

---

**Predicting the suitability of New Zealand rivers and lakes for colonisation and growth of the invasive, non-indigenous diatom, *Didymosphenia geminata*: an update**

---

NIWA Client Report: CHC2007-062  
September 2007

NIWA Project: MAF07505

**Predicting the suitability of New Zealand rivers  
and lakes for colonisation and growth of the  
invasive, non-indigenous diatom, *Didymosphenia  
geminata*: an update**

---

Cathy Kilroy  
John Leathwick  
Katie Dey  
Neil Blair  
Helen Roulston  
Julian Sykes  
Donna Sutherland

*Prepared for*

**MAF Biosecurity New Zealand**

NIWA Client Report: CHC2007-062  
September 2007  
NIWA Project: MAF07505

National Institute of Water & Atmospheric Research Ltd  
10 Kyle Street, Riccarton, Christchurch  
P O Box 8602, Christchurch, New Zealand  
Phone +64-3-348 8987, Fax +64-3-348 5548  
[www.niwa.co.nz](http://www.niwa.co.nz)

# Contents

---

Executive Summary	iv
1. Introduction	1
2. Evaluation of the accuracy of the first likely environments map (LEM)	3
2.1. Introduction	3
2.2. Methods	4
2.2.1. Data	4
2.2.2. Evaluation procedure	4
2.3. Results and discussion	4
3. Modelling suitable environmental conditions for the <i>D. geminata</i> predictive maps (DPM)	7
3.1. Introduction	7
3.2. Field methods	9
3.2.1. Site selection	9
3.2.2. Survey procedure	9
3.3. Data	11
3.3.1. <i>D. geminata</i> data	11
3.3.2. Substrate data	11
3.3.3. Flood effects	14
3.4. Environmental predictors	17
3.5. Modelling procedure	21
3.6. Predicting <i>D. geminata</i> % cover and thickness	22
3.6.1. Range extrapolation	23
3.6.2. A potential test for the models	23
3.7. Temperature	23
3.8. Results	24
3.8.1. Variables important in determining the distribution of <i>D. geminata</i> in New Zealand	24
3.8.2. Predictions	27
3.8.3. A potential test for the models	36
3.8.4. Temperature	39
3.9. Discussion	39
3.9.1. Modelling approach	39
3.9.2. Limitations of the models	42
3.9.3. Environmental factors influencing <i>D. geminata</i> growth and blooms	43
3.9.4. Comparison with the LEM	44
3.9.5. A temperature limit for <i>D. geminata</i> ?	46
3.9.6. Concluding overview	50
4. Colonisation in lakes	51

4.1.	Survey methods and results	52
4.1.1.	Lake Manapouri	52
4.1.2.	Lake Te Anau	53
4.1.3.	Lake Wakatipu	53
4.1.4.	Lakes Roxburgh and Dunstan	53
4.2.	Discussion	54
4.2.1.	Causes of colonisation differences between lakes	54
4.2.2.	Identification of factors favouring <i>D. geminata</i>	56
5.	Use of and applications for the DPM	56
5.1.	DPM vs. LEM	57
5.2.	<i>D. geminata</i> cover / thickness and stream values	58
5.3.	Limitations of the DPM	60
5.4.	An example	61
5.5.	Other potential applications	66
5.6.	Model prediction formats	67
6.	Conclusions and recommendations	68
7.	Acknowledgements	70
8.	References	71
Appendix 1:	List of river sites surveyed for <i>D. geminata</i> % cover and thickness in March – April 2007	
Appendix 2:	Boosted regression tree models	
Appendix 3:	Response curves and interaction plots for variables important in determining distribution of <i>D. geminata</i> in New Zealand	
Appendix 4:	Lake sites surveyed for % cover and thickness of <i>D. geminata</i> in April – May 2007	

Reviewed by:

Maurice Duncan

Helen Rouse

Approved for release by:

Brian Sorrell

## Executive Summary

- This report describes work commissioned by MAF Biosecurity New Zealand to develop new models to predict the potential distribution of the invasive, non-indigenous, diatom, *Didymosphenia geminata* (*didymo*), throughout New Zealand rivers. The outcome of the new models and predictions are referred to as the “didymo prediction maps” (DPM). These DPM update an earlier “likely environments map” (LEM), which was based on limited information available when *D. geminata* was first discovered in New Zealand. The DPM are based on information not available at the time of the first mapping exercise, and use a different analytical approach based on statistical modelling.
- Other related work is also covered in this report. Major findings under eight objectives are summarised below.

### Objective 1: Evaluation of the accuracy of the LEM

- In terms of proximity to the assumed ideal environmental conditions for *D. geminata*, sites at which *D. geminata* subsequently appeared corresponded well with the LEM. The map showed five levels of likelihood for establishment of the species; 74 sites where *D. geminata* has been reported growing (up to 31 May 2007) fell into the top two levels. Seventy seven percent of the positive records were in the top level (closest to the ideal), compared with 45% of all South Island river reaches. Because all reports of *D. geminata* were from rivers with stream order  $\geq 4$ , we limited the evaluation of the LEM to these larger rivers only.

### Objectives 2, 3, 4: Updated models for *D. geminata* distribution in New Zealand

- The DPM were based on a survey of *D. geminata* cover and mat thickness in wadeable depths and velocities at 146 South Island sites, in rivers where the species was known to have been established for at least 9 months. We surveyed sites with no or very little *D. geminata* only if we were sure its absence was not because of lack of time for it to become established. The dataset was unavoidably biased towards river types to which *D. geminata* has spread via human activities, but we aimed to include as wide a range of rivers as possible within this constraint (e.g., we included sites in both lake-fed and rain-fed rivers).
- Our main source of environmental data was a Geographical Information System (GIS) database of over 120 variables, comprising mainly variables derived for FWENZ (Freshwater Environments of New Zealand, a Department of Conservation initiative), which have been associated with the New Zealand River Environment Classification (REC) network. Redundant or irrelevant variables were dropped prior to model fitting, leaving a set of 18 variables likely to have strong functional links with variation in the abundance of *D. geminata*. These described hydrological, climatic, physical (e.g., substrate) and chemical conditions in rivers. Two of the variables (substrate index and time elapsed since the most recent significant

flood) were determined locally for each of the 146 sites. Equivalent estimated variables were substituted to make the predictions.

- We fitted two models, separately predicting the potential percentage cover and thickness of *D. geminata* mats, and displayed them respectively as “DPM\_cover” and a “DPM\_thickness”. A combined biomass index was tried initially, but proved technically difficult to model because its two components (cover and thickness) had different statistical distributions, and were also driven by somewhat different environmental variables. Predictions for each model were run under three scenarios: “shorter than average”, “average” and “longer than average” periods of elapsed time since the most recent flood.
- Ten and 12 variables, respectively, were included in the final models predicting *D. geminata* percentage cover and thickness. In both cases lake influence was by far the most important predictive variable, followed by variables associated with substrate size and hardness, and the number of days since a flood. Additional variables included a measure of temperature seasonality (greater cover and thickness associated with more continental climates), reach slope, rainfall (= flow) variability, and measures of the amount of pastoral landuse in the catchment. A direct temperature measure contributed a minor component to both models.
- All predictions are for mean % cover and mat thickness in wadeable depths. The DPM illustrate that greatest cover and thickness of *D. geminata* are predicted for eastern and southern parts of the South Island, but with much of the southern and south-eastern coastal areas less susceptible. Rivers in inland Otago and south Canterbury are predicted to develop high cover in long periods without floods. Lake-fed rivers throughout the South Island have maximum cover under all scenarios. The North Island is predicted to be much less susceptible than the South Island, though some lake-fed rivers may develop high cover levels and thick mats. After long periods without floods, much of inland central and eastern North Island is predicted to develop moderately thick mats, but this is associated with low levels of cover.
- These predictions must be treated with caution because the range of environmental conditions represented in the 146 data points used to generate the models were narrower than that in the New Zealand-wide rivers and streams for which we made predictions. To quantify this we calculated for every reach the total departure of all model variables (as a proportion) from their ranges in the training dataset. This was mapped separately as “DPM\_unreliability” to indicate the relative unreliability of predictions in different areas. An unreliability index of 0 means that environmental variables in the reach being predicted all fall within the range of those in the 146 survey sites. North Island predictions are less reliable than those for the South Island, with 25% and 41%, respectively, of river reaches of order  $\geq 3$  having unreliability = 0.
- We stress that the unreliability index must be taken into consideration when using the DPM. The only way to improve unreliability is to extend the range of the model data by adding more sites, then re-running the models.

- The DPM for mean % cover and mat thickness in the South Island correspond reasonably well on a broad scale with the relative likelihood for establishment indicated in the LEM. Correspondence for the North Island is poor, even for river reaches for which predictions were most reliable. This is probably because temperature was heavily weighted in the first map, but contributed very little to the new DPM.
- We investigated the relationship between *D. geminata* presence and temperature by extracting terrestrial temperature data for locations positive for *D. geminata* in the Northern Hemisphere. We found no confirmed records from areas where the mean temperature in the coldest month exceeded approximately 5.5 °C, although summer temperature could be as high as 28 °C. A possible explanation is that sexual reproduction in *D. geminata* – which is essential for the long-term maintenance of populations – may require a cold-temperature trigger. This could be tested experimentally.
- Our models confirmed observations that the important environmental variables favouring *D. geminata* colonisation and blooms are: high lake influence; stable, hard substrates; low flow variability; longer time since a flood; and large seasonal temperature differences. High lake influence was the most powerful predictive variable presumably because lake influence represents a combination of stable flows and stable substrates. A direct temperature link was unclear (probably because of lack of variability in the data), and links to water chemistry / quality variables were weak.
- Downstream declines of *D. geminata* percentage cover and thickness seen in the original data could be driven by hydrological factors, changing substrate size and stability, and / or temperature. The temperature question may be resolved through further surveys and by obtaining information on water temperature regimes. No satisfactory water temperature dataset currently exists in New Zealand, but could be developed relatively easily. We suggest that such a dataset may be critical in future predictions and decision-making related to freshwater invasive species in general.

#### **Objective 5: Estimate likelihood for colonisation of lakes by *D. geminata***

- We undertook preliminary surveys in several South Island lakes exposed to *D. geminata* inputs for at least 9 months. Using the survey results, we discuss the differences in colonisation among the lakes in terms of six factors likely to influence colonisation and blooms of the species. It was concluded that at this stage we do not have enough information to make predictions of either the timing of blooms in the present lakes, or the likely impact in other lakes. More thorough surveys in lakes currently exposed to *D. geminata* will establish the real extent of the species in these lakes and help identify lake-scale factors favouring growth. Monitoring over time in a single lake may help to determine the triggers for blooms.

#### **Objective 6: Design a protocol for ongoing monitoring of *D. geminata* in New Zealand**

- A suggested monitoring protocol is based on the method used to collect field data for this study. This involves visual estimates of *D. geminata* cover and thickness using an underwater viewer. Sample collection is encouraged in any instances where the identity of algal mats is uncertain. Surveys in which all the monitoring forms are completed in full will provide suitable data for updating the model in the future. The protocol is outlined within this report, but a more detailed protocol can be found at <http://www.biosecurity.govt.nz/pest-and-disease-response/pests-and-diseases-watchlist/didymosphenia-geminata/protocols>.

## Objectives 7, 8: Use of and applications for the models

- In contrast to the LEM which mapped relative suitability for *D. geminata* of all New Zealand river reaches, the DPM predict values for percentage cover and mat thickness of *D. geminata*. Although the DPM at first glance suggest a lower overall susceptibility in New Zealand compared to that suggested in the LEM, in fact they predict that over 40% of all reaches with stream order  $\geq 3$  will attain at least 10% cover by *D. geminata* (calculated for stream reaches with unreliability = 0), which represents cover that is clearly visible.
- Existing guidelines for maximum algal cover to retain aesthetics and recreational values and trout habitat and angling values are equivalent to 30% and 40% cover by *D. geminata*, respectively. These levels are predicted to be attained in 9 and 7.5% of South Island reaches (stream order  $\geq 3$ , unreliability = 0) respectively, and 0.7 and 0.5% of North Island reaches.
- Recent work on the effects of *D. geminata* on higher trophic levels found increases in invertebrate abundance associated with increasing *D. geminata* biomass. It has proved difficult to identify any thresholds of *D. geminata* biomass above which ecological effects (e.g., community composition changes) are evident.
- For applications of the DPM, for example, to assess risk at high-value sites, we recommend that predictions should be examined at the scale of catchments or reasonably long stretches of river. This allows for uncertainties both in the data underlying the models and in the model predictions themselves.
- An example of an application for the models is outlined: this relates to the discovery of *D. geminata* in the Kakanui catchment, North Otago, and subsequent evaluation of potential management actions particularly because of the presence of the endangered Kauru longjawed galaxias in the Kauru River, a tributary of the Kakanui River. Further studies are needed to determine the effects of *D. geminata* on populations of native fish species.
- We anticipate that the DPM we have developed for percentage cover and thickness will have applications for management of sites with special values, for the design of further surveillance programmes, for allocation of resources in relation to maintenance of *D. geminata*-free sites, and in decision-making regarding management of *D. geminata* infestations.



## 1. Introduction

An early information requirement following the incursion of the non-indigenous diatom *Didymosphenia geminata* into New Zealand (Southland) rivers (Kilroy 2004) was to predict those catchments in which *D. geminata* was most likely to thrive in New Zealand. Such predictions were needed for assessments of the potential environmental and economic impact of the species in New Zealand (Campbell 2005, Branson 2006). The first attempt to produce such predictions was initiated within two months of discovery of the organism. Because of lack of data, the resulting “likely environments map” (LEM) was not a prediction, but showed estimated relative ecological suitability of all New Zealand river reaches for establishment and blooming of *D. geminata* (Kilroy et al. 2005). Subsequently, the LEM has been used to help identify high-risk sites for sampling in a continuing series of delimitation surveys.

When the LEM was produced in May 2005, *D. geminata* had been confirmed in a single river system in New Zealand (the Mararoa and lower Waiau, Southland). This meant that commonly used methods of habitat preference mapping could not be applied, because all these require data from multiple locations across a broad gradient of environmental conditions. Limited data from international sources also constrained the options for undertaking the analysis. Since May 2005, *D. geminata* has been detected in approximately 53 rivers (including tributaries in the same catchment) spanning the length of the South Island. Infestations have been observed to vary in intensity (i.e., thickness and cover of the growth) both spatially and temporally (e.g., Larned et al., in prep.). This provides an opportunity to re-visit the predictive mapping exercise and to potentially improve the original output by examining environmental conditions at sites where *D. geminata* is now established in New Zealand, then identifying similar conditions throughout the entire New Zealand river network. In other words, a more defensible modelling approach can be taken. In particular, this may improve our understanding of the susceptibility of North Island rivers to invasions by *D. geminata*. To date, the species has not been detected in the North Island.

In the present report, we describe the development of new predictions of the eventual extent of *D. geminata* in New Zealand rivers, assuming it could spread to all suitable reaches. These new predictions, known as the didymo prediction maps (DPM), are based on information not available at the time of the first mapping exercise. Thus, the initial LEM was **an assessment of the relative suitability of all New Zealand river reaches for establishment and blooms of *D. geminata***. In contrast, the DPM are **quantitative predictions of expected percentage cover and mat thickness of *D. geminata* in all New Zealand river reaches**.

In addition, MAF Biosecurity New Zealand, who requested this work, has identified several related objectives, which are also covered in this report. The complete list of objectives is as follows:

1. Evaluate accuracy of the original LEM.
2. Predict the potential growth of *D. geminata* (e.g., as percentage cover and mat thickness) in all New Zealand river reaches and present in colour-coded map format (including known sites).
3. Determine which river reaches will support a defined bloom of *D. geminata* and present in colour-coded map format.
4. Based on the model developed for rivers, identify the important environmental factors that influence *D. geminata* growth and lead to blooms (including limiting factors) in river reaches.
5. Estimate likelihood for colonisation of lakes by *D. geminata*, based on assessment of suitable conditions for establishment and blooms.
6. Based on the field data collected for this study, design a protocol for ongoing monitoring of *D. geminata* in New Zealand to provide suitable data that can be used to update the model in future. This will potentially improve our knowledge of the relationship between *D. geminata* and environmental variables and contribute to international understanding of this organism.
7. Provide an example of the practical application of the maps for assessing the risk to high-value sites, e.g., identify sites susceptible to *D. geminata* blooms, which also support populations of a selected species of native fish.
8. Discuss other potential applications of the models and maps.

This report deals with these objectives in the order given above. Objectives 2 to 4 are combined into a single section dealing with development of a model of *D. geminata* growth in New Zealand rivers, and use of that model to predict coverage and thickness of *D. geminata* in rivers throughout the country (DPM). The preliminary output for Objective 6 (the field protocol) is presented herein, but a more detailed protocol is covered in a separate report available at <http://www.biosecurity.govt.nz/pest-and-disease-response/pests-and-diseases-watchlist/didymosphenia-geminata/protocols>.

## 2. Evaluation of the accuracy of the first likely environments map (LEM)

### 2.1. Introduction

Development of the LEM for *D. geminata* in 2005 (Kilroy et al. 2005) comprised the following steps:

1. Environmental conditions favourable for establishment and growth of *D. geminata* were determined from available information (literature, personal communications, etc.).
2. A subset of environmental variables associated with the New Zealand River Environment Classification (REC) network,<sup>1</sup> was selected to represent the environmental factors relevant to *D. geminata* establishment and growth.
3. Weightings were applied to those variables considered to be most relevant for *D. geminata* distribution.
4. Values representing optimal conditions for *D. geminata* were estimated for each variable.
5. We then calculated the multivariate environmental distance of every river reach in New Zealand from the estimated optimal variable values; the river network was then mapped with colour coding to indicate the relative risk attached to each river reach<sup>2</sup>.
6. Estimates of the proportions of highest-risk river reaches in North and South Islands, and in sub-regions within each island (defined in the Department of Conservation Waters of National Importance project) were derived from these maps.

Unavoidably, subjective decisions were required for steps 1, 2 and 3 – and the method has been criticised for this reason – but the procedure had the advantage of being rapid and easily understood (Kilroy et al. in press). A more serious criticism was of our assumption of a cool-temperature preference. Given the subsequent appearance of the species in warm areas of the USA, this assumption appeared likely to have underestimated the risk to the North Island. Nevertheless, since this rapid method may

---

<sup>1</sup> The REC is a classification of all river reaches in New Zealand (from 1:50 000 maps). See Section 3.1 for more details. Most of the variables associated with the REC network were derived for the Department of Conservation's FWENZ (Freshwater Environments of New Zealand) project.

<sup>2</sup> A reach refers to any REC river segment, defined as a section of river between successive tributaries (mappable at 1:50,000).

need to be used in future incursions of unwanted freshwater organisms, it is of interest to determine how well the maps performed.

## 2.2. Methods

### 2.2.1. Data

*D. geminata* has now been confirmed in 14 separate catchments in the South Island in about 75 recorded locations. We also have records from sites surveyed for the DPM (Section 3, Appendix 1). We used all these records to test the accuracy of the LEM.

### 2.2.2. Evaluation procedure

We first identified REC reaches for all known sites positive for *D. geminata*. Then we extracted the environmental distance from the ideal for *D. geminata* for each of these river reaches (see Section 2.1). The distribution of these distances was then compared with the distribution of distances extracted for all South Island river reaches. To ensure that we were comparing equivalent rivers, we used only reaches with stream order greater than or equal to the minimum stream order of the positive sites. At the time of writing, there were no positive sites in streams with order less than 4.<sup>3</sup>

If the LEM performed well we would expect to see most of the positive sites in reaches with a low distance from the ideal for *D. geminata*, compared to a more widely spread distribution of distances for the whole South Island.

The survey data collected for development of the updated model (Section 3) comprises mean % cover of the river bed and mean mat thickness of *D. geminata* from 146 sites in rivers known to have been affected by *D. geminata*. If the original map had performed well, we would expect to see significant negative correlations between environmental distance and both % cover and thickness.

## 2.3. Results and discussion

The distribution of positive sites overplotted on the LEM shows that many sites are placed in reaches in the closest group to the ideal for *D. geminata* (Figure 2.1).

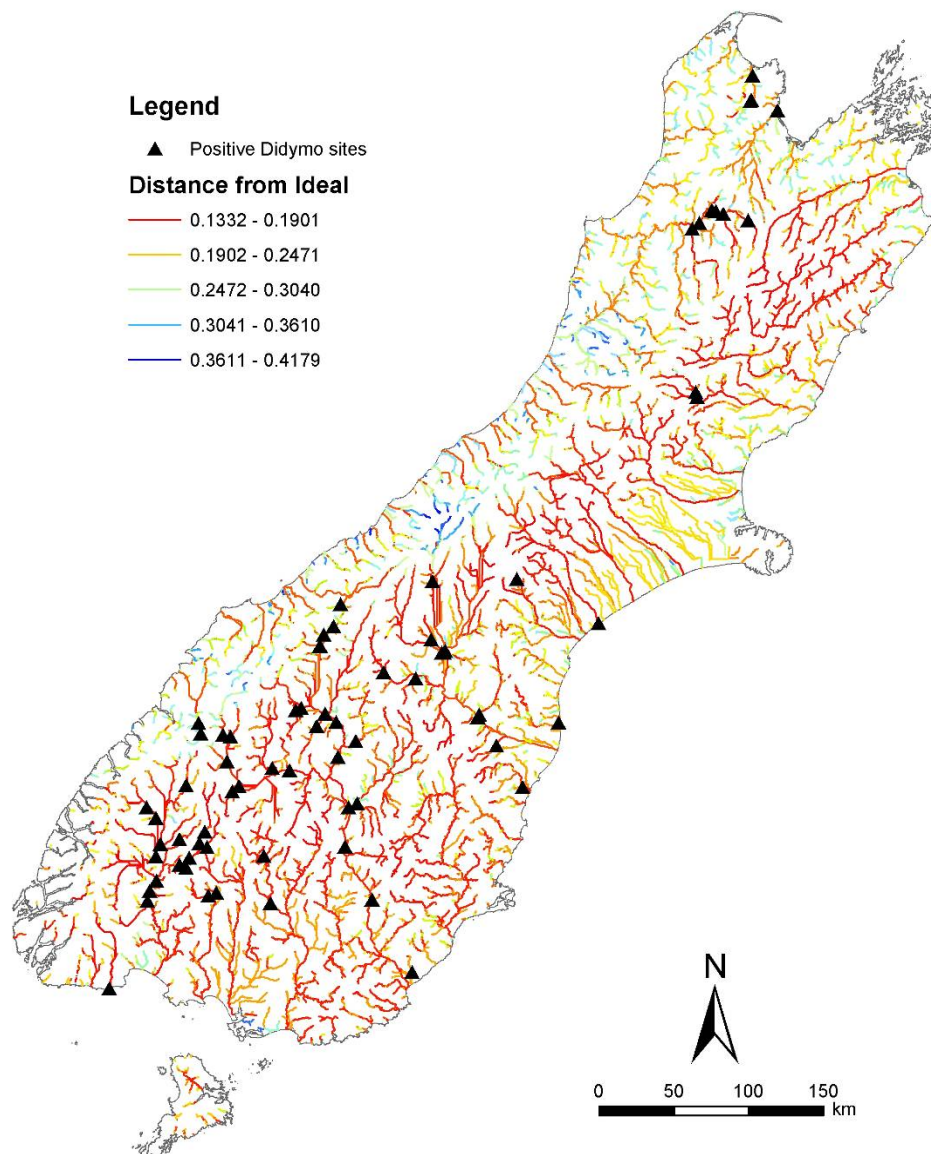
The range of distances from the ideal for *D. geminata* in all South Island rivers with stream order  $\geq 4$  was 0.133 (closest to ideal) to 0.418 (farthest from ideal). The range

---

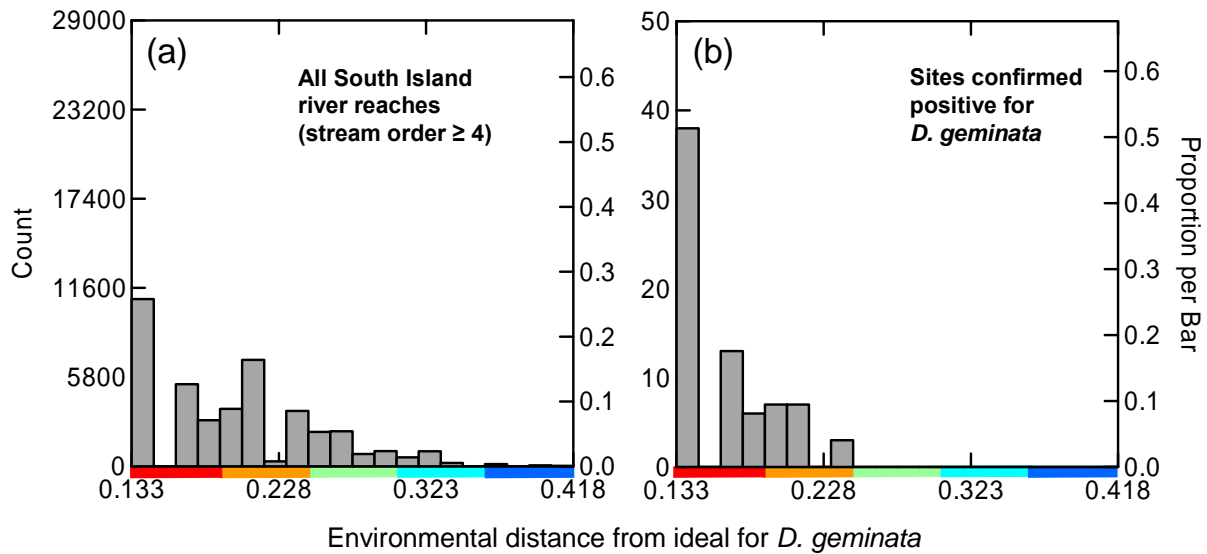
<sup>3</sup> A stream order of 1 refers to a headwater stream with no tributaries. At the first confluence with another headwater stream these become stream order 2, and so on. The maximum stream order in the REC is 8.

for positive sites was 0.133 to 0.246. Approximately 77% of the positive sites had distances in the bottom one-fifth of the range of distances (i.e., closest to the ideal), compared with approximately 45% of all comparable South Island river reaches (Figure 2.2). A two-sample Kolmogorov-Smirnoff (KS) non-parametric test confirmed that the two distributions are significantly different ( $P = 0.000$ ). (A KS test determines whether two populations come from the same distribution.)

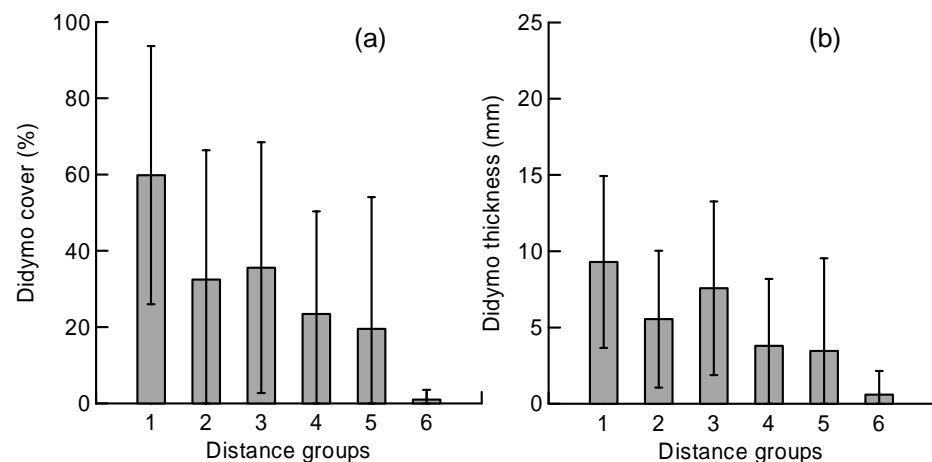
Distances were significantly and negatively correlated with % cover and thickness estimates from a survey of *D. geminata*-affected rivers ( $P < 0.001$ ). There was a great deal of scatter in the data (note large standard deviations in Figure 2.3), and the



**Figure 2.1:** River locations confirmed positive for *D. geminata* in the South Island, at 31 May 2007 plotted on the LEM. For clarity, only rivers with stream order  $\geq 4$  are plotted.



**Figure 2.2:** Distribution of river reaches among environmental distances from the ideal for *D. geminata* for (a) all South Island river reaches with stream order  $\geq 4$  (n = 42181) and (b) South Island river reaches in which *D. geminata* has been confirmed (n = 74). The colours along the x-axis correspond to colours used to map river reaches according to distance in Figure 2.1.



**Figure 2.3:** Mean *D. geminata* % cover and thickness plotted against distance from ideal for *D. geminata*, with distances split into six groups: 1, 0.133 – 0.152; 2, 0.1521 – 0.1708; 3, 0.171 – 0.189; 4, 0.190 – 0.208; 5, 0.2084 – 0.227; 6, 0.2275 – 0.246. Error bars are standard deviations.

predictive power of the relationships was low (% cover  $R^2 = 0.139$ ; thickness  $R^2 = 0.099$ ). However, high variability was not unexpected given the wide range of times since the most recent flood (see Appendix 1), and the strong association of *D. geminata* cover and thickness at individual sites with this variable (see Section 3.3.3, and Figure 3.3, below).



We conclude that the LEM performed reasonably well in identifying river reaches suitable for *D. geminata* establishment and growth. Even though South Island river reaches (stream order  $\geq 4$ ) are dominated by low distances from the ideal for *D. geminata* (i.e., suitable habitat, red and orange colours in Figure 2.1), the proportion of positive sites with these lower distances approached double that in all rivers. No positive sites had a distance greater than the red and orange groups in Figure 2.1.

### 3. Modelling suitable environmental conditions for the *D. geminata* predictive maps (DPM)

#### 3.1. Introduction

An acknowledged shortcoming of the methodology used for development of the LEM (Kilroy et al. 2005) was that ideal environmental conditions for *D. geminata* were described using single values for environmental variables. It was also necessary to make highly subjective decisions about the optimal value of each variable and the weighting applied. For example, because its known distribution indicated that *D. geminata* was a cool-climate species, we selected a relatively low optimal value for a temperature variable and also triple weighted this variable. Similarly, because this species appeared most likely to bloom in the stable conditions of lake-fed rivers, a variable for lake influence was set high and was double-weighted.

In this section we describe development of a predictive model derived from *D. geminata* data from all South Island rivers in which the organism is established. We estimated average percentage cover and mat thickness at ~150 sites, and collected associated field data on the environmental conditions known (from ongoing ecological studies) to influence its growth. These data, along with environmental variables associated with every site surveyed, were used to construct a statistical model that describes average relationships between environmental conditions and variation in *D. geminata* cover and mat thickness. We then used this model to make predictions of *D. geminata* cover and thickness for the river network in New Zealand.

Three important factors were taken into consideration:

1. Cover of *D. geminata*, like all river periphyton, changes over time as a result of flow variations, which cause periodic scouring or accumulation of growth. Therefore, we estimated the time since the most recent bed-moving flood event for every data point and included this as a variable in the modelling process.

2. The effect of flow variations itself varies depending on the stability of the river substrate. Thus, large, embedded substrates will remain stable in larger, faster water flows than smaller unconsolidated substrates. Periphyton is most likely to be scoured in flows that cause substrate movement. To account for this at a local level (i.e., at the scale of the riverbed areas we surveyed), we assessed substrate composition at every site surveyed.
3. Because *D. geminata* is still spreading in New Zealand, we can only define its current environmental limits, and these may not coincide with its potential limits. The South Island rivers in which it has established are unlikely to cover the entire range of conditions in which *D. geminata* can survive. In particular, recent records of *D. geminata* in the USA indicate that this diatom can grow well in rivers in warm areas (e.g., Arkansas). Therefore using only temperature data based on the South Island could underestimate the suitability of streams in the North Island. To address this uncertainty, we compared temperature data associated with records of *D. geminata* from other parts of the world with equivalent temperature data for the whole of New Zealand.

We also acknowledge that because *D. geminata* is still spreading in New Zealand, the sites where it is currently present likely represent only a subset of the river types in which it can potentially thrive. The species is appearing mainly in relatively large rivers that are popular for recreational activities (angling, kayaking, jet-boating, etc.) and from this we assume that it is being spread mainly by humans. Consequently, most of the following discussion and analysis applies only to high-order rivers and streams.

Our main source of environmental data was a GIS database associated with the River Environment Classification (REC) network, which was also the data source for the LEM (see Section 2.1). The REC was developed by NIWA as a tool to aid water resources management in New Zealand. The REC classifies individual sections of New Zealand's river networks on the basis of a hierarchy of controlling factors, including climate, topography, geology and land cover of the catchment of each section, and attributes of the section itself (Snelder & Biggs 2002). The system is based on a GIS layer that defines river sections with a mean unit size of approximately 1 km. Each river section is defined by its confluence with tributaries at its upstream and downstream ends. All sections are mappable at a 1:50 000 scale and the entire network comprises approximately 565,000 sections.

The REC itself is a categorical system. However, a parallel set of continuous environmental variables has been applied to the REC network, many of which were developed for the Freshwater Environments of New Zealand (FWENZ) dataset for the Department of Conservation (Wild et al. 2005). We used selected variables from the



FWENZ and other datasets (e.g., Leathwick et al. submitted) in the development of the present model for *D. geminata*.

### **3.2. Field methods**

#### **3.2.1. Site selection**

The dataset used for the modelling was obtained during a survey of *D. geminata*-affected rivers in March – April 2007. We estimated average *D. geminata* cover and thickness at multiple sites in rivers where the species was known to have been present for at least 9 months in most cases. In some smaller rivers we used a shorter time period (Table 3.1). Our aim was to include as wide a range of river types as possible, as well as the entire range of cover from 0 to 100%. We surveyed sites with no or very little *D. geminata* only if we were sure its absence was not because of lack of time for it to become established. Such sites were always downstream of locations where *D. geminata* was found at the time of its first detection.

As far as was practicable, sites were pre-selected to ensure the best coverage of each river. Ideally all sites sampled would be in separate REC reaches. In most cases numerous tributaries meant that this was not a problem. In fact the condition was not met for only one pair of sites (Twizel River). Site access by road was often possible, but limited access along the Waitaki River and parts of the lower Waiau and Hollyford Rivers necessitated use of a jet boat. The central section of the Buller River was accessed by kayak.

#### **3.2.2. Survey procedure**

At each site we worked in a river reach of 40 – 50 m. Survey areas were, as far as possible, representative of the REC reach segment. For example, if the river comprised a series of riffles and runs, we included areas of both water types in the survey area. We first marked the ends of four transects 10 – 20 m apart. Each transect extended into the river at right angles to a maximum depth of about 0.6 m (i.e., a wadeable depth). The first assessment was made about one-fifth of the way along the most downstream transect. To assess cover, we used an underwater viewer, which provides a clear view of the river bed and also defines a circle of approximately constant size (~ 50 cm diameter) at each survey point (equivalent to using a quadrat). Percentage cover of the substrate by *D. geminata* in the viewer circle was assessed by eye. Thickness was assessed by retrieving rocks and measuring/estimating mat thickness to the nearest 5 mm. Where a range of thicknesses were present an average was estimated. We repeated this procedure at five points along each transect (total of 20 points).

We assessed substrate composition by visually estimating the percentage of the surveyed area occupied by seven categories of substrate: bedrock, boulders, large cobbles, small cobbles, gravels, sand and silt/mud (see Section 5 for definitions). Other site information collected included water temperature, conductivity and pH, degree of shading (three categories), and water velocities (five categories, and measured in some cases).

**Table 3.1:** List of rivers included in the survey including relevant hydrological recording sites.

River	Date <i>D. geminata</i> detected	Date survey completed	No. of sites surveyed	Hydro. site no.	Hydro. site name
Lower Waiau	20-Oct-04	3-Apr-07	18	79719	Waiau, MLC
				79735	Waiau, Sunnyside
				79701	Waiau, Tuatapere
Mararoa	26-Oct-04	10-Apr-07	6	79737	Mararoa, Cliffs
Buller	24-Sep-05	28-Mar-07	25	93202	Buller, Longford
				93203	Buller, Te Kuha
Oreti	29-Sep-05	11-Apr-07	13	78608	Oreti, Three Kings
				78636	Oreti, Lumsden Cableway
Clutha	29-Sep-05	18-Apr-07	18	75282	Clutha, Cardrona confl
				75228	Clutha, Clyde
				75222	Clutha, Tuapeka
				75207	Clutha, Balclutha
Upper Waiau	29-Sep-05	4-Apr-07	3	79708	Waiau, Queens
Von	28-Oct-05	7-Apr-07	4	75276	Shotover, Bowen Peak
Whitestone	19-Dec-05	6-Apr-07	4	79737	Mararoa, Cliffs
Waitaki	21-Jan-06	22-Mar-07	18	71104	Waitaki, Kurow
Ahuriri	17-Feb-06	23-Mar-07	5	71116	Ahuriri, Sth Diadem
Aparima	26-Mar-06	11-Apr-07	7	78901	Aparima, Thornbury
				78905	Aparima, Otautau
				78906	Aparima, Dunrobin
Fraser	3-May-06	19-Apr-07	4	75261	Fraser, Laing Rd
Twizel	18-May-06	23-Mar-07	5	71129	Forks, Balmoral
Upukerora	18-May-06	6-Apr-07	1	79737	Mararoa, Cliffs
Makarora	1-Nov-06	23-Apr-07	3	75294	Matukituki, W. Wanaka
Eglinton	1-Nov-06	9-Apr-07	7	79737	Mararoa, Cliffs
Hollyford	22-Nov-06	12-Apr-07	5	86802	Haast, Roaring Billy

Because we were working only in rivers where *D. geminata* was known to be present, we did not expect significant problems distinguishing *D. geminata* from other types of algae, and no routine sampling was included as part of the survey protocol. However, to provide an identification check, we took photographs of the river bed and/or algae-covered rocks retrieved from the bed at every site. In a few cases where there was doubt about the identity of algal cover, samples were taken for later checking under the microscope.

A detailed protocol based on the above, and recommended for future surveys, is available as a separate report (see <http://www.biosecurity.govt.nz/pest-and-disease-response/pests-and-diseases-watchlist/didymosphenia-geminata/protocols> ).

### 3.3. Data

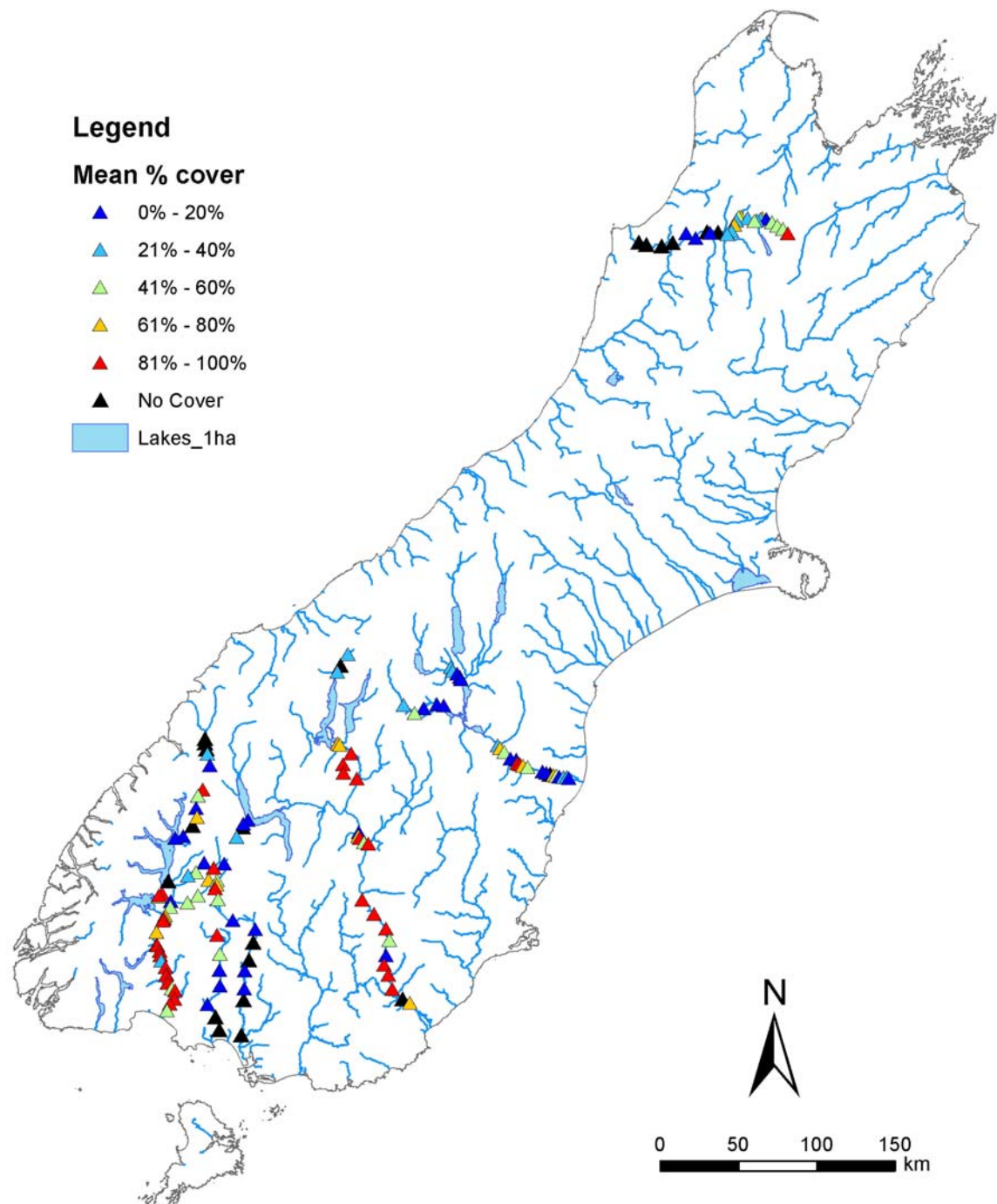
#### 3.3.1. *D. geminata* data

Sites surveyed are listed in Appendix 1. A total of 146 sites in 17 rivers (Table 3.1) were included in the dataset used for constructing the model. Mean % cover and mat thickness at each site were calculated from the 20 data points examined at each site. Mean cover and thickness of *D. geminata* ranged from 0 to 100% and 21 mm, respectively (Figure 3.1a, b). The two variables were highly correlated, although there was a wide range of thicknesses when % cover was high (Figure 3.2). Initially the variables were combined into a single index (mean % cover \* thickness). However, this index proved unsuitable for modelling because its error distribution could not be easily accommodated using standard model specifications, e.g., Gaussian or Poisson distributions. Therefore, separate models were constructed for % cover and thickness.

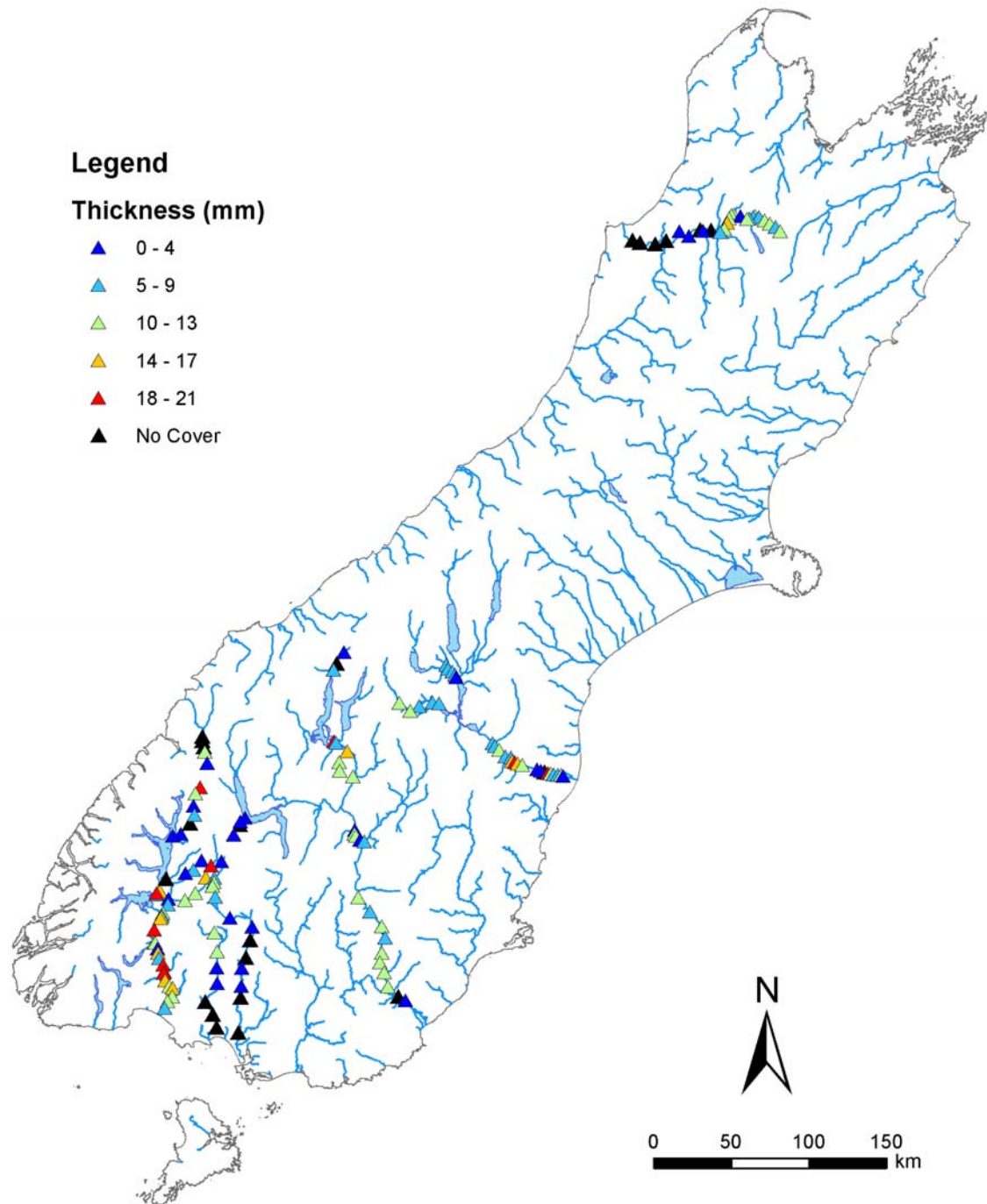
#### 3.3.2. Substrate data

We derived a single substrate index from our estimates of percentage cover by different substrate types. First we summed all the percentages, each weighted by 1 to 7 according to size (smallest to largest), then we divided by 100. For example, for a substrate comprising 50% gravel (weighting 3) and 50% small cobbles (weighting 4) the index would be:

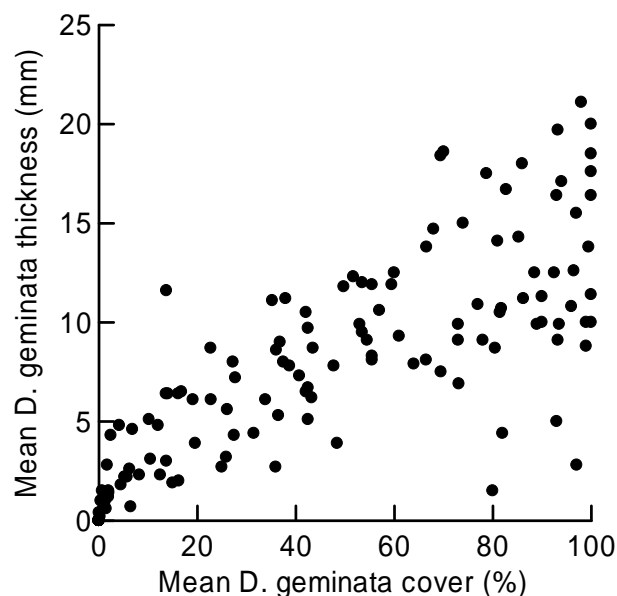
$$[(50 * 3) + (50 * 4)] / 100 = 3.5$$



**Figure 3.1(a):** River sites included in the survey, showing *D. geminata* % cover.



**Figure 3.1(b):** River sites included in the survey, showing *D. geminata* thickness.



**Figure 3.2:** Mean thickness of *D. geminata* mats plotted against mean % cover at 146 sites in South Island rivers.

### 3.3.3. Flood effects

As discussed above (Section 2.3), *D. geminata* cover varies in a river with flow. It is not known yet whether growth is also seasonal, though growth trials on artificial substrates have indicated that – as would be expected – growth rates are faster in summer than in winter (M.A. Lagerstedt, pers. comm.). Past ecological work on *D. geminata* has shown that this organism is quite resistant to floods. Indeed this is why it attains such high biomass and is a nuisance. It has been suggested that only floods capable of moving some or all of the bed substrate will scour a substantial proportion of *D. geminata* growth (Kilroy et al. 2006c). This implies that when *D. geminata* is present in a system, and all other things being equal, biomass will be positively correlated with the length of time without a bed-moving flood.

We accounted for the temporal component of *D. geminata* growth by adding, for each datapoint, the time (in days) since the most recent bed-moving flood. Estimating this variable presented several difficulties:

1. Only a few sites had flow recorders close by, and some rivers (e.g., Hollyford, Eglinton) had no flow recorder at all.
2. The size of a bed-moving flood depends on the bed substrate size as well as the flow. Small gravels will move in smaller flows than large cobbles. Since



substrate composition can vary considerably within a river, any estimate of a bed moving flow must be an average only.

3. It is extremely difficult (and dangerous) to measure the actual size of a bed-moving flood.

We relied on literature estimates based on standard flow statistics. For example, out of a wide range of hydrological variables, the frequency of floods greater than three times the median flow was found to correlate best with mean periphyton biomass in a river (Clausen and Biggs 1997). In a more recent study, Clausen and Plew (2004) found that the size of a theoretically determined bed moving flood in many New Zealand rivers was best predicted from a proportion of the mean annual maximum flow, although these authors cautioned that there is high variability in the relationship. In the absence of other validated suggestions, we used both of these approaches.

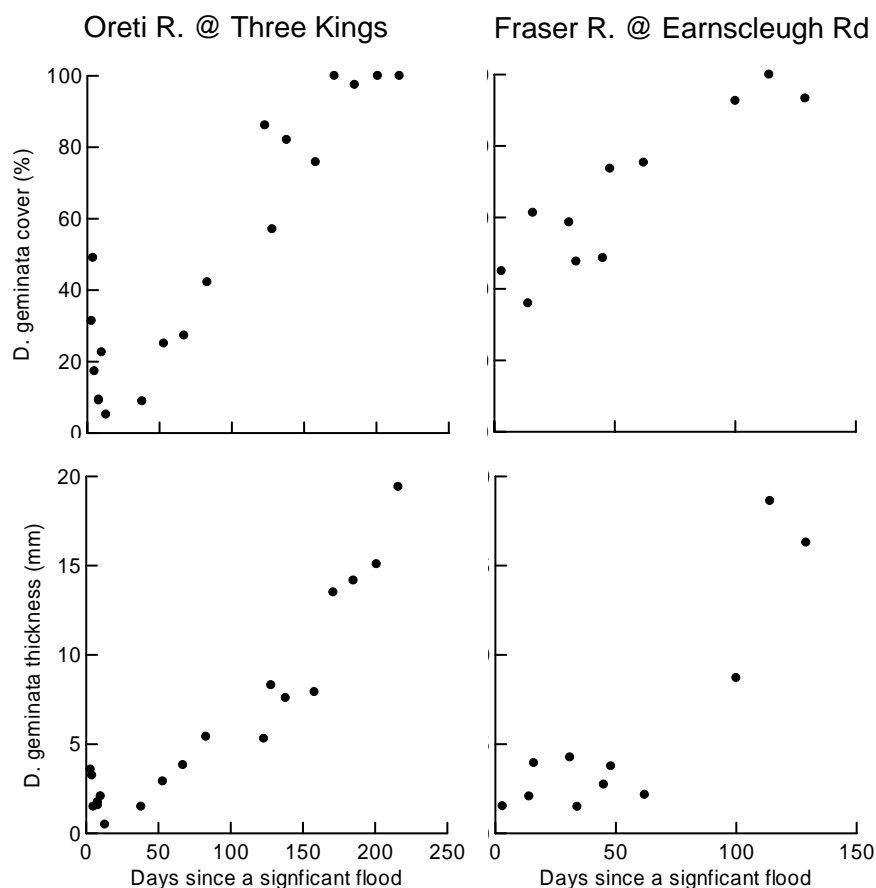
The procedure for deriving the number of days since the most recent significant (i.e., bed-moving) flood was as follows.

- Data were obtained from operating flow recorders on rivers included in the survey. Where rivers had no recorder, data were used from the nearest recorder on a river estimated to have similar hydrological characteristics (e.g., corresponding rainfall events). The 26 hydrological recorders used in the analysis are listed in Table 3.1.
- Flow statistics were calculated from each flow record using both the entire record and data for the past 10 years. In all cases the statistics were similar therefore the full record was used in the calculations. Statistics were calculated from mean daily flows.
- The size of significant floods was estimated in two ways: (a) 3 x median flow (3MED); (b) 0.18 x mean annual maximum flow (0.18MAM). The latter is the formula derived by Clausen and Plew (2004), which estimates the size of a flood capable of moving all particles on the river bed up to 50% of the median size. We assume that this degree of movement would cause sufficient abrasion to remove biomass from larger particles. For natural rivers with no or very low lake inputs, we found that these two estimates were normally similar. However, where flow recorders were close to lakes or in regulated rivers, 3MED was usually much higher than 0.18MAM. An exception was the Waiau at Manapouri Structure TW, where the low minimum flow and occasional very large lake spills caused 0.18MAM to be over twice 3MED. In all cases, we used the higher of the two estimates as the significant flood level because

the lower estimate was always unrealistically low (e.g., close to the mean flow). With few exceptions, the flood level used was 3MED.

- Every survey site was associated with the flow recorder determined to best reflect the flow conditions at the site.
- Using tables of mean daily flows, we counted the number of days back from the date of the survey to the designated flood level. This was then entered as a variable in the model dataset. Values (days) are shown in Appendix 1.

We have been able to verify the validity of this approach by plotting fortnightly records of cover and thickness at sites on the Oreti, Mararoa, Lower Waiau and Fraser rivers against days since the most recent bed-moving flood (data from Larned et al. in prep., and M. Dale, Otago Regional Council, pers. comm.) At all sites we found significant positive relationships between days since flood and both thickness and % cover. Examples are shown in Figure 3.3. Note that these data were not used in model construction.



**Figure 3.3** Examples of *D. geminata* average percentage cover and thickness at two sites plotted against days since a “significant” flood, calculated as three times the median flow. Note the obvious relationships. (Fraser River data from M. Dale, Otago Regional Council.)



### 3.4. Environmental predictors

The environmental dataset associated with the REC network includes at least 120 variables, developed during the production of a multivariate river classification for the Department of Conservation (Wild et al. 2005; and see examples in Leathwick et al. 2005, submitted; Leathwick and Julian 2007). Variables considered either redundant or irrelevant were dropped prior to model fitting. Because of the large number of available variables an element of subjectivity was unavoidable. The procedure for selecting variables included: (1) first dropping variables that were clearly biologically irrelevant to *D. geminata* (e.g., variables applicable downstream of stream sections); (2) producing a correlation matrix of all remaining variables, and dropping one or more variables in sets that were closely intercorrelated for both the whole of New Zealand and the data from the 146 sites; (3) selecting flow-weighted or unweighted variables, as appropriate (see Table 3.2); and (4) running trial models using progressively smaller sets of variables to check the effect of dropping variables. This procedure produced a set of 18 variables likely to have strong functional links with variation in the abundance of *D. geminata*. These described hydrological, climatic, physical (e.g., substrate) and chemical conditions in rivers (Table 3.2).

Hydrological variability has long been recognised as a major determinant of algal biomass in rivers (Biggs and Close 1989) because of the scouring effects of high flows through direct shear stresses on attached algae in high water velocities, and the abrasive effect of suspended sediment and mobilised bedload. The opposite to flood scouring is enhanced algal growth in stable or prolonged low flows (Suren et al. 2003). *D. geminata* would be expected to respond to these factors in the same way as other attached algae. Although *D. geminata* appears to tolerate a very wide range of water velocities (Kilroy et al. 2006c), more rapid local velocities appear to be required for successful colonisation (personal observations). A range of variables accounted for these hydrological features.

Climatic variables mainly relate to temperature. *D. geminata* has often been reported to be a cool-water diatom. However, the species has been reported in a number of North American, European and Asian rivers in very warm areas, suggesting a much wider range of temperature tolerance than traditionally assumed.<sup>4</sup> Temperature variables were included in the modelling, but we also considered temperature data from overseas (Section 3.7). A further climate-related variable differentiated light intensities through stream section shading by riparian vegetation.

Substrate composition is a significant determinant of biomass potential for *D. geminata*: large, hard, stable substrate particles favour larger and more persistent

<sup>4</sup> Note that in some cases at least where *D. geminata* has bloomed in areas with extremely hot summers, the affected rivers have been quite cold because of their origin from deep-release reservoirs (e.g., the Colorado River) (S. Spaulding, pers. comm.).

blooms than smaller more mobile substrates. Therefore, variables associated with rock hardness and upstream erosion and sediment transport were included.

Finally, we considered water chemistry variables. Although the chemical requirements for *D. geminata* are poorly known, the species has traditionally been thought to prefer low-nutrient waters (Krammer and Lange-Bertalot 1997, but see Kawecka and Sanecki 2003) and does appear to require some calcium (E.-A. Lindstrom, pers. comm.).

The variables used fall into two main groups:

*Segment variables:* these describe conditions within each river section. All such variables start with the prefix *seg...* For example, the variable *segShade* indicates the amount of shading received within each section.

*Upstream variables:* these describe conditions integrated from the entire catchment area starting at the most upstream point of the river section, and are prefixed *us....* For example, *usAvTWarm* is the mean air temperature in the warmest month averaged over the entire upstream catchment area.

One variable (*Warm\_us-seg\_stan*) was a combination of an upstream and a segment variable, calculated to remove the correlation between these, while preserving the unique information contained in one but not the other.

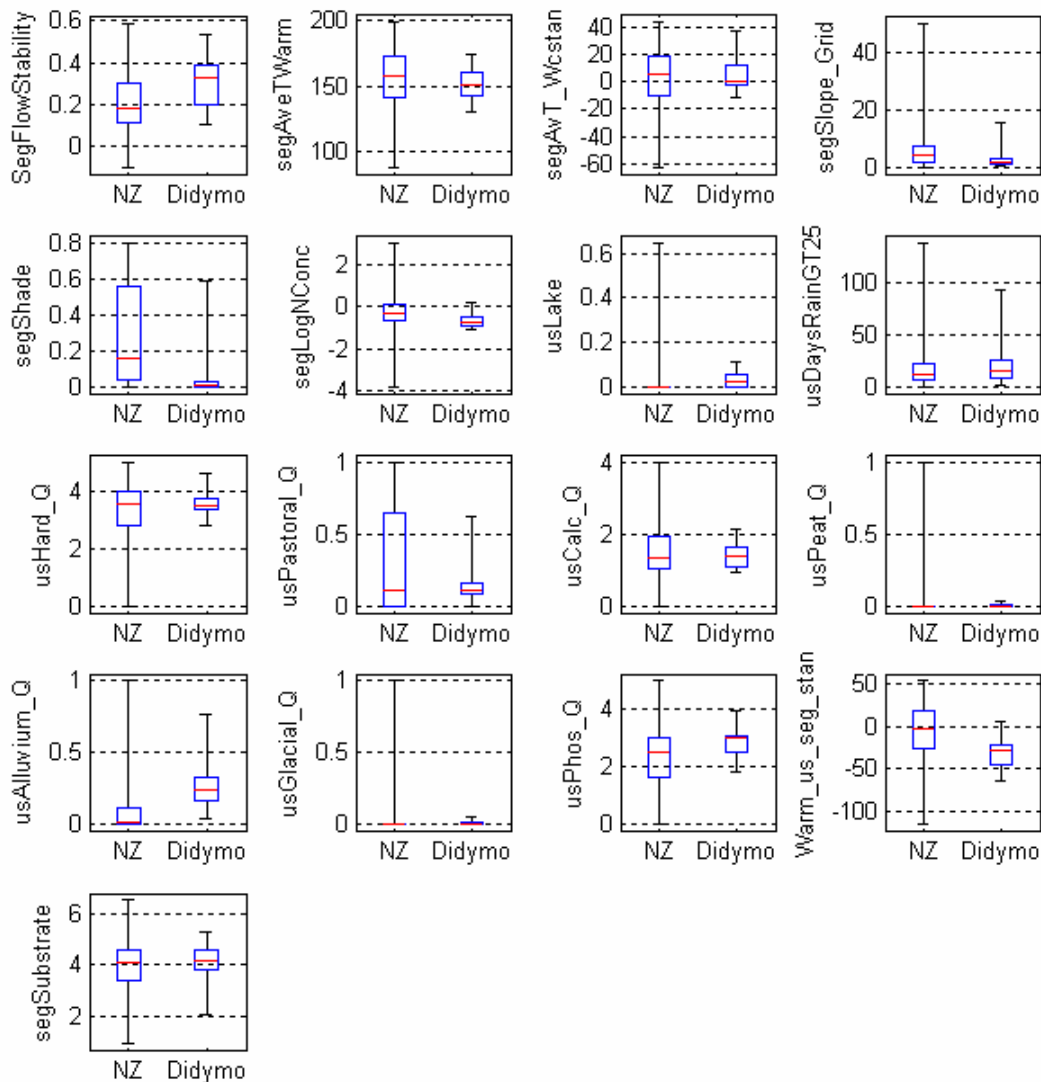
We also included the two local variables, *Subs\_index\_jl* (Section 3.3.2) and *Time\_sig\_flood* (Section 3.3.3)

All variables used in development of the models are summarised in Table 3.2.

**Table 3.2:** Variables used in development of models for predicting mean % cover and thickness of *D. geminata* in New Zealand rivers. Means and ranges for each variable for the modelled reaches and for the whole of New Zealand are shown in Figure 3.4. Variable names ending in \_Q are values have been weighted by the mean annual runoff at the reach, for a better estimate of the effect of catchment characteristics on instream characteristics.

Variable	Description, units	Significance
Segment variables		
<i>segFlowStability</i>	Ratio of the annual low flow to the annual mean flow.	Low values suggest fluctuating flows; values closer to 1 indicate more stable flows. High algal biomass is more likely to develop under stable flows.
<i>segAvTWarm</i>	Mean air temperature at the segment in the warmest month (January) (°C)	Temperature thought to be a major determinant of <i>D. geminata</i> distribution. This variable is related to water temperature, but also comparable with available international temperature data.
<i>segAvT_Wcstan</i>	Summer air temperature (warmest month) standardised with respect to winter air temperature (coldest month) (see Leathwick et al. submitted for details) (°C)	Indicates annual variability in temperature. Values are positive in areas with strong seasonal variation, e.g. continental climates. More equitable climates have negative values.
<i>segSlope_Grid</i>	Average within-segment slope based on a 30-m slope grid (angle)	Local slope can determine water velocity and river type (riffles, runs), both of which affect algal growth.
<i>segShade</i>	Amount of riparian shading estimated from the river size, solar angles, and expected vegetation height (Leathwick et al. 2005) (proportion)	Riparian shading restricts light at the stream bed, and thus can influence periphyton community composition.
<i>segLogNConc</i>	Stream nitrogen load estimated from CLUES, a leaching model combined with a regionally-based regression model (Woods et al. 2006) (ppm)	Nitrogen is an important plant nutrient. Algal growth may be stimulated by high concentrations, with different algae responding in different ways.
Upstream variables		
<i>usLake</i>	The ratio of rainfall within lake-buffered subcatchments to rainfall in the whole catchment (proportion)	The degree to which a river section is influenced by lake-fed waters. Lake inputs tend to stabilise river flows.
<i>usDaysRainGT25</i>	Mean number of days per year when rainfall in the upstream catchment exceeds 25 mm (days)	Low values indicate rivers prone to long stable-flow or low-flow periods, and high values indicate more flood-prone rivers.
<i>usHard_Q</i>	The proportion of hard rock in a catchment, derived from a 5-step scale of rock hardness adopted by the LENZ (1 = very low to 5 = very	High values are associated with larger, harder and more stable river substrates. Low hardness indicates high proportions of erodible, more mobile soft rocks (e.g., mudstones)

Variable	Description, units	Significance
	high)	
<i>usPastoral_Q</i>	Area of upstream catchment in pastoral (farming) land-use (proportion)	Increased nutrient inputs from farming affect both biomass and community composition of algae in stream (Biggs 1995)
<i>usCalc_Q</i>	Availability in surface rocks of calcium (Leathwick et al. 2005) (1 = low to 4 = very high)	Calcium is involved in adhesion in diatoms (Geesey et al 2000) and may be linked to presence / absence of <i>D. geminata</i> (E.-A. Lindstrom, pers. comm.)
<i>usPeat_Q</i>	Area of catchment occupied by peat (proportion)	High proportions of peat may lead to humic, acidic waters, which may be unfavourable for <i>D. geminata</i> . Low proportions may be more favourable than absence (M. Kelly, pers. comm.)
<i>usAlluvium_Q</i>	Area of catchment occupied by alluvium (proportion)	Alluvium (e.g., gravels) in the catchment suggest high erodibility and likely inputs of mobile substrate.
<i>usGlacial_Q</i>	Area of glaciers in upstream catchment (proportion)	While cool water may favour <i>D. geminata</i> , the fine silt in glacial waters may be unfavourable (attenuates light and covers surfaces)
<i>usPhos_Q</i>	Availability in surface rocks of phosphorus (Leathwick et al. 2005) (1 = very low to 5 = very high)	Phosphorus is an important plant nutrient. Algal growth may be stimulated by high concentrations, with different algae responding in different ways
Combined variables		
<i>Warm_us_seg_stan</i>	Upstream catchment mean air temperature in the warmest month standardised with respect to air temperature at the stream segment (°C).	Negative values indicate catchments that are colder (higher elevation) than average given the segment temperature; positive values indicate catchments with warmer than average (low elevation) headwaters, i.e., an indicator of the temperature/elevation difference between a segment and its upstream catchment.
REC surrogates for local measured variables		
<i>segSubstrate</i>	Weighted average of proportional cover of bed sediment in 7 categories, predicted from a model fitted to field data for 8600 sites from the New Zealand Freshwater Fish Database (Leathwick et al. 2007).	See Sections 5.7.2. and 3. 6. Our local substrate index ( <i>Subs_index_jl</i> ) was calculated in exactly the same way as <i>segSubstrate</i> . To run the predictions we substituted REC-wide <i>segSubstrate</i> for substrate index.
<i>365 / segFRE3</i>	<i>segFRE3</i> (annual frequency of floods > 3 x median flow, estimated for REC classes (Snelder et al. 2005)) converted to mean days between floods (days)	See Sections 3.3.2 and 3.6. Substituted for <i>Days_sig_flood</i> for the predictions.



**Figure 3.4.** Box plots showing the ranges of variables used in model development, for all of New Zealand (NZ) and for the surveyed river sections used to develop the models (Didymo). The boxes show the median (red line) and quartiles. The whiskers extend to the minimum and maximum of each range. The 18<sup>th</sup> variable is *days\_sig\_flood*. For this variable, the survey range was 1 to 400, and the NZ range (average scenario – Section 3.6) was 13.5 to 182.5 days (Snelder et al. 2005).

### 3.5. Modelling procedure

We developed models for *D. geminata* % cover and thickness separately because the error distribution of a combined index could not easily be accommodated using standard model specifications, e.g., Gaussian or Poisson distributions. For model development, we used boosted regression trees (BRT), a type of regression model (see Appendix 2 for details) that has its origins in machine learning. In recent comparative

studies (e.g., Elith et al. in review, Leathwick et al. 2006), BRT has substantially outperformed more conventional techniques such as Generalised Linear Models (McCullagh and Nelder 1989) and Generalised Additive Models (Hastie and Tibshirani 1990). Another novel technique, GARP (Genetic Algorithm for Rule-set Production), which has been used for this purpose elsewhere, was not considered because of its variable performance (Elith et al. 2006). Bayesian techniques might provide a useful alternative but, in our view, they are unlikely to offer any substantial advantages in performance over BRT.

All analyses were carried out in R (version 2.0.1 R Development Core Team, 2004) using the ‘gbm’ library of Ridgeway (2004), supplemented with custom-written functions (Elith et al., in review). For the thickness data, a BRT model was fit using the Poisson model family. For the cover data, a Gaussian family BRT was fit to the arcsin of the square root of the proportional cover. All models were fitted to allow interactions between predictors, using a tree complexity of 3, and with the learning rate set to fit 800–1200 trees (see Appendix 2). Ten-fold cross validation was used to determine the optimal number of trees for each model, i.e., that giving the maximum predictive performance. Initial models were first fitted for % cover and thickness using all 18 predictors. The predictor set was then simplified where possible by removing redundant predictors (Miller 2002). Because of the tendency of BRT models to over-fit the training data, the performance of all models was not assessed on the training data, but on predictions to sites that were withheld during cross-validation. For each model the predictive deviance was calculated, these values providing a measure of the goodness of fit between predicted and raw values when predicting to independent data, and we expressed this as a percentage of the null deviance.

The relative importance of the individual predictors was determined using a script in the ‘gbm’ library that sums, by predictor, reductions in error across all the individual regression tree rules. The influence of interactions between predictor variables was evaluated using custom-written functions (Appendix 2).

### 3.6. Predicting *D. geminata* % cover and thickness

Predictions were run in R on all reaches covering the entire REC network. We initially restricted the predictions to reaches with a stream order of 3 or greater since no rivers in the model dataset had a reach order lower than this (see comment in Section 3.1). As mentioned, we have fitted two models, separately predicting the potential cover and thickness of *D. geminata* mats.

Making predictions was complicated by issues with two variables, substrate index (*Subs\_index\_jl*) and days since flood (*Days\_sig\_flood*), that were measured in the field and calculated for each site, respectively. To run the predictions, we used the REC

variable *segSubstrate* as a surrogate for substrate index (Table 3.2). To transfer *Days\_sig\_flood* to the entire REC network, we used estimates of the annual frequency of floods at least three times the median flow (FRE3) derived by Snelder et al. (2005) for Climate and Source-of-flow levels of the REC. We converted the estimated mean FRE3 to mean days since flood ( $365 / \text{FRE3}$ ) then ran the model predictions for three scenarios:

1. 0.5 x mean days since a flood at least 3 times the median flow (shorter than average period of elapsed time since the most recent flood)
2. mean days since a flood at least 3 times the median flow (average ...)
3. 2 x mean days since a flood at least 3 times the median flow (longer than average ...).

### 3.6.1. Range extrapolation

An important constraint to generating predictions from the developed models was that the range of environmental conditions represented in the 146 data points used to generate the model was narrower than that in the New Zealand-wide rivers and streams for which we made predictions (Figure 3.4). We dealt with this problem by first running predictions allowing extrapolation beyond the training set range, recognising that the predictions would be more reliable in some areas than in others. To indicate where predictions are likely to be less reliable, we calculated for every reach the total departure of all model variables (as proportions) from their ranges in the original training dataset. All reaches falling within the survey range for all variables have values of 0 and have the most reliable predictions. Values greater than 0 should be treated as less reliable, with the degree of unreliability increasing as these values increase.

### 3.6.2. A potential test for the models

To check the models we have records from 36 sites where *D. geminata* has been found in the past 8 months. No % cover and thickness data are currently available from these sites, but this information may be available in the future. We extracted predictions for % cover and thickness for each days-since-flood scenario.

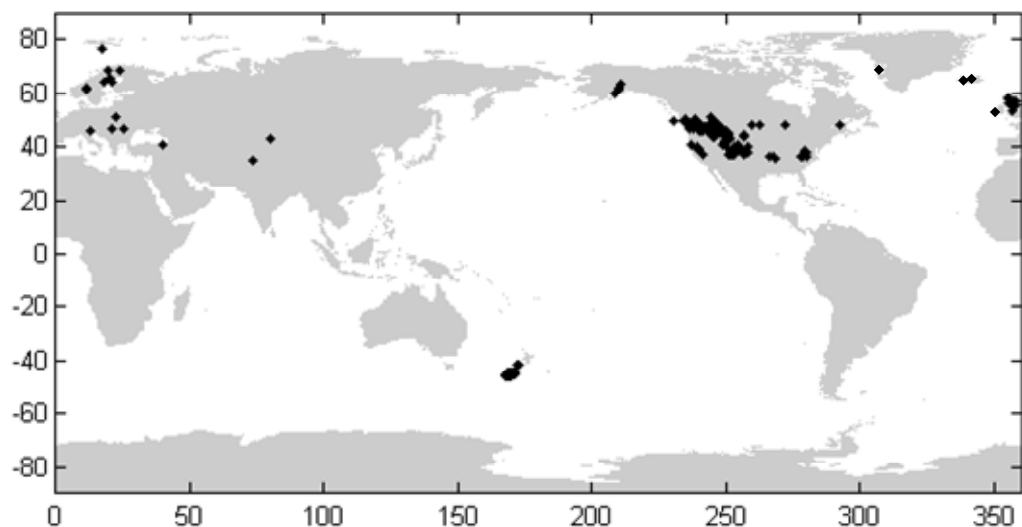
## 3.7. Temperature

As indicated in Section 3.1, because temperature is believed to be an important determinant in the distribution of *D. geminata*, we treated it as a special case when dealing with range extrapolation.



We obtained records of *D. geminata* presence from North America, Great Britain, and a range of other locations in Europe and Asia. We then retrieved a global dataset of terrestrial air temperatures from [http://www.jisao.washington.edu/data\\_sets/](http://www.jisao.washington.edu/data_sets/). The dataset used gave mean monthly air temperatures from 1950 to 1999 on a grid of 0.5 degrees of latitude/longitude. At mid latitudes this is a resolution of approximately 40 (lat) x 55 (long) km. All the overseas datapoints were then linked to their grid squares in the global dataset (Figure 3.5), and mean temperatures in the warmest and coldest months were extracted. We extracted the same data for the New Zealand data to compare the sites directly, and plotted all the locations (grid squares) on a scatter plot of warmest month vs. coldest month temperatures.

We also checked the entire range of temperatures in North American (our most comprehensive and geographically widespread dataset) to identify possible temperature limits for *D. geminata*. A similar exercise was carried out with a smaller dataset from Great Britain. The North American and British ranges were compared with the entire range of temperatures within New Zealand as well as to the range for other overseas and South Island sites positive for *D. geminata*.



**Figure 3.5:** Map of the world showing locations of international presence data for *D. geminata*.

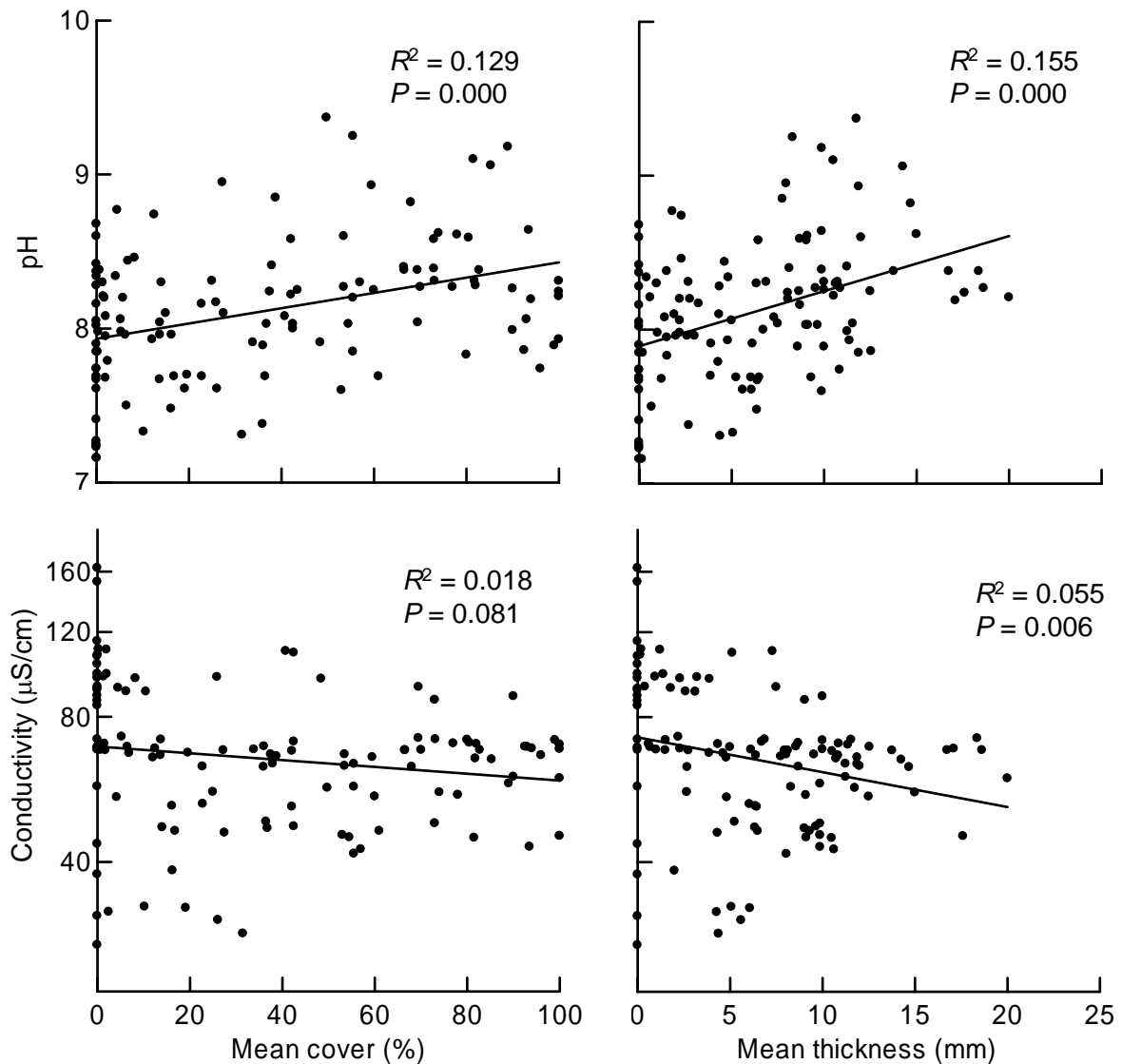
### 3.8. Results

#### 3.8.1. Variables important in determining the distribution of *D. geminata* in New Zealand

We first examined data collected at each site and looked for relationships with *D. geminata* % cover and thickness. We could not use some of these variables to develop the model because no surrogate variables exist in the REC dataset. However, any clear



relationships between, for example, *D. geminata* cover and conductivity or pH, would be worth reporting as these variables are very simple to measure. We found only weak positive relationships between pH and both *D. geminata* measures and very weak negative relationships between conductivity and *D. geminata* (Figure 3.6).



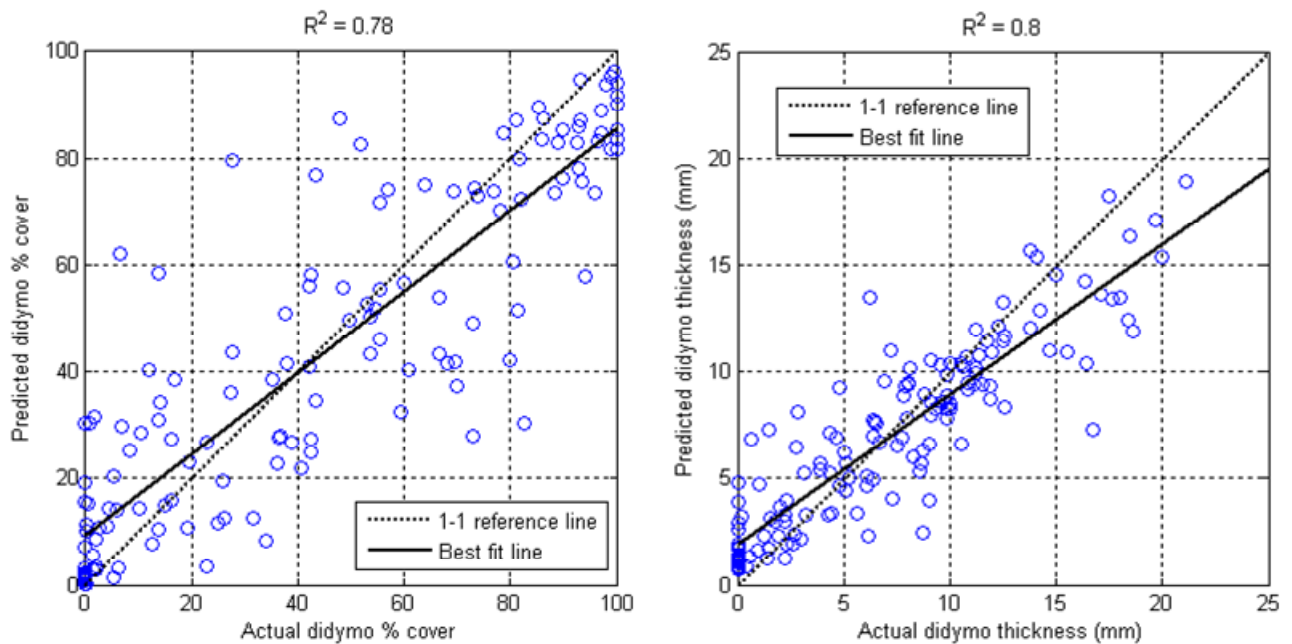
**Figure 3.6:** *D. geminata* % cover and thickness plotted against measured conductivity and pH at each site.

Predictive variables included in the final models for *D. geminata* % cover and thickness are listed in Table 3.3. The shapes of response curves for all variables included in the models are shown in Appendix 3. Figure 3.7 shows model performance, as observed values versus predicted values for the training data set. Ten-

fold cross-validation was used to estimate the predictive performance of the models. The cross-validated  $R^2$  values of 0.52 (% cover) and 0.48 (thickness), indicate that the models can be expected to account for about half of the variability in *D. geminata* % cover and thickness when predicting to new sites.

**Table 3.3:** Percentage contributions of predictors of *D. geminata* % cover and thickness used in the final models

Cover model			Thickness model		
		%			%
1	usLake	32.3	1	usLake	29.0
2	usHard_Q	15.1	2	subs_index_jl	10.6
3	days_sig_flood	13.2	3	segAvT_Wcstan	9.8
4	subs_index_jl	9.9	4	days_sig_flood	8.9
5	segAvT_Wcstan	6.7	5	usDaysRainGT25	7.9
6	segSlope_Grid	5.7	6	usHard_Q	5.6
7	usPastoral_Q	5.0	7	segLogNConc	5.3
8	usDaysRainGT25	4.8	8	segSlope_Grid	5.3
9	usCalc_Q	4.2	9	Warm_us_seg_stan	5.3
10	segAveTWarm	3.2	10	segAveTWarm	4.6
			11	usPastoral_Q	4.2
			12	usCalc_Q	3.6



**Figure 3.7:** Observed vs. predicted values for the two models.

In both models *usLake* was by far the most important predictor (Table 3.3), with *D. geminata* % cover and thickness strongly and positively associated with the amount of lake area in the upstream catchment (see Appendix 3). A substrate-related variable was next most important in both models but was *usHard* for % cover and was *sub\_index\_jl* for thickness. The latter contributed about 10% to both models. Days since a significant flood contributed 13% and 9% to the % cover and thickness models, respectively. A more general descriptor of flow variability (*usDaysRainGT25*) also featured in both models. The variable *segAvT\_Wcstan* was, respectively, fifth and third most important and suggested that higher cover and thicker growth are associated with more continental climates (i.e., hotter summers and colder winters). The direct temperature variable *segAveTWarm* contributed only a small percentage to both models. Reach slope contributed approximately 5% to each model, with lower slopes more favourable (see Appendix 3).

As expected, no single water chemistry variable contributed a high percentage to the model. However *segLogNConc* and *usPastoral* together accounted for almost 10% of the thickness model. Higher values of the former variable were associated with low thickness, while the response to *usPastoral* was unclear.

Interactions between variables were identified in both models. For % cover, *usLake* interacted with *days\_sig\_flood* so that very low % cover was found at sites with low or no lake influence and a low number of days since a flood, but a strong lake influence tended to counteract the effects of recent floods, i.e., high cover could persist even relatively soon after a flood. There was also a tendency for more extensive cover at sites with high lake influence if *usHard* values were also high (i.e., harder rock substrates). For thickness, we found a similar interaction between *usLake* and *days\_sig\_flood*, and also *segAvT\_Wcstan*. High thickness was associated with high lake influence but this effect was reduced in areas where *segAvT\_Wcstan* was lower (i.e., in more equitable climates). See Appendix 3 for examples of interaction plots..

### 3.8.2. Predictions

The DPM on the following pages show predicted average % cover (DPM\_cover) and thickness (DPM\_thickness) for the three days-since-flood scenarios (Section 3.6). Note that the predictions are for average % cover and mat thickness in wadeable parts of river reaches, i.e., in areas where data collection is possible.

The DPM (Figures 3.8 – 3.13) show predictions of greatest % cover and thickness in eastern and southern parts of the South Island, but with lower % cover and thickness in many coastal areas. Rivers in inland Otago and south Canterbury are predicted to develop high cover in long periods without floods (orange areas in Figure 3.10). Lake-fed rivers throughout the South Island have maximum cover under all scenarios. The

North Island is predicted to be much less susceptible than the South Island, though some lake-fed rivers may develop high cover levels and thick mats. After long periods without floods, much of inland central and eastern North Island is predicted to develop moderately thick mats (light green area in Figure 3.13), but this is associated with low levels of cover (less than 25% and often less than 10%).

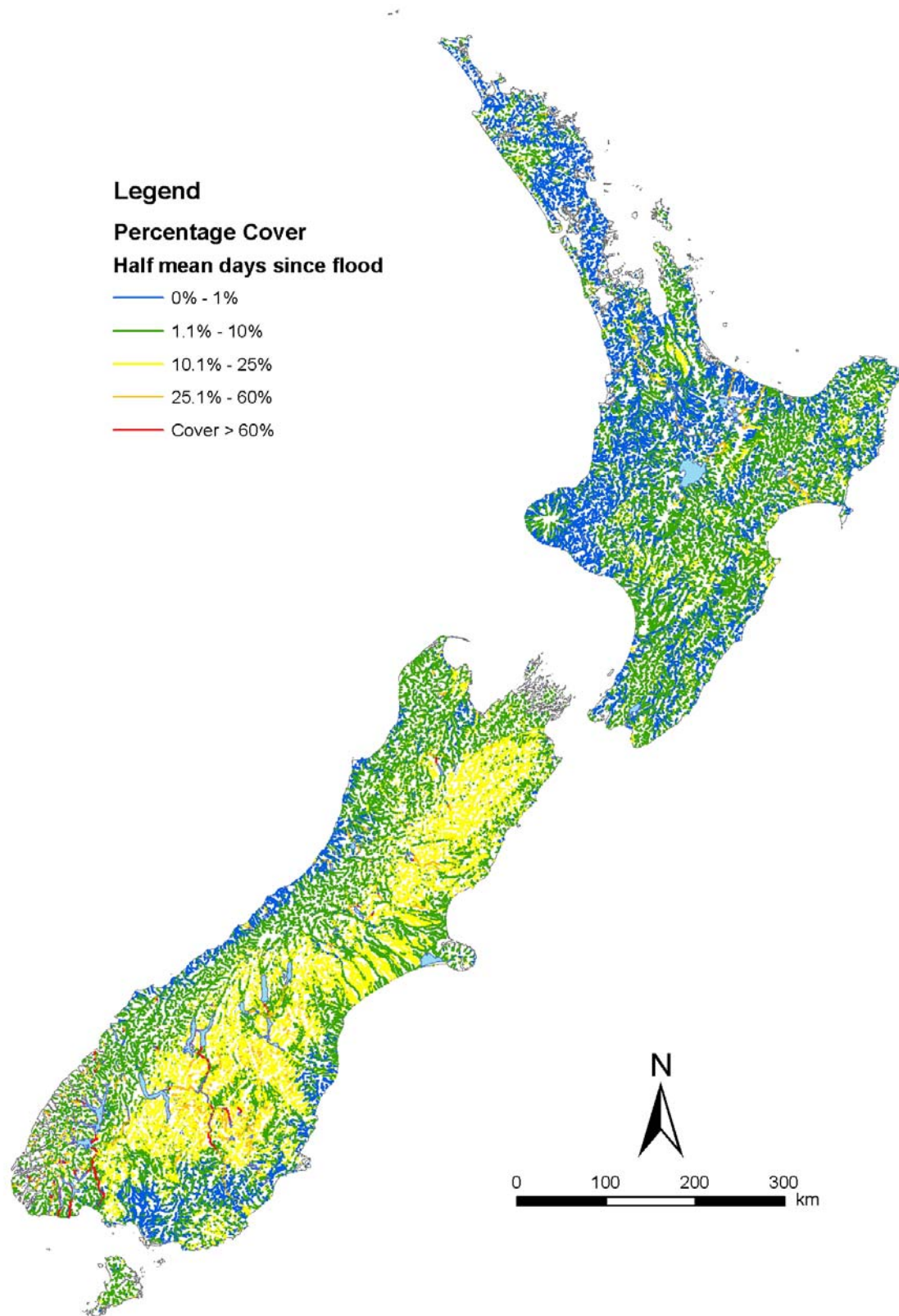
In rivers approaching the highest cover and thickness, predicted values were often slightly higher under the mean flood scenario than the twice mean flood scenario, although not significantly different (e.g., as illustrated in Section 5.4). A plausible explanation for this, supported by observations of the sloughing behaviour of heavy didymo mats, is that there is a maximum mat thickness above which increasing thickness is self-limited due to the mass of the mat exceeding the strength of the stalks. When these maxima are reached, mat extensions slough away and temporarily decrease mat cover and thickness until regrowth occurs.

Predicted mat thickness appears low (maximum of 21 mm) compared to anecdotal reports of mats 10 – 20 cm thick. This was due to two factors: (1) we measured average thickness at 20 points across a 30 – 40 m reach, whereas maximum thickness was usually ~1.5 x average thickness; and (2) mat thickness was measured or estimated from the upper surface of rocks lifted from the water (i.e., colonies somewhat collapsed), whereas anecdotal thicknesses may refer to the trailing edges of colonies. When interpreting the predictions it should be noted that the top category of mat thickness represents very substantial growth.

A single unreliability map (DPM\_unreliability) applies to both models because 10 predictive variables are common to both (Tables 3.3) and the extra two variables in the thickness model made minimal difference to the distribution of reliability. Maximum unreliability was 5.37, meaning that all 12 predictive variables were, on average, almost 45% outside the range of the model data. Any predictions in reaches where unreliability >0 (all except the green reaches in Figure 3.14) should be treated with caution.

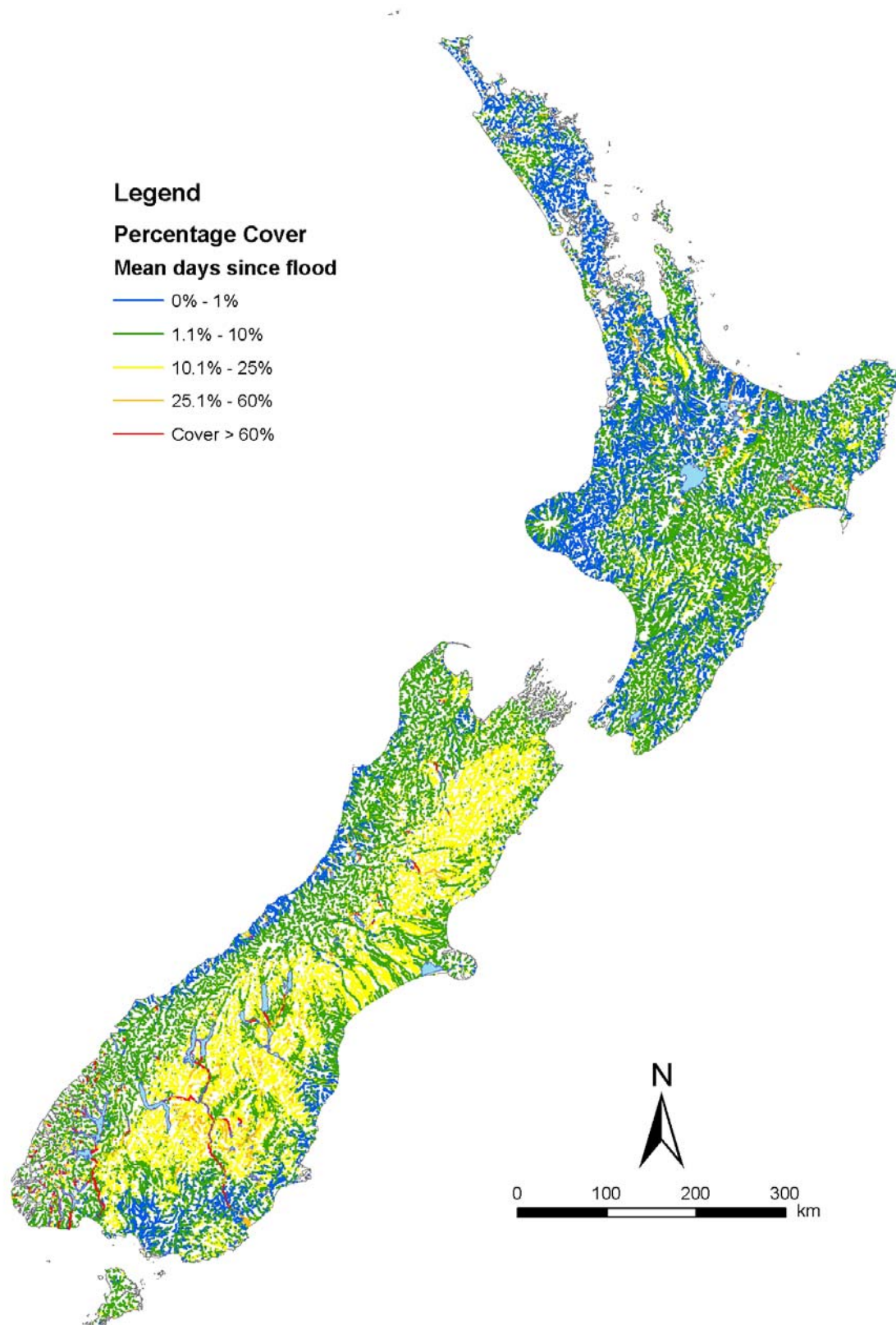
Greatest unreliability of predictions occurs in some coastal areas of the North Island and parts of the central, south-eastern and southern coastal areas in the South Island. Predictions in the central part of the Southern Alps are also unreliable (lilac and red reaches in Figure 3.14). To quantify this, 30.0% of all reaches fell within the model range (unreliability = 0), 16.1% in the North Island and 40.6% in the South Island.

Note that DPM shown in Figure 3.8 – 3.14 are examples only. Larger scale maps can be produced for practical applications. See Section 5 for an example. It is also possible to run the predictions for all stream orders. In all cases, the unreliability index should be considered along with the predictions.

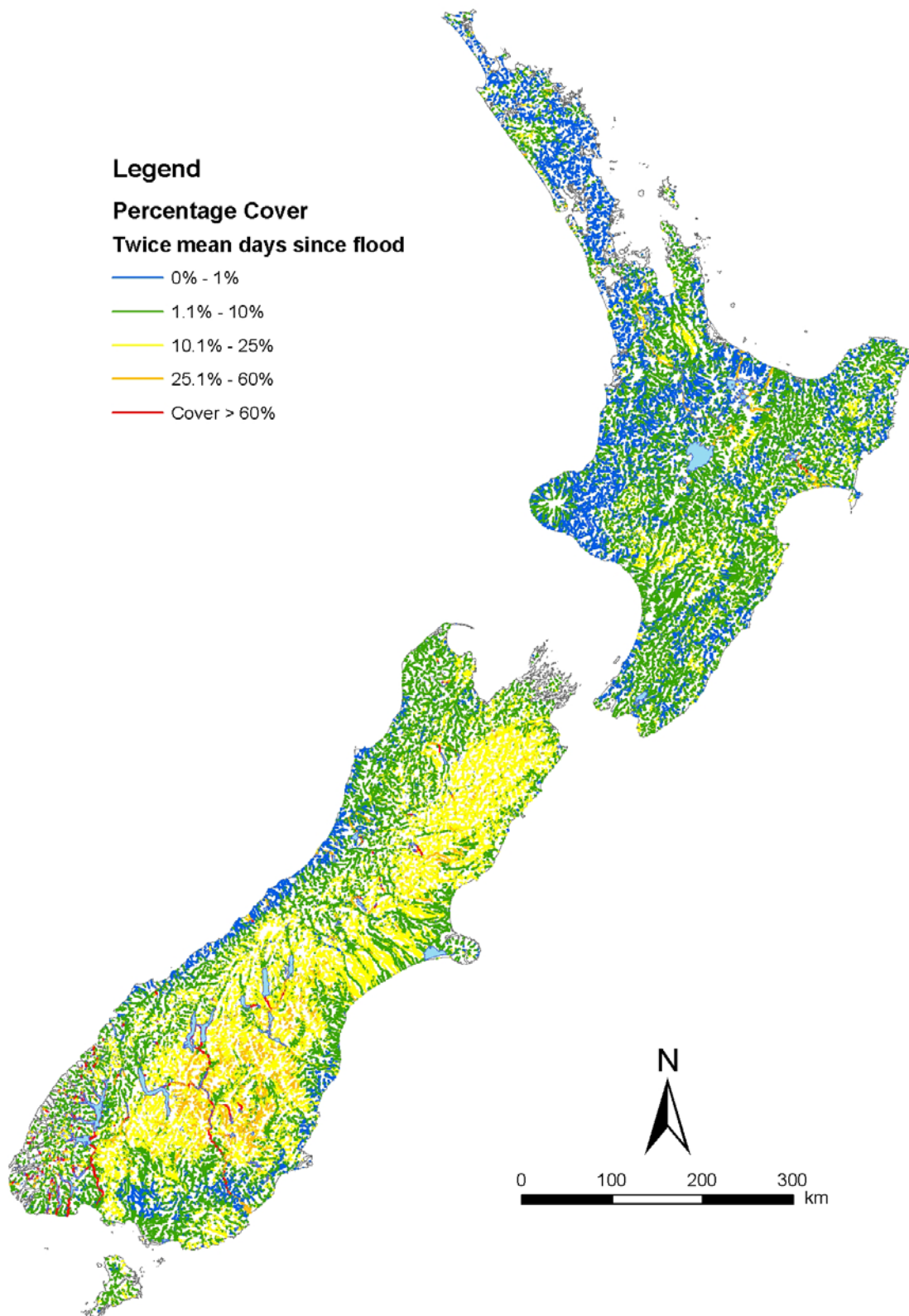


**Figure 3.8:** **DPM\_cover\_half mean flood** = predictions of % cover of *D. geminata* at half the estimated mean days since a flood of at least three times the median flow (best-case scenario)

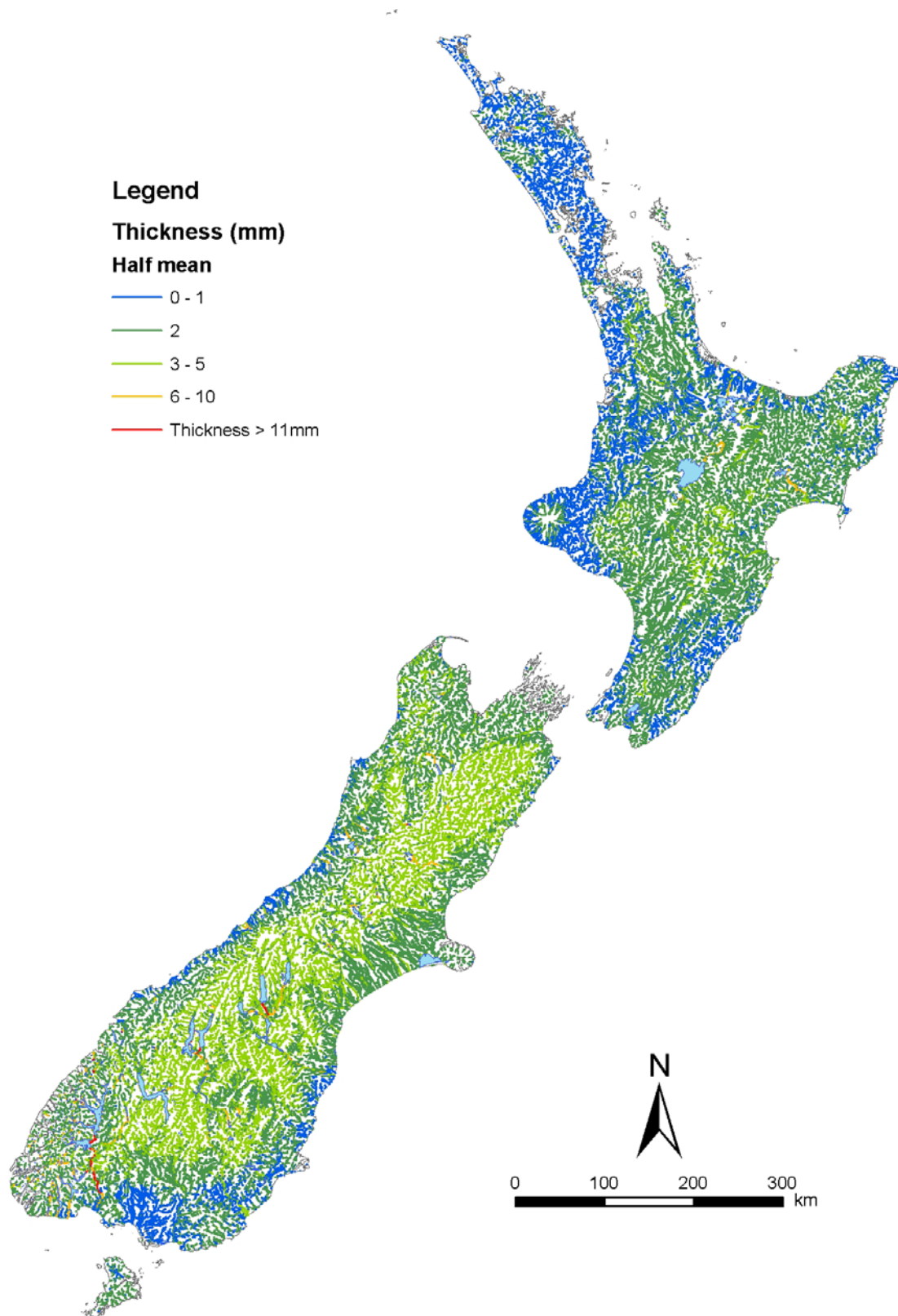




**Figure 3.9:** **DPM\_cover\_mean flood** = Predictions of % cover of *D. geminata* at the estimated mean number of days since a flood of at least three times the median flow (average scenario)

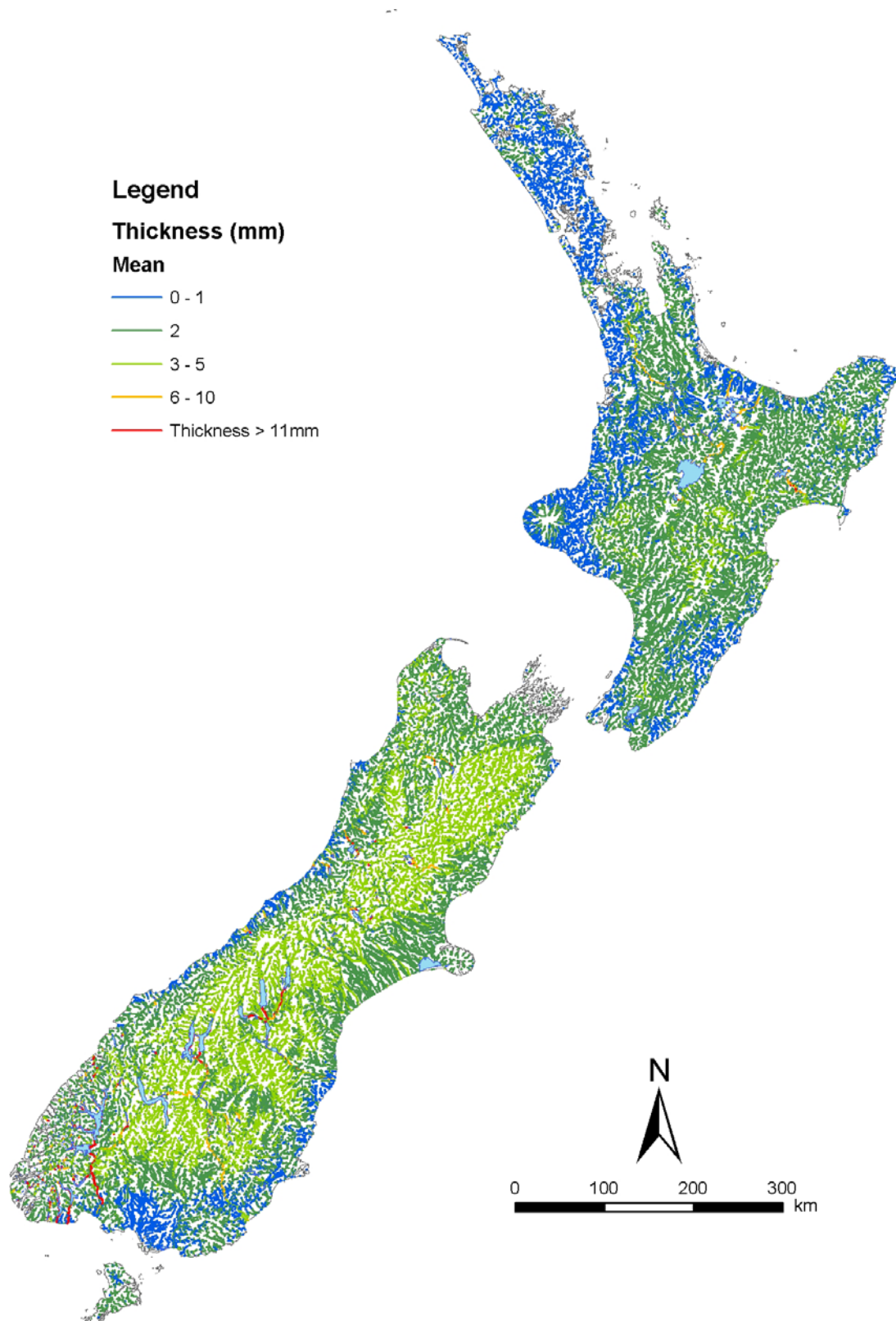


**Figure 3.10:** **DPM\_cover\_twice mean flood** = Predictions of % cover of *D. geminata* at twice the estimated mean number of days since a flood of at least three times the median flow (worst-case scenario).

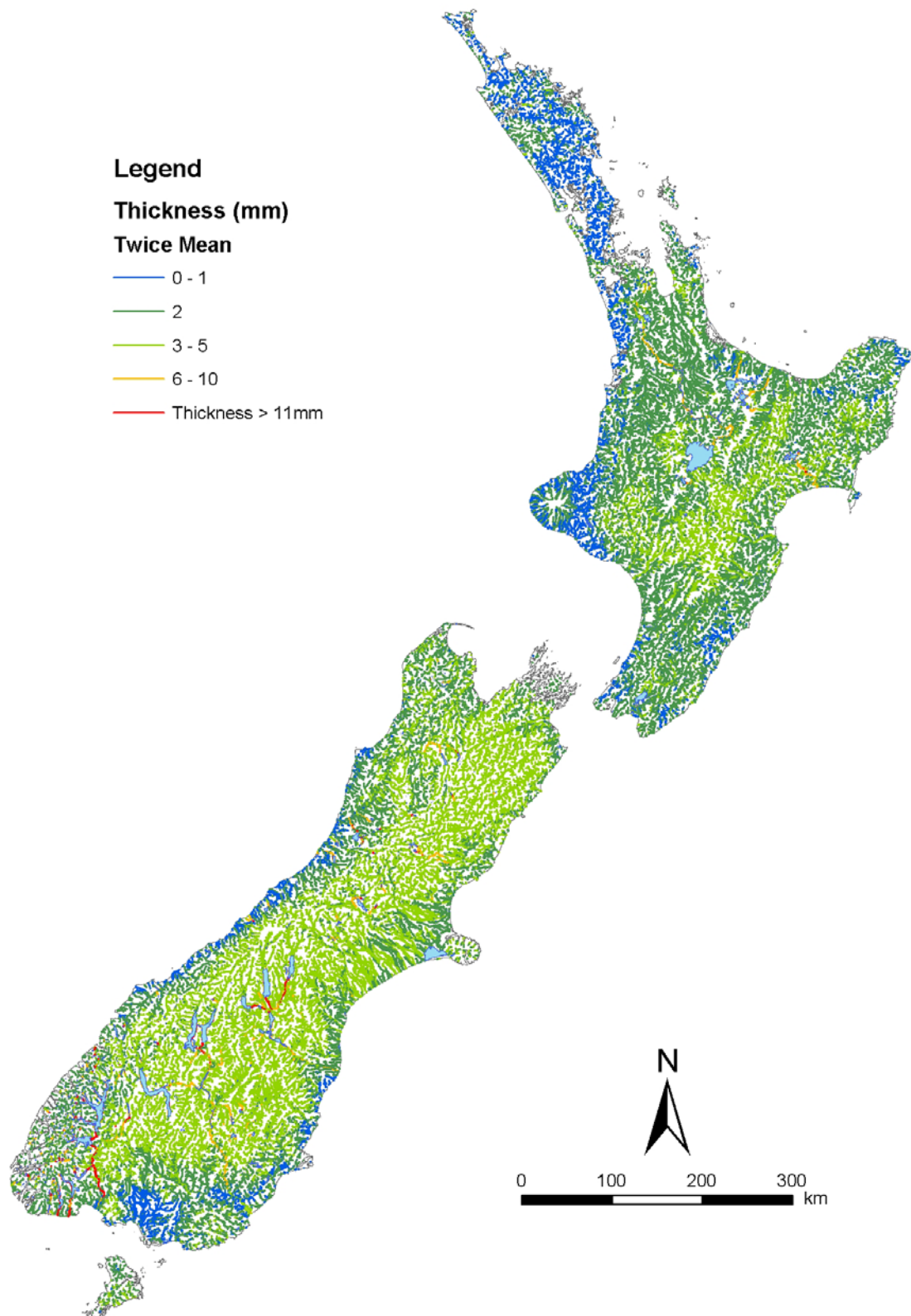


**Figure 3.11:** **DPM\_thickness\_half mean flood** = Predictions of thickness (mm) of *D. geminata* at half the estimated mean number of days since a flood of at least three times the median flow (best-case scenario)

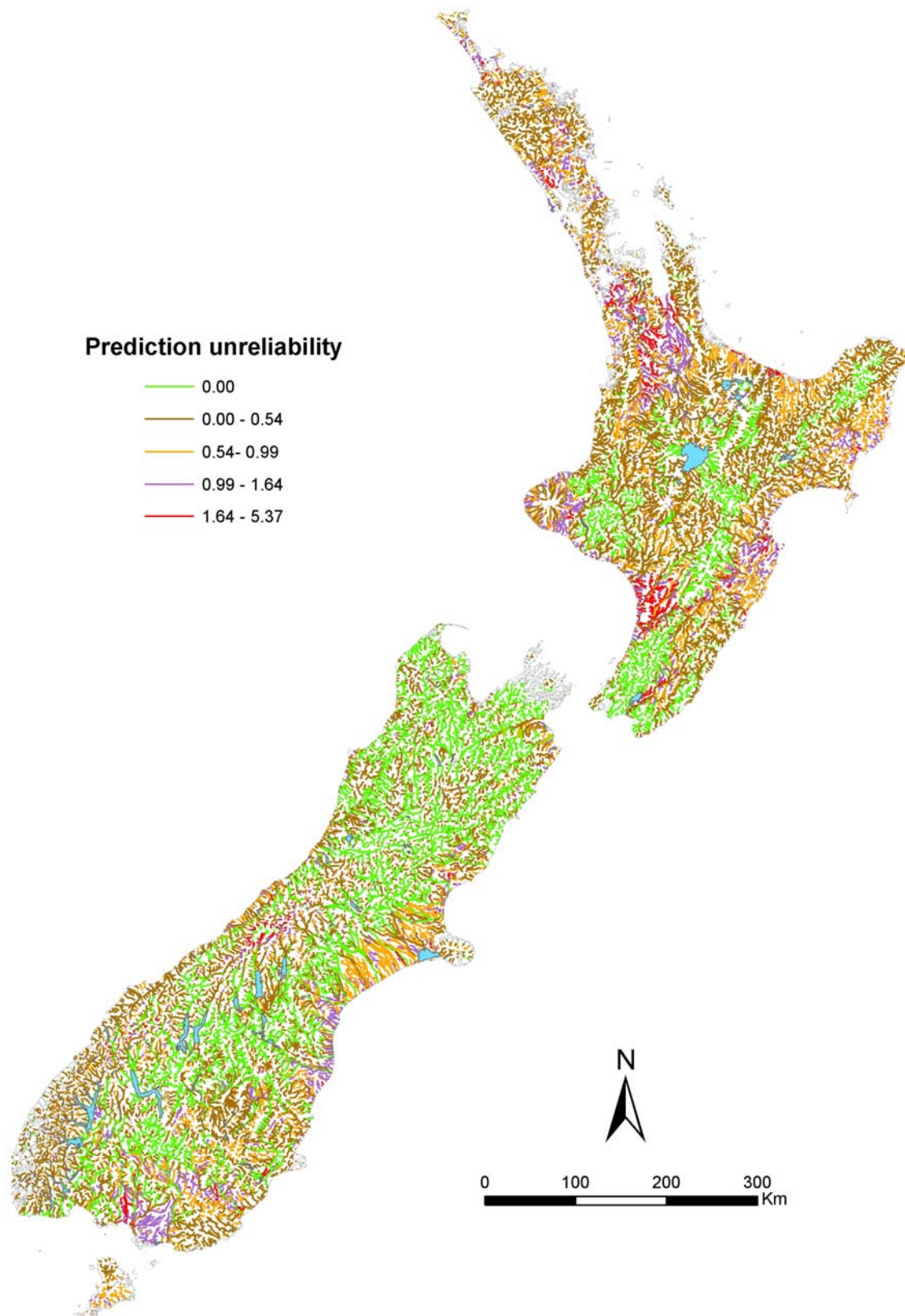




**Figure 3.12:** **DPM\_thickness\_mean flood** = Predictions of thickness (mm) of *D. geminata* at the estimated mean number of days since a flood of at least three times the median flow (average scenario)



**Figure 3.13:** **DPM\_thickness\_twice mean flood** = Predictions of thickness (mm) of *D. geminata* at twice the estimated mean number of days since a flood of at least three times the median flow (worst-case scenario)



**Figure 3.14:** **DPM\_unreliability** = map the sum of proportions of all variables lying outside the range of the modelled reaches. Unreliability is 0 if values of all variables lie within the range of the reaches used to construct the models. Higher values are increasingly unreliable.



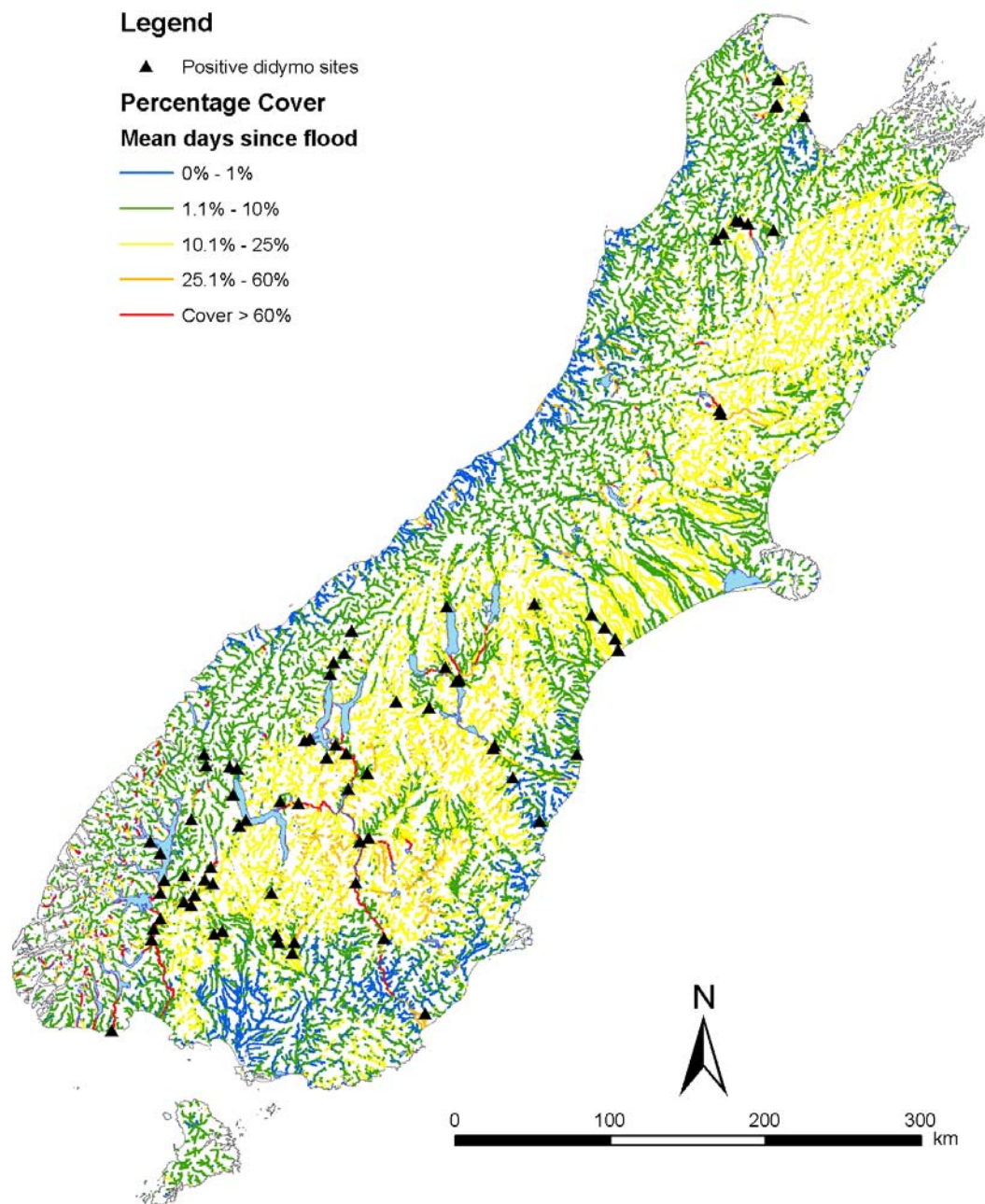


Figure 3.15 Locations of all South Island sites positive for *D. geminata* overlaid on the DPM\_cover\_mean flood (predictions for % cover at the estimated mean number of days since a flood of at least three times the median flow).

### 3.8.3. A potential test for the models

Table 3.4 shows model predictions at 36 sites where *D. geminata* has been identified the past ~8 months. All positive site locations are shown in Figure 3.15. Comments on the ual cover can be added when the information becomes available.

**Table 3.4:** Predictions of *D. geminata* % cover and thickness at 36 sites where its presence has been confirmed from observations of visible mats, or from examination of samples microscopically. Sites are arranged alphabetically by catchment. The date refers to the first confirmation of *D. geminata* at the site. Predictions are given for three scenarios based on the estimated mean number of days at that site between floods greater than three times the median flow (see Table 3.2): half the mean, mean, and twice the mean. Sites shaded grey had environmental variables outside the range in the model dataset, so the predictions are less reliable than those with unreliability = 0. We do not yet have field measurements of actual cover and thickness at these sites, but this can be added into the “Comments” column as the information comes to hand.

Catch.	River	Site	Date	Unreliability	Percent cover			Thickness (mm)			Comments
					mean/2	mean	2mean	mean/2	mean	2mean	
Aparima	Hamilton Burn	Hamilton Burn bridge	7-Aug-06	0.17	10	10	12	2	2	3	
Buller	Gowan	1 km u/s Buller confluence	11-May-06	0.00	52	65	64	6	9	8	
Buller	Matakitaki	1 km u/s SH 6	15-May-07	0.00	5	5	5	2	2	2	
Clutha	Motatapu	nr Mt Aspiring Rd, ~2km	1-Nov-06	0.00	24	24	26	4	4	5	
Clutha	Diamond Cree	Glenorchy	11-Dec-06	0.00	46	46	48	4	4	5	
Clutha	Matukituki	U/s of Cattle Flat	3-Aug-06	0.00	10	10	12	2	2	3	
Clutha	Lindis	Ardgour	29-Jan-07	0.00	32	32	38	4	4	6	
Clutha	Greenstone	between fishtrap and creek	26-Jan-07	0.00	1	1	2	1	1	1	
Clutha	Kawarau	Chards Rd	29-Jan-07	0.00	54	62	60	5	7	7	
Clutha	Manuherikia	Galloway Road	13-Feb-07	0.00	5	5	8	2	2	3	
Clutha	Shotover	Tuckers Beach	18-Apr-07	0.00	13	13	15	3	3	4	
Clutha	Young	Ram Flat	3-May-07	0.00	8	8	10	3	3	3	
Clutha	Cardrona	Mt Barker	3-May-07	0.00	28	29	34	3	3	5	
Clutha	Dart	Main Road bridge	4-May-07	0.00	1	1	2	2	2	2	
Haast	Haast	Burke Flat airstrip d/s Burke	16-Feb-07	0.00	9	9	11	2	2	3	
Hurunui	Hurunui	u/s Sth Branch confluence	15-May-07	0.00	55	69	68	7	10	9	
Hurunui	Hurunui	Jollie brook Access	15-May-07	0.00	56	70	68	7	10	9	
Kakanui	Kakanui	u/s SH1 BR	10-May-07	0.83	1	1	1	1	1	1	
Kakanui	Maerewhenua	Maerewhenua	10-May-07	0.00	9	9	13	2	2	4	
Mataura	Mataura	Ardlussa	8-May-07	0.00	15	15	17	2	2	2	
Motueka	Motueka	300 m u/s SH60 br	7-Feb-07	0.02	5	5	6	2	2	2	

Catch.	River	Site	Date	Unreliability	Percent cover			Thickness (mm)			Comments
					mean/2	mean	2mean	mean/2	mean	2mean	
Opihi	Opuha	Stony Creek above n_opuha confl	24-Apr-07	0.00	21	21	24	2	2	3	
Rangitata	Rangitata	u/s of lagoon	11-May-07	0.00	6	6	7	2	2	3	
Takaka	Takaka	at Paynes ford	25-Jan-07	0.00	13	13	14	2	2	2	
Takaka	Takaka	bridge 2km u/s Harwoods	25-Jan-07	0.00	4	4	4	2	2	2	
Takaka	Takaka	at Harwoods	25-Jan-07	0.00	6	6	7	2	2	2	
Waiau	Ettrick River	above first rapid u/s river m	15-Jan-07	0.09	6	6	6	2	2	2	
Waiau	Snag Burn	between 1st and 2nd rapids	25-Jan-07	0.18	9	9	9	2	2	2	
Waiau	Wairaurahiri	next to private lodge	1-Feb-07	0.00	67	75	73	8	11	10	
Waihao	Waihao	Bradshaws Rd Br	23-Nov-06	0.00	2	2	4	1	1	2	
Waitaki	Tekapo	below vehicle crossing to Haldons	9-Feb-07	0.23	59	69	67	10	13	13	
Waitaki	Tekapo	Roughly halfway between Haldo	9-Mar-07	0.21	61	70	68	11	14	13	
Waitaki	Ohau	D/s Twizel confl	18-May-06	0.00	34	46	44	8	11	10	
Waitaki	Omarama Stm	Omarama	14-Aug-06	0.00	15	15	20	2	2	4	
Waitaki	Tasman	Glentanner Stn	30-Apr-07	0.69	34	34	34	5	5	5	
Waitaki	Hakataramea	1.3 km from SH82	10-May-07	0.00	11	12	16	2	2	4	

### 3.8.4. Temperature

The range of air temperatures (defined by mean temperatures in the warmest and coldest months) in New Zealand is very small compared to that in North America (Figure 3.16(a)), but is larger than that in Great Britain. *D. geminata* occurs only in the central part of the whole range in North America, and in the lower part of the range in Great Britain (Figure 3.16(b)). There are no records from locations in North America or Great Britain where the mean air temperature in the coldest month exceeds ~5 °C. However, mean summer air temperatures can be as high as 27 °C in North America. Other presence data from overseas fall within or below the range for North America, and the New Zealand presence data coincides with (and slightly exceeds) the range from Great Britain (Figure 3.16(c)). While the North Island range partly overlaps that of the South Island, there are many locations with both warmer summers and winters (Figure 3.16(d)).

No *D. geminata* records from New Zealand have mean coldest month air temperatures greater than ~5.5 °C. Interestingly, in our survey the only sites visited where mean temperatures in the coldest month exceeded 5.5 °C were in the lower Buller River, which had a mean temperature in the coldest month of 7.2 °C. No *D. geminata* was visible at any of the four sites in the lower Buller (“NZ absence” in Figure 3.16(c),(d), and sites 50 – 53 in Appendix 1).

## 3.9. Discussion

### 3.9.1. Modelling approach

In this section we have addressed the following objectives (listed in Section 1):

2. Predict the potential growth of *D. geminata* (e.g., as percentage cover and mat thickness) in all New Zealand river reaches and present in colour-coded map format (including known sites).
3. Determine which river reaches will support a defined bloom of *D. geminata* and present in colour-coded map format.
4. Based on the model developed for rivers, identify the important environmental factors that influence *D. geminata* growth and lead to blooms (including limiting factors) in river reaches.



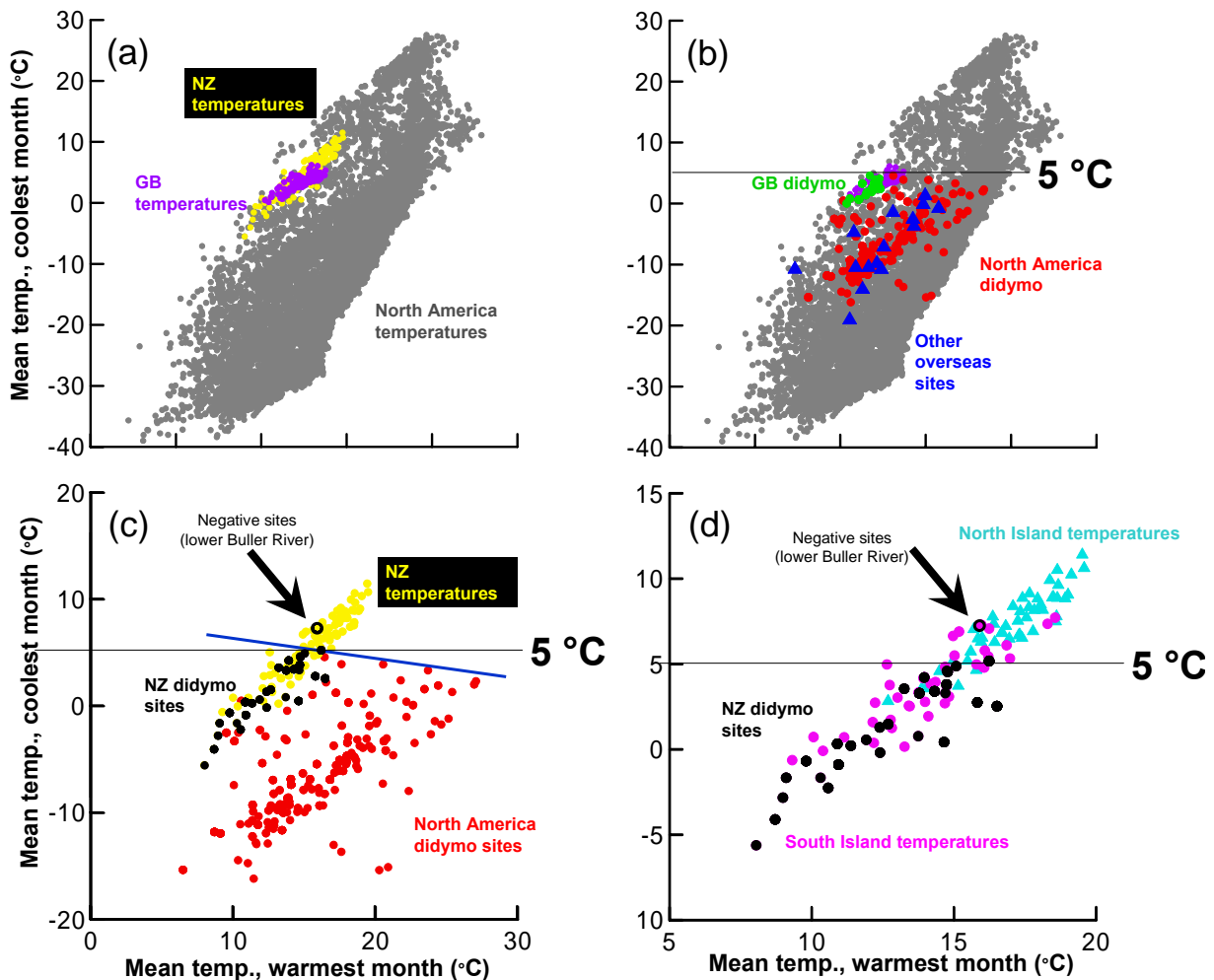


Figure 3.16: Temperature comparisons using data from a global dataset of terrestrial air temperatures. Mean temperatures in the warmest month are plotted against mean air temperatures in the coldest month. (a) Data from the whole of North America (including Mexico and Canada) compared with the range for the whole of New Zealand (yellow) and Great Britain (purple). (b) Known *D. geminata* range in North America (red), occupying central part of the range only. The known *D. geminata* ranges for Great Britain (green) and other global sites (blue triangles, see Figure 3.5) are included for comparison. (c) Detail showing *D. geminata* range in North America (red) and New Zealand (black) and other locations compared with the temperature range for the whole of New Zealand. (d) Current range of *D. geminata* in New Zealand (survey data and other positive sites) compared to the temperature ranges for South Island (pink) and North Island (turquoise).

It is common practice in modelling species distribution to focus on probabilities of presence vs. absence of the target organism. We have taken a different approach in that our models directly predict both average % cover and average thickness of *D. geminata*, rather than presence versus absence. This approach yields directly usable predictions rather than probabilities, which can be difficult to interpret. Further advantages are as follows.

- Determining presence vs. absence of a micro-organism presents problems not encountered in larger organisms. Sessile macroscopic organisms are either present (i.e., they can be observed) or they are absent. For a micro-organism, does presence of a small number of cells (invisible to the naked eye) constitute presence? Or does presence imply the existence of visible colonies? If we take the former view, then for presence to be established, we must collect samples from all sites and examine them microscopically for *D. geminata* cells. This has been the approach taken in the presence data for *D. geminata* from other parts of the world (Section 3.7). Thus many of the records from North America are records of the presence of *D. geminata* cells in composite periphyton samples. How many of these samples were from locations where *D. geminata* was abundant enough to be visible or problematic is unknown. The existence of visible colonies is arguably more relevant to management because it is only once colonies become visible that this micro-organism becomes a problem. This is dealt with further in Section 5.2.
- Determining presence in a riverine micro-organism presents particular problems. For example, we found in our survey that *D. geminata* could be established upstream in a river, but be visually absent downstream, except as dead fragments not attached to the substrate. Previous work has shown that once colonies are established upstream, cells can be detected microscopically in samples collected farther downstream even in the early stages of colonisation (Kilroy and Dale 2006). Thus microscopic presence of the species at a river site does not necessarily imply presence, in the sense that the local conditions are suitable for establishment and growth. In other words, the required environmental conditions for survival are not met, even though cells are present. This is the same problem as differentiating between a species' "source populations" and "sink populations" when modelling species distributions (Austin 2002). For *D. geminata* (or other similar microorganisms), a distribution model based on microscopic presence would predict presence in environments where the species could never grow.
- From our observations of *D. geminata* in rivers we believe that it is reasonable to assume that habitat suitability is *directly related* to percentage cover and thickness of colonies. For example, we established that % cover and mat thickness are positively correlated to time since the most recent flood (Figure 3.3). Therefore in flood-prone, fast-flowing, unstable mountain streams (and there are many of these in New Zealand), there will rarely be suitable habitat for *D. geminata* to establish and grow. In other words, the opportunities for *D. geminata* cells to encounter a suitable substrate surface, and also have time to make the initial attachment to that surface, will be limited. Among the small literature on surface attachment in diatoms, no information has been located on time required for initial attachment. However, it is clear that several steps are involved, all of which are reversible (Wetherbee et al. 1998). If a cell does succeed in making a firm attachment to a

substrate, the chances of progression to large colonies will also be limited because frequent floods will scour off growth at regular intervals. Therefore, even though these streams may be in all other respects (e.g., temperature, light, water chemistry) favourable for *D. geminata* growth, the fact that cells can rarely establish and grow into colonies renders the habitat largely unsuitable. This means that percentage cover and colony thickness will remain low. On the other hand, lake-fed rivers have stable flows and stable substrates. In our analysis of time since a flood, it was clear that some lake-fed rivers almost never experience “bed-moving floods”. In this situation there is a much higher chance that once a *D. geminata* cell has made a surface attachment, it will continue to divide and grow, and generate propagules which themselves can spread downstream. Our observations (see above) that percentage cover and mat thickness tend to decline in a downstream direction in some rivers (Figure 3.1) further supports our proposal that cover and thickness must be related to habitat suitability.

Overall, direct predictions of cover and thickness in *D. geminata* are more closely suited to the particular ecology of this species and its high levels of downstream transport into unsuitable habitat. A possible exception is that there may be some overriding environmental factor that absolutely precludes presence of *D. geminata* in any abundance (even microscopically) from some areas. One way to determine this would be to use global datasets of environmental data in combination with overseas records of *D. geminata* presence to generate a *D. geminata* presence model, then to use that model to predict geographical coverage in New Zealand. The practical problem with this is that global and New Zealand environmental variables are different, therefore it would be difficult to transfer any predictions to the REC network. In fact, we suspect that temperature is probably the only significant broad-scale limiting environmental factor. This has been dealt with separately (Sections 3.7, 3.8.4). However, additional verification (water temperature data) is needed before we can confidently transfer this environmental limitation to our predictions (see Section 3.9.5, below).

### 3.9.2. Limitations of the models

Modelling suitable environmental conditions for *D. geminata* and predicting distributions in New Zealand presented challenges arising from the fact that this organism is still spreading in New Zealand.

First, datasets for species distribution modelling should ideally be a random sample of the species population. Our sampling was not random, and *could* not be random, because we would then have sampled locations where *D. geminata* was not present only because propagules had not yet arrived there. Our strategy therefore was to include in our survey all rivers in which *D. geminata* was known to be established,

and to visit sites along the entire continuum of those rivers. Although not ideal, this strategy probably complies with the general principle that: “in practice, modellers would be reasonably happy if they could stratify sampling in order to obtain data that were representative of the species’ frequencies of occurrence in a given region and the kinds of environments in which species can live” (Araújo and Guisan 2006).

Second, 70% of river reaches had one or more variables beyond the range of the modelled data. Unavoidably, unreliability of predictions was higher overall in the North Island – again reflecting that *D. geminata* is still spreading in New Zealand. We could not include the full range of North Island environmental conditions in the model because all sites would have been negative for *D. geminata* and this would have biased the model. Indeed our model dataset was already biased towards environmental conditions in the southern part of the South Island. Quantifying unreliability is probably the only way to deal with this at present. In the future, re-running the models including data from additional sites should reduce the proportion of reaches outside the model range. The monitoring protocol used in this survey may facilitate this (Section 3.2.2, and see also <http://www.biosecurity.govt.nz/pest-and-disease-response/pests-and-diseases-watchlist/didymosphenia-geminata/protocols>). Refer to Section 5.3 for further discussion about model limitations in relation to use of the model predictions.

Third, the *D. geminata* dataset was biased towards large rivers because the alga is currently only present in those environments, as a result of its spread by human vectors among larger rivers that tend to be used for recreation. The predictions have therefore been made only for rivers with stream order equal to or greater than the lowest order represented in the dataset. As stated in Section 3.8.2 the predictions can be run for all stream orders, but unreliabilities must be considered.

### 3.9.3. Environmental factors influencing *D. geminata* growth and blooms

The variables included in the final models confirmed observations that *D. geminata* forms the most extensive cover and thickest mats in rivers with stable flows, particularly those fed by lakes. Determining the variable *days\_sig\_flood* highlighted the fact that a significant flood recorded far downstream in a river was often not recorded near the lake outlet: this was seen on some of the rivers which had multiple flow recorders (Table 3.1). This variable is strongly and positively linked to *D. geminata* % cover and thickness at individual sites (Larned et al., in prep.; Figure 3.3). Substrate appeared to be the next most important factor: larger-sized, hard substrates are most favourable. The interaction between *usLake* and *usHard* highlights the very favourable conditions created by the combination of high lake influence (i.e., stable flows) and high substrate stability. Lake-fed rivers are characterised by both stable flows *and* stable substrates, as are dam-regulated rivers (Young et al. 2004).

Temperature, particularly the summer–winter differential, was also associated with *D. geminata* growth in our dataset, with a strong preference for locations with wide seasonal fluctuations. Cooler summer air temperatures were associated with higher % cover and thickness, but the effect was small. Finally, water chemistry may have some influence, though the ranges of both nitrogen and pastoral influence in our dataset were probably too small for definite trends to be identified.

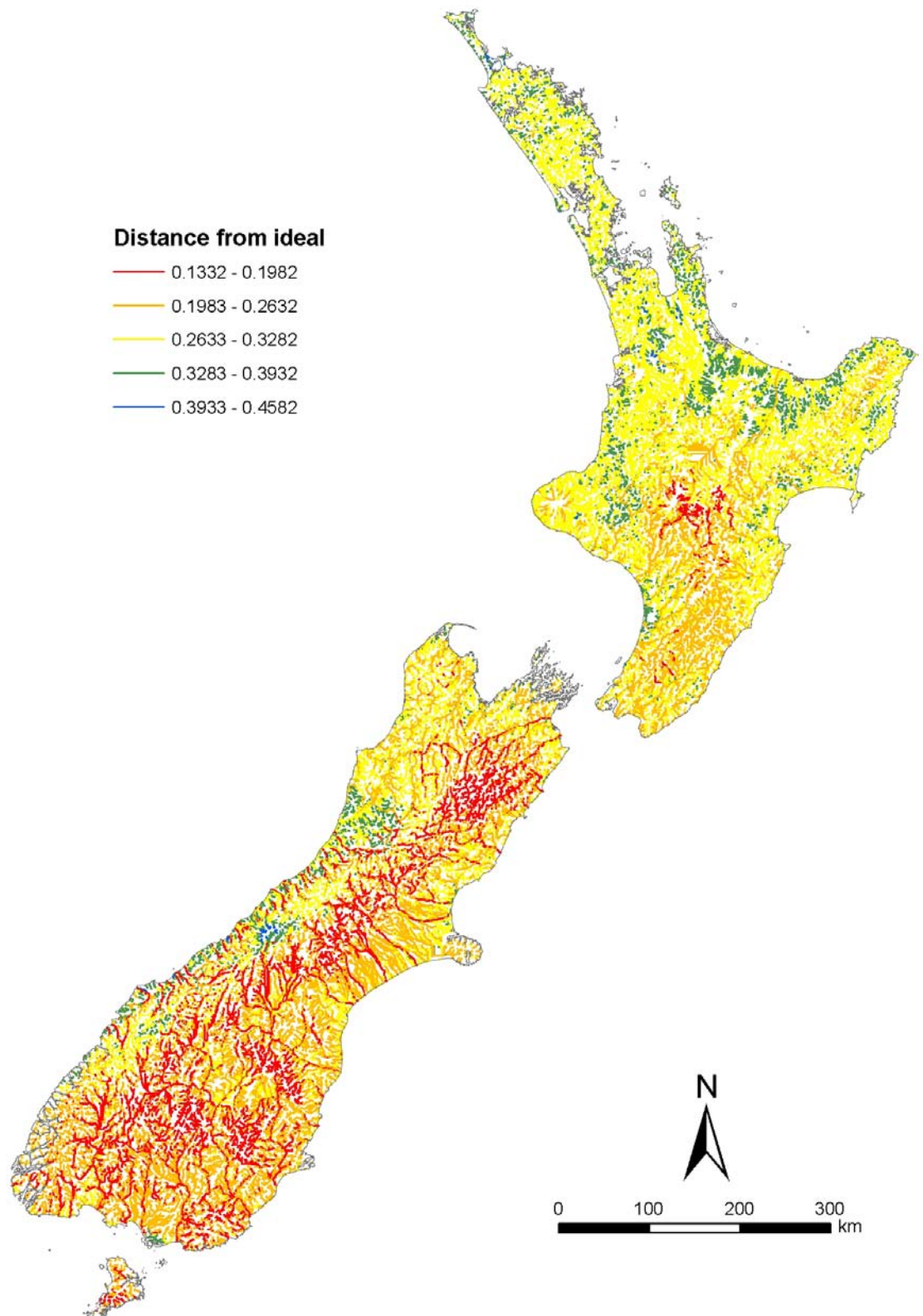
It was interesting that, despite a reasonably strong correlation, % cover and thickness predictions used different variables, suggesting that different environmental factors control these two aspects of growth. For example, a higher contribution by *segAvT\_Wcstan* (annual temperature range), and the inclusion of *Warm\_us\_seg\_stan* (segment temperature subtracted from upstream temperature, with negative values indicating cooler (mountainous) headwaters) in the model for thickness indicates that cooler waters associated with a more continental climate and proximity to high elevations are associated with development of thicker mats. Since mat thickness, especially at high % cover, is likely to be linked to the most severe ecological effects, it is important to consider both models when considering potential impacts.

#### 3.9.4. Comparison with the LEM

The variables included in the DPM described similar environmental factors to those used in the LEM. Thus, the original map used variables for flow stability (*usLake*) and flow variability (*usDaysRainGT25*), substrate size (*us Hard*), and temperature (*usAveTWarm*) (Kilroy et al. 2005). Other variables used in the first map included *usPeat* and *usGlacial*, neither of which featured in the present models. A major difference was that in the original map, a preference for cool temperatures was assumed to be the most important factor determining *D. geminata* distribution and was therefore assigned triple weighting.

High unreliability of the model predictions for much of the North Island and some of the South Island (Figure 3.14) makes a detailed comparison difficult. However, in spite of the fact that temperature (*segAvTWarm*) contributed only a small proportion to both the % cover and thickness models, the general patterns of predictions were consistent with patterns of relative suitability in the LEM in that both methods suggested much higher susceptibility in South Island than in the North Island (compare Figure 3.17 with Figures 3.8 to 3.10). Given the heavy temperature weighting in the LEM, and strong inter-island temperature differences (see, for example, Figure 3.20), a North Island – South Island difference was not surprising. In the current models, direct temperature (*segAvTWarm*) was a minor component in both models. The variable describing seasonality of temperature (*segAvT\_Wcstan*) was more important in both cases, and this may have differentiated the two islands to some extent.





**Figure 3.17:** The LEM for *D. geminata* (Kilroy et al. 2005), showing rivers with stream order  $\geq 3$ , for a direct comparison with the maps predicting % cover and thickness (Figures 3.8 to 3.13). Reaches coloured red are closest to estimated optimal conditions for *D. geminata* (i.e., assumed most susceptible) and blue reaches are farthest away (assumed least susceptible).



In the South Island, a pattern of higher susceptibility in the central mountainous areas and lower susceptibility on the coasts (especially the West Coast) shows up in both the old and new maps. One difference is that the LEM placed many mountain-fed rivers in the most susceptible (red) category all the way to the coast (east and west), whereas the DPM often predict lower cover and thickness in these major rivers. This may result partly from inclusion of a substrate size variable in the new predictions, which was not available earlier.

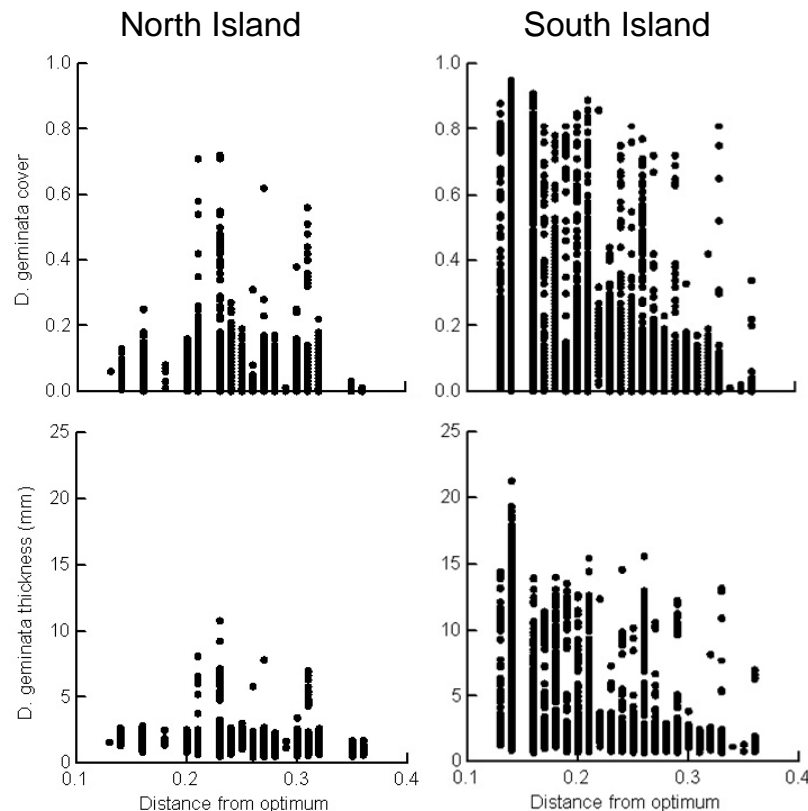
One way to test whether the distances in the LEM and DPM are identifying susceptibility in the same way is to plot the latter against the former. Good correspondence between the methods would produce strong negative correlations: river reaches predicted to have high cover and thickness should have low distances from the estimated ideal. As Figure 3.18 shows, correspondence of the two models is poor in the North Island, but better in the South Island, though there is considerable variability in the data. Note that only reaches with unreliability = 0 were plotted.

It should be appreciated that the LEM and DPM show very different information. The LEM shows relative distance from the estimated optimal environmental conditions for *D. geminata*. The colour bands are arbitrary and do not imply any particular cover. The DPM allow us to say that a particular river is expected to develop *D. geminata* mats of a defined thickness and coverage. Poor correspondence between the two methods is likely to partly reflect that temperature was triple-weighted in the LEM, and was based on a single ideal value. The species is more likely to grow equally well over a range of temperatures, but departure from the single ideal would have resulted in a range of distances. Note that predicted cover and thickness range from near zero at every distance (Figure 3.18).

For more discussion of the LEM vs. the new predictions refer to Section 5.1.

### 3.9.5. A temperature limit for *D. geminata*?

Our examination of overseas presence records of *D. geminata* in relation to temperature suggests that temperature may limit *D. geminata* distribution. All the presence records we mapped suggest an upper limit of approximately 5.5 °C for the mean air temperature in the coldest month. One possible explanation for an upper temperature limit in winter is that sexual reproduction in *D. geminata* may be triggered by a low temperature. This is hypothetical but feasible because in the absence of sexual reproduction, cells will eventually die (Chepurnov et al. 1990). In one of the very few studies on diatom life-cycles in natural population, Potapova and Snoeijs (1997) observed the onset of population-wide sexual reproduction in *Diatoma moniliformis* at low temperatures, but lack of such a response in a neighbouring population in warmer waters.

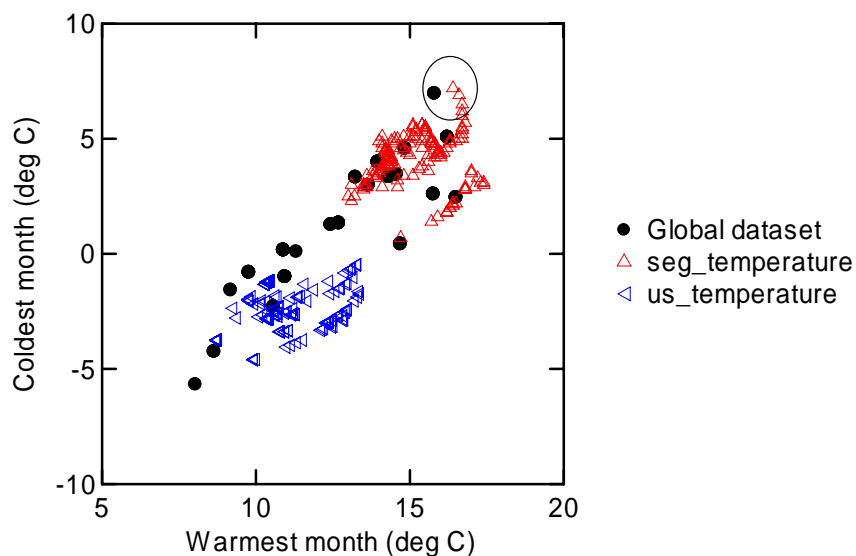


**Figure 3.18:** Predictions of mean *D. geminata* cover (as a proportion) and thickness from the current model plotted against the environmental distance from the estimated optimum conditions for *D. geminata* (in the LEM). Predictions are for the mean number of days since a flood greater than three times the median flow, with North Island and South Island separated. Only reaches within the range of the variables used in the model dataset are included (i.e., unreliability = 0).

An alternative is that the apparent winter air temperature maximum for *D. geminata* is simply a correlation with rivers whose waters are cooler than expected from summer air temperature measurements. For example, the records from warmer areas of North America and Asia were all from inland continental locations, where rivers are less likely to be in equilibrium with air temperature in the summertime than their more coastal counterparts. These areas not only have broad seasonal temperature ranges, but diurnal temperature ranges are also broad. As a result, we speculate that in summer the cooling during the night is sufficient to prevent as much heating of the water during the day as occurs further downstream. The climates of both New Zealand and Great Britain are more equitable with cooler summers than much of North America, suggesting sufficiently cool river-waters in summer in many places. Interestingly, the upper part of the temperature range in Great Britain has no *D. geminata* records (Figure 3.16(b)).

This scenario also requires that *D. geminata* has a physiological requirement for cooler waters (see below).

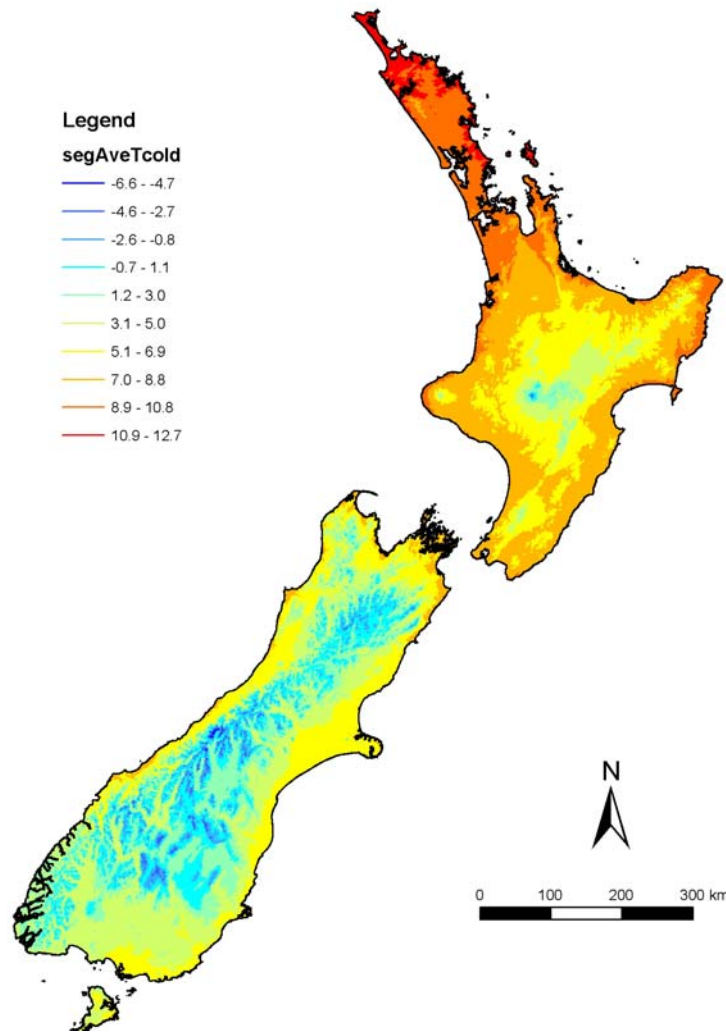
To which areas do winter maximum temperatures of  $\sim 5.5^{\circ}\text{C}$  apply in New Zealand? A difficulty is to transfer mean temperatures from a relatively coarse-scaled global dataset to local data derived on a much finer scale. Figure 3.19 shows that the segment temperature variables associated with the REC correspond well with the higher range temperature in the global dataset, whereas the upstream temperature variables correspond with the lower range global temperatures (See Section 3.1.4 for an explanation of the segment and upstream variables.) Thus reach-scale NZ temperature variables are underpredicted by the global data at low temperatures and the catchment-scale NZ variables are over-predicted at higher temperatures. Within the range of interest (around  $5^{\circ}\text{C}$ ), segment temperatures correspond quite well with the global estimates. The highest global temperature of  $7^{\circ}\text{C}$  almost exactly corresponds with the segment temperatures for the same sites (circled on Figure 3.19).



**Figure 3.19:** Global air temperature data for the New Zealand survey sites (black spots) compared with segment (reach-scale) and upstream (catchment-scale) temperature estimates for the same areas in New Zealand. The circled points show data for four sites on the lower Buller River where *D. geminata* was not present.

Therefore a mean segment temperature in the coldest month (the REC variable *segAveTcold*) of between  $5.5$  and  $7^{\circ}\text{C}$  (allowing a margin of error over the  $5.5^{\circ}\text{C}$  suggested by the global data) roughly corresponds to the observed  $5.5^{\circ}\text{C}$  winter temperature limit suggested by the overseas data. *If* (and this is a very large “if”) this represents a limit for *D. geminata* survival in New Zealand, we would not expect to find *D. geminata* in any areas coloured orange to red in Figure 3.20.

This assessment of a possible limit for *D. geminata* distribution is speculative. First, further overseas presence records of *D. geminata* may show that the limit does not hold. We know of two possible records, both of which will be difficult to verify.



**Figure 3.20:** Mean temperature in the coldest month (segAveTCold) plotted for all river reaches in the REC.

1. Carter and Denny (1982) reported *D. geminata* “in every sample” of material (5 samples) collected from the River Jong (Gaia), Sierra Leone, West Africa. This is a tropical river arising from a lake at 549 m a.s.l. Water temperature at the collection site was reported as 31 °C (1976). No details of cell size or shape were provided and the species was not illustrated. However, the authors are experienced diatomists and this is unlikely to be a mis-identification. From the description of sample preparation, cross-contamination seems unlikely (the authors have also published on diatoms from Scotland). Clearly, some further investigation of this record is needed.
2. Cleve (1894-96) listed Spain and Aragonia among localities for *D. geminata*. No site details are provided and no specimens are referred to.

Second, water temperature data are needed to confirm whether *D. geminata* distributions are limited by temperature. Such data do not exist either for New Zealand, or globally. Most easily accessible river water temperature data in New Zealand comprises spot temperatures readings taken during the daytime. These are only roughly indicative of mean conditions, but are still useful. For example, many rivers in the Hawkes Bay region (east coast of North Island) regularly reach 28 °C or higher during summer and rarely fall below 10 °C, whereas Southland Rivers barely attain 20 °C in summer and fall to <5 °C in winter (Mosley 1982; NIWA National Rivers Water Quality Network data; Environment Southland water temperature records). Laboratory experiments have indicated that *D. geminata* can survive well at 20 °C but cannot survive at 28 °C for more than ~36 h (Kilroy et al. 2006a), suggesting that some North Island rivers may well be too warm for this species.

### 3.9.6. Concluding overview

In conclusion, our models confirmed observations that the important environmental variables favouring *D. geminata* colonisation and blooms are: high lake influence; stable, hard substrates; low flow variability; and longer time since a flood. The three hydrological variables – lake, flow variability and time since floods – are interlinked. For example, a high lake influence is normally associated with low flow variability and longer intervals between floods. Segment temperature contributed only a small proportion to the model but large seasonal temperature differences at a segment scale were associated with high values of both % cover and thickness. The survey data show downstream declines of *D. geminata* % cover and /or thickness in several rivers (Figure 3.1). These declines could be driven by a combination of the important variables. For example, lake-fed rivers become less lake influenced with consequent increasing flow variability as tributaries join the main stem; substrate size tends to decline at the same time, and seasonal (and diurnal) temperature ranges become more equitable nearer to the coast.

Currently the predictions in many areas of New Zealand are unreliable to differing degrees because these areas lie outside the environmental range of the sites used to generate the models. Future observations in a wider range of rivers will reduce the level of unreliability. The monitoring protocol outlined in Section 3.2.2, and detailed in a separate report (<http://www.biosecurity.govt.nz/pest-and-disease-response/pests-and-diseases-watchlist/didymosphenia-geminata/protocols>), should facilitate this.

Whether temperature is a limiting variable in these downstream gradients – and possibly over whole areas in the North Island – remains to be proven. The absence of *D. geminata* in the lower Buller River could be a result of temperature limitation, since this area had the warmest winter temperatures in our dataset, which were also above the observed limit in the international data. However, the other variables

associated with progression downstream could equally contribute to its absence in the lower Buller. These include a slight decrease in sediment particle size, and increasing flow variability due to the contribution of large tributaries such as the Matakaiti, Maruia and Inangahua Rivers. Further observations in rivers flowing into warm areas (e.g., Motueka, Takaka) may help resolve the question. For example, thin *D. geminata* mats have been observed in the lower reaches of the Takaka River, which has a winter air temperature of >7 °C according to Figure 3.20 (Greg Napp, DOC, pers. comm.). Most importantly, we need information on river water temperature regimes. No satisfactory water temperature dataset currently exists in New Zealand, yet this is likely to be critical information in making future predictions about freshwater invasive species: *Didymosphenia geminata* will not be the last. Development of such a dataset requires deployment of temperature loggers at multiple strategic river sites over several seasons. With sufficient data, modelling to cover the entire country would be relatively straightforward.

Finally, the DPM provide quantitative predictions of average % cover and mat thickness of *D. geminata*, in contrast to the relative suitability of rivers for *D. geminata* provided by the LEM. This approach was taken partly because of the requirement to predict occurrences of a defined bloom. However, direct predictions of average % cover and thickness under different hydrological conditions (days since a flood) are straightforward to interpret, and arguably more useful in management applications (Section 3.9.1). It should be remembered that the predictions are for *average* % cover and thickness in *wadeable* parts of rivers.

## 4. Colonisation in lakes

Since *D. geminata* is now known to be establishing in lakes, it would also be desirable to include predictions for lakes. However, because colonisation in lakes has been recorded only recently (e.g., Lakes Wakatipu and Te Anau since November 2006), at this stage we can comment qualitatively only. Our approach in this project was to consider three questions:

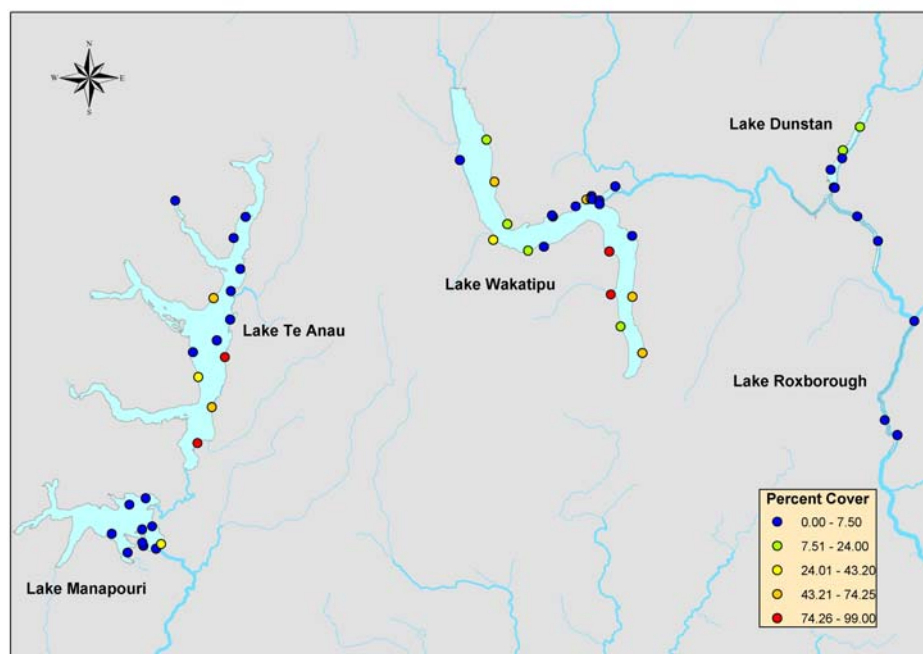
1. In the lakes exposed to *D. geminata* (i.e., with inflows from *D. geminata*-affected rivers), where are colonies visible?
2. Given that there are differences in the extent of *D. geminata* in different lakes, what factors might explain these differences?



3. In light of the extent and occurrence of *D. geminata* in lakes at present, can any features of lakes be identified as particularly favourable or unfavourable for establishment and growth of *D. geminata*?

#### 4.1. Survey methods and results

Lake-shore sites were surveyed for visible *D. geminata* mats using the same method as used in rivers (Section 3.2.2, Section 5). Lakes included were Manapouri, Te Anau, Wakatipu, Roxburgh and Dunstan. Lakes Manapouri, Wakatipu, Roxburgh and Dunstan have had known inputs of *D. geminata* since September/October 2005. Known exposure in Lake Te Anau has been shorter (since May 2006). Time constraints meant that only a limited number of sites could be visited in each lake. Site details are given in Appendix 4, and locations showing five categories of % cover are shown in Figure 4.1. The following commentary addresses the first question above.



**Figure 4.1:** Lake sites surveyed for *D. geminata* in April – May 2007. Refer to Appendix 4 for site details.

##### 4.1.1. Lake Manapouri

Of ten sites surveyed on Lake Manapouri, *D. geminata* colonies were visible at only one (Fraser's Beach, Manapouri village), and here even an experienced person might not have made a positive identification. The identity of the colonies was verified microscopically. At other sites there was thick algal growth, but this was determined

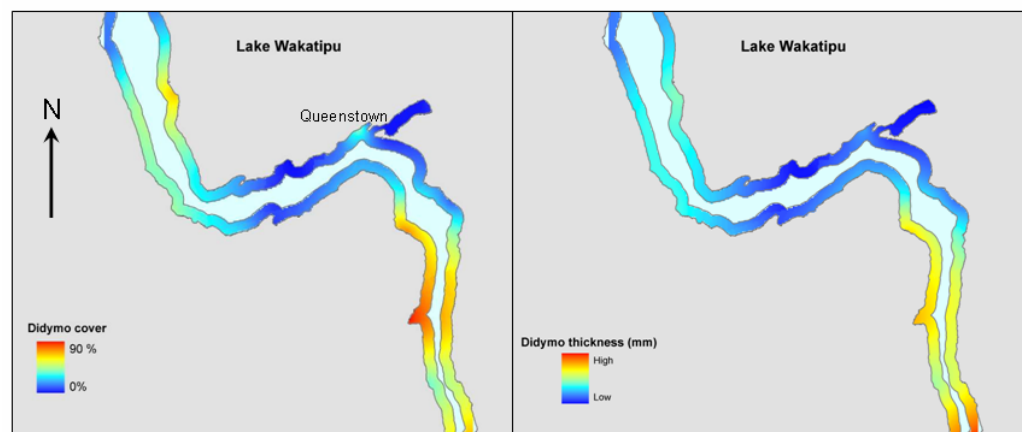
from photographs and two samples (Appendix 4) to be typical lake algae. A complete survey of Lake Manapouri was curtailed because of bad weather.

#### 4.1.2. Lake Te Anau

Of thirteen sites surveyed in Lake Te Anau, five were positive. Four of these sites were in the southern part of the lake, including Te Anau township (Bluegum Point), and the Te Anau caves wharf. At this stage we realised that, in lakes, microscopic examination of samples is critical because *D. geminata* mats and colonies in lakes do not always have the typical “look” of *D. geminata* seen in rivers (e.g., Sutherland et al. 2005).

#### 4.1.3. Lake Wakatipu

Samples were collected at 12 of the 22 sites surveyed, making this our most reliable survey. Sixteen sites were positive, with cover ranging from very slight (<1%) to over 90%. The western shore of the south arm had highest % cover, but the thickest mats were towards the southern end of the lake (Figure 4.2). Despite recent reports (March 2007) of high cover in the Queenstown area, this was not evident at the time of this survey (May 2007).



**Figure 4.2:** Distribution of *D. geminata* (as % cover and mat thickness) along the shoreline of Lake Wakatipu estimated from survey data from 22 sites. The north and south ends of the lake are cut off because there were no data at those locations. The width of the affected marginal area is exaggerated for clarity.

#### 4.1.4. Lakes Roxburgh and Dunstan

Of eleven sites checked in these above-dam sections of the Clutha River, visible *D. geminata* was recorded at three in Lake Dunstan (above the northern dam), in moderate amounts only (<25% cover).

## 4.2. Discussion

### 4.2.1. Causes of colonisation differences between lakes

Of the lakes examined, and from our limited surveys, Lake Wakatipu had the most extensive cover of *D. geminata*, followed by Lake Te Anau. Only limited growth was recorded in Lakes Manapouri and Roxburgh / Dunstan. Note that survey intensities differed among lakes, therefore these are general impressions only. Possibilities leading to these differences (our second question, above) are discussed in terms of six factors likely to influence blooms of the species in lakes.

1. *Wave exposure*: locations suitably oriented to the prevailing wind, with large fetch may be most susceptible because of greater water exchange. In rivers *D. geminata* has been shown to have weak biomass peaks at water velocities of ~0.5 m/s, but grows across a wide range of velocities (Kilroy et al. 2005).

*D. geminata* was most widespread in Lake Wakatipu, with a higher proportion of visible growth than in any of the other lakes. Lake Wakatipu is very exposed for most of its shoreline. Most growth was recorded on the western shore of the south arm and there was moderate growth on the eastern shore, and on both shores of the north arm (Figure 4.2). Strong winds from the north would be expected to create a dynamic wave climate in these areas, and possibly more so than in the central part of the lake, but this would need to be verified, e.g., by using a NIWA model for predicting the wave climate in large lakes. The smaller size of Lake Manapouri could restrict wind-wave potential in some areas. However, this is still a large lake and it seems unlikely that lack of wind-waves alone is preventing *D. geminata* from thriving there. We expect only limited wind-wave action in the river-lakes Roxburgh and Dunstan.

2. *Substrate*: as in rivers, stable, hard substrates are expected to be most favourable for colonisation.

There was no clear relationship between *D. geminata* % cover and thickness, and substrate size, either across all lakes, or within individual lakes. However, most of the negative Roxburgh / Dunstan sites had small substrates (index <4, indicating predominantly “small cobbles” (see Section 5) or smaller). Dense aquatic weed beds in parts of Lake Dunstan may prevent establishment of *D. geminata*, though all sites surveyed were largely free of weeds.

3. *Lake edge profile*: places where there is a steep drop into the water are expected to be less favourable because of declines in both water movement and light availability in deeper waters.

It is possible that steep shore profiles also created unfavourable conditions in parts of Lakes Roxburgh and Dunstan. For example, site 1 (see Appendix 4) was largely bedrock dropping steeply into the water. Several negative sites in Lake Manapouri had similarly steep profiles.

4. *Lake-level fluctuations*: lakes prone to large fluctuations in water level may be less susceptible to continuous blooms because of periodic desiccation.

All the lakes except Lake Wakatipu are regulated as part of hydroelectric power schemes, and all undergo natural fluctuations. Therefore a lake-level change may explain why blooms reported from Queenstown in early March were not seen during our survey in mid-May – in the intervening 2 months, the lake level had fallen 0.5 m, then risen again. However, lake level fluctuations are unlikely to explain the differences between lakes Wakaitipu, Manapouri and Te Anau, since levels in the three lakes follow similar patterns.

5. *Water quality / chemistry*: humic (tannin-stained) waters may be less favourable than clear waters because of greater light attenuation (especially UVB), and possibly unfavourably low pH. Other water chemistry features (e.g., nutrient loadings) may also be significant.

Both Lakes Te Anau and Manapouri have slightly tannin-stained waters (Rae et al. 2001), circumneutral pH, and their clarities are very similar (Schwarz et al. 2000; Environment Southland water monitoring data), therefore these aspects of water quality cannot explain different colonisation levels in the two lakes. Lake Wakatipu has higher clarity (Schwarz et al. 2000), and thus may be more favourable for *D. geminata*. pH is also slightly higher. Lake Dunstan is stated to have water quality (including clarity) similar to that of Lake Wanaka, which is similar to Wakatipu (Otago Regional Council 2002). All the lakes in which *D. geminata* is known at present are oligotrophic, therefore no comment can be made about the effect of different nutrient concentrations.

6. *Regular boat traffic*.

Anecdotal evidence suggests that in both Lakes Wakatipu and Te Anau, boat traffic could be playing an important role in transporting *D. geminata* within the lakes, and therefore the pattern of colonisation. For example, in Lake Te Anau, there is a popular tourist trip between Bluegum Point (Te Anau township) and Te Anau caves wharf, and obvious *D. geminata* growth at both locations (B. Jarvie, D. Sutherland, pers. comm.)

#### 4.2.2. Identification of factors favouring *D. geminata*

Because we could not definitively identify any factors that explained differences in colonisation among the lakes, it is difficult to speculate about the third question (... can any features of lakes be identified as particularly favourable or unfavourable for establishment and growth of *D. geminata*?). No single factor explains why Lake Wakatipu appears to be the most heavily affected lake, followed by Te Anau, and then Lakes Manapouri and Roxburgh / Dunstan. Note that a single site on Lake Benmore (downstream of the Ahuriri River, Waitaki system) was also inspected during the river survey. The site had large cobble substrate and was exposed to waves, but no there was no sign of *D. geminata* colonies.

It should be remembered that all these lakes have been exposed to *D. geminata* for a relatively short time and are probably still in the process of being colonised. Unlike river periphyton, lake periphyton tends to be very stable: once colonised, an area can stay unchanged for long periods because disruptive events similar to flood events in rivers occur much less frequently. Therefore the establishment of *D. geminata* at a particular site depends not only on the suitability of conditions there, but also on whether there is opportunity to colonise amongst the resident communities of well-adapted algal or aquatic plant species.

We conclude that further investigations are needed to improve understanding of the potential for *D. geminata* colonisation in lakes. These should comprise sample-based surveys, including dive surveys. More thorough surveys in lakes currently exposed to *D. geminata* will establish the real extent and growth of the species in these lakes and help identify whole-lake-scale factors favouring growth. Monitoring over time in a single lake may help to determine the triggers for localised or widespread blooms within lakes such as that reported near Queenstown in March 2007.

## 5. Use of and applications for the DPM

This section covers what the DPM mean in terms of impacts or lack of impacts on stream ecosystems, impacts on recreational or economic uses of waterways, and guidance on use of the model predictions. In other words, what are the implications of a prediction of a particular level of *D. geminata* growth, what are the limitations of the predictions, and how can the predictions be used to assist in stream management?

We first ask:

1. How should the DPM be interpreted compared to the relative levels of risk provided in the LEM?
2. What levels of *D. geminata* % cover and thickness are relevant for ecological, recreational (including aesthetic) and economic stream values?

An example is then provided of how the DPM can be used to assist in stream management related to *D. geminata*, and other possible scenarios where the maps may be useful are discussed.

Note that the following discussion mainly relates to model predictions for which unreliability = 0, which applies to about 30% of all river reaches (stream order  $\geq 3$ ) in New Zealand (see Section 3.6.1). It is reiterated that the unreliability measure *must* be taken into account when using the predictions.

### 5.1. DPM vs. LEM

At a first glance the DPM for *D. geminata* average cover and thickness suggest a lower overall susceptibility in New Zealand compared to that suggested in the LEM (e.g., compare Figures 3.9 and 3.17). In the LEM, large areas of the South Island coloured red indicated rivers falling into the environmental range closest to that estimated to be optimal for *D. geminata*. Orange, yellow, green and blue reaches were progressively farther away from the optimum. However, the scale was relative and did not imply any particular level of cover by *D. geminata*. Since we did not know of any limits to *D. geminata* colonisation and growth, reaches farthest from the optimum were not assumed to be unsuitable for colonisation, though they may have been interpreted in that way.

In contrast, the DPM show high proportions of the river network over all of New Zealand in the third (yellow) and fourth (green) categories of % cover and thickness, and only low proportions in the two highest (red, orange) categories. Nevertheless, the models predict that over 40% of all reaches with stream order  $\geq 3$  and unreliability = 0 will attain at least 10% cover by *D. geminata*, at an average number of days since a significant flood (over 50% in the South Island, Table 5.1). Cover of 10% corresponds to patchy cover which is normally easily seen. Only about 7% of river reaches are predicted to attain average cover of less than 1% of *D. geminata*. In South Island only, this falls to ~4% (Table 5.1).

This implies that *almost all* river reaches with stream order  $\geq 3$  in New Zealand for which we can make reliable predictions, are to some extent susceptible to colonisation by *D. geminata*: that is, all reaches except those coloured blue on the maps (Figures 3.8 to 3.10) are predicted to support cover of  $>1\%$ , on average, provided unreliability



is low. Just 1% cover sounds very small, but is in fact easily visible (Figure A1.2), particularly in streams where other algal growth is normally low.

**Table 5.1:** Proportions (%) of reaches in North and South Islands with stream order  $\geq 3$  predicted to support *D. geminata* above and below various thresholds of average cover in reaches with unreliability = 0 (30% of all reaches), at the average length of time since a flood. Average predicted mat thickness (mm) is also shown (means of all thicknesses above/below threshold).

Cover	North Island		South Island		Description (at threshold level) / Value
	% Reaches	Thickness (mm)	% Reaches	Thickness (mm)	
Less than 0.1%	8.7	0.9	0.7	1.0	Visually absent
Less than 1%	26.0	1.9	4.1	2.2	Just noticeable
Less than 10%	92.0	1.4	47.4	1.8	Becoming obvious
Greater than 10%	8.0	2.4	52.6	3.6	Patchy cover, easily seen
Greater than 25%	0.8	5.6	11.1	7.8	Moderate to thick cover
Greater than 60%	0.04	9.0	5.5	10.4	Extreme cover
Greater than 40%	0.5	6.6	7.5	9.6	Corresponds to trout habitat / angling guideline
Greater than 30%	0.7	6.0	9.1	8.7	Corresponds to aesthetics / recreation guideline

## 5.2. *D. geminata* cover / thickness and stream values

At least two thresholds of *D. geminata* growth are important for assessing potential impacts on stream ecology. First, at what level can we say that *D. geminata* is effectively absent, and second, at what level of cover / thickness do potential impacts become significant? The latter threshold could be split into two or more, each applicable to a different value.

As mentioned above, 1% cover of *D. geminata* can be easily seen. Thus a realistic level for visually undetectable cover must be set at around 0.1% or less. This low level of cover is predicted for less than 1% of South Island reaches, and only about 9% of North Island reaches (where stream order  $\geq 3$  and unreliability = 0). In other words, over 90% of these reaches in North Island and 99% of South Island reaches are predicted to be capable of supporting visible cover of *D. geminata*. (But note that this does not take into account any temperature limitation for *D. geminata* – see Section 3.9.5).

At the higher growth end of the scale, the first important threshold should align with current guidelines for assessing undesirable growth of periphyton in rivers (Biggs 2000). Two relevant guidelines are:

1. For aesthetics and recreation, the maximum algal cover on the visible stream bed should not be greater than 60% of diatoms or cyanobacteria (>0.3 cm thick) or 30% of green filamentous algae (>2 cm long).
2. For trout habitat and angling, ash-free dry mass (AFDM)<sup>5</sup> per square metre of stream bed should be no greater than 35 g.

These periphyton guidelines were developed following a review of previous studies, including public perceptions of algal cover in rivers, and relationships between algal biomass and invertebrate communities and water quality (Biggs 2000).

Figures 3.8 to 3.13 indicate that under all days-since-flood scenarios, the proportion of rivers where *D. geminata* is predicted to attain >60% cover is quite small. However, mean predicted mat thickness is three times the 0.3 cm minimum stated for diatoms or cyanobacteria (Table 5.1). In view of the thickness of the mats, the cover threshold for green filamentous algae (30%) may be more appropriate.

A *D. geminata* equivalent to the trout habitat and angling guideline requires conversion of cover and /or thickness to ash-free dry mass. A suitable conversion was determined from a field study on the lower Waiau and Mararoa Rivers, where ADFM was found to be strongly correlated with a *D. geminata* biovolume index (% cover x mat thickness) (Kilroy et al. 2006c). The criterion of 35 g/m<sup>2</sup> AFDM is equivalent to an index of approximately 220, which translates to a predicted cover of ~40% which is predicted to have mean thickness of about 5.5 mm at that level of cover.

The proportions of reaches predicted to exceed these two criteria following colonization by *D. geminata* are shown in Table 5.1, for reaches with unreliability = 0.

A more difficult task is to identify a threshold of *D. geminata* % cover and/or thickness that corresponds with measurable impacts on the stream ecosystem. The first question is whether *D. geminata* has measurable impacts. Larned et al. (2007) analysed invertebrate and associated periphyton samples from 12 sites in the Mararoa, Oreti and lower Waiau rivers from April 2006 to March 2007. They found generally weak (but often significant) relationships between periphyton biomass (as AFDM) and

---

<sup>5</sup> AFDM is a measure of the total amount of organic material in an algae sample.

a range of invertebrate indices (density, taxonomic richness, %EPT<sup>6</sup> abundance, % EPT biomass, and invertebrate biomass), but no clear thresholds for effects. There were significant negative relationships between the EPT variables and periphyton biomass. A threshold of 50% EPT biomass was taken (from previous studies) to differentiate between “clean water” invertebrate communities, and communities dominated by more tolerant taxa such as molluscs, crustaceans and midge larvae. Unfortunately, the periphyton biomass corresponding to this threshold varied widely among rivers – from 12.5 to 167 g/m<sup>2</sup> AFDM (Larned et al. 2007). There was also considerably inter-site variability. The high variability arises because different rivers support different kinds of invertebrate communities, presumably as a result of variability in environmental factors such as water quality, temperature, flow regime, local hydraulics, substrate, etc. Therefore, the commonly used EPT indices do not appear to be useful universal indicators of effects on invertebrates in *D. geminata*-affected rivers. Consequently, at this stage it is not possible to generalise about thresholds of *D. geminata* biomass that might affect invertebrate communities, except to observe that the existing guideline of 35 g/m<sup>2</sup> AFDM (Biggs 2000) might reasonably be taken as an interim threshold in the absence of further information.

At this stage there is also no information about levels of *D. geminata* coverage or thickness that might impact trout, native fish, or bird habitats. We can speculate that high cover (greater than 50%) seems likely to have effects on populations of some small native freshwater fish, for example, by smothering the marginal cobbles and gravels where they typically forage, spawn and take cover. However, targeted studies are needed if the predictions are to be used for assessing the impacts on native fish. The results of work commissioned by MAF Biosecurity New Zealand investigating impacts on trout are to be released shortly.

### 5.3. Limitations of the DPM

We anticipate that organisations involved in river management will be able to use the maps to answer questions about the potential impact of *D. geminata* on particular river systems and their values. However, potential users should first understand the limitations of this type of model.

1. As with all such models there is uncertainty in the predictions. In this case, our analysis showed that the models explain about 50% of the variability in *D. geminata* cover and thickness at any particular site.

---

<sup>6</sup> EPT refers to Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddis flies), which are orders of aquatic invertebrates generally accepted to be indicative of unpolluted waterways. For more details on the indices, see Larned et al. (2007).

2. The predictions were made using variables that are themselves modelled (i.e., extrapolated from data at a smaller number of points). This adds further uncertainty to the predictions, especially when considering individual sites or reaches.
3. Finally, as explained at length above, because of the way these models were derived, approximately 70% of all river reaches in New Zealand (stream order  $\geq 3$ ) fall outside the range of one or more of the variables in reaches used to develop the model. Predictions in these reaches are therefore unreliable, and our unreliability index *must* be taken into account when any predictions are made.

In view of the first two limitations, we recommend that the DPM are not appropriate for assessing risk on a very small scale (e.g., single segments or small groups of less than about 25 segments). We suggest that they are best applied to regions, catchments or long stretches of river, and that the five thickness categories shown on the maps provide an appropriate level of detail in most cases. It should also be remembered that models such as these cannot account for local environmental conditions. Therefore local knowledge about, for example, river shading, substrate mobility, and water turbidity, should be considered along with the predictions.

It should also be remembered that the predictions of thickness are conservative for the reasons explained in Section 3.8.2. Thus, the top two colour categories in Figure 3.11 to 3.13 (predicted thickness of  $> 6$  mm) represent very visible and prolific growth in at least some places in a reach.

#### 5.4. An example

*D. geminata* was discovered in the Kakanui River, North Otago, during a MAF Biosecurity New Zealand delimiting survey in May 2007. The Kauru River, a tributary of the Kakanui, is home of the endangered Kauru (lowland) long-jawed galaxias (*Galaxias cobitinis*). (McDowall and Waters 2002). The discovery of *D. geminata* therefore caused some concern since thick benthic algal mats could potentially smother the cobble/gravel substrates which these small, non-migratory fish inhabit.

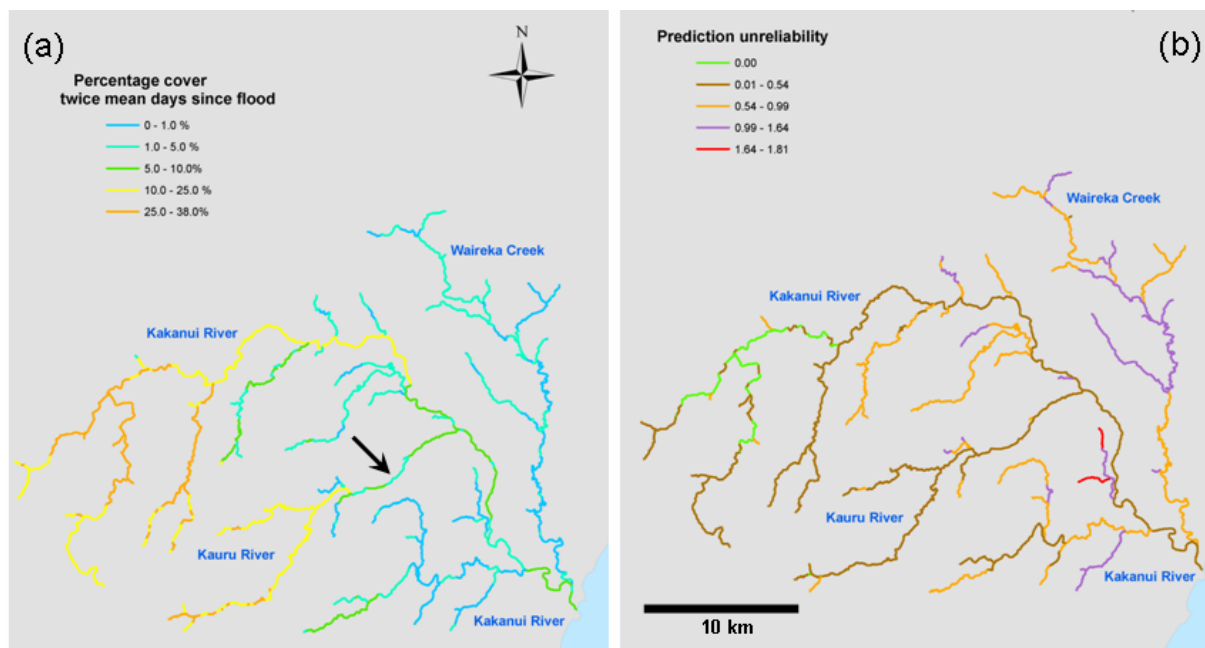
Management responses to this discovery are still being considered by MAF Biosecurity New Zealand, Department of Conservation, and other agencies, but the DPM of cover and thickness could potentially contribute to decision-making. Questions to consider would be:

1. To what extent is *D. geminata* predicted to grow in the lower Kakanui River? This information could assist in assessing the feasibility of any

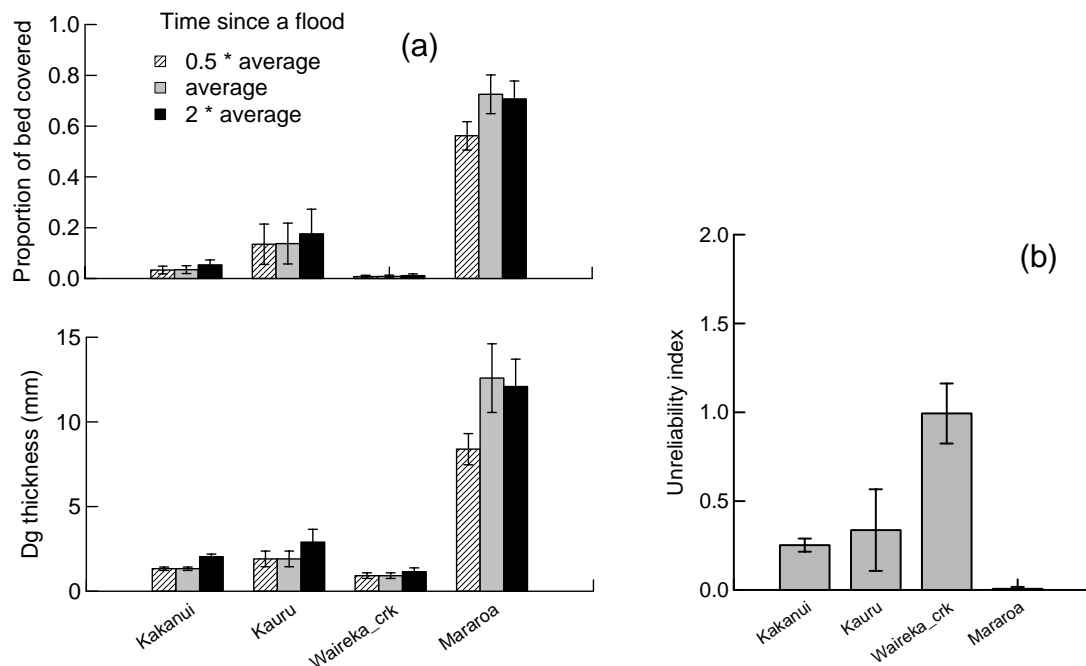
potential treatment of *D. geminata* (e.g., using Gemex® - Clearwater et al. 2007). The information would also help assess the likelihood of the species spreading upstream (thicker growth downstream will increase the chances of propagules spreading upstream).

2. To what extent is *D. geminata* predicted to grow in the Kauru River? In other words, how much of a threat is it to the galaxias?
3. How likely is it that *D. geminata* will be re-introduced to the Kakanui catchment? One possible avenue for re-introduction is via the Waireka Stream which takes overflow from irrigation waters originating from the Waitaki catchment. How suitable is this stream for harbouring an established population of *D. geminata*?

To answer all three questions, we first identified REC reach IDs for all stream segments in the lower Kakanui (downstream of the Kauru confluence), the Kauru mainstem, and the Waireka mainstem (Figure 5.1). The predictions for these reaches were then extracted from the table of model predictions and unreliabilities.



**Figure 5.1:** (a) Predictions of *D. geminata* cover in the Kakanui catchment for the worst-case scenario of twice the mean number of days since a flood. (b) Unreliabilities for these reaches: Note that only one reach in the headwaters of the Kauru is in the most reliable category (green). The black arrow on (a) indicates the approximate location of the upstream limit of stronghold populations of the lowland long-jawed galaxias. The downstream limit is at the confluence with the Kakanui. The arrow also marks the location of the flow recorder.



**Figure 5.2:** (a) Predictions of *D. geminata* cover and thickness for three stream reaches in the Kakanui catchment. (b) Unreliabilities for these reaches: 0 = most reliable. The Kakanui predictions apply to the mainstem downstream of the Kauru confluence, and the Mararoa predictions to the mainstem from South Mavora Lake to Kiwi Burn. Error bars are standard deviations.

To place the predictions for the Kakanui / Kauru system in the context of a known *D. geminata* infestation, we also extracted the predictions for the upper part of the Mararoa River, Southland. We assume that the first successful establishment of *D. geminata* was in this section of the Mararoa, and blooms have been present in the river since late 2004. The predictions are shown in Figure 5.2(a) as means for each waterway. Unreliability indices are shown in Figure 5.2(b).

Given the uncertainties of the DPM (Section 5.3), reference to the five % cover or thickness categories on the maps is usually an appropriate level of detail. In this example, we want to look at the overall means for the three areas, and the upper Mararoa reference area, so we go back to the raw predictions.

The Kakanui river downstream of the Kauru confluence is predicted to support quite low cover of didymo – <1% to 7.8 %, on average, depending on the time since a significant flood. The Kauru is predicted to be more susceptible with ~18% cover on average, reaching over 30% in some areas. Highest cover and thickness (not shown) are predicted for the top of the catchment (i.e., not in the area where the lowland long-jawed galaxias is found) (Figure 5.1a). The models predict that no part of Waireka Creek will attain high cover or thickness of *D. geminata*. Predicted cover in the upper Mararoa is 44% to 81% with thickness up to about 15 mm (but see Section 3.8.2).



Figure 5.2 (b) shows that the predictions in the Kakanui catchment are all to some extent unreliable (see Sections 3.6.1 and 3.8.2). For example, in the Kakanui below the Kauru confluence, environmental variables of these segments on average have a total proportion outside the range of the model data of ~25%, which is just under 2% per variable. Even if a single variable accounted for the whole 25%, this is not a large departure from the range, so we conclude that the prediction unreliability is low: i.e. the model predictions are reasonably reliable. Unreliability in the Kauru was more variable (maximum 1.2) but most reaches throughout the whole stream length had unreliability less than 0.3. Values in Waireka Creek, although high, were still well below the maximum unreliability in New Zealand of over 5.

At this stage local knowledge about these streams would also need to be considered. For example, in the area of interest, do the flow regimes, substrate stability, and light availability in these streams seem suitable for *D. geminata*? Do the streams support occasional growth of other stalked diatoms, such as *Gomphoneis*, which we know does well in similar habitats? To illustrate how this might be approached, we consider two of these factors: flow variability and previous information on algae.

### *Flow variability*

Flow variability is an important predictor in the models through the variables *Days\_sig\_flood* and *usDaysRainGT25* (Tables 3.2, 3.3). By checking the *actual* flow variability we are asking the question: is actual variability in the river consistent with the predictions for *D. geminata* cover and thickness? There are several ways of assessing flow variability in a system, all of which require access to hydrological records. For example, we could calculate the mean number of times in a year when the flow exceeds 3 x median (FRE3). This has been shown to be the best predictor of long-term algal biomass in a system (Clausen and Biggs 1998, and see Section 3.3.3). Another option is to calculate the coefficient of variation of flow. For the present example, we calculated both indices for some of the rivers included in the survey (Table 3.1), as well as the Kauru and Kakanui Rivers (Table 5.2).

Because regulated rivers have unusual hydrological characteristics compared to natural rivers, these are distinguished in Table 5.2. Lake-fed rivers are also indicated. Flow variability in the river being considered should be compared with variability in similar rivers (i.e., rain-fed, lake-fed, or regulated). As Table 5.2 shows, flow variability in the Kauru River, as %CV of flow, is relatively high compared to other rain-fed rivers. Variability in the Kakanui River is very high. This is consistent with model predictions of low cover of *D. geminata* in these reaches. FRE3 is similar to that in most other rain-fed rivers, but higher than that in the lake-fed and regulated rivers.

**Table 5.2.** Flow variability statistics (CV% and FRE3) for a selection of rivers in which *D. geminata* is known to be established, for comparison with statistics for the Kakanui and Kauru Rivers (highlighted). %CV = standard deviation of daily flows / mean daily flow; FRE3 = mean annual frequency of floods greater than 3 x median flow. Rivers are listed in order of increasing %CV.

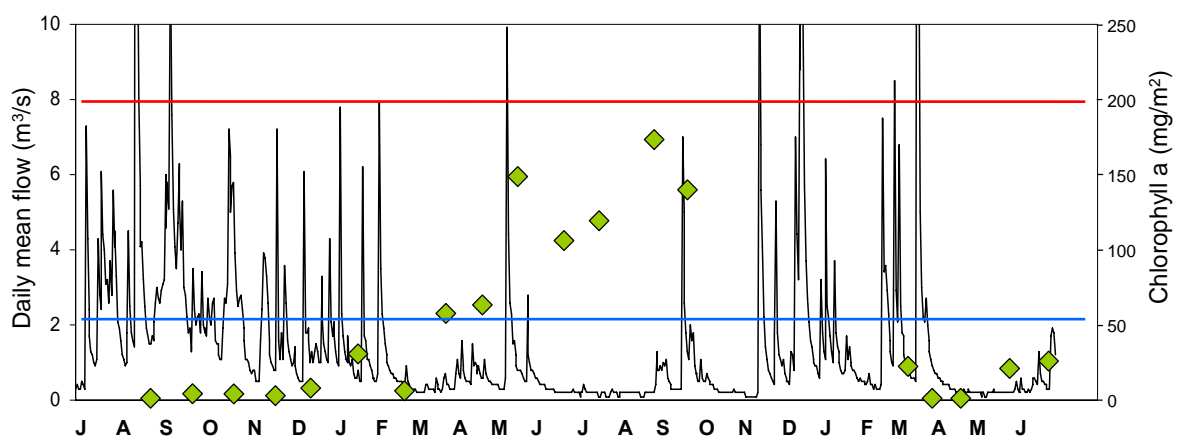
River	Recorder site	Site no.	Flows (m <sup>3</sup> /s)		%CV	FRE3
			Mean	Median		
Rain-fed rivers						
Ahuriri	Sth Diadem	71116	22.6	18.0	78	7
Oreti	Three Kings	78608	8.2	6.3	95	15
Aparima	Dunrobin	78906	7.4	5.4	116	20
Oreti	Lumsden	78636	27.0	18.3	124	19
Aparima	Thornbury	78901	22.6	15.3	132	15
Kauru	Ewings	71706	1.0	0.5	186	18
Kakanui	Clifton	71703	2.3	1.3	247	15
Kakanui	Mill Dam	71713	4.7	1.7	436	16
Lake-fed rivers						
Buller	Longford	93202	70.5	53.2	85	11
Mararoa	Cliffs	79737	31.2	22.9	96	9
Regulated rivers						
Waitaki	Kurow	71104	351	342	30	0.3
Clutha	Balclutha	75207	565	543	42	1
Clutha	Clyde	75213	497	464	47	0.1
Lower Waiau	Sunnyside	79735	93.6	48.8	115	14
Lower Waiau	MLC tailrace	79719	67.1	17.6	167	13

Note that the flow variability statistics apply only to the river near the flow recorder. In rain-fed rivers %CV generally increases in a downstream direction as tributaries join the system and catchment size increases. This is seen in all three pairs of recorders on the same river in Table 5.2. In all cases, the recorder with the lower %CV is the upstream site. In the Kauru, the flow recorder is in the area of interest (Figure 5.1), therefore the statistics are accurate for that area.

### *Other algal growth*

Data may be available from some rivers on previous periphyton growth and may help to evaluate whether algal blooms already occur in the system from time to time. This information could be used to assess whether an endangered fish species (for example) is already exposed to high algal biomass from time to time, and may help to determine impacts from the predicted *D. geminata* cover and thickness. Information on periphyton might be from, for example, previous State-of-the-Environment monitoring, or from research programmes.

In the present example, we know that periphyton sampling was undertaken in the Kakanui and Kauru catchments from 1992 to 1994 (Biggs et al. 1998). Periphyton biomass was estimated monthly (as chlorophyll *a*) from September 1992 to June 1994 at a site just upstream of Kauru Hill Road bridge, which is thought to mark the upstream boundary for stronghold populations of the endangered lowland longjaw galaxiids (Figure 5.1). Mean chlorophyll *a* overplotted on the hydrograph for the same period shows that biomass attained high levels during low flows from May to September 1993 (Figure 5.3). Cyanobacteria dominated the periphyton during this period, with diatoms secondarily important (Biggs et al. 1998). Note that the periphyton guideline for maintenance of trout habitat and angling values is a maximum of 200 mg/m<sup>2</sup> (for diatoms/cyanobacteria). This level was almost attained in September 1993. Note also that in rivers where *D. geminata* has established, algal mats can greatly exceed this level (e.g., Kilroy et al. 2006c, 2006b).



**Figure 5.3.** Hydrograph from the Kauru River (site 71706) from July 1992 to June 1994. Green symbols are approximate mean chlorophyll *a* from monthly collections in a research programme (Biggs et al. 1998). The upper red line is the MfE guideline for maximum algal biomass (as chlorophyll *a*) for maintenance of angling and trout habitat values. The lower blue line is the magnitude of a FRE3 flood.

### 5.5. Other potential applications

We anticipate that the DPM for average % cover and thickness of *D. geminata* will have applications for management of sites with special values, for the design of further surveillance programmes, for allocation of resources in relation to maintenance of *D. geminata*-free sites, and in decision-making over management of *D. geminata* infestations. Such decision-making could relate to prevention measures (e.g., stopping *D. geminata* arriving at places with special values), control measures (minimising nuisance growths), or eradication attempts. Hypothetical examples follow.

- Susceptibility of areas known to support populations of rare or endangered fish species can be checked. There is currently no information on what levels of *D. geminata* might adversely affect particular fish species, but locations predicted to favour high % cover, for example, could be targeted for extra protection over sites predicted to support lower cover.
- On a regional scale, the relative susceptibility of catchments can be examined and those catchments predicted to support highest % cover and thickness could be assigned highest resource expenditure, e.g., for extra signage and cleaning stations. This might be done in conjunction with a review of usage of the catchments by, for example, anglers and trampers.
- The predictions may assist with decision-making over the necessity for and/or feasibility of control measures. For example, in identified high-risk catchments with special values, it may be possible to determine in advance whether control attempts would be feasible, and then put resources into preparing a catchment or river specific treatment plan. This scenario is similar to the Kakanui example above, but assumes that *D. geminata* is not yet present.
- If *D. geminata* arrives in a river predicted to support very high % cover and thickness, where chemical eradication is not considered feasible (due to cost, low likelihood of success, for example), the likelihood of continuing blooms may be used to drive decision-making about implementing alternative control measures, such as flow manipulation (below dams).
- The cost-effectiveness of surveillance programmes may also be assessed in light of the expected impact of *D. geminata*. This again could be in conjunction with a possible control plan, to ensure that control could be attempted at the earliest stage of an infestation.

## 5.6. Model prediction formats

The predictions described in this report can be made available in map format (size up to A0), as a text file that can be imported into software of choice, as a .SHP file, or in other formats. Both maps and files can be generated to cover the entire country or specific regions or catchments. The files will contain REC reach IDs (so that they can be joined to the REC itself, Snelder et al. 2004), followed by the predictions of *D. geminata* cover and thickness for each of the three days-since-flood scenarios, and the unreliability index for that reach.

Use of the DPM (colour-coded according to % cover and thickness) will be more approximate than using the tables, but in many cases this may be more appropriate, given the uncertainties associated with the models discussed above.

## 6. Conclusions and recommendations

Key conclusions from this study are:

- The LEM performed reasonably well, given that it was produced at short notice and on the basis of very little information. This technique could be considered as an interim management tool in any future incursions of freshwater organisms about which little is known.
- New models based on a survey of *D. geminata* % cover and thickness at 146 South Island river sites suggest that environmental variables favouring *D. geminata* establishment and growth are: high lake influence; stable, hard substrates; low flow variability; longer time since a flood; and large seasonal temperature differences.
- Examination of overseas *D. geminata* presence records and associated temperatures suggest that a cool water temperature may also be important for *D. geminata*, with a possible survival limit at a mean air temperature in the coolest month of ~5.5 °C. However, water temperature data are needed to identify physiologically-based temperature limits. We suggest that development of a national water temperature database is an important goal for management of other freshwater organisms as well as *D. geminata*.
- A limitation of the new models is that 70% of river reaches (stream order  $\geq 3$ ) have at least one environmental variable outside the range in the reaches used to develop the model. Predictions for these reaches are therefore not as reliable as for reaches within the same range. Unreliability scores should be considered with all predictions.
- Predictions are presented in a series of didymo prediction maps (DPM), which indicate that, at an average number of days since a significant flood, 8% of river reaches (stream order  $\geq 3$ , unreliability = 0) in the North Island and over 50% in the South Island will attain cover of at least 10% of *D. geminata*. This represents easily visible cover. Over 5% of South Island river reaches are predicted to

support % cover of more than 60%, but this falls to only 0.04% in the North Island.

- The DPM show predictions are for mean % cover and mat thickness of *D. geminata* in wadeable parts of rivers.
- The only way to reduce the problem of unreliability of predictions is to re-run the models using new data from a wider range of affected rivers. The monitoring protocol outlined in Section 3.2.2, and detailed in a separate report (<http://www.biosecurity.govt.nz/pest-and-disease-response/pests-and-diseases-watchlist/didymosphenia-geminata/protocols>) should facilitate this.
- There is insufficient information to make predictions about the extent of *D. geminata* growth in lakes. No obvious drivers of colonisation/growth patterns were identified from surveys in several lakes. *D. geminata* may still be spreading in some of these lakes. Further monitoring is needed.
- Even with the stated limitations, the DPM should prove useful in addressing management issues that may arise as *D. geminata* continues to spread in rivers. Predictions of % cover and thickness are easier to interpret than probabilities of presence / absence, and provide a clear picture of what growth levels to expect in different rivers.

To improve the DPM, and our understanding of the environmental limitations of *D. geminata*, we recommend:

1. further data collection to extend the environmental range of the data used to generate the models, followed by generation of updated models;
2. development of a water temperature variable, to help to determine any physiological temperature limit for the species, such as a low-temperature trigger for sexual reproduction;
3. more thorough monitoring in lakes to help determine the factors influencing growth and distribution in these environments, possibly leading to a separate model for lakes.



## 7. Acknowledgements

This project has been undertaken over a relatively short time period, and many people helped to make it happen. First, thanks to Christina Vieglaiss and the didymo team at MAF Biosecurity New Zealand for funding the project and for facilitating the start of the work while negotiations were still in progress. The administration team at NIWA in Christchurch are also thanked for their support. We acknowledge the great cooperation from Regional Councils in participating in this work and thank Chris Arbuckle, Steve Ledington and Kirsten Meijer (Environment Southland), Trevor James and Rob Smith (Tasman District Council), Matt Dale and Matt Hickey (Otago Regional Council), and Shirley Hayward (Environment Canterbury) for their interest, support and data. Shirley, Trevor and Kirsten also participated in the field work, which was much appreciated. West Coast Regional Council were not directly involved, but a discussion with Johnny Horrocks on the project was also greatly appreciated. Ian Halstead (NIWA), Kieran Scott (NIWA), Gary Tong, Karen Bellew, George Hixson, and Kevin Trainor (SouthTech Ltd) also assisted with the field work, often at very short notice. Marty Black (Queenstown Harbourmaster) provided local knowledge for the Wakatipu survey, and acted as safety contact for boating; Lloyd Matherson (Te Anau) acted as safety contact for the Hollyford survey. Lindsay Hawke (NIWA) and Jim Drury (NIWA) are thanked for fast-tracking the necessary boating safety plans. Sarah Spaulding (USEPA) and Martyn Kelly (UK) kindly provided *D. geminata* presence data from North America (and other locations) and the UK, respectively. Kathy Walter assembled all the hydrological data, and the following organisations granted permission for our use of the data: Contact Energy, Environment Canterbury, Environment Southland, FRST, Meridian Energy, Otago Regional Council, The Market Place (M-co). The REC river network GIS layer has been used under an agreement with the Ministry for the Environment. We thank the Department of Conservation for use of the FWENZ database, which contains environmental variables for the REC river network. Finally, we thank Ton Snelder (CEMAGREF, France), Maurice Duncan (NIWA), Helen Rouse (NIWA), Kris McNyset (US Environmental Protection Agency), Lindsay Chadderton, and Sarah Spaulding (USEPA) for their reviews of earlier versions of some or all of this report.

## 8. References

- Araújo, M.B.; Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33: 1677-1688.
- Austin, M.P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157: 101-118.
- Biggs, B. (1995). The contribution of flood disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. *Freshwater Biology* 33: 419 - 438.
- Biggs, B.J.F. (2000). New Zealand periphyton guideline: detecting, monitoring and managing enrichment in streams. Ministry for the Environment, Wellington.
- Biggs, B.J.F.; Close, M.E. (1989). Periphyton biomass dynamics in gravel bed rivers: the relative effects of flows and nutrients. *Freshwater Biology* 22: 209 - 231.
- Biggs, B.J.F.; Kilroy, C.; Lowe, R.L. (1998). Periphyton development in three valley segments of a New Zealand grassland river: test of a habitat matrix conceptual model within a catchment. *Archiv für Hydrobiologie* 143(2): 147 - 177.
- Biggs, B.; Smith, R.A.; Duncan, M.J. (1999). Velocity and sediment disturbance of periphyton in headwater streams: biomass and metabolism. *Journal of the North American Benthological Society* 18(2): 222-241.
- Branson, J. (2006). *Didymosphenia geminata* economic impact assessment. Final report to Biosecurity New Zealand. New Zealand Institute of Economic Research. 20 p.
- Campbell, M.L. (2005). Organism Impact Assessment for potential impact of *Didymosphenia geminata*. *All Oceans Ecology*. 82 p.
- Carter, J.R.; Denny, P. 1982. Freshwater algae of Sierra Leone. IV. Bacillariophyceae: Part (i) Diatoms from the River Jong (Taia) at Naja. *Beihefte zur Nova Hedwigia* 73: 281-331.

- Chepurnov, V.A.; Mann, D.G.; Sabbe, K.; Vyverman, W. (2004). Experimental studies on sexual reproduction in diatoms. *International Review of Cytology* 237: 91-153.
- Clausen, B.; Plew, D. (2004). How high are bed-moving flows in New Zealand rivers? *Journal of Hydrology (NZ)* 43: 19-37.
- Clausen, B.; Biggs, B.J.F. (1997). Relationships between benthic biota and hydrological indices in New Zealand streams. *Freshwater Biology* 38: 327 - 342.
- Clearwater, S.J.; Jellyman, P.G.; Biggs, B.J.F.; Hickey, C.W.; Blair, N.; Clayton, J.S. (2007 in review). *Didymosphenia geminata* experimental control trials: Stage Three, Phase One (assessment of the effectiveness of Gemex™, a chelated copper formulation, in a small river). NIWA Client Report HAM2007-060 for MAF Biosecurity New Zealand. Project No: MAF07203.
- Cleve, P.T. (1894-1896). Synopsis of the Naviculoid diatoms. Konliga Svenska Vetenskaps-Akademiens Handlingar. Stockholm (Reprint A. Asher & Co., Amsterdam 1965).
- Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A.; Li, J.; Lohmann, L.G.; Loiselle, B.A.; Manion, G.; Moritz, C.; Nakamura, M.; Nakazawa, Y.; Overton, J.M.; Peterson, A.T.; Phillips, S.J.; Richardson, K.S.; Scachetti-Pereira, R.; Schapire, R.E.; Soberón, J.; Williams, S.; Wisz, M.S.; Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.
- Elith, J., Leathwick, J.R.; Hastie, T. (in review) Boosted regression trees – a new technique for modelling ecological data. Submitted to *Journal of Animal Ecology*.
- Geesey, G.G.; Wigglesworth-Cooksey, B.; Cooksey, K.E. (2000). Influence of calcium and other cations on surface adhesion of bacteria and diatoms: A review. *Biofouling* 15(1-3): 195-205.
- Hastie, T.; Tibshirani, R.J. (1990). Generalized Additive Models. Monographs on Statistics and Applied Probability, vol. 43. Chapman and Hall, London. 335 p.
- Kawecka, B.; Sanecki, J. (2003). *Didymosphenia geminata* in running waters of southern Poland - symptoms of change in water quality? *Hydrobiologia* 495: 193-201.

- Kilroy, C. (2004). A new alien diatom *Didymosphenia geminata* (Lyngbye) Schmidt: its biology, distribution, effects and potential risks for New Zealand fresh waters. NIWA Client Report CHC2004-128. For Environment Southland. 40 p.
- Kilroy, C.; Dale, M. (2006). A comparison of sampling methods for the detection of the invasive alga *Didymosphenia geminata* in New Zealand rivers. NIWA Client Report 2006-078. For Biosecurity New Zealand. 42 p.
- Kilroy, C.; Lagerstedt, A.; Davey, A.; Robinson, K. (2006a). Studies on the survivability of the invasive diatom *Didymosphenia geminata* under a range of environmental and chemical conditions. NIWA Client Report CHC2006-116. For Biosecurity New Zealand. 110 p. Revised May 2007.
- Kilroy, C.; Lambert, P.; Robinson, K.; Blair, N. (2006b). Periphyton and invertebrate monitoring programme, lower Waiau River. Results of the 2006 survey. NIWA Client Report 2006-113. For Meridian Energy. 45 p.
- Kilroy, C.; Biggs, B.J.F.; Blair, N.; Lambert, P.; Jarvie, B.; Dey, K.; Robinson, K.; Smale, D. (2006c). Ecological studies on *Didymosphenia geminata*. NIWA Client Report CHC2005-128. For Biosecurity New Zealand. 77 p.
- Kilroy, C.; Snelder, T.; Floerl, O.; Vieglais, C.C.; Dey, K. (2007). A rapid technique for assessing the suitability of areas for invasive species, applied to New Zealand's rivers. *Diversity and Distributions*
- Kilroy, C.; Snelder, T.; Sykes, J. (2005). Likely environments in which the non-indigenous freshwater diatom, *Didymosphenia geminata*, can survive, in New Zealand. NIWA Client Report 2005-043. For Biosecurity New Zealand. 43 p.
- Larned, S.; Arscott, D.; Blair, N.; Jarvie, W.; Jellyman, D.; Lister, K.; Schallenberg, M.; Sutherland, S.; Vopel, K.; Wilcock, R. (2007 in review). Ecological studies of *Didymosphenia geminata* in New Zealand, 2006 - 2007. NIWA Client Report: CHC2007-070. For MAF Biosecurity New Zealand.
- Leathwick, J.R.; Elith, J.; Hastie, T. (2006). Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling* 199: 188-196.

- Leathwick, J.; Julian, K. (2007). Production of pressure estimates for New Zealand river catchments. NIWA Client Report: HAM2007-027. For: Department of Conservation.
- Leathwick, J.; Julian, K.; Elith, J.; Ferrier, S. Snelder, T. (2007). A biologically-optimised environmental classification of New Zealand rivers and streams. NIWA Client Report: HAM2007-043. For: Department of Conservation.
- Leathwick, J.R.; Elith, J.; Chadderton, W.L.; Rowe, D.; Hastie, T. (submitted). Dispersal, disturbance, and the contrasting biogeographies of New Zealand's diadromous and non-diadromous fish species. *Journal of Biogeography*.
- Leathwick, J.R.; Rowe, D.; Richardson, J.; Elith, J.; Hastie, T. (2005). Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology* 50: 2034–2052.
- McCullagh, P.; Nelder, J.A. (1989). Generalized Linear Models. Monographs on Statistics and Applied Probability. Chapman and Hall, London. 261 p.
- McDowall, R.M.; Waters, J.M. (2002). A new species of *Galaxias* from the Kauru River, North Otago, New Zealand. *New Zealand Journal of Zoology* 29: 41-52.
- Miller, A.J. (2002) *Subset Selection in Regression*. 2<sup>nd</sup> edn. Chapman and Hall, London.
- Mosley, M.P. (1982). New Zealand River Temperature Regimes. Water & Soil Miscellaneous Publication No. 36. Ministry of Works and Development, Wellington.
- Potapova, M.; Snoeijs, P. (1997). The natural life cycle in wild populations of *Diatoma moniliformis* (Bacillariophyceae) and its disruption in an aberrant environment. *Journal of Phycology* 33: 924-937.
- Rae, R.; Hawes, I.; Hanelt, D.; Howard-Williams, C. (2001). Ultraviolet light: is it harming New Zealand's lakes? *Water & Atmosphere* 9(1): 10-12.
- Ridgeway, G. (2004). gbm: Generalized Boosted Regression Models. R package, version 1.3-5. <http://www.i-pensieri.com/gregr/gbm.shtml>
- Schwarz, A.-M.; Howard-Williams, C.; Clayton, J. (2000). Analysis of relationships between maximum depth limits of aquatic plants and underwater light in 63 New

- Zealand lakes. *New Zealand Journal of Marine and Freshwater Research* 34: 157-174.
- Snelder, T.H.; Biggs, B.J.F. (2002). Multi-scale river environment classification for water resources management. *Journal of the American Water Resources Association* 38: 1225–1240.
- Snelder, T.H.; Biggs, B.J.F.; Weatherhead, M. (2004). New Zealand River Environment Classification User Guide. MfE Report 499. Ministry for the Environment, Wellington
- Snelder, T.H.; Woods, R.; Biggs, B.J.F. (2005). Improved eco-hydrological classification of rivers. *River Research and Applications* 21: 609-628.
- Suren, A.; Biggs, B.J.F.; Kilroy, C.; Bergey, E.A. (2003). Benthic community dynamics during summer low-flows in two rivers of contrasting enrichment 1. Periphyton. *New Zealand Journal of Marine and Freshwater Research* (37): 53-70.
- Sutherland, D.; Biggs, B.J.F.; Blair, N.; Boustead, N.B.; Kelly, G.; Sykes, J. (2005). Benthic survey of the Waiau Arm to determine the distributional extent of *Didymosphenia geminata*. NIWA Client Report: CHC2005-118. For: Meridian Energy and Biosecurity New Zealand. 26 p.
- Wetherbee, R.; Lind, J.; Burke, J. (1998). The first kiss: establishment and control of initial adhesion by raphid diatoms. *Journal of Phycology* 34: 9-15.
- Wild, M.; Snelder, T.; Leathwick, J.; Shankar, U.; Hurren, H. (2005). Environmental variables for the Freshwater Environments of New Zealand River Classification. No. p. NIWA Client Report: CHC2004-086 prepared for the Department of Conservation.
- Woods, R.; Elliott, S.; Shankar, U.; Bidwell, V.; Harris, S.; Wheeler, D.; Clothier, B.; Green, S.; Hewitt, A.; Gibb, R.; Parfitt, R. (2006). The CLUES Project: Predicting the Effects of Land-use on Water Quality – Stage II. NIWA Client Report MAF05502.
- Young, R.; Smart, G.M.; Harding, J. (2004) Impacts of hydro-dams, irrigation schemes and river control works. Chapter 37 in Harding, J et al. (eds). *Freshwaters of New Zealand*. NZ Hydrological Society/NZ Limnological Society.



## Appendix 1

List of river sites surveyed for average *D. geminata* % cover and thickness (mm) in March – April 2007, and included in the development of a predictive model. Note that 7 sites surveyed were omitted from the model dataset, usually because they were suspected to have been recently invaded by *D. geminata*. These are indicated by breaks in the site number sequence.

Site	River	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness (mm)	Shade	Substr. index	Hydrological site	Days since flood
1	Waitaki	2354707	5585971	22-Mar-07	15.2	67.4	7.7	19.6	3.9	unshaded	3.50	Waitaki, Kurow	400
2	Waitaki	2353488	5586440	22-Mar-07	15.3	69.5	7.89	36.1	8.6	unshaded	3.18	Waitaki, Kurow	400
3	Waitaki	2351301	5586501	22-Mar-07	15.6	66.3	8.85	38.8	7.8	unshaded	3.80	Waitaki, Kurow	400
4	Waitaki	2348492	5586882	22-Mar-07	15.6	65.9	7.93	12.1	4.8	unshaded	4.50	Waitaki, Kurow	400
5	Waitaki	2346609	5587491	22-Mar-07	16.9	66.0	8.93	59.5	11.9	unshaded	3.75	Waitaki, Kurow	400
6	Waitaki	2344679	5587873	22-Mar-07	16.4	68.3	8.27	70.1	18.6	unshaded	4.00	Waitaki, Kurow	400
7	Waitaki	2342300	5588500	22-Mar-07	16.7	68.4	8.05	0.0	0.0	unshaded	3.90	Waitaki, Kurow	400
8	Waitaki	2340117	5589156	22-Mar-07	17.2	68.5	7.98	0.4	1.0	unshaded	3.90	Waitaki, Kurow	400
9	Waitaki	2337658	5590161	22-Mar-07	17.6	68.2	8.2	1.7	2.75	unshaded	4.50	Waitaki, Kurow	400
10	Waitaki	2328329	5592969	22-Mar-07	17.9	68.0	8.22	42.1	10.53	unshaded	4.14	Waitaki, Kurow	400
11	Waitaki	2325362	5594002	22-Mar-07	18.5	68.2	8.38	66.6	13.77	unshaded	4.00	Waitaki, Kurow	400
12	Waitaki	2323578	5594729	22-Mar-07	18.9	68.8	8.19	94.0	17.10	unshaded	4.26	Waitaki, Kurow	400
13	Waitaki	2321673	5595316	22-Mar-07	18.2	68.3	8.38	0.7	1.50	unshaded	3.60	Waitaki, Kurow	400
13	Waitaki	2321673	5595316	22-Mar-07	18.2	68.3	8.38	82.8	16.74	unshaded	3.70	Waitaki, Kurow	400
14	Waitaki	2320969	5597832	22-Mar-07	18.0	67.3	8.44	6.9	4.63	unshaded	3.68	Waitaki, Kurow	400
15	Waitaki	2317522	5598365	22-Mar-07	17.2	66.6	7.67	13.7	6.40	unshaded	3.53	Waitaki, Kurow	400
16	Waitaki	2313316	5602696	22-Mar-07	17.1	66.9	8.27	53.5	9.54	unshaded	4.50	Waitaki, Kurow	400
17	Waitaki	2310587	5605212	22-Mar-07	18.9	68.2	8.4	66.5	8.13	unshaded	4.40	Waitaki, Kurow	400
18	Waitaki	2309271	5606871	22-Mar-07	16.6	66.9	8.24	37.5	8.04	unshaded	4.85	Waitaki, Kurow	400
19	Twizel	2279279	5657359	23-Mar-07	10.4	28.4	7.31	31.5	4.38	partial	3.45	Forks, Balmoral	83
20	Twizel	2285660	5649313	23-Mar-07	13.5	31.5	7.79	2.5	4.28	unshaded	3.45	Forks, Balmoral	83
21	Twizel	2285014	5650929	23-Mar-07	13.3	32.1	7.61	19.1	6.08	unshaded	4.10	Forks, Balmoral	83

Site	River	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness (mm)	Shade	Substr. index	Hydrological site	Days since flood
22	Twizel	2283351	5652970	23-Mar-07	13.4	32.3	7.33	10.2	5.08	unshaded	4.05	Forks, Balmoral	83
23	Twizel	2280605	5654944	23-Mar-07	13.2	30.3	7.61	26.1	5.6	unshaded	3.26	Forks, Balmoral	83
24	Ahuriri	2270524	5633069	23-Mar-07	16.7	52.3	7.48	16.2	6.4	unshaded	3.70	Ahuriri, Sth Diadem	111
25	Ahuriri	2274650	5632716	23-Mar-07	18.3	54.5	8.34	4.2	4.8	unshaded	3.80	Ahuriri, Sth Diadem	111
26	Ahuriri	2261931	5630680	23-Mar-07	17.4	47.2	8.3	14.1	6.4	unshaded	4.40	Ahuriri, Sth Diadem	111
27	Ahuriri	2256064	5627755	23-Mar-07	16.6	47.4	8.03	42.5	9.7	unshaded	4.10	Ahuriri, Sth Diadem	111
28	Ahuriri	2248972	5633037	23-Mar-07	15.2	47.0	8.03	36.8	9.0	unshaded	4.30	Ahuriri, Sth Diadem	111
29	Buller	2494496	5935370	26-Mar-07	17.7	43.0	8.64	93.5	9.9	partial	5.55	Buller, Longford	116
30	Buller	2491395	5938236	26-Mar-07	19.2	41.6	8.2	55.5	8.1	unshaded	5.25	Buller, Longford	116
31	Buller	2487843	5940345	26-Mar-07	18.9	42.5	8.3	57.0	10.6	unshaded	5.20	Buller, Longford	116
32	Buller	2484428	5942509	26-Mar-07	18.1	45.5	7.6	53.0	9.9	unshaded	5.44	Buller, Longford	116
33	Buller	2480575	5944754	26-Mar-07	18.2	46.4	7.69	16.8	6.5	partial	4.95	Buller, Longford	116
34	Buller	2478416	5945172	26-Mar-07	18.1	46.4	7.69	61.0	9.3	partial	4.30	Buller, Longford	116
35	Buller	2476986	5944706	27-Mar-07	14.1	48.5	7.69	36.5	5.3	unshaded	4.19	Buller, Longford	117
36	Buller	2474529	5943769	27-Mar-07	14.0	52.8	7.69	22.8	6.1	unshaded	4.60	Buller, Longford	117
37	Buller	2473115	5943511	27-Mar-07	17.0	45.0	8.03	54.5	9.1	unshaded	4.99	Buller, Longford	117
38	Buller	2468749	5945090	27-Mar-07	18.0	46.0	8.1	27.5	4.3	unshaded	4.60	Buller, Longford	117
39	Buller	2464642	5946209	27-Mar-07	18.9	48.1	8.39	73.0	9.9	unshaded	4.85	Buller, Longford	117
40	Buller	2463110	5945905	27-Mar-07	20.0	63.3	8.6	53.5	12.0	unshaded	4.32	Buller, Longford	117
41	Buller	2461409	5943719	27-Mar-07				35.3	11.1	partial	4.30	Buller, Longford	117
42	Buller	2460392	5940872	27-Mar-07	18.6	63.0	8.82	68.0	14.7	partial	4.15	Buller, Longford	117
43	Buller	2458372	5936871	27-Mar-07	18.2	64.0	8.41	38.0	11.2	partial	4.85	Buller, Longford	117
44	Buller	2455757	5934903	27-Mar-07	18.8	68.2	8.95	27.3	8.0	partial	4.60	Buller, Longford	117
45	Buller	2449492	5936073	28-Mar-07	15.3	107.0	8.6	0.0	0.0	unshaded	3.64	Buller, Woolfs	10
46	Buller	2444620	5935838	28-Mar-07	16.2	97.0	8.3	1.4	1.0	unshaded	4.20	Buller, Woolfs	10
47	Buller	2442656	5936479	28-Mar-07	15.9	96.4	8.42	0.0	0.0	partial	3.84	Buller, Woolfs	10

Site	River	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness (mm)	Shade	Substr. index	Hydrological site	Days since flood
48	Buller	2435515	5932417	28-Mar-07	16.6	92.5	8.34	0.0	0.4	unshaded	3.45	Buller, Woolfs	10
49	Buller	2429468	5935450	28-Mar-07	16.5	98.2	8.08	2.0	1.4	unshaded	4.95	Buller, Woolfs	10
50	Buller	2420583	5929481	28-Mar-07	18.2	91.6	8.28	0.0	0.0	partial	3.45	Buller, TeKuha	10
51	Buller	2413794	5927307	28-Mar-07	17.9	88.5	7.9	0.0	0.0	unshaded	4.85	Buller, TeKuha	10
52	Buller	2398991	5929587	28-Mar-07	14.4	43.6	7.67	0.0	0.0	partial	4.75	Buller, TeKuha	10
53	Buller	2404073	5928088	28-Mar-07	16.5	86.4	7.69	0.0	0.0	partial	3.46	Buller, TeKuha	10
54	Lower Waiau	2098038	5436858	3-Apr-07	11.8	103.2	8.1	47.7	7.8	unshaded	4.40	Waiau, Tuatapere	114
55	Lower Waiau	2099820	5441307	3-Apr-07	11.5	103.1	7.96	99.0	8.8	unshaded	5.80	Waiau, Tuatapere	114
56	Lower Waiau	2102774	5444248	3-Apr-07	11.2	94.2	8.27	93.3	9.1	unshaded	7.00	Waiau, Tuatapere	114
57	Lower Waiau	2103217	5449704	3-Apr-07	11.7	90.0	8.78	97.0	15.5	unshaded	3.05	Waiau, Tuatapere	114
58	Lower Waiau	2100876	5451106	3-Apr-07	11.6	88.7	8.44	51.7	12.3	unshaded	3.32	Waiau, Tuatapere	114
59	Lower Waiau	2098273	5454722	3-Apr-07	12.7	81.6	9.3	99.5	13.8	unshaded	5.00	Waiau, Tuatapere	114
60	Lower Waiau	2097471	5459385	3-Apr-07	12.9	81.3	9	98.0	21.1	unshaded	4.60	Waiau, Tuatapere	114
61	Lower Waiau	2097642	5460240	3-Apr-07	13.1	76.5	9.2	100.0	18.5	unshaded	3.50	Waiau, Tuatapere	114
62	Lower Waiau	2097164	5464709	3-Apr-07	12.3	75.1	8.95	93.3	19.7	partial	4.41	Waiau, Tuatapere	114
63	Lower Waiau	2094389	5468794	3-Apr-07	12.5	73.4	9.1	27.8	7.2	partial	4.75	Waiau, Sunnyside	94
64	Lower Waiau	2093738	5472425	3-Apr-07	12.8	71.0	9.2	100.0	16.4	unshaded	4.70	Waiau, Sunnyside	94
65	Lower Waiau	2093395	5474935	3-Apr-07	11.9	68.4	8.9	97.1	2.8	unshaded	5.60	Waiau, Sunnyside	95
66	Lower Waiau	2091822	5478826	4-Apr-07	11.6	57.9	8.12	96.5	12.6	partial	5.25	Waiau, Sunnyside	95
67	Lower Waiau	2091260	5478897	4-Apr-07	9.2	87.9	8.15	86.3	11.2	partial	5.05	Waiau, Sunnyside	95
68	Lower Waiau	2091298	5487148	4-Apr-07	10.3	86.3	8.47	78.8	17.5	unshaded	4.05	Waiau, MLC	111
69	Lower Waiau	2095288	5494654	4-Apr-07	11.3	75.1	9.04	81.0	14.1	unshaded	4.65	Waiau, MLC	111
70	Lower Waiau	2095699	5495419	4-Apr-07	11.6	84.5	8.76	88.5	12.5	unshaded	3.90	Waiau, MLC	111
71	Lower Waiau	2096051	5496485	4-Apr-07	11.6	81.2	8.65	64.0	7.9	unshaded	4.70	Waiau, MLC	111
72	Upper Waiau	2092335	5509822	4-Apr-07	13.4	36.8	8.15	43.3	6.2	unshaded	4.25	Waiau, Queens	300
73	Upper Waiau	2092892	5510539	4-Apr-07	13.4	39.7	8.09	86.0	18.0	partial	5.05	Waiau, Queens	300

Site	River	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness (mm)	Shade	Substr. index	Hydrological site	Days since flood
74	Upper Waiau	2094839	5511772	4-Apr-07	13.2	39.7	8.11	93.0	16.4	partial	4.20	Waiau,Queens	300
75	Von	2149879	5559174	7-Apr-07	7.5	68.8	8.74	12.5	2.3	full	3.80	Shotover, Bowen Peak	31
76	Von	2146686	5556844	7-Apr-07	8.7	70.4	8.21	1.5	0.6	unshaded	3.65	Shotover, Bowen Peak	31
77	Von	2146221	5554599	7-Apr-07	9.3	57.4	8.02	0.0	0.0	unshaded	4.00	Shotover, Bowen Peak	31
78	Von	2142570	5548226	7-Apr-07	7.6	55.9	8.31	25.0	2.7	unshaded	4.40	Shotover, Bowen Peak	31
79	Whitestone	2100453	5506947	6-Apr-07	11.6	110.6	7.85	0.3	0.2	partial	4.00	Mararoa, Cliffs	125
81	Whitestone	2111286	5523492	6-Apr-07	9.6	96.8	8.17	25.9	3.2	unshaded	3.85	Mararoa, Cliffs	125
82	Whitestone	2116607	5525553	2-Apr-07	16.0	108.7	.	42.5	5.1	unshaded	4.50	Mararoa, Cliffs	121
83	Whitestone	2121721	5531676	2-Apr-07	11.4	90.3	.	10.5	3.1	unshaded	4.36	Mararoa, Cliffs	121
84	Eglinton	2103171	5547649	9-Apr-07	7.3	91.9	8.77	4.5	1.8	unshaded	4.40	Mararoa, Cliffs	128
85	Eglinton	2108482	5548872	9-Apr-07	7.1	96.2	8.46	8.3	2.3	unshaded	4.50	Mararoa, Cliffs	128
86	Eglinton	2113994	5555409	9-Apr-07	7.1	91.1	8.37	0.0	0.0	unshaded	4.30	Mararoa, Cliffs	128
87	Eglinton	2116604	5566659	9-Apr-07	7.3	72.8	8.06	5.3	2.2	unshaded	4.15	Mararoa, Cliffs	128
88	Eglinton	2117672	5574812	9-Apr-07	9.1	63.9	7.85	55.5	11.9	unshaded	4.60	Mararoa, Cliffs	128
89	Eglinton	2120970	5578694	9-Apr-07	11	59.7	8.21	100.0	20.0	unshaded	4.30	Mararoa, Cliffs	128
91	Eglinton	2117173	5561097	9-Apr-07	6.5	92.3	8.04	69.5	7.5	unshaded	5.15	Mararoa, Cliffs	128
93	Oreti	2134541	5531178	7-Apr-07	9.3	38.4	7.96	16.3	2.0	unshaded	2.98	Oreti, Three Kings	123
94	Oreti	2129966	5520379	10-Apr-07	6.9	52.1	8.58	42.1	6.5	unshaded	4.43	Oreti, Three Kings	128
95	Oreti	2129672	5517702	10-Apr-07	7.1	55.1	8.61	78.0	9.1	unshaded	3.65	Oreti, Three Kings	128
96	Oreti	2128699	5515134	10-Apr-07	7.9	44.9	9.1	81.5	10.5	unshaded	3.59	Oreti, Three Kings	128
97	Oreti	2130499	5508179	10-Apr-07	9.4	57.3	9.25	55.5	8.3	unshaded	3.80	Oreti, Three Kings	128
98	Oreti	2140120	5494697	11-Apr-07	8.4	.	7.95	2.0	1.5	unshaded	3.97	Oreti, Lumsden Cableway	34
99	Oreti	2154322	5489010	11-Apr-07	9.6	.	7.96	13.8	3.0	unshaded	4.17	Oreti, Lumsden Cableway	34
100	Oreti	2152783	5480406	11-Apr-07	9.7	.	7.85	0.0	0.0	unshaded	3.70	Oreti, Lumsden Cableway	34
101	Oreti	2150011	5469200	11-Apr-07	10.3	84.5	7.74	0.0	0.0	unshaded	3.95	Oreti, Lumsden Cableway	34
102	Oreti	2147632	5462546	11-Apr-07	10.6	90.4	7.96	6.3	2.6	unshaded	3.70	Oreti, Wallacetown	101

Site	River	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness (mm)	Shade	Substr. index	Hydrological site	Days since flood
103	Oreti	2147265	5451034	11-Apr-07	10.7	.	7.98	5.4	2.2	unshaded	3.75	Oreti, Wallacetown	101
104	Oreti	2146733	5443675	11-Apr-07	10.8	98.3	8.16	0.0	0.0	unshaded	3.90	Oreti, Wallacetown	101
105	Oreti	2145287	5421038	11-Apr-07	12.2	103	7.23	0.0	0.0	unshaded	3.10	Oreti, Wallacetown	101
106	Mararoa	2096871	5497962	10-Apr-07	10.7	86.8	8.58	73.0	9.1	partial	3.28	Mararoa, Cliffs	33
107	Mararoa	2100005	5503449	10-Apr-07	11.4	109.6	8.08	40.8	7.3	partial	3.59	Mararoa, Cliffs	33
108	Mararoa	2111005	5506096	6-Apr-07	11.5	57	9.37	49.8	11.8	unshaded	4.40	Mararoa, Cliffs	29
109	Mararoa	2117649	5510812	6-Apr-07	10.8	54.7	8.25	60.0	12.5	unshaded	4.20	Mararoa, Cliffs	29
110	Mararoa	2124418	5520811	10-Apr-07	10.4	55.8	8.62	74.0	15.0	unshaded	4.19	Mararoa, Cliffs	33
111	Mararoa	2127951	5528216	7-Apr-07	10.6	45.3	8.24	100.0	17.6	unshaded	6.48	Mararoa, Cliffs	30
113	Aparima	2131066	5424447	11-Apr-07	11.8	163	7.41	0.0	0.0	unshaded	3.02	Aparima, Thornbury	102
114	Aparima	2128800	5432564	11-Apr-07	11.8	152.5	7.24	0.0	0.0	partial	2.75	Aparima, Thornbury	102
115	Aparima	2123745	5441026	11-Apr-06	12.5	107.7	7.16	0.1	0.0	unshaded	3.15	Aparima, Otautau	132
116	Aparima	2131824	5452698	11-Apr-07	10.6	.	8.2	5.8	2.2	unshaded	4.10	Aparima, Otautau	132
117	Aparima	2131545	5462705	11-Apr-07	9.5	.	8.1	15.0	1.9	unshaded	3.40	Aparima, Otautau	132
118	Aparima	2131862	5473218	11-Apr-07	8.8	.	8.25	43.5	8.7	unshaded	4.20	Aparima, Dunrobin	132
119	Aparima	2129981	5485461	11-Apr-07	8.3	88.3	8.26	90.0	10.0	unshaded	6.47	Aparima, Dunrobin	132
121	Hollyford	2122053	5611398	12-Apr-07	7.1	26.9	7.16	0.0	0.0	unshaded	2.80	Haast, Roaring Billy	1
122	Hollyford	2122095	5608509	12-Apr-07	7.7	37.7	7.25	0.0	0.0	partial	2.90	Haast, Roaring Billy	1
123	Hollyford	2122944	5604520	12-Apr-07	7.73	30.9	7.27	0.0	0.0	unshaded	3.30	Haast, Roaring Billy	1
124	Hollyford	2123884	5601456	10-Apr-07	8.2	63.1	8.16	22.8	8.7	unshaded	4.15	Haast, Roaring Billy	12
125	Hollyford	2125361	5594118	10-Apr-07	7.9	110.3	7.68	2.0	1.2	partial	5.45	Haast, Roaring Billy	12
127	Clutha	2253037	5441872	17-Apr-07	12.5	71.7	7.83	80.0	1.5	unshaded	2.70	Clutha, Balclutha	127
128	Clutha	2248124	5444276	17-Apr-07	12.8	71.9	7.74	0.0	0.0	unshaded	1.80	Clutha, Balclutha	127
129	Clutha	2241787	5450561	17-Apr-07	12.5	71.6	7.89	99.0	10.0	unshaded	4.15	Clutha, Tuapeka	127
130	Clutha	2239422	5459751	17-Apr-07	12.9	69.4	7.86	92.5	12.5	unshaded	4.09	Clutha, Tuapeka	127
131	Clutha	2236733	5466044	17-Apr-07	11.8	60.1	7.99	90.0	11.3	unshaded	4.20	Clutha, Tuapeka	127

Site	River	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness (mm)	Shade	Substr. index	Hydrological site	Days since flood
132	Clutha	2237756	5472428	17-Apr-07	13.2	71.8	8.04	13.8	11.6	unshaded	4.10	Clutha, Tuapeka	127
133	Clutha	2239967	5481934	17-Apr-07	13.4	71.1	8	42.5	6.7	unshaded	3.60	Clutha, Tuapeka	127
134	Clutha	2237741	5489276	17-Apr-07	13.9	70.1	7.93	100.0	11.4	partial	4.45	Clutha, Tuapeka	127
135	Clutha	2230145	5498851	17-Apr-07	14	70.3	8.28	82.0	4.4	unshaded	2.90	Clutha, Tuapeka	127
136	Clutha	2222558	5507861	17-Apr-07	14.1	68.6	8.31	100.0	10.0	unshaded	4.40	Clutha, Tuapeka	127
137	Clutha	2226551	5543849	17-Apr-07	13.6	69.3	8.06	93.0	5.0	partial	4.90	Clutha, Clyde	300
138	Clutha	2220394	5551342	17-Apr-07	13.3	69.3	7.5	6.5	0.7	unshaded	4.10	Clutha, Clyde	300
139	Clutha	2210742	5589201	18-Apr-07	13.2	66.6	7.74	96.0	10.8	unshaded	4.09	Clutha, Cardrona confl	300
140	Clutha	2210534	5594698	18-Apr-07	13.5	65.6	8.31	81.8	10.7	unshaded	4.00	Clutha, Cardrona confl	300
141	Clutha	2215543	5601889	18-Apr-07	14.4	65.3	9.06	85.3	14.3	unshaded	4.20	Clutha, Cardrona confl	300
142	Clutha	2208150	5607181	18-Apr-07	14.3	71.9	8.31	73.1	6.9	partial	3.70	Clutha, Cardrona confl	300
143	Clutha	2206931	5608185	18-Apr-07	14.3	72.3	8.38	69.4	18.4	partial	4.30	Clutha, Cardrona confl	300
144	Clutha	2219402	5585606	18-Apr-07	14.5	58.2	9.18	89.0	9.9	unshaded	4.50	Clutha, Cardrona confl	300
147	Fraser	2220196	5548802	19-Apr-07	13.3	70.5	8.27	77.0	10.9	partial	4.55	Fraser, Laing Rd	140
148	Fraser	2221199	5547438	19-Apr-07	13.5	70.7	8.59	80.5	8.7	unshaded	2.90	Fraser, Laing Rd	140
149	Fraser	2223964	5544861	19-Apr-07	13.3	96	7.91	48.4	3.9	partial	3.60	Fraser, Laing Rd	140
150	Upukerora	2098490	5519837	6-Apr-07	9.6	114.9	8.68	0.0	0.0	unshaded	3.50	Mararoa, Cliffs	125
151	Makarora	2213419	5665484	23-Apr-07	9.8	63	7.38	36.0	2.7	partial	4.00	Matukituki, W. Wanaka	41
152	Makarora	2208581	5658122	23-Apr-07	9.3	69.2	7.61	0.0	0.0	partial	2.70	Matukituki, W. Wanaka	41
153	Makarora	2206723	5654365	23-Apr-07	9.6	68.5	7.91	33.9	6.1	unshaded	3.15	Matukituki, W. Wanaka	41



## Appendix 2

### Boosted regression tree models

From Appendix 1 in Leathwick et al. (submitted)

Boosted regression tree methodology (BRT), also known as stochastic gradient boosting (Friedman et al. 2000, Friedman 2001, 2002) is a relatively new addition to the range of tools available for modelling relationships between species distributions and environment (Elith et al. in review). Boosted methods such as BRT differ substantially from regression-based methods such as Generalised Linear Models (GLM – McCullagh & Nelder 1989) and Generalised Additive Models (GAM – Hastie & Tibshirani 1990) that have been used widely over the last decade for such analyses (Guisan and Zimmerman 2000). While these latter methods seek to identify a single “best” model describing relationships between the response and predictor variables, a boosted model progressively builds a sequence of models of increasing complexity, each one fitting the training data slightly better than its predecessor. The model-building process is referred to as 'forward stagewise', which reflects the fact that at each step a term is added to the model to slightly decrease the deviance. The terms added are in the form of small regression trees, which are fit to the gradient of the deviance (generalized residuals). Results from this ensemble of models are then averaged to form a final prediction. While the strong performance of boosted methods has been known for a number of years, we are aware of only a few instances in which they have been applied to the analysis of ecological data (Kawakita et al. 2005, Leathwick et al. 2006, Moisen et al. 2006, Caputo et al. 2005).

While in theory, boosting can be applied to any model fitting method, the use of regression trees as the individual model terms in a BRT model has particular advantages. These include their resistance to outliers, tolerance of extraneous predictors, capacity to accommodate missing values, ability to automatically fit interactions between predictors, and flexibility for modelling a variety of responses (Friedman & Meulman 2003). Fitting a BRT model requires optimisation of a number of parameters, of which the most important controls the complexity of the individual regression trees fitted as model terms, typically taking an integer value in the range from 1 to 10. This parameter is sometimes referred to as the interaction depth, reflecting its control over the potential fitting of interactions between predictors. An individual tree that consists of only a single decision rule will fit a purely additive model (Hastie et al. 2001), but the potential for fitting interactions between predictors increases as the complexity of the individual trees is increased, with the maximum degree of interaction that can be accommodated by each model term equalling the number of splits. In practice, the use of more complex individual trees does not force fitting of interactions, as even quite complex trees are capable of fitting purely additive effects where these predominate in the data being analyzed. As a consequence, the degree of interaction between predictors fitted by any model must be determined using diagnostic procedures such as those that we describe below. All models in this analysis were fitted using individual trees consisting of three nodes or decision rules. This value was chosen as giving the best overall predictive performance for the dataset size available in this study, after inspection of results using various tree sizes ranging from one through to seven.

The other parameter that affects model outcome is the learning rate, i.e., the weighting given to individual model terms. Small values give a low weight to individual model terms, so that model complexity is built up gradually, generally giving better model fit but at the expense of higher

computational demands. Our experience suggests that when the learning rate is kept constant, more trees are required to fit species of high prevalence than those of low prevalence (Elith et al., in review). We therefore set the learning rate differently for the two models so as to achieve a model containing 800-1200 trees each.

Models were fitted using a purpose-written script that implemented a 10-fold cross-validation procedure in which models of increasing complexity were fitted to 10 temporary data sets, each of which comprised 90% of the full dataset. Each of these datasets was created by withholding one of ten mutually exclusive subsets of the full dataset, each containing 10% of the data selected randomly. For each model, trees were added in sets of 50, after which predictions were formed for the points withheld from that dataset. These predictions were then compared with the withheld data by calculating the predictive deviance for each subset. Model complexity was increased until there was no further decrease in the deviance when predicting to the withheld data, averaged across all subsets, which typically reached a minimum before increasing again at higher levels of model complexity, indicating over-fitting of the data; by contrast, the training deviance typically continues to decline as the models become progressively over-fitted to the training data. We then identified the level of model complexity that gave the lowest predictive error and fitted a final model to the full dataset using this number of trees.

Because BRT provides no indication of the statistical significance of the contribution from individual predictors, and simply ignores those that have no relevance, we developed a procedure to identify those predictors that could be removed without degrading predictive performance. This was achieved using a 10-fold cross-validation procedure in which ten models were fitted to separate folds, each consisting of 90% of the dataset. In contrast to the application of cross-validation described in the previous paragraph, here it is used to robustly estimate the minimum number of predictors that will achieve similar predictive performance to that achieved when using the original predictor set. The number of predictors was then progressively reduced in each of these models by removing at each step, the predictor with the lowest contribution in the previous model (Miller 1990). Predictions were then made to the withheld data, and the average predictive performance was calculated across all folds as a function of the number of predictors removed. A final model was then fitted using the smallest number of predictors that retained predictive performance equivalent to a model fitted with all predictors, with the predictors to be dropped identified using a 'voting' system that calculates their average order of removal during the cross-validation analysis.

We detected the presence of interaction effects between predictor variables using purpose written code that examined the relationship between predicted values and all possible pair-wise combinations of predictors. This was achieved for each pair of predictors by creating two variables ( $x_1$ ,  $x_2$ ) that consisted of values at constant intervals along their ranges, and forming predictions ( $y'$ ) on the logit scale for all possible combinations of these. In making these predictions, values for all other variables were set at their mean for the dataset. We then used a linear model to relate the predicted values to the values of the two marginal variables, i.e.  $y' \sim x_1 + x_2$ , with the two predictor variables fitted as factors. Where the predicted values are formed by a largely additive combination of the two predictors, this regression object will have very low residual variance. However, where interaction effects increase the complexity of the predicted surface, a significant amount of variance in  $y'$  will be left unexplained. We recorded the residual deviance for linear models fitted separately to predictions formed using all pair-wise combinations of predictors for all species, and used these to assess the

relative strength of interaction effects fitted for each predictor pair by calculating its median value across all species.

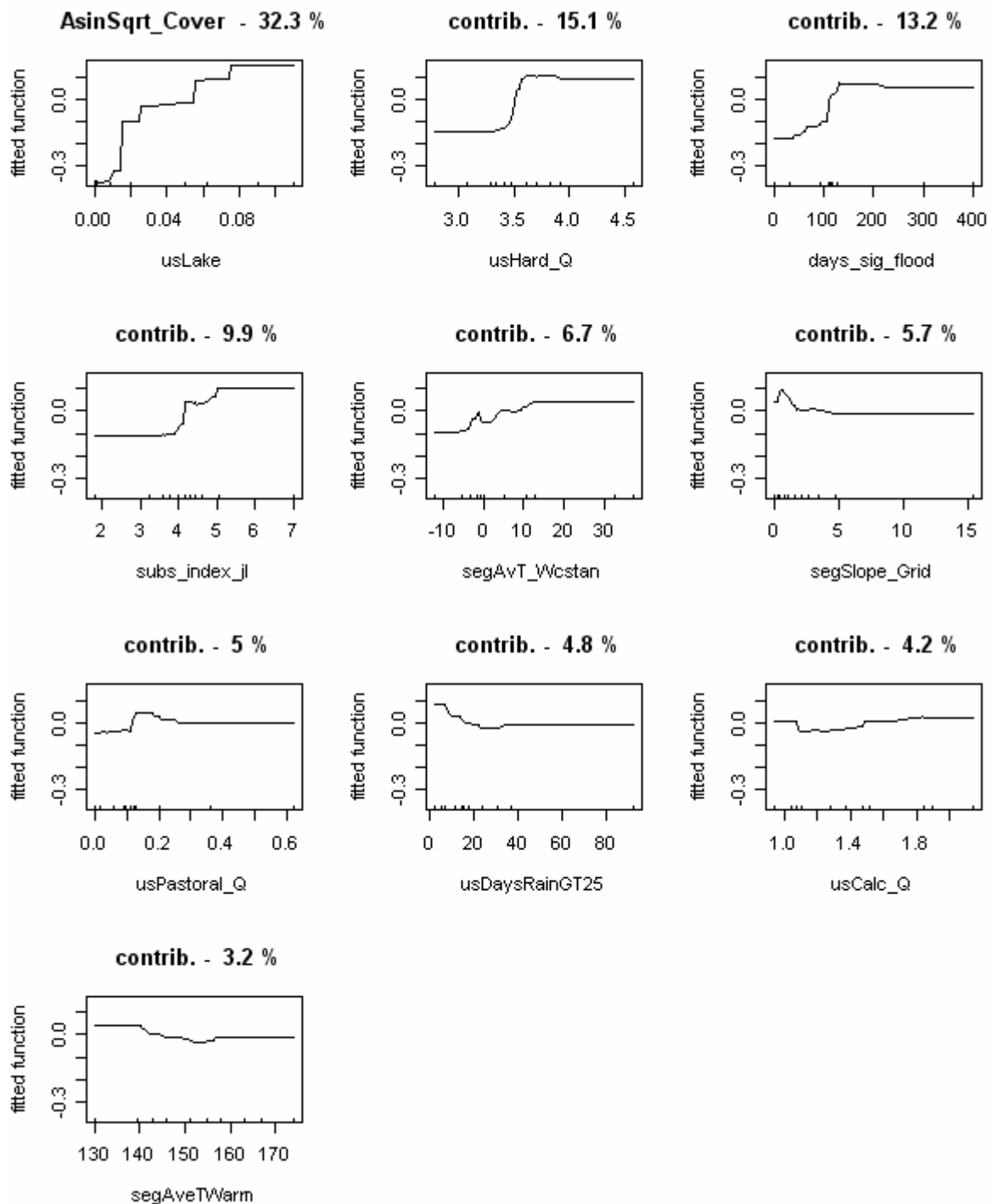
## References

- Cappo, M.; De'ath, G.; Boyle, S.; Aumend, J.; Olbrich, R.; Hoedt, F.; Colton, P.; Brunskill, G. (2005). Development of a robust classifier of freshwater residence in barramundi (*Lates calcarifer*) life histories using elemental ratios in scales and boosted regression trees. *Marine and Freshwater Research* 56: 713–723.
- Elith, J.; Leathwick, J.R.; Hastie, T. (in review). Boosted regression trees – a new technique for modelling ecological data. Submitted to *Journal of Animal Ecology*.
- Friedman, J.H. (2001). Greedy function approximation: the gradient boosting machine. *Annals of Statistics* 29: 1189–1232.
- Friedman, J.H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis* 38: 367–378.
- Friedman, J.H.; Hastie, T.; Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *The Annals of Statistics* 28: 337–407.
- Friedman, J.H.; Meulman, J.J. (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine* 22: 1365–1381
- Guisan, A.; Zimmerman, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.
- Hastie, T.; Tibshirani, R.J. (1990). Generalized Additive Models. Chapman and Hall, London.
- Hastie, T.; Tibshirani, R.; Friedman, J.H. (2001). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York.
- Kawakita, M.; Minami, M.; Eguchi, S.; Lennert-Cody, C.E. (2005). An introduction to the predictive technique AdaBoost with a comparison to generalized additive models. *Fisheries Research* 76: 328–343.
- Leathwick, J.R.; Elith, J.; Francis, M.P.; Hastie, T.; Taylor, P. (2006). Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series* 321: 267–281.
- McCullagh, P.; Nelder, J.A. (1989). Generalized linear models. 2nd edn. Chapman and Hall, London.
- Miller, A.J. (2002). Subset Selection in Regression. 2nd edn. Chapman and Hall, London.
- Moisen, G.G.; Freeman, E.; Blackard, J.; Frescino, T.; Zimmermann, N.E.; Edwards Jr., T.C. (2006). Predicting tree species presence in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecological Modelling* 199: 176–187.

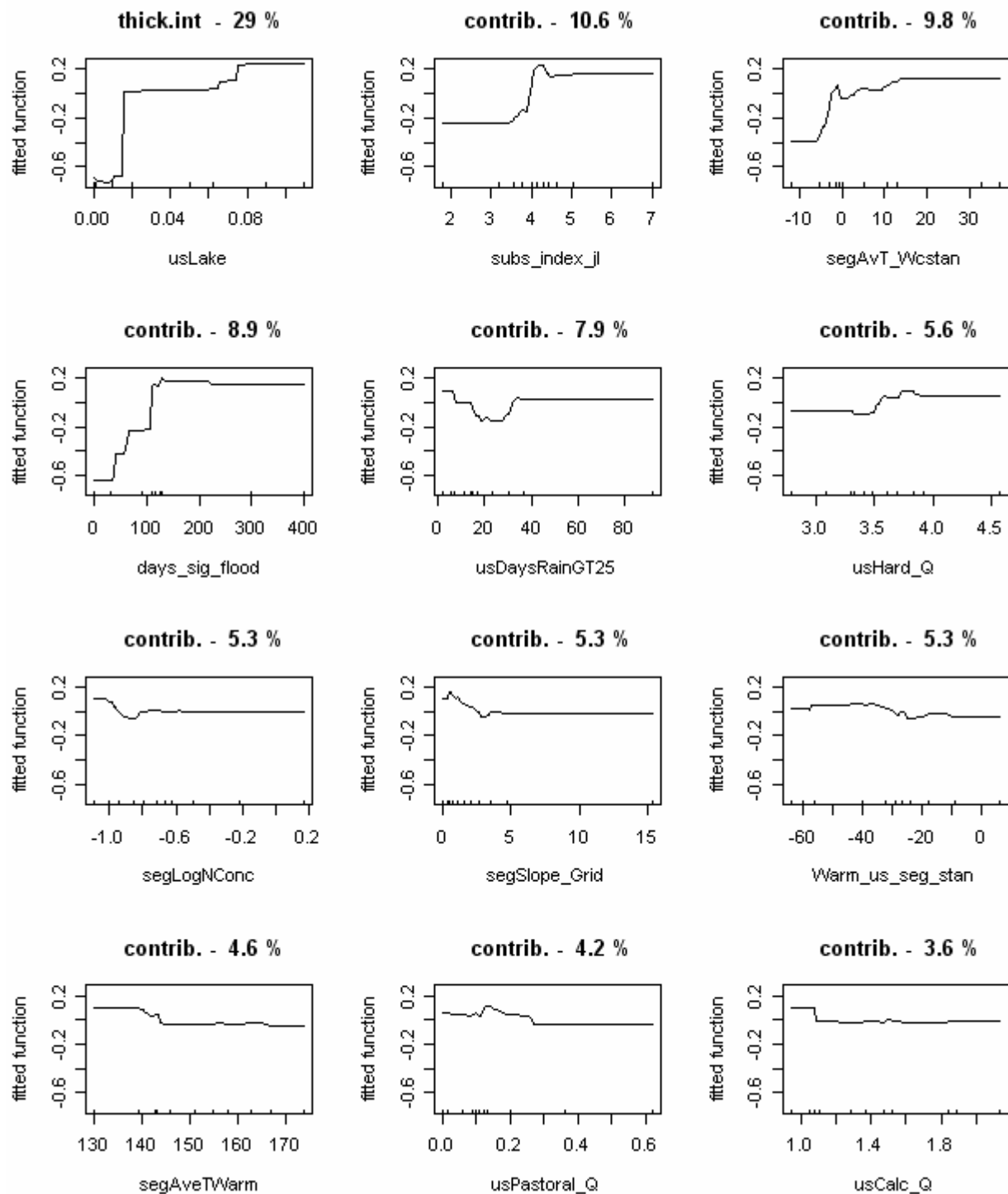
## Appendix 3

Response curves (a, b) and interaction plots (c, d) for variables important in determining distribution of *D. geminata* in New Zealand.

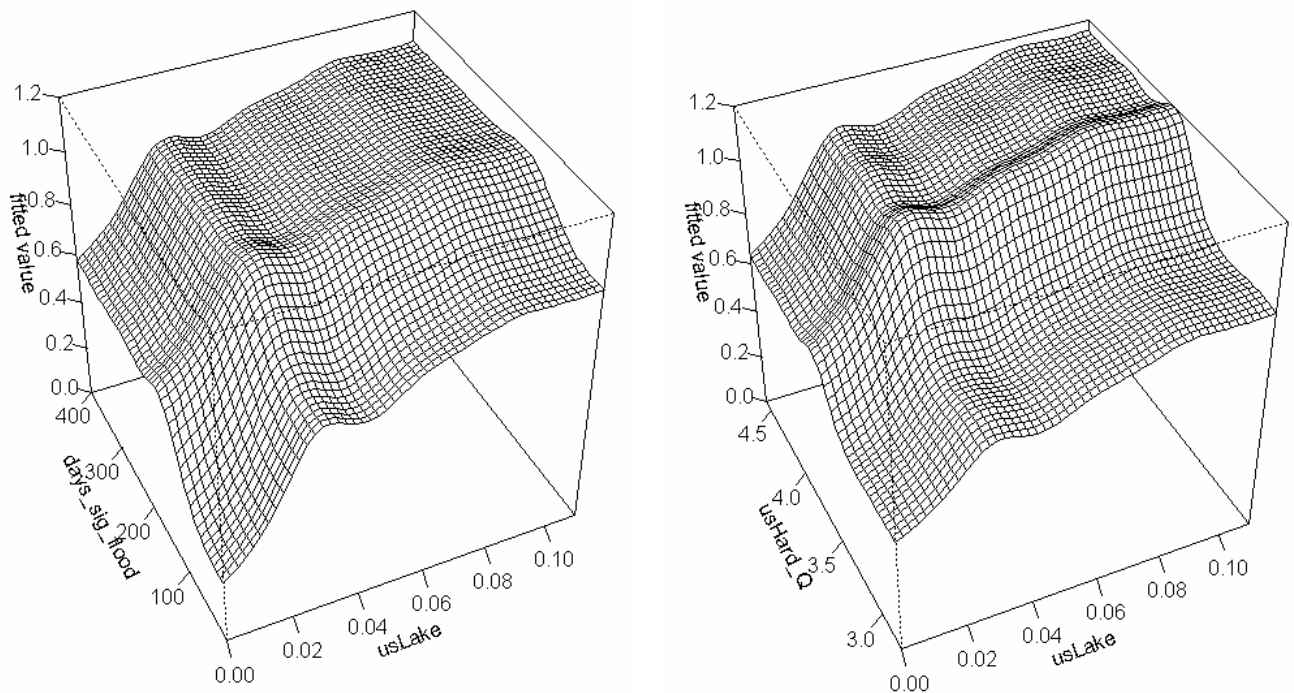
### (a) Response curves: model for % cover



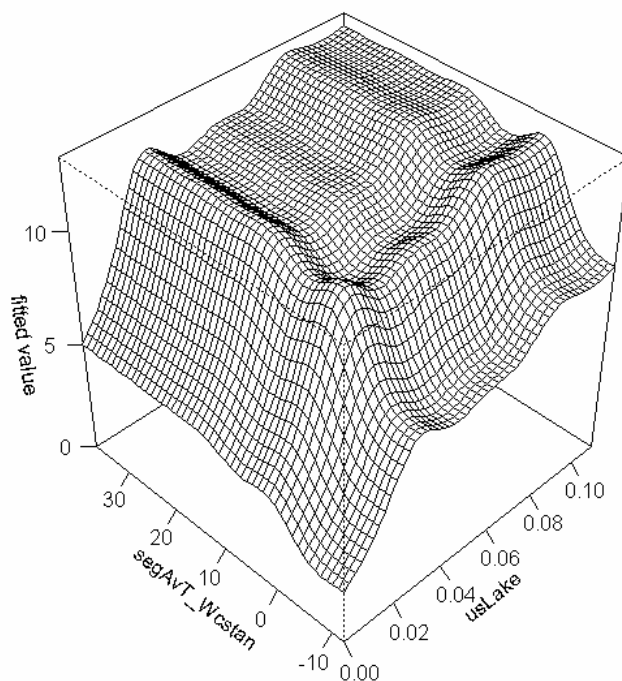
**(b) Response curves: model for *D. geminata* mat thickness (m)**



**(c) Interactions between *usLake* and *days\_sig\_flood*, *usLake* and *usHard*, for % cover**



**(d) Interaction between *usLake* and *segAvT\_Wcstan*, for thickness**





## Appendix 4

Lake sites surveyed for average % cover and thickness (mm) of *D. geminata*, in April – May 2007.

Site	Lake	site	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness	Sample checked?	shade	Substrate index
1	Roxburgh	u/s Musterer's Hut Bay	2223740	5522591	19-Apr-07	13.7	71.8	8.05	0.0	0.0	No	unshaded	6.5
2	Roxburgh	Frenchman's Track	2226852	5543295	19-Apr-07	14	69.5	8.12	0.0	0.0	No	unshaded	2.8
3	Roxburgh	Shingle Creek	2221467	5525269	19-Apr-07	14.2	71.5	8.21	0.0	0.0	No	unshaded	1.8
4	Dunstan	Champagne Gully	2220246	5557836	19-Apr-07	11.4	70.1	8.14	0.0	0.0	No	unshaded	3.62
5	Dunstan	Nine Mile Creek, Jacksons	2216425	5562319	19-Apr-07	11.2	72.3	8.15	0.0	0.0	No	unshaded	3.65
6	Dunstan	Bridge, right bank	2212212	5567631	19-Apr-07	10.6	72.9	8.27	2.5	1.5	No	unshaded	3.6
7	Dunstan	Bridge, left bank	2212404	5567560	19-Apr-07	12.3	71.1	8.15	0.0	0.0	No	full	3.6
8	Dunstan	Lowburn	2211626	5570836	18-Apr-07	11.8	70.5	7.97	0.0	0.0	No	unshaded	4.15
9	Dunstan	Pisa Moorings	2213805	5574345	18-Apr-07	12.5	69.2	8.7	15.5	1.6	Yes	unshaded	4.17
10	Dunstan	Bendigo Boat Ramp	2216955	5578602	18-Apr-07	14.6	68.9	9.3	20.3	1.4	Yes	unshaded	4.5
11	Dunstan	John Bull Creek	2213742	5572879	18-Apr-07	14.4	72.9	8.64	0.0	0.0	No	unshaded	3.8
12	Manapouri	Richters Rock	2086619	5502370	14-Apr-07	13.4	32.5	6.4	0.0	0.0	No	unshaded	2.95
13	Manapouri	Surprise Bay	2086452	5503004	14-Apr-07	13.6	30.7	7.39	0.0	0.0	No	unshaded	7
14	Manapouri	Pearl Harbour Entrance	2088905	5501881	14-Apr-07	14	31.3	7.56	0.0	0.0	No	unshaded	4.45
15	Manapouri	Frazers Beach	2089879	5502757	14-Apr-07	13.9	31.7	7.57	34.5	5.0	Yes	unshaded	4.6
16	Manapouri	Supply Bay	2088259	5505968	14-Apr-07	13.9	31.4	7.61	0.0	0.0	No	partial	7
17	Manapouri	Iris Burn	2087051	5511091	14-Apr-07	14	31.5	7.7	0.0	0.0	No	unshaded	3.96
18	Manapouri	Rona Island, Lee Side	2084079	5509876	14-Apr-07	14	31.2	7.73	0.0	0.0	Yes	partial	6.9
19	Manapouri	Stockyard	2080871	5504597	14-Apr-07	14.1	31.5	7.65	0.0	0.0	No	partial	6.75
20	Manapouri	Monument	2083724	5501149	14-Apr-07	14.1	31	7.71	0.0	0.0	Yes	unshaded	6.05
21	Manapouri	Stoney Point	2086422	5505402	14-Apr-07	14.8	30.8	7.61	0.0	0.0	No	unshaded	6.3
22	Te Anau	Glaisnock	2092429	5565173	22-Apr-07	13.8	42.6	.	0.0	0.0	No	unshaded	4.25
23	Te Anau	Cathy's Creek	2103010	5558398	22-Apr-07	14.2	41.4	.	0.0	0.0	No	partial	5.1
24	Te Anau	Billy Burn	2105223	5562219	22-Apr-07	13.8	40.1	.	0.0	0.0	No	unshaded	3.7
25	Te Anau	Dot Island	2104288	5552745	22-Apr-07	13.6	42.3	.	0.0	0.0	No	unshaded	5.2
26	Te Anau	Te Anau Downs	2102414	5543583	22-Apr-07	13.6	39.9	.	0.0	0.0	No	unshaded	4.7
27	Te Anau	Centre Island	2099971	5539802	22-Apr-07	13.7	39.9	.	0.0	0.0	No	unshaded	3.4

Site	Lake	site	E	N	Survey date	Temp	Cond	pH	D. gem % cover	D. gem thickness	Sample checked?	shade	Substrate index
28	Te Anau	Ettrick Burn	2095657	5537593	22-Apr-07	14	40.7	.	0.0	0.0	No	unshaded	4.35
29	Te Anau	Manuka Bay	2101447	5536704	22-Apr-07	13.8	42.5	.	81.0	5.0	No	unshaded	4.4
30	Te Anau	Ewe Burn	2099086	5527665	22-Apr-07	13.7	46.8	.	70.0	2.0	Yes	unshaded	4.2
31	Te Anau	Welcome Point	2102556	5548757	22-Apr-07	14.4	41.6	.	0.0	0.0	No	unshaded	3.85
32	Te Anau	MidBurn	2099359	5547459	22-Apr-07	14.2	41.3	.	74.3	5.0	No	unshaded	4.25
33	Te Anau	Te Anau Caves Wharf	2096643	5533106	22-Apr-07	14.3	41.5	.	43.0	4.0	No	unshaded	4.6
34	Te Anau	Bluegum Point	2096497	5521097	22-Apr-07	13.5	73.2	.	99.0	7.5	No	unshaded	4.75
35	Wakatipu	Wilson Bay/7-mile track	2161110	5562263	9-May-07	11.8	62.3	.	0.0	0.0	Yes	unshaded	3.8
36	Wakatipu	Wilson Bay/7-mile track	2160953	5562539	9-May-07	11.8	62.3	.	0.0	0.0	Yes	unshaded	3.8
37	Wakatipu	Sunshine Bay	2165252	5564171	9-May-07	12.2	59.6	.	5.0	0.5	Yes	partial	3.8
38	Wakatipu	Fernhill Roundabout	2167227	5565392	9-May-07	9.9	69.5	.	62.8	5.9	No	unshaded	4.61
39	Wakatipu	Town Pier	2168135	5566007	9-May-07	7.9	66	.	0.0	0.0	No	unshaded	2.85
40	Wakatipu	Gardens	2168010	5565510	9-May-07	12.6	63.8	.	17.5	0.5	Yes	unshaded	2.95
41	Wakatipu	Skating Rink	2168293	5565417	9-May-07		65	.	7.0	0.9	Yes	partial	2.25
42	Wakatipu	Walter Peak	2159452	5556781	10-May-07	11.8	67	.	0.0	0.0	Yes	unshaded	2.65
43	Wakatipu	Betw. Mt Nic and Walter Peak	2156596	5556104	10-May-07	12	65.3	.	23.8	2.3	Yes	unshaded	3.1
44	Wakatipu	Mt Nicholas Stn	2150239	5558065	10-May-07	12.3	65.6	.	43.2	4.4	No	unshaded	2.91
45	Wakatipu	Elfin Bay Stream	2144183	5572569	10-May-07	12.4	67.6	.	7.5	0.8	Yes	unshaded	4.35
46	Wakatipu	Opposite Islands east side	2149009	5576303	10-May-07	12.2	78.7	.	11.3	2.4	Yes	unshaded	3.49
47	Wakatipu	Mt Creighton Stn	2150490	5568663	10-May-07	12.9	67.1	.	72.3	5.6	No	partial	2.8
48	Wakatipu	Near Rat Pt	2152796	5560967	10-May-07	12.6	68.5	.	22.0	2.8	No	unshaded	3.55
49	Wakatipu	Southwest Bay	2173406	5542287	10-May-07	11.7	67.5	.	24.0	2.5	No	partial	4.45
50	Wakatipu	Halfway Bay	2171664	5548115	10-May-07	11.8	66.3	.	91.0	9.5	No	partial	4.15
51	Wakatipu	Collins Bay	2171344	5555983	10-May-07	11.9	67.6	.	81.3	9.5	No	unshaded	3.8
52	Wakatipu	Picnic area South Arm	2177397	5537455	11-May-07	11.4	65.4	.	70.8	11.5	No	unshaded	3.85
53	Wakatipu	Halfway B. boat ramp (E. side)	2175555	5547720	11-May-07	11.9	68	.	60.3	5.5	No	unshaded	4.4
54	Wakatipu	Lakeside Estate Wharf area	2175513	5558782	11-May-07	12	71.1	.	0.0	0.0	Yes	unshaded	4.15
55	Wakatipu	Earnslaw Boat slip	2169524	5565262	11-May-07	12.9	63.8	.	0.0	0.0	Yes	partial	4.7
56	Wakatipu	Bay by Kelvin Hts	2169595	5564525	11-May-07	11.6	68.1	.	0.0	0.0	Yes	unshaded	4.6
57	Wakatipu	Frankton Harbour	2172429	5567814	11-May-07	11.9	68.5	.	0.5	0.2	Yes	partial	3.2

