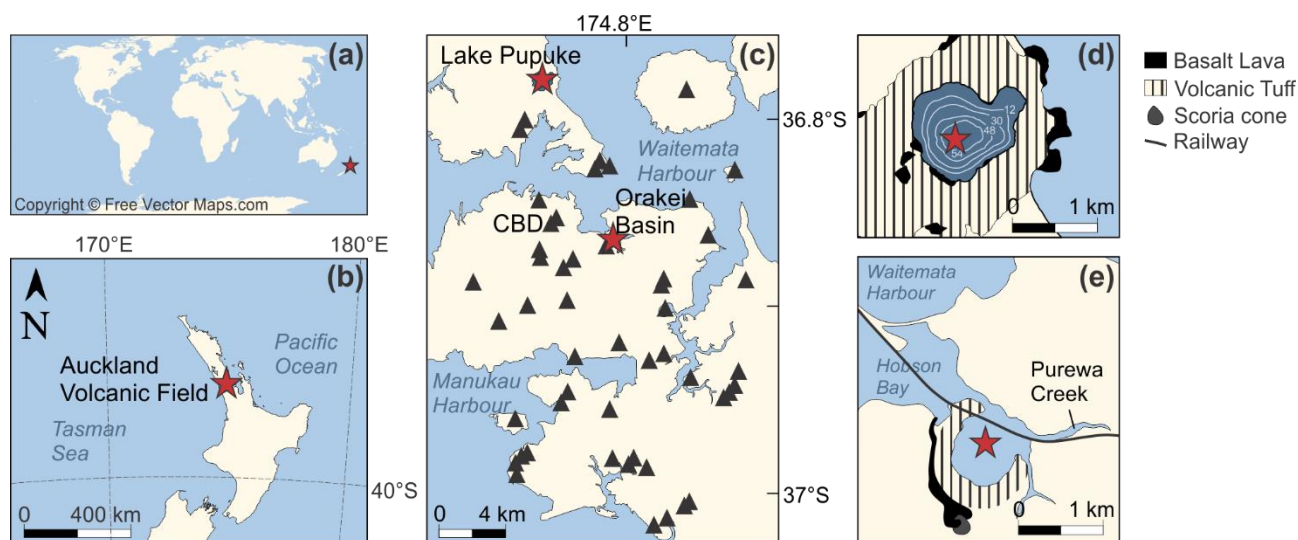
Figure 1.1: Position of New Zealand in the world (a), location of AVF within New Zealand (b), basaltic centres within the AVF (adapted from Hopkins et al. 2017) (c), map of Lake Pupuke (d) and Orakei Basin (e). Stars indicates study site/core location. Triangles indicate volcanic centres. CBD = Auckland's central business district.

1.2 The Auckland Maars Marsden project

This study is part of a broader investigation, funded by the Royal Society of New Zealand's Marsden fund (project 14-UOA-040), into the Auckland maars and the nature and drivers of Late Quaternary climate change in northern New Zealand (hereafter the Auckland Maars Marsden project). This project is divided into five steps: extracting and describing the cores (1), dating the records (2), producing multi-proxy records of past climate (3), sediment geochemistry (4), and studying the lake sediment laminae (5). This thesis fulfils part of the third step. In addition, I participated in part of step 1 (i.e. extracting the Orakei Basin core).

Unfortunately, due to delays from operational problems, not all steps of the broader Marsden project had been met at the time this thesis was required to be submitted. In particular, slow progress with step 2 (dating the records) has presented a fundamental challenge to the timely completion of this thesis. To some extent we have been able to overcome this problem through the application of tephrochronology, drawing on previous published work on the Auckland maar sediments. For the Holocene Lake Pupuke record, this solution was appropriate as the tephrochronology for the core was already well-established. For the Orakei Basin record, however, it was necessary to develop a preliminary tephrochronology

as part of this thesis, with the expectation that a better resolved chronology will eventually be produced by the wider Marsden team in the fullness of time. For this reason, it was decided not to submit the Orakei Basin palynology record for publication until the chronology of the sequence is properly resolved. This strategy has enabled me to complete my PhD thesis on time.

It is also necessary to clearly indicate which parts of the thesis are not my own work. In subsequent sections and chapters I will indicate which information was received from other researchers and which information is unavailable at present but is anticipated in the near future.

1.3 Thesis objectives

The overarching goal of this thesis is to elucidate patterns of climate change in northern New Zealand during the late Quaternary in relation to globally recognised patterns of climate change. This goal is addressed under the first three thesis objectives:

1. Produce pollen and chironomid records from the Pupuke (a) and Orakei (b) sediments

2. Reconstruct MAAT and SmT from the pollen and chironomid results respectively for Pupuke (a) and Orakei (b)

3. Compare the Pupuke (a) and Orakei (b) records to other records from the region, from southern New Zealand and to global climate reconstructions, with special interest for potential analogues of future climate.

It is rare for a doctoral thesis to proceed in strict accordance with its original expectations. In addition to the delays in receiving necessary chronological information from the Auckland Maars Marsden project discussed above (section 1.2), two further unanticipated problems have necessitated revisions to these original thesis objectives. The first arises from an important finding of this thesis: that the Orakei Basin core is unlikely to extend back to the LIG. This unanticipated outcome means that I cannot assess the LIG as a potential analogue for future climate. The second problem arose from another important finding of this thesis:

that the concentration of chironomid head capsules is too low in the Orakei Basin sediments to develop a viable chironomid record for that site. Potential causes for this low concentration are discussed in Appendix E. In addition, I discovered two significant limitations in the methodologies being employed during the course of the thesis work: that a key step in my pollen preparation procedure, following recommended practice, was flawed; and that the existing New Zealand chironomid-SmT model was not optimal for application to northern New Zealand Holocene assemblages.

As a result of these discoveries, considerable effort was applied to finding viable solutions and the following thesis objectives were added to the original three:

4. Review the density separation procedure in pollen slide preparation and find the minimum required specific gravity that results in representative samples

5. Develop an improved chironomid-SmT model based on an extension of the modern training set, using both traditional and novel methods

1.4 Thesis structure

Another consequence of these unanticipated methodological issues is that the order of chapters in this thesis is not the chronological order in which the research was conducted. Section 1.4.1 describes the content of each chapter in this thesis. The timeline of my PhD research and the rationale behind each step, resulting in the sequence of chapters in this thesis, is described in section 1.4.2. Publications that came out of the work on this project, and the contributions of other authors are mentioned in section 1.4.3.

1.4.1 Description of chapters

This thesis consists of 6 chapters. Chapter 1 explains the overall rationale and motivation for the thesis and presents explicit thesis objectives in the context of previous work. Chapters 2 to 5 are written in "publication style", meaning that they can be understood independently of the context of the rest of the thesis. This format reflects the importance of publishing the original results of a doctoral thesis in the context of time and funding constraints. Chapter 6 ("Synthesis") starts with a compilation and summary of all produced

results, followed by a discussion of the objectives and goal of this thesis in light of these results, and recommendations for future research that follow from this discussion.

Each of the publication-style chapters includes an introduction, methods section, results section and discussion. Therefore, there may be some repetition here and there, e.g. a description of the Auckland Volcanic Field is included in both chapters 3 and 5. To avoid even more repetition, it was decided to not include a general "Methods" chapter to describe coring, sampling, laboratory procedures, etc.; these methods are described in each respective chapter as appropriate.

Chapters 2 and 4 can be defined as methodological, i.e. Chapter 2 seeks to improve the pollen preparation method, and Chapter 4 develops the temperature models derived from the chironomid training set. Chapter 3 describes the results of pollen and chironomid analyses conducted on the Holocene sediments of Lake Pupuke, and Chapter 5 describes the results of pollen analyses conducted on the sediments of Orakei Basin that reach back to MIS 5.

1.4.2 Explanation of the order of chapters

This section explains the order in which the four publication-style chapters are presented in the thesis and why this is different from the order in which the work was undertaken. Of the four publication-style chapters, the work presented in Chapter 3 on Lake Pupuke was completed first, followed by the work presented in Chapter 2 on the pollen preparation method. The need to revise the pollen preparation method was not originally anticipated and arose from concerns about the initial pollen results for Lake Pupuke. After some tests were conducted, as described in Chapter 2, all Pupuke pollen samples were re-analysed using the new method. Only the results of the new analyses are described in Chapter 3.

Work on the chironomid temperature models, presented in Chapter 4, was begun after the Holocene Pupuke record (presented in Chapter 3) had been submitted for publication. During the process of publishing that work and responding to reviewers' comments, the need to improve chironomid-SmT modelling for northern New Zealand sites was highlighted. The original model had been developed primarily for cooler regions in New Zealand and was not able to accurately reconstruct the higher temperatures that were likely to have been

experienced in the study area. This realisation generated the impetus for the work presented in Chapter 4. The impact of the new chironomid-SmT model on the interpretation of the Lake Pupuke results is discussed in Chapter 4. Chapter 5, on the Orakei Basin record, was written after Chapters 2 to 4 were drafted. The synthesis, Chapter 6, was written last.

1.4.3 Publications and contributions from other authors

I am the primary author of all chapters in this thesis and the sole author of Chapters 1 and 6 and all Appendices. My contributions include: drilling of Orakei Basin and preliminary description of cores, collection and analysis of all fossil pollen and chironomid data from Orakei Basin and Lake Pupuke, testing of different pollen preparation methods, analysis of 15 chironomid training set samples, and data analysis of the extended chironomid training set. Although the results of this work are my personal contributions, I will use the pronoun "we" throughout Chapters 2 to 5 in favour of "I" to acknowledge the contributions of my co-authors on these publication-style chapters.

My supervisors are Rewi Newnham and Andrew Rees, who both reviewed and edited all chapters and provided general guidance throughout the project, with Rewi Newnham focussing mainly on pollen analyses and Andrew Rees focussing mainly on chironomid analyses. As mentioned earlier, this thesis is part of a greater project on the Auckland Maars, funded by the Royal Society of New Zealand's Marsden Fund. Associate Professor Paul Augustinus (The University of Auckland) is the primary investigator for this larger project. Other researchers that contributed to this project in relation to my thesis are PhD candidate Leonie Peti (The University of Auckland), who wrote the description of Orakei Basin's lithology, and Dr Jenni Hopkins (Victoria University of Wellington), who analysed the tephra layers). Below I describe the contributions from other authors and status of publication per chapter.

Chapter 2: Density separation in pollen preparation: how low can you go?
Valerie van den Bos, Rewi Newnham, Andrew Rees, Lisa Woods

Submitted to Journal of Paleolimnology in November 2018. Rewi Newnham and Andrew Rees gave general advice and reviewed the text. Lisa Woods provided advice on statistical methods and reviewed all text relating to the data analysis.

Chapter 3: Holocene temperature, humidity and seasonality in northern New Zealand linked to Southern Hemisphere summer insolation

Valerie van den Bos, Andrew Rees, Rewi Newnham, Marcus Vandergoes, Janet Wilmshurst, Paul Augustinus

Published in Quaternary Science Reviews, Volume 201, pp. 77-88 (2018). Rewi Newnham and Andrew Rees gave general advice and reviewed the text. Marcus Vandergoes provided the original chironomid training set data. Janet Wilmshurst performed MAAT modelling. Paul Augustinus provided the Pupuke sediment cores.

Chapter 4: An extended chironomid training set for reconstructing New Zealand summer air temperatures

Not yet submitted for publication

Rewi Newnham and Andrew Rees gave general advice and reviewed the text. In addition, Andrew Rees provided extra guidance concerning data analysis and counted several chironomid samples. Marcus Vandergoes and his colleagues at GNS Science provided most of the chironomid training set data. Lizette Reyes (GNS Science) calculated the temperature and precipitation data used in this chapter. Paul Augustinus provided three lake sediment surface samples.

Chapter 5: A terrestrial palynological record for the last ~85 kyr from Orakei Basin, northern New Zealand

Not yet submitted for publication

This chapter will be submitted for publication after the age-depth model (to be produced by Leonie Peti) is finalised. Rewi Newnham and Andrew Rees gave general advice and reviewed the text. Leonie Peti provided lithological information and compiled the composite depths for the cores. Jenni Hopkins assisted with drilling and identified the rhyolitic tephra layers to support the preliminary age model for the record presented in this thesis. Janet Wilmshurst provided the pollen pre-deforestation training set. Paul Augustinus funded drilling.

Chapter 2

Density separation in pollen preparation:
how low can you go?
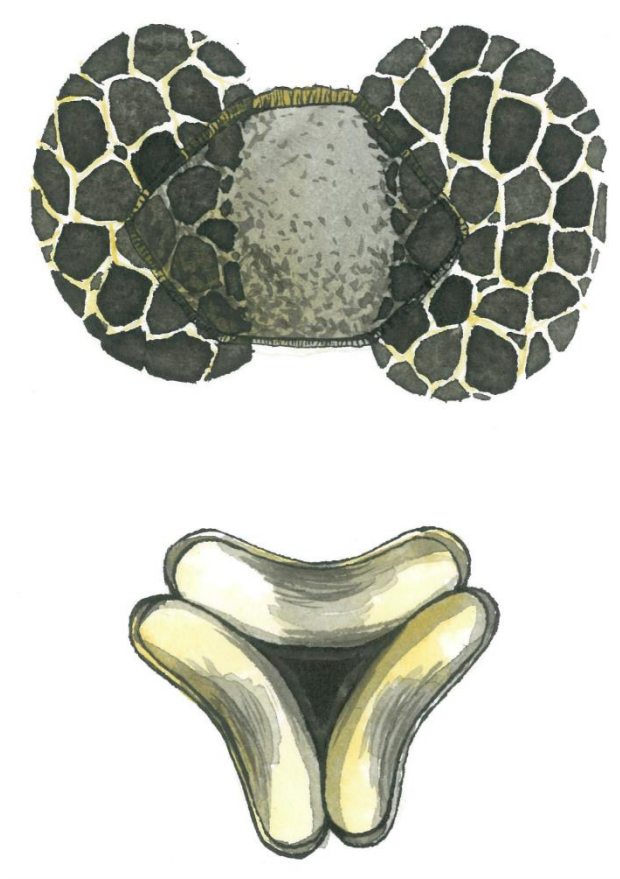
Preface

This chapter was submitted as a research paper to the Journal of Paleolimnology on 16-11-2018 and is currently under review. The results described in this chapter relate to objective 4: Review the density separation procedure in pollen slide preparation and find the minimum required specific gravity that results in representative samples. As mentioned in Chapter 1, there are three co-authors on this paper: Andrew Rees, Rewi Newnham and Lisa Woods.

Abstract

Palynological methods have evolved over time, but there is little literature on some of the newer practices that are now considered standard, such as density separation. In a study on the Holocene-age sediments of Lake Pupuke, northern New Zealand, we find that varying the density of the heavy liquid used to separate pollen from the minerogenic fraction within the range of reported practice affects outcomes: when a relatively low density is used (2.0 g/cm^3) buoyant pollen grains (such as *Prumnopitys taxifolia* and *Dacrydium cupressinum*) are overrepresented, while small, compact pollen grains (such as *Libocedrus* and *Metrosideros*) are underrepresented. In addition, exotic *Lycopodium* spores (which are added to allow calculation of concentrations) are underrepresented, leading to calculated concentrations that are too high. This result should raise wider concerns, as heavy liquid densities reported in literature range from 1.88 to 2.1 g/cm^3. We draw attention to this problem and recommend steps palynologists can take to ensure that their pollen assemblages are representative and do not lead to spurious interpretations.

2.1 Introduction

Palynology is a well-established method for reconstructing past vegetation and has been in use since the late 1800s (Faegri and Iversen 1989; Moore et al. 1991). Although palynology methods in general are widely reported, the literature has not always kept pace with evolving practice and this is currently the case for the density separation step in the procedure for preparing pollen slides. Density separation uses a heavy liquid to separate organic material including pollen by floatation from the minerogenic fraction of the sedimentary matrix. It is used increasingly instead of hydrofluoric acid (HF) (which dissolves silicates), which can pose a serious health and safety risk. Bromoform or zinc chloride were originally used to prepare the heavy liquid (Faegri and Iversen 1989; Moore et al. 1991), but most labs have shifted to the even safer sodium polytungstate (SPT).

To apply density separation successfully, the density of the heavy liquid should be sufficiently high for all pollen to float above the minerogenic fraction, but not so high as to retain a lot of extraneous organic matter. Although some studies show density separation with SPT produces the same or better results compared to HF treatment (Munsterman and Kerstholt 1996; Zabenskie 2006; Campbell et al. 2016), there have been no rigorous tests of the effect of varying heavy liquid density on the pollen yield. Reported densities range from 1.88 to 2.40 g/cm^3, which equates to a specific gravity (SG) of the same values (no unit). The recommended SG is typically relatively low, namely 1.88 to 2.1 (Munsterman and Kerstholt 1996; Zabenskie 2006; Campbell et al. 2016), to avoid retention of non-palynomorph organic matter. However, we are not aware of any reported testing of the pollen yields across this reported SG range.

Here, we describe the results of a study on the effect of varying heavy liquid density on pollen abundances. Surprisingly, we find that some pollen taxa are underrepresented when a low density within the reported range is used. As a result, relative abundances were skewed significantly, which could lead to erroneous interpretations of the vegetation composition. Exotic *Lycopodium* spores that are typically added to facilitate the calculation of concentrations are also affected, making all the concentrations erroneous. The purpose of this paper is to draw attention to this potentially widespread problem and to provide recommendations for palynologists to ensure that they are using the appropriate heavy liquid density and are not losing a portion of the pollen in their samples.

2.2 Methods

2.2.1 Material

For this study we used Holocene age sediment samples from Lake Pupuke, northern New Zealand (36°46'49"S, 174°45'57"E; City of Auckland, northern New Zealand; Fig. 2.1). Lake Pupuke is a well-studied crater lake (Horrocks et al. 2005; Augustinus et al. 2006; Striewski et al. 2009, 2013; Stephens et al. 2012a, b; Heyng et al. 2012, 2014, 2015; Newnham et al. 2018, see also Chapter 3): one of 53 volcanic centres in the active Auckland Volcanic Field (Fig. 2.1b). A 16.1-m long core was extracted from the lake in February 2007 using an UWITEC drilling platform and fixed-piston coring system. The sediments are laminated throughout, although the laminae are not considered to be annually deposited (Striewski et al. 2013). Nineteen samples between 220 and 590 cm depth were selected for this study. The sediments in this interval are composed of on average 15% clay, 76% silt and 9% sand. The organic content of the sediments is high, with an average organic carbon content of 17.8%. To provide a simple test of whether pollen yield of the same sample varies with differing SG within the reported range, each of the nineteen samples was analysed at a SG of 2.0 (hereafter treatment 1) and 2.2 (hereafter treatment 2) and their pollen abundances compared.
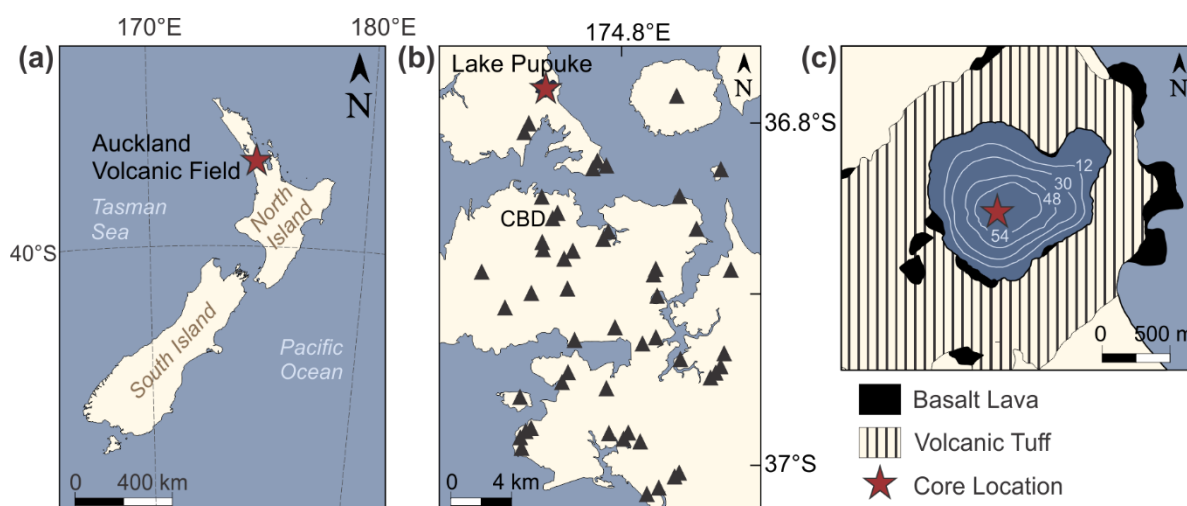
Figure 2.1: Map of northern New Zealand (a) with location of Lake Pupuke (adapted from Hopkins et al. 2017) (b) and the core site (adapted from Newnham et al. 2018) (c). Triangles indicate volcanic centres within the Auckland Volcanic Field. CBD = Auckland's central business district.

2.2.2 Sample preparation

A tablet with a known amount of exotic *Lycopodium* spores was added to each sample, with measured volume of 0.5 to 0.75 cm^3, to facilitate calculation of absolute pollen concentrations. We adapted the pollen preparation method described by Faegri and Iversen (1989) and Moore et al. (1991) as follows.

5 mL of 10% hydrochloric acid were added and heated for 5 min at 80ºC to remove any carbonates from the sediments. Then, samples were centrifuged for 5 min at 3000 RPM and decanted. 5 mL of reverse osmosis water were added and then decanted after centrifuging for 5 min at 3000 RPM. 5 mL of 10% potassium hydroxide were added and heated for 15 min at 80ºC to remove humic acids. Samples were centrifuged for 5 min at 3000 RPM and decanted, then dissolved in reverse osmosis water and poured over a 90-µm nylon sieve to remove coarse particles. The > 90-µm fraction was discarded and the smaller fraction centrifuged for 5 min at 3000 RPM and decanted.

5 mL of SPT solution were added (SG of solution dependent on treatment iteration: 2.0 for treatment 1 and 2.2 for treatment 2) to separate the organic from the minerogenic fraction of the sediments. Samples were centrifuged for 10 min at 2000 RPM and the floating portion of the sample transferred to a 6-µm sieve using a plastic pipette and washed with reverse osmosis water to remove clay particles. The non-floating residue was stored for later inspection. The > 6-µm fraction was transferred to a new tube using reverse osmosis water, centrifuged for 5 min at 3000 RPM and decanted. The smaller fraction was discarded.

The organic fraction was washed in 5 mL glacial acetic acid, centrifuged for 5 min at 3000 RPM and decanted to remove any water from the sample before acetolysis. 5 mL of acetolysis mixture were added (acetic anhydrite and sulfuric acid, ratio 9:1 by volume) and heated for 5 min at 90ºC to remove any cellulose. Then samples were centrifuged for 5 min at 3000 RPM and decanted. Samples were subjected to another wash with glacial acetic acid before two more reverse-osmosis-water washes, before mounting.

One droplet of sediment residue was added to a droplet of heated glycerine jelly on a microscope slide and mixed with a toothpick. Water was evaporated from the mix by heating

on a hot plate, before placing a cover slip. Pollen counting proceeded after slides had cooled to room temperature.

The pollen slides were analysed under a light microscope at 400x or 1000x magnification. Grains were identified with the aid of pollen floras (Moar 1993; Pocknall 1981) and modern reference slides, applying the naming convention suggested by Moar et al. (2011). We counted 250 pollen sum grains per sample, where the pollen sum includes dryland pollen but excludes wetland taxa and spores. Pollen taxon abundances are expressed as percentages of the dryland pollen sum.

2.2.3 Statistical analysis

We set out to test the difference in pollen abundances between treatment 1 and 2, but the dataset resulting from the analyses described above is complex. It comprises two explanatory variables, i.e. depth and treatment, and a multitude of dependent variables, i.e. the abundances of up to 63 identified taxa and their total concentration. Although depth is not a variable of interest in this study, it cannot be discounted, because it results in temporal autocorrelation between the nineteen samples. Samples that are close to each other in depth (and therefore age) are likely to be more similar with respect to taxon abundances than samples that are more distant along the depth scale. This rules out the use of simple paired tests between treatments 1 and 2, such as the paired *t*-test or non-parametric Wilcoxon signed-rank test, because it violates the assumption of independence among observations. Autocorrelation tends to lead to type I errors (false positives) in statistical tests (Legendre and Legendre, 1998). In addition, the different pollen taxa are not independent from each other, as they are part of an ecosystem, i.e. the abundances of some taxa will covary with others, made worse by the practice to calculate relative abundances. These multiple related, covarying variables in our dataset ensure that the assumptions of most statistical tests are violated. In addition, the non-normal distribution of pollen abundances rules out most parametric tests, such as multivariate analysis of variance (MANOVA). Therefore, we decided to compute the association between our samples using multivariate ordination, i.e. redundancy analysis (RDA), following recommendations in Borcard et al. (2011). Only taxa that occur in ≥ 5 samples with an abundance of ≥ 1% in at least two samples were included.

RDA was used to quantify the effects of depth and treatment on the taxon abundances in the samples. RDA preserves the Euclidean distance between samples and pollen taxa, although pollen taxa counts had to be transformed before they could be subjected to RDA. Appropriate transformation of count data will avoid the use of double-zeros as indications of resemblance among samples and will reduce the importance of more abundant taxa; otherwise RDA would give the same importance to absolute differences in abundance irrespective of the order of magnitude of the abundances. For this study, the Hellinger transformation (Legendre and Legendre 1998) was chosen, but other transformations (e.g. chi-square) produced similar results.

After RDA, a permutation test was applied to test the significance of the canonical axes. Partial RDA (pRDA) was used to partition the variance of the response data into the proportion explained by SG treatment and depth, along with the proportion of unexplained variance, by comparing the canonical eigenvalues of the RDA and pRDAs to the sum of all eigenvalues (Borcard et al. 1992). This results in four fractions of variation, i.e. the proportion of variance explained by SG treatment (fraction a), the proportion explained by depth (c), the proportion explained jointly by treatment and depth (b), and the unexplained variance (d). These proportions (or R^2 values) are subsequently adjusted, taking into account the number of samples and pollen taxa (degrees of freedom), following Peres-Neto et al. (2006).

Finally, to further assess how the effect of treatment may vary over depth, we calculated the Hellinger distances between the samples that underwent treatment 1 and treatment 2 from the same depth, and attempted to relate these distances to the trends of the pollen taxa abundances over depth.

All statistical tests were performed in the program R (R Core Team 2015). We use the vegan package (Oksanen et al. 2016) for ordinations and the analogue package (Simpson and Oksanen 2016; Simpson 2007) for calculating distances. All figures were plotted using the ggplot2 package (Wickham 2009).

2.3 Results

2.3.1 Concentration and taxon abundances

The sediments of Lake Pupuke contain very high abundances of pollen, i.e. up to 0.7×10^6 grains/cm^3 (overall mean = ~264,000 grains/cm^3). The mean measured concentration of samples that underwent treatment 1 is ~267,000 grains/cm^3 and the mean measured concentration of samples that underwent treatment 2 is ~261,000 grains/cm^3. The calculation of these concentrations is based on the abundance of *Lycopodium* spores.

The pollen abundances of the 38 samples are visualised in Figure 2.2, which shows that some taxa are always more abundant under treatment 1 (e.g. *Dacrydium cupressinum* and *Prumnopitys taxifolia*), while other taxa are always more abundant under treatment 2 (e.g. *Libocedrus* and *Leptospermum* type). The main trends over depth evident for both treatments involve the increase of *Agathis australis* and *Phyllocladus* above ~400 cm depth, concomitant with the decline of *Dacrydium cupressinum* and *Cyathea dealbata*.
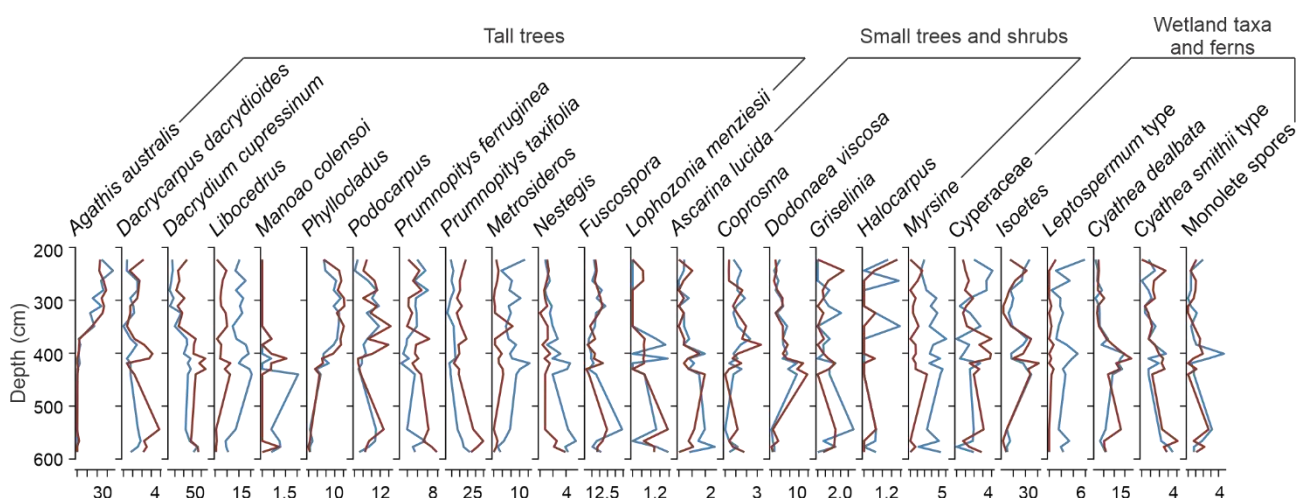
Figure 2.2: Pollen percentage diagram of Lake Pupuke: nineteen samples prepared with heavy liquid (SPT) both at an SG of 2.0 (treatment 1: red) and an SG of 2.2 (treatment 2: blue). Only taxa that occur in ≥ 5 samples with a maximum abundance of ≥ 1% in at least two samples are shown. Note that the x-axes are scaled differently between taxa; all graphs were plotted with the same width to improve legibility.

2.3.2 Ordination

Out of the 63 plant taxa present in the dataset, 25 occur in ≥ 5 samples with an abundance of ≥ 1% in at least two samples. The RDA and pRDAs were therefore based on the counts of these 25 taxa. Variation among samples is measured as the variance of the variables, where the sum of the variances is called "inertia". The first two canonical axes of the RDA explain 54.1% of the total inertia (Table 2.1), meaning that the two explanatory variables explain more than half of the variance in the dataset (see also section "partitioning of variance" below). RDA splits the 38 samples into two main groups over the first axis, which correlates strongly with depth: the samples at > 350 cm depth are positioned on the right side of the first axis (Fig. 2.3a). The second RDA axis correlates with treatment and broadly splits samples that underwent treatment 1 (mostly negative axis 2 values) from those that underwent treatment 2 (mostly positive axis 2 values). The first axis of the pRDA (Fig. 2.3b), where the effect of depth is partialled out, is similar to the second axis of the RDA, which is unsurprising as the vectors of depth and treatment in the RDA are positioned orthogonally (see also section "partitioning of variance" below).

Table 2.1: Total inertia and eigenvalues of the axes of three RDAs on pollen counts: ordination constrained by treatment and depth, constrained by treatment with depth partialled out, and constrained by depth with treatment partialled out.

	Axis 1	Axis 2	Axis 3	Axis 4	Total inertia
RDA: Treatment and depth					
Eigenvalue	0.0501	0.0186	0.0242	0.0062	0.1270
Proportion Explained	0.3947	0.1463	0.1903	0.0490	1.0000
Cumulative Proportion	0.3947	0.5410	0.7313	0.7803	
pRDA: Treatment					
Eigenvalue	0.0199	0.0242	0.0062	0.0048	0.0782
Proportion Explained	0.2543	0.3092	0.0795	0.0611	1.0000
Cumulative Proportion	0.2543	0.5634	0.6430	0.7041	
pRDA: Depth					
Eigenvalue	0.0488	0.0242	0.0062	0.0048	0.1071
Proportion Explained	0.4559	0.2256	0.0580	0.0446	1.0000
Cumulative Proportion	0.4559	0.6815	0.7395	0.7841	
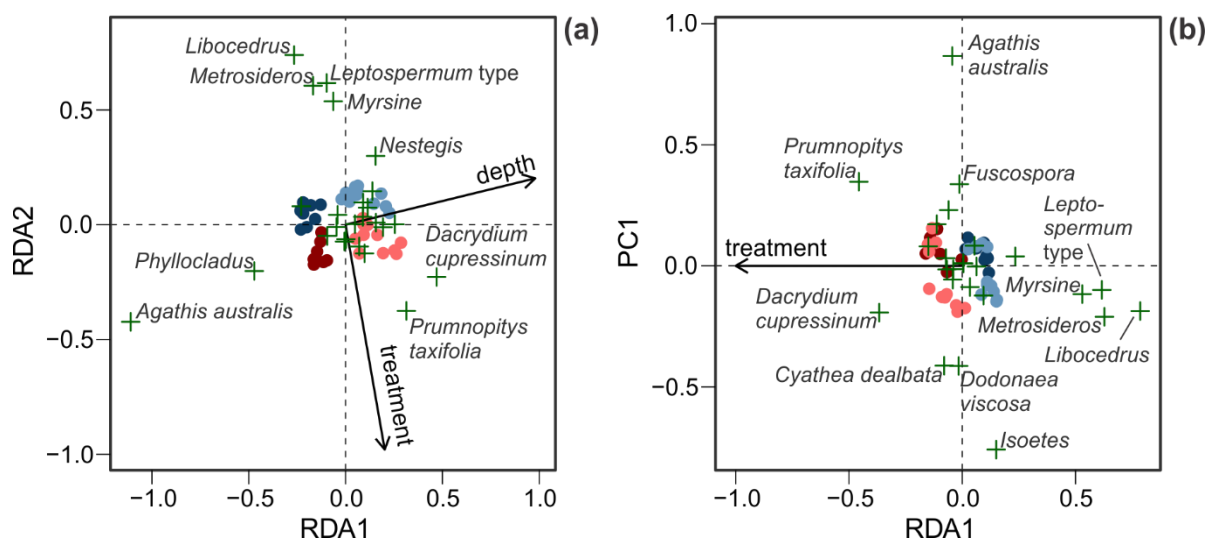
Figure 2.3: Triplots of the RDA (a) and pRDA with the effect of depth partialled out (b). Sample scores are denoted with circles and their depth and the treatment they underwent are indicated with colour: reds for treatment 1, blues for treatment 2; darker colours for samples at < 350 cm depth, lighter colours > 350 cm depth. Pollen taxon scores are denoted with a +-sign and only taxa away from the origin of the graph are labelled; for scores of all taxa, see Table 2.3. Vectors of independent variables are indicated with arrows. PC1 in (b) refers to principle component 1 of the variance not explained by treatment.

2.3.3 Partitioning of variance

The total variance in the dataset is 0.127, which is divided into 51.5% variation explained by SG treatment and depth and 48.5% unexplained variation (Table 2.2). The unique contribution of treatment is 14.7%. The permutation tests show that both treatment and depth are significant contributors (p-value < 0.01 for all tests). Interestingly, the fraction of variation explained jointly by treatment and depth is ~0. This is consistent with their orthogonal nature as seen in the RDA triplot (Fig. 2.3a), which means that treatment and depth do not co-vary.

Table 2.2: Partitioning of variance in the pollen dataset explained by treatment and depth.

Fraction	Variance explained	R^2	R^2_{adj}	F	p-value
(a) Treatment	0.0199	0.1565	0.1473	11.9	0.001***
(b) Joint	0.0000	0.0000	-0.0143[1]		
(c) Depth	0.0488	0.3846	0.3818	29.3	0.001***
(d) Unexplained	0.0583	0.4590	0.4852		
a+b+c	0.0687	0.5410	0.5148	20.6	0.001***
Total inertia	0.1270				

[1] Negative R^2_{adj} values can be ignored or considered as null for interpretation

2.3.4 Effect of treatment on pollen taxa

The triplot of the RDA (Fig. 2.3a) shows that depth is mostly positively correlated with *Dacrydium cupressinum*, and negatively with *Agathis australis* and *Phyllocladus*. The relationship between these taxa with depth is also evident in the pollen percentage diagram (Fig. 2.2). The triplot of the pRDA (Fig. 2.3b) indicates that treatment is negatively correlated with *Libocedrus*, *Metrosideros*, *Leptospermum* type and *Myrsine*, and positively with *Prumnopitys taxifolia* and *Dacrydium cupressinum*. Table 2.3 shows all 25 taxa sorted by their score on the first pRDA (treatment) axis. Positive (negative) scores indicate that the taxon may be underrepresented (overrepresented) under treatment 1: preparation with a low heavy liquid density.

Table 2.3: RDA (constrained by treatment and depth) and pRDA (constrained by treatment with depth partialled out) axis 1 and 2 scores of all 25 taxa included in the analysis, sorted from highest to lowest pRDA axis 1 score.

	RDA		pRDA	
	axis 1	**axis 2**	**axis 1**	**axis 2**
Libocedrus	-0.27	0.74	0.79	-0.19
Metrosideros	-0.17	0.61	0.63	-0.21
Leptospermum type	-0.10	0.62	0.62	-0.10
Myrsine	-0.06	0.54	0.53	-0.12
Nestegis	0.15	0.30	0.23	0.04
Isoetes	-0.23	0.08	0.15	-0.76
Ascarina lucida	0.14	0.15	0.09	-0.12
Monolete spores	0.09	0.10	0.06	0.00
Halocarpus	-0.04	0.04	0.05	0.08
Manoao	0.11	0.07	0.03	-0.09
Coprosma	-0.05	-0.01	0.01	0.01
Griselinia	0.09	0.03	0.00	0.01
Fuscospora	0.05	0.00	-0.01	0.34
Dodonaea viscosa	-0.10	-0.05	-0.02	-0.41
Phyllocladus	-0.47	-0.20	-0.04	-0.01
Lophozonia menziesii	0.16	0.01	-0.04	-0.06
Agathis australis	-1.11	-0.42	-0.04	0.87
Podocarpus	0.00	-0.06	-0.06	0.23
Cyperaceae	-0.01	-0.07	-0.07	-0.01
Cyathea smithii type	0.19	-0.01	-0.07	0.03
Cyathea dealbata	0.25	0.00	-0.08	-0.41
Prumnopitys ferruginea	0.07	-0.09	-0.11	0.17
Dacrycarpus dacrydioides	0.10	-0.12	-0.15	0.08
Dacrydium cupressinum	0.47	-0.23	-0.37	-0.19
Prumnopitys taxifolia	0.31	-0.38	-0.46	0.35

2.3.5 Distance over depth

The distance between samples prepared according to treatment 1 vs. treatment 2 varies over depth, between 0.23 and 0.48, but there is no clear trend (Fig. 2.4). There is also no apparent relationship between distances and vegetation community.
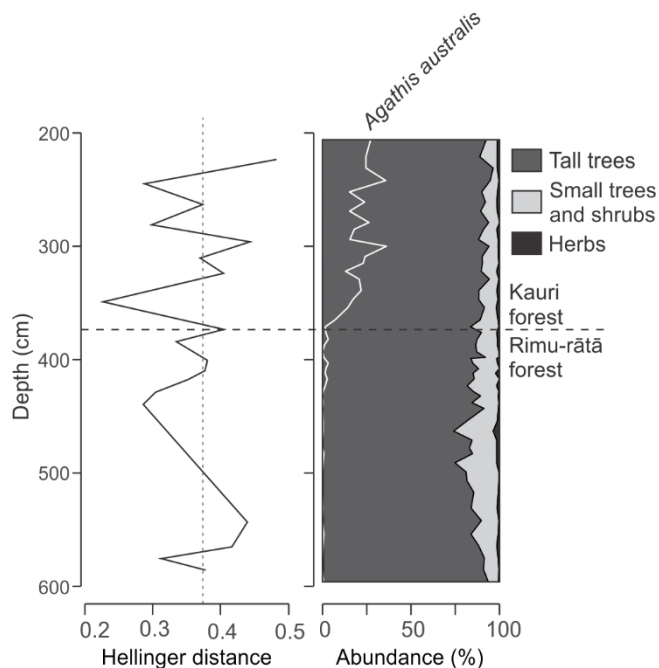
Figure 2.4: Hellinger distances between treatment 1 and 2 over depth (left) compared to summary of pollen abundances (right). Median distance (0.37) indicated with vertical dotted line. Transition from rimu-rātā to kauri forest indicated with horizontal line.

2.4 Discussion

2.4.1 Treatment affects taxon abundances

It is clear from the pollen diagrams (Fig. 2.2), ordinations (Fig. 2.3) and partitioning of the variance (Table 2.2) that taxon abundances are affected by the SG of the heavy liquid used during preparation. We conclude that an SG of 2.0 (treatment 1) is not sufficient to obtain a representative sample from the Lake Pupuke sediments. This was confirmed by the inspection of the non-floating fractions, which contained substantial numbers of grains of mainly *Metrosideros*, *Libocedrus* and *Myrsine.* No pollen grains were found in the non-

floating fractions of samples that underwent treatment 2 (SG = 2.2). Although a higher SG results in more extraneous material in the slides, this should not have precedent over reliability of sample yield which otherwise could result in erroneous interpretations. Furthermore, if heavy liquid SG is not tightly controlled during pollen preparation, spurious fluctuations in pollen abundances between samples, resulting from fluctuations in SG, may also be incorrectly interpreted in ecological or climatological terms.

Less clear is which taxa are affected, because we lack appropriate statistical methods to test each taxon. However, looking at the patterns in abundances in the pollen diagram and pRDA axis 1 scores (Table 2.3), we tentatively conclude that *Libocedrus*, *Metrosideros*, *Leptospermum* type, *Myrsine* and *Nestegis* are underrepresented under treatment 1, while *Prumnopitys taxifolia* and *Dacrydium cupressinum* are overrepresented. *D. cupressinum* and *P. taxifolia* are both saccate grains (Fig. 2.5a-b), meaning they have air bladders or "sacci". These sacci make the pollen grains buoyant, which is essential for their germination to fertilise ovules (Leslie, 2010) and may also aid in their aerial dispersal. It is therefore unsurprising that the grains of saccate pollen should be overrepresented compared to non-saccate grains in a situation where the separation medium is of relatively low density. Two other saccate grains included in this study, i.e. *Dacrycarpus dacrydioides* and *P. ferruginea*, also have negative values on pRDA axis 1. In contrast, the underrepresented taxa are all small, compact grains with no obvious features that might promote their buoyancy (Fig. 2.5c-g).

2.4.2 The effect of treatment on Lycopodium *spores*

The pollen concentrations for treatment 1, calculated from added *Lycopodium* abundances, are unusable. The absolute concentration of pollen grains appears higher in samples that underwent treatment 1 (SG of 2.0) compared to samples that underwent treatment 2 (SG of 2.2). However, this is highly unlikely to reflect the true concentrations, because more organic material appears to float at an SG of 2.2 compared to 2.0. Moreover, preliminary tests showed that the minerogenic fraction of samples prepared under treatment 1 still contained pollen grains, while that of samples prepared under treatment 2 did not. More likely, *Lycopodium* is underrepresented under treatment 1. If fewer *Lycopodium* spores are found in a sample, the calculated concentration will be higher. Therefore, if an SPT solution is used

for pollen preparation with an SG that is too low, calculated pollen concentrations will be biased, i.e. too high.
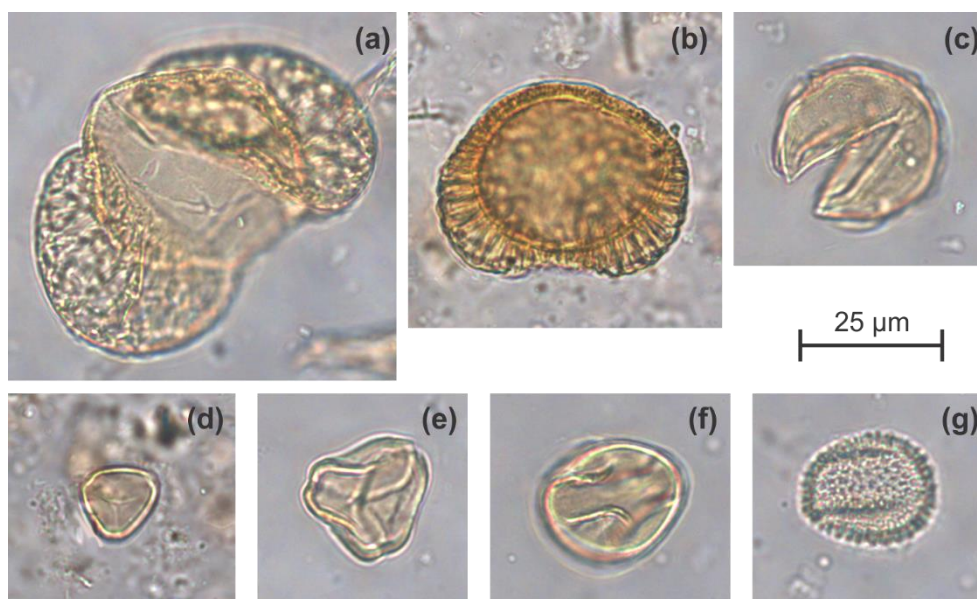
Figure 2.5: Photos of pollen taxa affected by a low SG: Prumnopitys taxifolia (a), Dacrydium cupressinum (b), Libocedrus (c), Leptospermum type (d), Metrosideros (e), Myrsine (f) and Nestegis (g).

2.4.3 Site-specific factors

We must emphasise that our results are obtained from the highly organic Pupuke Holocene lake sediments and that sediments from other sites may respond differently. Nevertheless, these results should sound a general warning about the uncritical use of density separation even within the reportedly acceptable SG range and we think that some generalisations are appropriate from our study. We think the sediment matrix of our particular core is a factor in the retention of small, compact grains during density separation. The organic content of the Lake Pupuke Holocene sediments is high (mean organic carbon content = 17.8%) and some grains may be trapped beneath a layer of material that does not float on SPT solution with a density of 2.0 g/cm^3 but does float at a density of 2.2 g/cm^3. Alternatively, organic particles may be stuck to the pollen grains. Saccate pollen may be buoyant enough to overcome the extra mass, while smaller grains are more likely to sink. At an SG of 2.2, all organic material appears to float. This phenomenon is unlikely to be specific to the Lake Pupuke sediments and might occur in all highly organic sediments. Other factors, such as the "stickiness" of

the sediment matrix or propensity for clumping of organic material and pollen, may play a role as well.

In addition, the effect of treatment may be dependent on the reconstructed vegetation community and associated pollen morphological characteristics and some taxon combinations might not be affected. In this study, only two main types of vegetation communities were covered, i.e. two types of conifer-broadleaf forest (Wardle, 1991) either dominated by *Dacrydium cupressinum* and *Metrosideros* (rimu-rātā forest) or by *Agathis australis* (kauri forest) (see Chapter 3 for full pollen diagram). However, there is no obvious pattern in the calculated Hellinger distances over depth (Fig. 2.4), indicating that, for both of these two forest types at least, treatment affects pollen assemblage.

2.4.4 Conclusions and Recommendations

The advantages of density separation with SPT over HF treatment for removal of silicates from pollen samples are now well established (Campbell et al. 2016). However, less attention has been paid in the literature to quality control issues with the use of SPT SG treatment. What is the optimum SG value and recommended range for a particular sediment type? Are pollen yields reliable across the recommended range? What checks should palynologists employ routinely to ensure they are not losing pollen in the process, effectively throwing the baby out with the bathwater? This study has not attempted to provide comprehensive answers to these questions, but rather we have set out to draw attention to the importance of quality control in the density separation procedure by illustrating the potential dangers of applying the procedure indiscriminately. For highly organic sediments, such as those in this study, the comparatively high levels of unwanted organic material may encourage the employment of SG towards the lower end of the recommended range but, as we have shown, this heightens the risk of pollen loss for certain morphological types.

We urge palynologists who are using density separation to consider what the most appropriate SG for their sedimentary material is. SG should be high enough to ensure that all pollen is collected from the sample, but as low as possible to minimise the retention of extraneous organic matter. We recommend routine application of a quick and simple test to at least one sample per batch: prepare a smear slide of the minerogenic fraction of the sample that is retained after density separation. Sometimes it is immediately obvious that

much pollen is retained in this fraction. However, if the minerogenic fraction is large and the pollen concentration relatively low, retained pollen may not be noticed on the smear slide and this test may not be sufficient to identify preparation problems. To be certain you are getting a representative sample of the pollen assemblage, one or more samples from your sediment core should be prepared at different densities, and the differences in outcome compared. It may be necessary to repeat this procedure for several sections of the same core, if lithological changes over depth occur.

Acknowledgements

Chapter 3

Holocene temperature, humidity and seasonality in northern New Zealand linked to Southern Hemisphere summer insolation
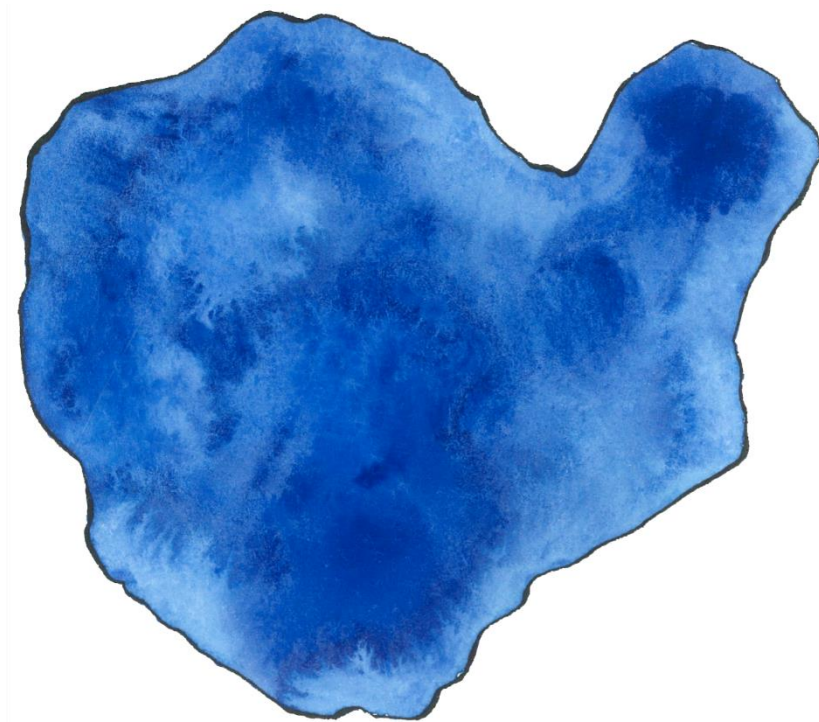
Preface

This chapter has been published as a research paper in Quaternary Science Reviews (reference: van den Bos, V., Rees, A., Newnham, R., Vandergoes, M., Wilmshurst, J., Augustinus, P. (2018). Holocene temperature, humidity and seasonality in northern New Zealand linked to Southern Hemisphere summer insolation. *Quaternary Science Reviews*, 201, 77-88.), as was mentioned in Chapter 1. The only changes made to the manuscript's text are the minor revisions as suggested by the thesis examiners and the replacement of abbreviation MAT for "modern analogue technique" to "MA technique", so all abbreviations would be consistent across the chapters in this thesis.

Abstract

The Holocene thermal maximum (HTM) is a spatio-temporally variable period of generally warmer conditions during the early and middle Holocene that is often used as an analogue for future climate change. Global scale climate reconstructions and models tend to smooth out the variations and complexity of the HTM and inconsistencies between reconstructions from different locations and proxies are often attributed to bias arising from different locations or proxies. We use these differences as a source of information about seasonality and precipitation during the Holocene in a multi-proxy investigation of the sediments of Lake Pupuke, northern New Zealand. The sediments, spanning the last 16 kyr, were analysed for pollen, from which mean annual air temperatures (MAAT) and effective precipitation were estimated, and chironomids, from which summer air temperature (SmT) was estimated. We found no evidence for an HTM in the MAAT reconstruction, questioning the validity of treating the early-to-mid Holocene as an analogue for future climate change in northern New Zealand. SmT increases between 10 and 3 cal kyr BP, correlating strongly with integrated local summer insolation. Early-Holocene low seasonality (from 12 to 9.3 cal kyr BP) was likely driven by low local summer insolation intensity. An early-to-mid-Holocene wet period (9.6–7.5 cal kyr BP) corresponds to relatively high southern westerly wind (SWW) strength. Mid-to-late-Holocene summers following the wet period were hot and dry, especially 4.0–2.4 cal kyr BP, allowing the tall conifer, kauri (*Agathis australis*) to expand throughout northern New Zealand. Low effective precipitation at this time is consistent with increased evapotranspiration due to higher SmT, although reduced precipitation due to southward displaced SWW or increased El Niño frequency may also have contributed.

3.1 Introduction

Proxy records provide valuable information on past climate changes and a test for climate model simulations that aim to improve our understanding of the mechanisms that drove these changes (Renssen et al., 2005). Yet, model-data comparisons do not always agree (e.g. Liu et al., 2014) and improvements in both domains are necessary. The inability of models to simulate reconstructed past warm climate intervals in particular, and therefore future climate change, highlights the importance of exploring more proxy reconstructions of those warm intervals that may be analogous to climate scenarios projected for the coming decades (Masson-Delmotte et al., 2013; Fischer et al., 2018).

One such period is the Holocene thermal maximum (HTM), which is loosely defined as an interval of generally warmer conditions spanning 11-5 kyr (Fischer et al., 2018). However, direct comparisons are complicated, because the HTM was asynchronous in its spatio-temporal expression. The warm period is best expressed in the higher northern latitudes, earliest in the North Pacific marine sector and delayed on land areas influenced by residual ice, especially the North Atlantic and Greenland regions. These regional complexities are typically smoothed out in global climate reconstructions, where inconsistencies between reconstructions from different proxies and locations are regarded as biases to be avoided (e.g. Marcott et al., 2013). On the other hand, the differences between proxy signals and records from different regions can be informative (Sikes et al., 2013). For instance, the seasonal bias in the differential response of multiple proxies to an environmental factor such as temperature may be used to reconstruct past seasonality.

We must also consider that the various spatio-temporal inconsistencies in global climate reconstructions may represent variable climate drivers in different areas of the globe. Although Northern Hemisphere summer insolation is often invoked to explain global climate signals (e.g. Marcott et al., 2013), the effect of local insolation parameters in other regions should not be overlooked (Huybers & Denton, 2008).

Here we aim to explore some of the complexity of the HTM and the structure of Holocene climate change in general using a suite of paleoclimate proxies developed from the sediments of Lake Pupuke, northern New Zealand (36°46'S; Fig. 3.1). The 16-kyr record combines mean summer air temperatures (SmT) reconstructed from chironomid remains

(Dieffenbacher-Krall et al., 2007) and mean annual air temperatures (MAAT) reconstructed from pollen assemblages (Wilmshurst et al., 2007), as well as an estimate of effective precipitation from those pollen assemblages (Jara et al., 2017).
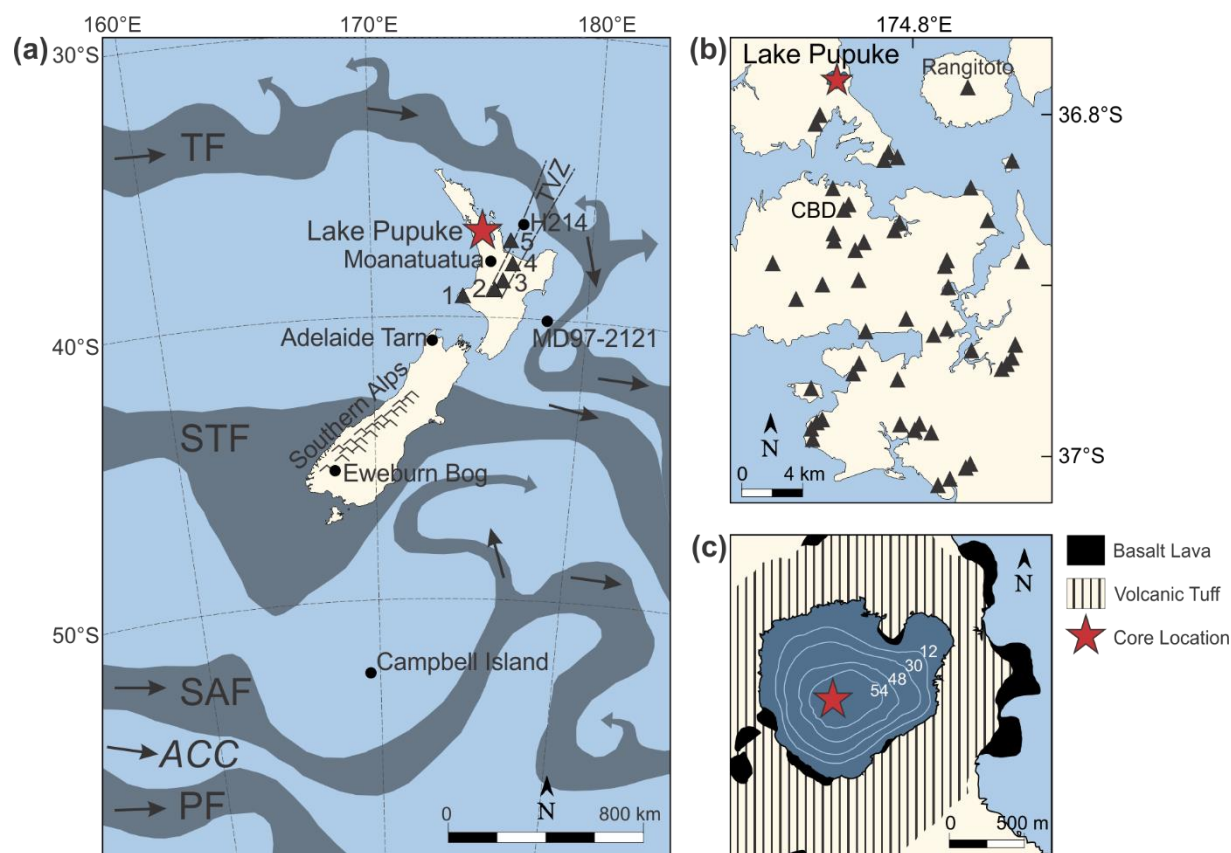
Figure 3.1: Map of New Zealand, including surface water circulation (based on Prebble et al., 2017) and locations and core sites mentioned in the text (a), map of Auckland Volcanic Field (adapted from Hopkins et al., 2017) (b) and map of Lake Pupuke with bathymetry (adapted from Newnham et al., 2018) and the location of the core site (c). TF = Tasman Front; STF = subtropical front; SAF = subantarctic front; PF = polar front; ACC = Antarctic circumpolar current. Star indicates our study site, Lake Pupuke; circles indicate other core sites; triangles indicate volcanic centres (1 = Egmont Volcanic Centre, 2 = Tongariro Volcanic Centre, 3 = Taupo Volcanic Centre, 4 = Okataina Volcanic Centre, 5 = Tuhua Volcanic Centre, TVZ = Taupo Volcanic Zone). CBD = Auckland's central business district.

The relevance of our record follows, first, from Lake Pupuke's southern mid-latitude position, which is a region that is strongly underrepresented in Holocene paleoclimate research (Marcott et al., 2013). Second, our multi-proxy approach involving both vegetation and aquatic insects sheds new light on changes in seasonality and effective precipitation. Third, we compare our temperature reconstructions to insolation intensity at different latitudes in light of the ongoing debate over the role of local insolation in Holocene temperature change

and the HTM in particular (Huybers & Denton, 2008; Liu et al., 2014; Marsicek et al., 2018). In addition, we draw attention to limitations in the current methodology for reconstructing past climate in New Zealand and suggest how these might be addressed with future work.

3.2 Study Region

Auckland isthmus, together with immediately adjacent land areas to the north and south, is the location of both an active volcanic field (the Auckland Volcanic Field: AVF) and New Zealand's largest city (Fig. 3.1). Auckland region experiences a warm temperate, oceanic climate with few temperature extremes (Chappell, 2014). MAAT usually lies between 14 and 16°C, with mean temperatures in summer (Dec-Feb) around 19.5°C and in winter (Jun-Aug) around 12°C (data from the National Climate Database CLIDB: https://cliflo.niwa.co.nz). The distribution of precipitation in the Auckland region (totalling around 1100-1200 mm per year) shows a winter maximum, consistent with a more northern position of the high-pressure belt north of New Zealand, which increases westerly airflow, compared to summer (Chappell, 2014). Unless otherwise specified, when we mention seasonality in subsequent text, we refer to the amplitude of temperature minima and maxima over a year (or the difference between mean summer and mean winter temperature), rather than this seasonality of precipitation.

New Zealand weather is strongly influenced by atmospheric circulation patterns that are dependent on a combination of broad scale climate modes, i.e. the Southern Annular Mode (SAM) and El Niño-Southern Oscillation (ENSO). Kidston et al. (2009) report a negative correlation between SAM and wind speed in New Zealand: during positive SAM, the southern westerly wind (SWW) belt contracts southwards, decreasing westerlies over New Zealand, but strengthening them in the core of the SWW belt. In the Auckland region, positive SAM is associated with fewer cyclones and more anticyclones, decreased precipitation, and increased temperatures (Kidston et al., 2009). Regarding ENSO, southwesterly airflow is enhanced during El Niño years, due to the northward displacement of the westerlies (Ummenhofer & England, 2007). El Niño reduces rainfall in the north of New Zealand, whereas during La Niña years rainfall is generally increased, although there are strong inter-region heterogeneities and anomalous patterns that differ from season to season for both phases (Kidson & Renwick, 2002; Ummenhofer & England, 2007).

Although the region is now urbanised, the natural vegetation cover can be deduced from isolated remnants (Wardle, 1991) and Holocene palynological records (Newnham & Lowe, 1991; Newnham et al., 2018). Prior to human settlement, ~750 yr BP (Newnham et al., 2018), the area supported extensive conifer-broadleaf forests, which are temperate rainforests that are dominated by either podocarps (rimu-rātā or podocarp-broadleaf forest) or *Agathis australis* (kauri forest). Both types of forest are multi-tiered with abundant lianas, epiphytes and tree ferns in the canopy and subcanopy layers (McGlone et al., 2016). In rimu-rātā forest, *Dacrydium cupressinum* (rimu) and *Metrosideros robusta* (northern rātā) are the major emergent species with, among others, *Beilschmiedia tawa* (tawa), *Knightia excelsa* (rewarewa), *Elaeocarpus dentatus* (hīnau) and *Weinmannia silvicola* (tōwai) in the canopy layer. Treeferns (*Cyathea* spp. and *Dicksonia* spp.) characterise the subcanopy and ferns the forest floor. Kauri forest is similar in structure and composition, but *Agathis australis* (kauri) and *Phyllocladus trichomanoides* (tanekaha) dominate the canopy, and the proportion of species that are part of the northern endemic group (e.g. *Beilschmiedia tarairi*, *Ixerba brexioides*) is much higher. Coastal forests in the Auckland region support *Metrosideros excelsa* (pōhutukawa), *B. taraire* and *Dodonaea viscosa* (akeake). Beech forests (with *Fuscospora* spp. and *Lophozonia menziesii*) are generally absent from Auckland northwards, apart from some isolated stands (Wardle, 1991).

The AVF comprises 53 Quaternary basaltic centres located within a 600 km^2 area (Hopkins et al., 2017; Fig. 3.1b) and includes several maar craters that have provided a depositional environment for the accumulation of lacustrine sediments. Lake Pupuke (36°46'49"S, 174°45'57"E; Fig. 3.1c) is the only such crater that still contains a freshwater lake today and, as a consequence, is unique in having a sedimentary record covering the entire Holocene. The crater's estimated age is c. 193 kyr (Leonard et al., 2017), making it one of the oldest, if not the oldest, AVF volcanoes. The lake is a closed system, fed only by precipitation, with no significant inflows or outflows, a total surface area of 1.1 km^2, volume of 2.9 km^3 and maximum depth of 57 m (Horrocks et al., 2005). The lake supports submerged macrophytes (Coffey & Clayton, 1987), is classed as moderately eutrophic (Cassie, 1989) and is stratified during summer (Horrocks et al., 2005).

3.3 Methods

3.3.1 Drilling, chronology, and sampling

Lake Pupuke was cored in February 2007 using an UWITEC drilling platform and fixed-piston coring system. The resulting 16 short, overlapping cores span 16.1 m of sediment with a diameter of 59 mm (Stephens et al., 2012a, b). Only the top 7.1 m representing the last ~16 kyr were analysed in this study. Sediment cores were split lengthwise, after which one half was cut into 1-cm wide wedges and bagged. Both working and archive halves were stored at 4°C and subsampled further between 2015 and 2017 for the palynological and chironomid analyses presented here.

The 16.1-m sequence is estimated to extend to 48.2 ± 4.7 cal kyr BP (Stephens et al., 2012a) and is laminated throughout, although the laminae are not considered to be annually deposited (Striewski et al., 2013). The sequence contains basaltic tephra layers with an AVF source and distal tephra layers from central North Island sources: rhyolitic tephra from Taupo Volcanic Zone, and andesitic tephra from Egmont volcano and the Tongariro Volcanic Centre (Fig. 3.1a, b).

In this study, a revised chronology was established for the top 7.1 m of the core, incorporating the youngest basaltic tephra layer: Rangitoto-1/2 (Needham et al., 2011; Newnham et al., 2018) and a newly modelled age for Tuhua tephra (Lowe, D.J., pers. comm.), while ages reported by Lowe et al. (2013) were used for the other rhyolitic tephra layers (Table 3.1). We also modelled the radiocarbon dates from Stephens et al. (2012a, b), which were recalibrated in OxCal v.4.3 (Bronk Ramsey, 2009a) using the SHCal13 (Hogg et al., 2013) calibration curve (Table 3.2). However, as with previously reported ^{14}C dating at Lake Pupuke, these proved to be problematic due to old carbon effects (Horrocks et al., 2005; Newnham et al., 2018) and the chronology of the core was instead based solely on tephrostratigraphy. The age-depth model was calculated using the *P_Sequence* model (Bronk Ramsey, 2008; Bronk Ramsey & Lee, 2013) with 1 interpolation per cm. The "*TSimple*" outlier model was used (Bronk Ramsey, 2009b), where the outlier distribution is normal ($\mu = 0$, $\sigma = 100$) and each measurement has a prior probability to be an outlier of 5%. This outlier model is appropriate for non-radiocarbon measurements and allows for a

temporal offset of ~ one hundred years between primary deposition of the tephra and final location in the core.

Table 3.1: Tephra layers included in the age-depth model and their sources (see Fig. 3.1a, b)

Tephra name	Type	Source	Depth (cm)	Age (cal yr BP ± 2σ)		Reference
Rangitoto-1/2	Basaltic	AVF	84	530	± 31	Needham et al., 2011
Taupo	Rhyolitic	Taupo VC	205	1718	± 10	Lowe et al., 2013
Tuhua	Rhyolitic	Tuhua VC	399	7637	± 100	Lowe, D.J., pers. comm.
Rotoma	Rhyolitic	Okataina VC	465	9423	± 120	Lowe et al., 2013
Waiohau	Rhyolitic	Okataina VC	618	14009	± 155	Lowe et al., 2013
Rotorua	Rhyolitic	Okataina VC	653	15635	± 412	Lowe et al., 2013
Rerewhakaaitu	Rhyolitic	Okataina VC	701	17496	± 462	Lowe et al., 2013

Table 3.2: AMS radiocarbon ages from Lake Pupuke sediments.

Laboratory number*	Depth (cm)	Conventional radiocarbon age (yr BP ± σ)		Calibrated median age (cal yr BP)	95.4% confidence intervals (cal yr BP)
NZA33095	442	8346	± 35	9306	9040-9043
					9089-9475
NZA33683	483	10124	± 45	11638	11292-11299
					11305-11986
NZA33096	512	10604	± 45	12514	12108-12227
					12266-12340
					12367-12704
OZK258	554	10662	± 104	12475	11834-11884
					11933-12984
NZA33097	585	11673	± 45	13458	13275-13634
					13669-17709
NZA33098	640	13126	± 50	15678	15316-16006

*NZ: Rafter Radiocarbon Lab, New Zealand; OZ: ANSTO-ANTARES, Australia

Samples for the period from 16 cal kyr BP to the present were selected for pollen and chironomid analyses from the stored 1-cm wide wedges that were cut from the cores in 2007. Additional subsamples were taken from the archived half of the core where necessary. This sampling strategy was aimed at deriving regularly spaced time intervals, although subsequent age modelling resulted in irregularly spaced intervals. Sample resolution averages 230 cal yr BP (range: 13–794 cal yr BP).

3.3.2 Palynology

For pollen analysis, 0.5-to-0.75-mL subsamples were obtained and prepared according to standard procedures (Moore et al., 1991), which include a hydrochloric acid step to remove carbonates, potassium hydroxide step to remove humic acids, density separation with a sodium polytungstate solution (specific gravity: 2.2), and acetolysis to remove cellulose. One *Lycopodium* tablet with a known number of spores was added to each sample to allow for the calculation of pollen concentrations. The processed material was mounted on microscope slides in glycerine jelly and analysed under a light microscope at 400x or 1000x magnification. Grains were identified with the aid of Moar (1993), Pocknall (1981) and modern reference slides, applying the naming convention suggested by Moar et al. (2011). We counted 250 pollen sum grains per sample, where the pollen sum includes dryland pollen and *Pteridium* spores but excludes wetland taxa and other spores. Taxa were expressed as percentages of the pollen sum.

The pollen diagram was plotted in R (version 3.3.1; R Core Team, 2016) using the package "rioja" (Juggins, 2015) and zonation was established by visual assessment and aided by applying stratigraphically constrained cluster analysis (CONISS method) to all pollen sum taxa. Rare taxa (with a maximum abundance of < 1%) were excluded from the figure to improve legibility, but included in the cluster analysis.

We chose pollen as a proxy appropriate for the reconstruction of MAAT and effective precipitation. For quantitative MAAT reconstructions, the weighted averaging modern analogue (MA) technique (10 closest analogues; $RMSEP_{boot} = 1.51$, $r^2_{boot} = 0.82$) and partial least squares (PLS) model with 2 components ($RMSEP_{boot} = 1.50$, $r^2_{boot} = 0.77$) were applied to the pollen percentage data, using the New Zealand pre-deforestation pollen dataset described in detail by Wilmshurst et al. (2007). MAAT was found by Wilmshurst et al. (2007) to have the best predictive success in the pollen dataset, and has been reconstructed in many other studies, e.g. McGlone et al. (2010a), Newnham et al. (2012), Sikes et al. (2013) and Jara et al. (2017). Newnham et al. (2013) discuss the strengths and weaknesses of the pre-deforestation pollen training set extensively. The pollen database covers a MAAT gradient from 3.2 to 16.1°C, encompassing all temperatures reconstructed for Lake Pupuke. MAAT was not reconstructed for samples younger than 1000 cal yr BP as anthropogenic impacts have compromised the modelled pollen-climate reconstructions (Wilmshurst et al.,

2007). A Generalised Additive Model (GAM) was fitted to the reconstructed MAAT over time to show a best fit through the data, using R package "mgcv" (Wood, 2006; 2011). Although we are mindful that certain plant species are sensitive to minimum temperatures, mean annual temperature has been found to have better predictive success than seasonal temperature in the training set (Wilmshurst et al., 2007). Therefore, the transfer function was calibrated to MAAT and we assume that pollen assemblages are representative of MAAT but recognise that variability in some taxa may correspond to minimum temperature and the prevalence of frosts, despite these being rare in the region today.

In addition, some key New Zealand pollen taxa show sensitivity to precipitation and drought (Wilmshurst et al., 2007). We calculated the Pollen Moisture Index (PMI) as the normalised, base-10 logarithm of the ratio of *Dacrydium cupressinum* percentages over the percentage of *Podocarpus* and *Prumnopitys* spp., following Jara et al. (2017). The PMI is based on the "rimu ratio" (the abundance of *D. cupressinum* pollen over other tree podocarps), which has previously been used to infer moisture conditions (McGlone & Topping, 1977; Newnham et al., 1995). *D. cupressinum* is less tolerant of droughts and more tolerant of waterlogged soils than other tree podocarps (Franklin, 1968), and its abundance correlates negatively with October vapour pressure deficit, a measure of spring humidity (Wilmshurst et al., 2007). Therefore, PMI can be considered an indicator of effective precipitation: the amount of rainfall that gets absorbed into the soil. Effective precipitation is a function of total precipitation and evaporation and reflects available moisture for plant growth. A GAM was fitted to the reconstructed PMI to show a best fit through the data.

3.3.3 Chironomid analysis

For chironomid analysis, 0.5–6.5 mL of sediment was prepared according to standard procedures (Walker et al., 1991). Samples were deflocculated in hot 10% KOH for 15 minutes and then sieved over a 90-µm nylon mesh. The exact volume of sediment was selected to retrieve at least 50 whole head capsules per sample (Heiri & Lotter, 2001). After sieving, head capsules were picked from the residue in a Bogorov tray under 50x magnification and mounted onto slides, which were fixed with "Norland Optical Adhesive 61". Chironomid taxa were then identified using a light microscope at 400x magnification. In addition, occurrences of mites, bryozoan oocytes, Ceratopogonidae, Characeae oospores,

cladoceran ephippia and Ephemeroptera mandibles were counted. All taxa were expressed as percentages of the total number of chironomids.

The chironomid diagram was plotted in R (version 3.3.1; R Core Team, 2016) using package "rioja" (Juggins, 2015) and zonation was established by visual assessment and aided by applying stratigraphically constrained cluster analysis (CONISS method) to all chironomid taxa. To aid in interpretation of the chironomid record, taxa were assigned to either "cold-adapted taxa" or "warm-adapted taxa", based on their affinity (β coefficient; see Appendix A1, Table A-1) with SmT (Dieffenbacher-Krall et al., 2007). Four taxa, i.e. *Cryptochironomus*, *Harrisius*, Pentaneurini 1 and *Xenochironomus*, do not have a β coefficient (too rare in training set) and were therefore plotted under "unassigned".

For summer temperature (SmT) reconstruction, we used the Dieffenbacher-Krall et al. (2007) chironomid inference model. This model was developed from South Island lakes and covers an SmT gradient from ~6 to 16ºC, which falls short of the modern SmT at Lake Pupuke (~19.5°C). As a consequence, reconstructions of SmT from the northern North Island may be underestimated and, whilst trends and relative differences should be robust, in absolute reconstructed temperatures should only be used with this limitation in mind.

The original training set was adapted, i.e. we did not distinguish between "short" or "long" antennal pedestal variants of both *Tanytarsus funebris* types A and C. To select the best performing model, the weighted averaging-partial least squares (WA-PLS) approach was compared to the PLS method; for a description of the model diagnostics, see Supplementary Information 1; WA-PLS was chosen as it was shown to perform better than PLS. The model was run in R (version 3.3.1; R Core Team, 2016) using package "rioja" (Juggins, 2015). Rare taxa (number of occurrences < 2 or maximum abundance < 1%) were excluded; percentages were square root transformed. After the best performing model was selected, past SmT was reconstructed from our fossil chironomid assemblages. To test the reliability of the inferred temperatures, the goodness of fit of the samples compared to the training set was assessed (Birks et al., 1990), as well as the abundance of rare (Hill's N2 ≤ 5; Hill, 1973; Brooks & Birks, 2001) and absent taxa in each sample. A GAM was fitted to the reconstructed SmT over time to show a best fit through the data, using R package "mgcv" (Wood, 2006; 2011).

3.3.4 Goodness of fit

To test the reliability of the temperature reconstructions, the goodness of fit of the samples compared to the training sets was assessed for both the MAAT and SmT reconstructions (Birks et al., 1990), as well as the abundance of rare taxa. Goodness of fit was determined by comparing the squared residual distance of the training set samples from an ordination axis constrained by temperature to the squared residual distance of the fossil samples that were passively fitted to the same axis (using R package "analogue"; Simpson, 2007; Simpson & Oksanen, 2016). The fit is considered to be poor if the squared residual distance of a sample is greater than the 90th percentile of the squared residual distances of the training set samples and very poor if it is greater than the 95th percentile (Birks et al., 1990). Rare taxa are defined as taxa that have a Hill's N2 (Hill, 1973) of 5 or less in the training set (Appendix A1, Table A-1 and A-2; Brooks & Birks, 2001). In addition, to test the reliability of the MA technique, the dissimilarity between pollen assemblages from the Pupuke samples and the closest modern analogue in the training set was calculated. If samples have a dissimilarity greater than the 5th percentile of the distribution of dissimilarities between training set samples, they are considered to have no close modern analogue (Simpson, 2007; Simpson & Oksanen, 2016).

3.4 Results

3.4.1 Chronology

The new age-depth model of the Lake Pupuke core ranges from 711 to 0 cm depth, and from 17,500 to -57 cal yr BP (Fig. 3.2). It includes 7 tephra ages (Table 3.1) but excludes the radiocarbon dates (Table 3.2), as many of them had older than expected ages (875 yr older on average), likely caused by a previously described old carbon effect in the Lake Pupuke sediments (Horrocks et al., 2005).

The 95.4% range for modelled ages varies, but averages around 2130 years, which is equivalent to a c. 1000 yr estimated 2-σ error. The 95.4% range is narrower during the last 2 cal kyr BP, i.e. c. 700 years (350 yr 2-σ error). The model predicts three time intervals with different sedimentation rates, i.e.: ~0.35 mm.yr^{-1} for the period between 18.0 to 1.8 cal kyr BP, ~1.0 mm.yr^{-1} for 1.8 to 0.6 cal kyr BP, and ~1.4 mm.yr^{-1} for the last 600 years.
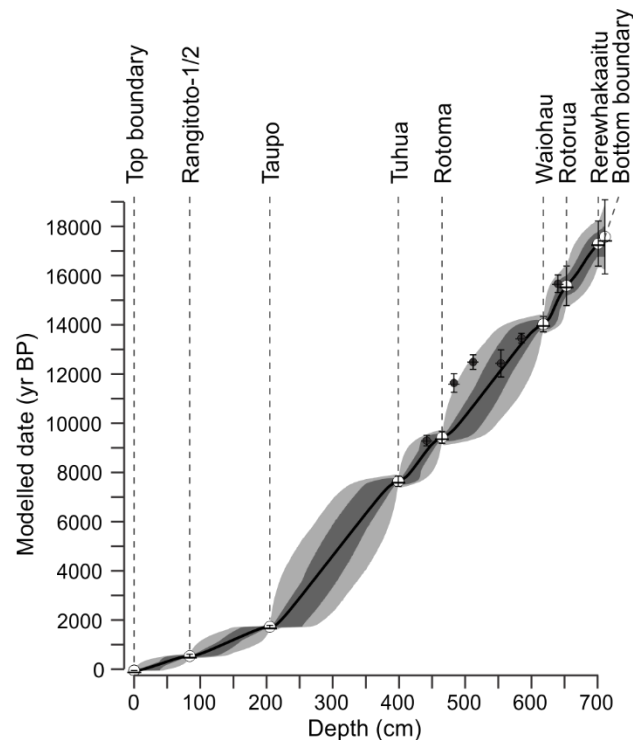
Figure 3.2: Age-depth model for the Lake Pupuke core. The line indicates the median age per depth, dark grey area indicates the 68.2% confidence interval and light grey area indicates the 95.4% confidence interval. Tephra dates (names at top) are plotted with a white symbol at their mean age and a dash at their median age; error bars indicate the 2-σ error. Radiocarbon dates were plotted passively (black symbols).

3.4.2 Pollen record

The pollen record (Fig. 3.3a) is divided into four CONISS zones (PZ1-4) with boundaries at 12.3, 6.7 and 0.5 cal kyr BP. PZ1 to PZ3 show persistently high percentages of forest taxa (~84%) with varying composition. PZ1 shows the gradual decline of conifer-beech forest as *Dacrydium cupressinum* rises to dominance (from 24 to ~32%). During PZ2a (from 12.3 cal kyr BP), *Prumnopitys taxifolia* declines (from ~15 to 5%) while other taxa, notably *Metrosideros* and *Ascarina lucida*, become more abundant (on average 7% and 3% respectively), indicating the development of podocarp-broadleaf forest (rimu-rātā forest); PZ2b (from ~9.3 cal kyr BP) shows further development of this forest along with an increase in *Dodonaea viscosa* (from 1 to 4%). The onset of PZ3 (~6.7 cal kyr BP) coincides with the sudden rise to prominence of *Agathis australis* (from 1 up to 36%) at the expense of *Dacrydium cupressinum* (which declines to ~18%), signalling the advance of kauri forest.

42

(a)

(b)

Figure 3.3: Caption on next page
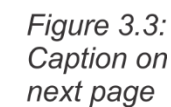
Figure 3.3: Pollen and spore (a) and chironomid (b) percentage diagrams. Each scale interval represents 10%. Pollen diagram excludes rare taxa (taxa with a maximum abundance of < 1%), which are shown in a table in Appendix A-2. Botryococcus values were divided by 5 to increase legibility. Chironomid taxa are ordered according to their affinity with SmT. Acari, bryozoan statoblast and Cladocera ephippium percentages were divided by 10 to increase legibility. Shading represents 5x (pollen) or 10x (chironomids) exaggeration. Zone numbers are indicated on the right. Note: Pteridium (bracken) is included in the dryland pollen sum, as is conventional, but plotted here in the group "ferns and allies"; ferns are normally excluded from the dryland pollen sum.

_____

PZ4 (0.5 cal kyr BP to present) represents the beginning of human influences at the site and is characterised by a sustained increase in *Pteridium* and shrub and herb taxa (e.g. *Coprosma*, *Coriaria*, Poaceae) to ~40%, followed by the appearance of exotic taxa (e.g. *Pinus*, *Plantago lanceolata*) signalling the onset of the European era. For rare taxa (maximum abundance < 1%), see Appendix A2.

3.4.3 Chironomid record

The chironomid record (Fig. 3.3b) is divided into five CONISS zones with boundaries at 9.6, 6.7, 0.9 and 0.7 cal kyr BP. The three samples older than 14.1 cal kyr BP contained virtually no chironomid remains. The samples at 14.1 and 13.6 cal kyr BP yielded sums of just 13.5 and 29.5, respectively. Nevertheless, they were included in all analyses, but interpretations of them are made with caution as these concentrations fall below the $n = 50$ threshold. Chironomid Zone (CZ) 1 is dominated by cold-associated taxa (54%) such as Macropelopini types (combined ~20%) and *Tanytarsus funebris* type C (~10%). CZ2 (from 9.6 cal kyr BP) displays the rise of warm-associated *Cricotopus aucklandensis* from 3 to 18%, which continues to rise through CZ3a (from 6.7 cal kyr BP) up to 47% as cold-associated taxa (notably *Tanytarsus vespertinus*) decline. Zone CZ3b (0.9 to 0.7 cal kyr BP) contains only two samples and is characterised by a spike in *Polypedilum* (42-60% vs. on average 8% in other zones); *Polypedilum* is a cosmopolitan, speciose taxon whose high abundance is difficult to interpret. CZ4 (0.7 cal kyr BP to present) shows relatively high percentages of cold-associated taxa, such as *Tanytarsus funebris* type A and C (~14% combined vs. ~4% in CZ3a), and *Chironomus* (7% vs. 1% in CZ3a); its timing overlaps with the human era.

3.4.4 Climate reconstructions

All pollen samples have close modern analogues in the training set (all have dissimilarities < 5th percentile value of the dissimilarity distribution of the training set samples; Fig. 3.4a) and all but two show a good fit to MAAT (Fig. 3.4b). Only the samples at 1669 and 8528 cal yr BP show a squared residual higher than the 90th percentile value of the training set squared residuals (= 5.30) and are therefore considered poorly fitted to MAAT. There are 8 samples with a high abundance (> 5%) of rare taxa and/or taxa that are not included in the model (Fig. 3.4a, b). These are samples with high abundances of *Dodonaea viscosa*, which is rare in the training set (Appendix A1, Table A-1). The two MAAT reconstructions show similar trends throughout the record (Fig. 3.5a, b). Temperatures increase rapidly, i.e. from ~9 to 13°C between 16.0 and 12.3 cal kyr BP, and more gradually between 12.3 and 10.0 cal kyr BP. After 10.0 cal kyr BP, the MA technique shows a stable MAATof ~14.5°C, while the partial least squares method produces a lower Holocene MAAT (~13.6°C).
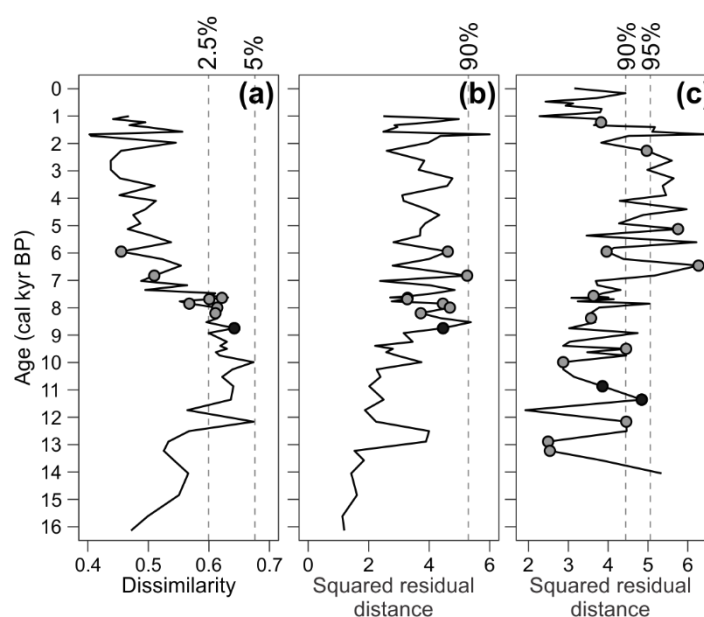
Figure 3.4: Reconstruction diagnostics. Dissimilarity of pollen assemblages to modern analogues (a) with vertical lines representing percentiles of the distribution of the pair-wise dissimilarities for the training set samples. Goodness of fit over time of pollen samples (b) and chironomid samples (c) to MAAT and SmT respectively, with vertical lines representing percentiles of the squared residual distances of training set samples. Samples with > 5% rare taxa (Hill's N2 in training set ≤ 5) and those taxa not included in the inference model are indicated with grey symbols, > 7.5% with black symbols.
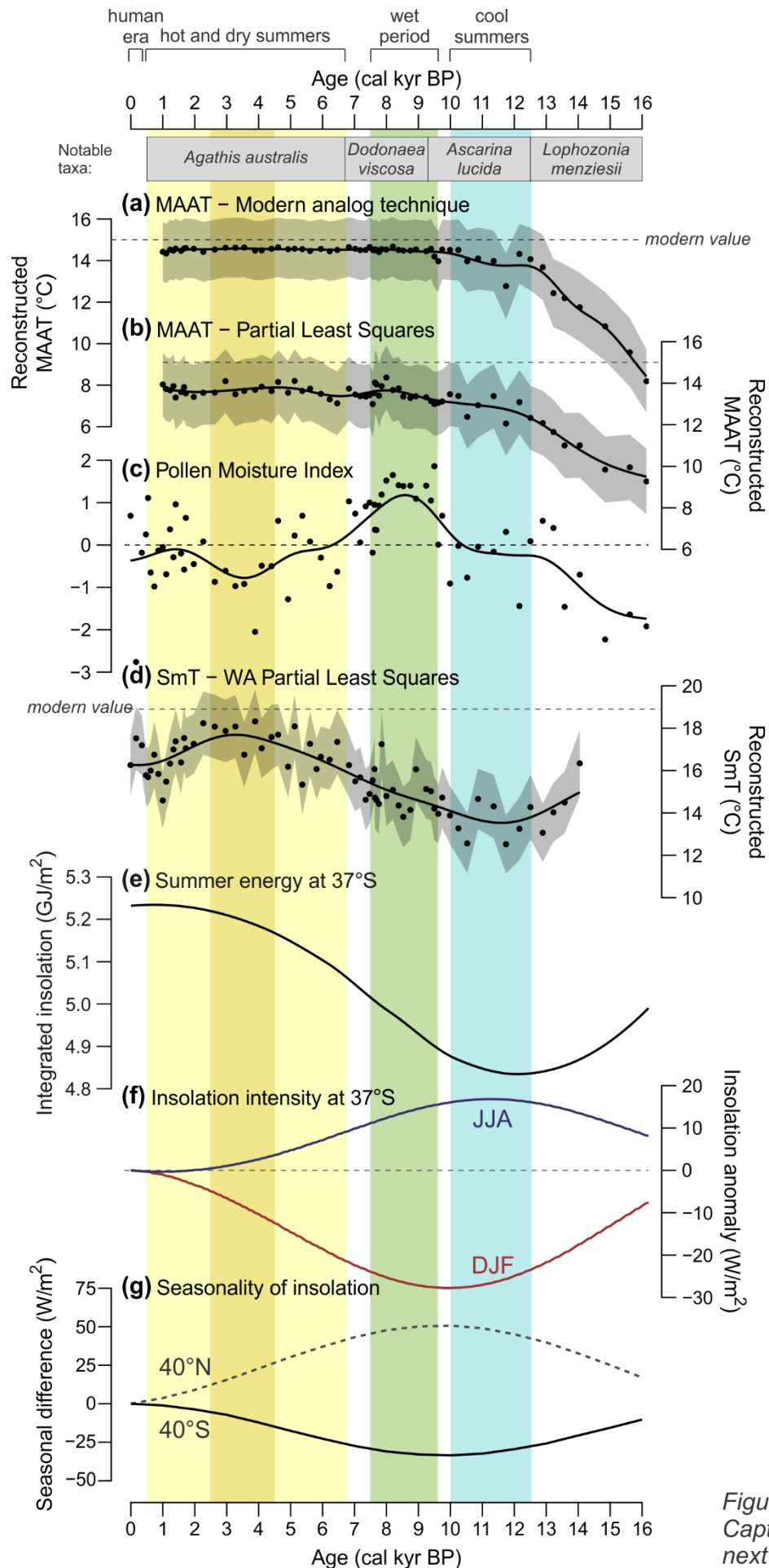
Figure 3.5:
Caption on
next page

Figure 3.5: Synthesis record including reconstructions of MAAT (MA technique (a) and PLS method (b)), effective precipitation (PMI) (c), and SmT (d); points are reconstructed values, line is generalised additive model, shaded areas represent transfer function error (RMSEP). Comparison to insolation parameters: integrated summer insolation ("summer energy") at 37 ^{0}S with a threshold τ of 425 W/m^2 (e), average insolation intensity anomaly at 37^{0}S of December, January, February (DJF) and July, June, August (JJA) (f), and seasonal insolation differences: insolation in DJF less insolation in JJA at 40^{0}S and insolation in JJA less insolation in DJF at 40°N (g). Insolation parameters from: Berger and Loutre (1991), Huybers (2006), and Huybers and Eisenman (2006).

_____

The PMI, a non-quantitative reconstruction of effective precipitation, shows a phase from dry to increasingly wet conditions between 16 and ~13 cal kyr BP, followed by a plateau between 13 and 10 cal kyr BP (Fig. 3.5c). A wet interval occurs between 9.6 and 7.5 cal kyr BP, before PMI gradually decreases, indicating a drying trend, with a minimum around 3.5 cal kyr BP, before increasing again. PMI values are most variable after ~2 cal kyr BP.

For chironomid-derived SmT estimates, we chose the 3-component WA-PLS model (RMSEP$_{jack}$ = 1.21, r^2_{jack} = 0.81) as it displayed lower bias at the higher end of the SmT gradient compared to the PLS model (see Appendix A1). The goodness-of-fit test shows that the fossil chironomid record contains 11 samples that are poorly fitted to SmT (~16% of all samples), and 13 samples that are very poorly fitted (~19% of all samples; Fig. 3.4c). There are 8 samples with a high abundance (> 5%) of rare taxa and/or taxa that are not included in the model. The distribution of these samples seems unrelated to the goodness of fit (Fig. 3.4c). Instead, the samples with a very poor fit include the oldest sample (which had a count sum of only 13.5) and samples with a high reconstructed SmT. The reconstructed SmT is lowest between c. 12.5 and 10.0 cal kyr BP (mean = 14.4°C; Fig. 3.5d). From ~11.5 cal kyr BP, SmT increases until the highest temperature is reached around 3 cal kyr BP (~18°C). Afterwards it rapidly declines to ~16°C. As expected, because the modern training set is based on South Island lakes, reconstructed temperatures are all well below modern day SmT at Lake Pupuke (which is ~19.5°C).

3.5 Discussion

3.5.1 Palaeoecological reconstructions

3.5.1.1 Pollen

The earliest phase of the pollen record (16-12.3 cal kyr BP) depicts a progressive transition from more open mixed conifer-beech forest with prominent shrub understorey to a more closed forest. Previous pollen records from Auckland that trace the vegetation changes of the region during the last glacial maximum and subsequent deglacial transition (Sandiford et al., 2002, 2003; Newnham et al., 2007b) show very similar trends for this period, and in particular the progressive decline of beech and shrub components. During the early-to-mid Holocene, beech virtually disappears as rimu-rātā forest becomes the dominant vegetation in the region. The expansion of the small, predominantly coastal tree *Dodonaea viscosa* occurs from around 9 cal kyr BP, and is consistent with the approaching culmination of postglacial sea level rise (~7-8 cal kyr BP; Clement, 2016) and the establishment of coastal conditions at the study site. The most prominent change in the vegetation record occurs at ~7 cal kyr BP with the expansion of kauri forest, presumably extending into areas previously supporting rimu-rātā forest. Previous pollen records from the AVF region have shown this kauri expansion, along with the other Holocene vegetation changes depicted in the Lake Pupuke record (e.g. at Lake Waiatarua, Newnham & Lowe, 1991; Mt Richmond, Sandiford et al., 2002; Pukaki Crater, Sandiford et al., 2003; Onepoto maar, Augustinus et al., 2011, 2012; and a previous study on Lake Pupuke that focussed on the last ~9 cal kyr BP, Horrocks et al., 2005), although exactly why this profound change developed from this time remains unclear.

These pollen-vegetation reconstructions are important in helping to evaluate the MAAT reconstructions derived from them. Despite the marked changed in vegetation composition, MAATs are comparatively stable after ~10 cal kyr BP (see section 3.5.2 for comparison to other records). We note that the dominant pollen taxa involved are all correlated strongly with MAAT in the training set (i.e. *r* values: 0.59 for *Agathis australis*, 0.57 for *Metrosideros*, 0.51 for *Dacrydium cupressinum*; from Wilmshurst et al., 2007). In addition, all samples display pollen assemblages with close modern analogues in the pollen training set (Fig. 3.4a), and all but two show a good fit with MAAT (Fig. 3.4b). We suggest therefore that

MAAT by itself was not a factor in these marked vegetation changes and in section 3.5.3 explore other climate parameters that may have been important.

3.5.1.2 Chironomids

The chironomid record shows that warm-associated taxa gradually replace cold-associated taxa over time, with a peak in warm-associated taxa between 4 and 2 cal kyr BP (Fig. 3.3b). This pattern is mostly driven by the increase of *Cricotopus aucklandensis* (*β* coefficient: 26.6) and decline of *Tanytarsus funebris*, type C (*β* coefficient: 10.8). Dieffenbacher-Krall et al. (2007) found that SmT is the dominant environmental variable that shows a strong relationship with chironomid abundances. Other variables that were found to be associated with chironomid abundances include conductivity, chlorophyll *a* and organic content of the sediments. There is no stratigraphic or palaeoecological evidence to suggest that these other factors were important in the Lake Pupuke chironomid record. The sediments are laminated consistently throughout, with no visible variations. Lake depth is another variable that has been linked to changes in chironomid assemblages (e.g. Engels et al., 2016), but as Lake Pupuke is currently ~57 m deep any changes in lake level throughout the Holocene were likely to have been relatively small. In addition, because the basin is steep sided, changes in the size of the littoral area due to lake level changes are also unlikely to have been significant. These considerations support our confidence in the SmT reconstructions from Lake Pupuke, but we acknowledge that these other local site variables may also have influenced the chironomid assemblages.

The SmT transfer function we applied to our chironomid data was developed from South Island lakes and covers an SmT gradient from ~6 to 16°C (Dieffenbacher-Krall et al., 2007). This is not ideal for reconstructing SmT at Lake Pupuke, where modern day mean temperature for the summer months is ~19.5°C. The absence of training set lakes from the warmer North Island may shift taxon optima towards the cooler end of the SmT gradient and reconstructions from Lake Pupuke may be underestimated. This is confirmed by our goodness-of-fit test, which shows that samples with a high reconstructed SmT display a very poor fit (Fig. 3.5b). This limitation should be borne in mind for any consideration of the absolute reconstructed temperatures above 16°C from Lake Pupuke. Nonetheless, we suggest that the trends and relative changes in SmT are robust reconstructions. We also point out that the poorly fitted samples that display high temperatures are also the samples

with high abundances of *Cricotopus aucklandensis*, which is one of the taxa most strongly associated with high SmT.

Our SmT reconstruction from Lake Pupuke highlights the lack of North Island lakes in the current chironomid training set and associated difficulty the model has with reconstructing high SmT. This work demonstrates the importance of undertaking future work aimed at expanding the chironomid training set with data from northern New Zealand sites.

3.5.2 Temperature reconstructions in a regional and global context

Our Lake Pupuke record shows that after the Last Glacial–Interglacial transition, from ~12 cal kyr BP, MAAT stabilises around 14.5°C, within the range of modern MAAT at the site (Fig. 3.5a, b). There is no evidence for a Holocene Thermal Maximum (HTM), which is consistent with mean annual sea surface temperatures (SST) inferred from nearby marine records (i.e. H214 and MD97-2121, Fig. 3.1a; Prebble et al., 2017). Prebble et al. (2017), studied seven marine records on a meridional transect of the southwest Pacific Ocean and found that only records in the region immediately south of the subtropical front display an early-Holocene warm peak, while records along the east coast of the North Island show stable SSTs throughout the Holocene. The stability of our reconstructed Holocene MAAT is inconsistent with the regional temperature reconstruction for the southern mid latitudes reported by Marcott et al. (2013), which shows an early-Holocene maximum (as does the global mean temperature reconstruction). In light of our results and the observations of Prebble et al. (2017), we suggest that the regional temperature reconstruction for the southern mid latitudes reported by Marcott et al. (2013) may reflect records from sites that are predominantly located between the subtropical front and polar front. Alternatively, the modern relationship between vegetation and MAAT may have broken down during the early Holocene under the influence of secondary, confounding variables, such as minimum temperatures or precipitation (see section 3.5.3). However, we have no way to test or control for this and must therefore keep the limitations of the model in mind.

In the Northern Hemisphere, North Atlantic SST records show a very clear HTM during the early Holocene, followed by a cooling trend (Marcott et al., 2013), in line with declining northern summer insolation throughout the Holocene. Marsicek et al. (2018) find a similar cooling trend in European and North American pollen records but only after ~5.5 kyr BP.

They explain that the HTM and subsequent cooling trend is delayed in continental records compared to the SST records in this region, because remnant ice sheets caused continental air temperature depression during the early Holocene. Subsequent warming is in line with model simulations forced by declining albedo, elevation, and meltwater of ice sheets. The cooling trend, consistent with declining northern summer insolation, only commences after 5.5 kyr in the continental records (cf. Renssen et al., 2009).

This match between summer cooling and declining summer insolation in the Northern Hemisphere mirrors our findings for northern New Zealand. The SmT record from Lake Pupuke indicates that summers were cool during the early Holocene until ~10 cal kyr BP, and then warmed progressively until ~3 cal kyr BP (Fig. 3.5d). This warming trend is opposite to the Northern Hemisphere records, but in line with increasing summer insolation intensity at the southern mid-latitudes (Fig. 3.5f) and correlates well with total summer energy (integrated summer insolation of days whose diurnal average insolation > 425 W/m^2; Fig. 3.5e) (Berger & Loutre, 1991; Huybers, 2006; Huybers & Eisenman, 2006). The similarity of the SmT and local summer insolation records implies a causal relationship throughout the Holocene. Putnam et al. (2012) describe a similar inter-hemispheric mirroring pattern, where glacier behaviour in New Zealand's Southern Alps (Fig. 3.1a) correlates with southern summer insolation, whereas glaciers in the European Alps were found to respond to northern summer insolation forcing.

The observation that the expression of the HTM in different regions might be influenced by local insolation forcing and the presence or absence of ice sheets is not a new idea (e.g. Renssen et al., 2009; Fischer et al., 2018; Marsicek et al., 2018). Fischer et al. (2018) admit that the complex regional expression of the HTM makes it difficult to draw a unique global picture but argue that it can still serve as an example of regionally warmer conditions. However, our findings indicate that early Holocene climate might not offer an instructive analogue for the impact of future warming in northern New Zealand, because we do not find temperatures higher than present.

3.5.3 Seasonality and effective precipitation

Although the MAAT reconstruction from the pollen assemblage data is stable from 10 cal kyr BP onwards, there are distinct changes in vegetation throughout the Holocene (Fig.

3.3a), as has been observed previously in northern New Zealand (e.g. Newnham & Lowe, 1991; Sandiford et al., 2002, 2003; Horrocks et al., 2005; Augustinus et al., 2011, 2012). Specifically, *Metrosideros* and *Ascarina lucida* are abundant during the early Holocene, while *Dodonaea viscosa* rises from 9.3 cal kyr BP. The most pronounced change occurs at 6.7 cal kyr BP, when *Agathis australis* (kauri) abruptly becomes the dominant taxon at the expense of, among others, *Dacrydium cupressinum*. We propose these shifts are driven by changes in seasonality and effective precipitation (cf. Sikes et al., 2013).

The early Holocene (12.3–9.3 cal kyr BP) corresponds to the period of lowest reconstructed SmT, while MAAT has reached close to modern values. Our combined temperature records thus indicate low seasonality (cool summers and warm winters), conditions favourable to *Metrosideros* and *Ascarina lucida*, which have reached maximum abundances at this time. These angiosperm trees are frost sensitive and require high humidity (McGlone & Moar, 1977; Wilmshurst et al., 2007) so would have benefited from mild, moist winters. The PMI rises during this period to reach a peak at ~9 cal kyr BP, indicating moisture availability was unlikely to be a limiting factor. Low seasonality has previously been invoked to explain early Holocene vegetation patterns in New Zealand (e.g. Newnham & Lowe, 1991; Wilmshurst et al., 2002; McGlone and Basher, 2010; McGlone et al., 2010b; Jara et al., 2015) as well as other local trends (e.g. Stephens et al., 2012b). Our combined MAAT and SmT reconstructions support this hypothesis.

After the early Holocene, effective precipitation (PMI) increases, signalling a wet period between 9.6 and 7.5 cal kyr BP concomitant with a peak in *Botryococcus* algae and the rise of *Dodonaea viscosa*. Summer insolation intensity and seasonality are still relatively low at this time. The increase in *Botryococcus* is consistent with an expanded lake area and higher lake levels in response to increased precipitation. Jara et al. (2017) reconstruct a very similar period of wettest conditions at Moanatuatua (c. 140 km southeast of Lake Pupuke, at 38°S; Fig. 3.1a) between 10 and 6 cal kyr BP, confirming that the increase in effective precipitation is likely a regional phenomenon, rather than a local one.

Effective precipitation starts to decline along with rising summer temperatures from 7.5 cal kyr BP onwards. The warmest and driest conditions occur during the late Holocene, between ~4.0 and 2.4 cal kyr BP, when summer energy reaches peak values. The warm, dry summer conditions overlap with the rise in kauri abundance, which is a widespread phenomenon

observed in northern New Zealand pollen records (e.g. Newnham & Lowe, 1991; Newnham, 1999; Ogden et al., 1992; Newnham et al., 1995). Less clear is why this profound shift occurred at this time, from around 7 cal kyr BP. Climatic explanations are confounded by the severely modified modern distribution of kauri, its longevity (often > 600 years) and periodic regeneration (Ogden et al., 1992). Nevertheless, it has been established that kauri growth is enhanced during sunny, dry summers (September to April) (Ogden & Ahmed, 1989), conditions that are clearly shown in our temperature and effective precipitation reconstructions (Fig. 3.5c-d) as coinciding with the expansion of kauri during the mid-to-late Holocene. Previous climate inferences drawn from this kauri expansion could be questioned on grounds of circularity of argument, but in this case the SmT and PMI reconstructions are independent of the kauri pollen record.

3.5.4 Potential causes of wet and dry periods

The early-to-mid Holocene wet period followed by increasingly dry conditions as apparent from the PMI and the appearance of kauri, indicate that effective precipitation declined over time, which is in line with increasing evapotranspiration from 7.5 cal kyr BP onwards due to higher SmT. Alternatively, the decline in effective precipitation could have been caused by a decline in rainfall, consistent with a progressively southwards shifting SWW belt, causing westerly winds over New Zealand to decline (Kidston et al., 2009), while strengthening in the core. This is broadly in line with the findings of Moros et al. (2009) who report weaker westerlies and poleward shifting fronts in South Australia (~36°S; in the northern margin of the SWW belt) and Saunders et al. (2018), who report generally higher wind intensity after ~7 cal kyr BP at Macquarie Island (54°S; in the core of the SWW belt). The southward shift of the SWW belt would have allowed more influence from the high pressure belt in northern New Zealand, which is associated with lighter winds in general and clearer skies. Strong influence from the high-pressure belt is associated with a positive SAM at the present day, causing warmer and drier conditions in the Auckland region (Ummenhofer & England, 2007; Kidston et al., 2009), such as we find at Lake Pupuke in the mid-to-late Holocene.

Dry conditions in northern New Zealand during the mid-to-late Holocene also concur with increased ENSO activity and more frequent El Niño events reported after 5 kyr BP (Donders et al., 2005, 2007). In the South Island, Wilmshurst et al. (2002) and Jara et al. (2015) invoke increased ENSO activity during the mid-to-late Holocene to explain climate patterns, i.e. wet

episodes at Eweburn Bog (southern South Island) and cooling at Adelaide Tarn (northern South Island) respectively. Although the climate shifts are different between the sites, they are consistent with an increasing El Niño influence, because these conditions cause reduced precipitation over the North Island and increased precipitation over the southwestern edge of the South Island, due to enhanced southwesterly airflow (Ummenhofer & England, 2007). However, we acknowledge that in invoking increased El Niño influence during the late Holocene, the modern ENSO spatiotemporal patterns have only been observed at much shorter timescales (decadal-to-interannual) than the long term, centennial-scale drying trend we find at Lake Pupuke.

Generally declining westerly wind strength in New Zealand (the northern margin of the SWW belt) and increasing El Niño frequency over the course of the Holocene are both consistent with findings of a southwards shifting intertropical convergence zone (ITCZ) during this period, causing a southward shift of the subtropical front (Haug et al., 2001, Putnam et al., 2012). Haug et al. (2001) postulate that this shift was caused by changes in the seasonality of insolation: Southern Hemisphere insolation is more seasonal during the late Holocene, while Northern Hemisphere insolation is less seasonal (Berger & Loutre, 1991; Fig. 3.5g). However, the observed wet period at Lake Pupuke (9.6–7.5 cal kyr BP) does not fully overlap with the period of lowest seasonality (~12.5–7.5 kyr BP). Other drivers may be at play here. Jara et al. (2017) found a similar PMI trend at Moanatuatua and argued that SWW were weaker in northern New Zealand between 12 and 8 cal kyr BP (see also McGlone et al. (2010a) on Campbell Island and Fletcher & Moreno's (2012) review of SWW reconstructions). The wet period at Moanatuatua (10–7 cal yr BP) was attributed to an intensification of the subtropical circulation resembling the present-day La Niña state, bringing more subtropical moisture (Jara et al., 2017). However, Kidston et al. (2009) show that more northerly influence is associated with fewer cyclones and more anticyclones in most of the North Island, including the Auckland area, which is therefore inconsistent. In addition, as with arguments for stronger El Niño influence after 5 kyr BP, there is an implicit assumption that the spatial patterns observed in modern ENSO cycles were played out over longer, centennial-millennial timescales during the early Holocene.

Whether the trends in effective precipitation at Lake Pupuke are related to atmospheric circulation (precipitation amount), SmT (evapotranspiration), or a combination of both, the apparent relationship between SmT and Southern Hemisphere summer insolation

parameters is strong, as is their similarity to the PMI record after ~8 cal kyr BP. After ~2 cal kyr BP the SmT and insolation trends diverge, which in part could be explained by anthropogenic disturbance after 700 cal yr BP. Nonetheless, insolation seems to have been a dominant guide for most of the Holocene and a perfect match is impractical with any driver or cause. A further question remains as to whether this association reflects direct radiative forcing of local temperature or indirect forcing via regional-scale climate processes such as changing atmospheric or ocean circulation patterns.

3.6 Summary and conclusions

We reconstructed MAAT, SmT and effective precipitation over the last 16,000 years in northern New Zealand from the sediments of Lake Pupuke. SmT was inferred from chironomid assemblages, independent from MAAT and effective precipitation, which were reconstructed from pollen analysis. Pupuke MAAT increases coming out of the Last Glacial and stabilises after 10 cal kyr BP at close to modern values, consistent with nearby marine records of SST. There is no evidence for an early HTM. The vegetation changes that occur during the MAAT plateau are likely the result of changing humidity and seasonality. SmT increases between 10 and 3 cal kyr BP, in correspondence with local summer insolation parameters. Our results highlight the importance of undertaking future work aimed at expanding the chironomid training set with data from northern New Zealand sites.

Early-Holocene low seasonality (12–9.3 cal kyr BP) is indicated by cool summers reconstructed from the chironomid assemblages and mild winters inferred from the vegetation community and was likely driven by low local summer insolation intensity or summer duration. An early-to-mid-Holocene wet period (9.6–7.5 cal kyr BP) corresponds to relatively high westerly strength, consistent with a steep latitudinal insolation gradient. Mid-to-late-Holocene summers following the wet period were hot and dry (especially 4.0–2.4 cal kyr BP), allowing kauri to expand throughout northern New Zealand. High SmT would have increased evapotranspiration and decreased effective precipitation at the time. Whether low effective precipitation was primarily caused by increased evapotranspiration, southward displaced SWW, increased El Niño frequency, or some combination of these factors remains unclear. However, a strong correspondence is observed between SmT, Southern Hemisphere summer insolation parameters and PMI after 8 cal kyr BP.

The interpretation of our integrated MAAT, SmT and effective precipitation record leads to two main conclusions. First, stable MAAT and increasing SmT at Lake Pupuke are consistent with increasing local summer insolation intensity and total summer energy during the course of the Holocene, mirroring the pattern observed in the Northern Hemisphere: there is no evidence for an early or mid-Holocene HTM, but rather, an early Holocene with cool summers and mild winters. Therefore, the early-to-mid Holocene does not appear to be a good analogue for future projected global warming impacts in northern New Zealand. Second, although MAATs were stable, vegetation at Lake Pupuke responded to changing seasonality and effective precipitation, highlighting the importance of multiproxy studies in addressing the complexities of climate patterns superimposed on MAAT reconstructions. Where Marcott et al. (2013) could not rule out seasonal and precipitation bias (cf. Liu et al., 2014), we actively searched for the influence of these factors on our proxies (cf. Sikes et al., 2013), leading to a better understanding of Holocene climate change complexities. While global compilations of temperature observations and global climate model simulations may be useful for the mitigation of future climate change, they may not be representative of a particular region and may overlook important aspects of Earth's climate system.

Supplementary Information

Appendix A1: Temperature inference models diagnostics
- Table A-1: Pollen taxa included in the MAAT inference model and their Hill's N2
- Table A-2: Chironomid taxa included in the SmT inference model, their Hill's N2, and their β coefficients in the Weighted Average-Partial Least Squares model.
- Table A-3: Performance of WA-PLS and PLS SmT inference models.
- Fig. A-1: Residual plots of the SmT inference models

Appendix A2: Rare pollen taxa
- Table A-3

Acknowledgements

on an earlier version of our discussion, and two anonymous reviewers for their constructive criticism. Postgraduate study of VB is funded through Marsden Fund project 14-UOA-040.

Chapter 4

An extended chironomid training set for reconstructing
New Zealand summer air temperatures
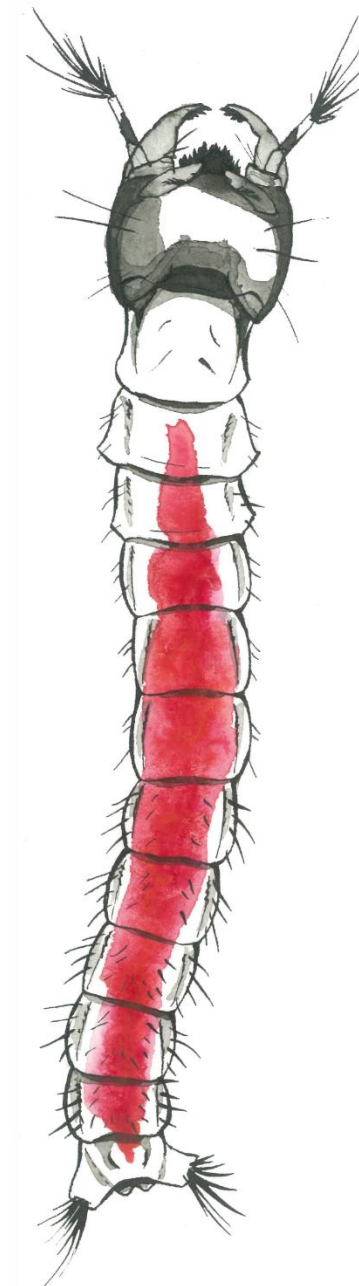
Preface

This chapter was written in publication style but will probably not be submitted for publication before more North Island lakes are added to the training set (see Chapter 6, section 6.3). The results described in this chapter relate to objective 5: Develop an improved chironomid-SmT model based on an extension of the modern training set, using both traditional and novel methods. As mentioned in Chapter 1, much of the data described in this chapter was provided by Marcus Vandergoes. I analysed several sediment surface samples to add to the training set, and the data analysis and interpretation are my personal contributions.

Abstract

While the number of quantitative temperature reconstructions in the Southern Hemisphere is increasing at a steady pace, many regions remain underrepresented. In this study we analyse a 104-lake chironomid training set from New Zealand, an extension of an earlier published 60-lake training set used to infer summer air temperature (SmT) (Dieffenbacher-Krall et al., 2007). Chironomids have been used to reconstruct a variety of environmental variables but have been found to respond most strongly to summer air temperature (SmT). If used in combination with mean annual air temperature proxies, chironomids can give insights into the seasonal aspects of past climate change, providing new insights into long-term climate dynamics. The original model covered an SmT gradient from ~6.5 to 15°C and was developed primarily for southern New Zealand. The model was not able to reconstruct Holocene SmT at Lake Pupuke in northern New Zealand, where modern day SmT is 18.9°C. We confirm that SmT is the most important variable to explain chironomid assemblage variance in the extended training set, but find potential confounding factors, some of which correlate with SmT. We apply seven different transfer function techniques: one assemblage approach, two multivariate regression models, and four machine learning techniques. The latter are supposedly more robust in the presence of confounding variables. We discuss the differences between the seven models and apply them to the Lake Pupuke chironomid record to recalculate SmT. Although the best-performing model (weighted-averaging partial least squares) based on the extended training set is better able to reconstruct temperatures close to modern day values at Lake Pupuke, the general trend is the same as the reconstruction based on the 60-lake training set. The higher end of the temperature gradient is still underrepresented in the extended training set, leaving room for further improvement.

4.1: Introduction

Quantitative climate reconstruction from downcore proxy records has become a standard tool in paleoclimate research. Sophisticated proxy-climate models have helped increase our understanding of past climates. For instance, quantitative temperature reconstructions have helped us identify potential past analogues for future warming scenarios (Fisher et al., 2018), allowing us to address some critical environmental questions of today, such as what effect a 2°C increase in temperatures will have on ecosystems.

Chironomids, or non-biting midges, are excellent palaeoenvironmental proxies, because their larval remains (head capsules) are taxon diagnostic, preserve well and are generally abundant in lake sediments. In addition, most chironomid taxa live within the lake and respond to environmental change rapidly, as the winged adults disperse easily and can produce one or more generations per season (Ruiz et al., 2006). Chironomids have been used to reconstruct a variety of environmental variables, including hypolimnetic oxygen (Little & Smol, 2001), chlorophyll-*a* (Woodward & Shulmeister, 2006) and water depth (Engels et al., 2016). In New Zealand (and other locations, e.g. Brooks & Birks, 2000; Heiri et al., 2003; Rees & Cwynar, 2010) they have been found to respond most strongly to temperature, summer air temperature (SmT) specifically (Woodward & Shulmeister, 2006; Dieffenbacher-Krall et al., 2007).

As a proxy for SmT, chironomids work especially well in combination with mean annual air temperature (MAAT) proxies such as pollen (Wilmshurst et al., 2007) to reveal seasonal aspects of past climate change and ecosystem response that cannot be explained by MAAT alone (see Chapter 3). As chironomids respond faster to environmental stimuli than many other common proxies, they also provide a tool to identify abrupt, short term, or subtle climatic changes that might not appear in other records.

A chironomid-SmT inference model using the transfer function approach (Dieffenbacher-Krall et al., 2007) has been applied in southern New Zealand (Vandergoes et al., 2008), providing insight into changing seasonality between 17.5 and 10 kyr BP. However, the training set was not designed to reconstruct temperatures above ~15.2°C. This limitation became apparent when the model was applied to sub-fossil chironomid remains in a Holocene sequence at Lake Pupuke, northern North Island (Chapter 3) where the modern

SmT exceeds the range of the training set. Nevertheless, the Pupuke results showed that this approach can provide valuable new insight into Holocene climate dynamics at the same time as demonstrating the need to extend the range of the modern training set.

Here we analyse an extension of the Dieffenbacher-Krall (2007) training set, which consists of 60 lakes, covering a gradient from ~6.5 to 15°C. Our new, extended training set contains 104 lakes and covers a gradient from 6.5 to 19.5°C. In addition to recalculating a multivariate regression-based model similar to the one previously developed, we apply an assemblage approach and machine learning (ML) algorithms (see section 4.2 for rationale). We also test the impact of count size for both training set and fossil samples on temperature estimates. Finally, we present a revised Holocene SmT record from Lake Pupuke and compare it with the original reconstruction.

4.2: Model types and assumptions

4.2.1 The problem of confounding variables

All models discussed in this chapter are based on the transfer-function approach: using the modern distribution of taxa and their relationship with the environment to reconstruct the environment of the past from fossil assemblages of these taxa. To infer the relationship between taxa and environmental variables, we use a training set of modern lakes. The assumptions underlying this approach have been discussed and refined by many authors (Imbrie & Kipp, 1971; Imbrie & Webb, 1981; Birks et al., 1990, 2010), but Juggins (2013) points out that two key assumptions are often ignored and sometimes may be violated. They are: the environmental variable(s) to be reconstructed is, or is linearly related to, an ecologically important determinant in the system of interest; and environmental variables other than the one of interest have negligible influence, or their joint distribution with the environmental variable does not change with time.

In practice, Juggins (2013) elaborates, advocates of the transfer function approach often identify the environmental variable that has the best explanatory power in their training set, but fail to consider or critically discuss the effect of potential confounding variables. Often the exact mechanism by which the environmental variable influences the organism of interest is unknown or poorly understood, and it actually might not have a direct ecological

effect. Instead, the statistical relationship identified reflects complex, underlying environmental gradients. While the model based on this variable might show good performance (as determined by cross-validation; see section 4.3.4.2) and consistent reconstructions, the interpretation might be spurious if the joint distribution of the variable and underlying confounding factors in the training set has changed over the period of study (Velle et al., 2010).

In the case of chironomids, many in-lake variables have been found to influence taxon assemblages (see Introduction). Even if temperature is the main explanatory variable and affects chironomid larvae directly (e.g. through growth or development), it is likely that temperature also affects limnological variables such as nutrient conditions or deep-water oxygen concentrations which in turn affect the survival and reproduction of aquatic biota (Eggermont and Heiri, 2012). Moreover, these limnological variables may also be influenced by factors other than temperature and may change independently of temperature. The exact nature of the mechanism responsible for the observed relationship between chironomid assemblages and air temperature remains uncertain, a situation that Juggins (2013) warns can lead to problems when interpreting results.

We cannot resolve the matter of potential confounding variables in this study. However, we will critically examine our results, attempt to identify possible secondary variables in our explanatory analyses, and consider the potential effect of these variables on our reconstructions. We follow the recommendations of Juggins (2013), to explicitly consider the possibility of confounding factors and related issues with our dataset and where appropriate apply caution to interpretation. In addition, we apply several different transfer function techniques, to get a better grip on the source of bias in our dataset.

4.2.2 Model descriptions

We tested seven different transfer functions for quantitative climate reconstruction: one assemblage approach, two multivariate regression models, and four ML algorithms. The assemblage approach, or modern analogue (MA) technique, involves calculating the (weighted) mean SmT of the closest analogous proxy assemblages in the training set to the sample of interest. In this way, the MA technique considers the fossil assemblage as a whole and the relative abundances of all different taxa (Birks et al., 2010). The main benefit of the

MA technique is that it encourages the intuitive interpretation of results. If there are no close modern analogues in the training set, the standard error will be high, highlighting the problem (ter Braak, 1995). The standard error will also be high if the fossil assemblage is similar to several modern analogues that are dissimilar in terms of their modern climate affiliation (i.e. due to confounding variables or bimodal taxon-climate relationships). Although MA models are adept at revealing such problems, they are not well suited to reconstructing past climate if there are no close modern analogues, or if there are multiple analogues with dissimilar climate conditions.

The multivariate calibration approach estimates parametric functions for all taxa present. It allows for more extrapolation than the MA technique in the absence of modern analogues. We attempt two calibration types, i.e. partial least squares (PLS), which calculates linear parametric functions, and weighted averaging-PLS (WA-PLS), which calculates unimodal parametric functions. PLS and WA-PLS are more susceptible to problems arising from confounding variables than the MA technique. Two other potential weaknesses of transfer functions are sensitivity to the distribution of training set samples along the gradient of interest and the underlying assumption that taxon responses are linear or unimodal (ter Braak & Looman, 1986; Birks et al., 2010). The former is a problem when training set lakes are not evenly distributed over the environmental gradient of interest. The latter can be a problem when multiple species with different climatic preferences are captured within one taxon, due to limitations of identification. Such amalgamated taxa may show a bimodal response instead.

ML algorithms, specifically tree-ensemble models, are supposedly more robust to the effects of confounding variables and noise (Elith et al., 2006; Velle et al., 2010; Juggins et al., 2015; Salonen et al., 2016), and are therefore especially promising in light of Juggins' (2013) warnings concerning transfer functions. Tree-ensemble models can also overcome some of the other potential problems of multivariate regression models, as they make minimal assumptions about the taxon-climate relationships. Salonen et al. (2016) list several of their potential advantages: their ability to model complex (including bimodal) taxon responses, their ability to deal with a diverse combination of taxon response shapes (as opposed to only linear or unimodal), their ability to model interactions between predictors (taxa), their selectivity in only using taxa that show an important response along the climatic gradient, and their sensitivity to rare taxa. A drawback of ML techniques is that they require large

training sets (how large depends on the nature and strength of the relationship between proxy and environmental variable) and so may suffer from bias-towards-the-mean problems when this condition is not met (see section 4.4.2.2).

Here, we specifically attempt to create ensembles of regression trees and explore four methods of calculating and compiling these regression trees: bagging, random forest (RaF), rotation forest (RoF) and boosted regression trees (BRT). The four methods differ in the way they introduce diversity into the tree ensemble. Bagging is the simplest implementation of aggregated regression trees: each tree is constructed independently using a bootstrap sample of the training set and the mean of all tree outcomes is taken for prediction (Breiman, 1996). RaF is similar to bagging, but a layer of randomness is added as the regression trees are constructed using a subset M of chironomid taxa (the predictors) at each node (Breiman, 2001; Liaw & Wiener, 2002). For RoF, the chironomid taxa are divided into K random subsets for each tree and PCA is applied to each subset, creating a random rotation of the predictor space (Rodríguez et al., 2006). BRT is different from the other three tree-ensemble models, as the individual trees are not independent from each other: successive trees give extra weight to points incorrectly predicted by earlier predictors and a weighted mean is calculated for prediction at the end (De'ath, 2007).

4.3: Methods

4.3.1 Sample collection and analysis

4.3.1.1 Surface sediment samples

During the austral summers of 2003 to 2012, surface sediment samples were collected from 104 lakes and tarns throughout New Zealand (Fig. 4.1, Table 4.1); this includes the 60 lakes of the original training set (Dieffenbacher-Krall et al., 2007). Lakes were selected with conditions as near to natural as possible to minimise anthropogenic disturbance, along both an altitudinal and latitudinal range to cover a broad summer temperature (SmT) gradient. Sediment-water interface samples were extracted from the deepest part of each lake using a 0.5-m Hongve-style gravity corer (Wright, 1990; Dieffenbacher-Krall et al., 2007). Subsamples from 0-1 cm depth and 1-2 cm depth were transferred to polythene bags and stored at 4°C until processing.
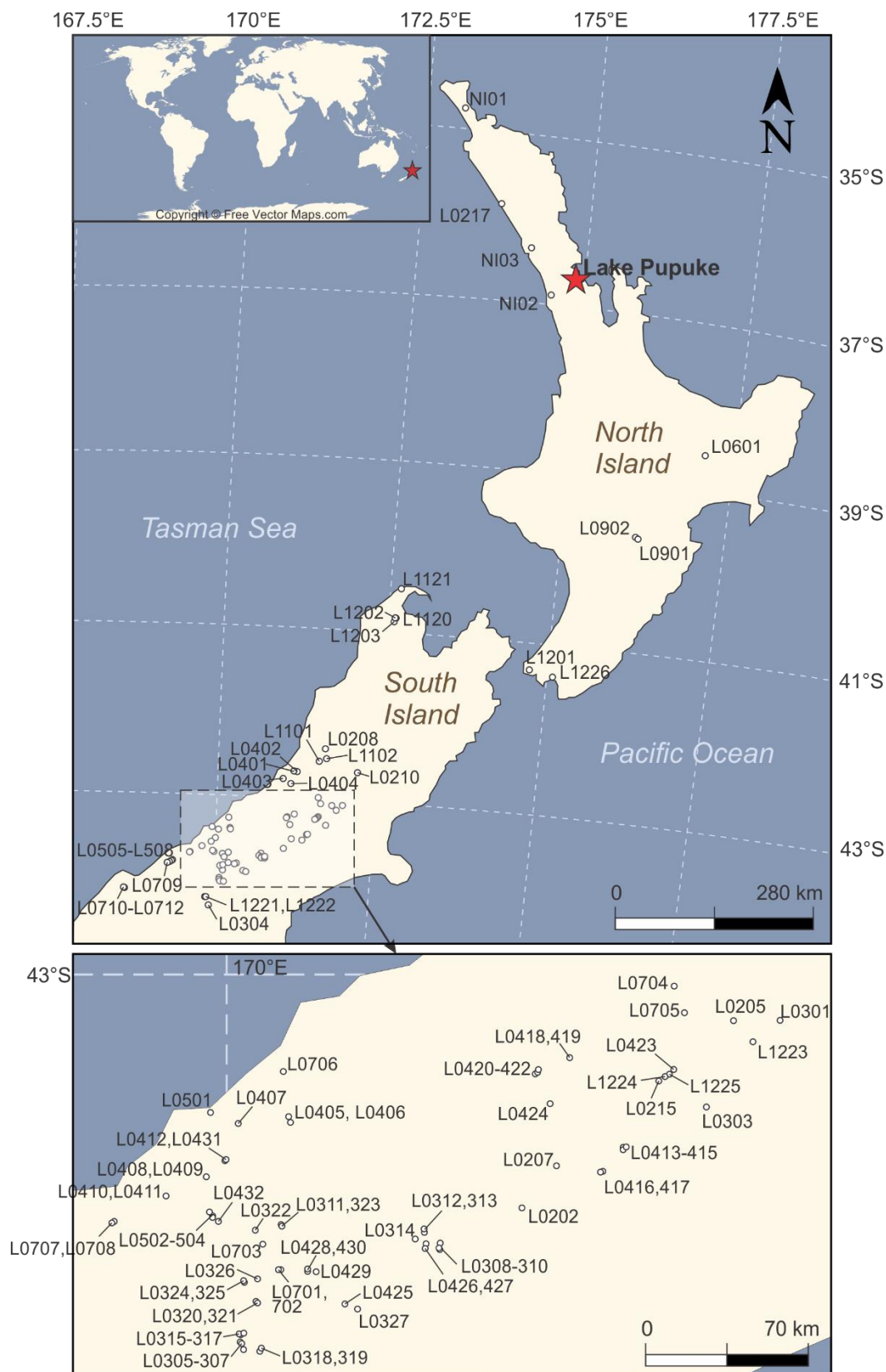
Figure 4.1: Map of New Zealand with locations of training set lakes and Lake Pupuke. See Table 4.1 for details, including names, for each training set lake.

Table 4.1: All 104 training set lakes sorted by latitude (north to south). See Figure 4.1 for locations. See Table 4.2 for the full suite of measured variables; column "# variable measured" refers to the number of these variables that were measured at each lake. Columns "# head capsules", "Taxon richness" and "Shannon's diversity" refer to parameters of the chironomid assemblages found in the lakes' sediments.

Lake ID	Name	Latitude	Longitude	Altitude (m)	#variables measured	#head capsules	Taxon richness	Shannon's diversity
NI01	Te Kahika	-34.625	172.999	16	7	142.5	13	1.38
L0217	Waikere	-35.799	173.629	79	19	124.0	11	1.23
NI03	Humuhumu	-36.328	174.122	51	7	127.5	14	2.26
NI02	Kawaupaka	-36.894	174.459	69	7	100.5	12	1.34
L0601	Kanohirua Lake	-38.592	177.023	775	30	149.5	16	1.82
L0901	Lake Colenso	-39.670	176.136	712	27	117.5	12	1.75
L0902	Tupari Tarn	-39.688	176.182	1377	27	89.0	9	1.76
L1121	Kaihoka Lake	-40.554	172.602	34	13	122.5	16	1.98
L1202	Lake Clara	-40.896	172.558	1330	27	99.5	11	1.64
L1120	Darby Pond	-40.900	172.569	1002	13	608.0	24	2.43
L1203	Adelaide Tarn	-40.942	172.543	1250	27	120.0	14	2.20
L1201	Karori	-41.298	174.745	170	27	77.0	17	2.47
L1226	Lake Pounui	-41.345	175.113	10	27	111.5	14	1.98
L0208	Hochstetter	-42.448	171.670	251	22	145.5	20	2.21
L1102	Haupiri	-42.568	171.693	195	19	96.5	24	2.68
L1101	Lady	-42.600	171.575	119	19	53.0	21	2.87
L0210	Katrine*	-42.719	172.201	523	22	40.5	12	2.10
L0401	Lake Mudgie	-42.723	171.176	150	27	122.0	10	1.66
L0402	Okuku Reservoir	-42.731	171.230	200	27	109.0	17	2.40
L0403	Overlook Hill	-42.820	171.006	150	27	119.5	14	1.47
L0404	Manks Tarn	-42.873	171.130	120	27	123.5	19	2.31
L0704	Bealy Spur 1	-43.029	171.579	1041	29	123.5	10	1.64
L0705	Lagoon Saddle	-43.050	171.596	1175	29	126.5	8	1.31
L0301	Letitia	-43.055	171.949	589	30	120.5	17	2.06
L0205	Grasmere	-43.064	171.776	584	22	63.0	8	1.07
L1223	Lake Marymere	-43.117	171.854	621	26	213.0	14	1.48
L0418	Mistake Hill 1	-43.196	171.189	1850	30	98.5	9	1.51
L0419	Mistake Hill 2	-43.196	171.183	1660	30	134.0	19	1.77
L0423	Mystery Tarn	-43.207	171.572	850	30	138.0	16	2.07
L1225	Lake Catherine	-43.221	171.560	670	26	137.0	14	2.18
L0421	Totara Stm 2	-43.233	171.075	1710	30	101.5	9	1.46
L1224	Lake Ida	-43.235	171.539	670	28	157.1	19	2.23
L0422	Totara Stm 3	-43.240	171.076	1620	30	169.0	9	1.29
L0215	Selfe	-43.242	171.519	574	22	107.5	9	1.45
L0420	Totara Stm 1	-43.243	171.063	1730	30	84.0	7	0.61
L0706	Alpine Lake	-43.288	170.138	101	30	132.5	14	1.11
L0303	Lyndon	-43.304	171.700	841	30	290.0	17	2.12
L0424	Quagmire Tarn	-43.324	171.122	750	30	149.5	12	2.05
L0405	Lake Wombat	-43.405	170.167	250	30	129.5	16	2.19
L0501	Galway Tarn	-43.410	169.878	132	30	138.0	19	2.34
L0415	Middle Ck 3	-43.421	171.423	1620	30	133.5	6	0.85
L0406	Peters Pool	-43.421	170.172	230	30	157.0	14	1.94
L0414	Middle Ck 1	-43.422	171.410	1820	30	123.5	3	0.28
L0413	Middle Ck 2	-43.425	171.407	1800	30	124.5	9	0.80
L0407	Lake Gault	-43.432	169.984	330	30	132.5	14	1.80
L0207	Heron	-43.486	171.165	691	22	139.5	14	2.27
L0416	Taylor Range 1	-43.492	171.334	1865	30	147.5	7	1.33
L0417	Taylor Range 2	-43.492	171.341	1810	30	166.0	9	1.84
L0412	Copland Range 2	-43.534	169.945	1330	30	167.0	8	1.08
L0431	Copland Range 1	-43.537	169.942	1440	30	108.5	8	1.02
L0409	Copland Range 4	-43.582	169.876	1180	30	103.5	8	1.51

Table 4.1 - continued

Lake ID	Name	Latitude	Longitude	Altitude (m)	#variables measured	#head capsules	Taxon richness	Shannon's diversity
L0408	Copland Range 3	-43.585	169.881	1270	30	131.5	6	1.10
L0202	Clearwater	-43.605	171.049	667	19	118.0	12	2.18
L0411	Mt Jacob 2	-43.639	169.731	1200	30	170.0	13	1.36
L0410	Mt Jacob 1	-43.642	169.738	1199	30	113.5	13	1.98
L0312	Second Waterfall Stm 1	-43.678	170.691	1810	27	159.9	8	1.07
L0502	Horace Walker 1	-43.681	169.909	1036	30	137.0	10	1.29
L0313	Second Waterfall Stm 2	-43.685	170.695	1870	27	112.5	8	1.21
L0503	Horace Walker 2	-43.685	169.910	1022	30	157.5	10	1.52
L0504	Horace Walker 3	-43.686	169.910	1016	30	131.5	10	1.57
L0311	Blue Lakes 1	-43.694	170.164	840	27	127.0	13	2.24
L0323	Blue Lakes 2	-43.698	170.166	840	27	139.5	15	2.15
L0432	Horace Walker 4	-43.700	169.936	1004	26	147.5	8	1.00
L0314	Stone Hut Stm	-43.707	170.665	1550	27	194.5	11	1.75
L0310	Royal Hut 3	-43.714	170.753	1270	27	161.5	10	0.91
L0322	Sealy Tarn	-43.715	170.074	1310	27	135.5	10	1.46
L0427	Mt Hope 2	-43.717	170.700	1730	30	128.5	8	1.49
L0707	Douglas Range 1	-43.720	169.552	1082	30	108.0	7	1.11
L0708	Douglas Range 2	-43.722	169.542	1054	30	120.5	10	1.63
L0309	Royal Hut 2	-43.725	170.749	1450	27	125.5	11	1.50
L0426	Mt Hope 1	-43.731	170.701	1775	30	154.0	10	1.79
L0308	Royal Hut 1	-43.731	170.755	1400	27	166.0	14	1.82
L0703	Sebastopol Tarns 1	-43.751	170.104	1203	29	125.5	7	1.40
L0428	Gammack Range 2	-43.810	170.293	2010	30	127.5	3	0.25
L0429	Gammack Range 1	-43.817	170.293	2020	30	164.5	4	0.57
L0702	Barnet Range 2	-43.823	170.194	1200	29	106.0	7	1.77
L0701	Barnet Range 1	-43.823	170.195	1216	29	118.0	7	1.75
L0505	Lake Dime	-43.826	169.279	1172	30	104.0	8	0.88
L0430	Mt Stevenson 1	-43.829	170.312	2070	30	146.0	6	0.43
L0506	Mataketake Range 1	-43.834	169.273	1143	30	108.5	9	1.43
L0507	Mataketake Range 2	-43.836	169.260	1148	30	138.5	14	1.86
L0326	Acland Lagoon	-43.837	170.111	590	27	117.0	10	1.95
L0325	Lagoon Basin 2	-43.846	170.074	1560	27	153.0	7	1.18
L0324	Lagoon Basin 1	-43.848	170.073	1560	27	131.0	11	1.56
L0508	Mataketake Range 3	-43.851	169.231	1210	30	115.5	11	1.59
L0709	Mataketake Range 4	-43.863	169.196	1051	30	113.0	8	1.47
L0425	Glenmore Tarn	-43.894	170.420	900	30	132.5	12	2.09
L0320	Glentanner 1	-43.895	170.097	1060	27	178.5	15	2.21
L0321	Glentanner 2	-43.900	170.099	1060	27	127.0	10	1.74
L0327	Lake Murray	-43.905	170.470	720	25	159.5	8	0.70
L0316	Irishman Stm 2	-43.995	170.041	1840	27	161.5	8	0.93
L0315	Irishman Stm 1	-43.996	170.051	2000	27	96.5	4	1.01
L0317	Irishman Stm 3	-43.999	170.037	1790	27	197.0	13	1.26
L0305	Duncan Stm 1	-44.013	170.043	1840	27	364.0	8	1.87
L0306	Duncan Stm 2	-44.019	170.052	1690	27	165.0	9	1.32
L0319	Boundary Stm 2	-44.034	170.121	830	27	510.0	16	1.83
L0307	Duncan Stm 3	-44.035	170.060	1340	27	177.5	15	1.77
L0318	Boundary Stm 1	-44.038	170.119	830	27	104.5	13	2.08
L0712	Martyr 3	-44.184	168.520	789	30	102.0	6	1.21
L0710	Martyr 1	-44.186	168.538	1164	30	142.0	10	1.52
L0711	Martyr 2	-44.189	168.532	1038	30	126.5	12	1.80
L1222	Ohau Skifield 2	-44.228	169.837	1504	26	156.5	17	1.43
L1221	Ohau Skifield 1	-44.228	169.837	1465	26	222.0	19	2.18
L0304	Raupo Lagoon	-44.320	169.903	586	27	112.0	15	2.26

*Excluded from analysis because of low head capsule count

Sediment samples from each lake were prepared according to standard procedures (Walker et al., 1991). Samples were deflocculated in hot 5% KOH for 20 minutes and then sieved over a 90-µm nylon mesh. Sample volume was selected with the aim to retrieve at least 100 whole head capsules per sample, starting with the top subsample and progressing to 1-2 cm as required. After sieving, head capsules were picked from the residue in a Bogorov tray under 50x magnification and mounted onto slides, which were fixed with either Pro-Texx® or "Norland Optical Adhesive 61". Chironomid taxa were then identified using a light microscope at 400x magnification and the identification guide compiled by Dieffenbacher-Krall and Vandergoes (2007). We did not distinguish between "short" or "long" antennal pedestals in *Tanytarsus funebris* types, because they have been found to display the same environmental preferences (Dieffenbacher-Krall et al., 2007), the pedestals rarely preserve, and it is difficult to distinguish between short pedestals and broken-off pedestals. Unidentifiable head capsules were removed from consideration.

4.3.1.2 Environmental variables

Average secchi depth (SEC), surface water temperature (SrT) and bottom water temperature (BT) were recorded at each sediment sampling point, except for NI01, 02, and 03. Water chemistry analyses included: pH, conductivity (CON), dissolved oxygen (DO), measured at the site using a TPS Con/pH/mV and TPS DO meter. Chlorophyll a (ChlA), total carbon (TC), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), total nitrogen (TN), total phosphorus (TP), chloride (Cl^-), sulphate (SO_4^{2-}), ammonia (NH_4^+), reactive nitrogen (NOx), reactive phosphorus (RP), and cations calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^+), and potassium (K^+) were measured in water samples taken at the sites. ChlA analysis was carried out at the National Institute of Water and Atmospheric Research (NIWA) Chemistry Laboratory, Hamilton, New Zealand. The other water chemistry analyses were carried out at Landcare Research, Environmental Chemistry Laboratory, Massey University, Palmerston North, New Zealand. In addition, the organic (ORG) and carbonate (CRB) content of the sediments was estimated by loss-on-ignition (LOI). Not all variables were measured for all sites (Table 4.2). Climate data (air temperatures and precipitation) were compiled for all 104 sites from the NIWA climate surface model derived from recording station data spanning the period from 1981 to 2010 (Table 4.2).

Table 4.2. Summary of environmental variables

Variable	Unit	Code	Mean	Median	Maximum	Minimum	# Lakes measured
Mean annual temperature	°C	MAT	7.4	7.1	16.4	2.5	103
Mean summer temperature	°C	SmT	11.7	11.5	19.5	6.9	103
Mean winter temperature	°C	WT	2.5	2.3	13.4	-3.8	103
Mean February temperature	°C	FT	12.3	12.1	20.4	7.7	103
Mean July temperature	°C	JT	1.8	1.7	13.1	-4.7	103
Mean annual precipitation	mm	MAP	3275	1770	9309	608	103
Altitude	m a.s.l.	ALT	1045	1060	2070	10	103
Lake area	m^2	ARE	2.01	1.25	6.60	0.14	11
Maximum depth	m	DEP	6.9	3.7	37.0	0.5	100
Secchi depth	m	SEC	3.5	2.4	19.3	0.5	100
Surface temperature	°C	SrT	15.0	15.9	23.8	4.0	100
Bottom temperature	°C	BT	13.2	13.5	23.8	4.6	96
pH		pH	6.84	7.09	9.50	3.80	96
Conductivity	ms/cm	CON	28.006	13.260	224.000	1.000	95
Dissolved oxygen	mg/L	DO	9.50	9.57	15.42	5.30	100
Chlorophyll a	mg/m^3	ChlA	1.1	0.6	5.9	0.1	88
Total carbon	mg/L	TC	4.4	2.8	19.5	0.8	61
Inorganic carbon	mg/L	DIC	1.6	0.7	16.0	0.2	61
Organic carbon	mg/L	DOC	3.9	2.6	38.7	0.2	89
Total nitrogen	mg/L	TN	0.203	0.170	1.690	0.040	69
Total phosphorus	mg/L	TP	0.0301	0.0139	0.2552	0.0000	98
Chloride	mg/L	Cl$^-$	2.2	0.8	42.4	0.4	90
Sulphate	mg/L	SO$_4^{2-}$	1.1	0.3	11.8	0.1	90
Ammonium	mg/L	NH$_4^+$	0.010	0.005	0.074	0.002	98
NOx - nitrogen	mg/L	NOx	0.012	0.008	0.150	0.001	98
Reactive phosphorus	mg/L	RP	0.008	0.006	0.209	0.002	98
Calcium	mg/L	Ca^{2+}	2.88	1.79	40.70	0.01	90
Magnesium	mg/L	Mg^{2+}	0.45	0.15	5.02	0.03	90
Sodium	mg/L	Na$^+$	1.98	0.77	28.10	0.15	90
Potassium	mg/L	K$^+$	0.35	0.10	15.62	0.01	90
LOI organic	%	ORG	21.332	16.005	85.996	2.113	82
LOI carbonate	%	CRB	2.071	1.693	14.865	0.754	82

4.3.2 Minimum required count sum of training set samples

To determine the minimum number of head capsules required to obtain a representative sample, we conducted several tests. First, we calculated taxon richness and Shannon's diversity (H; Shannon & Weaver, 1949) of simulated low count sums of training set samples using rarefaction methods (Birks & Line, 1992) and compared them to actual richness and H of the training-set sample in question. For each training-set lake we simulated 100 iterations each for count sums of 5 to 160 head capsules at intervals of 5 (i.e. count sum of 5, 10, 15 … 155, 160) by sampling without replacement. This range was chosen because we wanted to find the count sum from which changes in richness and diversity (and other measures described below) level off, and other studies showed most of the improvement happens between 5 and 50 head capsules. We did not extend our simulations beyond 160 head capsules, because very few (*n* = 19) training-set samples originally contain that many head capsules.

Second, we calculated correspondence analysis (CA) axis-1 values of simulated low-count samples of six selected lakes by passive ordination based on a CA of all training set lakes except the lake in question (see also Quinlan & Smol, 2001). These six lakes were selected based on a range of observed SmT and sample H and having an original count sum > 150: L0429 (low SmT, low H), L0305 (low SmT, high H), L0310 (intermediate SmT, low H), L1120 (intermediate SmT, high H), L0327 (high SmT, low H) and L1224 (high SmT, high H). We simulated count sums of 5 to 150 head capsules at intervals of 5 by sampling with replacement for each of the six lakes. We chose CA rather than linear methods because CA is sensitive to the abundances of rare taxa, which are likely to be misrepresented at low count sums.

Finally, we tested the difference in model diagnostics for models run including only training set samples with a relatively high count sum vs. those run including samples with a relatively low count sum ("high" and "low" based on the results of previous test). While samples with higher count sums may result in lower prediction errors, there is a trade-off with the number of samples included in the model. Not only can it be very time consuming to reach a high chironomid count sum, sometimes it is impossible to obtain minimum counts due to a low concentration of head capsules. This last test will show whether, for our training set, it is better to be conservative and remove samples with low count sums, or whether too much useful information is lost when you do so.

4.3.3 Data exploration

Before we can model the relationship between chironomid assemblages and SmT, we must explore which environmental or climate variables are ecologically important in the dataset. Although Dieffenbacher-Krall et al. (2007) found that SmT accounted for the greatest amount of variance in their training set limited to the Southern Alps, this relationship may have changed with a changed SmT gradient length. The extended SmT gradient may have made the relationship stronger, but there is also potential for introducing confounding variables (that may or may not covary with SmT) that were not significant in the original training set, due to the greater spatial distance between sites in the extended training set.

For subsequent analyses, only taxa that occur in at least 2 samples, with a maximum abundance of at least 1% were included (Table 4.3). Ordinations were performed to assess

the relationships between sites and taxa. Where necessary (i.e. in ordinations constrained by more than one variable), missing values were replaced with column means; in all other cases lakes with missing data were discarded. Both unimodal and linear techniques were applied: for CA and its canonical counterpart (CCA), raw count data were used; for principle component analysis (PCA) and its canonical counterpart (redundancy analysis, RDA) Hellinger-transformed data were used. Monte Carlo permutation tests (1000 permutations) were applied where appropriate to estimate whether variables explained a significant amount of variance ($P < 0.05$). Redundant variables were removed from the dataset; redundancy was determined based on a PCA of all variables, Spearman's correlation coefficients between variables, variable inflation factors (VIFs) and ecological interpretation. Constrained ordination (CCA and RDA) was used to identify which variables are important in the dataset and to evaluate how the different environmental variables relate to each other. Ordination with forward selection was used to identify which variables independently explain a significant amount of variance and was performed twice: with all lakes using only variables that were available for all lakes and with the subset of lakes for which all variables were available (See Appendix B). Partial CCAs and RDAs were used to estimate the independent explanatory power of each significant variable.

Table 4.3: Chironomid taxa found in the surface samples of the training set lakes. Abb. = abbreviations used in plots. The number of lakes each type occurs in, their maximum abundance in the training set (in % of the total chironomid sum), Hill's N2, and best-fitting HOF models for each chironomid taxon (with at least five occurrences) to SmT are listed. Taxa with a Hill's N2 < 5 are considered rare. Taxa are grouped by subfamily and tribe.

Abb.	Taxon	# of occur-rences	Maximum abundance (%)	Hill's N2	Preferred HOF model bimodal excluded	bimodal allowed
Diamesinae						
Ma	*Maoridiamesa*	6	11.6	2.32	Unimodal, skewed	Unimodal
Chironominae - tribe Chironomini						
Che	Chironomini early instar	68	24.2	30.29	Sigmoidal	Bimodal, skewed
Ch	*Chironomus*	101	94.5	49.23	Sigmoidal	Bimodal, skewed
Cl	*Cladopelma*	30	16.1	19.08	Sigmoidal, skewed	Bimodal, skewed
Cr[*1]	*Cryptochironomus*	1	3.3	N/A	N/A[*2]	N/A
Ha[*1]	*Harrisius*	1	0.7	N/A	N/A	N/A
K	*Kiefferulus*	6	7.4	3.08	Sigmoidal, skewed	Bimodal
La	*Lauterborniella*	3	2.4	2.57	N/A	N/A
Mt	*Microtendipes*	3	5.3	1.98	N/A	N/A
Pau	*Paucispinigera*	36	75.5	13.3	Sigmoidal, skewed	Sigmoidal, skewed
Pol	*Polypedilum*	46	33.8	24.45	Sigmoidal, skewed	Bimodal, skewed
Prc	*Parachironomus*	9	1.6	8	Sigmoidal	Sigmoidal
Ri[*1]	*Riethia/Pseudochironomini*	2	0.4	N/A	N/A	N/A
Xe	*Xenochironomus*	3	1.9	2.67	N/A	N/A

Table 4.3 - continued

Abb.	Taxon	# of occur-rences	Maximum abundance (%)	Hill's N2	Preferred HOF model bimodal excluded	Preferred HOF model bimodal allowed
Chironominae - tribe Tanytarsini						
Co	*Corynocera*	33	82.8	12.28	Unimodal, skewed	Bimodal, skewed
Pt	*Paratanytarsus*	7	3.5	2.94	N/A	N/A
St[*1]	*Stempellina*	1	1.3	N/A	N/A	N/A
TA	*Tanytarsus funebris* type A	74	68	29.2	Sigmoidal	Bimodal, skewed
TB	*Tanytarsus funebris* type B	20	9.3	9.8	Unimodal, skewed	Bimodal, skewed
TC	*Tanytarsus funebris* type C	85	73.6	40.61	Unimodal, skewed	Bimodal, skewed
Tvs	*Tanytarsus vespertinus*	23	20.8	7.89	Unimodal, skewed	Bimodal, skewed
Orthocladiinae						
Cn	*Corynoneura*	4	2.1	3.57	N/A	N/A
Ca	*Cricotopus aucklandensis*	46	63.7	16.48	Unimodal, skewed	Unimodal, skewed
Cp	*Cricotopus planus*	8	4.7	3.61	Sigmoidal	Sigmoidal
Cz	*Cricotopus zealandicus*	9	2.9	5.17	None[*3]	None
Eu	*Eukiefferiella*	9	4.1	6.74	Sigmoidal	Sigmoidal
Gy	*Gymnometriocnemus* type	20	5.7	14.71	Unimodal, skewed	Bimodal, skewed
H	*Hevelius*	4	1.2	3.86	N/A	N/A
Ka[*1]	*Kaniwhaniwhanus*	1	0.7	N/A	N/A	N/A
Li	*Limnophyes/Paralimnophyes*	39	19	12.67	Unimodal, skewed	Bimodal, skewed
Nf	*Naonella forsythi*	65	53.3	16.76	Sigmoidal	Bimodal, skewed
Nk	*Naonella kimihia*	69	43.2	28.42	Unimodal, skewed	Bimodal, skewed
N3	*Naonella* type 305	19	7.9	8.61	Sigmoidal	Sigmoidal
N4	*Naonella* type 419	16	57.1	4.88	Sigmoidal	Bimodal, skewed
Pk	*Parakiefferiella*	6	3.8	5.4	Unimodal	Unimodal
Sm[*1]	*Smittia*	3	0.7	N/A	N/A	N/A
OE	*Pirara*/Orthocladiinae E	8	20.2	2.59	N/A	N/A
SO[*1]	SO4 (*sensu* Cranston 2000)	1	2.6	N/A	N/A	N/A
O1b	Orthocladiinae 1b	12	5.9	6.23	Sigmoidal	Sigmoidal
O3	Orthocladiinae 3	12	3.5	6.82	Unimodal, skewed	None
O4[*1]	Orthocladiinae 4	1	0.7	N/A	N/A	N/A
O5[*1]	Orthocladiinae 5	3	0.8	N/A	N/A	N/A
O7	Orthocladiinae 7	2	1.5	2	N/A	N/A
O8	Orthocladiinae 8	7	2.8	6.53	None	None
O9	Orthocladiinae 9	2	2	1.8	N/A	N/A
O10[*1]	Orthocladiinae 10	1	2.2	N/A	N/A	N/A
O15[*1]	Orthocladiinae 15	1	0.8	N/A	N/A	N/A
O16	Orthocladiinae 16	4	2.8	3.2	N/A	N/A
O19	Orthocladiinae 19	3	1.9	3	N/A	N/A
Podonominae						
Pcl	*Parochlus*	31	14.5	13.91	Sigmoidal	Bimodal, skewed
Tanypodinae						
Ab	*Ablabesmyia*	29	19.7	13.14	Sigmoidal, skewed	Sigmoidal, skewed
P1	Pentaneurini type 1	2	1	1.71	N/A	N/A
Ap	*Apsectrotanypus* type	71	40.3	22.65	Sigmoidal	Bimodal, skewed
M1	Macropelopini 1	62	23.5	13.63	Unimodal	Unimodal
M2	Macropelopini 2	65	37.3	17.03	Sigmoidal	Bimodal, skewed

[*1] Excluded from analyses
[*2] HOF-model only determined for taxa present at ≥ 5 sites
[*3] "None" means no significant response, or HOF type I

Taxon response models were calculated for each variable that independently explained a significant portion of the variance. To calculate the best fit, Huisman-Olff-Fresco (HOF) models were chosen (Huisman et al., 1993): increasingly complex parametric response models are fitted to determine the most complex model that gives a statistically significant

improvement in fit. This analysis was performed twice: once restricting the analysis to unimodal skewed responses as the most complex (to aid in the decision to choose either linear or unimodal methods) and once allowing for bimodal response models (Jansen & Oksanen, 2013).

4.3.4 Model development

4.3.4.1 Model parameters

There are three settings to specify for MA modelling, i.e. which dissimilarity coefficient D to use (e.g. Euclidean, chi-square or chord), whether to calculate a simple mean or weighted mean, and the number of analogues (a) used to calculate the mean. The best configuration was determined by bootstrapping the training set and finding the lowest root-mean-square error of prediction (RMSEP; see section 4.3.4.2). For PLS and WA-PLS, the number of components C to include needs to be selected. A component was considered important if its addition resulted in > 5% reduction of the RMSEP after bootstrapping (Birks, 2003). Percentage abundance data was square-root transformed before fitting these two model types.

For three of the four ML techniques, bagging, RaF and RoF, the total number of trees nt needs to be selected. Model performance generally increases with ensemble size (as does processing time). Although performance often tends to stabilise from around 50 trees (e.g. Salonen et al., 2016), we chose $nt = 500$ to be conservative. For RaF, we additionally need to set parameter M, the subset of chironomid taxa that are tried at each node. The default value for M is $p/3$, where p is the number of chironomid taxa. We additionally tried $M = p/6$ and $M = p/1.5$ (see Liaw & Wiener, 2002), and picked the best (lowest RMSEP). For RoF, we need to set K, the number of random subsets of chironomid taxa on which PCA is applied for each tree. We tried a range of values for K from 1 to 35 and picked the best (lowest RMSEP).

BRT is more complex than the other ML techniques and requires several parameters to be set: learning rate (lr), tree complexity (tc), bagging fraction (bf) and the maximum number of trees (mt). To find the best combination of parameters for the BRT model, we followed the recommendations by Elith et al. (2008). The learning rate is used to shrink the contribution

of each tree as it is added to the model, where a lower learning rate generally increases the number of trees required. We tested four different learning rates, i.e. 0.01. 0.005, 0.001 and 0.0005. Slower learning rates are preferable to faster ones, as they shrink the contribution of each tree more and help the final model to reliably estimate the response. As a rule of thumb, models should be fitted with at least 1000 trees. Tree complexity, or the number of nodes per tree, also affects the number of trees required. Generally, as *tc* is increased *lr* should be decreased to fit enough (> 1000) trees. We tried *tc* values from 1 to 20 at intervals of 1. The bagging fraction determines the amount of stochasticity in the model by specifying the proportion of data to be selected at each step. The default bagging fraction is 0.5 (50% of data are drawn at random, without replacement, from the full training set); additionally we tried *bf* = 0.75. We set the maximum number of trees to 30,000.

4.3.4.2 Model comparisons

To compare the performance of all the different model types to each other we chose *k*-fold cross validation (CV), where the training set samples are randomly divided into *k* groups to evaluate the reliability of the proxy calibrations; we chose *k* = 10. The 10-fold CV is repeated 10 times to stabilise results. Assessment of performance is based on the RMSEP and the maximum bias of each model. The *k*-fold CV scheme is slightly more robust to spatial autocorrelation than commonly used leave-one-out CV, as pseudoreplicate samples have a 1/*k* chance of being removed from the training set data during CV (Trachsel & Telford, 2015; Salonen et al., 2016).

To compare the outcome of the models based on the extended training set to the original training set, we used Spearman's rank-order tests. We applied this test to the beta coefficients of chironomid taxa (an output from PLS and WA-PLS indicating the relationship of each taxon with the relevant environmental variable) and to taxon importance measures (an output of tree-ensemble models indicating which taxa contribute most information to the algorithm). Inconsistencies in the rank order of beta coefficients or taxon importance measures could indicate that the relationship between chironomid taxa and the environmental variable of interest was misrepresented or that certain taxa were badly sampled in the original training set.

4.3.5 Reconstruction from fossil samples

4.3.5.1 Minimum required count sum of fossil samples

Whereas the minimum required count sum for training set samples can be tested by cross-validation on the basis of their fit with the model, this is not possible for downcore reconstructions, as we do not know the "true" value of a given sample. Therefore, we instead use training set samples with simulated low count sums, following the methods of Heiri and Lotter (2001). For this analysis we used the same six lakes as for our previous test of the minimum count sum that results in stable CA axis 1 values (section 4.3.2; L0429, L0305, L0310, L1120, L0327 and L1224), using the same simulated count sums (5 to 150 head capsules at intervals of 5).

A chironomid-SmT inference model was calibrated separately for each of the six selected lakes, based on the extended training set without the lake in question, using the model type that was found to work best on the full, extended training set. SmT was inferred for all simulated samples, creating a data set of 100 temperatures for each simulated count sum for each of the six lakes. In addition, the SmT_{tot} and $RMSEP_{tot}$ for the original sample were calculated. Finally, we calculated the standard deviation (SD) of the 100 reconstructions per count sum, per lake, and the mean error (ME) of the reconstructed temperatures in relation to that of the original sample. The simulated count sum that results in an SD < 50% of the $RMSEP_{tot}$ and/or an ME that approaches 0, is considered to be the minimum required count sum for reliable fossil samples.

4.3.5.2 Lake Pupuke record

SmT was reconstructed from the Lake Pupuke fossil samples using all seven model types (see Appendix B). Lake Pupuke is a maar in the Auckland Volcanic Field (Fig. 4.1), where SmT presently lies around ~18.9°C. The analysed sediments cover the period from ~14 cal kyr BP to present, from which 68 subsamples were taken. For more information about the site and sediments, see Chapter 3.

Only the best-performing non-ML model and best-performing ML model are shown and discussed in the main text. Goodness-of-fit and dissimilarity of fossil samples to training set

samples were calculated along the age scale as well. A generalised additive model (GAM) is fitted through the reconstructed values to show the trends in the data. Results from the original training set were compared to results from the new, extended training set.

4.3.6 Software

All data analysis was performed in R version 3.5.1 (R Core Team, 2018). Data transformations were performed using the analogue package (Simpson & Oksanen, 2016; Simpson, 2007), ordinations were performed using the vegan package (Oksanen et al., 2018), HOF models were calculated using the eHOF package (Jansen & Oksanen, 2013), and GAMs were fitted using the mgcv package (Wood, 2011; 2017). All plots were made using packages ggplot2 (Wickham, 2016) and cowplot (Wilke, 2018).

Calibration models were implemented using the analogue package for the MA model (Simpson & Oksanen, 2016; Simpson, 2007), rioja for transfer functions (Juggins, 2017), rpart for bagging (Therneau & Atkinson, 2018), randomForest for RaF (Liaw & Wiener, 2002) and gbm for BRTs (Greenwell et al., 2018). RoF was implemented using the functions provided by Salonen et al. (2016), which also included the code for k-fold CV. We used the code from Elith et al. (2008) to find the best parameters for BRT modelling.

4.4: Results & discussion

4.4.1 Data exploration

Of 55 chironomid taxa recorded, 11 were excluded because they occur in < 2 samples and/or have a max. abundance < 1%: *Cryptochironomus*, *Harrisius*, *Riethia*/Pseudochironomini, *Stempellina*, *Kaniwhaniwhanus*, *Smittia* and Orthocladiinae types SO4 (*sensu* Cranston 2000), 4, 5, 10 and 15 (Table 4.3). SmT values cover a gradient from 6.9 to 19.5°C, mostly spread over altitude rather than latitude (Fig. 4.2a). Lakes with an SmT > 16°C are strongly underrepresented in the dataset (Fig. 4.2b).
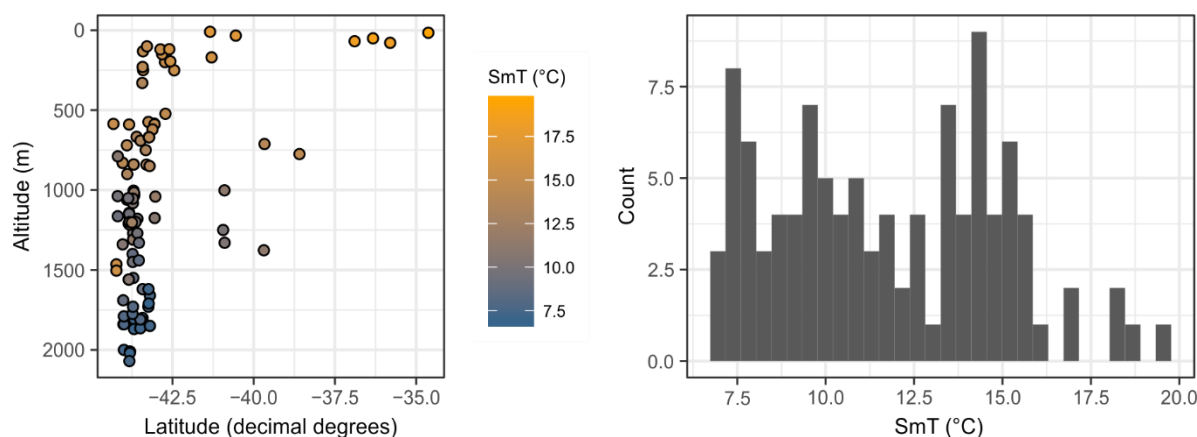
Figure 4.2: Summary of SmT data in the training set: temperature by latitude and altitude (a) and histogram of the spread of SmT values in the training set (b).

4.4.1.1 Minimum required count sum of training set samples

Rarefaction analysis shows that a simulated count sum of 130 head capsules, which is the median count sum of the 104 training set samples, captures 96.5% of the actual taxonomic richness on average (Fig. 4.3). For count sums of 100, 50 and 40 head capsules 93.1%, 76.5% and 71.2% of richness is captured on average, respectively. There are ten training set samples with a count sum of < 100, and one with a count sum of < 50. As the percentage of actual taxonomic richness captured at count sums lower than 100 steeply declines, it might be prudent to exclude those ten samples in the training set from analysis. On the other hand, the captured diversity appears to be less sensitive to low count sums than taxonomic richness (Fig. 4.3), with on average 94.4% of diversity captured at a simulated count sum of 40 head capsules. This could mean that the most important or abundant taxa are captured at representative abundances at low count sums, with only the rarest taxa missing. This is confirmed by the CA analyses on six selected lakes (Fig. 4.4). CA axis-1 scores show increasing stability with increasing sample size, but the range of scores decreases most strongly between count sums of 5 to ~30 to 40 head capsules. Moreover, the lakes with a count sum < 100 include several of the lakes at the higher end of the SmT gradient (i.e. L0205, L1101, L1102 and L1202), which is already underrepresented. We decided to include all training set lakes with a count sum ≥ 50 in subsequent analyses, i.e. removing only one lake (L0210, Lake Katrine) which has a count sum of 42 head capsules. In section 4.4.2 we compare the performance of the model based on all sites with a count sum ≥ 50 to the model based on sites with a count sum ≥ 100.
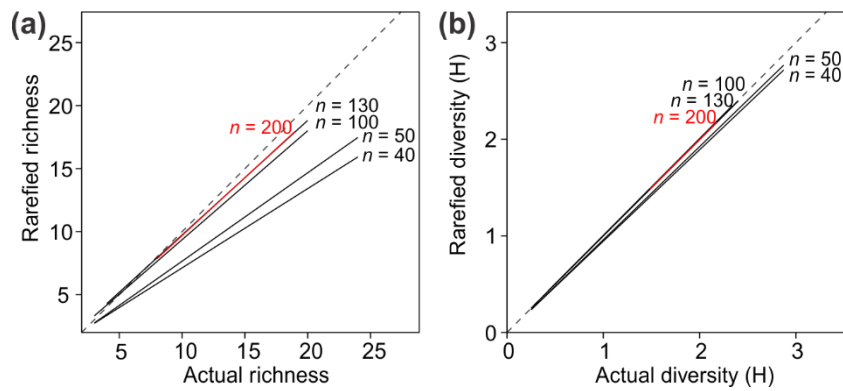
Figure 4.3: Richness (a) and diversity (b) of simulated samples with varying count sums. Lines represent linear models through simulated points based on a count sum (n) of 40, 50, 100, 130 (= median count sum of training set samples) and 200. Linear model through points with n = 200 in red to make it visible, as it overlaps with the linear model of n = 130.
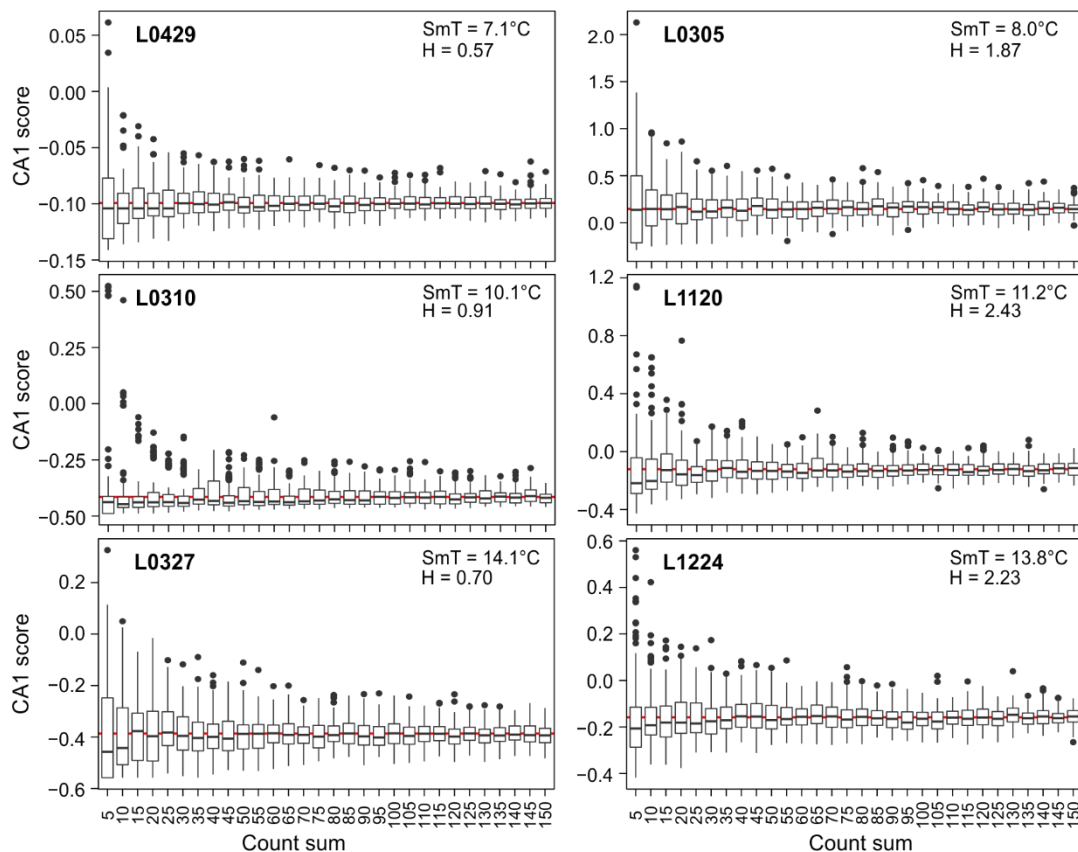
Figure 4.4: CA axis 1 scores of 100 simulated samples with varying count sums for six selected lakes. Samples passively ordinated in the 102-lake training set (excluding the passive sample and L0210 which has a count sum < 50), based on the abundances of 44 chironomid taxa. The box limits represent the first and third quartiles, horizontal line within the box is the median value, the whiskers encompass all values within 1.5 times the interquartile distance from the box and extreme values are indicated with points. Red line represents CA1 score of the original sample passively ordinated.

4.4.1.2 Ordination results

Detrended correspondence analysis on the chironomid counts showed a long first axis or gradient length, i.e. 4.43, which is substantially longer than the original training set (3.17; Dieffenbacher-Krall et al., 2007). Both linear and unimodal methods might be appropriate. The HOF models (see section 4.4.1.4) indicate that linear methods should be more appropriate, as most taxa show a sigmoidal response to various variables (including SmT; Table 4.3), although many taxa show a (skewed) unimodal response. In the following paragraphs we discuss both CCA and RDA in recognition of the varied taxon responses.

Separate ordinations of the 103 lakes based on taxon abundances in one run and environmental variables in another showed several lakes as outliers (see Appendix B, Fig. B-1 and B-2): L0310, L0327, L0409, L0418, L0419, L0421, L0422, L0601, L0705, L0901 and NI03, and L0327, L0901, L1201 and L1226 respectively. Lakes L0901 and L0327 were identified as outliers in both ordinations and should therefore be treated with special interest.

Ordinations constrained by one variable each showed that 23 variables have a significant impact on chironomid abundances when unimodal methods are used, or 25 when linear methods are used (Table 4.4). Of those, DIC, BT and Mg explained the largest percentage of inertia (> 6%) under CCA, whereas DIC, pH and SmT explained the largest percentage (> 8%) under RDA. The $\lambda 1$ to $\lambda 2$ ratio is < 1 for all variables, which indicates that there may be multiple (uncorrelated) determinants of chironomid assemblages in the training set.

Many of the variables show strong correlations with each other, evident from both the PCA on all variables (Appendix B, Fig. B-2) and their Spearman's correlation coefficients. We removed several variables to reduce redundancy in the dataset, guided by the results of the PCA and correlation coefficients, as well as ecological considerations. All temperature variables correlate strongly with each other. We opted to retain SmT only, because it explains the highest percentage of inertia (8.1%; Table 4.4) and because chironomid larvae have a demonstrated affinity with summer heat balance compared to e.g. FT or MAT (Livingstone and Lotter, 1998; Dieffenbacher-Krall et al., 2007; Armitage et al., 1995). Many cations correlate with each other, and with CON and DIC. Of these we opted to retain CON, as it is a measurement of dissolved ions, including the cations in question (ergo Ca^{2+}, Mg^{2+}, Na^+, Cl^-, K^+ and DIC were removed). Of DEP and SEC we chose to retain DEP because

SEC was often equal to or deeper than DEP, and thus unmeasurable. Of ChlA, ORG and TC, we chose to retain ChlA, because it was measured in more lakes than LOI and TC (i.e. 88 lakes vs. 82 and 61 respectively).

Table 4.4: Results of canonical ordinations (CCA and RDA) and test of significance on individual variables. "# Lakes" refers to the number of lakes in which each respective variable was measured or modelled in the case of climate variables.

	CCA				RDA				
	p Value	$\lambda 1$	% Inertia explained	$\lambda 1/\lambda 2$	p Value	$\lambda 1$	% Inertia explained	$\lambda 1/\lambda 2$	# Lakes
MAT	0.001	0.209	5.69	0.362	0.001	0.038	7.88	0.353	103
SmT	0.001	0.209	5.69	0.364	0.001	0.039	8.06	0.358	103
WT	0.001	0.199	5.42	0.342	0.001	0.036	7.41	0.334	103
FT	0.001	0.206	5.61	0.358	0.001	0.038	7.91	0.352	103
JT	0.001	0.194	5.29	0.334	0.001	0.035	7.26	0.328	103
MAP	0.002	0.100	2.72	0.168	0.001	0.027	5.60	0.274	103
ARE	0.106	0.309	15.13	0.589	0.078	0.068	15.49	0.628	11
DEP	0.019	0.079	2.19	0.133	0.001	0.021	4.35	0.212	100
SEC	0.016	0.111	3.09	0.194	0.001	0.019	4.02	0.198	100
SrT	0.001	0.199	5.53	0.370	0.001	0.030	6.23	0.293	100
BT	0.001	0.217	6.08	0.406	0.001	0.035	7.23	0.342	96
pH	0.001	0.145	4.01	0.239	0.001	0.039	8.07	0.425	96
CON	0.003	0.162	4.47	0.273	0.001	0.027	5.70	0.263	95
DO	0.077	0.057	1.59	0.098	0.008	0.014	2.86	0.135	100
ChlA	0.004	0.134	3.56	0.222	0.001	0.020	4.17	0.173	88
TC	0.002	0.208	5.87	0.302	0.001	0.038	7.85	0.372	61
DIC	0.014	0.235	6.63	0.341	0.001	0.043	8.94	0.460	61
DOC	0.012	0.165	4.46	0.274	0.002	0.018	3.70	0.157	89
TN	0.259	0.056	1.59	0.084	0.053	0.013	2.70	0.127	69
TP	0.801	0.022	0.61	0.037	0.249	0.006	1.22	0.053	98
Cl$^-$	0.118	0.072	1.95	0.119	0.037	0.011	2.21	0.094	90
SO$_4^{2-}$	0.008	0.130	3.50	0.216	0.001	0.030	6.17	0.291	90
NH$_4^+$	0.017	0.091	2.48	0.152	0.014	0.011	2.32	0.101	98
NOx	0.124	0.067	1.81	0.112	0.076	0.008	1.76	0.076	98
RP	0.939	0.008	0.21	0.013	0.832	0.003	0.62	0.027	98
Ca^{2+}	0.009	0.187	5.04	0.306	0.001	0.028	5.85	0.257	90
Mg^{2+}	0.002	0.222	6.00	0.365	0.001	0.028	5.83	0.249	90
Na$^+$	0.017	0.121	3.26	0.198	0.006	0.016	3.29	0.140	90
K$^+$	0.017	0.171	4.62	0.281	0.001	0.015	3.14	0.133	90
ORG	0.002	0.168	4.64	0.273	0.001	0.023	4.78	0.201	82
CRB	0.225	0.053	1.47	0.084	0.185	0.008	1.68	0.066	82

Constrained ordination showed that RDA had better explanatory power than CCA, with a higher percentage of inertia explained by the ten variables (27.7% vs. 20.2%; Table 4.5). Therefore, we focus on the RDA triplot here and in section 4.4.1.3. The first RDA axis positively correlates with MAP, and negatively correlates with pH, SO$_4^{2-}$, DEP and DO (Fig. 4.5b). This axis partially, but not fully, separates lakes located east (left) and west (right) of the Southern Alps ridge. The second RDA axis correlates strongly negatively with SmT and weaker with ChlA, DOC and NH$_4^+$. The CCA shows broadly the same variable relationships

as the RDA (Fig. 4.5a), although axis 1 and 2 appear flipped and chironomid taxa are more scattered, with none standing out.

Table 4.5: Summary of CCAs and RDAs of 103 lakes constrained by ten non-redundant variables and five significant variables (forward selected). Only the values of the first five axes are shown.

	CCA					RDA				
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
All non-redundant variables:										
$\lambda 1$	0.244	0.220	0.087	0.080	0.040	0.056	0.043	0.012	0.009	0.005
% inertia	6.7	6.0	2.4	2.2	1.1	11.5	8.9	2.4	1.8	1.0
cumulative:	6.7	12.7	15.0	17.2	18.3	11.5	20.4	22.8	24.6	25.6
% explained inertia	32.9	29.6	11.7	10.7	5.4	41.7	32.0	8.7	6.5	3.6
cumulative:	32.9	62.6	74.3	85.0	90.4	41.7	73.7	82.4	88.9	92.5
Canonical coefficients:										
DEP	-0.114	0.235	-0.259	-0.869	-0.447	-0.036	0.012	0.087	0.002	0.019
pH	0.131	0.248	0.365	-0.190	-0.078	-0.038	-0.003	-0.009	0.026	0.041
CON	-0.093	0.221	0.541	0.145	0.402	-0.009	0.014	-0.007	0.060	-0.016
DO	0.024	0.083	0.070	0.253	0.680	-0.009	0.005	-0.046	-0.012	0.008
ChlA	0.118	0.030	0.239	0.172	-0.409	-0.002	-0.015	-0.006	0.017	0.020
DOC	0.402	0.265	-0.995	0.276	-0.174	-0.015	-0.021	-0.009	-0.097	-0.030
SO_4^{2-}	0.127	0.512	-0.125	-0.078	0.276	-0.042	-0.009	0.010	-0.015	-0.100
NH_4^+	0.208	0.077	-0.508	-0.302	0.322	-0.009	-0.009	0.032	-0.057	0.020
SmT	0.794	-0.519	0.495	-0.159	0.202	0.021	-0.089	-0.011	0.036	0.012
MAP	0.085	-0.269	-0.029	-0.283	0.599	0.026	-0.006	0.031	0.033	-0.036
With forward selection:										
$\lambda 1$	0.234	0.204	0.078	0.047	0.020	0.053	0.041	0.008	0.006	0.004
% inertia	6.4	5.6	2.1	1.3	0.5	10.9	8.5	1.7	1.2	0.7
cumulative:	6.4	11.9	14.1	15.3	15.9	10.9	19.4	21.1	22.3	23.1
% explained inertia	40.1	35.0	13.5	8.0	3.4	47.3	37.0	7.4	5.1	3.2
cumulative:	40.1	75.1	88.6	96.6	100.0	47.3	84.3	91.7	96.8	100.0
Canonical coefficients:										
SmT	0.829	-0.482	0.529	-0.396	-0.061	0.012	0.094	-0.011	-0.061	-0.020
pH	0.083	0.424	0.395	-0.884	0.642	-0.054	-0.006	0.041	-0.091	0.027
DEP	-0.112	0.377	0.408	0.957	0.328	-0.045	-0.019	-0.081	0.035	0.042
SO_4^{2-}	0.170	0.635	-0.093	0.177	-1.022	-0.050	0.003	0.009	0.068	-0.080
DOC	0.449	0.400	-0.852	0.522	0.206	-0.022	0.020	0.039	0.093	0.058

Constrained ordination with forward selection showed that of the ten remaining variables, five accounted for a significant independent portion of inertia (Table 4.5; Fig. 4.5c, d), i.e. SmT, pH, DEP, SO_4^{2-} and DOC. RDA again had better explanatory power (23.1% of inertia explained) than CCA (15.9% explained; Table 4.5). Partial RDA show that SmT explains the largest amount of inertia when the effects of other variables are partialled out (i.e. 7.5%; Table 4.6).

The RDA triplots show a cluster of sites separate from the other training set lakes (Fig. 4.5b, d). We could not find any measured variable that correlates with these sites, but they are all located on the South Island, east of the Southern Alps, and appear to contain relatively high abundances of *Chironomus*.
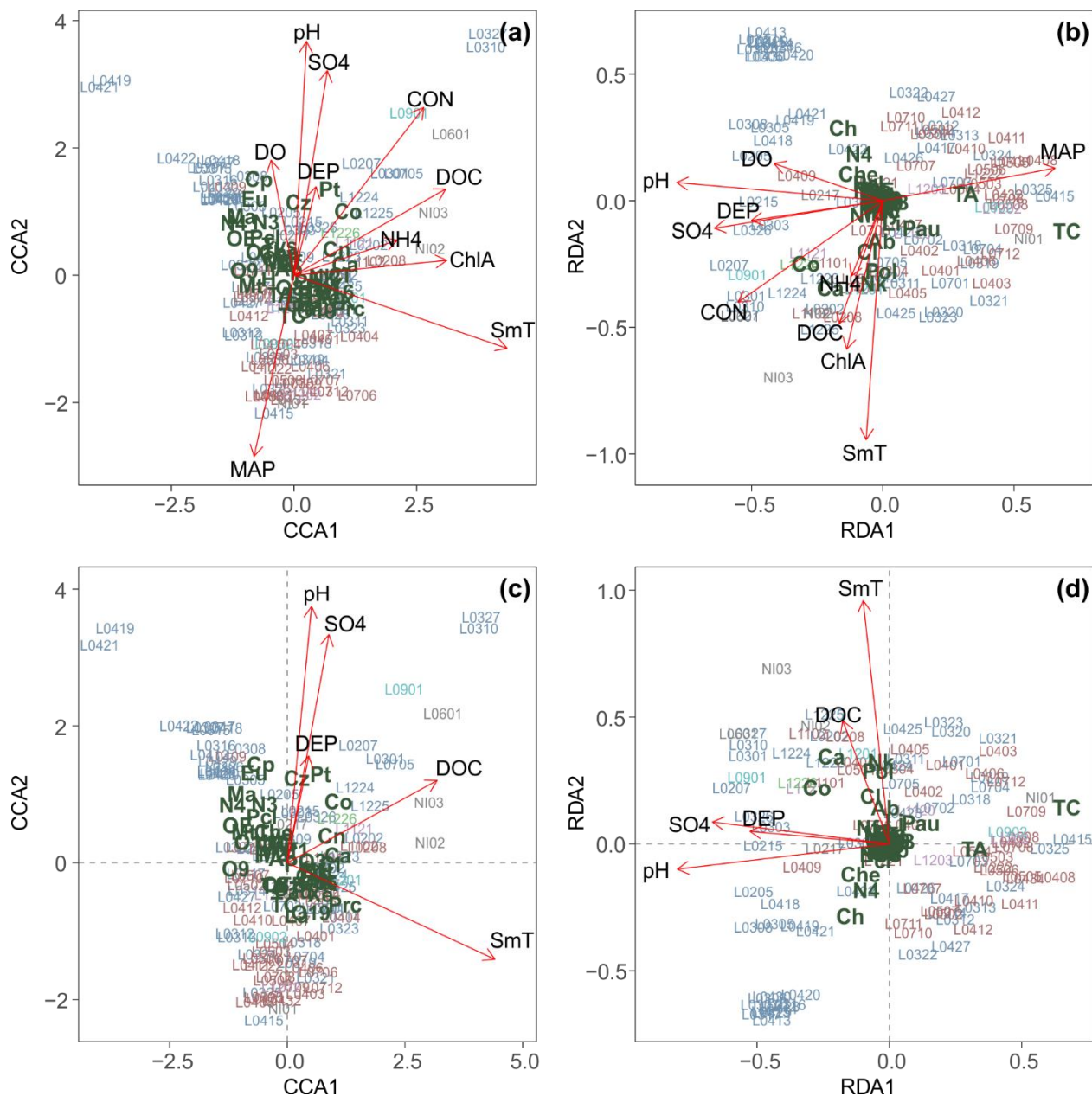
Figure 4.5: Triplots of the CCAs (a and c) and RDAs (b and d) of 103 lakes constrained by ten variables (a and b) or five, forward selected variables (c and d). Taxon scores in green. Variables plotted as vectors. Site scores indicated with coloured label: purple = northern North Island, green = eastern North Island, light blue = western North Island, pink = northern South Island, red = western South Island, blue = eastern South Island. Cluster of sites in top-left corner in (b) and bottom-left corner in (d) consists of: L0306, L0307, L0315, L0316, L0317, L0413, L0414, L0420, L0428, L0429 and L0430.

4.4.1.3 Ecological interpretation

The combination of ordinations and partial ordinations show that, although SmT explains the largest amount of variation on its own, other variables also play an important role. SmT is

correlated with the second RDA axis, whereas MAP and a combination of variables that are related to the lakes' chemistry are correlated with the first. Regarding chironomid taxa, *Tanytarsus funebris* types seem most affected by these variables. Near to nothing is known of the ecology of *T. funebris*; Forsyth (1978) found this species in four out of seven studied lakes near Rotorua (North Island), with the highest abundance (2.7%) in the eutrophic Lake Okaro.

Table 4.6: Results of partial canonical ordinations (CCA and RDA) of five variables separately and with the effects of remaining variables partialled out.

		CCA				RDA			
	Covariable(s)	$\lambda 1$	$\lambda 1/\lambda 2$	p value	% explained	$\lambda 1$	$\lambda 1/\lambda 2$	p value	% explained
SmT	None	0.209	0.364	0.001	5.7	0.039	0.358	0.001	8.1
	pH	0.209	0.615	0.001	5.9	0.039	0.882	0.001	8.8
	DEP	0.218	0.511	0.001	6.1	0.040	0.649	0.001	8.7
	SO_4^{2-}	0.210	0.906	0.001	5.9	0.038	1.157	0.001	8.4
	DOC	0.179	0.555	0.001	5.1	0.032	0.804	0.001	6.8
	All	0.174	0.959	0.001	5.4	0.030	1.235	0.001	7.5
pH	None	0.145	0.239	0.001	4.0	0.039	0.425	0.001	8.1
	SmT	0.146	0.687	0.001	4.3	0.039	1.254	0.001	8.9
	DEP	0.134	0.292	0.001	3.8	0.034	0.568	0.001	7.3
	SO_4^{2-}	0.091	0.318	0.007	2.6	0.025	0.637	0.001	5.5
	DOC	0.138	0.406	0.001	4.0	0.038	0.908	0.001	8.2
	All	0.062	0.411	0.036	2.0	0.016	0.712	0.001	4.2
DEP	None	0.079	0.133	0.023	2.2	0.021	0.212	0.001	4.4
	SmT	0.088	0.388	0.010	2.6	0.022	0.662	0.001	5.0
	pH	0.068	0.151	0.033	2.0	0.016	0.261	0.002	3.6
	SO_4^{2-}	0.069	0.236	0.036	2.0	0.017	0.455	0.001	3.8
	DOC	0.082	0.244	0.013	2.4	0.022	0.511	0.001	4.7
	All	0.067	0.390	0.026	2.2	0.016	0.656	0.001	4.0
SO_4^{2-}	None	0.130	0.216	0.010	3.5	0.030	0.291	0.001	6.2
	SmT	0.132	0.526	0.003	3.8	0.031	0.918	0.001	7.0
	pH	0.074	0.206	0.076	2.1	0.015	0.325	0.003	3.5
	DEP	0.123	0.265	0.009	3.4	0.026	0.417	0.001	5.7
	DOC	0.142	0.468	0.002	4.0	0.032	0.800	0.001	6.9
	All	0.086	0.457	0.020	2.7	0.015	0.597	0.001	3.8
DOC	None	0.165	0.274	0.019	4.5	0.018	0.157	0.002	3.7
	SmT	0.128	0.513	0.036	3.7	0.010	0.284	0.069	2.2
	pH	0.155	0.445	0.013	4.4	0.017	0.403	0.001	3.9
	DEP	0.162	0.413	0.015	4.5	0.018	0.262	0.001	3.8
	SO_4^{2-}	0.178	0.583	0.011	5.0	0.020	0.497	0.002	4.4
	All	0.109	0.583	0.042	3.4	0.007	0.308	0.120	1.9

The partial separation of eastern and western lakes over the first axis indicates that spatial location might be an important factor. DEP and SO_4^{2-} strongly correlate. This could be due to deeper lakes' higher likelihood to be stratified, resulting in anoxic conditions favourable to photosynthetic sulphur bacteria that oxidise sulphur to sulphate (Wetzel, 2001). The separation of deep, anoxic, alkaline lakes on the east vs. west of a mountain range in combination with a rainfall pattern resulting from orographic lift coinciding with this boundary,

brings to mind Tyler's Line in Tasmania (Tyler, 1992). Tyler's Line separates two limnological areas that reflect differences in climate, geology and vegetation in Tasmania, where western lakes are typically acidic and dystrophic and eastern lakes are oligotrophic. Rees and Cwynar (2010) observed a strong separation in their chironomid training set based on this spatial pattern and found that chironomid abundances were mainly related to pH. In our training set, pH might be driving the separation over the first axis too, with the other variables correlating due to spatial correlation.

In Rees and Cwynar's Tasmania training set, temperature correlates with pH. Fortunately, in our training set SmT appears to be independent of pH (vectors positioned orthogonally), though in fossil samples with high amounts of *T. funebris*, care must be taken while interpreting any trends. Of more concern is the correlation between SmT, ChlA, DOC and NH_4^+, which indicates that lake productivity may have a potential confounding effect on SmT. In subsequent analysis we must consider whether any perceived changes in SmT might actually be forced by changes in productivity. This is especially relevant for samples of the anthropogenic era, where eutrophication often plays a role. Although the relationship between SmT and lake productivity is strong and might in part explain chironomids' sensitivity to temperature (Eggermont & Heiri, 2012), we cannot be certain that this relationship has been linear in the past at any one location. Moreover, while productivity may increase with temperature, temperature does not change with productivity and it may be difficult to separate these two influences (Velle et al., 2010).

Little is known of the ecology of the chironomid taxa with the highest loadings in the RDAs (Fig. 4.5): *Tanytarus funebris* (mentioned above), *Cricotopus aucklandensis*, *Corynocera, Naonella kimihia*, *Chironomus* and *Polypedilum*. Virtually nothing is known of the New Zealand Orthocladiinae; *Cricotopus aucklandensis* and *Naonella* spp. have mainly been described in literature in morphological terms (Boothroyd, 1989; 1990; 2002; 2004). Both appear to be positively correlated with SmT and/or productivity in the training set (Fig. 4.5b, d). The *Corynocera* specimens found in New Zealand are morphologically identical to the Holarctic taxon *C. ambigua*, the distribution of which is thought to be unrelated to temperature. In the Northern Hemisphere, *C. ambigua* is associated with charophytes, although the presence of both is possibly related to shallow, clear-water lakes with stable sediments (Brodersen & Lindegaard, 1999). Forsyth (1971) described three species of *Chironomus*. Woodward and Shulmeister (2006) were able to distinguish two larval types,

one of which appears to be associated with high altitude, cool, oligotrophic lakes and the other with lowland, warm, eutrophic lakes; the latter is unlikely to occur in our training set, because eutrophic lakes were actively avoided during sampling. This is consistent with the negative correlation between *Chironomus* and SmT in the RDA triplots (Fig. 4.5b, d). There are at least six described species of *Polypedilum* in New Zealand (e.g. Forsyth, 1971), although they are only distinguishable by examination of adult males rather than larval head capsules. The genus is represented in a wide range of freshwater habitats (Winterbourn and Gregson, 1981). In our training set, *Polypedilum* appears to be associated with relatively high SmT.

4.4.1.4 HOF models

The predominant response among chironomid taxa to all of the five significant variables is sigmoidal if complex responses are excluded, but bimodal skewed if complex responses are allowed (Table 4.7, 4.3). However, there are also many chironomid taxa that show a unimodal or skewed unimodal response (if bimodal responses are excluded).

Table 4.7: Distribution of best-fitting HOF models for the species percentage data to each significant variable (only taxa that are present in at least 5 sites are included). HOF-types: I = no significant response, II = sigmoidal, III = sigmoidal with plateau, IV = unimodal, V = unimodal skewed, VI = bimodal symmetric, VII = bimodal skewed. For responses to SmT per taxon, see Table 4.3.

Variable	Allowed HOF-types	I	II	III	IV	V	VI	VII	Preferred
SmT	I-V	2	13	5	2	10			II (43.3%)
	I-VII	3	5	2	3	1	1	17	VII (58.6%)
pH	I-V	2	11	6	3	9			II (37.9%)
	I-VII	2	1	3	3	4	1	17	VII (58.6%)
DEP	I-V	3	10	6	3	10			II or V (34.5%)
	I-VII	1	2	1	4	2	0	22	VII (71.0%)
SO_4^{2-}	I-V	4	9	4	7	5			II (36.0%)
	I-VII	1	1	4	4	0	0	19	VII (67.9%)
DOC	I-V	2	12	3	6	6			II (44.4%)
	I-VII	1	5	3	6	3	0	11	VII (39.3%)

4.4.2 Model performance

4.4.2.1 Performance statistics and outliers

We calibrated 7 chironomid-SmT models for five different training set configurations. The configurations are: the original training set (Dieffenbacher-Krall et al., 2007) containing 60 sites and 36 (non-rare) taxa, the full, extended training set containing 103 sites and 44 taxa, the extended training set without the cluster identified in Fig. 4.5b and d ("Without cluster") containing 92 sites and 43 taxa, the extended training set restricted to sites with an SmT lower than 16°C ("SmT < 16°C") containing 97 sites and 44 taxa, and finally the extended training set including only sites with a count sum ≥ 100 ("Count sum ≥ 100") containing 94 sites and 38 taxa. Table 4.8 shows the parameterisation for each configuration that produces the best-performing models.

Table 4.8: Parameterisation of calibration models for five different configurations of the training set (TS), where model types are: MA = modern analogue technique, PLS = partial least squares, WA-PLS = weighted-averaging PLS, RaF = random forest, RoF = rotation forest and BRT = boosted regression tree. Parameters are: a = number of analogues, D = dissimilarity coefficient, C = number of components, M = number of chironomid taxa tried at each node, p = number of predictors or taxa, K = number of subsets the chironomid taxa are divided over, lr = learning rate, tc = tree complexity and bf = bagging fraction.

Training set configuration	# sites	# taxa	MA a	MA D	MA weigh[1]	PLS C	WA-PLS C	RaF M	RoF K	BRT lr	BRT tc	BRT bf
Original TS	60	36	2	chord	yes	2	1	$p/1.5$ (= 24)	4	0.001	11	0.75
Extended TS	103	44	10	chord	yes	2	2	$p/3$ (= 15)	4	0.0005	4	0.75
Without cluster	92	43	11	chord	yes	2	2	$p/6$ (= 7)	4	0.0005	13	0.75
SmT < 16°C	97	44	6	chord	yes	2	2	$p/3$ (= 15)	7	0.01	6	0.75
Count sum ≥100	94	38	5	chord	yes	2	2	$p/3$ (= 13)	6	0.0005	6	0.75

[1] Use weighted averaging

The performance statistics are similar between model types. For the extended training set, RMSEP lies between 1.78 and 2.13°C (Fig. 4.6). The MA technique performs worst for all training set configurations with high maximum biases, except for the original training set, and is therefore not considered a viable option for the extended training set. Poor performance of MA models applied to chironomid assemblages has been recognised by other authors as well (e.g. Rees & Cwynar, 2010), whereas MA often works well for pollen assemblages (e.g. Wilmshurst et al., 2007). Perhaps this is because different chironomid species do not interact much with each other, and as such do not form a "community" in the

way that plant species form vegetation communities. If this is the case, multivariate regression models (PLS and WA-PLS), which calculate an SmT response for each taxon individually, might be more appropriate.

WA-PLS produces the best performance statistics of the seven model types, with an RMSEP of 1.78°C and maximum bias of 3.42°C for the full training set configuration (Fig. 4.6). The best ML technique is RaF, with a RMSEP of 1.88°C and maximum bias of 5.14°C. Errors are particularly high at the high end of the SmT gradient, above ~16°C; this is the case for all model types, but is least pronounced for WA-PLS. The ML models do not appear to be able to reconstruct SmT above ~16°C at all. The reconstructed temperatures of all lakes with an SmT > 15°C are strongly underestimated. In the WA-PLS model, only lakes L0217 and NI01, the two northernmost lakes, appear as outliers on the warm end of the SmT gradient (residual temperature > 4°C).

At the other end of the spectrum, coolest temperatures tend to be overestimated in all seven model types. Of note are L0310 and L0416, which have high residuals in the parametric models and most ML models, and L0308 and L0419, which have residuals > 4°C in all four ML models (See Appendix B, Fig. B-9) for residual plots with labelled points). Of these, L310 and L0419 had outlying chironomid assemblages (high abundance of *Corynocera* or *Naonella* type 419 respectively) but were not distinctive in terms of lake chemistry or other measured environmental variables. Therefore, we could not justify removing them from the training set.

Removal of the 13-lake cluster as identified in Figure 4.5 does not significantly alter model performance over the SmT gradient (Fig. 4.6). If anything, the models perform worse than their full training set counterparts with higher RMSEPs and maximum biases. Similarly, the removal of training set lakes with a count sum of < 100 worsens performance statistics. Therefore, these two configurations of the training set are hereafter discarded. The higher RMSEPs of these two reduced training sets is most likely due to the reduction of information caused by removing lakes from the training set.
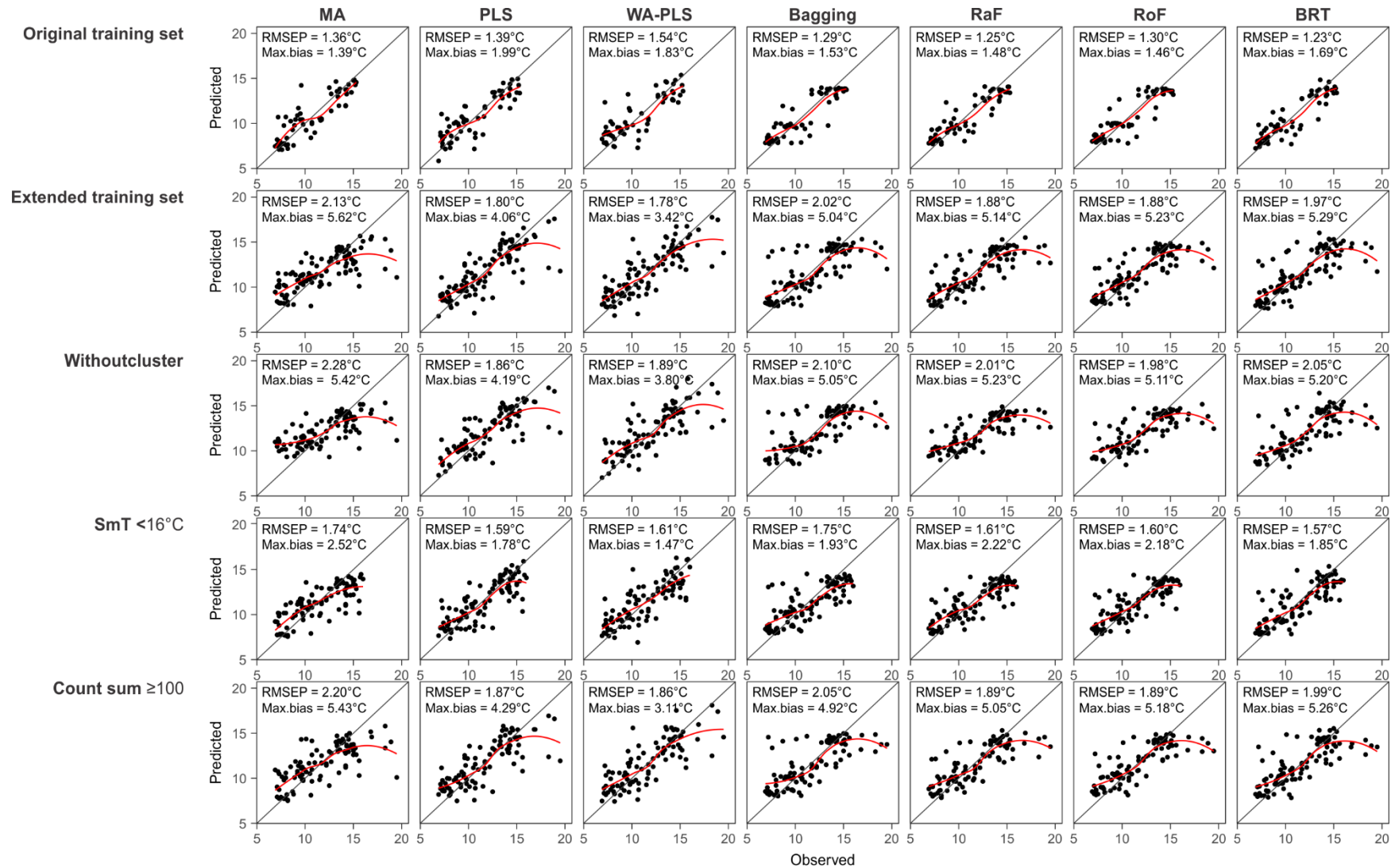
Figure 4.6: Model performance of seven model types in columns and five training set configurations in rows. Predicted SmT values vs. observed. Red line is the loess-smoothed trend through the points. Within each plot, the RMSEP and maximum bias are indicated, which are the means of performance statistics from 10 runs of 10-fold cross validation.

The original training set (Dieffenbacher-Krall et al., 2007) seems to perform better than the new, extended training set: RMSEP for WA-PLS on the original training set is 1.54°C. The beta coefficients and variable importance values are consistent between the original and extended training set (ρ(36) = 0.86 to 0.91, P = 0.000***; Fig. 4.7), indicating there are no major overhauls in the relationship between each taxon and SmT. However, the SmT gradient of the original training set is limited and has been shown to be insufficient to reconstruct Holocene temperatures in northern New Zealand (Chapter 3). The longer SmT gradient of the extended training set is likely to result in a better estimation of each taxon's SmT optimum, especially at the ends of the gradient. The higher RMSEP and especially the high maximum bias of the extended training set is exaggerated by the lakes towards the high end of the training set. Moreover, the better performance statistics of the original training set might indicate higher precision, but not necessarily higher accuracy.

As an additional test to compare the old and new training set, we fitted the seven models to a reduced training set, only including the lakes of the extended training set with an SmT < 16°C. For this reduced training set, BRT produces the best performance statistics, with an RMSEP of 1.57°C.

4.4.2.2 Transfer functions vs. tree-ensemble models

The poor performance of the ML algorithms compared to the transfer functions is surprising, as so many of the chironomid taxa appear to have a bimodal response to SmT (Table 4.7). ML algorithms are supposed to be able to model bimodal responses, as well as deal with a variety of different responses, compared to traditional transfer function approaches (Elith et al., 2008). This mismatch between expectation and outcome regarding tree-ensemble modelling was also found by Juggins et al. (2015), who compared parametric methods to BRT. In their datasets, the dominant taxa that are most influential in the BRT model are taxa with complex multivariate responses to the variable of interest and one or more confounding variables. Traditional parametric methods implement information from many taxa as opposed to several selected, dominant taxa, possibly modulating the reconstructed values when dominant taxa show complex responses. In our case, the two most important taxa in the best-performing ML model (RaF), *Cricotopus aucklandensis* and *Naonella kimihia* (Fig. 4.7c), show a unimodal response (Table 4.3). Therefore, WA-PLS may be just as good at modelling SmT based on the abundances of these two dominant taxa as RaF. Interestingly,

C. aucklandensis, *N. kimihia* and the next three most important taxa in the RaF model (*Polypedilum*, *Cladopelma* and *Ablabesmyia*) all have high beta coefficients (Fig. 4.7a, b) and are thus associated with high SmT. Due to the lack of sites at the high end of the SmT gradient, the relationship between SmT and these five taxa may not have been captured very well in our training set. Parametric methods are thought to be better at extrapolation beyond calibration data than ML techniques (Birks et al., 2010; Salonen et al., 2016). ML techniques are thought to be more sensitive to bias towards dataset mean in general, as towards gradient ends trees become increasingly likely to predict using samples located towards the centre of the gradient.
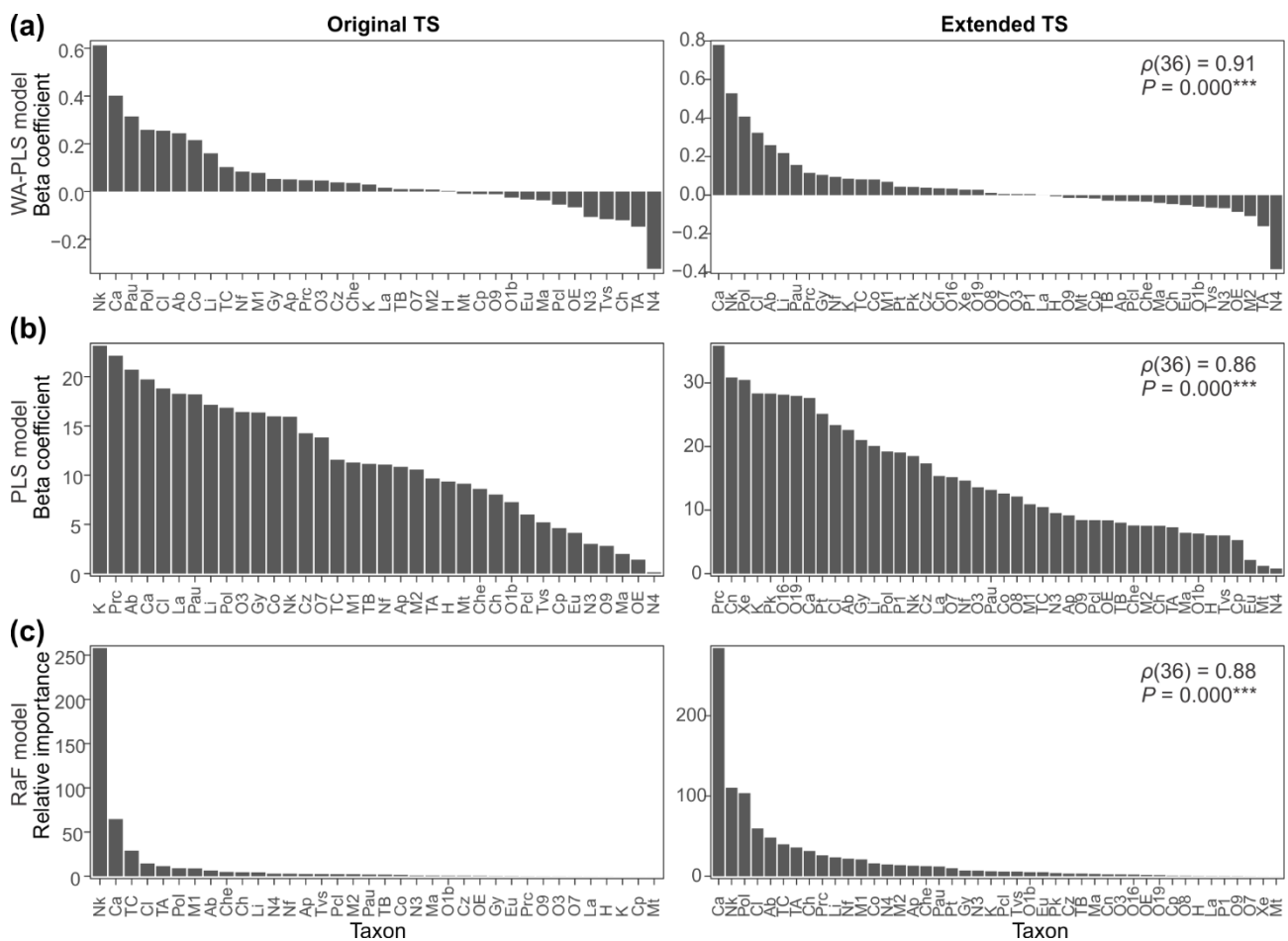
Figure 4.7: Beta coefficients of taxa ordered from highest to lowest for PLS model (a) and WA-PLS model (b), and relative importance of taxa in the RaF tree-ensemble model (c). Original training set on left, extended training set on right. Results of Spearman's rank-order test indicated in right-hand corner.

Related to this bias-towards-the-mean problem, ML algorithms are sensitive to small samples. ML algorithms may be at a disadvantage compared to transfer functions in relatively small datasets, as each prediction is based on a small subset of samples residing

in a terminal tree node (Salonen et al., 2016). If the SmT gradient has not been sampled densely and evenly, there might not be enough information at terminal nodes for accurate prediction across the gradient (Hjort & Marmion, 2008; Wisz et al., 2008). As shown in Figure 4.2, the SmT gradient is not sampled evenly in our training set, with sites > 16°C strongly underrepresented and another gap in the distribution around ~13°C.

Finally, the performance of the transfer function models may be overestimated relative to the ML techniques. It is possible that performance is not captured accurately with cross validation, especially in the presence of secondary gradients such as productivity. These confounding factors may only affect downcore reconstructions and not the training set lake reconstructions on which cross validation is based (Juggins, 2013; Juggins et al. 2015). It is not implausible that secondary gradients might play a role in the extended training set, as $\lambda1/\lambda2$ was less than 1 for SmT in constrained ordination (Table 4.6), and three variables related to productivity (ChlA, DOC and NH_4^+) correlate with SmT. ML techniques might perform better in the presence of confounding factors, but this is difficult to test.

Concluding, WA-PLS may be the best model choice for reconstructing SmT from chironomid records, especially when analysing North Island lakes where temperatures above 16°C may be expected. However, for more southern or alpine lakes, or when reconstructing temperatures for glacial periods, other options might be considered. BRT seems to perform better than WA-PLS when lakes at the high end of the SmT gradient are removed from the training set. We recommend examining your chironomid assemblages in detail before choosing the most appropriate model type. Which taxa are most abundant in your dataset? Which taxa show the most dramatic changes in abundance? What sort of relationship do these taxa have with SmT: are the response curves primarily sigmoidal, unimodal or bimodal (Table 4.3, 4.7), or a mix? Have these taxa been shown to respond to other environmental variables, such as pH? Consider using ML techniques when taxa with a bimodal response are abundant in your dataset, abundant taxa display a variety of responses, or when you suspect confounding variables may play a role. The best approach may be to apply all model types and critically examine the differences in reconstructed values.

4.4.3 Reconstruction from subfossil chironomid remains

4.4.3.1 Minimum count sum required for downcore samples

The range of reconstructed SmTs of simulated samples for the six selected lakes stabilises at a count sum of ~40-60 head capsules (Fig. 4.8). At a simulated count sum of 55, reconstructed SmT of all lakes shows a SD < 50% the value of the $RMSEP_{tot}$ for each respective lake (Fig. 4.9a). The ME of reconstructions does not approach 0 for all six lakes even at a count sum of 150 head capsules (Fig. 4.9b). However, ME changes most between 5 and ~60 head capsules. Therefore, the minimum required count sum of fossil samples is 60 head capsules.
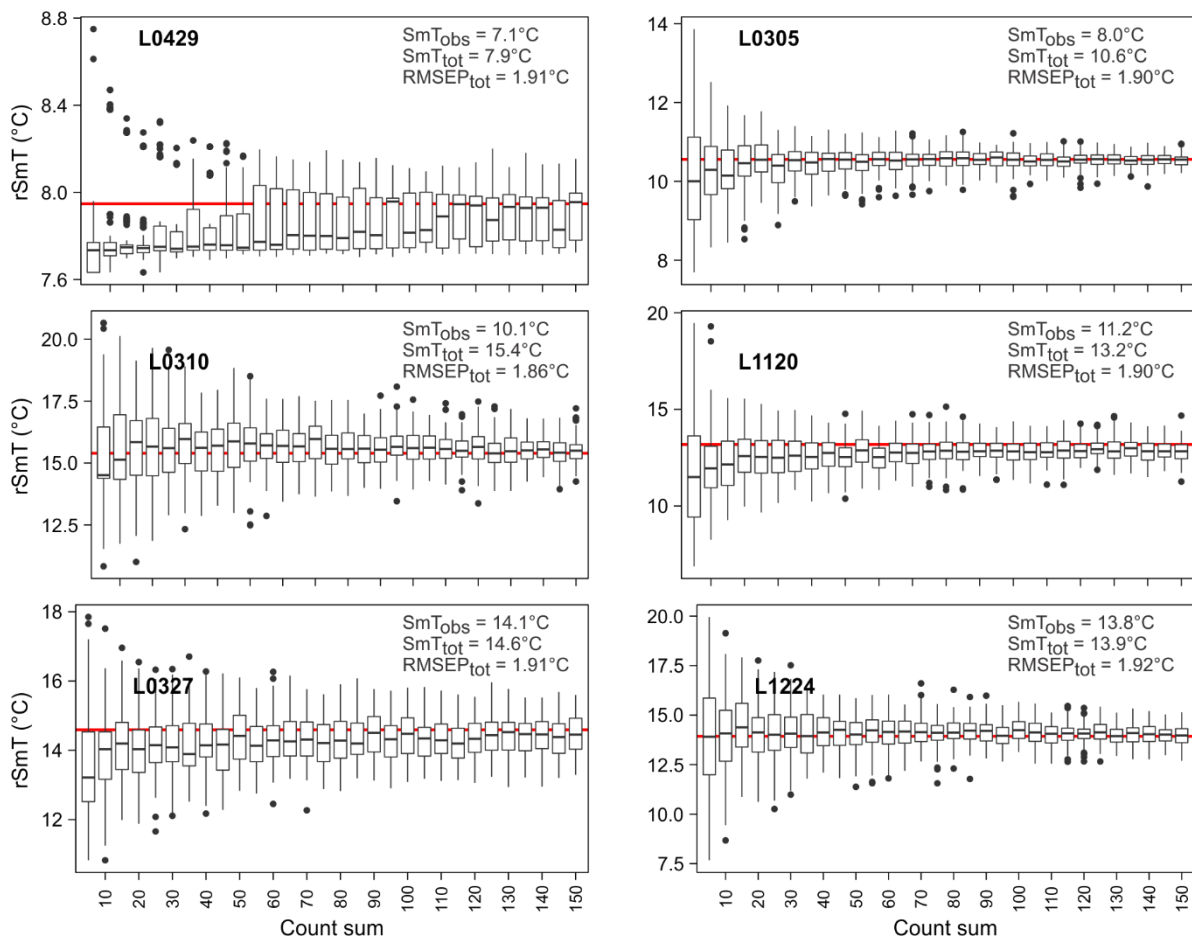
Figure 4.8: Inferred temperatures for different simulated count sums for six selected lakes; each box plot represents 100 iterations for a given count sum. The box limits represent the first and third quartiles, horizontal line within the box is the median value, the whiskers encompass all values within 1.5 times the interquartile distance from the box and extreme values are indicated with points. The horizontal red line represents the reconstructed temperature SmT_{tot} of the original sample.
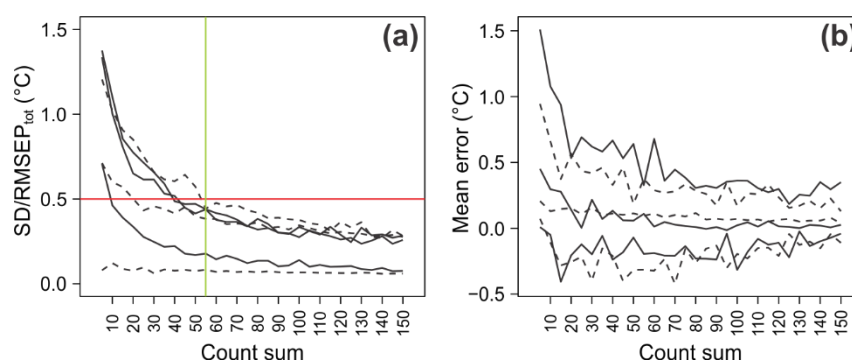
Figure 4.9: Standard deviation (SD) of reconstructed SmT for 100 iterations of each simulated count sum in relation to the RMSEP of the original sample (RMSEP$_{tot}$) (a) and mean error of inferred temperatures (b). Each line represents one of the six selected lakes, where dashed lines represent lakes with a low taxon diversity. Red, horizontal line in (a) represents the value where SD is 50% of RMSEP$_{tot}$; green, vertical line represents count sum for which SD is < 50% of RMSEP$_{tot}$ for all six lakes.

4.4.3.2 Lake Pupuke SmT reconstruction

Very few of the Lake Pupuke fossil chironomid assemblages have close modern analogues in the training set (Fig. 4.10a), indicating that the MA technique is probably not appropriate for this dataset. Of the 68 Lake Pupuke samples, 29% show a very poor fit to SmT (Fig. 4.10b). Interestingly, this number is higher than for the original training set, where 19% of Pupuke samples show a very poor fit (Chapter 3). The poor fit may be caused by the abundance of rare taxa (Hill's N2 < 5; Table 4.3), or taxa that are excluded from the models (Fig. 4.10c), such as *Harrisius*. The presence of rare taxon *Xenochironomus* in the Pupuke samples appears to overlap with poorly fitted samples. *Xenochironomus* is only present in three training set lakes at the higher end of the SmT gradient, and consequently has a low importance in the RaF model and high beta coefficient in the WA-PLS model (Fig. 4.7). The only described *Xenochironomus* species in New Zealand is *X. canterburyensis* (Freeman), which is a commensal of the freshwater mussel *Hyridella menziesi* (Gray) (Forsyth & McCallum, 1978a, b; Fosyth, 1979). The presence of *Xenochironomus* is therefore likely to be more related to the presence of its host than to SmT.

The most abundant taxa are *Corynocera*, *Ablabesmyia* and *Cricotopus aucklandensis* (Chapter 3), which display a bimodal skewed, sigmoidal skewed and unimodal skewed response, respectively. Despite these dissimilar responses, we think WA-PLS is the most

appropriate model type for the Lake Pupuke dataset, because modern day SmT lies around 18.9°C and ML models do not appear to be able to reconstruct temperatures above 16°C (Fig. 4.6, Fig. 4.11c). As modern day SmT at Lake Pupuke lies around ~18.9°C it is highly unlikely that temperatures throughout the Holocene were below 16°C.
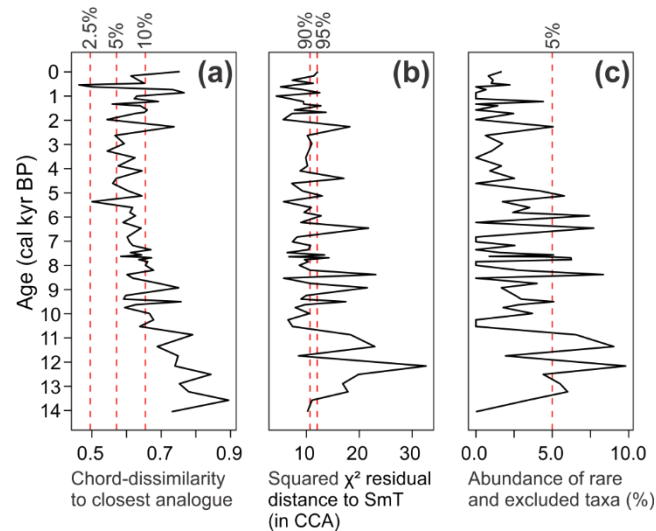
Figure 4.10: Reconstruction diagnostics for Lake Pupuke fossil chironomid samples. Distance of sample to closest modern analogue in the training set (a), fit of fossil samples to SmT ordinated passively in a CCA based on 103 training set lakes and 44 chironomid taxa (b), and the abundance of rare taxa and taxa excluded from modelling in the fossil assemblages (c). Red, dashed lines indicate: 2.5th, 5th and 10th percentiles of dissimilarities within the training set (a), 90th and 95th percentiles of residual distances of training set samples to SmT canonical axis (b) and the boundary above which the abundance of rare and excluded taxa is considered high (= 5%) (c).

Regarding confounding variables, first, *Tanytarsus funebris* taxa are present in reasonable (> 10%) abundances in the Lake Pupuke record. *T. funebris* type C abundances vary throughout, with the lowest values between ~6.5 and 1.0 cal kyr BP, which overlaps with the highest reconstructed SmT values. We have no reason to believe that the lake became more alkaline during this interval, but we can also not rule it out. Lake depth (negatively associated with *T. funebris*, Fig. 4.5) is unlikely to have increased at this time, as we have independent evidence for decreased effective precipitation after ~7 cal kyr BP (Chapter 3). Second, we have some evidence for fluctuating lake productivity (i.e. changing *Botryococcus*

abundance) throughout the records, but the trend in *Botryococcus* abundances does not correspond to the reconstructed SmT trend.
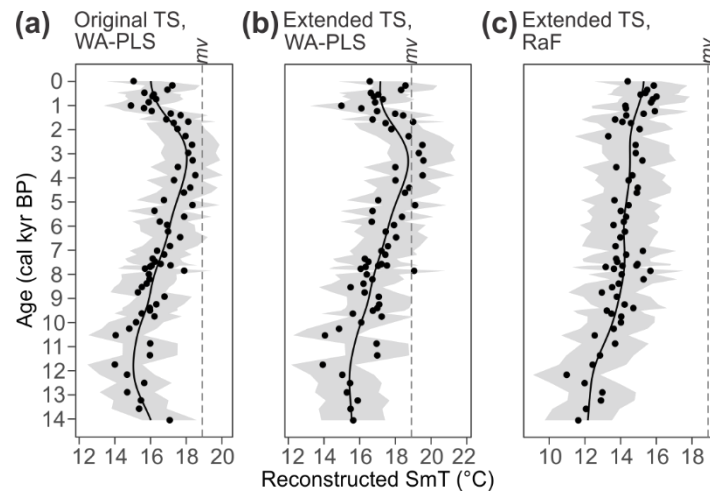
Figure 4.11: SmT reconstructions for Lake Pupuke, using the WA-PLS transfer function of the original training set (Dieffenbacher-Krall et al., 2007) and the new, extended training set, and the best-performing ML algorithm for the extended training set (RaF). Points are reconstructed values, shaded area represents RMSEP and curve is the result of a GAM fitted through the points. The dashed, vertical line represents the modern SmT value (mv) at the site (18.9°C).

In comparison to the original training set, the extended training set is better able to predict temperatures close to modern day values at Lake Pupuke (Fig. 4.11). Nonetheless, the SmT trend is very similar between the original and extended training set. This implies that our original interpretation of SmT (and apparent relationship with local summer insolation intensity) still stands. Nonetheless, the extended training set gives us more confidence in the reconstructed trend and allows us to speculate about absolute temperatures.

4.5: Conclusion

We have analysed a 104-lake chironomid training set from New Zealand, which is an extension of the training set published by Dieffenbacher-Krall et al. (2007). We determined that for our dataset, the minimum required count sum of training set samples is 50 head capsules; for downcore samples the minimum required count sum is 60 head capsules. We confirm that SmT is the most important variable to explain variance in chironomid

assemblages, but we find potentially confounding factors, some of which do not correlate with SmT (i.e. pH, lake depth) and some of which do (i.e. chlorophyll-*a*, DOC and NH_4^+). We find that the WA-PLS technique performs best and displays the lowest error when reconstructing temperatures above 16°C. This is likely due to the bias-towards-the-mean problem ML algorithms demonstrate, exaggerated by the underrepresentation of training set lakes at the high end of the SmT gradient. However, if reconstructing cooler temperatures that are unlikely to exceed 16°C, ML models such as BRT might be more appropriate, especially in the presence of confounding variables. We advise any analyst to consider their dataset and the taxa that are most abundant or display the most pronounced changes. Interpretation of reconstructed SmT should always incorporate both model statistics and underlying chironomid ecology.

We recalculated Lake Pupuke SmTs based on the WA-PLS model calibrated from the extended training set. The new reconstruction shows the same trend as the reconstruction based on the original training set, thus our original conclusions still stand (see Chapter 3). However, the new model is better able to reconstruct temperatures close to modern day SmT at Lake Pupuke, giving us more confidence in the absolute values.

Further improvements can be made to the chironomid-SmT inference model as follows:

1. Investigating the relationship between chironomids and their environment. We need a more comprehensive understanding of all environmental variables that affect chironomids, especially focussing on pH, lake depth and productivity. If we have a better grasp on the effect of these confounding variables, we will be better able to recognise these effects in model reconstructions.

2. Identification of taxa to a lower taxonomic level. If we are able to distinguish between taxa that are now lumped into one, we may be able to eliminate bimodal response models and increase our understanding of the relationship between SmT and chironomids. Progress in the environmental DNA space may facilitate this objective.

3. Increase the size of the training set and fill in the gaps. Our results have shown that there is a large bias-towards-the-mean problem in our ML models. Extending the training set, especially at the high end of the SmT gradient, will reduce this problem and ML techniques might become more appropriate.

Chapter 5

A terrestrial palynological record for the last ~85 kyr
from Orakei Basin, northern New Zealand

Preface

Chapter 5 is written in "publication-style", although it has not yet been submitted to a journal. Several people have contributed data/information for this chapter in the context of the Auckland Maars Marsden project (see Chapter 1), i.e. the lithological description (Leonie Peti) and preliminary chronological constraints (Jenni Hopkins), which will be acknowledged in the relevant sections. The age-depth model for the Orakei Basin sediment core has not been finalised yet. Therefore, the pollen record will be presented on a depth scale in this chapter instead of on an age scale, with chronological tie points provided by preliminary tephra identifications. Our intention is to publish the work presented in this chapter as a research paper once the finalised age-depth model has been completed by the wider Auckland Maars Marsden project.

Chapter 5 relates to objectives 1b, 2b and 3b. It was not possible to complete objective 1b (to produce pollen and chironomid records from the Orakei sediments), because the concentration of chironomid head capsules was too low to produce a record within the timeframe of this thesis (see Appendix E). For the same reason, objective 2b (to reconstruct MAAT and SmT from the pollen and chironomid results respectively) could not be completed with regards to SmT. Objective 3b is to compare the Orakei records to other records from the region, from southern New Zealand and to global climate reconstructions, with special interest for potential analogues of future climate. It was initially expected that the Orakei Basin core would reach to the Last Interglacial period (~129-116 kyr BP), one of the potential analogues, but as will be explained in this chapter, our results indicate that this period was probably not recorded. However, the Orakei Basin pollen record adds much value to the existing body of research, as it can tie together the scattered, fragmentary records of the last ~85 kyr from northern New Zealand and is the first continuous terrestrial record of vegetation and MAAT from the North Island.

Abstract

Orakei Basin is a ~126-kyr old maar in the Auckland Volcanic Field that has accumulated lacustrine sediments until its crater rim was breached by postglacial marine transgression ~9 kyr BP. The 85 meters of lacustrine sediments therefore are expected to cover the Last Interglacial-Glacial cycle to the early Holocene. The five zones identified in the pollen record of Orakei Basin (resolution 1 sample per 0.5 m) are broadly consistent with Marine Isotope Stages (MIS) 5 to 1. This interpretation is supported by both the mean annual air temperature (MAAT) reconstruction that was developed from the pollen assemblages using a transfer function, and the preliminary tephrostratigraphy. It is argued, however, that the pollen record does not cover the entire MIS 5 and commences in substage MIS 5a or 5c. MIS 4 appears to be a brief cold period at Orakei Basin, but not as cold as MIS 2, consistent with other pollen records from New Zealand. MIS 3 shows an earlier warm phase, followed by a progressive cooling trend concurrent with the expansion of beech taxa above the Rotoehu tephra (~45.2 cal kyr BP) culminating in MIS 2. An early onset of glacial conditions is consistent with previously reported records for the Last Glacial coldest period (LGCP). Similarly, the pollen-vegetation signature of the Last Glacial/Interglacial transition is typical for northern New Zealand, with an early dominance of *Prumnopitys taxifolia*, which is replaced by *Dacrydium cupressinum*, followed by *Metrosideros* and *Ascarina lucida* towards the top of the record. Forest taxa persisted throughout the Orakei Basin record, although grassland indicators increase during the LGCP and to a lesser extent during MIS 4, confirming the hypothesis that Auckland lies on or near the boundary of continuous forest cover at that time. The broad similarity of the Orakei Basin pollen record to the MIS subdivisions and the Okarito pollen record in south Westland imply a common driver: Northern Hemisphere high-latitude summer insolation. However, climate changes superimposed on the global trend were likely forced by local to regional summer insolation intensity.

5.1: Introduction

Continuous terrestrial records that reach beyond the Last Glacial are uncommon wordwide. In regions prone to glaciation, older depositional records are often overprinted by the most recent glacial advances. In other regions, records tend to be discontinuous due to strong climate variability with increased likelihood of disturbance from in-wash and erosion (e.g. Banks Peninsula record, Shulmeister et al., 1999; Lake Poukawa, Shulmeister et al., 2001) or pauses in deposition during dry periods (e.g. Lake Omapere; Newnham et al., 2004) (Fig. 5.1a). The terrestrial southern mid-latitudes are especially underrepresented in paleoclimate research and global compilations (e.g. Marcott et al., 2013), because of the relative lack of landmasses compared to the Northern Hemisphere, and lower levels of research activity in these regions.

Here we present a new palynological record and climate reconstruction from Orakei Basin, a maar in the Auckland Volcanic Field (AVF), northern New Zealand (Fig. 5.1), that covers much of the Last Interglacial-Glacial cycle. The vegetation reconstruction, along with reconstructed mean annual air temperatures (MAAT) and the pollen moisture index (PMI), offers new insight into regional differences in climate change and possible driving mechanisms. The Orakei pollen record provides a climato-stratigraphic framework that may be used to tie together the numerous fragmentary records from the northern North Island of New Zealand and it also enables comparable Holocene records from the region to be set in a longer time context. As the AVF lies approximately on the boundary of continuous forest cover during the last glacial maximum (LGM) (Newnham et al., 2013) the Orakei pollen record should be sensitive to broadscale vegetation shifts throughout the Last Interglacial-Glacial cycle.

In addition to the wider contributions discussed above, the Orakei Basin record can also serve to address several important research questions. First, is there merit in the model proposed by Newnham (1999) for northern New Zealand, that kauri and beech forest are indicative of warm and cool climate phases, respectively? Previous research has shown that conifer-angiosperm forest was present in Northland (Fig. 5.1a) throughout the Last Glacial cycle, but that kauri and beech alternated in importance, possibly in relation to temperature. However, temperature is not the only control on the presence of these taxa, with precipitation and local site factors thought to play a role as well. Recent research showing spatial variation

in the dominant phases of beech and kauri in Northland that does not appear to follow a time-transgressive pattern suggests that local site factors may indeed be important (Newnham et al., 2017a).
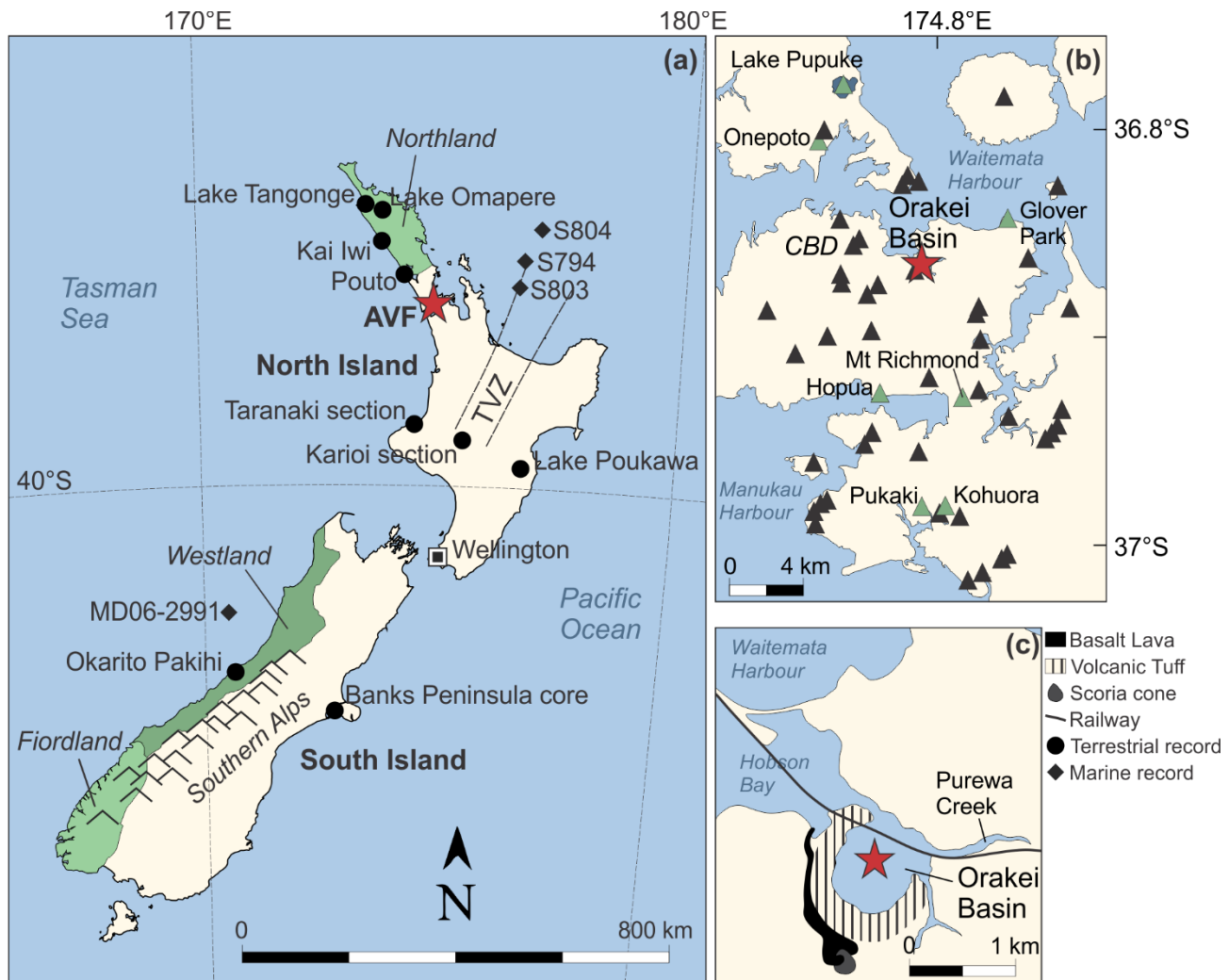
Figure 5.1: Map of New Zealand with locations and core sites mentioned in the text (a), map of Auckland Volcanic Field (adapted from Hopkins et al., 2017) (b), and map of Orakei Basin area showing the breached tuff ring (c). Green areas in (a) indicate the extent of Northland, Westland and Fiordland regions. Volcanic centres in (b) indicated with triangles, those mentioned in text coloured green. Star indicates study site/core location. TVZ = Taupo Volcanic Zone, CBD = Auckland's central business district.

Second, how does stadial MIS 4 compare to stadial MIS 2 in this record from Auckland? Recent research from a marine sequence offshore South Australia (MD03-2607) indicates that the climate of MIS 4 was almost as cold as the LGM, the coldest phase of MIS 2 (De Deckker et al., 2019). In New Zealand, glacial advances were more extensive during MIS 4 compared to MIS 2 (Williams et al., 2015). On the other hand, pollen records of MIS 4 from

Northland (Newnham et al., 2017a), south Westland (Newnham et al., 2007a) and a marine record off the coast of Westland (Ryan et al., 2012) show vegetation did not change as drastically as during MIS 2. The Orakei Basin record affords an opportunity to further investigate this conundrum.

Third, to what extent can the vegetation-climate patterns depicted in the Orakei Basin pollen record be explained by the classical MIS subdivision of the Late Quaternary (Lisiecki and Raymo, 2005)? Does the record display stadial and interstadial phases during the Last Glacial equivalent to those in the marine records? In addition, although the Orakei Basin record seems unlikely to encompass all of MIS 5, can any of the MIS 5 substages be observed in the record? There are a few fragmentary records of MIS 5 in the North Island, e.g. from Taranaki (Newnham and Alloway, 2004) and the Karioi section in the central North Island (Newnham et al., 2017b) (Fig. 5.1a), but we do not know whether the five-fold subdivision that is normally applied to MIS 5 (Lisiecki and Raymo, 2005) also applies further north.

5.2: Study Region

Orakei Basin (36°51'59"S, 174°48'46"E) is a large explosion crater and one of 53 basaltic centres of the Auckland Volcanic Field (AVF), which occupies a 600 km^2 area on the Auckland isthmus and the adjacent land areas to the north and south (Hopkins et al., 2017; Fig. 5.1). Several of the maars in the AVF have proven excellent depositional environments that have accumulated lacustrine sediments over the Last Interglacial-Glacial cycle, some dating back to 100 kyr BP and beyond (Leonard et al., 2017). These sediments have been used to reconstruct the eruptive history of the AVF (e.g. Molloy et al., 2009; Hopkins et al., 2017) and some palaeoecological work has been performed as well (e.g. Sandiford et al., 2003; Horrocks et al., 2005, Augustinus et al., 2011, 2012; Chapter 3).

Orakei Basin is one of the maars in the central part of the AVF, located ~4 km southeast of Auckland's central business district (Fig. 5.1b, c) with an estimated age of 126.0 kyr BP (Hopkins et al., 2017; 1 standard deviation error: 3.0 kyr). The basin features in several tephrostratigraphical studies (i.e. Molloy et al., 2009; Hopkins et al., 2015, 2017), but no palaeoecological work has been performed on the sediments. The basin was a lake for most of its history (Hayward et al., 2011), slowly filling up with mud carried in by Purewa Creek

(Fig. 5.1c), until it eventually became a swamp (see "Lithology"). Rising sea level caused the tuff ring enclosing the maar to be breached by the ocean around 9.0 – 8.5 kyr BP (Hayward et al., 2008), transforming it into a tidal lagoon, and later intertidal sand flats. Today, Orakei Basin is a shallow (~2 m deep), saltwater lake, its water level controlled through gates since a railway embankment was built across the flats in the mid-1920s. The crater has a 0.8-km diameter and is surrounded by a tuff ring up to 40 m high on the west, south and east sides. The crater wall is steep, and the basin is surrounded by hilly terrain.

The oceanic climate of the Auckland region is classified as warm temperate (Chappell, 2014) and exhibits few temperature extremes, with a mean annual temperature between 14 and 16°C and mean summer and winter temperatures around 18.9°C and 12°C respectively. Auckland is wet throughout the year, but with slightly elevated precipitation delivered by the westerlies in winter; annual rainfall is around 1100-1200 mm (Chappell, 2014). Modern New Zealand weather is strongly influenced by atmospheric circulation patterns that are dependent on a combination of broad scale climate modes, i.e. the Southern Annular Mode and El Niño-Southern Oscillation.

Prior to human arrival, the Auckland area supported extensive conifer-broadleaf forests, although the region is now almost completely urbanised. The natural vegetation cover was deduced from isolated remnants and Holocene palynological records (Wardle, 1991; Newnham et al., 2018). Conifer-broadleaf forests are temperate rainforest that can be classified into two main types, i.e. rimu-rātā or podocarp-broadleaf forest and kauri forest. In rimu-rātā forest, *Dacrydium cupressinum* (rimu) and *Metrosideros robusta* (northern rātā) are the principal canopy and emergent species. Other common canopy species in rimu-rātā forest include the broadleaved trees *Beilschmiedia tawa* (tawa), *Knightia excelsa* (rewarewa), *Elaeocarpus dentatus* (hīnau) and *Weinmannia silvicola* (tōwai) and the tall podocarps *Prumnopitys* spp. and *Podocarpus* sp. In kauri forest, *Agathis australis* (kauri) and *Phyllocladus trichomanoides* (tanekaha) are the major emergent trees, and the proportion of species that are part of the northern endemic group (e.g. *Beilschmiedia tarairi*, *Ixerba brexioides*) is higher. Temperate rainforests are multi-tiered; tree ferns (*Cyathea* spp. and *Dicksonia* spp.) characterise the subcanopy, ferns the herb tier, and lianas and epiphytes are abundant throughout. Interspersed throughout the conifer-broadleaf-dominated forests are isolated beech stands. *Nothofagus truncata* (hard beech; included in *Fuscospora* pollen taxon) prevails on sandstone ridges and kauri/hard beech communities

grow locally on shallow clay soils, from sea level to 600 m altitude (Wardle, 1991). It is possible that the modern presence of beech, primarily a taxon adapted to subalpine environments, is due to chance dispersal to suitable sites or that beech stands are relict from a time when climate was more favourable to beech.

5.3: Methods

5.3.1 Drilling

Orakei Basin was cored in February 2016 from an anchored barge, using a standard wire-line drilling system. Two cores were drilled roughly eight meters apart, i.e. OB16A (36.86647°S, 174.81278°E) and OB16B (36.86653°S, 174.81273°E) (Fig. 5.1c). The first c. 42.5 meters of OB16A were drilled in 1.5-m sections, but because of extreme expansion (up to 90%), it was decided to switch to 1-m sections for the remainder of OB16A and OB16B. For OB16A, we drilled until we reached several meters of local volcanic material between c. 100 and 105 m depth, i.e. coarse basalt gravel, lapilli and scoria chunks, prompting us to terminate. OB16A consequently consists of 91 core sections. For OB16B, we drilled to 101.5 m depth, but the top 16.5 m of marine sediments were discarded, resulting in a total of 85 meters of core. While drilling OB16B, the base of each section was chosen such that it was in the middle of the equivalent OB16A section; the resulting core sections from OB16A and OB16B therefore overlap, which ensures full sediment recovery even if the top and base of each individual section are disturbed. Core sections (6 cm diameter) were split lengthwise into one "working half" and one "archive half" (kept pristine) and stored at 4°C.

5.3.2 Composite depth and sampling

A composite depth scale was developed from the OB16A and OB16B cores by PhD candidate Leonie Peti based on visual comparison and x-ray fluorescence data, where the marine/freshwater transition (see Table 5.1) was set at 0 m depth. Working halves were sampled at 10 cm intervals by taking 1-cm wide wedges. For the top 56 m (composite depth) of the core, OB16B was sampled, while for 56 to 84 m composite depth, OB16A was sampled. The reason for the switch from OB16B to OB16A was logistic: the lower sections of OB16B were kept unsplit and dark in anticipation of optically stimulated luminescence dating.

Table 5.1: Simplified lithology.

Code	Depth (m)	Description	Structure	Pollen zone
0	-17.0 – 0.00	Marine mud	massive	-
1 a	0.00 – 0.30	Peat	massive	5
b	0.30 – 0.59	Beige, massive clay	massive	5
c	0.59 – 1.39	Peat	massive	5
2 a	1.39 – 1.76	Beige, massive clay	massive	5
b	1.76 – 1.96	Sand	massive	5
c	1.96 – 9.35	Light-brown, massive clay with bioturbation	massive	5
3 a	9.35 – 10.16	Finely banded, greyish sand and light-brown silt	banded	4
b	10.16 – 13.81	Disturbed, banded sand, including many organics	banded	4
c	13.81 – 16.97	Fine, light sand bands in brown sand/silt, with some larger organics	banded	4
4	16.97 – 26.60	Light-brown clay with fine laminations	laminated	4
5	26.60 – 27.92	Brown, greyish, very fine clay with light, wavy laminations	laminated	4
6	27.92 – 29.26	Massive, light-brown clay with basaltic tephra and disturbance	massive	"no-pollen zone"
7	29.26 – 31.41	Reddish-brown clay with fine laminations and abundant basaltic tephra	laminated	3, "no-pollen zone"
8 a	31.41 – 45.14	Dark-brown, fine laminations (colour contrast weak)	laminated	3
b	45.14 – 50.95	Dark-brown, fine laminations with thin turbidites (colour contrast strong)	laminated	3
9 a	50.95 – 51.18	Dark silt	massive	3
b	51.18 – 55.74	Dark, banded sand and silt	banded	3
c	55.74 – 56.45	Dominantly sand	massive	3
10	56.45 – 65.68	Fine, brown laminations with frequent thin turbidites	laminated	1 – 3
11 a	65.68 – 67.89	Silt with thin sand layers	banded	1
b	67.89 – 68.29	Abundant sand bands in silty matrix	banded	1
c	68.29 – 70.81	Coarse sand with big wood pieces	massive	1
d	70.81 – 73.46	Thick sand bands in silty matrix with some wood	banded	1
12 a	73.46 – 75.35	Finely laminated clay with abundant silt and sand layers, with some larger organics	laminated	1
b	75.35 – 76.70	Finely laminated, dark-brown clay (strong colour contrast) with very abundant turbidites	laminated	1
13 a	76.70 – 77.87	Dark-brown clay with abundant turbidites and fine laminations	laminated	1
b	77.87 – 82.55	Finely laminated, dark-brown clay (strong colour contrast) with very abundant turbidites	laminated	1
14 a	82.55 – 82.87	Massive, light-brown clay with layered basaltic ash	massive	1
b	82.87 – 84.08	Massive, light-brown clay	massive	1
15	84.08 – 86.33	Basaltic ash (some wood debris)	massive	-
16	86.33 – 88.40	Country rock, sandstone	massive	-

For pollen analysis 0.5-to-0.75-mL of material was subsampled from the 1-cm wedges. Subsamples were selected every 0.5 m, except for the sections between 11 and 14 m composite depth (because of extremely low pollen concentrations) and 66 and 75 m composite depth (because sediments were considered disturbed), where subsamples were chosen every 1 m. The final distribution of samples over depth is not as constant as described here, because samples were selected based on the original core section depths, before the composite stratigraphy was established.

For chironomid analysis, 1-mL of material was subsampled from the 1-cm wedges at 2-m intervals. However, chironomid head capsule concentrations were found to be low throughout most of the Orakei Basin core. It was therefore not possible to develop a chironomid record for this study, although alternative methods to concentrate head capsules from sediments will be tested in the future (see Appendix E).

5.3.3 Chronology

The strategy for developing the chronology of the Orakei Basin core relies upon work undertaken as part of the Marsden-funded Auckland maars project. The goal was to develop an integrated chronology for the Orakei Basin core, based on tephrostratigraphy, radiocarbon dating, luminescence dating on feldspars, paleomagnetic variation (direction and intensity) and Beryllium-10 variation. However, at the time of writing the analyses were still in progress and the age-depth model that will result from these analyses is not yet available. As a result, for this thesis, the decision was taken to develop the Orakei Basin pollen record presented here with preliminary ages based on tephrostratigraphy.

Twelve tephra layers were investigated (Table 5.2). At The University of Auckland, bulk tephra samples were washed and sieved to remove marine sediments and organics, and the 125-250 µm size fraction was retained for in-situ glass shard analysis. Samples were mounted into epoxy rounds, polished and carbon coated. This sample preparation was undertaken by PhD candidate Leonie Peti. The prepared samples were sent to Victoria University of Wellington for electron microprobe analysis (EMPA) for major elements using a JEOL JXA 8230 Superprobe (EMPA). This work was undertaken by Dr Jenni Hopkins. Assignment of the tephra sources was based primarily on their major element composition

in comparison to tephra layers identified by Molloy (2008) and Allan (2008), in combination with their stratigraphic position and thickness.

Table 5.2: Analysed tephra layers and preliminary assignments

Sample	Composite depth (m)	Thickness (mm)	Composition	Preliminary assignment	Age (cal yr BP)[4]	Age reference
T15	5.235	20	Rhyolitic	Rotorua	15,635 ± 412	Lowe et al. (2013)
T01	20.770	3	Rhyolitic	Okareka	21,858 ± 290	Lowe et al. (2013)
T02	24.643	25	Rhyolitic	cf. KOT[1]	25,360 ± 160	Vandergoes et al. (2013)
T04	31.495	5	Rhyolitic	cf. Okaia	28,621 ± 1428	Lowe et al. (2013)
T12	35.434	3	Andesitic	Eg25[2]	30,700[5]	Molloy et al. (2009)
T71	41.082	15	Rhyolitic	Hauparu	36,000[5]	Molloy et al. (2009)
T72	42.658	25	Rhyolitic	Maketu	36,320 ± 1150	Molloy et al. (2009)
T73	42.725	2	Rhyolitic	Tahuna	39,268 ± 2386	Molloy et al. (2009)
T74	46.142	315	Rhyolitic	Rotoehu	45,170 ± 6600	Danišik et al. (2012)
T66	77.260	15	Basaltic	N/A		
T68	82.322	10	Basaltic	Orakei[3]		
T69	82.643	65	Basaltic	Orakei[3]		

[1] Kawakawa/Oruanui Tephra
[2] Source: Egmont Volcano
[3] Reworked
[4] Age ± 95% probability range for ages from Lowe et al. (2013), ± 2 standard deviations for Molloy et al. (2009), Danišik et al. (2012) and Vandergoes et al. (2013)
[5] Inferred from sedimentation rates in old Orakei core, drilled in 2007

For EMPA a defocussed beam of 10 µm at 8nA was used to analyse major elements as oxides (SiO_2, TiO_2, Al_2O_3, FeO_t, MnO, MgO, CaO, Na_2O, K_2O) and Cl^-. Matrix-matched standards VG-586 and VG-A99 were used as bracketing standards and run twice at the end of each sample to monitor instrument drift. Accuracy of the analyses was within 6% of the recommended values for all elements (except TiO_2 and Ca) that were ≥ 10%) for the internal standards, and analytical precision (2 standard deviations) is < 0.3 for all elements. Oxide concentrations were determined using the ZAF correction method, with analytical totals for glass shards 95-99%. Deviations from 100% are attributed to variable degrees of post eruption hydration and all reported data are corrected to 100% on this basis.

5.3.4 Palynology

Pollen samples were prepared according to standard procedures (Moore et al., 1991): hydrochloric acid was used to remove carbonates, potassium hydroxide to remove humic acids, density separation with sodium polytungstate (density: 2.2 g/cm³) to separate pollen from the minerogenic fraction, and acetolysis to remove cellulose. To allow for the calculation of concentrations, one *Lycopodium* tablet was added per sample. The processed material was mounted on microscope slides in glycerine jelly and analysed under a light

microscope at 400x or 1000x magnification. We identified 250 pollen sum grains per sample with the aid of pollen guides (Moar, 1993; Pocknall, 1981), where the pollen sum includes dryland pollen and excludes wetland taxa and spores. We additionally recorded 37 non-pollen palynomorph types, which are described in Appendix D. Taxa were expressed as percentages of the pollen sum and the naming convention suggested by Moar et al. (2011) was applied. Zonation of the diagram was based on visual comparison aided by clustering using the CONISS method and principal component analysis based on all dryland taxa.

To estimate effective precipitation, the Pollen Moisture Index (PMI) was calculated as the normalised, base-10 logarithm of the ratio of *Dacrydium cupressinum* percentages over the percentage of *Podocarpus* and *Prumnopitys* spp. (Jara et al., 2017). The relative abundance of *D. cupressinum* has previously been used to infer moisture conditions during the Holocene, as this tall tree is less tolerant of drought and more tolerant of waterlogged soils than *Podocarpus* and *Prumnopitys* spp (Franklin, 1968; McGlone & Topping, 1977). The PMI has not been used on pollen assemblages under glacial conditions before (see section 5.5.1.2). A Generalised Additive Model (GAM) was fitted to PMI over time using the mgcv package (Wood, 2006; 2011).

Mean annual air temperatures (MAAT) were reconstructed from the pollen percentages using the weighted averaging modern analogue (MA) technique and partial least squares (PLS) model described by Wilmshurst et al. (2007), using the analogue (Simpson, 2007; Simpson & Oksanen, 2016) and rioja (Juggins, 2015) packages in R (version 3.5.1; R Core Team, 2018). The pre-deforestation training set these models are based on was slightly adapted to match the Orakei Basin pollen dataset (see Appendix C1). A GAM was fitted to the reconstructed MAAT over time using the mgcv package (Wood, 2006; 2011).

5.4: Results

5.4.1 Lithology

Large sections of the Orakei Basin core are laminated, though some parts are banded or massive (Table 5.1). Laminations are variable in thickness and colour. The banded sections (i.e. code 3a-c, 9b and 11a, b and d) contain sand layers, which may be indicative of storm events and/or increased erosion and in-wash. The development of peat at 1.39 m depth

(lithology 1c) indicates that the Orakei Basin lake was transitioning to terrestrial wetland at the time. The basin may have returned to its lake state between 0.59 and 0.30 m depth (lithology 1b), before returning to a wetland state (lithology 1a). As mentioned in the Study Region section, Orakei's crater rim was breached by the ocean around 9.0 – 8.5 kyr BP (Hayward et al., 2008), transforming the maar into a tidal lagoon, over-topping these pollen-bearing sediments with ~17 m of marine muds (lithology 0).

5.4.2 Chronology

The primary age constraints of the Orakei Basin pollen record include the age of the crater and the breaching of the crater rim by the ocean. The age of the crater, 126 ± 3.0 kyr BP, was inferred from the depth of its associated tephra layer in the Glover Park core (another maar in the AVF; Fig. 5.1b) and calculated from sedimentation rates (Hopkins et al., 2017). This age constitutes the maximum age of the basal sediments; as we do not know when consistent sediment deposition, undisturbed by slumps of native volcanic material, commenced, this age is highly uncertain. The marine/freshwater transition (0 m composite depth) occurred ~9.0 – 8.5 cal kyr BP (based on the radiocarbon date of a sample taken 1 cm below the transition; Hayward et al., 2008). Therefore, the 85 m of sediments cover a maximum of 117.5 kyr.

Additional age constraints were derived from seven of the twelve analysed tephra layers. Eight were analysed by EMPA as rhyolitic in composition and one as andesitic, most of which could be preliminarily assigned to a source volcano (Table 5.2, Fig. 5.2). Tephra layers T66, T68 and T69 are basaltic and derived from AVF eruptions, the bottom two of which are reworked material from Orakei Basin itself. Two of the rhyolitic tephra layers, T04 and T02 have very similar major element signatures (Fig. 5.2). Based on their stratigraphic position, thickness (Table 5.2) and elimination of other options, it is likely that T02 is the Okaia tephra and T04 is the Kawakawa/Oruanui tephra (KOT). The Okaia tephra is older than the KOT (Lowe et al., 2013), and the KOT represents a much larger eruption than Okaia. Other know rhyolitic tephra deposits of the same approximate age include Te Rere and Poihipi, but their geochemical data does not match the geochemistry of T02 and T04 (Molloy et al., 2008). However, to acknowledge the uncertainty of these assignments, we indicate T04 as "cf. Okaia tephra" and T02 as "cf. KOT", and they were excluded from any

age or sedimentation rate calculations. This uncertainty will be resolved in the future when more information (e.g. additional age constraints, trace element data) becomes available.
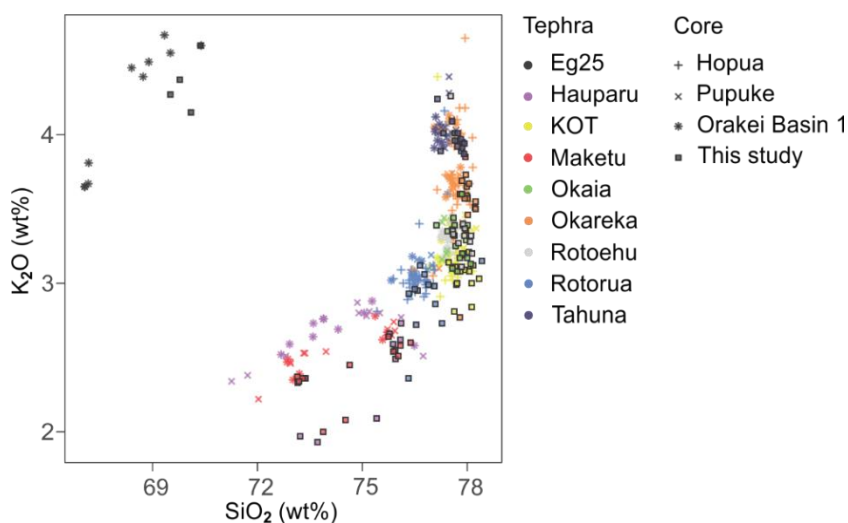
Figure 5.2: Major element composition of the tephra samples (Table 5.2). Comparison data from three AVF cores (Hopua, Pupuke and an old Orakei core, see Fig. 5.1a) from Molloy (2008); squares indicate data from this study. Colours indicate different tephra layers.

Based on the limited age information we calculated sedimentation rates throughout the record (Fig. 5.3b). Sedimentation rates appear lower (~0.48 mm/yr) before the deposition of the Rotoehu tephra (45.2 cal kyr BP) than after (average ~1.28 mm/yr). There is one extreme peak in sedimentation between the Maketu and Hauparu tephra layers, which could be due to errors in their ages; the age of the Hauparu tephra is inferred from sedimentation rates in Molloy et al.'s (2009) record. Excluding this extreme peak, highest sedimentation rates occur between ~20 and 5 m depth (2.50 mm/yr), which concurs with low pollen concentrations in this part of the record.

5.4.3 Palynology

The pollen record shows a five-fold division based mainly on the abundance of: tall trees, which are most abundant in Zones 1, 3 and 5; small trees and shrubs, which show peak abundance in Zone 2; and herbs, which are most abundant in Zone 4 (Fig. 5.4). A detailed description of each zone follows.
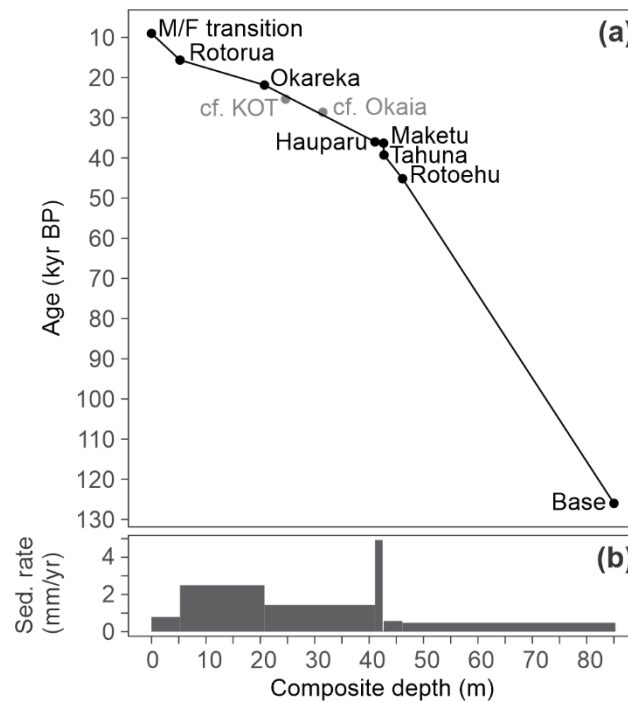
Figure 5.3: Simple age-depth model based (a) and sedimentation rates (b) based on linear interpretation between age constraints. Tephra layers are indicated (see Table 5.2), base set at 126 kyr BP (maximum age) and top (marine/freshwater transition) set at 9 kyr BP. The age and possible depth of the KOT and Okaia tephra are indicated in grey.

5.4.3.1 Zone 1 (83.7 – 62.1 m)

Tall trees dominate; mostly *Dacrydium cupressinum*, *Prumnopitys taxifolia*, *Metrosideros* and *Agathis australis*, indicative of conifer-broadleaf or kauri forest. *Alectryon excelsus* is present throughout, with a sudden peak of 37.6% at 69.65 m, which is within lithology 11c. *Cyathea dealbata* is variable, with a peak at 69 and 66 m (within lithology 11c and 11a respectively). *Cyathea smithii* type increases throughout the zone. Monoletes increase too, until ~67 m, broadly coinciding with the boundary between lithology 11a and 10, after which their abundance declines sharply and stays low until the top of the zone. Algae are almost absent until ~65 m, after which *Botryococcus* increases, with a peak of 83.0% at 62.61 m.
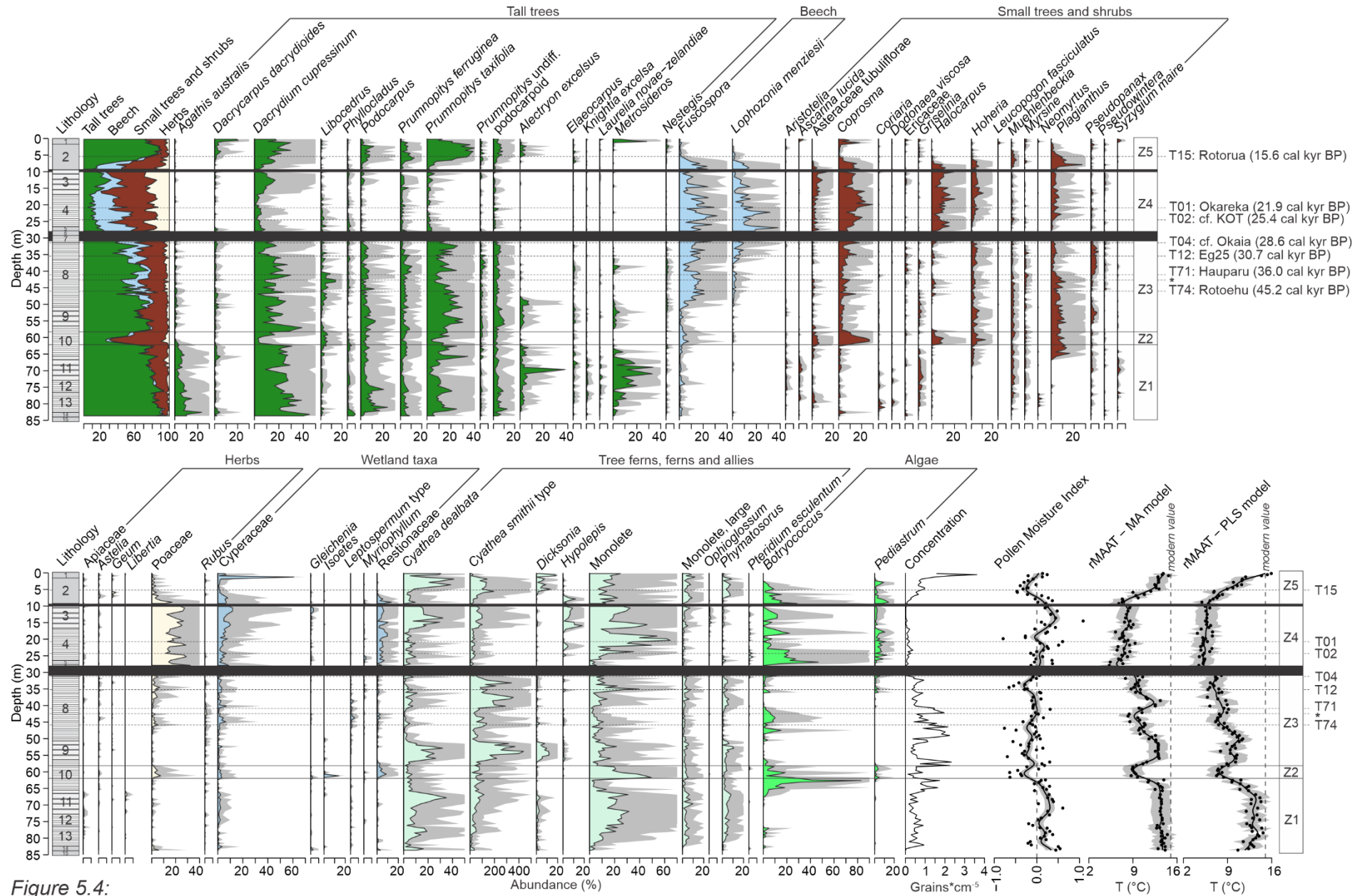
Figure 5.4:
Caption on next page

Figure 5.4: Pollen and spore percentage diagrams. Scale intervals of all taxa represent 10%, except for Cyathea smithii *type, for which each scale interval represents 100%. Diagram excludes rare taxa (taxa that occur in < 3 samples and/or with a maximum abundance of < 2%), which are shown in Appendix C2. Shading represents 5x exaggeration. Also included are estimated effective precipitation (PMI) and reconstructed MAAT (both from modern analogue model and PLS method). Points are reconstructed values, lines represent the GAM through points, shaded area reflects RMSEP of transfer function (except for PMI, where shaded areas represent GAM error). Zone numbers indicated on the right of the diagram, tephra layers indicated on the right of the pollen zones; asterisk indicates tephra layers Maketu and Tahuna (Table 5.2). Lithology numbers refer to Table 5.1.*

_____

5.4.3.2 Zone 2 (62.1 – 58.3 m)

Tall trees decline suddenly and *Metrosideros* disappears completely, while small trees and shrubs rise. The vegetation is mostly composed of Asteraceae, *Coprosma*, *Halocarpus*, *Hoheria* and *Plagianthus*, while *Libocedrus*, *Fuscospora* and Poaceae also increase. These changes indicate an expansion of subalpine shrubland vegetation, although conifer-broadleaf forest is likely still present in the wider region or locally at reduced levels. Regarding wetland taxa, Restionaceae are more abundant, indicating the presence of wetlands nearby. *Cyathea* spp. spores are relatively rare, while monolete spores are very abundant. *Botryococcus* declines in abundance compared to the latter part of Zone 1.

5.4.3.3 Zone 3 (58.3 – 29.4 m)

Tall trees return, i.e. mostly *D. cupressinum* and *Prumnipitys taxifolia*, while *Agathis australis* is only present at very low abundances. *Metrosideros*, indicative of rimu-rātā forest, is present in low abundances until ~48.5 m depth (= approximate depth of Rotoehu tephra) and then disappears. *Fuscospora* and *Lophozonia menziesii* appear a bit earlier, from ~50.4 m depth, at the expense of *Prumnopitys taxifolia*, signalling the advance of beech forest or mixed beech/conifer forest. *Cyathea* spp. and *Dicksonia* are very abundant during the rimu-rātā phase (*C. smithii* up to 438%), which broadly overlaps with lithology 9b: banded sand and silt. *Botryococcus* is absent at this time. *C. smithii* is additionally very abundant in the latter part of the zone, ~40.2 and 32.4 m depth. Between Zones 3 and 4 there is a 3-m interval (from 30.5 to 28.5 m) where no pollen was found.

5.4.3.4 Zone 4 (29.4 – 9.7 m)

Pollen concentrations are very low, i.e. < 50k grains/cm^3. Tall trees appear to decline drastically, while *Lophozonia menziesii* and small trees and shrubs (mostly *Coprosma* and *Halocarpus*) and Poaceae increase. These taxa are indicative of a beech-subalpine scrub complex or beech forest mosaic. Wetland taxa Cyperaceae and Restionaceae are abundant throughout, signalling the return of wetland hanitat nearby. Algae are abundant as well. *Cyathea* spp. are relatively subdued but increase in the second half of the zone, i.e. from ~27.5 m. Monolete spores are abundant but highly variable.

5.4.3.5 Zone 5 (9.7 – 0 m)

Tall trees increase during the first half of this zone, especially *P. taxifolia* and *Dacrydium cupressinum*, while beech trees, small trees and shrubs and Poaceae decline. Near the top of the sequence, from 0.9 m, *Metrosideros* and *Ascarina lucida* increase sharply. Regarding local indicators, Restionaceae declines in the first phase, as does Cyperaceae, but Cyperaceae suddenly increases again at 1.3 m, coinciding with peat development (lithology 1c). Algae are present until the start of peat development (1.3 m).

5.4.4 Climate reconstructions

Effective precipitation, as inferred from PMI values (Fig. 5.4), appears to be high during pollen Zone 1, especially between 73.0 and 66.9 m. PMI declines during Zone 2 and stays low in the lower part of Zone 3, until 43.7 m when values rise. There is another minimum in PMI between 36 and 30 m. During Zone 4, PMI initially displays intermediate values, then values peak between 16.5 and 8.5 m. During Zone 5, PMI drops to a minimum at ~5 m depth and then increases again.

To reconstruct MAAT, the MA technique performed slightly better than the partial least squares (PLS) model, with an RMSEP$_{boot}$ of 1.52 vs. 1.59 and an r^2_{boot} of 0.78 vs. 0.76. However, in the analysis of modern analogues for the fossil dataset it was found that 138 of 143 fossil samples do not have a close modern analogue in the pre-deforestation training set (Fig. 5.5b). On the other hand, most of the fossil samples do show a good fit to MAAT (Fig. 5.5a). Both models disproportionately overestimate temperatures at the lower end of

the MAAT gradient (< 8°C; Appendix C1, Fig. C-3). PLS models in general perform better than the MA technique in the absence of modern analogues for the fossil samples. Therefore, subsequent reported temperatures are based on the PLS model, unless specified otherwise. Reconstructed MAAT is on average 1.8°C higher for the modern analogue model compared to the PLS model (Fig. 5.4). See Appendix C1 for more details on the development of the MAAT reconstruction.
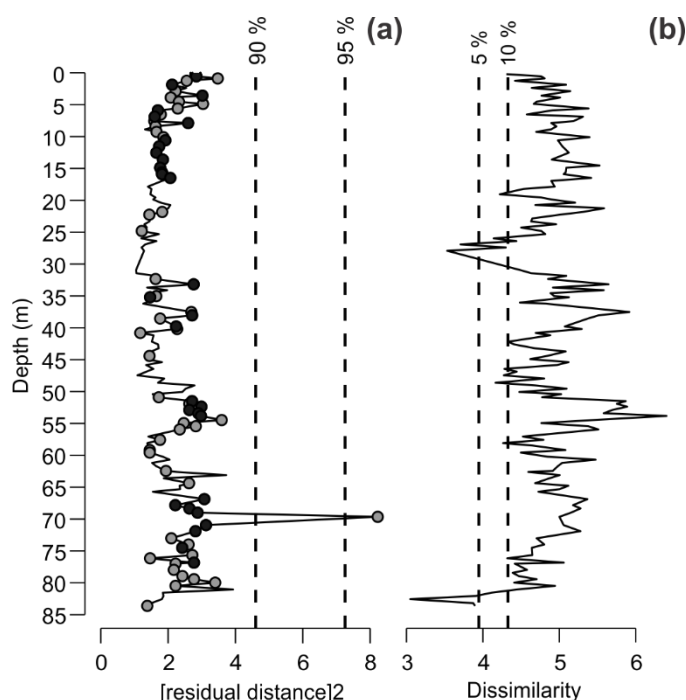
Fig. 5.5: Goodness of fit of fossil samples to MAAT (a) and dissimilarity of fossil samples to modern analogues in the training set (b). vertical lines represent percentiles of the squared residual distances of training set samples in (a) and percentiles of the distribution of pair-wise dissimilarities for the training set samples in (b). In (a), samples with > 5% rare taxa (Hill's N2 in training set ≤ 5) and those taxa not included in the inference model are indicated with grey symbols, > 7.5% with black symbols.

Reconstructed MAAT (Fig. 5.4) is high during pollen Zone 1, i.e. ~12.9°C, then decreases at the end of Zone 1 to a minimum in Zone 2 of ~7.2°C. During Zone 3, reconstructed MAAT is highly variable with a slight downward trend from ~10 to 8°C (average: 8.6°C). The reconstruction from the modern analogue model shows a cooler period between 44.9 and 42.6 m commencing just before the deposition of the Rotoehu tephra (minimum ~8.2°C), and a maximum at ~38.5 m (~13.3°C). Zone 4 displays the coolest temperatures of the record, i.e. average MAAT lies around 5.4°C. Temperatures increase rapidly throughout Zone 5 from ~6 to 16°C. According to the modern analogue model, warming plateaus around 4.9 m depth, while the PLS model shows MAAT rising towards the top of the core.

5.5: Discussion

5.5.1 Reliability of climate reconstructions

5.5.1.1 Mean Annual Air Temperatures

The main changes in reconstructed MAAT match the main shifts in pollen assemblages (Fig. 5.6a), indicating that either MAAT is the primary driver of vegetation change throughout the record, or both are driven by an underlying confounding factor. This is in contrast to the Lake Pupuke MAAT reconstruction (Chapter 3), where shifts in vegetation are superimposed on a stable MAAT plateau and vegetation change appears to be driven by seasonality and effective precipitation. However, the lack of modern analogues in the pre-deforestation dataset for the fossil assemblages (Fig. 5.5b) is a concern. This is probably because the Orakei samples contain high abundances of rare taxa and taxa that are not included in the model, such as *Hypolepis* (rare) and *Phymatosorus* (not modelled). On the other hand, the good fit of the fossil assemblages to MAAT in constrained ordination (Fig. 5.5a) strengthens our confidence in the PLS-based MAAT reconstruction. Our reanalysis of the adapted pre-deforestation dataset shows that the MAAT models are less reliable at the lower end of the MAAT gradient (Appendix C1, Fig. C-3), which has also been pointed out in previous papers (e.g. Wilmshurst et al., 2007; Newnham et al., 2013; Jara et al., 2015). In addition, high percentages of *Fuscospora* have been shown to distort temperature reconstructions (e.g. Jara et al., 2015) as *Fuscospora* pollen is often overrepresented. However, this is mainly a concern for southern or alpine sites, where *Fuscospora* is overrepresented against a background of open grassland vegetation or scrub, leading to abundances over 50% (e.g. Jara et al., 2015). In the Orakei pollen record, *Fuscospora* is always present in conjunction with other forest taxa and never reaches over 32%.

5.5.1.2 Pollen Moisture Index

The PMI (Jara et al., 2017), which we use to qualitatively reconstruct effective precipitation (see also Chapter 3), was based on the "rimu ratio", which has previously been used to infer moisture conditions in the North Island (McGlone & Topping, 1977; Newnham et al., 1995). Rimu (*Dacrydium cupressinum*) is less tolerant of droughts and more tolerant of waterlogged soils than other tree podocarps (Franklin, 1968). *Dacrydium cupressinum* and *Prumnopitys*

spp. are prominent tall trees in the modern vegetation of the North Island. However, it is unclear whether calculation of the PMI is appropriate for records beyond the transition from the Last Glacial to the Holocene. In particular during the Last Glacial, when forest levels were reduced substantially, precipitation controls on the distribution of forest species and on forest dynamics may have differed substantially from the Holocene. Whereas the PMI may accurately reflect effective precipitation in interglacials other than the Holocene, we must be careful when interpreting PMI during cooler intervals, when temperature may have been a more important control on tall tree podocarp (relative) abundances.

5.5.2 Vegetation and climate history at Orakei Basin

Although we have some indication of ages throughout the core, i.e. maximum age of 126 kyr BP, top dated at ~9 kyr BP and nine identified tephra layers (Table 5.2), our chronology for the Orakei Basin pollen record is uncertain. Below the Rotoehu tephra (at ~46 m depth, 45.2 cal kyr BP), we have no age indication at all, except for inference from assumed constant sedimentation rates (Fig. 5.3). While a chronology that integrates several different methods (see section 5.3.3) is anticipated in the near future, for this thesis we must be cautious with our interpretation. Based on the limited information we have, we can tentatively assign the five pollen zones (Fig. 5.4) to glacial and interglacial or stadial and interstadial conditions that correspond to MIS (Fig. 5.6c; Lisiecki & Raymo, 2005), as will be laid out in the following section.

5.5.2.1 Comparison to Marine Isotope Stages

Pollen Zone 1 shows the highest abundance of tall trees and includes interglacial taxa such as *Agathis australis* and *Ascarina lucida*, the former indicative of northern kauri forest. Reconstructed temperatures (Fig. 5.6a) are high, but not as high as present-day temperatures at Orakei Basin (~2°C cooler). In the absence of any independent age markers, we tentatively assign Zone 1 to MIS 5, probably substage MIS 5a (see section 5.5.2.2 for alternative explanations). The stable MAAT and high tall tree abundance give us confidence in the effective precipitation reconstruction for this zone, as conditions are similar to the present-day situation on which the index is based. The PMI indicates generally wet conditions, especially above ~75 m depth.
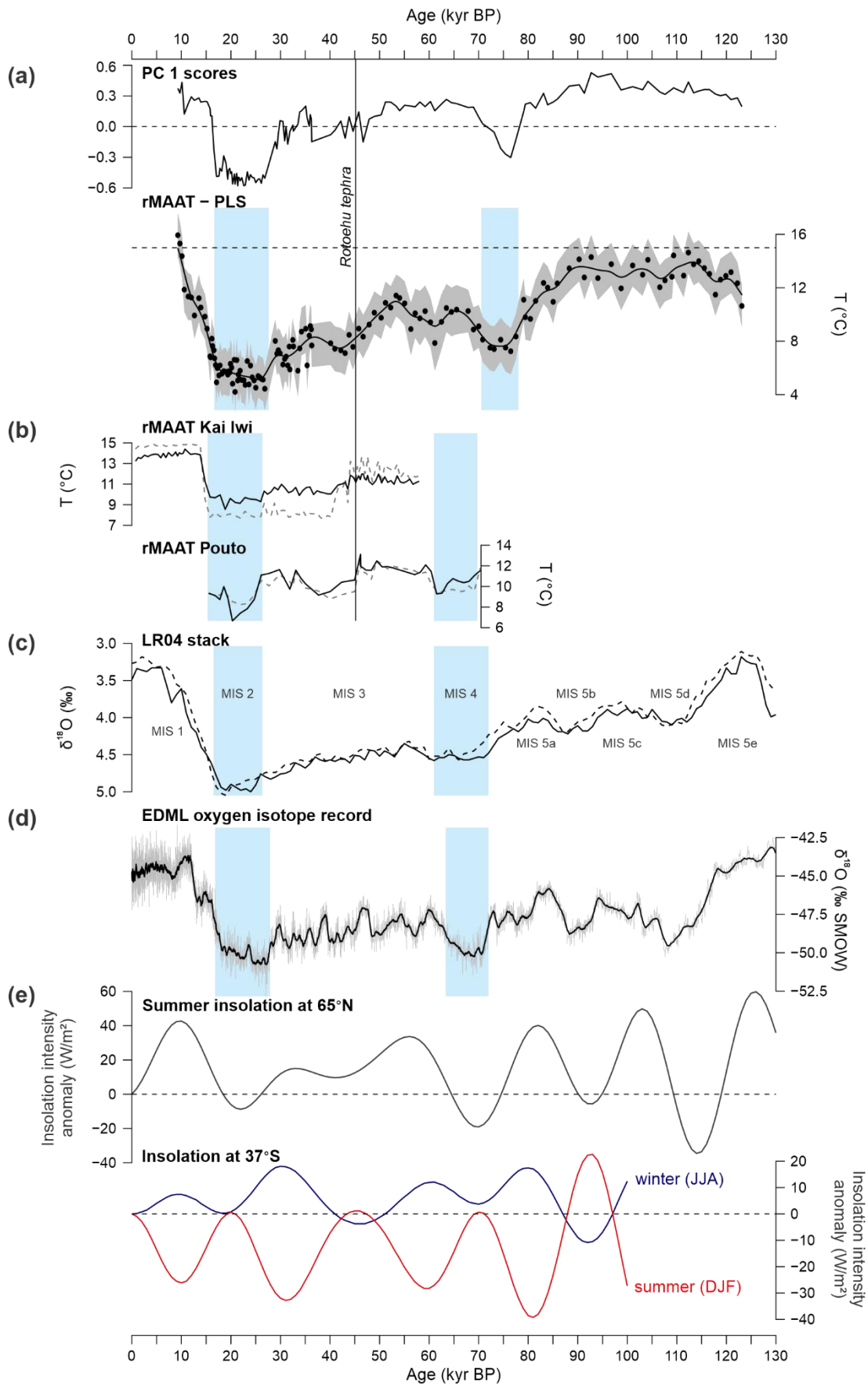
Figure 5.6:
Caption on next page

Figure 5.6: Orakei Basin axis 1 values from a PCA on all dryland pollen taxa and MAAT record (PLS model) (a) compared to Northland MAAT records Kai Iwi and Pouto (b), MIS (Lisiecki and Raymo, 2005) (c), the oxygen isotope record of Antarctic ice core EDML (Bazin et al., 2013) (d), and insolation intensity curves (Berger & Loutre, 1991) (e) . In (a), Points are reconstructed values, line represents the GAM through points, shaded area reflects RMSEP of transfer function. In (b), dashed line represents reconstructed MAAT using the MA model, solid line the PLS model. Both the Pacific (solid line) and Atlantic (dashed line) $\delta^{18}O$ stacks of MIS are indicated in (c). In (d), the running mean of 25 samples is indicated in black. In (e), grey line represents Northern Hemisphere summer insolation intensity, red line represents local summer (December-January-February) insolation anomaly and blue line represents local winter (June-July-August) insolation anomaly. The position of the Rotoehu tephra is indicated in (a) and (b). Blue zones represent the coldest phases.

Zone 2 shows an abrupt decrease in tall trees, while shrubs like Asteraceae and *Halocarpus* increase. *Fuscospora* expands as well, but *Lophozonia menziesii* does not. This period registers as a brief cold spell (~7°C cooler than present) in the MAAT reconstruction. Although our basic age-depth model based on linear intrapolation between the estimated basal age of the crater and the position of the Rotoehu tephra indicates the age of this zone is around ~80 to 70 kyr BP, we believe it represents stadial MIS 4 (~71 – 57 kyr BP). The PMI indicates dry conditions, although this interpretation is uncertain as tall podocarps are rare at this time.

Zone 3 is more heterogenous: tall trees return, initially mainly taxa that are indicative of rimu-rātā forest, such as *Metrosideros*, whereas beech appears from ~50.4 m depth. MAAT is also variable throughout this zone: temperatures vary between 2 to 7°C cooler than today with a long-term cooling trend. Effective precipitation is low, except for a period of intermediate PMI values between ~45 and 35 m depth. Zone 3 is consistent with the interstadial conditions of MIS 3 (~57 – 29 kyr BP), which is confirmed by the ages of the identified tephra layers: Rotoehu (~45.2 cal kyr BP), Tahuna (39.3 cal kyr BP), Maketu (36.3 cal kyr BP) and Hauparu (~36 cal kyr BP) (Table 5.2).

Between Zone 4 and 3, close above the cf. Okaia tephra (~28.6 cal kyr BP), lies a 3-m thick band of sediments (31 to 28 m depth) that does not contain pollen. Possible explanations for the absence of pollen in this part of the record are considered in section 5.5.2.3.

Zone 4 is characterised by a high abundance of Poaceae and other subalpine indicators, such as *Halocarpus*, and in combination with the lowest reconstructed MAAT of the record (9.5°C cooler than present), points towards MIS 2 (29 – 14 kyr BP). The cf. KOT (~25.4 cal kyr BP), Okareka tephra (21.9 cal kyr BP), and cf. Okaia tephra at the base confirm this assignment. PMI values are intermediate during Zone 4, until ~18 m depth when PMI increases at the same time *Dacrydium cupressinum* abundance increases, above the Okareka tephra. However, *Prumnopitys* spp. abundances are very low in Zone 4, so PMI values may not be reliable.

Zone 5 represents the tail end of MIS 2 and transition to MIS 1, also known as the Last Glacial/Interglacial transition (LGIT), which occurred between 18 and 11.7 kyr BP. Temperatures increase rapidly from the MIS 2 minimum (6°C) to modern day temperatures (16°C). By the time of the Rotorua deposition (15.6 cal kyr BP), *Prumnopitys taxifolia* has reached maximum abundances. *Dacrydium cupressinum* becomes dominant when *P. taxifolia* begins to decline. The top metre of the sequence shows the appearance of interglacial taxa, such as *Metrosideros* and *Ascarina lucida*. PMI decreases to a minimum at ~5 m and subsequently increases towards the top of the core, consistent with the PMI record from nearby Lake Pupuke (Chapter 3).

5.5.2.2 MIS 5: alternative interpretations

MIS 5 is conventionally divided into five substages (Fig. 5.6c; Lisiecki and Raymo, 2005), i.e. MIS 5e represents the Last Interglacial (LIG), MIS 5d and 5b are stadials, and MIS 5c and 5a interstadials. MIS 5e is thought to have been warmer than present in New Zealand (Williams et al., 2015; Lorrey & Bostock, 2017) and in many other parts of the world (Fischer et al., 2018), whereas the other MIS 5 substages were likely cooler, especially the stadials when subalpine shrubland vegetation expanded in Taranaki (Newnham and Alloway, 2004) and the central North Island (Karioi section; Newnham et al., 2017b). As Orakei Basin Zone 1 does not show evidence for cool episodes (Fig. 5.4), nor do reconstructed temperatures rise beyond modern MAAT (but see section 5.5.1.1 for a discussion of the reliability of the MAAT reconstruction), it appears that only substage 5a is represented in our record. However, this interpretation assumes that the Auckland vegetation responded in a similar fashion to other regions, which leads to a circular argument, as we do not know how MIS 5 manifested in Auckland. The MIS 5d and 5b stadials might not have been cool enough to be

registered in the Orakei Basin pollen record, and Zone 1 may represent the entire MIS 5. Mild stadials would be consistent with the less severe MIS 4 stadial period (see section 5.5.4.1).

Another potential explanation for the absence of MIS 5 stadials in the Orakei Basin record is that the sediment core contains hiatuses. One or more substages within MIS 5 might not have been registered because there was no sediment accumulation during these periods. However, there is no indication of a pause in sediment deposition in the stratigraphy, nor any obvious mechanism that could have resulted in a pause. The Orakei crater is very deep, i.e. > 88 m, and assuming sea level was comparable to present day or even higher, water depth must have been substantial and not very responsive to droughts. The PMI does not indicate dry spells either.

Whether the MIS 5 stadials are represented or not, there are some indicators of interglacial conditions in Zone 1 of the Orakei pollen record, which are an argument for MIS 5e/LIG representation. *Ascarina lucida*, *Metrosideros* and *Agathis australis* are present in substantial abundances, but are absent or occur at very low rates during interstadial MIS 3 (Zone 3). This seemingly contradicts the assignment of Zone 1 as interstadial MIS 5a. The pollen assemblages of MIS 5a in Taranaki (~240 km south of Auckland; Fig. 5.1a; Newnham & Alloway, 2004) show the return of lowland podocarp forest after stadial MIS 5b, but are missing some key interglacial taxa, like *Ascarina lucida*. However, *Metrosideros* and *A. lucida* are thought to be particularly sensitive to winter temperatures, as both are frost sensitive (McGlone & Moar, 1977). Their presence indicates that winters must have been mild, which is consistent with high local winter insolation intensity during MIS 5a (Fig. 5.6e) and not inconsistent with stable MAAT slightly lower than present day. At the more southern Taranaki site, frosts may have persisted through MIS 5a, or other climatic conditions (such as humidity) were not suitable for these taxa further south.

On the other hand, *Agathis australis* (kauri) is associated with dry, hot summers (Ogden & Ahmed, 1989) and high local summer insolation intensity (Chapter 3), but MIS 5a coincides with a summer insolation minimum (Fig. 5.6e). Kauri is absent from many early-Holocene North-Island records (Newnham, 1999; Ogden et al., 1992; Newnham et al., 1995), including from the Auckland region (Newnham & Lowe, 1991; Horrocks et al., 2005; also see Chapter 3) and only appears after ~7 cal kyr BP, when local summer insolation and seasonality have

increased. The presence of kauri during the supposed interstadial MIS 5a is inconsistent with this observation. Possibly, stadials 5b and 5d were not extreme enough (compared to the glacial conditions of MIS 2) to eliminate kauri from the Auckland area after MIS 5e; kauri's presence in MIS 5a may be a relic from MIS 5e or it was able to expand from nearby refugia after stadial MIS 5b.

Finally, a strong argument against MIS 5e representation in the Orakei Basin record relates to sea level. Regional sea level was above modern height during MIS 5e (specifically 128-122 kyr BP; Williams et al., 2015), so would likely have overtopped Orakei's crater rim, as happened during the early Holocene. The absence of marine deposits at the base of the core indicates that the Orakei crater may be several kyr younger than was inferred from the depth of its tephra deposit in the Glover Park core (126.2 ± 3.3 kyr BP; Hopkins et al., 2017).

In conclusion, it is unlikely that MIS 5e is represented in the Orakei Basin sediments, because the vegetation patterns and temperature reconstructions are inconsistent with a warm interglacial and because high sea level at the time (Williams et al., 2015) would have overtopped the crater rim. The stable conditions and MAAT plateau during Zone 1 seem to indicate that only MIS 5a is represented. However, we cannot rule out that stadials MIS 5d and 5b are covered by the sediment core but were too mild or brief to show up in the pollen diagram. Ultimately, the anticipated improved age-depth model (see section 5.3.3) will provide further insight into these correlations. In addition, a chironomid record for this part of the core might indicate whether stadial conditions occurred, as chironomids are more sensitive to small changes in temperature and respond faster than vegetation (Ruiz et al., 2006).

5.5.2.3 The "no-pollen zone"

Just above the cf. Okaia tephra (~28.6 cal kyr BP, Table 5.2) sits a 3-m thick zone (from ~31 to 28 m depth) that contains no pollen. Four subsamples from this zone were prepared for pollen analysis twice: once using 1 cm^3 of sediment and once using 2.5 cm^3. Although we recovered the *Lycopodium* spores that were added as a positive control, we did not find any pollen grains and the samples are conspicuously devoid of organic remains. The retention of *Lycopodium* spores rules out an aberration in the pollen preparation procedure as the cause of pollen absence. There is one sample between the cf. Okaia and this "no-pollen

zone" at 31.0 m depth that does contain pollen (concentration: ~9300 grains/cm^3). Pollen concentrations remain low throughout Zone 4/MIS 2 (around 10,000 grains/cm^3). The low pollen concentrations during MIS 2 are likely due to a combination of increased sedimentation rate (Fig. 5.3) and glacial conditions unfavourable to pollen production. However, these two factors on their own cannot explain the complete absence of pollen between 31 and 28 m.

Here we consider alternative explanations for the absence of pollen in this part of the record: first, the three meters of sediments could have been deposited in a very short time, comprising one "event", such as a storm event or slump; second, the three meters represent a period of time when there was no vegetation present in the area (e.g. due to an environmental catastrophe such as a volcanic event), or the vegetation that was there did not produce pollen; third, pollen entered the lake but was not preserved due to unusual depositional conditions.

The first possibility is unlikely, because of the nature of the sediments in this no-pollen zone (Table 5.1). Just above the cf. Okaia tephra, between 31.4 and 29.3 m depth, the stratigraphy consists of a finely laminated reddish-brown clay (lithology 7), and from 29.3 to 27.9 m a massive, light-brown clay (lithology 6). Although the laminations visually appear to be different from the laminations in other parts of the core, and their nature is unclear, it is unlikely that they were deposited in one event. There is no obvious mechanism that could explain the deposition of hundreds of distinct sedimentary layers in a short period of time, although a study is underway to identify the source/cause of these laminations that could resolve this mystery.

It is notable that the sediments in this section contain a lot of basaltic tephra (Table 5.1). The AVF was more active during some periods than others, although the most significant "flare-up" time occurred between 34 and 30 kyr BP (Molloy et al., 2009; Leonard et al., 2017), a little earlier than the no-pollen zone. If volcanic activity was heightened during the no-pollen zone, vegetation may have been disturbed, resulting in low pollen production in conjunction with the harsh climatic conditions at the time. However, the AVF eruptions are not expected to be substantial enough to eliminate vegetation for a prolonged period of time, and other AVF pollen records show no evidence of catastrophic volcanic disturbance of the

vegetation (e.g. Shane & Sandiford, 2003; Newnham et al., 2007b). At Kohuora (Fig. 5.1b), for instance, pollen concentrations are intermediate to high (> 100k grains/cm^3) at the time.

The third possible explanation for the lack of pollen could have been the combination of low pollen concentrations caused by above-mentioned effects, and poor preservative conditions. Destruction of pollen grains can be mechanical, although the lack of abundant sand in this section makes that option unlikely. Oxidation or microbial attack (or a combination of both) are more probable culprits. Autoxidation may occur when pollen is exposed to air (Havinga, 1967). Lowered lake levels and/or increased turbulence (e.g. due to increased windiness) could have oxygenated the water, which would both increase the rate of autoxidation and the ability of microorganisms to be active in the benthic zone. Preliminary x-ray fluorescence data from Orakei Basin, and the ratio of Fe to Mn specifically, is not inconsistent with this interpretation (L. Peti, pers. comm.). A drier and windier climate is also consistent with other records of MIS 2 in the region (e.g. Sandiford et al., 2002; Augustinus et al., 2011; 2012).

In summary, the cause of the absence of pollen between 31 and 28 m depth is likely due to a combination of effects: increased sedimentation rate lowered the pollen concentration (Fig. 5.3), the sudden onset of harsh climatic conditions and heightened local volcanic activity disturbed the vegetation, and oxidative conditions in the lake caused a suboptimal preservative environment. The latter might be related to climatic conditions as well: MIS 2 is associated with decreased humidity and increased windiness, lowering the lake level and increasing turbulence and water column mixing.

5.5.3 The local environment

In pollen assemblages, (tree) fern spores and wetland taxa are typically not included in the dryland-pollen sum, because they are thought to be representative of the local flora rather than the wider regional vegetation community (Wilmshurst et al., 2007). In the Orakei Basin record, Cyperaceae and Restionaceae are the most abundant wetland taxa. Restionaceae (probably mostly or exclusively *Empodisma* type) is only abundant during cold periods, i.e. MIS 4 and 2. The expansion of swamp/wetlands during MIS 2 was found in other AVF records too (Sandiford et al., 2003; Augustinus et al., 2011) and was previously explained by the lowering of the lake level and expansion of lake margin areas. However, it is unlikely

that a lowering of the lake level would increase the available habitat for wetland taxa at Orakei Basin due to the steepness of the crater wall and surrounding hilly terrain. We suggest that it is more likely that lowering sea levels during glacials and stadials (Clark and Mix, 2002) allowed the expansion of wetland habitats over the shelf that is now Hobson Bay, on the shoreward edge of Orakei Basin (Fig. 5.1c). Cyperaceae is also more abundant during colder periods, especially MIS 2, but is additionally abundant when the lake transitions to swamp (lithology 1c and a) and was likely part of the local swamp vegetation at this time.

Fern spores are unusually abundant in the Orakei Basin record, especially *Cyathea smithii* type, which reaches abundances of up to 438%. The fern spore abundances are partially correlated to each other, but not throughout and it is difficult to discern patterns. Tree ferns *Cyathea dealbata* and *Dicksonia* appear to be most abundant in sections of the core that are not laminated, i.e. lithology 11 and 9 (banded silt and sand to massive sand), lithology 3 (banded silt and sand; *C. dealbata* only) and lithology 1 (mostly peat). Tree fern taxon *C. smithii* type is most abundant during MIS 3 (lithology 9 and 8), but highly variable. Monolete fern spores show peaks mostly during MIS 4 and 2 but are also abundant during most of MIS 5a. Tree ferns, especially *C. dealbata*, are part of the podocarp-broadleaf forest. Monolete spores are produced by a variety of fern species, which include taxa that grow in more open vegetation communities (consistent with their abundance in MIS 4 and 2) and taxa that grow in the understory of forests (consistent with their abundance in MIS 5a). In other AVF records (Sandiford et al., 2002, 2003; Newnham et al., 2007b; Augustinus et al., 2011, 2012), tree ferns (mostly *C. dealbata*) are abundant during warm phases and monolete spores during cool phases, but spore percentages do not fluctuate as much as in the Orakei record. We propose that fern spore peaks in the Orakei sediments represent periods of increased erosion, wash-in or stream input (from e.g. the Purewa Creek; Fig. 5.1c), exaggerating the local fern presence and vegetation that grows along streams. This input would have been amplified even more by the hilly terrain around Orakei Basin. Many of the spore peaks overlap with parts of the lithology with visibly increased disturbance and sand layer occurrence and peaks that do not overlap with visible sandy intervals may still be the result of increased stream inflow; increased coarse grained input might not always be visible with the naked eye. Grain-size analysis should confirm or reject this hypothesis.

5.5.4 Orakei Basin in a regional context

5.5.4.1 Northern New Zealand

Newnham et al. (2013) describe the contrast between glacial and interglacial vegetation communities over New Zealand. The Northland region (extending to the north of Auckland) is characterised by continuous forest cover, with the main variation being a larger beech forest component relative to conifer-hardwood forest during the LGM compared to the Holocene. South of Auckland, vegetation is more open during glacials, i.e. shrubland with scattered stands of beech forest and more dominant shrub and Poaceae components progressively southwards. Thus, Auckland appears to be near the boundary of continuous forest cover during the LGM and past changes in climate are likely to be registered as changes in the relative abundance of tree pollen in our record. This is confirmed by our pollen record, where tall podocarps and beech trees remain present at all times, though at lower abundances in some zones with an expansion of shrubs and herbs, which presumably represent glacial or stadial conditions.

In section 5.5.2.2 we discussed MIS 5 in northern New Zealand to determine which substages are represented in the Orakei Basin record. According to our current interpretation, the Orakei Basin record does not extend to the Last Interglacial, instead registering interstadial MIS 5a, traditionally named the Otamangakau interstadial (McGlone & Topping, 1983; Newnham, 1999), at its base.

MIS 4 appears in the Orakei Basin record as a brief cold period, with MAAT on average 5°C cooler than the preceding zone, but not as cold as MIS 2 (Fig. 5.6a). The general trend of cool vs. cold temperatures during MIS 4 and MIS 2 respectively is consistent with the few records from northern New Zealand that cover the last ~70 kyr, which seem to indicate that any vegetation changes during MIS 4 were not as extreme as during MIS 2 (Shane and Sandiford, 2003; Elliot et al., 2005; Newnham et al., 2017a). The pollen record from Onepoto (AVF, Fig. 5.1b; Shane and Sandiford, 2003) resembles the Orakei Basin record, with the expansion of *Coprosma* and *Halocarpus* indicating an opening of the forest canopy during MIS 4, but no other major changes in the forest composition. There is, for instance, no great expansion of beech forest or *Lophozonia menziesii* (silver beech, a species with a southern distribution in the present day), such as occurred during MIS 2. The Northland Pouto record

(~78 km NW of Auckland, Fig. 5.1a; Newnham et al., 2017a) similarly displays low reconstructed temperatures for MIS 4 (Fig. 5.6b), but no beech expansion.

The broad similarity of the Orakei Basin pollen record to other AVF and Northland records continues throughout MIS 3 and 2: MIS 3 displays interstadial conditions and MIS 2 displays glacial conditions (see also Lorrey & Bostock, 2017). One notable difference is that the PMI indicates generally dry conditions during MIS 3 at Orakei Basin, whereas Newnham (1999) describes this period as a very wet interstadial. The PMI is likely to be inaccurate during MIS 3, because the total abundance of tall tree podocarps is low. In addition, the relative abundance of tall podocarps may be more strongly influenced by temperature than precipitation during an interstadial (see section 5.5.1.2).

The MAAT record shows an earlier warm phase during MIS 3 (before the Rotoehu tephra deposit), followed by a progressive cooling starting at ~50 m depth and culminating in MIS 2. This pattern is similar to the Kai Iwi and Pouto MAAT records (Fig. 5.1a, Fig. 5.6b) and offshore records east of Northland (S804, S794 and S803, Fig. 5.1a; Wright et al., 1995). Wright et al. (1995) designate the earlier warm phase as substage 3b and the cooling phase as substage 3a. The progressive cooling trend of substage 3a correlates to decreasing local mean and summer insolation intensity and decreasing insolation seasonality (Fig. 5.6e). At Orakei Basin, the boundary between substage 3b and 3a coincides with the appearance of beech taxa *Fuscospora* and *Lophozonia menziesii*. Beech continues to be abundant until the start of MIS 1. Interestingly, all Northland records show a beech (*Fuscospora*) expansion during the Last Glacial, but the timing and duration of the beech phase varies between records. The Pukaki record shows the same pattern as Orakei, with beech appearing before the deposition of the Rotoehu tephra (Shane and Sandford, 2003). However, at Kai Iwi beech expansion starts instead shortly after the deposition of the Rotoehu tephra (at ~40 kyr BP), and at Pouto at ~27 kyr BP. At Lake Tangonge (northern Northland; Fig. 5.1a), *Fuscospora* percentages increase at 30 kyr BP (Elliot et al., 2005). Contrary to Newnham et al.'s (2017a) conclusions, there appears to be a time-transgressive pattern in Northland, with some exceptions (e.g. Pouto): the further north, the later the expansion of beech. This pattern is consistent with a progressive cooling trend, whereby the minimum absolute temperature suitable to beech is achieved progressively later northwards. The exceptions to this spatial pattern, like Pouto, may be due to local site factors or microclimate (Newnham et al., 2017a).

Kauri is another taxon that shows variation between sites within northern New Zealand. At Orakei and Onepoto, kauri is only present at very low abundances throughout MIS 3 (Shane and Sandiford, 2003) and almost absent during MIS 2 (Augustinus et al., 2011). At Pouto, kauri is virtually absent throughout the entire studied period, whereas at Kai Iwi kauri is present during MIS 3 and 2, but most dominant after ~15 kyr BP (Newnham et al., 2017a). In the far north kauri is present throughout MIS 4 to 1, with high abundances during MIS 3 (Elliot et al., 2005). The pollen records from offshore eastern Northland (Wright et al., 1995) show kauri is present continuously during MIS 3 to 1, while beech percentages are relatively low during the first half of MIS 3 but increasing. As Newnham et al. (2017a) explain, the offshore records are assumed to integrate pollen from the wider Northland region (with a bias towards eastern areas). It seems therefore likely that kauri was present in northern New Zealand throughout MIS 3 and 2, but with variable dominance in different regions and no clear spatial pattern.

Newnham (1999) suggested that the relative abundances of *Fuscospora* (representative of *Fuscospora truncata*, hard beech) and kauri might be indicative of stadial/glacial vs. interstadial/interglacial conditions respectively. This new record from Orakei Basin adds to the evidence that *Fuscospora* is indicative of cool conditions. However, the presence and absence of beech and kauri must be, at least in part, influenced by local site conditions as well. The imperfect spatial patterns for the beech expansion and lack of a pattern for kauri presence in northern New Zealand indicates that MAAT cannot be the only driving factor. Complicating the interpretation is the fact that kauri pollen is generally underrepresented when not locally present, whereas beech pollen has strong dispersal capabilities and can thus be overrepresented when present in the wider region (Newnham et al., 2017a). At Orakei, kauri was probably not present locally after MIS 5a due to temperature depression, whereas at Lake Tangonge (Elliot et al., 2005) kauri probably persisted due to generally higher temperatures than in other parts of the country.

Glacial MIS 2 conditions commence in the Orakei Basin record at ~31.5 m depth, just above the cf. Okaia tephra (28.6 cal kyr BP), indicating an early onset (see section 5.5.4.2), similar to nearby Kohuora (Newnham et al., 2007b) and Onepoto (Augustinus et al., 2011). The low PMI is consistent with charcoal records from Pukaki (Fig. 5.1b; Sandiford et al., 2003) and Onepoto; more microcharcoal during MIS 2 indicates a drier and windier climate. The

average MAAT reconstructed for MIS 2 at Orakei Basin (i.e. ~9.5°C cooler than present) is within the range of previously inferred MAAT during the LGM in northern New Zealand, although at the extreme end (Newnham et al., 2013). MIS 2 appears homogenous and consistently cold at Orakei Basin, whereas the other AVF records, specifically Kohuora, Pukaki and Mt Richmond (Fig. 5.1b; Sandiford et al., 2002), show an amelioration between 26 and 24 kyr BP, around the KOT (Lorrey & Bostock, 2017). Possibly, the resolution of the Orakei Basin pollen record is not high enough to pick up this warming interval. The extreme and homogenous temperature depression in combination with the mysterious "no-pollen zone" suggest that further investigation is warranted.

The Orakei Basin pollen record displays a very typical North Island signal for the LGIT: first Poaceae, then shrubs and beech make way for tall conifers. *Prumnopitys taxifolia* dominates early on, while *Dacrydium cupressinum* continues to rise. Near the top of the sequence, just before sea level rise causes the crater to be flooded by the ocean, *Metrosideros* appears, followed by *Ascarina lucida*, indicating interglacial conditions. The same or similar patterns are evident in other AVF records, i.e. Pukaki (Sandiford et al., 2003), Onepoto (Augustinus et al., 2011) and Kohuora (Newnham et al., 2007b).

5.5.4.2 Comparison to records beyond northern New Zealand

The last 30 kyr BP have been studied extensively in New Zealand and have been summarised into the New Zealand climate stratigraphy by Alloway et al. (2007) and Barrell et al. (2013). For a comparison of the older parts of the Orakei record, we turn to the pollen record of Okarito Pakihi bog in south Westland (Fig. 5.1a; Newnham et al., 2007a), which is the only continuous terrestrial pollen record for the Last Interglacial-Glacial cycle in New Zealand. There are offshore pollen records that cover the Last Interglacial-Glacial cycle, but although they are often better dated than terrestrial records (Ryan et al., 2012), their resolution is generally low, and their interpretation suffers from myriad problems caused by depositional processes (McGlone, 2001). We also refer to Williams et al. (2015), who reviewed pollen records, speleothem records and moraine ages from the Last Interglacial to the LGM.

The last glaciation at Okarito (MIS 4, 3 and 2) is characterised by subalpine and grassland taxa with variable relative abundances (Newnham et al., 2007a). Like at Orakei Basin, MIS

4 is a cold phase but not as severe as the LGM. Nevertheless, a relatively mild MIS 4 indicated by both these northern and southern pollen records is inconsistent with glacial geomorphological records from the Southern Alps (Fig. 5.1a), which show MIS 4 ice limits lie beyond MIS 2 ice limits (Suggate, 1990; Williams et al., 2015). The composite $\delta^{18}O$ series from speleothem records of the western North Island and South Island of New Zealand indicate MIS 4 was cold and wet (Williams et al., 2015), but it is hard to infer the amount of temperature change, as the relationship between $\delta^{18}O$ and temperature is not linear. The $\delta^{18}O$ curve of marine core MD06-2991 (off the west coast of the South Island; Fig. 5.1a) shows an excursion with values close to those achieved during the LGM, similar to the $\delta^{18}O$ curve of marine core MD03-2607 off the south coast of Australia (De Deckker et al., 2019), indicating sea surface temperatures were much reduced. On the other hand, the marine $\delta^{18}O$ LR04 stack (Lisiecki & Raymo, 2005) does not show such dramatic change at the time (Fig. 5.6c), nor do Antarctic ice core records (Fig. 5.6d). The variable response of the described proxies to MIS 4 cooling may be due to the seasonal distribution of energy. Local summer insolation intensity is low during MIS 4 (especially between ~65 and 57 kyr BP), whereas during the LGM, summer insolation intensity is high (Fig. 5.6e). Sea surface temperature records suffer from seasonal biases that relate to the production season of the proxy carrier (Liu et al., 2014) that are hard to pick apart. Vegetation is thought to be mostly responsive to MAAT and minimum temperatures (prevalence of frosts) (Wilmshurst et al., 2007). Glaciers may be more sensitive to temperature during the ablation season (November to March) or westerly wind strength (Hooker & Fitzharris, 1999; Putnam et al., 2012), explaining the discrepancy between pollen and glacial records in New Zealand. Loess records from the eastern South Island show accumulation around 65 – 60 kyr BP, indicating a dry and windy climate (Berger et al., 2002), while speleothem records in the west show $\delta^{13}C$ excursions indicative of increased precipitation (Whittaker et al., 2011; Williams et al., 2015). These observations are consistent with increased westerly wind strength during MIS 4.

MIS 3 and 2 appear more complex at Okarito than at Orakei, with five separated intervals of grassland expansion (cooling), of which two occur during MIS 3 and the three strongest cooling events occur during MIS 2 (Newnham et al., 2007a). As explained by Newnham et al. (2007a), the two moderate cooling events during MIS 3 have no equivalents in the glacial geomorphologic record (Suggate, 1990), but are consistent with the Westland loess record (Almond et al., 2001) and glacial and speleothem records from Fiordland (Fig. 5.1a;

Williams, 1996). The two biggest cooling intervals during late MIS 3 (30 – 27 kyr BP) and MIS 2 (21 – 18 kyr BP) do concur with ice advances (i.e. Larrikins and Moana Glaciations respectively) and a smaller intervening cool period (25 – 23 kyr BP) may represent the second phase of Larrikins ice advance (Suggate, 1990). See also the Lorrey and Bostock (2017) synthesis, which summarises the most recent evidence for the timing of ice advances during this period. Possibly, variations in climate superimposed on the Last Glacial cycle are disconnected between North and South, or the Orakei pollen record resolution is not sufficient to distinguish the separated cooling phases found at Okarito. The latter may be more likely, because (as mentioned in the previous section) other AVF records also show more variation within MIS 2 than the Orakei Basin record does. Even as far north as Queensland, Australia, two cold events during MIS 2 have been identified (Petherick et al., 2008). Three cooling events are also evident in the New Zealand climate stratigraphy (i.e. NZce-10, -8 and -6; Barrell et al., 2013). The climate amelioration identified in other AVF records to occur between 26 and 24 kyr BP (around the deposition of the KOT) is consistent with the mid-LGM warming phase identified at Okarito. Regarding the cool phases at Okarito that are not recognised in any AVF records, the proximity of Okarito to alpine environments and the Southern Ocean make that site potentially more sensitive to showing cold phases than northern New Zealand.

Despite the above-mentioned differences, the early onset of glacial conditions after MIS 3 at Orakei Basin is consistent with the first major grassland expansion at Okarito. Both concur with the Last Glacial coldest period from 28 to 18 kyr BP (Alloway et al., 2007). This early onset is evident in speleothem and glacial geomorphic records, terrestrial and marine pollen records, marine isotopic and carbonate records, sea surface temperature records from the southern mid latitudes, and Antarctic ice core records (Fig. 5.6d) as summarised by Vandergoes et al. (2005), Alloway et al. (2007) and Williams et al. (2015). In south-east Australia (e.g. Redhead Lagoon, New South Wales) extremely dry conditions occur between 30 and 20 cal kyr BP (Williams et al., 2006).

Finally, we find no evidence for a Late-Glacial Reversal at Orakei Basin. The latter is not unexpected, as previous studies have shown that the reversal is only weakly registered, if at all, at more northern sites of New Zealand (Newnham et al., 2012).

5.5.4.3 Forcing factors

The broad similarity of the Orakei Basin pollen and MAAT records to the main zones of the Okarito pollen record (Newnham et al., 2007a), MIS record (Lisiecki & Raymo, 2005) and EDML oxygen isotope record (Bazin et al., 2013) suggests a common driver: orbital forcing at high northern latitudes (Vandergoes et al., 2005). Summer insolation at 65°N correlates with the temperature variations reconstructed for the Last Interglacial-Glacial cycle (Fig. 5.6), in agreement with the Milankovitch orbital forcing hypothesis (Imbrie et al., 1992; 1993). A closely coupled ocean-atmosphere-terrestrial system in the Pacific at this time (Vandergoes et al., 2005) would have linked Northern Hemisphere insolation to vegetation change in New Zealand.

Nevertheless, local signals appear to be imprinted on the long-term, North Atlantic-driven trend. The early onset of the LGM at Orakei Basin and other southern mid-latitude sites is consistent with, and possibly forced by, a regional insolation minimum between 35 and 30 kyr BP, as suggested by Vandergoes et al. (2005). Local to regional insolation may also have played a role in the differential response of vegetation vs. glaciers to MIS 4 cooling, as explained in the previous section (5.5.4.2). Whether local insolation affected vegetation at Orakei Basin directly, through changing air temperatures, or indirectly, through changing westerly wind strength, remains unclear.

5.6: Conclusions and future work

In summary, the Orakei Basin pollen record paints a picture of the last ~85 kyr in northern New Zealand with warm MIS 5 interstadials, a brief MIS 4, a warm early MIS 3 with a cooler beech phase commencing just before the deposition of the Rotoehu tephra (45.2 cal kyr BP), an early onset MIS 2 displaying the lowest temperatures of the record, and a typical LGIT for northern New Zealand. These broad scale patterns are generally consistent with previous records from northern New Zealand, as discussed, but these other records are fragmentary and shorter and are now able to be viewed in a more continuous framework for the last ~85 kyr BP.

At present, we can draw several general conclusions:

1. Forest taxa persisted throughout the last 85 kyr at Orakei Basin, although grassland indicators increase during the LGM; the pollen record is consistent with more open or patchy forest vegetation than in Northland, but a more dominant forest community than in the central North Island, confirming the hypothesis that Auckland lies on or near the boundary of continuous forest cover during the LGM (Newnham et al., 2013).

2. The pre-Rotoehu beech expansion at Orakei Basin appears to be part of a time-transgressive pattern in Northland, consistent with a progressive cooling trend. There are several exceptions to the rule (e.g. Pouto), indicating that local site factors or microclimates also play a role. Newnham's (1999) hypothesis, stating that kauri- and hard beech-dominated communities are prominent during interstadials and stadials respectively, is not inconsistent with these findings.

3. MIS 4 appears to be a brief cold period at Orakei Basin and not as cold as MIS 2. This is consistent with other pollen records from New Zealand (Elliot et al., 2005; Newnham et al., 2007a; Newnham et al., 2017a), but inconsistent with glacial records from the Southern Alps (Suggate, 1990; Williams et al., 2015).

4. The broad patterns of vegetation change at Orakei Basin were likely driven by northern high-latitude summer insolation, consistent with the Milankovitch orbital forcing hypothesis (Imbrie et al., 1992; 1993), although the early-onset LGM, consistent with other southern mid-latitude records, indicates a role for regional summer insolation as well (Vandergoes et al., 2005).

There remain some distinctions between the Orakei Basin record and other records in the region, such as the apparent simplicity of MIS 2. Some of these differences may be attributable to local site factors and others to differences between chronological control or temporal resolution. We look forward to greater clarity when the age-depth model for Orakei Basin is finalised.

Future work on the Orakei Basin record will comprise the refinement of the timing of vegetation changes and inferred climate changes, following the planned finalisation of the age-depth model (see section 5.3.3). Other promising lines of research to pursue include:

1. Increasing the resolution of the Orakei Basin pollen record. This can address many uncertainties in past climate reconstructions, but especially the section around the cf. KOT could be of interest in the context of a possible mid-MIS 2 climate amelioration as identified in other records (e.g. Newnham et al., 2007b).

2. Constructing a chironomid record for Orakei Basin. It is preferable if a quicker, more time-efficient extraction method is developed first (see Appendix E). A summer temperature reconstruction based on chironomids could shed light on the seasonal aspect of climate change, for instance during MIS 4. We showed in Chapter 3 that chironomid-inferred temperatures closely match local summer insolation parameters during the last 14 cal kyr BP.

Chapter 6

Synthesis

6.1: Introduction

The overarching goal of this thesis was to elucidate patterns of climate change in northern New Zealand during the late Quaternary in relation to globally recognised patterns of climate change. This was achieved through the development of pollen and chironomid records from the Holocene-age sediments of Lake Pupuke, and a pollen record covering ~85 kyr BP to 9 kyr BP from Orakei Basin. Quantitative climate reconstructions derived from these records were compared to other records from the region, from southern New Zealand and to global climate reconstructions. In addition, two limitations in the pollen and chironomid methodologies were addressed. Here, I summarise the findings of this thesis in the context of the five objectives formulated in Chapter 1. I also discuss some recommendations for future research that follow from the results of this thesis.

6.2: Summary of findings

During the course of the development of the Lake Pupuke pollen record (Chapter 3), I found inconsistencies in relative pollen abundances related to the pollen preparation method, specifically density separation. This prompted me to test the difference in outcome between using heavy liquid with a specific gravity (SG) of 2.0 (within the recommended range from literature: Munsterman & Kerstholt, 1996; Zabenskie, 2006; Campbell et al., 2016) and a SG of 2.2 (density used at GNS Science, recommended by palynologist Dr Xun Li, pers. comm.), the results of which are presented in Chapter 2. Surprisingly, we found that small, compact pollen grains such as *Libocedrus* and *Metrosideros* are underrepresented when a low SG is used, while saccate grains such as *Prumnopitys taxifolia* and *Dacrydium cupressinum* are overrepresented, resulting in erroneous relative abundances. Small grains may be trapped under a layer of non-floating material during density separation with a low SG, or are weighted down by particles stuck to them, whereas saccate grains are probably buoyant enough to overcome the extra mass. The downside of using a higher SG is that a lot of extraneous material is incorporated into the pollen slides, making them more difficult to count. We believe that the issue with low SG has largely gone unnoticed and we urge palynologists to test what SG value is appropriate for their sedimentary material. By determining that an SG of 2.2 is required to get a representative pollen sample from the Lake Pupuke sediments, we fulfilled thesis objective 4: "Review the density separation

procedure in pollen slide preparation and find the minimum required specific gravity that results in representative samples".

Chapter 3 describes the results of the Lake Pupuke pollen and chironomid analyses, which satisfies objective 1a. The mean annual air temperature (MAAT) reconstruction for the last 16 cal kyr BP and summer air temperature (SmT) reconstruction for the last 14 cal kyr BP from pollen and chironomid results respectively, fulfil objective 2a. MAAT was found to be stable throughout the Holocene, while the vegetation community changes through time. Vegetation appears to be more strongly influenced by seasonality and effective precipitation during the Holocene than by MAAT. SmT increases between 10 and 3 cal kyr BP, showing a strong relationship with integrated local summer insolation. Comparison to other vegetation and climate records from New Zealand (objective 3a) led to the conclusion, first, that seasonality was low during the Early Holocene (12 to 9.3 cal kyr BP), likely driven by low local summer insolation intensity. Second, mid-to-late-Holocene summers were hot and dry, allowing the tall conifer kauri (*Agathis australis*) to expand throughout northern New Zealand. There is no evidence for the Holocene Thermal Maximum (HTM) as observed in global climate reconstructions to occur between 11 and 5 kyr BP (Fisher et al., 2018). Rather, Holocene temperatures at Lake Pupuke are apparently driven by local to regional factors. The HTM was identified as an event of special interest in Chapter 1, because it represents a potential analogue for future climate change. My results indicate that the HTM may not be informative in this context for the future of northern New Zealand's climate and environment.

The Lake Pupuke chironomid-SmT reconstruction from Chapter 3 highlighted an issue with the transfer function: the model, based on a 60-lake training set (Dieffenbacher-Krall et al., 2007), was not able to reconstruct values close to modern day SmT at Lake Pupuke (which is 18.9°C). Therefore, I explored an extended training set of 104 lakes, which includes lakes from northern New Zealand. The results of this exploratory analysis and the formulation of new transfer functions are described in Chapter 4 and fulfils objective 5: "Develop an improved chironomid-SmT model based on an extension of the modern training set, using both traditional and novel methods". New models were fitted using both traditional methods, such as the modern analogue technique and (weighted averaging) partial-least-squares (PLS) regression, and modern machine learning methods. Weighted-averaging PLS regression was found to perform best, being the only model to accurately predict SmT above

16°C. The disappointing performance of machine learning methods is probably due to the relatively low number of lakes in the training set, especially at the high end of the SmT gradient. The new model was used to improve the SmT reconstruction from the Lake Pupuke chironomid record. The SmT trend reconstructed by the new transfer function is the same as the original trend, leaving the conclusions drawn in Chapter 3 unchallenged. However, the new transfer function is able to reconstruct higher temperatures, including modern day SmT at Lake Pupuke: a vast improvement.

Chapter 5 describes the Orakei Basin pollen record, relating to objective 1b. The core was expected to cover the period from 126 kyr BP (estimated age of the eruption; Hopkins et al., 2017) to 9 kyr BP (when the crater wall was breached by the ocean; Hayward et al., 2008). As mentioned, chironomid analysis was unsuccessful on the sediments of Orakei, because head capsule numbers were low and consequently, I could not construct an informative record in a timely manner (also see Appendix E). I found a great diversity of non-pollen palynomorphs on the pollen slides, which are described in Appendix D. MAAT was reconstructed from the pollen record, in part fulfilment of objective 2b. The Orakei pollen record was compared to other, more fragmented records from northern New Zealand, tying them together, as well as to other New Zealand records and the Marine Isotope Stages (MIS) in fulfilment of objective 3b. The pollen record shows a five-fold division that likely corresponds to MIS 5 to 1 (Lisiecki & Raymo, 2005), corroborated by the preliminary tephrochronology, indicating MAAT exerts strong influence on vegetation throughout the Last Interglacial-Glacial cycle. The base of the core probably represents interstadial MIS 5a, indicating that the record does not reach beyond ~85 kyr BP. The broad similarity of the Orakei Basin pollen record and MAAT reconstruction to the MIS and the Okarito pollen record in south Westland (Newnham et al., 2007a) implies all were driven by northern high-latitude summer insolation, consistent with the Milankovitch orbital forcing hypothesis (Imbrie et al., 1992; 1993). Several patterns in the Orakei record superimposed on the general trend stand out: first, MIS 4 is brief and not as cold as MIS 2. While consistent with other pollen records, glaciers were more extensive in the Southern Alps during MIS 4 compared to MIS 2 (Suggate, 1990), possibly indicating a role for seasonality and/or westerly wind strength. Second, MIS 3 displays an earlier warm phase followed by a progressive cooling trend concurrent with the expansion of beech. This pattern appears to be time-transgressive in northern New Zealand records (e.g. Wright et al., 1995; Newnham et al., 2017a) and might be correlated to decreasing local summer insolation intensity over

this period (Berger & Loutre, 1991). MIS 2 commences just after the deposition of the cf. Okaia tephra (28.6 cal kyr BP; Lowe et al., 2013), consistent with the early onset of the LGM in the southern mid latitudes, which was likely driven by regional insolation intensity (Vandergoes et al., 2005).

In conclusion, this thesis has pointed out the importance of considering regional-to-local drivers of climate changes superimposed on global reconstructions. Whereas the MAAT reconstructions of both Orakei and Pupuke show that broad patterns of change during the last ~85 kyr BP follow northern high-latitude insolation intensity, summer temperature, effective precipitation and westerly wind strength, all of which are likely related to local-to-regional insolation parameters, are important factors impacting the regional environment. It follows that multi-proxy records can help disentangle the different aspects of the climate system, where especially chironomids can be helpful to elucidate the role of SmT and local summer insolation. Finally, this thesis has shown the importance of questioning the appropriateness of conventional methodologies and where possible, addressing their limitations.

As explained, during the course of this thesis a number of unexpected discoveries were made. Some of these discoveries posed significant challenges in relation to the original objectives of this thesis. Two of the challenges were addressed, as reported in Chapters 2 and 4 (relating to objectives 4 and 5 respectively). Other discoveries, whether in the form of methodological challenges or representing opportunities for fruitful further work are discussed in the following section.

6.3: Directives for further research

As explained in Chapter 1, this study is part of a broader, Marsden-funded project about the Auckland maars. The Pupuke and Orakei records are part of the third step of this greater project: producing multi-proxy records of past climate. When the age-depth model for the Orakei Basin core is finalised (part of step 2: dating the records) the Orakei Basin pollen record, described in Chapter 5, will be submitted for publication. Other than this obvious next step in the Auckland maars project, there are three other avenues of research that I think future studies should focus on: to incorporate more lakes with a modern SmT value above 16°C into the chironomid training set, to develop a chironomid record for the last ~85 kyr

from Orakei Basin sediments, and to develop a non-pollen palynomorph database for New Zealand. All three of these objectives can contribute to the original aim of this thesis, i.e. to elucidate patterns of climate change in northern New Zealand during the late Quaternary in relation to globally recognised climate change events.

6.3.1 Incorporate more lakes with a modern SmT value above 16°C into the chironomid training set

Although the extended chironomid training set has improved the accuracy of temperature reconstructions, lakes with a modern SmT above 16°C are still strongly underrepresented (see Chapter 4). This causes a high bias for reconstructions at the high end of the gradient. It is also possible that the chironomid taxa with high beta coefficients (e.g. *Parachironomus*, *Corynoneura*, *Xenochironomus*, *Cricotopus aucklandensis*) are not well sampled.

The main challenge for extending the chironomid training set is to find lakes with conditions as near to natural as possible. Many lakes in the North Island are eutrophic due to human activity, affecting within-lake food webs and thus natural chironomid assemblages. Fortunately, the recently begun "Lakes380" project provides the perfect opportunity to acquire lake sediment surface samples from a large number of clean lakes. Lakes380 is a project aimed at characterising the health and environmental history of New Zealand's lakes by analysing 380 lakes throughout the country (see: https://lakes380.com/). The project involves taking lake core-top or sediment surface samples along with environmental data across the full climatic gradient of New Zealand, which are ideal for extending the chironomid training set.

Another potential improvement of the chironomid-SmT model lies in the enhancement of our ability to identify different taxa. As shown in Chapter 4, many of the currently identified chironomid taxa show a bimodal response to SmT. This is an indication that two or more species with different environmental requirements are amalgamated into one taxon. Woodward and Shulmeister (2007) were able to distinguish a *Chironomus* type (*Chironomus* sp. a) that appeared distinct from other *Chironomus* specimens by the number of striations on the ventromental plate. However, this particular distinction is probably not useful for the training set described in this thesis, as one of the *Chironomus* morphotypes is associated with eutrophic lakes, which were actively avoided. Moreover, the preservative quality and

size of head capsules are often not high enough to be able to count the ventromental plate striations. In the preliminary chironomid samples from Orakei Basin, we found a potential distinct morphotype of *Polypedilum* with a more curved and chitinised mentum than the original *Polypedilum* taxon. However, as the ratio of "normal" to "abnormal" *Polypedilum* head capsules was constant and because the development of a chironomid record for Orakei Basin was not pursued for this study, this morphotype was not further analysed.

6.3.2 Develop chironomid record for the last ~85 kyr from Orakei Basin sediments

I explain in Appendix E why it was not possible to construct a chironomid record for the Orakei Basin core. In summary: the concentration of chironomid head capsules in the Orakei Basin sediments is low, especially between 60 and 10 m depth. The lower the head capsule concentration, the more sedimentary material is required to reach the minimum head capsule count of 60 (see Chapter 4). If a lot of sedimentary material needs to be sieved and picked through, the amount of time spent per sample increases markedly. However, as also mention in Appendix E, Rolland and Larocque (2007) describe a method for kerosene flotation of head capsules, which could greatly decrease the amount of time spent per sample. This method should be tested on the Orakei Basin sediments before it is applied, as previously it has just been tested for ten Canadian lakes (Rolland & Larocque, 2007). Kerosene floatation was not pursued for this thesis due to time constraints. Two other methodologies were deemed more important (i.e. the pollen preparation method, Chapter 2, and chironomid transfer function, Chapter 4).

If the kerosene flotation method proves fruitful for the Orakei Basin sediments, constructing a chironomid record for the last ~85 kyr BP would be a valuable contribution to the palaeoclimatological research on the Late Quaternary in northern New Zealand. As was shown in Chapter 3 on Lake Pupuke, combined MAAT and SmT reconstructions can provide valuable insight into past seasonality and drivers of change. It will be interesting to investigate whether the relationship between local summer insolation parameters and reconstructed SmT (as found for the last 14 cal kyr BP at Lake Pupuke) holds during the Last Glacial, or even as far back as MIS 5. Specifically, an SmT record can help confirm some of the hypotheses postulated in Chapter 5, such as: were summers relatively cool during MIS 4 and progressively warming during MIS 2? Does the early onset of the LGM in

the southern mid latitudes concur with a regional summer insolation minimum (= cold summers) at this time?

6.3.3 Develop non-pollen palynomorph database

During the development of the Orakei Basin pollen record, I found an abundance of distinct non-pollen palynomorphs (NPP) as described in Appendix D. In comparison to some other regions (e.g. Europe: van Geel et al., 2003; East Africa: Gelorini et al., 2011) NPP analysis is an underdeveloped technique in New Zealand. NPPs, which comprise any remains found in pollen slides other than pollen and plant spores, can offer valuable palaeoecological information. For instance, fungal spores can provide information about the presence of their substrate, e.g. host plant, decaying organic matter, dung, etc. Algal remains can be indicative of local conditions and productivity in lakes. Even remains that have not been traced back to a source organism can provide palaeoecological information if they have been associated with certain conditions or organisms in other records. I found 37 distinct NPP types in the Orakei Basin sediments, some of which are very similar to types found in other parts of the world. Some of the unidentified types seem to be associated with either tall trees, beech forest, or open vegetation.

To make NPP informative in New Zealand records, they simply have to be recorded. As recording NPP types alongside pollen requires little extra work or time, I encourage all palynologists to start adhering to this practice. Some research questions that are currently important and could be addressed with NPP analysis include: where and when did *Phytophtora agathidicida*, the pathogen responsible for kauri dieback disease (Waipara et al., 2013), originate; and what NPP types are associated with human impact in New Zealand and can they be used to trace human migration throughout the country?

Together, these three recommendations for future research can embody the next step into elucidating patterns of climate change in northern New Zealand during the late Quaternary, as our ever-increasing knowledge of past climates continues to lead us to new questions.

6.3.4 Beyond northern New Zealand

This thesis has highlighted the importance of summer insolation intensity as a driver of climate change, whether it be northern high-latitude summer insolation forcing of glacial cycles, or regional-to-local summer insolation driving variations from the general trend. However, the relative importance of local insolation is still unclear and even the importance of northern high-latitude insolation has been contested, because it covaries with Southern Hemisphere insolation parameters (e.g. Huybers & Denton, 2008). Although the Lake Pupuke chironomid-SmT record shows a strong relationship with local summer insolation, increasing the spatial resolution of chironomid-derived SmT reconstructions could show whether this connection is valid in other regions (e.g. southern low or high latitudes) as well.

Regarding analogues for future climate change, this thesis has shown that there are no phases during the last ~85 kyr BP that were warmer than present in northern New Zealand. This raises questions about the usability of global temperature reconstructions. Periods such as the HTM may have been warmer than present in some regions, but not all. Whereas past driving forces, such as orbital forcing through insolation intensity, may have resulted in spatially diverse responses, greenhouse gas emissions are likely to cause warming more globally. To investigate the effect of warming on the local environment, local climate reconstructions are required.

References

Allan, A.S.R. (2008). *An elemental and isotopic investigation of Quaternary silicic Taupo Volcanic Zone tephras from ODP Site 1123: chronostratigraphic and petrogenetic applications.* MSc thesis, Victoria University of Wellington, New Zealand.

Alley, R.B., Marotzke, J., Nordhaus, W.D., Overpeck, J.T., Peteet, D.M., Pielke Jr., R.A., Pierrehumbert, R.T., Rhines, P.B., Stocker, T.F., Talley, L.D., Wallace, J.M. (2003). Abrupt Climate Change. *Science*, 299, 2005-2010.

Alloway, B.V., Lowe, D.J., Barrell, D.J.A., Newnham, R.M., Almond, P.C., Augustinus, P.C., Bertler, N.A.N., Carter, L., Litchfield, N.J., McGlone, M.S., Shulmeister, J., Vandergoes, M.J., Williams, P.W., and NZ-INTIMATE members. (2007). Towards a climate event stratigraphy for New Zealand over the past 30 000 years (NZ-INTIMATE project). *Journal of Quaternary Science*, 22, 9-35.

Almond, P.C., Moar, N.T., Lian, O.B. (2001). Reinterpretation of the glacial chronology of south Westland. *New Zealand Journal of Geology and Geophysics*, 44, 1-15.

Armitage, P.D., Cranston, P.S., Pinder, LC.V. (Eds.) (1995). *The Chironomidae: Biology and ecology of non-biting midges.* Chapman and Hall, London, United Kingdom.

Augustinus, P.C., Reid, M., Andersson, S., Deng, Y., Horrocks, M. (2006). Biological and geochemical record of anthropogenic impacts in recent sediments from Lake Pupuke, Auckland City, New Zealand. *Journal of Paleolimnology*, 35, 789-805.

Augustinus, P.C., D'Costa, D., Deng, Y-B., Shane, P.A., Hagg, J., (2011). A multi-proxy record of changing environments from ca. 30,000 to 9,000 cal yr BP: Onepoto Maar paleolake, Auckland, New Zealand. *Journal of Quaternary Science*, 26, 389-401.

Augustinus, P.C., Cochran, U., Kattel, G., D'Costa, D. (2012). Late Quaternary paleolimnology of Onepoto maar, Auckland, New Zealand: implications for the drivers of regional paleoclimate. *Quaternary International*, 253, 18-31.

Barrell, D.J.A., Almond, P.C., Vandergoes, M.J. (2013). A composite pollen based stratotype for inter-regional evaluation of climatic events in New Zealand over the past 30 000 years. *Quaternary Science Reviews*, 74, 4-20.

Bazin, L., Landais, A., Lemieux-Dudon, B., Toyé Mahamadou Kele, H., Veres, D., Parrenin, F., Martinerie, P., Ritz, C., Capron, E., Lipenkov, V.Y., Loutre, M.-F., Raynaud, D., Vinther, B.M., Svensson, A.M., Rasmussen, S.O., Severi, M., Blunier, T., Leuenberger, M.C., Fischer, H., Masson-Delmotte, V., Chappellaz, J.A., Wolff, E.W. (2013). delta 18O measured on ice core EDML on AICC2012 chronology. *PANGAEA.* doi:10.1594/PANGAEA.824888

Berger, A., Loutre, M.F. (1991). Insolation values for the climate of the last 10 million years. *Quaternary Science Reviews*, 10, 297-317.

Berger, G.W., Pillans, B.J., Bruce, J.G., McIntosh, P.D. (2002). Luminescence chronology of loess-paleosol sequences from southern South Island, New Zealand. *Quaternary Science Reviews*, 21, 1899-1913.

Birks, H.J.B. (2003). Quantitative palaeoenvironmental reconstructions from Holocene biological data. In: Mackay, A., Battarbee, R.W., Birks, H.J.B., Oldfield, J.S. (Eds.), *Global Change in the Holocene*. Arnold, London, United Kingdom, pp. 107–123.

Birks, H.J.B., Line, J.M. (1992). The use of rarefaction analysis for estimating palynological richness from Quaternary pollen-analytical data. *The Holocene*, 2, 1-10.

Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C., ter Braak, C.J.F. (1990). Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society of London*, 327, 263-278.

Birks, H.J.B., Heiri, O., Seppä, H., Bjune, A.E. (2010). Strengths and weaknesses of quantitative climate reconstructions based on late-Quaternary biological proxies. *The Open Ecology Journal*, 3, 68-110.

Boothroyd, I.K.G. (1989). First record of *Cricotopus aucklandensis* Sublette and Wirth (Diptera: Chironomidae) from mainland New Zealand, with a description of pupa and larva. *New Zealand Natural Sciences*, 16, 51-56.

Boothroyd, I.K.G. (1990). Additions to the Chironomidae (Diptera: Insecta) of New Zealand: *Cricotopus* van der Wulp species from a North Island stream. *New Zealand Journal of Zoology*, 17, 419-436.

Boothroyd, I.K.G. (2002). *Cricotopus* and *Paratrichocladius* (Chironomidae: Insecta) in New Zealand, with description of *C. hollyfordensis* n. sp., and redescriptions of adult and immature stages of *C. zealandicus* and *P. pluriserialis*. *New Zealand Journal of Marine and Freshwater Research*, 36, 775-788.

Boothroyd, I.K.G. (2004). A new species of *Naonella* Boothroyd (Chironomidae: Orthocladiinae) from New Zealand. *New Zealand Entomologist*, 27, 11-15.

Borcard, D., Gillet, F., Legendre, P. (2011). *Numerical ecology with R*. Springer Science+Business Media: New York, United States of America.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24, 123-140.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.

Brodersen, K.P., Lindegaard, C. (1999). Mass occurrence and sporadic distribution of *Corynocera ambigua* Zetterstedt (Diptera, Chironomidae) in Danish lakes. Neo- and palaeolimnological records. *Journal of Paleolimnology*, 22, 41-52.

Bronk Ramsey, C. (2008). Deposition models for chronological records. *Quaternary Science Reviews*, 27, 42-60.

Bronk Ramsey, C. (2009a). Bayesian analysis of radiocarbon dates. *Radiocarbon*, 51, 337-360.

Bronk Ramsey, C. (2009b). Dealing with outliers and offsets in radiocarbon dating. *Radiocarbon*, 51, 1023-1045.

Bronk Ramsey, C., Lee, S. (2013). Recent and Planned Developments of the Program OxCal. *Radiocarbon*, 55, 720-730.

Brooks, S.J., Birks, H.J.B. (2001). Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. *Quaternary Science Reviews*, 20, 1723-1741.

Campbell, J.F.E., Fletcher, W.J., Hughes, P.D., Shuttleworth, E.L. (2016). A comparison of pollen extraction methods confirms dense-media separation as a reliable method of pollen preparation. *Journal of Quaternary Science*, 31, 631-640.

Cassie, V. (1989). Micro-algae of Lake Pupuke, Auckland, New Zealand. *New Zealand Natural Sciences*, 16, 39-50.

Chappell, P.R. (2014). *The Climate and weather of Auckland: NIWA Science and Technology Series 60*. National Institute of Water and Atmospheric Research Ltd, Wellington, New Zealand.

Clark, P.U., Mix, A.C. (2002). Ice sheets and sea level of the Last Glacial Maximum. *Quaternary Science Reviews*, 21, 1-7.

Clement, A.J.H., Whitehouse, P.L., Sloss, C.R. (2016). An examination of spatial variability in the timing and magnitude of Holocene relative sea-level changes in the New Zealand archipelago. *Quaternary Science Reviews*, 131, 73-101.

Coffey, B. T., Clayton, J. S. (1987). Submerged macrophytes of Lake Pupuke, Takapuna, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 21, 193-198.

Danišík, M., Shane, P., Schmitt, A.K., Hogg, A., Santos, G.M., Storm, S., Evans, N.J., Fifield, L.K., Lindsay, J.M. (2012). Re-anchoring the late Pleistocene tephrochronology of New

Zealand based on concordant radiocarbon ages and combined $^{238}U/^{230}Th$ disequilibrium and (U–Th)/He zircon ages. *Earth and Planetary Science Letters*, 349, 240-250.

De'ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology*, 88, 243-251.

De Deckker, P., Arnold, L.J., van der Kaars, S., Bayon, G., Stuut, J.-B. W., Perner, K., Lopes dos Santos, R., Uemura, R., Demuro, M. (2019). Marine Isotope Stage 4 in Australasia: A full glacial culminating 65,000 years ago – Global connections and implications for human dispersal. *Quaternary Science Reviews*, 204, 187-207.

Dieffenbacher-Krall, A., Vandergoes, M. (2007). Guide to New Zealand Subfossil Chironomids.
http://www.climatechange.umaine.edu/Research/facilities/perl/nzguide.htmlS.

Dieffenbacher-Krall, A.C., Vandergoes, M.J., Denton, G.H. (2007). An inference model for mean summer air temperatures in the Southern Alps, New Zealand, using subfossil chironomids. *Quaternary Science Reviews*, 26, 2487-2504.

Donders, T.H., Wagner, F., Dilcher, D.L., Visscher, H. (2005). Mid- to late-Holocene El Niño-Southern Oscillation dynamics reflected in the subtropical terrestrial realm. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 10904-10908.

Donders, T.H., Haberle, S.G., Jope, G., Wagner, F., Visscher, H. (2007). Pollen evidence for the transition of the Eastern Australian climate system from the post-glacial to the present-day ENSO mode. *Quaternary Science Reviews*, 26, 1621-1637.

Eggermont, H., Heiri, O. (2012). The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. *Biological Reviews*, 87, 430-456.

Elith, J., Graham, C.H., Anderson, R.P. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129-151.

Elith, J., Leathwick, J.R., Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802-813.

Elliot, M., Neall, V., Wallace, C. (2005). A Late Quaternary pollen record from Lake Tangonge, far northern New Zealand. *Review of Palaeobotany and Palynology*, 136, 143-158.

Engels, S., Bakker, M.A.J., Bohncke, S.J.P., Cerli, C., Hoek, W.Z., Jansen, B., Peters, T., Renssen, H., Sachse, D., van Aken, J.M., van den Bos, V., van Geel, B., van Oostrom, R., Winkels, T., Wolma, M. (2016). Centennial-scale lake-level lowstand at Lake Uddelermeer (The Netherlands) indicates changes in moisture source region prior to the 2.8-kyr event. *The Holocene*, 26, 1075-1091.

Faegri, K., Iversen, J. (1989). *Textbook of pollen analysis*. John Wiley & Sons: Chichester, United Kingdom.

Fischer, H., Meissner, K.J., Mix, A.C., Abram, N.J., Austermann, J., Brovkin, V., Capron, E., Colombaroli, D., Daniau, A.-L., Dyez, K.A., et al. (2018). Palaeoclimate constraints on the impact of 2 °C anthropogenic warming and beyond. *Nature Geoscience*, 11, 474-485.

Forsyth, D.J. (1971). Some New Zealand chironomidae (Diptera). *Journal of the Royal Society of New Zealand*, 1, 113-144.

Forsyth, D.J. (1978). Benthic macroinvertebrates in seven New Zealand lakes. *New Zealand Journal of Marine and Freshwater Research*, 12, 41-49.

Forsyth, D.J. (1979). Life stages and taxonomic affinities of *Xenochironomus canterburyensis* (Chironomidae: Diptera). *New Zealand Journal of Zoology*, 6, 467-472.

Forsyth, D.J., McCallum, I.D. (1978a). *Xenochironomus canterburyensis* (Chironomidae, Diptera) an insectan, inquiline commensal of *Hyridella menziesi* (Lamellibranchia, Mollusca). *Journal of Zoology (London)*, 186, 331-334.

Forsyth, D.J., McCallum, I.D. (1978b). *Xenochironomus canterburyensis* (Diptera: Chironomidae), a commensal of *Hyridella menziesi* (Lamellibranchia) in Lake Taupo; features of pre-adult life history. *New Zealand Journal of Zoology*, 5, 795-800.

Franklin, D.A. (1968). Biological flora of New Zealand 3. *Dacrydium cupressinum* Lamb. (Podocarpaceae) rimu. *New Zealand Journal of Botany*, 6, 493-513.

Gelorini, V., Verbeken, A., van Geel, B., Cocquyt, C., Verschuren, D. (2011). Modern non-pollen palynomorphs from East African lake sediments. *Review of Palaeobotany and Palynology*, 164, 143-173.

Greenwell, B., Boehmke, B., Cunningham, J., GBM Developers (2018). gbm: Generalized Boosted Regression Models. R package version 2.1.4. https://CRAN.R-project.org/package=gbm

Haug, G.H., Hughen, K.A., Sigman, D.M., Peterson, L.C., Röhl, U. (2001). Southward migration of the Intertropical Convergence Zone through the Holocene. *Science*, 293, 1304-1308.

Havinga, A.J. (1967). Palynology and pollen preservation. *Review of Palaeobotany and Palynology*, 2, 81-98.

Hayward, B.W., Morley, M.S., Sabaa, A.T., Grenfell, H.R., Daymond-King, R., Molloy, C., Shane, P.A., Augustinus, P.A. (2008). Fossil records of the post-glacial marine breaching of Auckland's volcanic maar craters. *Records of the Auckland Museum*, 45, 73-99.

Hayward, B.W., Murdoch, G., Maitland, G. (2011). *Volcanoes of Auckland*. Auckland University Press: Auckland, New Zealand.

Heiri, O., Lotter, A.F. (2001). Effect of low count sums on quantitative environmental reconstructions: an example using subfossil chironomids. *Journal of Paleolimnology*, 26, 343-350.

Heiri, O., Lotter, A.F., Hausmann, S., Kienast, F. (2003). A chironomid-based Holocene summer air temperature reconstruction from the Swiss Alps. *The Holocene*, 13, 477-484.

Heyng, A.M., Mayr, C., Lücke, A., Striewski, B., Wastegård, S., Wissel, H. (2012). Environmental changes in northern New Zealand since the Middle Holocene inferred from stable isotope records ($\delta^{15}N$, $\delta^{13}C$) of Lake Pupuke. *Journal of Paleolimnology*, 48, 351-366.

Heyng, A.M., Mayr, C., Lücke, A., Wissel, H., Striewski, B. (2014). Late Holocene hydrologic changes in northern New Zealand inferred from stable isotope values of aquatic cellulose in sediments from Lake Pupuke. *Journal of Paleolimnology*, 51, 485-497.

Heyng, A.M., Mayr, C., Lücke, A., Moschen, R., Wissel, H., Striewski, B., Bauersachs, T. (2015). Middle and Late Holocene paleotemperatures reconstructed from oxygen isotopes and GDGTs of sediments from Lake Pupuke, New Zealand. *Quaternary International*, 374, 3-14.

Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54, 427-432.

Hjort, J.K., Marmion, M. (2008). Effects of sample size on the accuracy of geomorphological models. *Geomorphology*, 102, 341-350.

Hogg, A.G., Hua, Q., Blackwell, P.G., Niu, M., Buck, C.E., Guilderson, T.P., Heaton, T.J., Palmer, J.G., Reimer, P.J., Reimer, R.W., Turner, C.S.M., Zimmerman, S.R.H. (2013). ShCal13 Southern Hemisphere Calibration, 0-50,000 years cal BP. *Radiocarbon*, 55, 1889-1903.

Hooker, B.L., Fitzharris, B.B. (1999). The correlation between climatic parameters and the retreat and advance of Franz Josef Glacier, New Zealand. *Global and Planetary Change*, 22, 39-48.

Hopkins, J.L., Millet, M.A., Timm, C., Wilson, C.J.N. (2015). Tools and techniques for developing tephra stratigraphies in lake cores: A case study from the basaltic Auckland Volcanic Field, New Zealand. *Quaternary Science Reviews*, 123, 58-75.

Hopkins, J.L., Wilson, C.J.N., Leonard, G.S., Timm, C., McGee, L.E., Smith, I.E.M., Smith, E.G.C. (2017). Multi-criteria correlation of tephra deposits to source centres applied in the Auckland Volcanic Field, New Zealand. *Bulletin of Volcanology*, 79: 55.

Horrocks, M., Augustinus, P.C., Deng, Y., Shane, P., Andersson, S. (2005). Holocene vegetation, environment, and tephra recorded from Lake Pupuke, Auckland, New Zealand. *New Zealand Journal of Geology and Geophysics*, 48, 85-94.

Huisman, J., Olff, H., Fresco, L.F.M. (1993). A hierarchical set of models for species response analysis. *Journal of Vegetation Science*, 4, 37-46.

Huybers, P. (2006). Early Pleistocene glacial cycles and the integrated summer insolation forcing. *Science*, 313, 508-511.

Huybers, P., Eisenman, I. (2006). Integrated summer insolation calculations. *NOAA/NCDC Paleoclimatology Program Data Contribution*, #2006-079.

Huybers, P., Denton, G. (2008). Antarctic temperature at orbital timescales controlled by local summer duration. *Nature Geosciences*, 1, 787-792.

Imbrie, J., Boyle, E.A., Clemens, S.C., Duffy, A., Howard, W.R., Kukla, G., Kutzbach, J., Martinson, D.G., McIntyre, A., Mix, A.C., Molfino, B., Morley, J.J., Peterson, L.C., Pisias, N.G., Prell, W.L., Raymo, M.E., Shackleton, N.J., Toggweiler, J.R. (1992). On the structure and origin of major glaciation cycles 1. Linear responses to Milankovitch forcing. *Paleoceanography*, 7, 701-738.

Imbrie, J., Kipp, N.G. (1971). A new micropaleontological method for quantitative paleoclimatology: application to a Late Pleistocene Caribbean core. In: Turekian, K.K. (Ed.), *The Late Cenozoic Glacial Ages*. Yale University Press, New Haven, United States of America, pp. 77-181.

Imbrie, J., Webb III, T., 1981. Transfer functions: calibrating micropaleontological data in climatic terms. In: Berger, A. (Ed.), *Climatic Variations and Variability: Facts and Theories*. D. Reidel, Dordrecht, The Netherlands, pp. 125-134.

Imbrie, J., Berger, A., Boyle, E.A., Clemens, S.C., Duffy, A., Howard, W.R., Kukla, G., Kutzbach, J., Martinson, D.G., McIntyre, A., Mix, A.C., Molfino, B., Morley, J.J., Peterson, L.C., Pisias, N.G., Prell, W.L., Raymo, M.E., Shackleton, N.J., Toggweiler, J.R. (1993). On the structure and origin of major glaciation cycles 2. The 100,000-year cycle. *Paleoceanography*, 8, 699-735.

Jansen, F., Oksanen, J. (2013). How to model species responses along ecological gradients – Huisman-Olff-Fresco models revisited. *Journal of Vegetation Science*, 24, 1108-1117.

Jara, I.A., Newnham, R.M., Vandergoes, M.J., Foster, C.R., Lowe, D.J., Wilmshurst, J.M., Moreno, P.I., Renwick, J.A., Homes, A.M. (2015). Pollen–climate reconstruction from northern South Island, New Zealand (41˚S), reveals varying high- and low-latitude teleconnections over the last 16 000 years. *Journal of Quaternary Science*, 30, 817-829.

Jara, I.A., Newnham, R.M., Alloway, B.V., Wilmshurst, J.M., Rees, A.B.H. (2017). Pollen-based temperature and precipitation records of the past 14,600 years in northern New Zealand (37°S) and their linkages with the Southern Hemisphere atmospheric circulation. *The Holocene*, 27, 1756-1768.

Juggins, S. (2013). Quantitative reconstructions in palaeolimnology: new paradigm or sick science? *Quaternary Science Reviews*, 64, 20-32.

Juggins, S. (2015). rioja: Analysis of Quaternary Science Data. R package version 0.9-9. http://cran.r-project.org/package=rioja

Juggins, S. (2017). rioja: Analysis of Quaternary Science Data. R package version 0.9-15.1. http://cran.r-project.org/package=rioja

Juggins, S., Simpson, G.L., Telford, R.J. (2015). Taxon selection using statistical learning techniques to improve transfer function prediction. *The Holocene*, 25, 130-136.

Kidson, J., Renwick, J.A. (2002). Patterns of convection in the tropical Pacific and their influence on New Zealand weather. *International Journal of Climatology*, 22, 151-174.

Kidston, J., Renwick, J.A., McGregor, J. (2009). Hemispheric-scale seasonality of the Southern Annular Mode and impacts on the climate of New Zealand. *Journal of Climate*, 22, 4759-1770.

Kindler, P., Guillevic, M., Baumgartner, M., Schwander, J., Landais, A., Leuenberger, M. (2014). Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core. *Climate of the Past*, 10, 887-902.

Legendre, L., Legendre, P. (1998). *Numerical ecology*. Elsevier: Amsterdam, The Netherlands.

Leonard, G.S., Calvert, A.T., Hopkins, J.L., Wilson, C.J.N., Smid, E.R., Lindsay, J.M., Champion, D.E. (2017). High-precision ^{40}Ar/^{39}Ar dating of Quaternary basalts from Auckland Volcanic Field, New Zealand, with implications for eruption rates and paleomagnetic correlations. *Journal of Volcanology and Geothermal Research*, 343, 60-74.

Leslie, A.B. (2010). Flotation preferentially selects saccate pollen during conifer pollination. *New Phytologist*, 188, 273-279.

Liaw, A., Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2, 18-22.

Lisiecki, L.E., Raymo, M.E. (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography*, 20, PA1003.

Little, J.L., Smol, J.P. (2001). A chironomid-based model for inferring late-summer hypolimnetic oxygen in southeastern Ontario lakes. *Journal of Paleolimnology*, 26, 259-270.

Liu, Z., Zhu, J., Rosenthal, Y., Zhang, X., Otto-Bliesner, B.L., Timmermann, A., Smith, R.S., Lohmann, G., Zheng, W., Elison Timm, O. (2014). The Holocene temperature conundrum.

Proceedings of the National Academy of Sciences of the United States of America, 111, E3501-E3505.

Livingstone, D.M., Lotter, A.F. (1998). The relationship between air and water temperatures in lakes of the Swiss Plateau: a case study with palaeolimnological implications. *Journal of Paleolimnology*, 9, 181-198.

Lorrey, A.M., Bostock, H. (2017). The climate of New Zealand through the Quaternary. In: Shulmeister J. (Ed.) *Landscape and Quaternary environmental change in New Zealand*. Atlantic Press, Paris, France, pp. 67-139.

Lowe, D.J., Blaauw, M., Hogg, A.G. (2013). Ages of 24 widespread tephras erupted since 30,000 years ago in New Zealand, with re-evaluation of the timing and palaeoclimatic implications of the Lateglacial cool episode recorded at Kaipo bog. *Quaternary Science Reviews*, 74, 170-194.

Lowe, D.J., Rees, A.B.H., Newnham, R.M., Hazell, Z.J., Gehrels, M.J., Charman, D.J., Amesbury, M.J. (in preparation). Isochron-informed Bayesian age modelling for tephras and cryototephras: application to mid-Holocene Tūhua tephra, northern New Zealand.

Marcott, S.A., Shakun, J.D., Clark, P.U., Mix, A.C. (2013). A reconstruction of regional and global temperature for the past 11,300 years. *Science*, 339, 1198-1201.

Marsicek, J., Shuman, B.N., Nartlein, P.J., Shafer, S.L., Brewer, S. (2018). Reconciling divergent trends and millennial variations in Holocene temperatures. *Nature*, 554, 92-96.

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J.F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., Timmermann, A. (2013). Information from Paleoclimate Archives. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge, United Kingdom and New York, NY, United States of America.

McGlone, M.S. (2001). A late Quaternary pollen record from marine core P69, southeastern North Island, New Zealand. *New Zealand Journal of Geology and Geophysics*, 44, 69-77.

McGlone, M.S., Moar, N.T. (1977). The *Ascarina* decline and post-glacial climatic change in New Zealand. *New Zealand Journal of Botany*, 15, 485-489.

McGlone, M.S., Topping, W.W. (1977). Aranuian (post-glacial) pollen diagrams from the Tongariro region, North Island, New Zealand. *New Zealand Journal of Botany*, 15, 749-760.

McGlone, M.S., Topping, W.W. (1983). Late Quaternary vegetation, Tongariro region, central North Island, New Zealand. *New Zealand Journal of Botany*, 21, 53-76.

McGlone, M.S., Turney, C.S.M., Wilmshurst, J.M., Renwick, J., Pahnke, K. (2010a). Divergent trends in land and ocean temperature in the Southern Ocean over the past 18,000 years. *Nature Geoscience*, 3, 622-626.

McGlone, M.S., Hall, G.M.J., Wilmshurst, J.M. (2010b). Seasonality in the early Holocene: Extending fossil-based estimates with a forest ecosystem process model. *The Holocene*, 21, 517-526.

McGlone, M.S., Basher, L. (2012). Holocene vegetation change at treeline, Cropp Valley, Southern Alps, New Zealand. *Terra Australis*, 34, 343-358.

McGlone, M.S., Buitenwerf, R., Richardson, S.J. (2016). The formation of the oceanic temperate forests of New Zealand. *New Zealand Journal of Botany*, 54, 128-155.

Moar, N.T. (1993). *Pollen grains of New Zealand dicotyledonous plants.* Manaaki Whenua Press: Lincoln, New Zealand.

Moar, N.T., Wilmshurst, J.M. (2011). Standardizing names applied to pollen and spores in New Zealand Quaternary palynology. *New Zealand Journal of Botany*, 49, 201-229.

Molloy, C.M. (2008). *Tephrostratigraphy of the Auckland maar craters*. MSc thesis, University of Auckland, New Zealand.

Molloy, C., Shane, P., Augustinus, P. (2009). Eruption recurrence rates in a basaltic volcanic field based on tephra layers in maar sediments: implications for hazards in the Auckland volcanic field. Geological Society of America Bulletin, 121, 1666-1677.

Moore, P.D., Collinson, M., Webb, J.A. (1991). *Pollen analysis*. Blackwell Scientific Publications: Oxford, United Kingdom.

Moros, M., De Deckker, P., Jansen, E., Perner, K., Telford, R.J. (2009). Holocene climate variability in the Southern Ocean recorded in a deep-sea sediment core off South Australia. *Quaternary Science Reviews*, 28, 1932-1940.

Munsterman, D., Kerstholt, S. (1996). Sodium polytungstate, a new non-toxic alternative to bromoform in heavy liquid separation. *Review of Palaeobotany and Palynology*, 91, 417-422.

Needham, A.J., Lindsay, J.M., Smith, I.E.M., Augustinus, P., Shane, P.A. (2011). Sequential eruption of alkaline and sub-alkaline magmas from a small monogenetic volcano in the Auckland Volcanic Field, New Zealand. *Journal of Volcanology and Geothermal Research*, 201, 126-142.

Newnham, R.M. (1999). Environmental change in Northland, New Zealand during the last glacial and Holocene. *Quaternary International*, 57/58 61-70.

Newnham, R.M., Lowe, D.J. (1991). Holocene vegetation and volcanic activity, Auckland Isthmus, New Zealand. *Journal of Quaternary Science*, 6, 177-193.

Newnham, R., Alloway, B. (2004). A terrestrial record of Last Interglacial climate preserved by voluminous debris avalanche inundation in Taranaki, New Zealand. *Journal of Quaternary Science*, 19, 299-314.

Newnham, R.M., de Lange, P.J., Lowe, D.J. (1995). Holocene vegetation, climate and history of a raised bog complex, northern New Zealand based on palynology, plant macrofossils and tephrochronology. *The Holocene*, 5, 267-282.

Newnham, R.M., Lowe. D.J., Green, J.D., Turner, G.M., Harper, M.A., McGlone, M.S., Stout, S.L., Horie, S., Froggatt, P.C. (2004). A discontinuous ca. 80 ka record of Late Quaternary environmental change from Lake Omapere, Northland, New Zealand. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 207, 165-198.

Newnham, R.M., Vandergoes, M.J., Hendy, C.H., Lowe, D.J., Preusser, F. (2007a). A terrestrial palynological record for the last two glacial cycles from southwestern New Zealand. *Quaternary Science Reviews*, 26, 517-535.

Newnham, R.M., Lowe, D.J., Giles, T., Alloway, B.V. (2007b). Vegetation and climate of Auckland, New Zealand, since ca. 32 000 cal. yr ago: support for an extended LGM. *Journal of Quaternary Science*, 22, 517-534.

Newnham, R.M., Vandergoes, M.J., Sikes, E., Carter, L., Wilmshurst, J.M., Lowe, D.J., McGlone, M.S., Sandiford, A. (2012). Does the bipolar seesaw extend to the terrestrial southern mid-latitudes? *Quaternary Science Reviews*, 36, 214-222.

Newnham, R., McGlone, M., Moar, N., Wilmshurst, J., Vandergoes, M. (2013). The vegetation cover of New Zealand at the Last Glacial Maximum. *Quaternary Science Reviews*, 74, 202-214.

Newnham, R.M., Alloway, B.V., Holt, K.A., Butler, K., Rees, A.B.H., Wilmshurst, J.M., Dunbar, G., Hajdas, I. (2017a). Last Glacial pollen-climate reconstructions from Northland, New Zealand. *Journal of Quaternary Science*, 32, 685-703.

Newnham, R., Alloway, B., McGlone, M., Juchnowicz, H., Rees, A., Wilmshurst, J. (2017b). A Last Interglacial pollen-temperature reconstruction, central North Island, New Zealand. *Quaternary Science Reviews*, 170, 136-151.

Newnham, R.M., Lowe, D.J., Gehrels, M., Augustinus, P. (2018). Two-step human–environmental impact history for northern New Zealand linked to late-Holocene climate change. *The Holocene*, 28, 1093-1106.

Ogden, J., Ahmed, M. (1989). Climate response function analyses of kauri (*Agathis australis*) tree-ring chronologies in northern New Zealand. *Journal of the Royal Society of New Zealand*, 19, 205-221.

Ogden, J., Wilson, A., Hendy, C., Newnham, R.M. (1992). The late Quaternary history of kauri (*Agathis australis*) in New Zealand and its climatic significance. *Journal of Biogeography*, 19, 611-622.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H. (2016). vegan: Community ecology package. R package version 23-3. https://CRANR-projectorg/package=vegan

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H. (2018). vegan: Community Ecology Package. R package version 2.5-2. https://CRAN.R-project.org/package=vegan

Peres-Neto, P.R., Legendre, P., Dray, S., Borcard, D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, 87, 2614-2625.

Petherick, L., McGowan, H., Moss, P. (2008). Climate variability during the Last Glacial Maximum in eastern Australia: evidence of two stadials? *Journal of Quaternary Science*, 23, 787-802.

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.-M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V.Y., Lorius, C., Pépin, L., Ritz, C., Saltzman, E., Stievenard, M. (1999). Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature*, 399, 429-436.

Pocknall, D.T. (1981). Pollen morphology of the New Zealand species of *Dacrydium* Selander, *Podocarpus* L'Heritier, and *Dacrycarpus* Endlicher (Podocarpaceae). *New Zealand Journal of Botany*, 19, 67-95.

Prebble, J.G., Bostock, H.C., Cortese, G., Lorrey, A.M., Hayward, B.W., Calvo, E., Northcote, L.C., Scott, G.H., Neil, H.L. (2017). Evidence for a Holocene Climatic Optimum in the southwest Pacific: A multiproxy study. *Paleoceanography*, 32, 763-779.

Putnam, A.E., Schaefer, J.M., Denton, G.H., Barrell, D.J.A., Finkel, R.C., Andersen, B.G., Schwartz, R., Chinn, T.J.H., Doughty, A.M. (2012). Regional climate control of glaciers in New Zealand and Europe during the pre-industrial Holocene. *Nature Geoscience*, 5, 627-630.

Quinlan, R., Smol, J.P. (2001). Setting minimum head capsule abundance and taxa deletion criteria in chironomid-based inference models. *Journal of Paleolimnology*, 26, 327-342.

R Core Team (2015) R: A language and environment for statistical computing R Foundation for Statistical Computing. Vienna, Austria. URL https://wwwR-project.org/

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rees, A.B.H., Cwynar, L.C. (2010). Evidence for early postglacial warming in Mount Field National Park, Tasmania. *Quaternary Science Reviews*, 29, 443-454.

Renssen, H., Goosse, H., Fichefet, T., Brovkin, V., Driesschaert, E., Wolk, F. (2005). Simulating the Holocene climate evolution at northern high latitudes using a coupled atmosphere-sea ice-ocean-vegetation model. *Climate Dynamics*, 24, 23-43.

Renssen, H., Seppä, H., Heiri, O., Roche, D.M., Goosse, H., Fichefet, T. (2009). The spatial and temporal complexity of the Holocene thermal maximum. *Nature Geoscience*, 2, 411-414.

Rodríguez, J.J., Kuncheva, L.I., Alonso, C.J. (2006). Rotation forest: A new classifier ensemble method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28, 1619-1630.

Rolland, N., Laroque, I. (2007). The efficiency of kerosene flotation for extraction of chironomid head capsules from lake sediments samples. *Journal of Paleolimnology*, 37, 565-572.

Ruiz, Z., Brown, A. G., Langdon, P. G. (2006). The potential of chironomid (Insecta: Diptera) larvae in archaeological investigations of floodplain and lake settlements. *Journal of Archaeological Science*, 33, 14-33.

Ryan, M.T., Dunbar, G.B., Vandergoes, M.J., Neil, H.L., Hannah, M.J., Newnham, R.M., Bostock, H., Alloway, B.V. (2012). Vegetation and climate in Southern Hemisphere mid-latitudes since 210 ka: new insights from marine and terrestrial pollen records from New Zealand. *Quaternary Science Reviews*, 48, 80-98.

Salonen, J.S., Verster, A.J., Engels, S., Soininen, J., Trachsel, M., Luoto, M. (2016). Calibrating aquatic microfossil proxies with regression-tree ensembles: Cross-validation with modern chironomid and diatom data. *The Holocene*, 26, 1040-1048.

Sandiford, A., Horrocks, M., Newnham, R.M., Ogden, J., Alloway, B.V. (2002). Environmental change during the last glacial maximum (ca 25 500-16 500 years BP) at Mt Richmond, Auckland Isthmus, New Zealand. *Journal of the Royal Society of New Zealand*, 32, 155-167.

Sandiford, A., Newnham, R.M., Alloway, B.V., Ogden, J. (2003). A 28 000–7600 cal yr BP pollen record of vegetation and climate change from Pukaki Crater, northern New Zealand. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 201, 235-247.

Saunders, K.M., Roberts, S.J., Perren, B., Butz, C., sime, L., Davies, S., Van Nieuwenhuyze, W., Grosjean, M., Hodgson, D.A. (2018). Holocene dynamics of the Southern Hemisphere westerly winds and possible links to CO_2 outgassing. *Nature Geoscience*, 11, 650-655.

Shane, P., Sandiford, A. (2003). Paleovegetation of marine isotope stages 4 and 3 in Northern New Zealand and the age of the widespread Rotoehu tephra. *Quaternary Research*, 59, 420-429.

Shannon, C. E., Weaver, W., (1949). *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, United States of America.

Shulmeister, J., Soons, J.M., Berger, G.W., Harper, M., Holt, S., Moar, N., Carter, J.A. (1999). Environmental and sea-level changes on Banks Peninsula (Canterbury, New Zealand) through three glaciation-interglaciation cycles. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 152, 101-127.

Shulmeister, J., Shane, P., Lian, O.B., Okuda, M., Carter, J.A., Harper, M., Dickinson, W.W., Augustinus, P., Heijnis, H. (2001). A long late-Quaternary record from Lake Poukawa, Hawkes bay, New Zealand. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 176, 81-107.

Sikes, E.L., Medeiros, P.M., Augustinus, P., Wilmshurst, J.M., Freeman, K.R. (2013). Seasonal variations in aridity and temperature characterize changing climate during the last deglaciation in New Zealand. *Quaternary Science Reviews*, 74, 245-256.

Simpson, G.L. (2007). Analogue methods in palaeoecology: Using the analogue package. *Journal of Statistical Software*, 22, 1-29.

Simpson, G.L., Oksanen, J. (2016). analogue: Analogue matching and modern analogue technique transfer function models, R package version 0.17-0. http://cran.r-project.org/package=analogue

Stephens, T., Atkin, D., Augustinus, P.C., Shane, P., Lorrey, A., Street-Perrott, A., Nilsson, A., Snowball, I. (2012a). A late glacial Antarctic climate teleconnection and variable Holocene seasonality at Lake Pupuke, Auckland, New Zealand. *Journal of Paleolimnology*, 48, 785-800.

Stephens, T., Atkin, D., Cochran, U., Augustinus, P.C., Reid, M., Lorrey, A., Shane, P., Street-Perrott, A. (2012b). A diatom-inferred record of reduced effective precipitation during the Last Glacial Coldest Phase (28.8–18.0 cal kyr BP) and increasing Holocene seasonality at Lake Pupuke, Auckland, New Zealand. *Journal of Paleolimnology*, 48, 801-817.

Striewski, B., Mayr, C., Flenley, J., Naumann, R., Turner, G., Lücke, A. (2009). Multi-proxy evidence of late Holocene human-induced environmental changes at Lake Pupuke, Auckland (New Zealand). *Quaternary International*, 202, 69-93.

Striewski, B., Shulmeister, J., Augustinus, P.C., Soderholm, J. (2013). Late Holocene climate variability from Lake Pupuke maar, Auckland, New Zealand. *Quaternary Science Reviews*, 77, 46-54.

Suggate, R.P. (1990). Late Pliocene and Quaternary glaciations of New Zealand. *Quaternary Science Reviews*, 9, 175-197.

ter Braak, C.J.F. (1995). Non-linear methods for multivariate statistical calibration and their use in palaeoecology: a comparison of inverse (*k*-nearest neighbours, partial least squares and weighted averaging partial least squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems*, 28, 165-180.

ter Braak, C.J.F., Looman, C.W.N. (1986). Weighted averaging, logistic regression and the Gaussian response model. *Vegetatio*, 65, 3-11.

Therneau, T., Atkinson, B. (2018). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-13. https://CRAN.R-project.org/package=rpart

Trachsel, M., Telford, R.J. (2015). Technical Note: Estimating unbiased transfer-function performances in spatially structured environments. *Climate of the Past Discussions*, 11, 4729-4749.

Tyler, P.A. (1992). A lakeland from the Dreamtime: the second founders' lecture. *European Journal of Phycology*, 27, 353-368.

Ummenhofer, C.C., England, M.H. (2007). Interannual extremes in New Zealand precipitation linked to modes of Southern Hemisphere climate variability. *Journal of Climate*, 20, 5418-5440.

van den Bos V, Rees A, Newnham R, Vandergoes M, Wilmshurst J, Augustinus P (2018) Holocene temperature, effective precipitation and seasonality in northern New Zealand linked to Southern Hemisphere summer insolation. Quat Sci Rev 201: 77-88

van Geel, B., Buurman, J., Brinkkemper, O., Schelvis, J., Aptroot, A., van Reenen, G., Hakbijl, T. (2003). Environmental reconstruction of a Roman Period settlement site in Uitgeest (The Netherlands), with special reference to coprophilous fungi. *Journal of Archaeological Science*, 30, 873-883.

Vandergoes, M.J., Newnham, R.M., Preusser, F., Hendy, C.H., Lowell, T.V., Fitzsimons, S.J., Hogg, A.G., Kasper, H.U., Schlüchter, C. (2005). Regional insolation forcing of late Quaternary climate change in the Southern Hemisphere. *Nature*, 436, 242-245.

Vandergoes, M.J., Dieffenbacher-Krall, A.C., Newnham, R.M., Denton, G.H., Blaauw, M. (2008). Cooling and changing seasonality in the Southern Alps, New Zealand during the Antarctic Cold Reversal. *Quaternary Science Reviews*, 27, 589-601.

Vandergoes, M.J., Hogg, A.G., Lowe, D.J., Newnham, R.M., Denton, G.H., Southon, J., Barrell, D.J.A., Wilson, C.J.N., McGlone, M.S., Allan, A.S.R., Almond, P.C., Petchey, F., Dabell, K., Dieffenbacher-Krall, A.C., Blaauw, M. (2013). A revised age for the Kawakawa/Oruanui tephra, a key marker for the Last Glacial Maximum in New Zealand. *Quaternary Science Reviews*, 74, 195-201.

Velle, G., Brodersen, K.P., Birks, H.J.B., Willassen, E. (2010). Midges as quantitative temperature indicator species: lessons for palaeoecology. *The Holocene*, 20, 989-1002.

Waipara, N.W., Hill, S., Hill, L.M.W., Hough, E.G., Horner, I.J. (2013). Surveillance methods to determine tree health, distribution of kauri dieback disease and associated pathogens. *New Zealand Plant Protection*, 66, 235-241.

Walker, I.R., Smol, J.P., Engstrom, D.R., Birks, H.J.B. (1991). An assessment of Chironomidae as quantitative indicators of past climatic change. *Canadian Journal of Fisheries and Aquatic Sciences*, 48, 975-987.

Wardle, P. (1991) *Vegetation of New Zealand*. Cambridge University Press: Cambridge, United Kingdom and New York, NY, United States of America.

Wetzel, R.G. (2001). *Limnology: Lake and river ecosystems* (3rd edition). Academic Press: San Diego, United States of America and London, United Kingdom.

Whittaker, T.E., Hendy, C.H., Hellstrom, J.C. (2011). Abrupt millennial-scale changes in intensity of Southern Hemisphere westerly winds during marine isotope stages 2–4. *Geology*, 39, 455-458.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag: New York, United States of America.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, United States of America.

Wilke, C.O. (2018). cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 0.9.3. https://CRAN.R-project.org/package=cowplot

Williams, P.W. (1996). A 230 ka record of glacial and interglacial events from Aurora Cave, Fiordland, New Zealand. *New Zealand Journal of Geology and Geophysics*, 39, 225-241.

Williams, N.J., Harle, K.J., Gale, S.J., Heijnis, H. (2006). The vegetation history of the last glacial-interglacial cycle in eastern New South Wales, Australia. *Journal of Quaternary Science*, 21, 735-750.

Williams, P.W., McGlone, M., Neil, H., Zhao, J.-X. (2015). A review of New Zealand palaeoclimate from the Last Interglacial to the global Last Glacial Maximum. *Quaternary Science Reviews*, 110, 92-106.

Wilmshurst, J.M., McGlone, M.D., Charman, D.J. (2002). Holocene vegetation and climate change in southern New Zealand: linkages between forest composition and quantitative surface moisture reconstructions from an ombrogenous bog. *Journal of Quaternary Science*, 17, 653-666.

Wilmshurst, J.M., McGlone, M.S., Leathwick, J.R., Newnham, R.M. (2007). A pre-deforestation pollen-climate calibration model for New Zealand and quantitative temperature reconstructions for the past 18 000 years BP. *Journal of Quaternary Science*, 22, 535-547.

Winterbourn, M.J., Gregson, K.L.D. (1981). Guide to the aquatic insects of New Zealand. *Bulletin of the Entomological Society of New Zealand*, 5, 1-80.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14, 763-773.

Wood, S.N. (2006). *Generalized Additive Models: An introduction with R*. Chapman and Hall/CRC, United States of America.

Wood, S.N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73, 3-36.

Wood, S.N. (2017) *Generalized Additive Models: An Introduction with R* (2nd edition). Chapman and Hall/CRC, United States of America.

Woodward, C.A., Shulmeister, J. (2006). New Zealand chironomids as proxies for human-induced and natural environmental change: transfer functions for temperature and lake production (chlorophyll a). *Journal of Paleolimnology*, 36, 407-429.

Wright Jr., H.H. (1990). An improved Hongve sampler for surface sediments. *Journal of Paleolimnology*, 4, 91-92.

Wright, I.C., McGlone, M.S., Nelson, C.S., Pillans, B.J. (1995). An integrated latest Quaternary (Stage 3 to Present) palaeoclimatic and paleoceanographic record from offshore northern New Zealand. *Quaternary Research*, 44, 283-293.

Zabenskie, S. (2006). *Post-glacial climatic change on Boothia Peninsula, Nunavut, Canada.* Thesis submitted to the Faculty of Graduate and Postdoctoral Studies of the University of Ottawa, Canada.

Appendices

Appendix A: Supplementary Information to Chapter 3

A1: Model diagnostics

For mean annual air temperature (MAAT) reconstruction, the pre-deforestation pollen calibration model of Wilmshurst et al. (2007) was used. The training set contains 83 taxa, 50 of which are present in the Lake Pupuke dataset (Table A-1).

Table A-1. Pollen taxa included in the MAAT inference models and their Hill's N2. Only taxa that are also present in the Lake Pupuke samples are shown.

Taxon	Hill's N2	Taxon	Hill's N2
Tall conifers		Small trees and shrubs - continued	
Agathis australis	11.5	*Griselinia*	14.1
Dacrydium cupressinum	71.6	*Halocarpus*	20.0
Dacrycarpus dacrydioides	49.5	*Hebe*	9.3
Libocedrus plumosa	7.8	*Leucopogon fasciculatus*	1.5*
Manoao colensoi	3.3*	*Muehlenbeckia*	13.0
Phyllocladus trichomanoides	16.9	*Myrsine*	33.9
Podocarpus	57.6	*Neomyrtus* type	7.4
Prumnopitys ferruginea	58.1	*Plagianthus*	21.0
Prumnopitys taxifolia	66.4	*Pseudopanax*	19.0
Tall angiosperms		*Pseudowintera*	6.2
Alectryon excelsus	6.3	*Rhopalostylis*	3.1*
Elaeocarpus	18.1	*Syzygium maire*	1.8*
Fuscospora	54.3	*Tupeia antarctica*	8.0
Knightia	11.9	*Weinmannia*	9.2
Laurelia novae-zelandiae	6.5	Herbs	
Lophozonia menziesii	22.6	*Dactylanthus taylorii*	1.0*
Metrosideros	22.3	*Passiflora tetrandra*	1.0*
Nestegis	18.1	Poaceae	26.2
Quintinia	7.6	*Ripogonum scandens*	2.0*
Small trees and shrubs		*Rubus*	15.3
Aristotelia	4.5*	Ferns and allies	
Ascarina lucida	8.8	*Cyathea dealbata*	13.2
Asteraceae	27.2	*Cyathea smithii* type	29.5
Coprosma	51.2	*Dicksonia fibrosa*	2.0*
Cordyline	2.7*	*Dicksonia squarrosa*	19.9
Coriaria	6.4	Monolete	20.1

*Hill's N2 ≤ 5; rare taxon

For summer temperature (SmT) reconstruction, an adapted version of the chironomid inference model of Dieffenbacher-Krall et al. (2007) was used. The original training dataset was slightly adapted, as we did not distinguish between "short" or "long" antennal pedestals in *Tanytarsus funebris* types A and C. The adapted training set therefore contains 38 taxa, vs. 40 in the original (Table A-2). The same number of lakes (60) were included and chironomid percentages were square root transformed to stabilise the variance. Dieffenbacher-Krall et al. found that the weighted averaging-partial least squares (WA-PLS) approach produced a stronger performing model, with a low jack-knifed root mean square error of prediction ($RMSEP_{jack} = 1.27$) and a high coefficient of determination ($r^2_{jack} = 0.80$),

compared to the partial least squares (PLS) technique (RMSEP$_{jack}$ = 1.30, r^2_{jack} = 0.79). To make sure that this was still the case with the adapted dataset, we repeated the analyses and checked the model performances using leave-one-out cross validation. The models were run in R (version 3.3.1; R Core Team, 2016) using package "rioja" (Juggins, 2015).

Table A-2. Chironomid taxa included in the SmT inference model, their Hill's N2, and their β coefficients in the Weighted Average–Partial Least Squares model.

Taxon	Hill's N2	*β* coefficient
Diamesinae		
Maoridiamesa	1.3*	-2.3
Podonominae		
Parochlus	6.3	8.1
Chironominae		
tribe Chironomini		
Chironomini early instar	24.4	9.6
Chironomus	32.3	6.7
Cladopelma	11.0	16.0
Kiefferulus	1.8*	46.4
Lauterborniella	2.4*	4.3
Microtendipes	1.2*	-2.9
Parachironomus	3.6*	34.0
Paucispinigera	7.2	18.7
Polypedilum	10.8	16.8
tribe Tanytarsini		
Corynocera	8.3	13.9
Paratanytarsus	2.4*	23.0
Tanytarsus funebris type A	13.6	4.5
Tanytarsus funebris type B	8.2	9.7
Tanytarsus funebris type C	29.5	10.8
Tanytarsus vespertinus	5.6	4.1
Orthocladiinae		
Cricotopus aucklandensis	10.3	26.6
Cricotopus planus	3.6*	9.9
Cricotopus zealandicus	3.8*	13.5
Eukiefferiella	4.8*	5.5
Gymnometriocnemus	5.3	24.9
Hevelius	2.9*	11.2
Limnophyes/Paralimnophyes	3.3*	25.0
Naonella type 305	7.3	3.7
Naonella type 419	3.8*	-0.4
Naonella forsythi	9.0	13.8
Naonella kimihia	24.1	20.9
Orthocladiinae 1b	4.1*	10.5
Orthocladiinae 3	4.9*	12.3
Orthocladiinae 7	2.0*	21.5
Orthocladiinae 9	1.7*	14.0
Pirara/Orthocladiinae E	1.4*	8.4
Tanypodinae		
Ablabesmyia	6.1	21.6
Macropelopini type 1	13.9	8.4
Macropelopini type 2	9.8	3.7
Macropelopini type 3**	16.2	7.3

*Hill's N2 ≤ 5; rare taxon
**named "*Apsectrotanypus*" in Dieffenbacher-Krall *et al.* (2007)

The results are shown in Table A-3 and are similar to the original inference models (Dieffenbacher-Krall et al., 2007), although PLS has a lower $RMSEP_{jack}$ (i.e. 1.19) and higher r^2_{jack} (i.e. 0.82) than the WA-PLS model. However, WA-PLS has very similar diagnostics to PLS ($RMSEP_{jack}$ = 1.22, r^2_{jack} = 0.81) and seems to perform better at relatively higher temperatures, with residuals closer to 0 than the PLS model at high temperatures (Fig. A-1). As our fossil dataset is expected to be on the higher side of the SmT gradient and even above it, we chose to use the WA-PLS model with 3 components for the SmT reconstruction from our fossil chironomid assemblages.

Table A-3. Performance of WA-PLS and PLS inference models.

	$RMSEP_{jack}$	r^2_{jack}	Average bias	Maximum bias	% Improvement $RMSEP_{jack}$ over previous component
Weighted averaging–partial least squares					
Component 1	1.54	0.70	-0.19	1.76	
Component 2	1.30	0.78	-0.08	1.63	15.3
Component 3	1.22	0.81	-0.10	1.05	6.4
Component 4	1.20	0.82	-0.10	0.89	1.5
Component 5	1.23	0.81	-0.11	0.97	-2.8
Partial least squares					
Component 1	1.76	0.60	-0.01	2.47	
Component 2	1.34	0.77	-0.03	1.68	24.1
Component 3	1.19	0.82	-0.02	1.32	11.0
Component 4	1.16	0.83	0.00	1.21	2.9
Component 5	1.16	0.83	0.02	1.09	-0.1
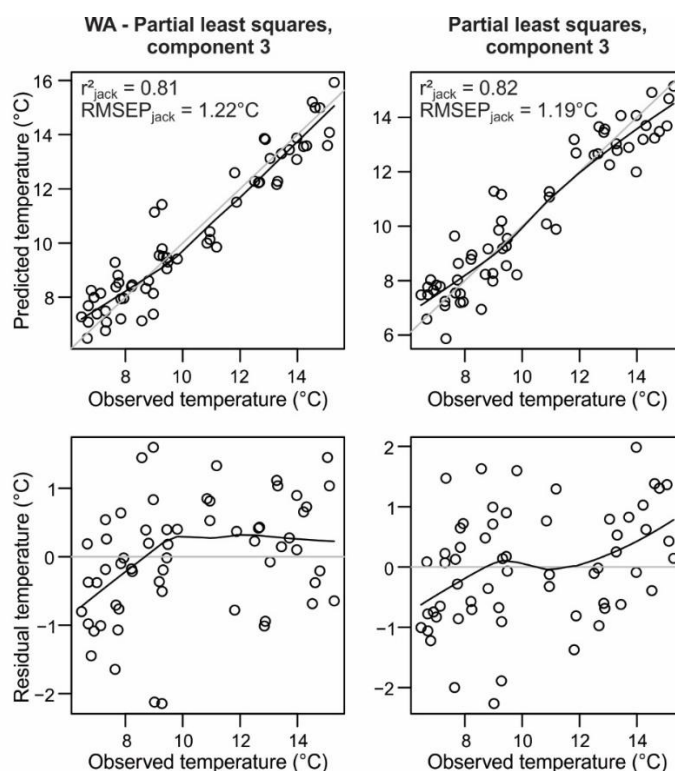
Figure A-1. Residual plots of the WA-PLS (left) and
PLS (right) SmT inference models.

References for Appendix A1:

Dieffenbacher-Krall, A.C., Vandergoes, M.J., Denton, G.H. (2007). An inference model for mean summer air temperatures in the Southern Alps, New Zealand, using subfossil chironomids. *Quaternary Science Reviews*, 26, 2487-2504.

Juggins, S. (2015) rioja: Analysis of Quaternary Science Data, R package version 0.9-9. (http://cran.r-project.org/package=rioja).

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Wilmshurst, J.M., McGlone, M.S., Leathwick, J.R., Newnham, R.M. (2007). A pre-deforestation pollen-climate calibration model for New Zealand and quantitative temperature reconstructions for the past 18 000 years BP. *Journal of Quaternary Science*, 22, 535-547.

A2: Rare pollen taxa

Table A-4: Table of rare pollen taxa (maximum abundance of < 1%) that are not included in the pollen diagram (Fig. 3.3a).

Taxon

Depth (cm)	Age (yr BP)	Tall trees				Small trees and shrubs												Herbs												Exotics				Ferns and allies						Wetl.*
		Alectryon excelsus	*Dysoxylum spectabile*	*Laurelia novae-zelandiae*	*Quintinia*	*Weinmannia*	*Aristotelia*	*Cordyline*	*Ericaceae*	*Geniostoma rupestre*	*Hebe*	*Ileostylus micranthus*	*Leucopogon fasciculatus*	*Macropiper*	*Pseudowintera*	*Rhopalostylis*	*Tupeia antarctica*	*Apiaceae*	*Asteraceae liguliflorae*	*Asteraceae tubuliflora*	*Brassicaceae*	*Dactylanthus taylorii*	*Forstera*	*Libertia grandiflora*	*Passiflora tetrandra*	*Phormium*	*Plantago*	*Ranunculus* type	*Ripogonum*	*Rubus*	*Quercus*	*Rumex*	*Taraxacum* type	*Blechnum*	*Dicksonia fibrosa*	*Dicksonia squarrosa*	*Ophioglossum*	*Phlegmariurus varius*	*Gleichenia*	*Typha*
13	-12																													X	X	X	X	X						X
34	164				X						X								X																					
54	343																		X	X																			X	
70	470			X																X																			X	
86	538					X						X	X							X																		X		
100	621	X			X																																			
111	734	X											X																	X										
122	861		X			X				X	X									X					X															
134	1001				X				X				X															X												
143	1108		X		X																																			
153	1227										X																			X										
162	1332				X						X																													
168	1402																			X																				
183	1571	X																																						
193	1669											X																												
206	1721	X																		X										X										
221	1970	X																		X										X										
231	2267																													X										
242	2628																																							
252	2963						X											X		X																				
261	3272	X																		X															X	X				
269	3548			X																																				
279	3889																																						X	
285	4095	X								X				X						X					X			X												

Table A-4 - continued

Depth (cm)	Age (yr BP)	Alectryon excelsus	Dysoxylum spectabile	Laurelia novae-zelandiae	Quintinia	Weinmannia	Aristotelia	Cordyline	Ericaceae	Geniostoma rupestre	Hebe	Ileostylus micranthus	Leucopogon fasciculatus	Macropiper	Pseudowintera	Rhopalostylis	Tupeia antarctica	Apiaceae	Asteraceae liguliflorae	Asteraceae tubuliflora	Brassicaceae	Dactylanthus taylorii	Forstera	Libertia grandiflora	Passiflora tetrandra	Phormium	Plantago	Ranunculus type	Ripogonum	Rubus	Quercus	Rumex	Taraxacum type	Blechnum	Dicksonia fibrosa	Dicksonia squarrosa	Ophioglossum	Phlegmariurus varius	Gleichenia	Typha
		Tall trees			**Small trees and shrubs**												**Herbs**															**Exotics**			**Ferns and allies**						**Wetl.***
294	4403																X			X									X											
300	4610																																							
309	4921		X												X					X																				
315	5126	X							X						X															X						X				
322	5366																			X	X																			
329	5608	X																																						
339	5948										X						X								X	X			X											
347	6223	X			X															X									X											X
354	6460			X																X																				
365	6827	X		X								X	X					X				X																		
371	7023																																						X	
376	7179						X								X														X											
382	7350												X												X															
387	7471				X													X												X						X				
393	7568												X		X															X						X				
398	7627				X																																			
399	7640	X			X												X								X				X											
403	7687												X												X					X										
408	7765	X			X								X																											X
412	7849							X																					X											
417	7996								X						X	X																								
423	8199																																							
428	8381										X	X													X															
432	8528											X			X										X												X			
438	8747	X			X						X						X								X															X
443	8924	X																							X															
454	9244																								X			X												

Table A-4 - continued

Depth (cm)	Age (yr BP)	Tall trees					Small trees and shrubs											Herbs													Exotics			Ferns and allies					Wetl.*	
		Alectryon excelsus	*Dysoxylum spectabile*	*Laurelia novae-zelandiae*	*Quintinia*	*Weinmannia*	*Aristotelia*	*Cordyline*	*Ericaceae*	*Geniostoma rupestre*	*Hebe*	*Ileostylus micranthus*	*Leucopogon fasciculatus*	*Macropiper*	*Pseudowintera*	*Rhopalostylis*	*Tupeia antarctica*	*Apiaceae*	*Asteraceae liguliflorae*	*Asteraceae tubuliflora*	*Brassicaceae*	*Dactylanthus taylorii*	*Forstera*	*Libertia grandiflora*	*Passiflora tetrandra*	*Phormium*	*Plantago*	*Ranunculus type*	*Ripogonum*	*Rubus*	*Quercus*	*Rumex*	*Taraxacum type*	*Blechnum*	*Dicksonia fibrosa*	*Dicksonia squarrosa*	*Ophioglossum*	*Phlegmariurus varius*	*Gleichenia*	*Typha*
463	9393									X						X				X				X			X													
471	9500	X										X								X										X									X	
478	9623																							X	X															
483	9745	X										X								X				X										X						
491	9988										X	X	X					X	X	X			X		X															
499	10252									X	X				X					X														X						
507	10526					X						X	X																X					X					X	
517	10867		X									X								X														X		X				
531	11356	X										X	X							X										X				X		X				
542	11741					X			X											X																				
554	12161	X										X							X															X						
564	12506					X																X			X									X						
575	12890					X														X					X															
585	13226											X								X																				
596	13574											X																X									X			
619	14050	X								X		X	X							X										X						X		X		
637	14844	X								X		X																X		X						X				
654	15616										X									X																				
670	16138																			X										X										

*Wetl. = Wetland and aquatic plant

Appendix B: Supplementary Information to Chapter 4

B1: Outliers

CA of 103 lakes with 44 chironomid taxa shows five lakes as outliers (Fig. B-1a), i.e. L0418, L0409, L0422, L0419 and L0421 on axis 1 related to the abundances of several rare taxa including *Pirara*/Orthocladiinae E, *Naonella* type 419 and Orthocladiinae 9, and L0901, L0705, L0601, L0310 and L0327 on axis 2 related to *Corynocera* abundances. PCA of 103 lakes with 44 chironomid taxa shows only three outliers (Fig. B-1b), i.e. L0310, NI03 and L0327, mostly related to *Corynocera* abundances. This PCA also shows that there is some clustering of lakes that lie in the same geographical region, i.e. lakes from the western South Island (WSI) on the left of PC 1 and lakes from the eastern South Island (ESI) on the right of PC1, although the clusters are not completely separated.
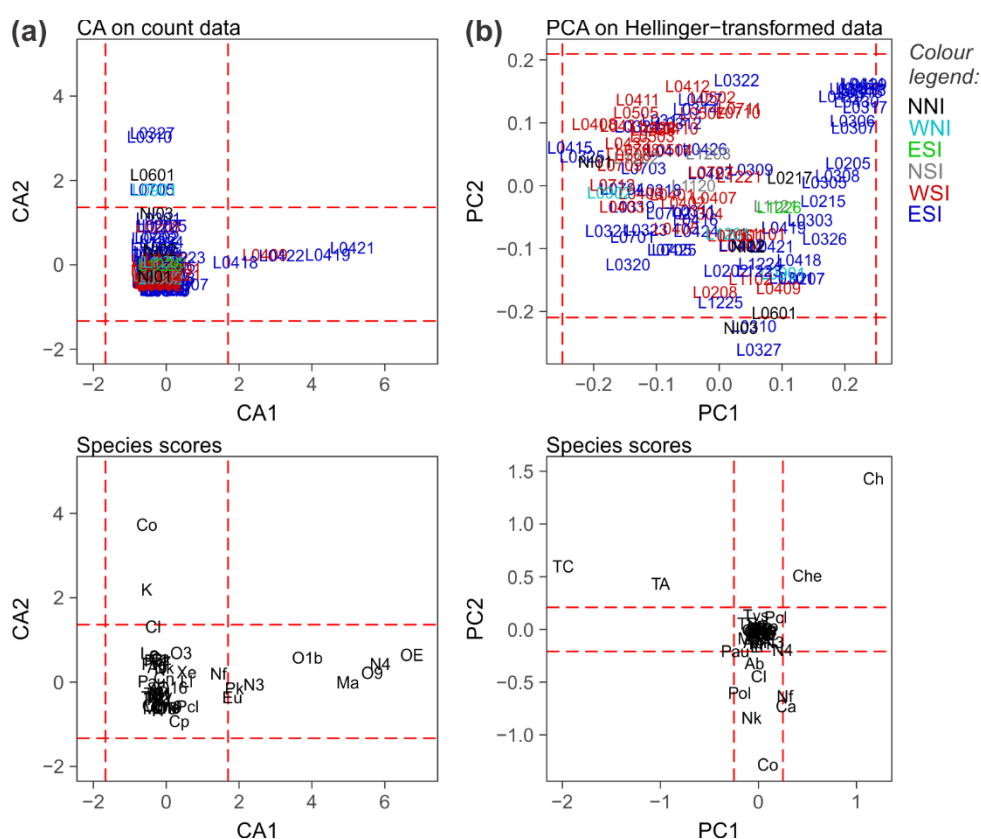
Figure B-1. Ordination of 103 lakes based on the abundances of 44 chironomid taxa. CA on left (a), PCA on Hellinger-transformed data on right (b). Site scores and species scores displayed separately. Sites are colour coded according to region: NNI = northern North Island, WNI = western NI, ENI = eastern NI, NSI = northern South Island, WSI = western SI, ESI = eastern SI. Red, dashed lines represent the mean site scores on axis 1 and 2 respectively, plus or minus two times the standard deviation on the same axis.

PCA of 103 lakes based on the standardised values of 31 environmental and climate variables shows four lakes as outliers (Fig. B-2), i.e. L1226, L0327, L1201 and L0901, correlating mostly with total carbon, conductivity, and several cations. There is also some clustering of sites by geographical region, but it is difficult to see which variables drive these clusters. Constrained ordinations make the relationships between environment and region more obvious. L0901 and L0327 in particular could be outliers as identified by ordination based on taxon abundances as well as the environmental variables.
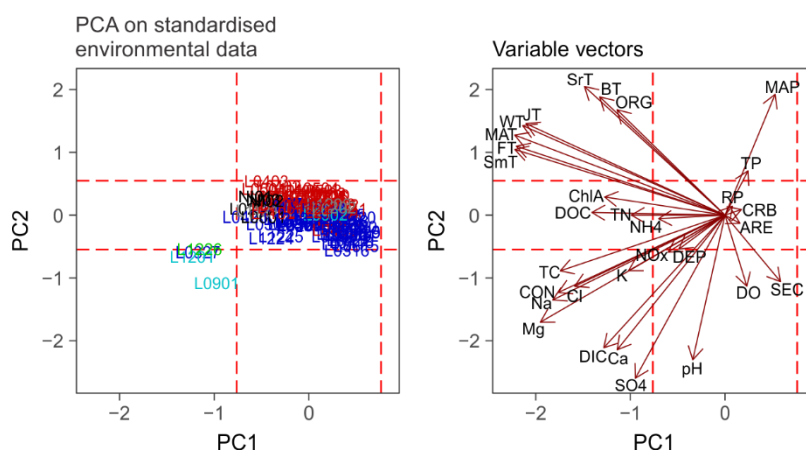
Figure B-2. Ordination of 103 lakes based on the standardised values of 31 environmental variables. Site scores appear on left, and variable vectors on right. Sites are colour coded according to region: see Fig. B-1 for legend of colours. Red, dashed lines represent the mean site scores on axis 1 and 2 respectively, plus or minus two times the standard deviation on the same axis.

B2: Additional ordination triplots

Fig. B-3 shows the results of the constrained ordinations using all 25 significant variables, compared to constrained ordinations after removing redundant variables (ten variables remain). Constrained ordination with forward selection showed that of the ten remaining variables, five accounted for a significant portion of inertia, i.e. SmT, pH, maximum depth, sulphate and dissolved organic carbon (Fig. B-4). We performed additional constrained ordinations including only lakes for which all five variables were measured (85 lakes; Fig. B-5). The triplots for this reduced training set are very similar to the full training set.
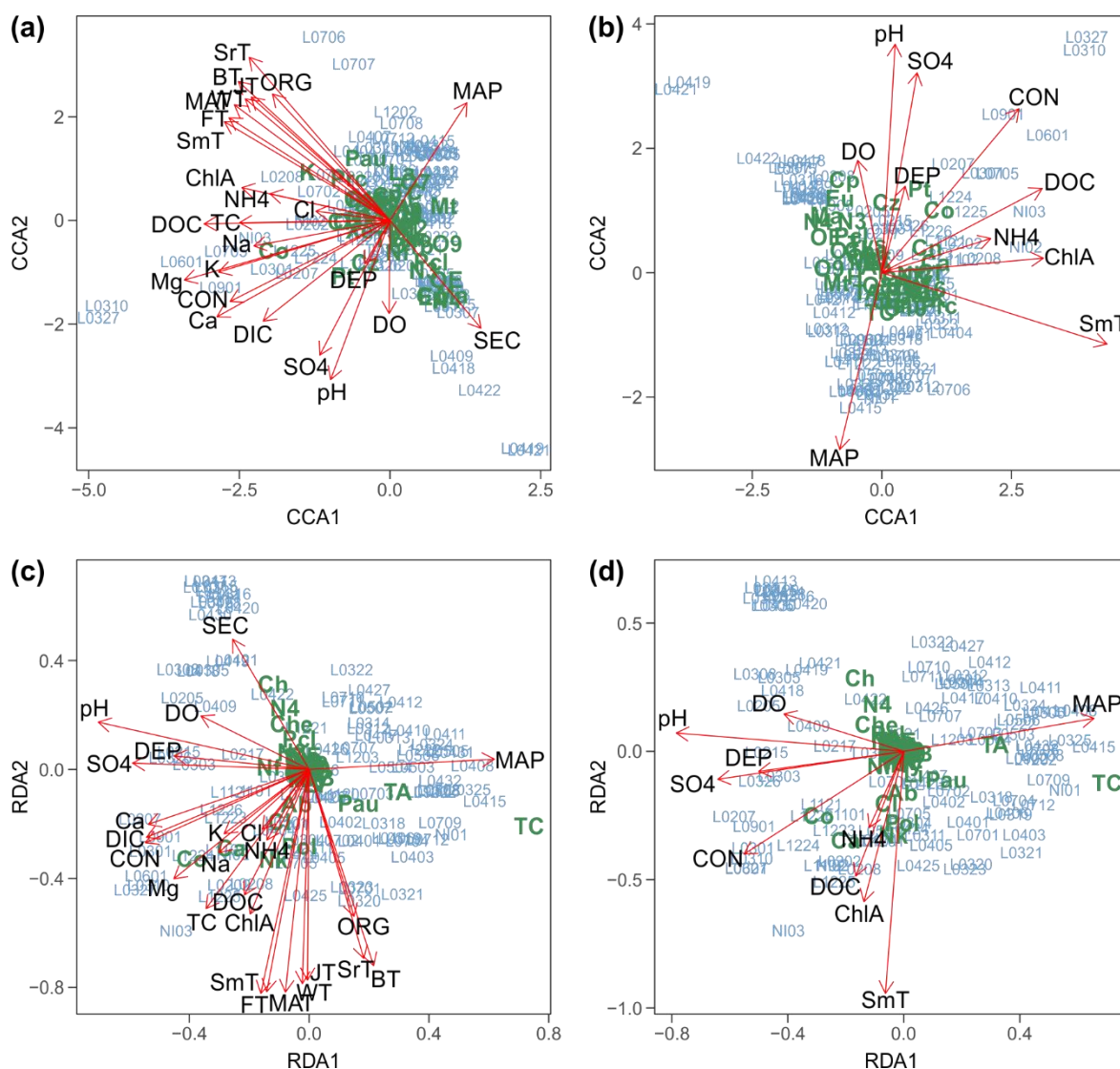
Figure B-3. Triplots of ordinations of 103 lakes constrained by environmental and climate variables. CCA at top (a, b), RDA at bottom (c, d). Constrained by all 25 significant variables on left (a, c). Constrained by selection of 10 variables (non-redundant) on right (b, d). Site scores indicated in blue, taxon scores in green. Cluster of sites in top-left corner of RDAs consists of: L0306, L0307, L0315, L0316, L0317, L0413, L0414, L0420, L0428, L0429 and L0430.
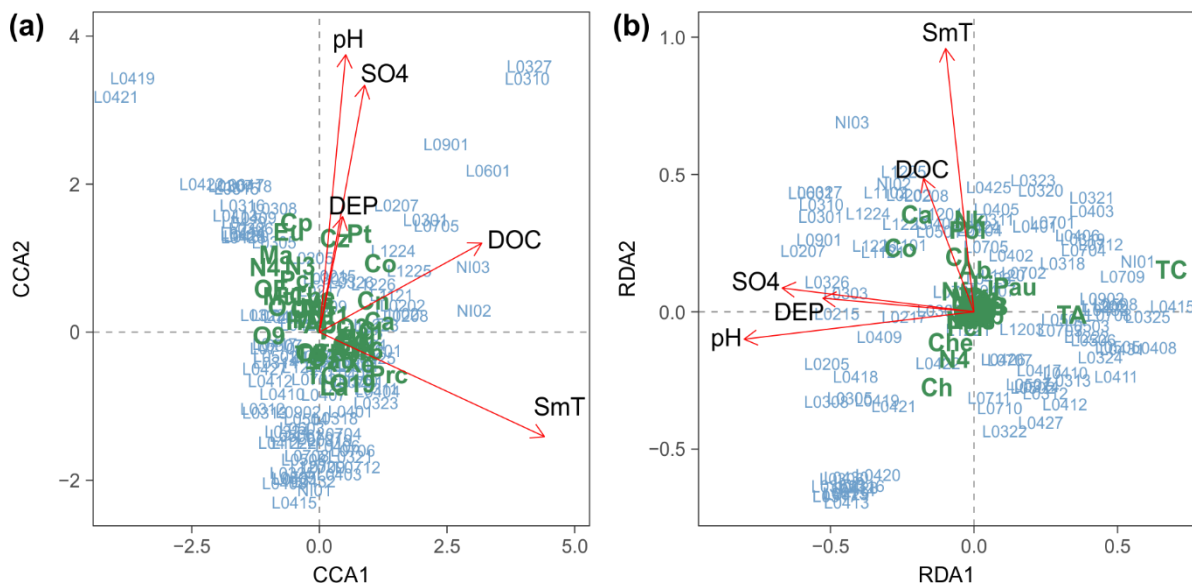
Figure B-4. Ordinations of 103 lakes constrained by variables that explain a significant portion of the variance, as determined by forward selection: CCA (a) and RDA (b). Site scores indicated in blue, taxon scores in green.
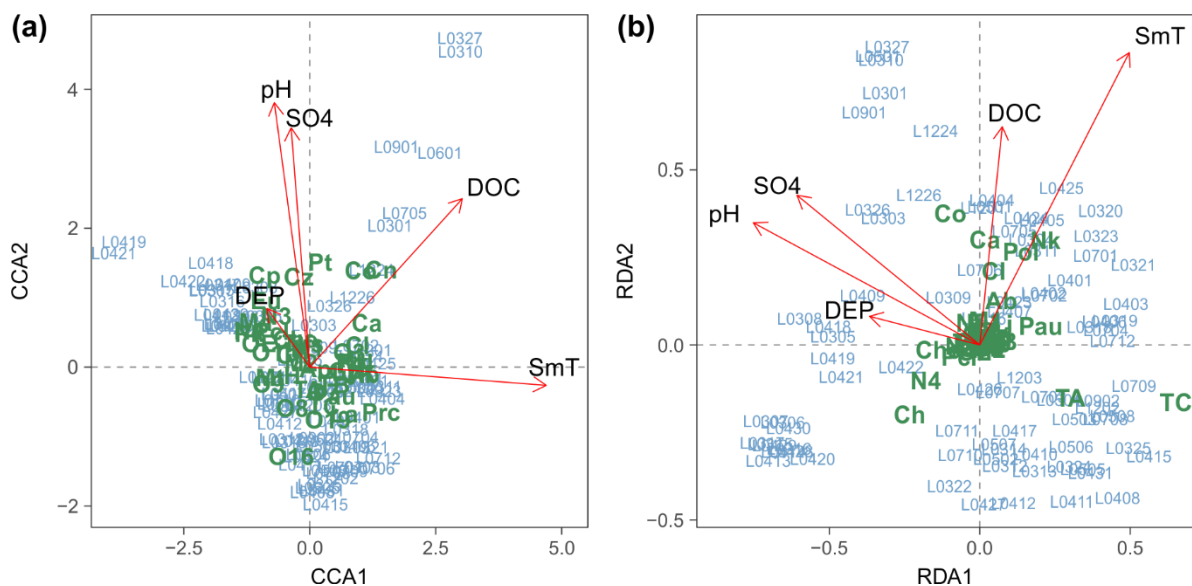
Figure B-5. Ordinations of 85 lakes for which all five variables that explain a significant portion of the variance were measured: CCA (a) and RDA (b). Site scores indicated in blue, taxon scores in green.

B3: Chironomid responses: best-fitting HOF models

Table B-1. Best-fitting HOF models for each taxon that occurs in ≥ 5 sites, to each significant variable. HOF-types: I = no significant response, II = sigmoidal, III = sigmoidal with plateau, IV = unimodal, V = unimodal skewed, VI = bimodal symmetric, VII = bimodal skewed.

	SmT		pH		DEP		SO$_4^{2-}$		DOC	
	I-V	*I-VII*	*I-V*	*I-VII*	*I-V*	*I-VII*	*I-V*	*I-VII*	*I-V*	*I-VII*
Ma	V	IV	IV	IV	IV	IV	IV	IV	II	II
Pcl	II	VII	V	VII	III	VII	IV	VII	II	II
Che	II	VII	II	VII	III	VII	III	III	IV	IV
Ch	II	VII	II	VII	V	V	V	VII	II	VII
Cl	III	VII	II	VII	III	VII	II	VII	III	III
K	III	VI	III	III	V	VII	IV	IV	IV	IV
La	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
Mt	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
Pau	III	III	III	VII	V	VII	V	VII	V	VII
Pol	III	VII	II	VII	II	VII	II	VII	V	V
Prc	II	II	I	I	II	II	<5	<5	<5	<5
Xe	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
Co	V	VII	V	VII	II	VII	II	VII	II	VII
Pt	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
TA	II	VII	V	V	II	VII	II	VII	V	V
TB	V	VII	III	III	II	VII	II	II	V	IV
TC	V	VII	V	VI	II	VII	II	VII	V	V
Tvs	V	VII	III	V	V	VII	I	VII	III	VII
Cn	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
Ca	V	V	II	VII	V	VII	IV	VII	III	III
Cp	II	II	IV	V	I	IV	IV	IV	IV	IV
Cz	I	I	III	III	II	VII	<5	<5	<5	<5
Eu	II	II	IV	IV	III	IV	III	III	II	II
Gy	V	VII	II	VII	V	VII	I	VII	IV	IV
H	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
Li	V	VII	III	VII	V	VII	III	III	I	VII
Nf	II	VII	V	VII	V	VII	V	VII	II	VII
Nk	V	VII	II	VII	II	VII	II	VII	II	VII
N3	II	II	V	VII	III	III	III	III	II	VII
N4	II	VII	V	V	V	V	V	VII	II	II
Pk	IV	IV	<5	<5	V	VII	<5	<5	<5	<5
OE	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
O3	V	I	I	I	III	II	I	I	V	VII
O7	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
O8	I	I	II	II	IV	VII	IV	IV	I	I
O9	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
O16	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
O19	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
O1b	II	II	V	IV	I	I	V	VII	II	II
Ab	III	III	V	VII	IV	IV	I	VII	IV	IV
P1	<5	<5	<5	<5	<5	<5	<5	<5	<5	<5
Ap	II	VII	II	VII	II	VII	II	VII	II	VII
M1	IV	IV	II	VII	I	VII	II	VII	IV	III
M2	II	VII	II	VII	II	VII	IV	VII	II	VII

B4: Model parameterisation
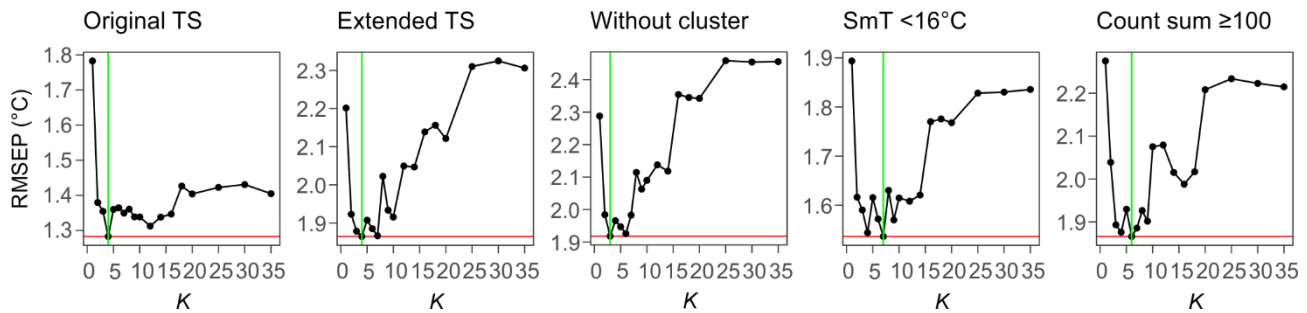
Figure B-6. The effect of K *(number of subsets the chironomid taxa are divided over) on the RMSEP of the RoF models. Lowest RMSEP indicated with red line,* K *value for which RMSEP is lowest indicated with green line.*
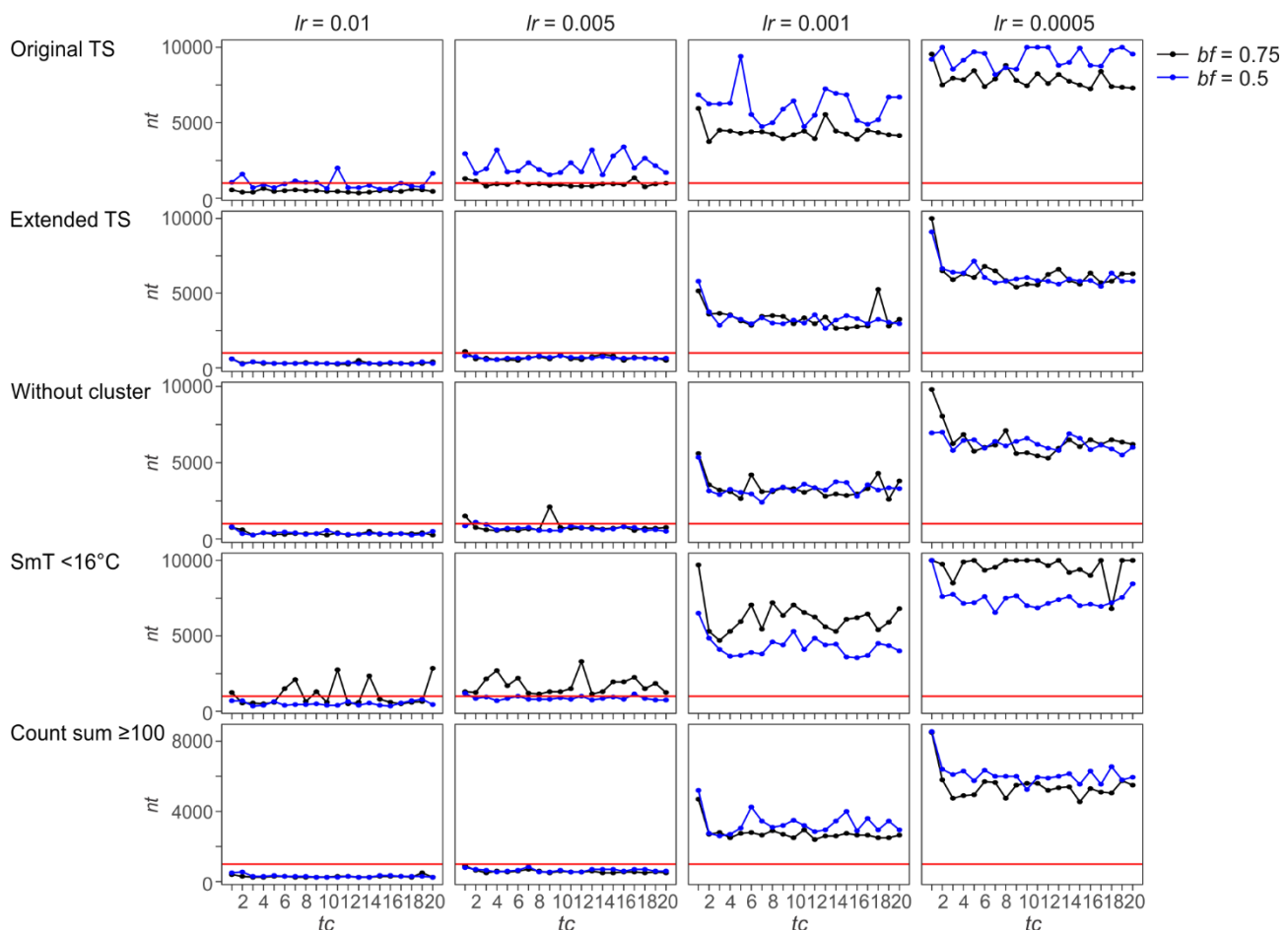
*Figure B-7. Number of trees (*nt*) in BRT model that results in the lowest deviance at different learning rates (*lr*) and tree complexities (*tc*) for two different bagging fractions (*bf*). Red line indicates the minimum required* nt*, i.e. 1000.*
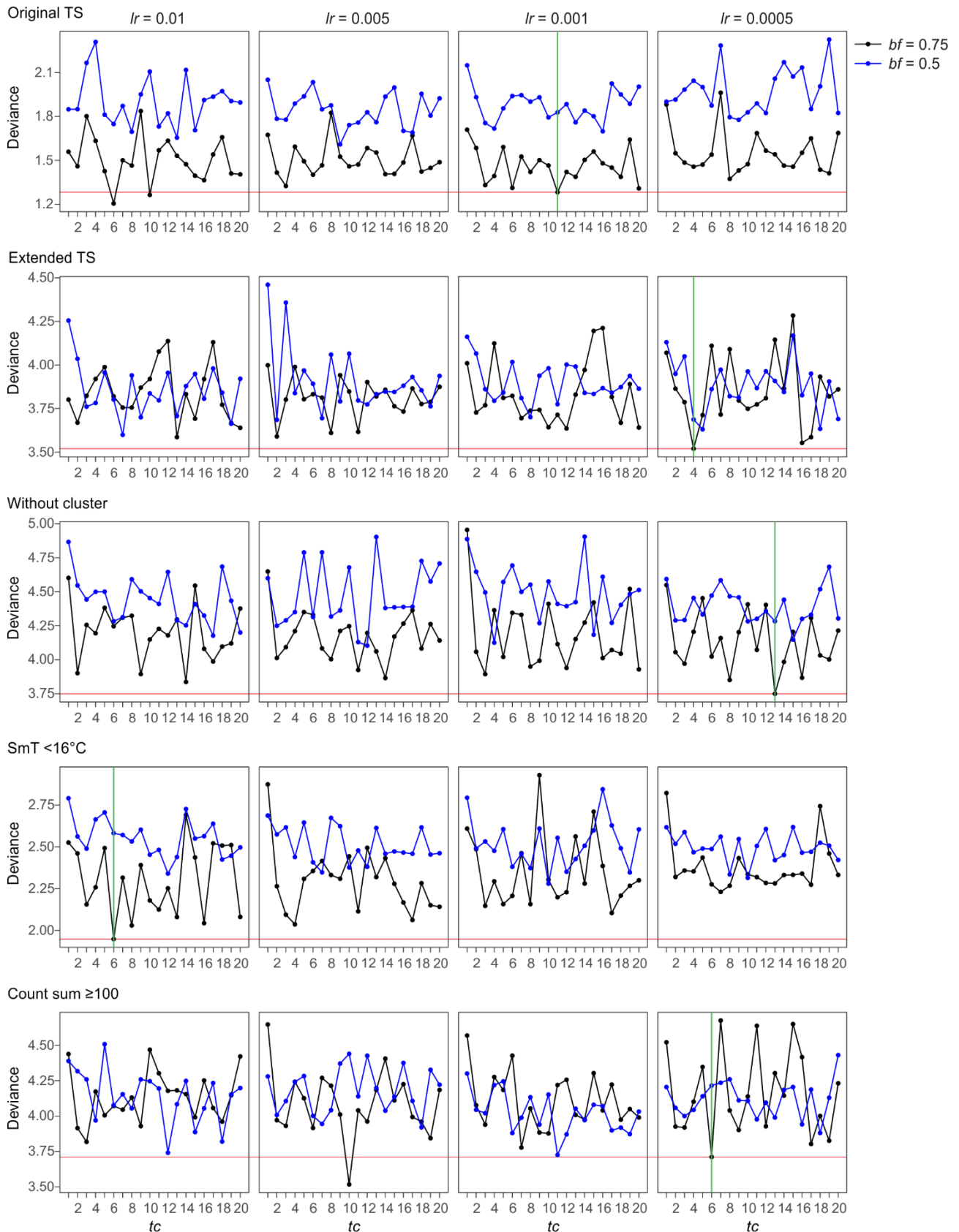
*Figure B-8. Deviance of BRT model for different learning rates (*lr*) and tree complexities (*tc*) for two different bagging fractions (*bf*). Red line indicates the lowest deviance for which* nt *> 1000. Green line indicates* lr *and* tc *for which deviance is lowest.*

B5: Labelled residual plots
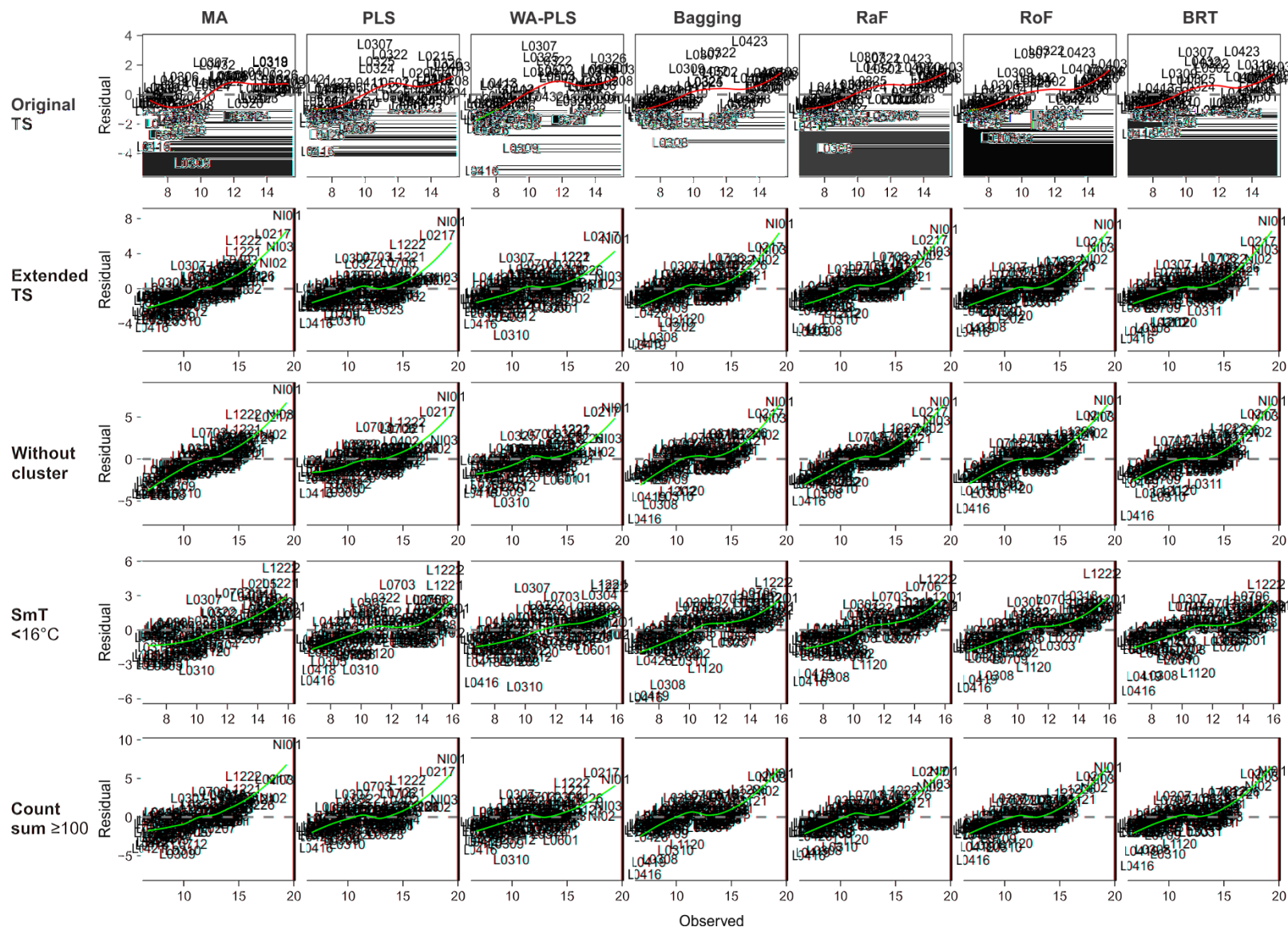
Figure B-9: Caption on next page

Figure B-9. Model performance of seven model types (MA = modern analogue technique, PLS = partial least squares, WA-PLS = weighted-averaging PLS, RaF = random forest, RoF = rotation forest and BRT = boosted regression tree) in columns and five TS configurations in rows. Predicted SmT values vs. residuals. Red line is the loess-smoothed trend through the points.

_____

B6: Downcore reconstructions
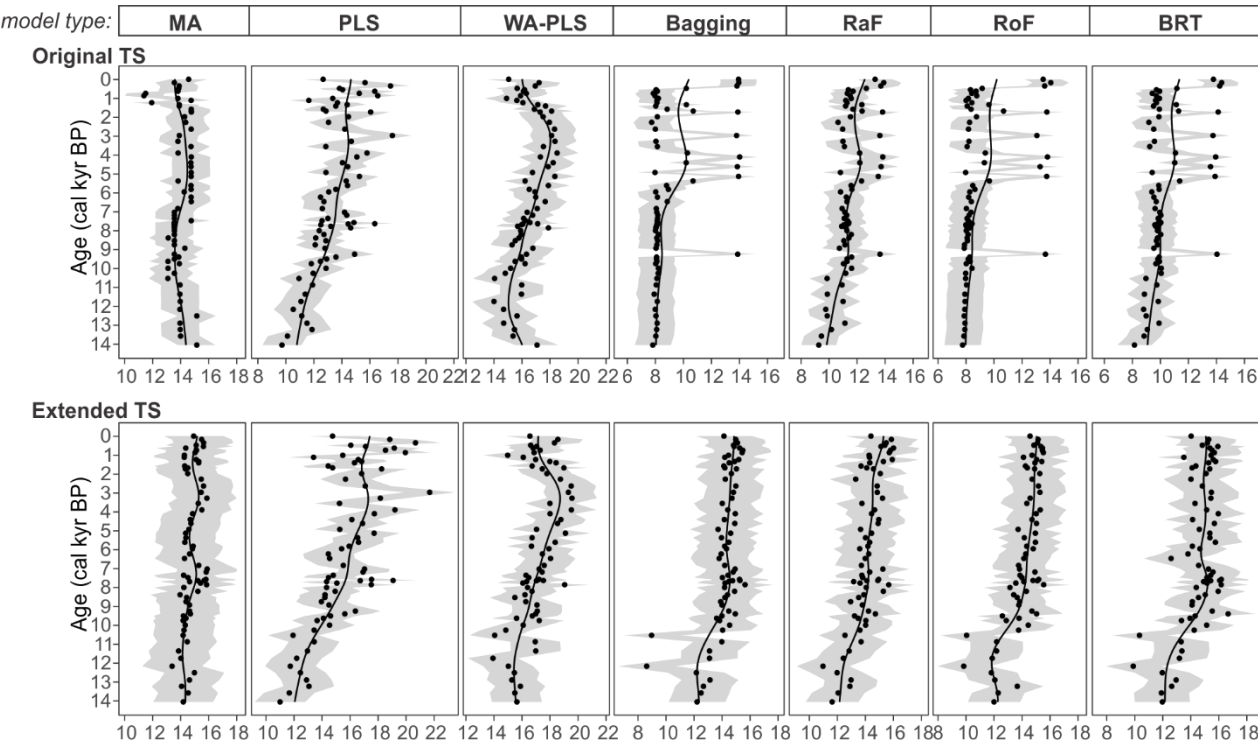
Figure B-10. Lake Pupuke SmT reconstructions using all seven model types. Training set configurations used: original 60-lake training set and extended 103-lake training set. Points are reconstructed values, shaded area represents RMSEP and curve is the result of a GAM fitted through the points.

Appendix C: Supplementary Information to Chapter 5

C1: MAAT model diagnostics

To model mean annual air temperatures (MAAT) from the pollen assemblages, we used the pre-deforestation pollen training set published by Wilmshurst et al. (2007). We slightly adapted the data, i.e. we did not distinguish between *Metrosideros excelsa* and *M. robusta*-type, neither between *Dicksonia fibrosa* and *D. squarrosa*, but otherwise followed their methods as closely as possible. The adapted training set comprises 79 taxa, 57 of which are present in the Orakei Basin fossil pollen dataset (Table C-1). All modelling and calculation of model diagnostics was performed in R (version 3.5.1; R Core Team, 2018), using package "analogue" (Simpson, 2007; Simpson & Oksanen, 2016) for the modern analogue technique (MA) technique and package "rioja" (Juggins, 2015) for the partial least squares (PLS) transfer function.

Pollen abundances of both the training set and fossil dataset were converted to percentages and square-root transformed (= Hellinger transformation). Although fossil sample abundances show a good fit to MAAT (Fig. C-1), most of them do not have a close modern analogue in the training set (Fig. C-2). Therefore, several configurations with different taxon compositions were tested in an attempt to reduce the dissimilarity between fossil and training set samples. Configuration 1 includes all 79 taxa, configuration 2 excludes taxa that are present in < 2 training set samples and taxa with a maximum abundance of < 1% (69 taxa retained) and configuration 3 excludes all taxa that produce spores (Pteridophytes) but includes all pollen-producing taxa (68 taxa retained). Configuration 2 shows very similar diagnostics compared to configuration 1 (Fig. C-2b, Table C-2, Table C-3) and little improvement. Configuration 3 shows decreased dissimilarities to modern analogues, but dissimilarities are still high for most samples (Fig. C-2c). Therefore, going forward, the configurations with the highest r^2 and lowest RMSEP were chosen.

Table C-1. Pollen taxa present in the Orakei Basin dataset and their Hill's N2 value in the pollen pre-deforestation training set.

Taxon	Hill's N2*	Taxon	Hill's N2*
Tall trees - conifers		Small trees and shrubs - continued	
Agathis australis	11.5	*Pseudopanax*	19.0
Dacrydium cupressinum	71.6	*Pseudowintera*	6.2
Dacrycarpus dacrydioides	49.5	*Schefflera digitata*	3.0
Lepidothamnus	3.0	*Syzygium maire*	1.8
Libocedrus	12.8	*Tupeia antarctica*	8.0
Manoao colensoi	3.3	Herbs	
Phyllocladus	40.7	*Acaena*	NA
Podocarpus	57.6	Apiaceae	NA
Prumnopitys ferruginea	58.1	*Astelia*	NA
Prumnopitys taxifolia	66.4	Brassicaceae	NA
Tall trees - angiosperms		*Bulbinella*	NA
Alectryon excelsus	6.3	Caryophyllaceae	NA
Corynocarpus laevigatus	NA	Chenopodiaceae	NA
Dysoxylum spectabile	NA	*Collospermum microsporum*	2.7
Elaeocarpus	18.1	*Dactylanthus taylorii*	1.0
Knightia excelsus	11.9	*Epilobium*	NA
Laurelia novae-zelandiae	6.5	*Euphrasia*	NA
Metrosideros	24.0	Fabaceae	NA
Nestegis	18.1	*Freycinetia banksii*	1.7
Quintinia serrata	7.6	*Geranium*	NA
Weinmannia	9.2	*Geum*	NA
Fuscospora	54.3	*Gunnera*	NA
Lophozonia menziesii	22.6	*Libertia*	NA
Small trees and shrubs		*Mentha*	NA
Aristotelia	4.5	*Passiflora tetranda*	1.0
Ascarina lucida	8.8	*Phormium*	NA
Asteraceae tubuliflorae	27.2	*Plantago*	NA
Coprosma	51.2	Poaceae	26.2
Cordyline	2.7	*Potentilla*	NA
Coriaria	6.4	*Ranunculus* type	NA
Dodonaea viscosa	3.3	*Ripogonum scandens*	2.0
Entelea arborescens	NA	*Rubus*	15.3
Ericaceae	NA	*Rumex*	NA
Fuchsia	1.0	*Sebaea ovata*	NA
Geniostoma	NA	*Urtica*	NA
Griselinia	14.1	*Wahlenbergia*	NA
Halocarpus	20.0	*Xeronema callistemon*	NA
Hebe type	9.3	Ferns and allies	
Hedycarya arborea	1.0	*Cyathea dealbata*	13.2
Hoheria	17.6	*Cyathea smithii* type	29.5
Ileostylus micranthus	NA	*Dicksonia*	20.3
Leptecophylla	NA	*Hymenophyllum*	6.4
Leucopogon fasciculatus	1.5	*Hypolepis*	1.0
Macropiper	NA	*Lycopodium scariosum*	NA
Melicytus	2.4	Monolete	20.1
Muehlenbeckia	13.0	*Ophioglossum*	NA
Myrsine	33.9	*Phlegmarius varius*	NA
Neomyrtus type	7.4	*Phymatosorus*	NA
Pennantia corymbosa	2.0	*Pteridium esculentum*	9.8
Pittosporum	7.4	*Pteris*	NA
Plagianthus	21.0	*Trichomanes*	NA

*If value for Hill's N2 = "NA", taxon is not included in the training set
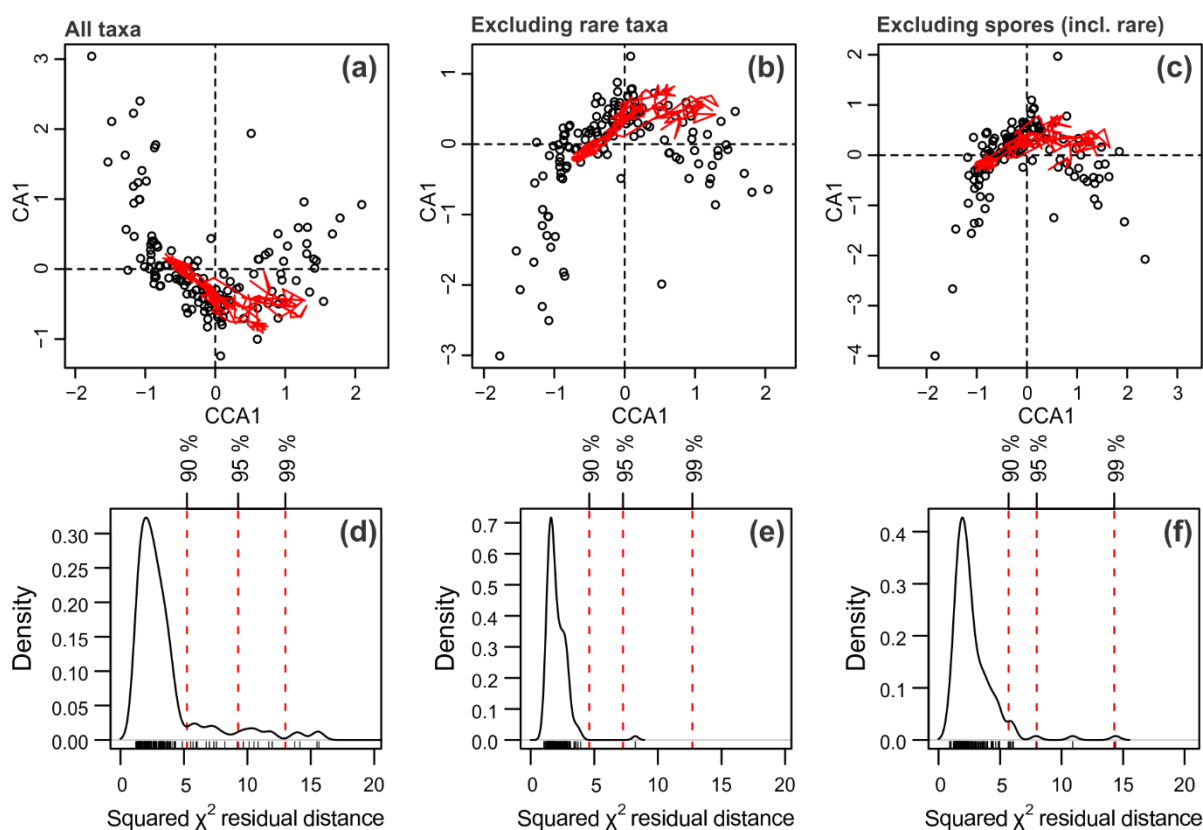
Figure C-1. Fit of fossil samples to MAAT: canonical correspondence analysis of training set samples with MAAT as constraining variable, with fossil samples plotted passively (a, b, c) and the distribution of squared residual distances of fossil samples to MAAT (d, e, f). Three configurations were tested: full dataset (a, d), dataset without rare taxa (b, e) and dataset without Pteridophytes (c, f). In (a), (b) and (c) points represent training set samples, and red lines represent "timetracks" of fossil samples. In (d), (e) and (f), curve shows distribution of squared residuals of fossil samples to MAAT and vertical, dashed lines represent 90th, 95th and 99th percentiles of the distribution of squared residuals of training set samples to MAAT. Fossil samples with a squared residual distance greater than the 90th percentile are considered poorly fitted to MAAT, while those with a squared residual distance greater than the 95th percentile are considered very poorly fitted to MAAT.
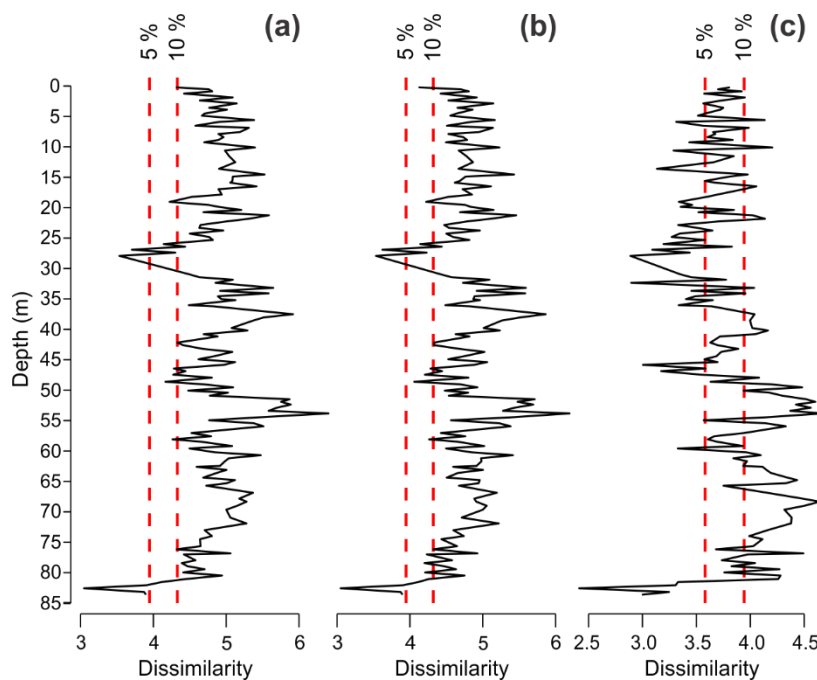
Figure C-2. Dissimilarity of fossil samples from their closest modern analogue for full dataset (a), dataset without rare taxa (b) and dataset without Pteridophytes (c). Vertical, dashed lines represent 5th and 10th percentiles of the distribution of dissimilarities among the training set samples. Fossil samples that display a dissimilarity higher than the 5th percentile are considered to have no close modern analogue in the training set.

For the model that uses modern analogues to calculate MAAT, configuration 1 (including all taxa) with the weighted mean of $k = 6$ closest analogues displays the best diagnostics (RMSEP$_{boot}$ = 1.52, r^2_{boot} = 0.78; Table C-2). For the model that uses the PLS method, configuration 2 (excluding taxa that occur < 2 samples and taxa with a maximum abundance of < 1%) with three components displays the best diagnostics (RMSEP$_{boot}$ = 1.59, r^2_{boot} = 0.76; Table C-3). The MA model has slightly better diagnostics than the PLS model. Both models disproportionately overestimate temperatures at the lower end of the MAAT gradient (< 8°C) and the MA model slightly underestimates temperatures > 8°C (Fig. C-3).

Of the 143 fossil samples, 77 (or 54%) contain high abundances (> 5%) of taxa that are poorly modelled (Hill's N2 of < 5 in training set; Table C-1) or taxa that are not included in the model at all; 32 fossil samples (or 22%) contain very high abundances (> 10%) of such taxa (Fig. C-4). The most abundant poorly modelled taxa include: *Hypolepis* (max. abundance 18.3%), *Freycinetia banksia* (max. 4.7%), *Syzygium maire* (max. 4.4%) and

Dodonaea viscosa (max. 3.7%). The most abundant excluded taxa include: *Phymatosorus* (10.9%), *Geum* (4.8%), Ericaceae (max. 4.2%).

Table C-2. Model performance modern analogue technique with weighted averaging. Three configurations were tested: All taxa included (Configuration 1), rare taxa (taxa that occur in < 2 samples with a maximum abundance of < 1%) excluded (Configuration 2), and Pteridophytes (spore plants) excluded (Configuration 3). The number of analogues (k) that produces the lowest RSMEP is indicated in bold.

	k	RMSEP	Average bias	Maximum bias	r^2	δRMSEP (%)
Configuration 1						
	1	1.86	-0.22	1.89	0.6896	
	2	1.62	-0.31	1.34	0.7539	-12.8
	3	1.57	-0.38	1.65	0.7660	-3.4
	4	1.56	-0.37	1.51	0.7661	-0.7
	5	**1.52**	**-0.39**	**1.58**	**0.7772**	**-2.4**
	6	1.59	-0.42	1.83	0.7583	4.4
	7	1.59	-0.46	1.75	0.7608	0.0
	8	1.58	-0.49	1.80	0.7656	-0.3
	9	1.61	-0.50	1.95	0.7592	1.5
	10	1.62	-0.49	2.19	0.7529	1.0
Configuration 2						
	1	1.85	-0.21	1.89	0.6956	
	2	1.62	-0.31	1.34	0.7557	-12.5
	3	1.57	-0.38	1.65	0.7669	-3.2
	4	1.56	-0.38	1.51	0.7666	-0.6
	5	**1.52**	**-0.40**	**1.58**	**0.7764**	**-2.0**
	6	1.59	-0.42	1.83	0.7585	4.0
	7	1.58	-0.45	1.75	0.7618	-0.3
	8	1.58	-0.48	1.80	0.7650	-0.1
	9	1.60	-0.50	1.95	0.7595	1.5
	10	1.63	-0.48	2.19	0.7500	1.5
Configuration 3						
	1	1.76	-0.11	1.43	0.7139	
	2	**1.51**	**-0.26**	**1.71**	**0.7760**	**-14.1**
	3	1.55	-0.32	1.58	0.7640	3.0
	4	1.58	-0.27	1.64	0.7512	1.8
	5	1.58	-0.27	2.01	0.7515	-0.4
	6	1.58	-0.31	2.00	0.7521	0.2
	7	1.58	-0.30	1.99	0.7500	0.2
	8	1.57	-0.34	2.02	0.7553	-0.5
	9	1.59	-0.35	2.05	0.7520	0.9
	10	1.61	-0.36	2.18	0.7456	1.4

Table C-3. Model performance Partial Least Squares technique. Three configurations were tested: All taxa included (Configuration 1), rare taxa (taxa that occur in < 2 samples or with a maximum abundance of < 1%) excluded (Configuration 2), and Pteridophytes (spore plants) excluded (Configuration 3). The number of components that was considered best is indicated in bold.

Configuration	Component	RMSEP$_{boot}$	Average bias	Maximum bias	r^2_{boot}	δRMSEP (%)
Configuration 1						
	1	2.11	0.02	4.15	0.55	-32.2
	2	1.73	0.02	2.84	0.72	-18.1
	3	**1.60**	**0.03**	**3.06**	**0.76**	**-7.4**
	4	1.61	0.00	2.65	0.77	0.4
	5	1.65	-0.02	2.48	0.76	2.9
Configuration 2						
	1	2.10	0.00	4.18	0.55	-32.3
	2	1.72	0.02	2.84	0.72	-18.3
	3	**1.59**	**0.03**	**3.04**	**0.76**	**-7.3**
	4	1.61	0.00	2.65	0.77	0.8
	5	1.65	-0.01	2.45	0.76	3.0
Configuration 3						
	1	2.11	0.00	4.22	0.55	-32.1
	2	1.77	0.02	2.72	0.70	-16.1
	3	**1.62**	**0.03**	**2.91**	**0.75**	**-8.5**
	4	1.61	0.01	2.68	0.76	-0.7
	5	1.64	0.01	2.48	0.76	1.8
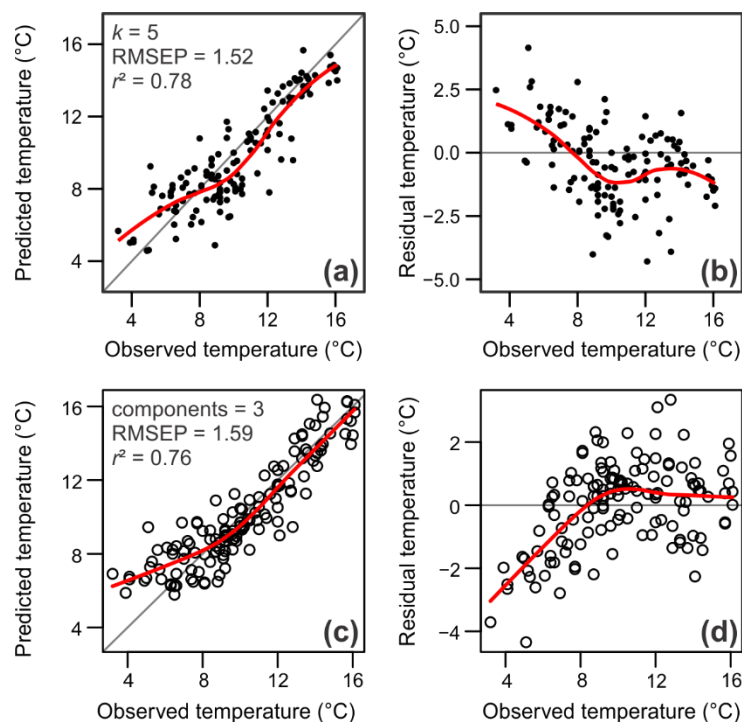
Figure C-3. Model performance: Observed vs. predicted temperatures (a, c) and residual temperatures (b, d) of training set samples modelled using the modern analogue technique (a, b) or partial least squares method (c, d).
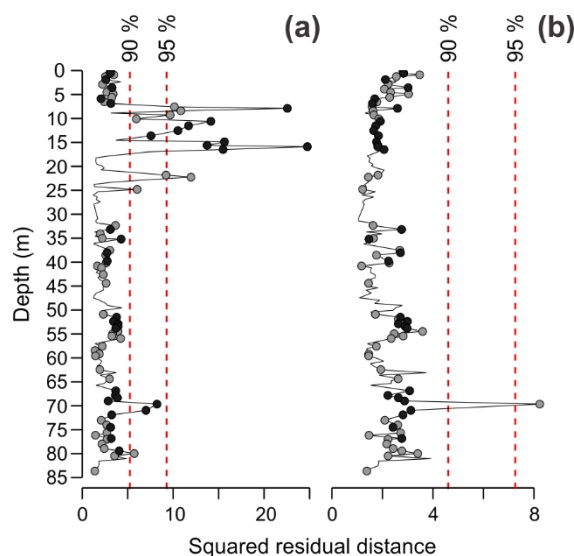
Figure C-4. Goodness of fit of fossil samples over depth for Configuration 1 (all taxa) (a) and Configuration 2 (taxa that occur in < 2 samples or with a maximum abundance of < 1% excluded) (b). The abundance of poorly modelled taxa (Hill's N2 < 5 in training set) plus excluded is indicated with points: grey for values > 5% and black for values > 10%. Vertical, dashed lines represent 90th and 95th percentiles of the distribution of squared residuals of training set samples to MAAT.

If all taxa are included (Configuration 1, used for MA model), 10 fossil samples show a poor fit to MAAT, and 12 samples show a very poor fit (Fig. C-4a). The samples with a very poor fit are mostly the samples with high percentages of *Hypolepis* spores. If taxa that are present in < 2 samples or with a maximum abundance < 1% are excluded from the model (Configuration 2, used for PLS model), only one sample shows a very poor fit to MAAT (Fig. C-4b). This sample, at 69.65 m depth, contains a very high abundance of *Alectryon excelsus* pollen (37.6%).

References for Appendix C1:

Juggins, S. (2015) rioja: Analysis of Quaternary Science Data, R package version 0.9-9. (http://cran.r-project.org/package=rioja).

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Simpson, G.L. (2007). Analogue methods in palaeoecology: Using the analogue package. *Journal of Statistical Software*, 22, 1-29.

Simpson, G.L., Oksanen, J. (2016). analogue: Analogue matching and modern analogue technique transfer function models, R package version 0.17-0. (http://cran.r-project.org/package=analogue)

Wilmshurst, J.M., McGlone, M.S., Leathwick, J.R., Newnham, R.M. (2007). A pre-deforestation pollen-climate calibration model for New Zealand and quantitative temperature reconstructions for the past 18 000 years BP. *Journal of Quaternary Science*, 22, 535-547.

C2: Rare pollen taxa

Table C-4a: Table of rare pollen taxa (occur in fewer than 3 sites or with a maximum abundance of < 1%) that are not included in the pollen diagram (Fig. 5.4). Part 1: Taxa 1 to 30.

Depth (m)	Tall trees			Small trees and shrubs													Herbs													
	Lepidothamnus	*Manoao colensoi*	*Corynocarpus laevigatus*	*Dysoxylum spectabile*	*Quintinia*	*Weinmannia*	*Cordyline*	*Entelea arborescens*	*Fuchsia*	*Geniostoma rupestre*	Hebe type	*Hedycarya arborea*	*Ileostylus micranthus*	*Leptecophylla juniperina*	*Macropiper excelsum*	*Melicytus*	Myrtaceae undifferentiated	*Pennantia corymbosa*	*Pittosporum*	*Pseudopanax rugulate*	*Schefflera digitata*	*Tupeia antarctica*	*Acaena*	Asteraceae liguliflorae	Brassicaceae	*Bulbinella*	Caryophyllaceae	Chenopodiaceae	*Collospermum microsporum*	*Dactylanthus taylorii*
0.215																1.5														
0.550																														
0.885														0.8																
1.250																														
1.875									0.4																					
2.370					0.4	0.4																								
2.870																														
3.580									1.2																					
3.880									0.4		0.4																			
4.535											0.8																			
4.880									1.1					0.7																
5.600									1.1		0.4			0.4																
5.885	0.7																													
6.530																														
6.880														1.1																
7.580		0.4							0.4												0.4									
7.885																														
8.405																														
8.870									0.4												0.4						0.4			
9.260									0.4																					
10.100									0.4							0.4					0.4									
10.600									0.4																					

Table C-4a – continued

Taxon

Depth (m)	Lepidothamnus	Manoao colensoi	Corynocarpus laevigatus	Dysoxylum spectabile	Quintinia	Weinmannia	Cordyline	Entelea arborescens	Fuchsia	Geniostoma rupestre	Hebe type	Hedycarya arborea	Ileostylus micranthus	Leptecophylla juniperina	Macropiper excelsum	Melicytus	Myrtaceae undifferentiated	Pennantia corymbosa	Pittosporum	Pseudopanax rugulate	Schefflera digitata	Tupeia antarctica	Acaena	Asteraceae liguliflorae	Brassicaceae	Bulbinella	Caryophyllaceae	Chenopodiaceae	Collospermum microsporum	Dactylanthus taylorii
	Tall trees		Small trees and shrubs																				Herbs							
11.535																							0.4							
12.530																														
13.620											0.4											0.8								
14.540						0.4				0.4						0.8							0.4							
14.885																							0.4							
15.605		0.4									0.4												0.4	0.4						
15.900																								0.4						
16.490										0.3	1.7											0.3								
16.960																														
17.850											0.4																			
18.220																												1.2		
19.050																												0.4		
19.555	0.4																													
19.875																														
20.375											0.8												0.4					0.4		
20.745											1.2																			
21.245											0.4																			
21.830											0.4												0.4					0.4		
22.255								0.4																				0.4		
22.845																														
23.765																														
24.270											1.1													0.4						
24.815											0.4																			
25.310											1.5													0.4						
25.980						0.4					0.8																	0.4		
26.405											1.2																			0.4

Table C-4a – continued

Depth (m)	Lepidothamnus	Manoao colensoi	Corynocarpus laevigatus	Dysoxylum spectabile	Quintinia	Weinmannia	Cordyline	Entelea arborescens	Fuchsia	Geniostoma rupestre	Hebe type	Hedycarya arborea	Ileostylus micranthus	Leptecophylla juniperina	Macropiper excelsum	Melicytus	Myrtaceae undifferentiated	Pennantia corymbosa	Pittosporum	Pseudopanax rugulate	Schefflera digitata	Tupeia antarctica	Acaena	Asteraceae liguliflorae	Brassicaceae	Bulbinella	Caryophyllaceae	Chenopodiaceae	Collospermum microsporum	Dactylanthus taylorii
			Tall trees						Small trees and shrubs																Herbs					
26.915									0.4		0.4																			
27.400											1.1	0.4																		
27.920																								0.4						
30.955						0.4																								
31.435														0.3										0.3					0.3	0.3
31.845						1.1																								
32.345								1.1														0.4	0.4							
33.180						0.7					0.4																			0.4
33.680											0.4																			
34.090					0.4	0.4				0.4																				
34.570											0.4																			
35.035																														0.4
35.200											0.8													0.4						
36.040																														
36.215								0.4			0.4																			
37.515						0.4																						0.4		0.4
38.045						0.4																								
38.545														0.4						0.7										0.7
39.785								0.4			0.4																			0.4
40.185		0.4									0.4			0.7								0.4								0.4
40.805																														0.3
41.175	0.4												0.4									0.4								
42.180																						0.4								
42.600	1.2										0.4																	0.4		
43.215	0.4									0.4																				
43.715											0.4	0.4																		

205

Table C-4a – continued

Taxon

Depth (m)	Tall trees			Small trees and shrubs																			Herbs							
	Lepidothamnus	*Manoao colensoi*	*Corynocarpus laevigatus*	*Dysoxylum spectabile*	*Quintinia*	*Weinmannia*	*Cordyline*	*Entelea arborescens*	*Fuchsia*	*Geniostoma rupestre*	*Hebe type*	*Hedycarya arborea*	*Ileostylus micranthus*	*Leptecophylla juniperina*	*Macropiper excelsum*	*Melicytus*	*Myrtaceae undifferentiated*	*Pennantia corymbosa*	*Pittosporum*	*Pseudopanax rugulate*	*Schefflera digitata*	*Tupeia antarctica*	*Acaena*	*Asteraceae liguliflorae*	*Brassicaceae*	*Bulbinella*	*Caryophyllaceae*	*Chenopodiaceae*	*Collospermum microsporum*	*Dactylanthus taylorii*
44.415											0.8																0.4			
44.915		0.7				0.4																								
45.380								0.4							0.4							0.8								
45.835																														
46.445		0.4																												
46.870		1.1																				0.4								
47.455		0.3			0.3				0.3													0.3								
47.955																														
48.620		0.4																												
49.040																														
49.545						0.4		0.8			0.4											0.4								
50.070											0.4																			
50.410					0.4	0.4																								0.4
50.900		0.8			0.4				0.4		0.4																			
51.500						1.6					0.8	0.4										1.2								
51.920								0.4	0.4		0.4				0.8							0.4	0.4							
52.375												0.4						0.4	1.2											
52.885						1.2			0.8						0.4															0.4
53.400						0.7			0.7		0.4				0.4				0.4									0.4		
53.825	0.4					0.7			0.4	0.4					0.4							1.1	0.4							
54.470						0.4								0.4								0.4								
54.950								0.4																						
55.440						0.8																								
55.940						1.2																0.4								
57.575																														
58.090																						0.4								

Table C-4a – continued

Taxon

Depth (m)	Tall trees				Small trees and shrubs																		Herbs							
	Lepidothamnus	*Manoao colensoi*	*Corynocarpus laevigatus*	*Dysoxylum spectabile*	*Quintinia*	*Weinmannia*	*Cordyline*	*Entelea arborescens*	*Fuchsia*	*Geniostoma rupestre*	*Hebe* type	*Hedycarya arborea*	*Ileostylus micranthus*	*Leptecophylla juniperina*	*Macropiper excelsum*	*Melicytus*	Myrtaceae undifferentiated	*Pennantia corymbosa*	*Pittosporum*	*Pseudopanax* rugulate	*Schefflera digitata*	*Tupeia antarctica*	*Acaena*	Asteraceae liguliflorae	Brassicaceae	*Bulbinella*	Caryophyllaceae	Chenopodiaceae	*Collospermum microsporum*	*Dactylanthus taylorii*
58.470		0.4																												
59.180						0.7																								
59.580															0.3							1.0								
60.220											1.9												0.4							
60.690									0.4		0.9																			
61.190					0.8						0.4																		0.4	
61.690							0.4																0.4		0.4					
62.450					0.4																								0.4	
62.610																														
63.100						1.5																								
63.605					0.4				0.4																					
64.375						0.4					1.3							0.4							0.8					
64.775											0.4				0.4															
65.310									0.4	0.4																				
65.710																														
66.870					0.4	0.4			0.4																					0.4
67.830									0.8	0.4		1.2																		
68.330					0.8	1.1			1.1		0.4																			0.4
69.000	0.4				0.4	0.4	0.8						0.8					2.3	0.4			0.4								
69.650																														
70.950					0.4	0.4			1.9		0.8	1.9	0.4		0.4															
71.905						1.6					0.4	0.4																		
73.020															0.4															
73.990						0.4			0.4		0.8																			
74.500						1.2								0.4																
75.655											0.4																			

Table C-4a – continued

Depth (m)	Lepidothamnus	Manoao colensoi	Corynocarpus laevigatus	Dysoxylum spectabile	Quintinia	Weinmannia	Cordyline	Entelea arborescens	Fuchsia	Geniostoma rupestre	Hebe type	Hedycarya arborea	Ileostylus micranthus	Leptecophylla juniperina	Macropiper excelsum	Melicytus	Myrtaceae undifferentiated	Pennantia corymbosa	Pittosporum	Pseudopanax rugulate	Schefflera digitata	Tupeia antarctica	Acaena	Asteraceae liguliflorae	Brassicaceae	Bulbinella	Caryophyllaceae	Chenopodiaceae	Collospermum microsporum	Dactylanthus taylorii
Taxon	Tall trees		Small trees and shrubs														Herbs													
76.160																								0.4						
76.835						1.1						0.4									0.8									
76.995																						0.8								0.4
77.970										0.8																	0.4			
78.430									0.4				0.4																	
78.940				1.1							0.4																			
79.435									0.4																					
79.975									0.4			1.2																		
80.475					0.8				0.4			0.4	0.4														0.4			
81.055					0.4	0.4			1.1	0.4																				
81.550																														
82.065																												0.4		
82.565		0.8																												
83.185																														
83.620																														

Table C-4b: Table of rare pollen taxa (occur in fewer than 3 sites or with a maximum abundance of < 1%) that are not included in the pollen diagram (Fig. 5.4). Part 2: Taxa 31 to 60.

Taxon — Herbs cont'd · Wetland taxa · Ferns and allies

Depth (m)	Epilobium	Euphrasia	Fabaceae	Freycinetia banksii	Geranium	Gunnera	Mentha cunninghamii	Passiflora tetrandra	Phormium	Plantago	Potentilla anserinoides	Ranunculaceae	Ranunculus type	Ripogonum scandens	Rumex	Sebaea ovata	Urtica	Wahlenbergia	Xeronema callistemon	Callitriche	Drosera arcturi	Potamogeton	Typha/Sparganium	Blechnum	Hymenophyllum	Lycopodium scariosum	Lycopodium unidentified	Phlegmariurus varius	Pteris	Trichomanes
0.215				4.7				0.4																						
0.550								0.3					0.7																0.3	
0.885													1.2							0.8					1.2					0.4
1.250									0.4				0.8																	
1.875																				0.7										
2.370								0.8					0.4																	
2.870																												0.4		
3.580													0.4	0.8														0.8	0.8	
3.880																							0.4							
4.535								0.4																				0.4		
4.880																							0.4					0.4		
5.600								0.4					0.4																0.4	
5.885								0.4																						
6.530													0.4																0.4	
6.880											0.4												0.4							
7.580													0.7									0.4								
7.885																							0.3							
8.405									0.4				0.8															0.4		
8.870																														
9.260													0.8																	
10.100								0.4					0.4																	
10.600						0.4		0.4																		0.7				
11.535	0.4					1.2			0.4				0.4															1.6	0.4	
12.530						0.7			0.4	0.4																		0.4		

Table C-4b – continued

	Herbs cont'd																			Wetland taxa				Ferns and allies						
Depth (m)	Epilobium	Euphrasia	Fabaceae	Freycinetia banksii	Geranium	Gunnera	Mentha cunninghamii	Passiflora tetrandra	Phormium	Plantago	Potentilla anserinoides	Ranunculaceae	Ranunculus type	Ripogonum scandens	Rumex	Sebaea ovata	Urtica	Wahlenbergia	Xeronema callistemon	Callitriche	Drosera arcturi	Potamogeton	Typha/Sparganium	Blechnum	Hymenophyllum	Lycopodium scariosum	Lycopodium unidentified	Phlegmariurus varius	Pteris	Trichomanes
13.620									0.8																		2.3	0.4	0.4	
14.540													0.8																	
14.885	0.4				0.4				0.4																			0.4	0.7	
15.605	0.4					0.8		0.4																				0.4		
15.900									0.4				0.4														1.2	1.2		
16.490						0.3			0.7				0.3															0.7		
16.960									0.8																					
17.850						0.4			0.4				0.4																0.4	
18.220						0.4			0.4																					
19.050													0.4																	
19.555									0.4						0.8													0.8		
19.875									0.4																				0.8	
20.375																					0.4							0.4		
20.745																														
21.245							0.4											0.4								0.8		0.4		
21.830																														
22.255			0.4																											
22.845													0.7																	
23.765									0.4				0.4																	
24.270	0.4								1.1																					
24.815									0.4				0.4															0.4		
25.310																												0.4		
25.980									0.4																					
26.405									0.4																					
26.915																														
27.400																										0.4				

Table C-4b – continued

Taxon

Depth (m)	Herbs cont'd																			Wetland taxa				Ferns and allies						
	Epilobium	*Euphrasia*	*Fabaceae*	*Freycinetia banksii*	*Geranium*	*Gunnera*	*Mentha cunninghamii*	*Passiflora tetrandra*	*Phormium*	*Plantago*	*Potentilla anserinoides*	*Ranunculaceae*	*Ranunculus* type	*Ripogonum scandens*	*Rumex*	*Sebaea ovata*	*Urtica*	*Wahlenbergia*	*Xeronema callistemon*	*Callitriche*	*Drosera arcturi*	*Potamogeton*	*Typha/Sparganium*	*Blechnum*	*Hymenophyllum*	*Lycopodium scariosum*	*Lycopodium* unidentified	*Phlegmariurus varius*	*Pteris*	*Trichomanes*
27.920																														
30.955																												0.8		
31.435									0.6																					
31.845													0.4																	
32.345																							0.4							
33.180													0.7				0.4											0.7		
33.680																														
34.090																												0.4		
34.570																														
35.035					0.4																								0.4	
35.200																														
36.040									0.7																					
36.215									0.8																		0.4	0.4		
37.515													0.4															0.4		
38.045													0.8								0.4							0.8		
38.545									0.4																					
39.785																														
40.185									0.7																0.4			0.4		
40.805									0.3								0.3					0.3							0.3	
41.175									0.4																			0.4	0.8	
42.180													0.7																	
42.600									0.4																					
43.215													0.8																	
43.715													0.4																	
44.415									0.4				0.8															0.4		
44.915									0.4				0.7																	

Table C-4b – continued

Taxon

Herbs cont'd   Wetland taxa   Ferns and allies

Depth (m)	Epilobium	Euphrasia	Fabaceae	Freycinetia banksii	Geranium	Gunnera	Mentha cunninghamii	Passiflora tetrandra	Phormium	Plantago	Potentilla anserinoides	Ranunculaceae	Ranunculus type	Ripogonum scandens	Rumex	Sebaea ovata	Urtica	Wahlenbergia	Xeronema callistemon	Callitriche	Drosera arcturi	Potamogeton	Typha/Sparganium	Blechnum	Hymenophyllum	Lycopodium scariosum	Lycopodium unidentified	Phlegmariurus varius	Pteris	Trichomanes
45.380									0.4																					
45.835													0.8																	
46.445													0.8																	
46.870				0.4					0.4																					
47.455													0.3																0.3	
47.955													0.8	0.4																
48.620								0.8																						
49.040													1.1																	
49.545								1.2																						
50.070													1.2																	
50.410													0.4															0.4	0.4	
50.900													0.4																0.4	
51.500								0.4					0.8															0.4	0.4	
51.920													1.1															1.1	0.8	
52.375								0.4					0.8					0.4										0.8	0.4	
52.885													1.9															0.8		
53.400																													0.4	
53.825													1.8	0.4														0.4	1.5	
54.470													1.2															0.4	0.4	
54.950													0.4															0.4	0.4	
55.440							0.4						1.2		0.4													0.8		
55.940													2.0																	
57.575								0.4				1.2																		
58.090																			0.4											
58.470																														
59.180																														

Table C-4b – continued

Taxon — Herbs cont'd (*Epilobium* … *Xeronema callistemon*) | Wetland taxa (*Callitriche* … *Typha/Sparganium*) | Ferns and allies (*Blechnum* … *Trichomanes*)

Depth (m)	*Epilobium*	*Euphrasia*	*Fabaceae*	*Freycinetia banksii*	*Geranium*	*Gunnera*	*Mentha cunninghamii*	*Passiflora tetrandra*	*Phormium*	*Plantago*	*Potentilla anserinoides*	*Ranunculaceae*	*Ranunculus* type	*Ripogonum scandens*	*Rumex*	*Sebaea ovata*	*Urtica*	*Wahlenbergia*	*Xeronema callistemon*	*Callitriche*	*Drosera arcturi*	*Potamogeton*	*Typha/Sparganium*	*Blechnum*	*Hymenophyllum*	*Lycopodium scariosum*	*Lycopodium* unidentified	*Phlegmariurus varius*	*Pteris*	*Trichomanes*
59.580																														
60.220																														
60.690																													0.4	
61.190						0.8								0.8																
61.690																														
62.450																													0.4	
62.610													0.4																	
63.100																														
63.605																														
64.375								0.4					0.4			0.4														
64.775			0.4					0.4					0.4																	
65.310																													0.4	
65.710							0.4																					0.4		
66.870								0.4						1.5					1.1											
67.830													0.4																	
68.330								1.1					1.1	0.4														0.8	0.4	
69.000													0.4										0.4							
69.650									0.4																			0.8	0.4	
70.950								0.4																			0.4		0.4	0.4
71.905																		0.4												
73.020																														
73.990																														
74.500								0.8					0.4																	
75.655																													0.4	0.4
76.160																														
76.835													0.8																	

Table C-4b – continued

Depth (m)	*Epilobium*	*Euphrasia*	*Fabaceae*	*Freycinetia banksii*	*Geranium*	*Gunnera*	*Mentha cunninghamii*	*Passiflora tetrandra*	*Phormium*	*Plantago*	*Potentilla anserinoides*	*Ranunculaceae*	*Ranunculus type*	*Ripogonum scandens*	*Rumex*	*Sebaea ovata*	*Urtica*	*Wahlenbergia*	*Xeronema callistemon*	*Callitriche*	*Drosera arcturi*	*Potamogeton*	*Typha/Sparganium*	*Blechnum*	*Hymenophyllum*	*Lycopodium scariosum*	*Lycopodium unidentified*	*Phlegmariurus varius*	*Pteris*	*Trichomanes*
76.995													0.4																0.4	
77.970														0.4					0.4										0.4	
78.430											0.4																			
78.940																														
79.435								1.1	0.4										1.5		0.4									
79.975																														
80.475								0.4																						
81.055								0.4																					0.7	
81.550																													0.4	
82.065													0.4														0.4			
82.565													0.8																	
83.185																												0.4		
83.620																												0.4	0.8	

Appendix D: Non-pollen palynomorphs

D1: Introduction

Non-pollen palynomorphs (NPP) are a variety of microfossil remains that can be found in palynological slides but are not pollen. The study of NPP was introduced by van Geel (1978), who started describing the non-pollen remains he found in Holocene peat bogs in Germany and the Netherlands and numbered the different morphotypes. Since then, many more types have been described and added to the list that now counts more than 1000 types from all over the world (Hooghiemstra, 2012; Gelorini et al., 2011). NPP can basically be anything, as long as they are found in pollen slides. The most commonly recorded NPP are fungal spores, but other types are also common, such as plant remains other than known pollen and spores, algal remains, cyanobacteria, animal remains and amoeba.

The indicative value of NPP is debatable. There are several limitations to their usefulness, i.e. lack of identification accuracy (remains are generally attributed to 'types' that may be more or less ambiguous), lack of linkage between fossil data and modern living organisms and their ecology, and the selective preservation of taxa and material (e.g. ascospores are strongly overrepresented) (van Geel & Aptroot, 2006). Selective preservation may be exaggerated even more due to different preparation methods in different labs. Samples from different sites may not be comparable to each other if they were prepared using a different protocol, e.g. different sieve mesh sizes. The first two limitations could be (partially) overcome if NPP were recorded systematically by palynologists. In New Zealand in particular, many palynologists simply ignore NPP.

In spite of their limitations, NPP have provided valuable palaeoecological information in different parts of the world. Fungal spores, for instance, can provide information about the presence of their substrate (including host plants, decaying organic matter and dung) as well as local conditions (e.g. wetness; Cugny et al., 2010). Not all fungi are indicative of specific conditions, but the ascospores of coprophilous ("dung-loving") fungi in particular haven been used extensively to indicate the presence or absence of large herbivores, whether wild of domesticated (e.g. van Geel et al., 2003; Bos et al., 2005; Ghosh et al., 2017). Apart from coprophilous fungi, the assignment of fungal remains to litter, dung or wood related fungi and ericoid mycorrhizal fungi can be helpful in finding patterns of change as well (van Geel,

2002; Mulder et al., 2003). Another important group of NPP is comprised of algal remains, mainly green algae of the family Zygnemataceae and colonial green algae genera *Pediastrum* and *Botryococcus*, which can be indicative of local limnological conditions (van Geel, 2001).

Here, I describe the NPP that were found in the lake sediments of Orakei Basin and compare them to known types from other regions. Next, I apply redundancy analysis (RDA) to explore the indicative value of the types found at Orakei Basin in conjunction with vegetation classes (e.g., trees, shrubs, herbs).

D2: Methods

147 samples from Orakei Basin were prepared according to standard pollen preparation methods. See Chapter 5 for information about Orakei Basin, its sediments, the distribution of samples and their preparation. NPP were recorded during regular pollen counting. Any remains that seemed distinct in their morphology were recorded initially, but only types that are similar to existing types or that occur in more than one sample were retained.

Orakei Basin NPP were numbered with the prefix "VUW" for Victoria University of Wellington, following the convention of previous studies (e.g. "HdV" for Hugo de Vries Laboratory, University of Amsterdam, or "UG" for Universiteit Gent, etc.). I numbered every type, even if they resemble previously described types, because it is difficult to verify that the taxa recorded here are the same as in other regions, especially considering the isolation of New Zealand from the rest of the world.

For comparison of the VUW-types to existing types, I preferentially referenced papers on HdV-types (van Geel, 1978; van Geel et al., 1981; van Geel et al., 1989). Miola (2012) compiled a list of HdV types from HdV-1 to -1402 (described between the years 1972 and 2011), which cites the first publication on each type. Additional important NPP publications include Cugny et al. (2010), who studied NPP in modern surface samples in the Basque mountains (France) and use the prefix "TM" (University of Toulous-le Mirail); Gelorini et al. (2011), who studied NPP in east African lake sediments and use the prefix "UG"; and Revelles et al. (2016), who studied a Neolithic settlement at La Draga (Spain) and use the prefix "UAB" (Universitat Autònoma de Barcelona).

NPP abundances were expressed as a percentage of the pollen sum (see Chapter 5). Redundancy analysis on the square-root transformed NPP percentages, constrained by the abundance of the main pollen groups (i.e. "Tall trees", "Beech trees", "Small trees and shrubs" and "Herbs") as explanatory variables, was used to assess the indicative value of the different types and their relation to each other.

D3: Results and discussion

D3.1 Recorded types

I recorded 37 distinct NPP types (and a few subtypes) in the Orakei Basin samples (Table D-1; Fig. D-1). The five algal types (VUW-1 to -5) could be assigned to genera and correspond to existing HdV types. Of the 30 fungal types (VUW-6 to -35), 24 were equivalent or similar to existing types. There are two types of unknown origin (VUW-36 and -37), i.e. it is unclear whether they are animal, fungal, algal or derive from plants. Type VUW-16 is divided into four subtypes (A-D); this multi-celled fungal spore is generally found with five cells (VUW-16C), but also occurs with three (A), four (B) or six (D) cells. Type VUW-25 has five subtypes (A-E); these multi-celled conidia (for explanation of terminology, see Table D-2) look similar, but the shape of the cells varies. During the first half of this study, these subtypes were not distinguished and were all counted as VUW-25.

D3.1.1 Algal remains

Type VUW-1: *Botryococcus*
Colonial green algae with densely-packed conical cells radiating and branching from the centre of the roughly-spherical colony. Can be present in a wide range of environments, is mainly associated with freshwater habitats but has also been recorded in high abundances in saline habitats (van Geel, 2001). Often co-occurs with *Pediastrum* and may indicate extreme conditions (e.g. very cold, clear, oligotrophic waters) if abundant in the absence of *Pediastrum* (Jankovská & Komárek, 2000).

Table D-1. NPP types recorded at Orakei Basin. See Table D-2 for terminology.

Type	Category	Subcategory	Name[1]	(Potential) equivalent type(s)	Reference(s)
VUW-1	Algal	Colonial green algae	*Botryococcus*	HdV-766	Jankovská & Komárek, 2000; van Geel, 2002
VUW-2	Algal	Colonial green algae	*Pediastrum*	HdV-760	Jankovská & Komárek, 2000; van Geel, 2002
VUW-3	Algal	Zygospores (Zygnemataceae)	*Zygnema* type	HdV-213; HdV-314	van Geel et al., 1989; van Geel, 2002
VUW-4	Algal	Zygospores (Zygnemataceae)	*Spirogyra*	HdV-212	van Geel et al., 1989; van Geel, 2002
VUW-5	Algal	Zygospores (Zygnemataceae)	*Mougeotia*	HdV-313	van Geel et al., 1981; van Geel, 2002
VUW-6	Fungal	Ascospores (Sordariales)	Sordariaceae	HdV-55B; UAB-55	van Geel, 1978; Cugny et al., 2010; Revelles et al., 2016
VUW-7	Fungal	Ascospores (Sordariales)	*Podospora*	HdV-368	van Geel et al., 1981; Cugny et al., 2010
VUW-8	Fungal	Ascospores (Sordariales)	*cf. Cercophora*	HdV-112; HdV-1013; UAB-33	van Geel et al., 1981; van Geel et al., 2011; Gelorini et al., 2011; Revelles et al., 2016
VUW-9	Fungal	Ascospores (Sordariales)	*cf. Neurospora*	HdV-55C; UG-1187	van Geel, 1978; Gelorini et al., 2011
VUW-10	Fungal	Ascospores (Xylariaceae)		UG-1065	Gelorini et al., 2011
VUW-11	Fungal	Ascospores (Xylariaceae)	*cf. Rosellinia*	UAB-11; UG-1174	Gelorini et al., 2011; Revelles et al., 2016
VUW-12	Fungal	Ascospores (Xylariaceae)	*Kretzschmaria deusta*	HdV-44	van Geel, 1978; Cugny et al., 2010; van Geel et al., 2013
VUW-13	Fungal	Ascospores (Xylariaceae)		HdV-1052	van Geel et al., 2011; Gelorini et al., 2011
VUW-14	Fungal	Ascospores	*cf. Coniochaeta*	TM-211; UAB-9	Cugny et al., 2010; Revelles et al., 2016
VUW-15	Fungal	Ascospores		UAB-10,	Revelles et al., 2016
VUW-16	Fungal	Ascospores	*cf. Meliola*	UG-1113	Gelorini et al., 2011
VUW-17	Fungal	Ascospores	*Meliola cf. niessliana*	HdV-14a; UG-1137	van Geel, 1978; Gelorini et al., 2011
VUW-18	Fungal	Ascospores	*cf. Sporormiella*	HdV-113	van Geel et al., 2003
VUW-19	Fungal	Ascospores		UAB-41A-C; UG-1106, 1155	Gelorini et al., 2011; Revelles et al., 2016

218

Table D-1 – continued

Type	Category	Subcategory	Name[1]	(Potential) equivalent type(s)	Reference(s)
VUW-20	Fungal	Ascospores	*cf. Delitschia* spp.	UAB-34A, B; UG-1066	Gelorini et al., 2011; Revelles et al., 2016
VUW-21	Fungal	Ascospores		UAB-35	Revelles et al., 2016
VUW-22	Fungal	Chlamydospores	*Glomus*	HdV-207	van Geel et al., 1989
VUW-23	Fungal	Conidia	*Tetraploa aristata*	HdV-89	van Geel, 1978; Gelorini et al., 2011
VUW-24	Fungal	Conidia	*Helicoon pluriseptatum*	HdV-30	van Geel, 1978
VUW-25	Fungal	Conidia	*cf. Bacterodesmium* spp.	cf.HdV-569; TM-011; UAB-18A; UG-1091	Cugny et al., 2010; Gelorini et al., 2011; Revelles et al., 2016
VUW-26	Fungal	Fungal spores		UAB-39	Revelles et al., 2016
VUW-27	Fungal	Fungal spores		UAB-7, 8, 15, 48, 49	Revelles et al., 2016
VUW-28	Fungal	Fungal spores		UAB-50	Revelles et al., 2016
VUW-29	Fungal	Fungal spores			
VUW-30	Fungal	Fungal spores			
VUW-31	Fungal	Fungal spores			
VUW-32	Fungal	Fungal spores			
VUW-33	Fungal	Fungal spores		UG-1107	Gelorini et al., 2011
VUW-34	Fungal	Fungal spores			
VUW-35	Fungal	Fungal spores			
VUW-36	Unknown	Unknown		HdV-91	van Geel, 1978
VUW-37	Unknown	Unknown		UG-1286	Gelorini et al., 2011

[1] If known, types are named; if name is suspected but uncertain, abbreviation *cf.* ("conferatur", or compare) is used

Type VUW-2: *Pediastrum*

Radially-symmetrical colonial green algae with the outermost cells displaying one or two horns. Indicates a wide range of ecological responses (van Geel, 2001), possibly because several different species are lumped together and are usually not distinguished. However, different species can be morphologically distinct, and identification to a lower taxonomic level may lead to more useful ecological interpretation (Jankovská & Komárek, 2000).
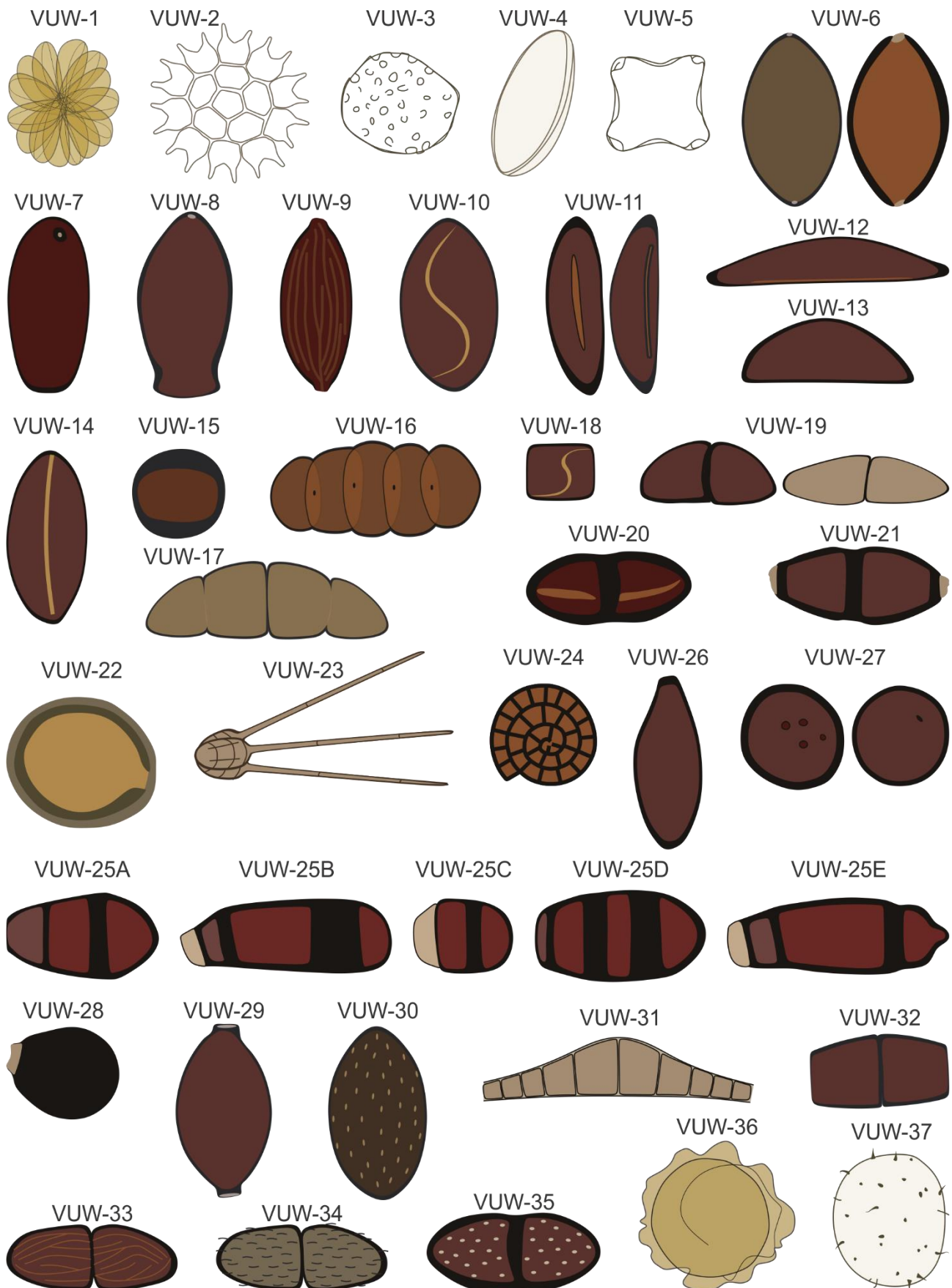
VUW-1 VUW-2 VUW-3 VUW-4 VUW-5 VUW-6

VUW-7 VUW-8 VUW-9 VUW-10 VUW-11

VUW-12

VUW-13

VUW-14 VUW-15 VUW-16 VUW-18 VUW-19

VUW-17 VUW-20 VUW-21

VUW-22 VUW-23 VUW-24 VUW-26 VUW-27

VUW-25A VUW-25B VUW-25C VUW-25D VUW-25E

VUW-28 VUW-29 VUW-30 VUW-31 VUW-32

VUW-36 VUW-37

VUW-33 VUW-34 VUW-35

Figure D-1. Images of VUW NPP types. Magnified ~1000x, except Type VUW-23, which is magnified ~500x.

Table D-2. Terminology

Term	Meaning
Ascospore	Fungal spore formed through meiosis followed by mitosis, specific to fungi classified as Ascomycetes
Chlamydospore	Fungal resting spore, usually the result of asexual reproduction
Conidium	Asexual, non-motile fungal spore generated through the process of mitosis. Type of chlamydospore
Coprophilous	"Dung-loving"; coprophilous fungi grow on animal dung
Endomycorrhizal	Pertains to the symbiotic association between a fungus and the roots of a vascular host plant, where the fungus penetrates the cortical cells of the roots
Hyaline	Composed of a clear substance, produced by the degeneration of connective tissues
Hyphen	Long, branching filamentous structure and the main mode of vegetative growth of a fungus
Zygospore	Diploid resting spore of algae or fungi. Chlamydospores are a Type of zygospore.

Type VUW-3: *Zygnema* type

Pitted zygospore, shape heterogenous, spheroidal to flattened. In temperate climates, Zygnemataceae produce their spores during spring in shallow, relatively warm water. Within Zygnemataceae, few ecological differences are known, but different types may indicate different trophic conditions (van Geel et al., 2002). *Zygnema* is possibly associated with high trophic state and/or increased run off (van Geel et al., 1989).

Type VUW-4: *Spirogyra*

Ellipsoidal zygospore with longitudinal furrow encircling almost the entire spore. See Type VUW-3 for Zygnemataceae ecology.

Type VUW-5: *Mougeotia*

Square zygospore with depressions in the centre of the angles. See Type VUW-3 for Zygnemataceae ecology. *Mougeotia* possibly indicates shallow, mesotrophic to eutrophic open water.

D3.1.2 Fungal ascospores

Type VUW-6: Sordariaceae

Spores one-celled, ellipsoidal, brown, smooth-walled, with two protruding apical pores. First found in the somewhat mesotrophic part of the Engbertsdijksveen section (van Geel, 1978).

Associated with coprophilous fungi (Cugny et al., 2010), but probably not restricted to dung substrate.

Type VUW-7: *Podospora*

One-celled, ellipsoidal, smooth, brown, with one protruding pore (surrounded by annulus) directly below the apex. Basal end bluntly conical. Coprophilous (Lundqvist 1972) and often encountered in archaeological context (van Geel & Aptroot, 2006).

Type VUW-8:

Truncate at basal side and tapering at the apical end, with one subapical pore. Possibly *Cercophora* sp. (van Geel et al., 1981) or *Cercophora* type (van Geel et al., 2011), which is known from woody substrates, herbaceous stems and leaves and dung (Lundqvist, 1972).

Type VUW-9:

One-celled, ellipsoid, brown, with longitudinal sub-parallel striae running over the entire spore-length; with two protruding apical pores. Similar to UG-1187 (Gelorini et al., 2011) and superficially resembles *Neurospora*, which is associated with fire and charred remains (van Geel, 1978).

Type VUW-10:

One-celled, slightly ellipsoid, dark brown, smooth, thick-walled and with tapering ends; germ slit sigmoidal and running over the entire spore-length. Similar to UG-1065 (Gelorini et al., 2011), probably a member of the Xylariaceae, which are widespread in temperate and tropical regions and are known as wood-rotting fungi, inhabitants of dung or litter, and pathogens of a range of plants (Gelorini et al., 2011).

Type VUW-11:

One-celled, fusiform, dark brown, smooth, with tapering ends and straight germ slit running over entire spore length near the flattened side. Similar to UG-1174 (Gelorini et al., 2011), which was identified as *Rosellinia* sp., a genus in the Xylariaceae (see Type VUW-10 for ecology).

Type VUW-12: *Kretzschmaria deusta*

One-celled, fusiform with flattened side, dark brown, smooth, wall thickened at the ends, and flattened side bearing a longitudinal germ slit. Tree pathogen that causes soft-rot of wood of living, deciduous trees (van Geel et al., 2013).

Type VUW-13:

One-celled, ellipsoid with flattened side, brown, smooth, thick-walled, with tapering ends and straight, longitudinal germ slit on flattened side. Similar to HdV-1052, a member of the Xylariaceae (see Type VUW-10 for ecology).

Type VUW-14:

One-celled, ellipsoid, brown, smooth, with straight, longitudinal germ slit. Similar to TM-211 (Cugny et al., 2010), who attributed this spore type to *Coniochaeta*; in the Basque mountains it was most common in fern fallows and heathland.

Type VUW-15:

One-celled, globose to ellipsoidal, dark brown with a longitudinal light brown zone. Similar to UAB-10 (Revelles et al., 2016), who suggest it might be a member of the Coniochaetaceae.

Type VUW-16:

Oblong, smooth, usually 5-celled (each cell globose), constricted at the septa, with one pore in each septum. Subcategories: VUW-16A: 3-celled, 16B: 4-celled, 16C: 5-celled, 16D: 6-celled. Similar to Type UG-1113 (Gelorini et al., 2011), which was attributed to *Meliola*, which are found as parasites on leaves and stems of a wide range of hosts.

Type VUW-17: *Meliola cf. niessliana*

Oblong, flattened on one side, 4-celled, slightly constricted at septa. Meliolaceae are obligate parasites on green plants (van Geel, 1978).

Type VUW-18:

Small, rounded-rectangular cell, with sigmoidal germ slit. Only encountered once (Fig. A.V-2). Resembles a single cell of *Sporormiella* (HdV-113), which is coprophilous (Lundqvist,1972).

Type VUW-19:

Two-celled, symmetrical ellipsoid, smooth, constricted at septum, with tapering ends. Probably encompasses many genera and species. Similar to e.g. Type UG-1106, UG-1155 (Gelorini et al., 2011), and UAB-41 (Revelles et al., 2016).

Type VUW-20:

Two-celled, ellipsoid, smooth, dark brown, constricted at septum, each cell with germ slit parallel to the long axis of the spore. May include *Delitschia* spp. (Gelorini et al., 2011; Revelles et al., 2016), which are mostly coprophilous.

Type VUW-21:

Four-celled, two angular, brown central cells and two hyaline cells (one at each end), but hyaline cells often not preserved. Similar to Type UAB-35 (Revelles et al., 2016).

D3.1.3 Other fungal spores

Type VUW-22: *Glomus*

Globose spore with thick wall and hyphal attachment. *Glomus* is an endomycorrhizal fungus that occurs on a variety of host plants (van Geel et al., 1989).

Type VUW-23: *Tetraploa aristata*

Verrucose conidia consisting of three columns, each column terminating in a septate appendage. *Tetraploa aristata* is a widespread fungus found of leaf bases and stems just above the soil. Recorded on a variety of Poaceae, Cyperaceae and Juncaceae (van Geel, 1978).

Type VUW-24: *Helicoon pluriseptatum*

Helically coiled conidia, multi-septate, with individual cells almost square. *Helicoon pluriseptatum* has previously been found in peat bogs on decaying plant material (van Geel, 1978).

Type VUW-25:

Unequally and asymmetrically multi-celled (5 subtypes differ in shape and number of cells), smooth, thick-walled, with dark septa. This type and subtypes may include *Bacterodesmium*

species (e.g. Type UG-1091; Gelorini et al., 2011), *Trichocladium* species (e.g. Type TM-011, similar to HdV-569; Cugny et al., 2010), and probably others. *Bacterodesmium* grows on wood and bark of various deciduous trees (Gelorini et al., 2011).

- Subtype VUW-25A: ellipsoid to club-shaped, 3-celled, more rounded at apical end, more pointed at basal end. Sometimes the remains of a basal hyaline cell are visible.
- Subtype VUW-25B: ellipsoid to oblong, 4-celled, apical cell short and rounded, second cell elongate.
- Subtype VUW-25C: ellipsoid to egg-shaped, shorter than other subtypes, 3-celled, apical cell rounded, basal cell subhyaline. Similar to Type UG-1084 (Gelorini et al., 2011).
- Subtype VUW-25D: ellipsoid, 4-celled, rounded at apical end. Very similar to Type UG-1091 (Gelorini et al., 2011) and Type UAB-18A (Revelles et al., 2016).
- Subtype VUW-25E: ellipsoid to oblong, 4-celled, similar to Subtype VUW-25B with elongate second cell, but apical cell has a nipple instead of being rounded.

Type VUW-26:

One-celled, ellipsoidal, smooth, brown, one end rounded and one end flattened, with protruding pore. Similar to Type UAB-39, which was found in waterlogged sedimentary layers at an Early Neolithic-age settlement (Revelles et al., 2016).

Type VUW-27:

One-celled, globular spores, some with one or more pores. This type likely encompasses many genera and species and it may be possible to distinguish some (e.g. based on number of pores) that were lumped together in this ambiguous type. Some similar types include: UAB-7, 8, 15 and 48 (Revelles et al., 2016).

Type VUW-28:

One-celled, round to ellipsoidal, smooth, dark brown, truncated at one end with hyaline remains. Similar to Type UAB-50, which was found in peaty layers at an Early Neolithic-age settlement (Revelles et al., 2016).

Type VUW-29:

One-celled, ellipsoidal, dark brown, smooth-walled, with two big, protruding apical pores. Similar to Sordariaceae (VUW-6) but pores bigger and more protruding.

Type VUW-30:

One-celled, ellipsoidal, dark brown, with many small pores covering entire spore.

Type VUW-31:

Multi-celled (10 to 13 cells), symmetrical, elongate spore, thin walled, individual cells angular, biggest in the middle and reducing in size toward ends, with cells at each end hyaline.

Type VUW-32:

Two-celled, symmetrical oblong with flat ends, smooth, dark brown, constricted at septum, individual cells angular.

Type VUW-33:

Two-celled, ellipsoid with rounded ends, striate, brown, constricted at septum. Similar to Type UG-1107 (Gelorini et al., 2011)?

Type VUW-34:

Two-celled, ellipsoid with rounded ends, brown, constricted at septum, with small spines or echinae over surface.

Type VUW-35:

Two-celled, ellipsoid with rounded ends, brown, constricted at septum, with small pores puncturing the wall.

D3.1.4 Unknowns

Type VUW-36:

Globose microfossil, with multi-layered wall, each layer thin and loosely arranged. Similar to Type HdV-91 (van Geel, 1978)?

Type VUW-37:

Globose, yellow, with small spines or appendages of uneven size and shape, irregularly arranged across the surface. Similar to Type UG-1286 (Gelorini et al., 2011)?

D3.2 Orakei Basin NPP record and ordination

Recorded NPP do not show coherent patterns that relate to the five pollen zones (Fig. D-2; see Chapter 5 for details about the pollen zones). Many occur only sporadically. Broadly speaking, fungal spores are relatively more abundant when algal remains are less abundant. Unknown Type VUW-37 seems most abundant when VUW-1 (*Botryococcus*) is also present. Several fungi show high abundances in (but are not restricted to) the sediments relating to Lithology 9, which is dark, banded sand and silt (see Chapter 5), i.e. Types VUW-6 (Sordariaceae), 10, 11, 12 (*Kretzschmaria deusta*), 13, 14, 16, 25A, 25B and 27, whereas algae are virtually absent at these depths.

Turning to the RDA, the abundances of tall trees, beech trees, small trees and shrubs, and herbs together explain 9.5% of the variance in the NPP dataset (Table D-3). A large proportion remains unexplained. The first unconstrained, principle component axis has a higher eigenvalue than the first constrained axis (3.78 vs. 2.54). This implies that NPPs do not necessarily respond to the same environmental variables as the dryland vegetation. Possibly NPPs are more sensitive to local conditions such as lake conditions or surrounding soils.

The RDA biplot (Fig. D-3) shows that the presence of tall trees vs. beech and open vegetation explains most of the variance in the NPP dataset (of the variables tested). This dichotomy is associated with temperature, although the presence and absence of the fungal NPP types is more likely related to the presence of their substrate or host. Fungal spores are mostly correlated to tall trees, whereas algae show a relationship with beech trees, although the association between algae and beech is likely related to a confounding variable causing increasing lake productivity or decreasing terrestrial input into the lake. The negative correlation between fungal spores and algae as identified in the NPP percentage record (Fig. D-2) is visible in the biplot as well. NPP Types VUW-11 (Xylariaceae), 12 (*Kretzschmaria deusta*), 14 and 19 correlate most strongly with tall trees. VUW-11 (Xylariaceae) is associated with decaying wood (Mulder et al., 2003) and VUW-12 (*Kretzschmaria deusta*) grows on living, deciduous trees (van Geel et al., 2013). Likely these types enter the lake through run-off at times when trees are locally present. Algae and the two unknown types (VUW-36 and 37) are negatively correlated with tall trees and may be indicative of decreased terrestrial input into the lake or increased in-situ production.
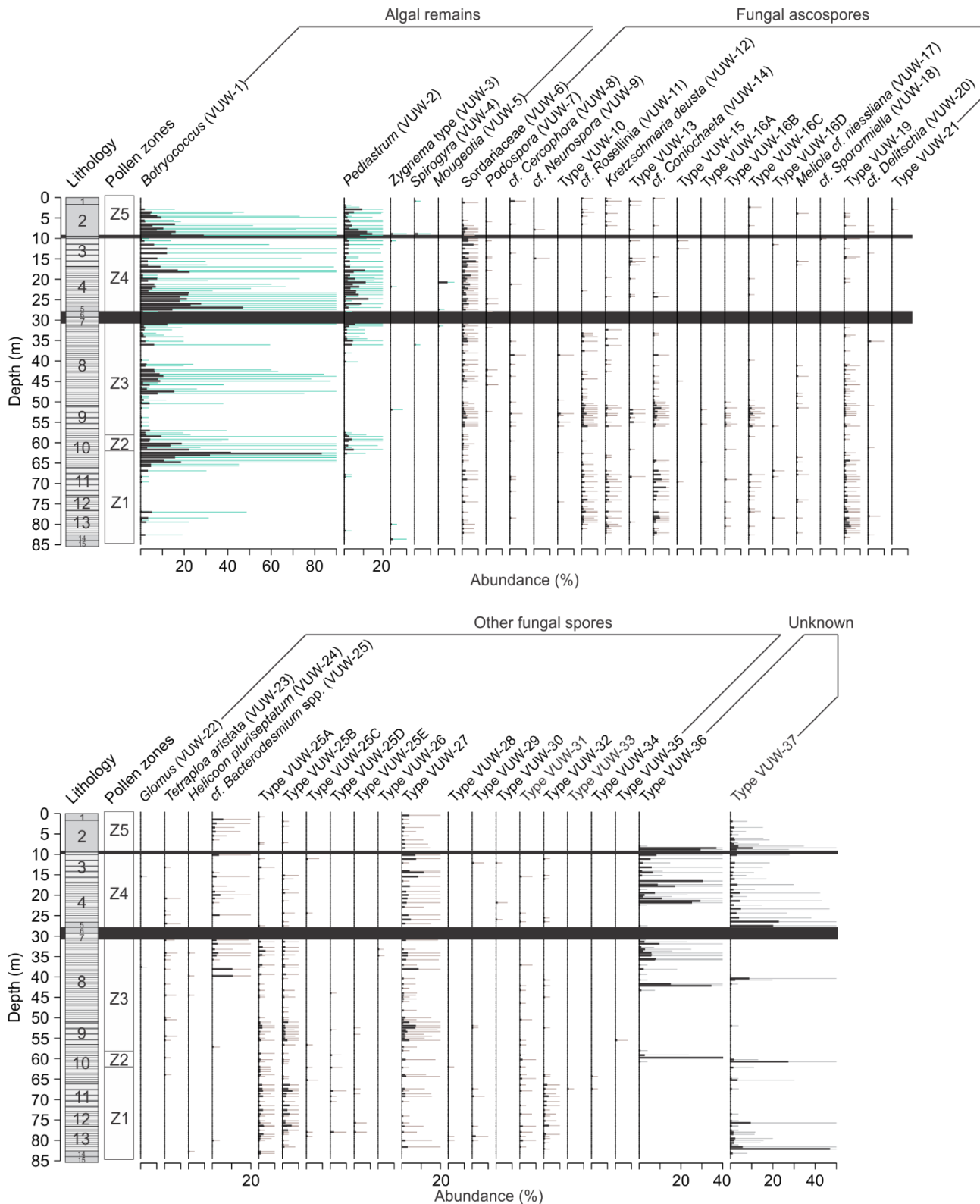
Figure D-2. NPP record for Orakei Basin. Abundances of the 37 distinguished NPP types as percentage of the pollen sum, indicated with black bars. Coloured bars represent 10x exaggeration with turquoise for algal remains, brown for fungal remains, and grey for remains of unknown origin. Lithology and pollen zones are also indicated (see Chapter 5).

Table D-3. Eigenvalues of RDA on NPP abundances constrained by the four major groups of pollen taxa.

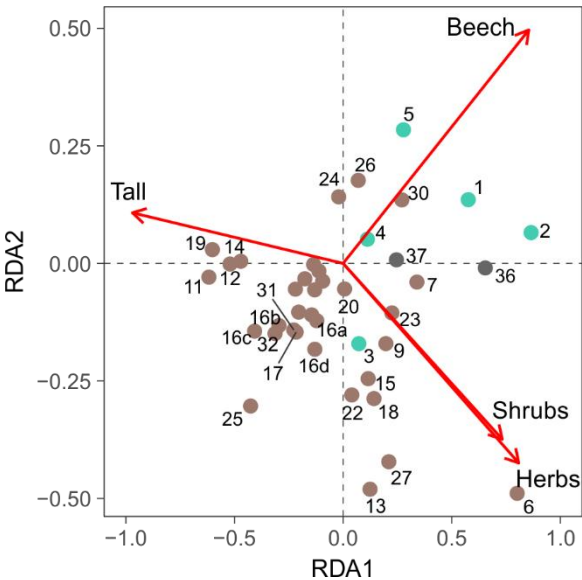|		Eigenvalue	Percentage of Inertia
Constrained	Axis 1	2.54	6.35
	Axis 2	0.74	1.85
	Axis 3	0.34	0.85
	Axis 4	0.17	0.43
Unconstrained	Axis 6	3.70	9.25
	Axis 7	2.38	5.95
Sum of constrained		3.78	9.45
Total Inertia		40.00	100

Figure D-3. Biplot of RDA on NPP abundances with major vegetation groups' pollen abundances as constraining variables. Taxon scores labelled with VUW-number. Turquoise points indicate algal types, brown fungal types, grey unknown types. Cluster of taxa down and left from the origin consists of: VUW-8, 10, 21, 28, 29, 33, 34 and 35. Tall = tall trees, Beech = beech trees, Shrubs = small trees and shrubs.

Type VUW-6 (Sordariaceae) is strongly associated with open vegetation in this dataset. Sordariaceae and other Sordariales (e.g. *Podospora* VUW-7, VUW-8) are often associated

with grasslands and grazing (e.g. Cugny et al., 2011) and many of its members include coprophilous species. Coprophilous fungi are normally spread via the digestive tract of large herbivores, which were not present in New Zealand before human arrival, although several coprophilous fungi genera have been found in avian dung samples in New Zealand (Bell, 1983). Sordariales also grow on herbaceous plants and decaying vegetative material (Mulder et al., 2003). I encountered only a handful of spores (VUW-7 and 18) that are very similar to the spores from obligate coprophilous fungi known from other parts of the world.

D4: Conclusions and NPP proxy potential

NPP are a neglected proxy for past environmental change in New Zealand. While the interpretation of NPP records is difficult at present, due to the lack of identification accuracy and linkage with modern living organisms, it could be improved over time if more palynologists would record the NPP they encounter. Here I present a first step towards crafting NPP into a useful proxy to be incorporated with pollen analysis. If developed, NPP records could be used to track the local presence of deciduous trees and other vegetation types, the extend of run-off and erosion, in-situ productivity of algae, etc. Another possible avenue is the tracking of certain tree pathogens, such as *Phytophtora agathidicida*, the cause of kauri dieback disease (Waipara et al., 2013) and answering questions such as: Where did *P. agathidicida* come from and how long has it been in New Zealand or infecting kauri trees?

References for Appendix D:

Bell, A. (1983). *Dung fungi: an illustrated guide to coprophilous fungi in New Zealand.* Victoria University Press, Wellington, New Zealand.

Bos, J.A.A., van Geel, B., Groenewoudt, B.J., Lauwerier, R.C.G.M. (2005). Early Holocene environmental change, the presence and disappearance of early Mesolithic habitation near Zutphen (The Netherlands). *Vegetation History and Archaeobotany*, 15, 27-43.

Cugny, C., Mazier, F., Galop, D. (2010). Modern and fossil non-pollen palynomorphs from the Basque mountains (western Pyrenees, France): the use of coprophilous fungi to reconstruct pastoral activity. *Vegetation History and Archaeobotany*, 19, 391-408.

Gelorini, V., Verbeken, A., van Geel, B., Cocquyt, C., Verschuren, D. (2011). Modern non-pollen palynomorphs from East African lake sediments. *Review of Palaeobotany and Palynology*, 164, 143-173.

Ghosh, R., Paruya, D.K., Acharya, K., Ghorai, N., Bera, S. (2017). How reliable are non-pollen palynomorphs in tracing vegetation changes and grazing activities? Study from the Darjeeling Himalaya, India. *Palaeogeography, Palaeoclimatology, Palaeoeology*, 475, 23-40.

Hooghiemstra, H. (2012). Non-pollen palynomorphs: From unknown curiosities to informative fossils. Celebrating the scientific career of Bas van Geel. *Review of Palaeobotany and Palynology*, 186, 2-4.

Jankovská, V., Komárek, J. (2000). Indicative value of *Pediastrum* and other coccal green algae in palaeoecology. *Folia Geobotanica*, 35, 59-82.

Lundqvist, N. (1972). Nordic Sordariaceae s. lat. *Symbolae Botanicae Upsalienses*, 20, 1-374.

Miola, A., (2012). Tools for non-pollen palynomorphs (NPPs) analysis: a list of Quaternary NPP types and reference literature in English language (1972–2011). *Review of Palaeobotany and Palynology*, 186, 142-161.

Mulder, C., Breure, A.M., Joosten, J.H.J. (2003). Fungal functional diversity inferred along Ellenberg's abiotic gradients: Palynological evidence from different soil microbiota. *Grana*, 42, 55-64.

Revelles, J., Burjachs, F., van Geel, B. (2016). Pollen and non-pollen palynomorphs from the Early Neolithic settlement of La Draga (Girona, Spain). *Review of Palaeobotany and Palynology*, 225, 1-20.

van Geel, B. (1978). A palaeoecological study of Holocene peat bog sections in Germany and The Netherlands. *Review of Palaeobotany and Palynology*, 25, 1-120.

van Geel, B. (2001). Non-pollen palynomorphs. In: J.P. Smol, H.J.B. Birks, W.M. Last (Eds.), Tracking environmental change using lake sediments, terrestrial, algal and silicaceous indicators, Vol. 3. Kluwer Academic Press: Dordrecht, The Netherlands, pp. 99-119.

van Geel, B., Aptroot, A. (2006). Fossil ascomycetes in Quaternary deposits. *Nova Hedwigia*, 82, 313-329.

van Geel, B., Bohncke, S.J.P., Dee, H. (1981). A palaeoecological study of an upper Late Glacial and Holocene sequence from "De Borchert", The Netherlands. *Review of Palaeobotany and Palynology*, 31, 367-448.

van Geel, B., Coope, G.R., van der Hammen, T. (1989). Palaeoecology and stratigraphy of the Lateglacial type section at Usselo (The Netherlands). *Review of Palaeobotany and Palynology*, 60, 25-129.

van Geel, B., Buurman, J., Brinkkemper, O., Schelvis, J., Aptroot, A., van Reenen, G., Hakbijl, T. (2003). Environmental reconstruction of a Roman Period settlement site in Uitgeest (The Netherlands), with special reference to coprophilous fungi. *Journal of Archaeological Science*, 30, 873-883.

van Geel, B., Engels, S., Martin-Puertas, C., Brauer, A. (2013). Ascospores of the parasitic fungus *Kretzschmaria deusta* as rainstorm indicators during a late Holocene beech-forest phase around lake Meerfelder Maar, Germany. *Journal of Paleolimnology*, 50, 33-40.

Waipara, N.W., Hill, S., Hill, L.M.W., Hough, E.G., Horner, I.J. (2013). Surveillance methods to determine tree health, distribution of kauri dieback disease and associated pathogens. *New Zealand Plant Protection*, 66, 235-241.

Appendix E: Chironomids: promising proxy or elusive larvae?

E1: Introduction

The Chironomidae are a family within the insect order Diptera. They are commonly referred to as non-biting midges or chironomids. Compared to pollen analysis, chironomid analysis is a relatively new approach to climate reconstruction. Chironomid head capsules are well suited to use as a palaeoenvironmental proxy for three main reasons (Ruiz et al., 2006): first, chironomid larvae live in lakes and are therefore generally abundant in lake sediments, lending themselves to statistical analysis; second, chironomids respond to environmental change rapidly (winged adults disperse easily and can produce one or more generations per season); third, the head capsules are taxon diagnostic (to genus or species level). Identification is based mainly on the shape and number of teeth on the mentum, shape of the ventromental plates, and the mandibles. Chironomids have been used for quantitative reconstructions of e.g. chlorophyll-a (Woodward & Shulmeister, 2006), water depth (Engels et al., 2016) and summer air temperature (SmT; Dieffenbacher-Krall et al., 2007; see Chapters 3 and 4). We showed in Chapter 3 that chironomids can providing new insights into environmental reconstructions through complementarity to pollen.

However, while chironomid analysis and subsequent summer temperature reconstructions were successfully developed from the sediments of Lake Pupuke, we encountered two major difficulties with this method: first, chironomids were rare or even absent in sediments of pre-Holocene age (Fig. E-1a); second, even in samples where chironomids were present at high abundances, isolating them from the sediments can be very time consuming. The latter was made worse by the Lake Pupuke sedimentary matrix, which consists of relatively large organic particles that tend to stick together, obscuring the chironomid head capsules in masses of extraneous material. At Orakei Basin (Chapter 5), there are reasonably high amounts of head capsules (> 20/cc) between 77 and 65 m depth and above 7 m depth but low concentrations at other depths (Fig. E-1b). These depths probably correspond to warm episodes (see Chapter 5). Here we discuss the potential causes of low concentrations at our study sites and discuss a possible methodological solution.
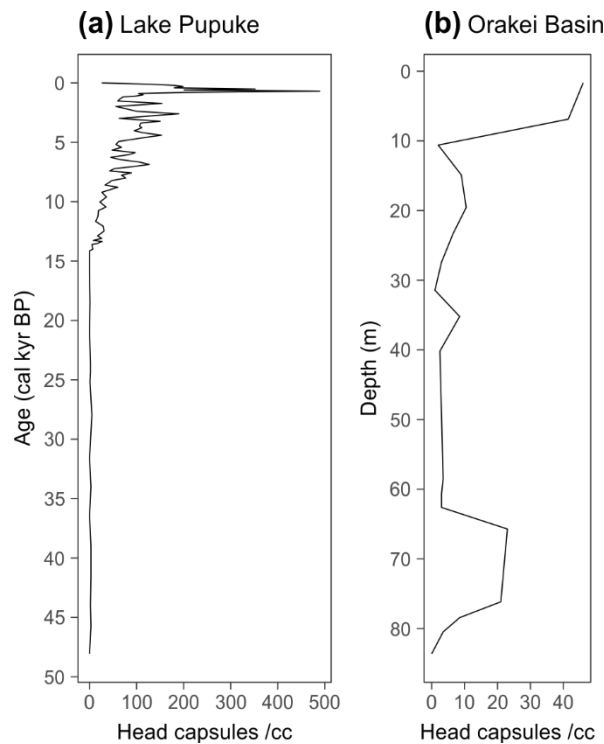
Figure E-1. Concentration of chironomid head capsules in the sediments of Lake Pupuke (a) and Orakei Basin (b).

E2: Speculation on the cause of low head capsule abundances

We cannot say with certainty what causes the low abundances of head capsules in the glacial sections of the Lake Pupuke and Orakei Basin sediments. Potential cuprits are: the lakes were too deep, too steep or too cold for larvae to thrive/survive, or the northern New Zealand environment was not suitable for the survival of chironomids during the Last Glacial. Climate is likely to be a factor, as chironomids are mainly absent in the core sections thought to represent glacial conditions (see Chapters 3 and 5).

Lake Pupuke is deep (57-58 m) and stratified, which implies that its benthic zone is anoxic. While Orakei Basin is shallow nowadays, it must have been deep in the past, as lake sediments reach deeper than 100 m in the crater. In combination with low temperatures during the Last Glacial, deep lakes may have produced unfavourable conditions for chironomids. However, Eggermont & Heiri (2012) mention results from very deep (100 m), stratified lakes in their review of the chironomid-temperature relationship. Moreover, Larocque et al. (2009) describe chironomid-temperature relations from Lake Silvaplana in the Swiss Alps, which has dimensions comparable to Lake Pupuke. Silvaplana has a

maximum depth of 77 m, and although the maximum summer air temperature in the region is only 10.8°C, they find high chironomid head capsule abundances (on average 30 capsules per 30 mg of dry weight). These results from other deep, cold lakes makes it unlikely that water depth and temperature are controlling head capsule abundances at Lake Pupuke and Orakei Basin.

Basin morphology may be what sets the Auckland maars apart: both the Lake Pupuke and Orakei Basin craters are steep sided with relatively narrow littoral zones. Presuming most chironomids live in the protected, warmer, oxygenated realms of shallow waters, a narrow littoral zone can explain low abundances of head capsules in the lake in general. In other deep lakes, head capsules from central cores may derive from the littoral zone, as discarded head capsules accumulate in the deepest part of the lake. Alternatively, chironomids do live in the anoxic benthic zones of other deep lakes but rely on nutrients produced in the littoral zone trickling down. However, Pupuke is deep and steep now, and its surface sediments contain high amounts of head capsules. Perhaps the combination of lake depth, crater steepness and low temperatures during the Last Glacial together eliminated chironomid larvae from Pupuke and Orakei.

Finally, it is possible that chironomids were truly rare in northern New Zealand during the Last Glacial. This may seem unlikely, as chironomids have been found to be abundant at other New Zealand sites at this time (Vandergoes et al., 2008) and are still abundant in alpine lakes today. On the other hand, New Zealand chironomid populations may be more sensitive to disturbance compared to populations in e.g. Europe and North America, where large land masses with myriad lakes are present to repopulate disturbed communities from. Perhaps cold-tolerant taxa were not able to migrate as far north as the Auckland region. The Cook Strait may also represent a physical barrier for South Island taxa and populations to reach the North Island during Glacials.

E3: Strategy

If there are no head capsules present in the sediments you are studying, chironomid analysis is obviously not possible. Low head capsule concentrations, on the other hand, pose a dilemma: large amounts of sedimentary material are needed for the analysis of such samples. In these cases, the decision to go ahead with analysis will depend on how much

material is available and how many other analyses are planned. The minimum required number of head capsules for downcore summer temperature reconstruction is 60 (see Chapter 4), so if the concentration of chironomids is < 6 capsules/cc (as throughout most of the pre-Holocene Lake Pupuke sediments), at least 10 cc of sedimentary material is required. In addition, the amount of time it will take to analyse these large samples needs to be considered. Head capsules are picked by hand after sieving the sample over a 90 µm nylon mesh; the larger the sample, the longer it takes to pick.

For Lake Pupuke, we decided to focus on the sediments covering the last 14 cal kyr BP, rather than investing time into analysing the older sediments. However, before this decision was made, we tested whether freeze drying or sonication of the sediment samples before sieving would improve the picking time by eliminating the larger organic particles. We did not dwell long on these tests: freeze drying and sonication for ~10 seconds did not improve picking time. Sonicating for more than 10 seconds did eliminate most extraneous organic matter, but additionally destroyed most head capsules (0 – 13 head capsules recovered per cc from a sample that yielded 50.5 head capsules per cc using standard preparation methods).

Preliminary analyses quickly made clear that producing a chironomid record for Orakei Basin would be very time consuming, as head capsule abundances were low for most of the record (Fig. E-1b). However, the preliminary results are promising, with identifiable remains found as deep as 80 m (Table E-1), which is thought to represent Marine Isotope Stage 5 (see Chapter 5). Although it was not possible to produce an Orakei Basin chironomid record for this thesis, we think additional methods should be tried to make it a possibility for future work.

Kerosene flotation, which is commonly used to extract beetle remains from sediments, appears to be a promising method to speed up chironomid picking (Rolland and Larocque, 2007). Kerosene flotation was found to be efficient for extracting chironomids from sediments of 10 different lakes in Canada, with low relative abundance errors. Larger head capsules (400-500 µm width) had a higher relative abundance error, due to the fact that they are often filled with sediment particles. Perhaps a combination of kerosene flotation and sieving of the non-floating residue over a coarser mesh (standard mesh size = 90 µm), could result in representative samples.

Table E-1. Preliminary results of Orakei Basin chironomid analysis: number of head capsules per taxon.

	Depth (m)																
	1.680	6.880	10.600	14.885	19.555	23.320	27.400	31.435	35.200	40.185	58.470	60.690	62.610	65.710	76.160	78.430	80.475
Unidentified	3.5														1		
Chironominae - tribe Chironomini																	
Chironomini early instar															3		1
Chironomus	1.5	2.5		2			1		1				1	2	2	1	
Cladopelma		2											1				
Cryptochironomus										1							
Paucispinigera	6			1					1.5	0.5				1.5		1.5	1
Polypedilum	15	4		2		1				2				1	6		1
Xenochironomus		1															
Chironominae - tribe Tanytarsini																	
Corynocera		5			1				1			1					
Tanytarsus funebris type C	8.5			4	1	1	1							2	4	1	
Tanytarsus vespertinus															1		
Unidentified Tanytarsini		1															
Orthocladiinae																	
Cricotopus aucklandensis					1.5		3.5					2		15.5	4	4	
Cricotopus zelandiae														1			
Gymnometriocnemus type	2.5																
Kaniwhaniwhanus			1														
Limnophyes/Paralimnophyes	1	1		0.5						0.5							0.5
Naonella forsythi	7.5			1	2				2	1						1	
Naonella kimihia	3.5	0.5		0.5													
Naonella type 315				1													
Naonella type 419	0.5			2													
Orthocladiinae 5								1									
Parakiefferiella									1								
Tanypodinae																	
Ablabesmyia		2	1			1							1		1		
Apsectrotanypus type		2															
Macropelopini 1		1				1			2	1							
Macropelopini 2		1.5															
Unidentified Tanypod														1			

References for Appendix E:

Dieffenbacher-Krall, A.C., Vandergoes, M.J., Denton, G.H. (2007). An inference model for mean summer air temperatures in the Southern Alps, New Zealand, using subfossil chironomids. *Quaternary Science Reviews*, 26, 2487-2504.

Eggermont, H., Heiri, O. (2012). The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. *Biological Reviews*, 87, 430-456.

Engels, S., Bakker, M.A.J., Bohncke, S.J.P., Cerli, C., Hoek, W.Z., Jansen, B., Peters, T., Renssen, H., Sachse, D., van Aken, J.M., van den Bos, V., van Geel, B., van Oostrom, R., Winkels, T., Wolma, M. (2016). Centennial-scale lake-level lowstand at Lake Uddelermeer (The Netherlands) indicates changes in moisture source region prior to the 2.8-kyr event. *The Holocene*, 26, 1075-1091.

Larocque, I., Grosjean, M., Heiri, O., Bigler, C., Blass, A. (2009). Comparison between chironomid-inferred July temperatures and meteorological data AD 1850–2001 from varved Lake Silvaplana, Switzerland. *Journal of Paleolimnology*, 41, 329-342.

Rolland, N., Laroque, I. (2007). The efficiency of kerosene flotation for extraction of chironomid head capsules from lake sediments samples. *Journal of Paleolimnology*, 37, 565-572.

Ruiz, Z., Brown, A. G., Langdon, P. G. (2006). The potential of chironomid (Insecta: Diptera) larvae in archaeological investigations of floodplain and lake settlements. *Journal of Archaeological Science*, 33, 14-33.

Vandergoes, M.J., Dieffenbacher-Krall, A.C., Newnham, R.M., Denton, G.H., Blaauw, M. (2008). Cooling and changing seasonality in the Southern Alps, New Zealand during the Antarctic Cold Reversal. *Quaternary Science Reviews*, 27, 589-601.

Woodward, C.A., Shulmeister, J. (2006). New Zealand chironomids as proxies for human-induced and natural environmental change: Transfer functions for temperature and lake production (chlorophyll *a*). *Journal of Paleolimnology*, 36, 407-429.