DEVELOPMENT OF NATIONAL EXTENT TERRAIN ATTRIBUTES (TANZ), SOIL WATER BALANCE SURFACES (SWATBAL), AND ENVIRONMENTAL SURFACES, AND THEIR APPLICATION FOR SPATIAL MODELLING OF *PINUS RADIATA* PRODUCTIVITY ACROSS NEW ZEALAND

A thesis submitted in fulfilment
of the requirements for the degree of

Doctor of Philosophy

by

David John Palmer

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Department of Earth and Ocean Sciences
The University of Waikato, Hamilton
New Zealand

2008

View of the Routeburn Flats with Mount Earnslaw in the background (photograph by Adam Vonk).

Toitū he whenua, whatungarongaro he tangata
The land is permanent, man disappears

Whakatauki
Māori proverb

ABSTRACT

The most widely distributed and commercially important forestry crop in New Zealand is *Pinus radiata* D. Don. Until recently foresters have focussed on maintaining plantation management systems that are highly productive, while remaining sustainable. However, the new era of reduced carbon emissions and carbon trading means forestry systems are now viewed as potential sinks for the sequestration of carbon. Never before has the need to quantify the productive capacity of New Zealand's plantation forests at the national extent been so great. Furthermore, regions of relatively low productivity may become increasingly desirable because these sites require lower capital outlay.

In this research, a series of spatial surfaces potentially useful in the modelling and mapping of forest productivity across the national extent of New Zealand have been developed. Modelled surfaces include 15 primary and four secondary terrain attributes; 13 shortwave radiation surfaces, topographically adjusted (one annual and 12 monthly surfaces); and 39 soil water balance model surfaces (one annual and 12 monthly surfaces for fraction of available root zone water storage, available root zone water storage, and drainage).

Terrain attributes were developed using a 25-m floating point DEM and are unique and currently the best comprehensive surfaces for the following reasons.

(1) Terrain attributes comprehensively encompass the entire country compared with previous piecemeal and site-specific surfaces.

(2) Terrain attributes were modelled using a macro-catchment concept that divides the New Zealand landscape into large, naturally draining catchments to avoid the modelling problems associated with edge effects at catchment boundaries.

(3) Upslope contributing areas were calculated by switching between an FD8 algorithm that modelled flow divergence in upland regions above defined stream channels and a D8 algorithm used in low-lying areas where modelling of flow convergence is appropriate.

(4) Where appropriate, terrain attributes were corrected for undesirable spurious sinks inherent in the 25-m floating point DEM, while retaining naturally occurring sinks in karst environments, depressional lakes and wetlands. This correction provided a continuous surface that modelled flow either to a sink or continuously across the surface until reaching the sea.

The soil water balance model, SWatBal, is a dynamic spatial model that can be updated over time as new and improved data become available. SWatbal calculates the fraction of available root-zone water content, available root-zone water content, and drainage for the *P. radiata* species at a 100-m resolution throughout New Zealand. SWatBal was applied in this study to derive monthly mean soil water balance values, but the model can easily be adjusted to calculate any spatial extent or period. A further advance of SWatBal is the development of "reasoned and allocated virtual" (RAV) rainfall data. RAV consist of 365 rainfall surfaces representing the "normal rainfall distribution" on a monthly basis. The

advantage RAV data have over monthly mean rainfall is that rainfall distribution of an actual month is used, making the data realistic, rather than assuming constant rainfall across each day for a month.

A shortwave-radiation model was developed for New Zealand at a 25-m cell-size resolution utilising a national extent DEM and a latitude surface. This shortwave radiation model encompassed slope and aspect adequately while simultaneously accounting for the influence of terrain shading. As a model it has simplicity, flexibility, and minimal computation time and storage requirements.

A partial least squares (PLS) regression technique was used to develop the surfaces of (i) stem volume mean annual increment at age thirty years for a defined reference regime of 300 stems ha^{-1} (300 Index), and (ii) mean top height at age twenty (Site Index) using TANZ, SWatBal and other developed and existing New Zealand spatial datasets. Together, (i) and (ii) provided the basis for a spatial model of *P. radiata* productivity. Initially, the 300 Index and Site Index values were calculated for 1698 permanent sampling plot (PSP) locations. For cross validation purposes, 552 PSP sites were withheld from all modelling procedures. PLS regression was used to model and predict 300 Index and Site Index values using previously developed and some existing datasets including climate, landuse, terrain, and their environmental surfaces. Best models explained 58% and 67 % of the variance for 300 Index and Site Index, respectively. The PLS models were also used to develop quantitative productivity maps across the national extent of New Zealand. In addition, a regression kriging (RK) technique was used, where

ordinary kriging (OK) of the PLS model residuals was undertaken to improve model outcomes by summing the PLS and OK surfaces. Cross validation showed that prediction precision increased for both the 300 Index and Site Index RK models. However, only Site Index predictions were considered less biased using the RK technique. Findings from the commonly used and relatively straight forward spatial interpolation technique, inverse distance weighting (IDW), were compared with those derived using the more complex RK, OK, and PLS techniques. Cross validation showed that all techniques performed better than their respective data means. OK, RK, and IDW techniques were similar in prediction precision with the IDW prediction precision best for the 300 Index, and RK best for the Site Index. However, OK predictions showed reduced prediction bias. Having stated that RK, OK, and IDW interpolation techniques provided overall better predictions than PLS, it is emphasised that cross validation locations only occur within currently forested landscapes. Beyond these forested regions PLS regression has an inordinate advantage over OK and IDW prediction techniques by utilising local environmental and landform information. Additionally, there is the potential of prediction improvement through the coupling of the PLS model with its kriged regression residuals. Indeed, the main purpose of producing the 300 Index and Site Index maps was to provide empirically based predictions of regions currently without forests as much as regions with forests through spatial interpolation of existing national extent observed PSP data. Possible drivers of *P. radiata* productivity across 14 broad LENZ-derived environmental regimes were also assessed. It was found that generally air temperature and water balance variables were the predominate drivers.

ACKNOWLEDGEMENTS

Postgraduate study for me was a wonderful journey providing opportunities to meet and work with many interesting people from a wide variety of backgrounds. Throughout this journey I was blessed to have so many who took the time to assist me in my quest for knowledge and, no matter how great or small their contribution, I extend my thanks and gratitude to them all. This project was a co-operative endeavor between Scion Research (formerly Ensis and its predecessor, Forest Research) and the University of Waikato, and I am grateful to many people for their support and encouragement throughout my project.

I was fortunate to have the guidance of three outstanding supervisors, David Lowe (University of Waikato), Barbara Höck (Scion Research), and Tim Payn (Scion Research), whose collective wide ranging skills and knowledge guided and greatly enhanced my research. Firstly, I wish to thank my chief supervisor, Professor David Lowe, for his unwavering support and never ending enthusiasm for my project. His continued advice, guidance, and editorial help were greatly appreciated. David also dealt with ongoing administrative facets including supporting scholarship applications, sorting out complex university contracts, and lobbied for much needed and expensive datasets related to my project. Most of all, thank you David for allowing me the freedom to explore the wide-ranging scientific disciplines including pedology, soil science, hydrology, pedometrics, geostatistics, and environmental modelling that eventually resulted in this thesis and associated scientific publications. We are a great team and I hope we can continue to work together and publish papers in the future.

I owe immense gratitude and thanks to Barbara Höck for her supervisory role, ongoing support, and enthusiasm throughout this project. Barbara's advice and guidance from the conceptual stages through to the more technical issues within this project were invaluable. Indeed, her experience, knowledge and skills were central to the successful outcomes of my project. No matter how great the technical difficulties were, Barbara always gave wide ranging discussion relating to the issues and then put forward logical solutions arising from the debate. I owe a great deal to Barbara, who first introduced me to the amazing world of geostatistics and spatial analysis. I greatly appreciate the time she gave me over the years, helping me develop and hone the skills of these disciplines. Because of Barbara's efforts, my view of the world has been changed forever – thank you.

I would like to express my gratitude and thanks to Dr Tim Payn for his continued support and guidance throughout my study. Tim has a great way of cutting to the chase and teasing out the basic raw components of a proposal and then giving the project a holistic perspective. This ability to provide an overview is unique and highly valued when those working constantly on the detail sometimes lose sight of the end point or outcome. Furthermore, Tim is thanked for giving me the key that opened the door to many of the resources of Scion Research. This generosity meant unlimited access not only to national datasets, software packages, laboratories and the national forestry library, but also to the substantial human resources available at Scion Research.

Projects like this can only be successful by collaborating with and pooling the knowledge and skills from a raft of highly skilled people. This is especially true for three people who were critically important contributors towards my research: Dr Michael Watt, Mark Kimberley, and Andrew Dunningham, all from Scion Research. Mike was instrumental in initiating the development of SWatBal and contributed greatly to the forthcoming publications from this work. Mike is especially thanked for enduring my endless questions and for having the patience to guide me through the myriad of possible modelling component alternatives. I enjoyed Mike's informal exchanging of ideas and yet I really appreciated his refreshingly frank comments when it came to the important stuff – writing scientific papers to international standards. Mike is sincerely thanked also for having the patience to allow me to develop the manuscripts at my own pace and in my own words.

Mark Kimberley was instrumental in developing statistical methods that could manage the project's large national datasets. We soon discovered it is one thing to work on data from one local region, yet quite another to deal with and manage the huge volumes of data when modelling across New Zealand's national extent. Mark is also thanked for the many hours of hard work in preprocessing and developing PSP data into forest productivity data. Mark treated me as a colleague and his door was always open to discuss my barrage of questions – thank you.

Andrew Dunningham played an instrumental role in the development of TANZ. I thank Andy for his assistance in compiling the TAPE-G code, developing extraction methods,

and helping to overcome the many technical issues related to national scale modelling. Andy is especially thanked for his good company and infectious good humour and laughter that always lifted our spirits.

I would like to express my thanks and gratitude to Dr David Whitehead from Landcare Research for his continued support and guidance in the development of SWatBal. David's advice from behind the scenes through Mike Watt was instrumental in the success of SWatBal. David Whitehead is also thanked for reading and commenting on the entire thesis, as is Dr Robin Thwaites (Queensland University of Technology).

I would like to express my gratitude to numerous staff of Scion Research who took the time out to discuss issues related to my work. I thank Mark Dean and Carolyn Anderson who sorted out the permissions needed, then extracted and preprocessed the PSP data. Furthermore, I thank the forestry companies who gave permission to access their PSP data. This was quite a leap of faith allowing an unknown, David Palmer, to use their private forest productivity data, especially in an uncertain forestry environment surrounded with issues of carbon credits.

I gratefully acknowledge FRST for awarding me their Top Achiever Scholarship, and the University of Waikato for awarding me a University of Waikato Doctoral Scholarship. A special mention to Scion's Forests and Environment group, led by Peter Clinton, for coming forward and supporting me with course-related costs in the last year of my project. Scion Research is also thanked for assisting with journal colour publishing costs.

Black and white could never provide enough detail and would have undersold what we have achieved. Similarly, Dr Dave Campbell (current COD) of the Department of Earth and Ocean Sciences at the University of Waikato is thanked for absorbing reasonable costs of colour figures within the thesis.

Finally, but most importantly I thank all my friends and family for their support, understanding and patience throughout my PhD study.

CONTENTS

LIST OF FIGURES

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

LIST OF TABLES

CHAPTER FIVE

CHAPTER SIX

View of “The Organs” and the interesting landscape carved by the Mohaka River, Hawke’s Bay (photograph by Dr Kyle Bland)

Examine each question in terms of what is ethically and aesthetically right, as well as what is economically expedient. A thing is right when it tends to preserve the integrity, stability and beauty of the biotic community. It is wrong when it tends otherwise.

Aldo Leopold - *Sand County Almanac*

CHAPTER ONE

GENERAL INTRODUCTION

1. BACKGROUND

Over recent years the forest industry has focussed on the productive capacity and sustainability of plantation crops (Payn and Thwaites, 1996; Watt et al., 2005). Today, in the light of predicted global climate change and carbon budgeting, there is a move in the forest industry towards modelling carbon sequestration (Coomes et al., 2002) and developing maps that can provide information related to forest productivity and site quality (Payn et al., 1999). The purchase of land for forestry purposes and carbon sequestration has historically targeted high-producing sites requiring large capital outlays. However, new plantation forest developed for carbon sequestration may focus on locations with lower capital outlay. It is therefore not surprising that the interest and demand for *Pinus radiata* productivity models and predictive maps have increased dramatically since the inception of this study.

In New Zealand the two main estimates of *P. radiata* productivity and site quality are the 300 Index, defined as the stem volume mean annual increment at age thirty years with a reference regime of 300 stems ha^{-1} (Kimberley et al., 2005), and Site Index, defined as the mean top height at age twenty (Goulding, 2005). Previous studies have modelled the 300 Index and Site Index at different locations across New Zealand (Jackson and Gilford,

1974; Hunter and Gibson, 1984; Woollons et al., 2002; Watt et al., 2005) with reasonable degrees of success. However, the work by Eyles (1986) is the only study to develop a site ranking map across the national extent of New Zealand. Here I define 'national' to mean the archipelago of New Zealand, whereas 'extent' refers to spatial extent, avoiding any confusion with definitions relating to scale. The main reasons these earlier productivity models were not developed into maps describing spatial variation are as follows.

1. 300 Index and Site Index data are developed from permanent sampling plots (PSP), which have remained in private ownership and therefore issues of data access have prevailed.
2. Logistical, measurement and collection costs for research other than large scale projects have meant national extent projects have remained cost-prohibitive.
3. The collection of environmental field data that correlate with the 300 Index and Site Index data has historically only been undertaken for specific, well-resourced studies. Therefore, national extent projects have tended to remain expensive because of logistical, measurement and collection costs.

Although high costs of data collection will likely continue to prevail into the future, new technologies including geographical information systems (GIS) and prediction methods including spatial interpolation are beginning to provide potential tools for the development of both predictive models and map development. Indeed, the use of GIS and geostatistics for New Zealand forestry systems is well recognised (Höck et al., 1994; Payn et al., 1999; Palmer et al., 2005). Like all computer technology, there has been a

huge advancement of GIS technology over the last decade. Because of these advancements, it is now possible to process and store large volumes of spatial data quickly and efficiently. Further, many reference books and thematic journal issues are now dedicated to the disciplines of GIS (e.g., Burrough and McDonnell, 1998; Heuvelink, 1998; Brimicombe, 2003) and geostatistics (e.g., Journel and Huijbregts, 1978; Webster and Oliver, 1990, 2007; Cressie, 1991; Goovaerts, 1997, Diggle and Ribeiro, 2007). Modelling of even large-scale studies were previously found to be problematic, but are now considered relatively straight forward with GIS technology. In addition, national extent modelling of environmental properties across countries or entire continents is now possible using GIS and associated database management systems (e.g., see Henderson et al., 2001; 2005).

In studies of the prediction of soils information in Australia, a national soils point database (Henderson et al., 2005) was used to fill gaps in soil survey information covering large spatial extents (Bui and Moran, 2003). This research used techniques including expert knowledge and rule-based induction (Bui and Moran, 2001). New Zealand's soil database (Landcare Research, 2008) comprises around 1300 profile descriptions with associated physical and chemical data dating back to the 1970s. This database has potential for amalgamation with new, more recently collected soil information including that arising from the 500 Soils project (Landcare Research, 2008). However, land ownership privacy issues restrict the identification of observation locations, thereby limiting the use of this resource. The fundamental soil layers (FSL) (Newsome et al., 2000) currently provide the best nationally available soils information

dataset. S-map (Landcare Research, 2008), although still in the development stages, will provide an important array of national soils information in digital form in the future. Numerous reference books and journal articles are dedicated to the prediction of soil attributes including those associated with the disciplines of pedology, pedometrics, and digital soil mapping (e.g., Mulla and McBratney, 2000; Guo et al., 2003; Hengl et al., 2003; Grunwald, 2005; Lagacherie et al., 2006; Carre et al., 2007). Indirect soil information includes a number of surfaces used in the development of land environments of New Zealand (LENZ) (Leathwick et al. 2002b; 2003). However, for the successful modelling and mapping of forest productivity, other supplementary information relating soils and their landscapes is required.

Landform influences the amount of solar energy reaching the Earth's surface, which impacts on the distribution and nature of organisms occurring there (Blaszcynski, 1997). Terrain attributes are intrinsically related to landform, and to some extent the soils and organisms that grow there. Primary (e.g., elevation, slope, aspect, curvature) and secondary terrain attributes (e.g., upslope contributing area, topographic wetness index, stream power index) provide essential information for environmental modelling in a wide range of disciplines. The digital elevation model (DEM) forms the foundation or building blocks needed for modelling all terrain attributes (Wilson and Gallant, 2000). Many papers discuss issues related to grid resolution and its sensitivity on terrain attributes (e.g., Florinsky, 1998; Wilson et al., 2000; Thompson et al., 2001; Claessens, 2005). Generally, the development of terrain attributes from a high-quality DEM of appropriate resolution is an excellent method of providing supplementary information related to

forests (Mummery et al., 1999). Primary terrain attributes are often derived directly from a DEM, whereas secondary terrain attributes are calculated from the primary terrain attributes. The development of appropriate terrain attributes corresponds to objective one in this thesis, as noted in Section 1.3 below.

Plantation productivity is often attributed to limiting environmental factors including temperature, water, light and nutrients (Nambiar and Sands, 1993). Notwithstanding cost, rainfall, solar radiation, and temperature data are routinely collected for the calculation of soil water balance modelling. Climate surfaces including rainfall, solar radiation, and temperature of a known accuracy are readily available for New Zealand (Leathwick et al., 2002a; Tait et al., 2006). Soil water balance has long been recognised as a major determinant of forest productivity at locations exhibiting seasonal water deficits (Arneth et al., 1998a, 1998b; McMurtrie et al., 1990; Richardson et al., 2002; Watt et al., 2003). However, to date, soil water balance has not been proven as a major determinant of forest productivity at the national extent. Historically, New Zealand models have substituted soil water balance with the relatively low cost variable, rainfall. Although spatial estimates of soil water balance have been calculated for New Zealand (Barringer et al. 1995; Barringer and Lilburne, 1999), these models are not applicable to plantation forestry systems because they were developed for agricultural systems. Therefore, the development of soil water balance surfaces across New Zealand's national extent would provide information potentially relevant in the modelling of plantation productivity. The development and assessment of a forestry soils water balance model for New Zealand corresponds with objectives two and three (see below).

Modelling *Pinus radiata* productivity and the development of spatially representative maps is not without its difficulties, especially at the national extent. There are many paths open for modelling forest productivity including geostatistical techniques, regression modelling, classification and regression trees, expert knowledge and rule-based induction, generalised linear models (GLM), generalised additive models (GAM), to name some. Scientists including Austin et al. (1995) have discussed the appropriateness of different modelling techniques. A New Zealand example of modelled native tree species distribution was derived using various environmental and climate variables (Leathwick, 1995; 1998; 2001; Leathwick et al., 1996). Land environments of New Zealand (LENZ) uses spatial layers including root zone water deficit, and other environmental properties related to climate and landform, to model and map environmental regions of similar character using GAMs (Leathwick et al., 2002b; 2003). Historically, national models researching the productivity of the *P. radiata* species across New Zealand have used multiple linear regression (MLR) (Jackson and Gilford, 1974; Hunter and Gibson, 1984; Woollons et al., 2002; Watt et al., 2005). Before the inferences made from MLR can be considered valid, a number of assumptions require verifying (Rogerson, 2006). Some assumptions of multiple linear regression are that (1) independent variables are few in number; (2) such variables are not significantly redundant (collinear); and (3) they have well understood relationships to the responses. Cross validated partial least squares (PLS) regression may provide an alternative modelling technique for situations where the first and second conditions above may have been compromised. Tegelmark (1998) illustrated the use of the PLS technique for

modelling the regeneration of Scots pine in Sweden. The development of PLS regression models and their associated maps for New Zealand corresponds with objectives four and five.

The use of global regression may benefit from the application of techniques such as geostatistics. For example, further improvements to multiple linear regression model predictions can be obtained using regression kriging techniques as described by Odeh et al. (1995). Inverse distance weighting (IDW), ordinary kriging (OK), regression kriging (RK), and regression models are commonly used interpolation techniques. Regression modelling and kriging techniques require practitioner experience and expertise when compared with using IDW which is a relatively simple method to use. The international literature for geostatistics is extensive, and some of the fundamental texts include Journel and Huijbregts (1978), Webster and Oliver (1990), Cressie (1991), and Goovaerts (1997). However, the direct application of interpolation methods to *P. radiata* plantations for New Zealand remains at the development stage. Therefore, gaining an understanding of the predictive performance of forest productivity models is an essential step towards the comprehensive evaluation of plantation forest environments. The comparisons of spatial interpolation methods for the 300 Index and Site Index surfaces correspond to objective seven.

2. AIM AND SCOPE OF THIS THESIS

The two main aims of this research were (1) to develop productivity and site quality models using ancillary maps and environmental surfaces; and (2) to develop these models into national extent maps capable of delineation, assessment and evaluation of the spatial variability of 300 Index and Site Index across New Zealand.

3. OBJECTIVES

The specific objectives of the research were as follows.

1. To produce robust terrain attributes for New Zealand using the TAPES-G program in conjunction with GIS technology.
2. To produce a spatial soil water balance model, SwatBal, specific to *Pinus radiata* for the assessment of spatial variability in soil moisture across New Zealand.
3. To establish the utility of interpolated long-term climate data and SWatbal as a possible surrogate for field measured data in determining *Pinus radiata* productivity across New Zealand.
4. To develop the 300 Index and Site Index models with a quantifiable level of certainty using the surfaces developed in objectives 1, 2, and 3.
5. From these productivity models, develop national extent maps capable of evaluating *P. radiata* productivity across New Zealand.

6. To compare spatial interpolation techniques, inverse distance weighting (IDW), ordinary kriging (OK), partial least squares regression (PLS), and regression kriging (RK), for predicting the 300 Index and Site Index across New Zealand; and to assess the precision and bias of IDW, OK, PLS, and RK spatial predictions.

4. THESIS STRUCTURE AND CHAPTER OUTLINE

This thesis comprises seven chapters including a general introductory chapter (Chapter 1), five chapters written as manuscripts or papers for publication (Chapters 2, 3, 4, 5, and 6), and an overarching concluding chapter (Chapter 7). A series of appendices are provided at the end of the thesis. All chapters except 1 and 7 are written in a stand-alone format to facilitate their publication as refereed Scion Research technical bulletins (Chapters 2 and 3) and as refereed journal publications (Chapters 4, 5, and 6). Therefore, chapters 2–6 are self contained, each having its own abstract, introduction, main body, conclusions, acknowledgements and references. Chapters 1 and 7 serve to provide structure and form while facilitating connectivity and flow throughout the thesis. Chapter 3, describing SWatBal, contains an additional list of symbols not provided in other chapters. Chapter 3 has an appendix of 39 water balance graphics associated with the SWatBal bulletin that have been placed at the end of the thesis to help reader fluency.

Chapter 2 is a published technical bulletin defining the protocols and processes involved in the development of 15 primary and four secondary terrain attributes, and 13 topographically adjusted shortwave radiation surfaces (one annual and 12 monthly

surfaces). The rationale behind the development of these surfaces is that although terrain attributes are frequently developed and used in New Zealand, they tend to remain piecemeal and site specific. Therefore, comprehensive surfaces covering New Zealand's national extent were developed from the best available data, using reliable and trusted methodologies and algorithms, in conjunction with soundly developed processes and protocols. The outcomes from this work are surfaces that can be utilised in modelling the productivity of *P. radiata* across New Zealand.

Chapter 3 is written as a technical bulletin that defines the processes and protocols in developing the national extent soil water balance (SWatBal) model for New Zealand. This bulletin describes a fully spatial model capable of defining available root-zone water storage, fractional available root-zone water storage, and drainage for any or all locations across New Zealand.

Chapter 4 describes an application of SWatBal and spatially interpolated long-term climate data. It compares the utility of developing *P. radiata* productivity models that use measured observations with models using SWatBal and long-term interpolated climate data.

Chapter 5 brings together all the national extent, independent data developed in chapters 2–4 for the modelling of *P. radiata* 300 Index and Site Index. Partial least squares (PLS) regression was used to develop models and their derived maps illustrating the spatial variability of forest productivity across New Zealand. Improvements were made to model

predictions through regression kriging (RK). Cross validation procedures defined model precision and bias, providing an estimation of improvements made to the PLS technique using RK.

Chapter 6 explores spatial interpolation techniques including inverse distance weighting (IDW), ordinary kriging (OK), regression kriging (RK), and regression techniques such as partial least squares (PLS). The purpose of this chapter is to compare the relatively easy-to-use interpolation technique, IDW, with techniques that require increased practitioner experience, expertise and cost, including OK, RK, and PLS regression. This chapter also looks at the issues of sparse data that are spread irregularly spatially, and the effectiveness of the range over which the kriging technique semi-variograms operate.

Chapter 7 provides a summary and synthesis of the results given in previous chapters and summarises the main conclusions.

5. THESIS CONTRIBUTIONS

This thesis comprises research led and written by the candidate. My supervisors and others, as noted in the acknowledgements, have made various contributions to the research and this contribution has been recognised by their inclusion as co-authors on the chapters that have been or will be submitted as publications.

6. REFERENCES

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998a. Assessment of annual carbon exchange in a water-stressed *Pinus radiata* plantation: an analysis based on eddy covariance measurements and an integrated biophysical model. Global Change Biology 5, 531-545.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998b. Fluxes of carbon and water in a *Pinus radiata* forest subject to soil water deficit. Australian Journal of Plant Physiology 25, 557-570.

Austin, M.P., Meyers, J.A., Belbin, L., Doherty, M.D., 1995. Simulated data case study, Sub-project 5, Modelling of landscape patterns and processes using biological data. Division of Wildlife and Ecology, Commonwealth Scientific and Industrial Research Organisation, 99 pp.

Barringer, J.R.F., Porteous, A., Salinger, M.J., Trangmar, B.B., 1995. Estimating spatial patterns of wilting point deficit using a water balance model and a geographic information system. Journal of Hydrology 34, 42-59.

Barringer, J., Lilburne, L., 1999. Scale issues in developing regional-scale soil water balance surfaces. In: Whigham, P.A., (Ed.), Proceedings of the Eleventh Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand, pp. 231-240.

Blaszcynski, J.S. 1997. Landform characterization with geographic information systems. Photogrammetric Engineering and Remote Sensing 63, 183-191.

Brimicombe, A., 2003. GIS, Environmental Modelling and Engineering. Taylor and Francis, London, pp. 312.

Bui, E.N., Moran, C.J., 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. Geoderma 103, 79–94.

Bui, E.N., Moran C.J., 2003. A strategy to fill gaps in soil survey over large spatial extents: an example from the Murray-Darling basin of Australia. Geoderma 111, 21-44.

Burrough P.A. 1986. Principles of Geographical Information Systems for Land Resources Assessment. Oxford University Press, New York.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographic Information Systems. Oxford University Press, New York.

Carre, F., McBratney, A.B., Mayr, T., Montanarella, L., 2007. Digital soil assessments: Beyond DSM. Geoderma 142, 69-79.

Claessens, L. 2005. Modelling Landslide Dynamics in Forested Landscapes: Addressing landscape evolution, landslide soil redistribution and vegetation patterns in the Waitakere Ranges, west Auckland, New Zealand. PhD thesis, Wageningen University, Netherlands, 143pp.

Coomes, D.A., Allen, R.B., Scott, N.A., Goulding, C., Beets, P., 2002. Designing systems to monitor carbon stocks in forests and shrublands. Forest Ecology and Management 164, 89-108.

Cressie, N. 1991. Statistics for Spatial Data. New York, Wiley.

Diggle, P. J., Ribeiro Jr., P.J. 2007. Model-based Geostatistics. New York, Springer, 228 pp.

Eyles, G.O., 1986. *Pinus radiata* site index rankings for New Zealand. New Zealand Forestry 31, 19-22.

Florinsky, I.V. 1998. Accuracy of local topographic variables derived from digital elevation models. International Journal of Geographical Information Science 12, 47-61.

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, USA, 483 pp.

Goulding, C.J., 2005. "Measurement of Trees"; Section 6.5 of the NZIF Forestry Handbook, 4th Edition, Mike Colley (Ed.), NZIF. 318 pp.

Grunwald, S. (ed.) 2005. Environmental Soil-landscape Modelling: Geographic Information Technologies and Pedometrics. Books in Soils, Plants, and the Environment 111. CRC Press, Boca Raton, 504 pp.

Guo, Y., Gong, P., Amundson, R., 2003. Pedodiversity in the United States of America. Geoderma 117, 99-115.

Henderson, B., Bui, E., Moran, C., Simon, D., 2001. ASRIS: Continental-scale soil property predictions from point data. CSIRO Land and Water, Technical Report 28/01, Canberra, Australia.

Henderson, B., Bui, E., Moran, C., Simon, D., 2005. Australia-wide predictions of soil properties using decision trees. Geoderma 124, 383-398.

Hengl, T., Rossiter, D.G., Sttein, A., 2003. Soil sampling strategies for spatial prediction by correlation with auxilliary maps. Australian Journal of Soil Research 41, 1403-1422.

Heuvelink, G.B.M. 1998. Error Propagation in Environmental Modelling with GIS. Taylor and Francis, London, UK, 127 pp.

Höck, B.K., Payn, T.W., Shirley, J.W., 1994. Using a geographic information system and geostatistics to estimate site index of *Pinus radiata* for Kaingaroa Forest, New Zealand. New Zealand Journal of Forestry Science 23, 264-277.

Hunter I.R., Gibson, A.R., 1984. Predicting *Pinus radiata* site index from environmental variables. New Zealand Journal of Forestry Science 14, 53-64.

Jackson, D.S., Gifford, H.H., 1974. Environmental variables influencing the increment of radiata pine (1) periodic volume increment. New Zealand Journal of Forestry Science 4**,** 3-26.

Journel, A.G., Huijbregts, Ch.J. 1978. Mining Geostatistics, Academic Press, London, 600 pp.

Kimberley, M.O., West, G., Dean, M., Knowles, L., 2005. Site Productivity: The 300 Index - a volume productivity index for radiata pine. New Zealand Journal of Forestry 50, 13-18.

Lagacherie, P., McBratney, A., Voltz, M., (eds) 2006. Digital Soil Mapping: An Introductory Perspective. Developments in Soil Science 31, Elsevier, Amsterdam, 350pp.

Landcare Research, 2008. http://www.landcareresearch.co.nz.

Leathwick, J.R., 1995. Climatic relationships of some New Zealand forest tree species. Journal of Vegetation Science 6, 237-248.

Leathwick, J.R., 1998. Are New Zealand's *Northofagus* species in equilibrium with their environment? Journal of Vegetation Science 9, 719-732.

Leathwick, J.R., 2001. New Zealand's potential forest pattern as predicted from current species-environment relationships. New Zealand Journal of Botany 39, 447-464.

Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K., 2002b. Land Environments of New Zealand: A technical guide. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 237 pp.

Leathwick, J.R. Whitehead, D. McLeod, M., 1996. Predicting changes in the composition of New Zealand's indigenous forests in response to global warming: a modelling approach. Environmental Software 11, 81-90.

Leathwick, J., Wilson, G., Rutledge, D., Wardle, P., Morgan, F., Johnston, K., McLeod, M., Kirkpatrick, R., 2003. Land Environments of New Zealand. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 184 pp.

Leathwick, J.R., Wilson, G., Stephens, R.T.T., 2002a. Climate surfaces for New Zealand. Landcare Research contract report: LC9798/126. Landcare Research, Hamilton, New Zealand, 22 pp.

McMurtrie, R.E., Rook, D.A., Kelliher, F.M., 1990. Modelling the yield of *Pinus radiata* on a site limited by water and nitrogen. Forest Ecology and Management 30, 381-413.

Mulla, D.J., McBratney, A.B. 2000. Soil spatial variability. In: Sumner, M.E. (editor-in-chief), Handbook of Soil Science. CRC Press, Boca Raton, A321-A352.

Mummery, D., Battaglia, M., Beadle, C. L., Turnbull, C. R. A., McLeod, R. 1999. An application of terrain and environmental modelling in a large-scale forestry experiment. Forest Ecology and Management 118, 149-159.

Nambiar, E.K.S., Sands, R., 1993. Competition for water and nutrients in forests. Canadian Journal of Forest Research 23, 1955-1968.

Newsome, P.F.J., Wilde, R.H., Willoughby, E.J., 2000. Land resource information system spatial data layers. Landcare Research, Palmerston North, New Zealand, 84 pp.

Odeh I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215-226.

Palmer, D.J., Lowe, D.J., Payn, T.W., Höck, B.K., McLay C.D.A., Kimberley, M.O., 2005. Soil and foliar phosphorus as indicators of sustainability for *Pinus radiata* plantation forestry in New Zealand. Forest Ecology and Management 220, 140-154.

Payn, T.W., Hill, R.B., Höck, B.K., Skinner, M.F., Thorn, A.J., Rijkse, W.C., 1999. Potential for the use of GIS and spatial analysis techniques as tools for monitoring changes in forest productivity and nutrition, a New Zealand example. Forest Ecology and Management 122, 187-196.

Payn, T.W., Thwaites, R.N., 1996. Classification and analysis of forest sites for productivity potential. Proceedings New Zealand Institute of Forestry Conference, April 29-May 1st Invercargill, pp. 43–48.

Richardson, B., Whitehead, D., McCracken, I.J., 2002. Root-zone water storage and growth of *Pinus radiata* in the presence of a broom understorey. New Zealand Journal of Forestry Science 32, 208-220.

Rodgerson, P.A., 2006. Statistical Methods for Geography: A Student's Guide. SAGE Publications, London. 304 pp.

Tait, A., Henderson, R., Turner, R., Zheng, X., 2006. Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26, 2097-2115.

Tegelmark, D.O., 1998. Site factors as multivariate predictors of the success of natural regeneration in Scots pine forests. Forest Ecology Management 109, 231-239.

Thompson, J.A.; Bell, J.C.; Butler, C.A. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma 100, 67-89.

Watt, M.S., Coker, G., Clinton, P.W., Davis, M.R., Parfitt, R., Simcock, R., Garrett, L., Payn, T., Richardson, B., Dunningham, A., 2005. Defining sustainability of plantation forests through identification of site quality indicators influencing productivity - A national view for New Zealand. Forest Ecology and Management 216, 51-63.

Watt, M.S., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C., 2003. Modelling the influence of weed competition on the growth of young *Pinus radiata* at a dryland site. Forest Ecology and Management 178, 271-286.

Webster, R., Oliver, M.A. 1990. Statistical Methods in Soil and Land Resource Survey. Oxford University Press, Oxford. 316 pp.

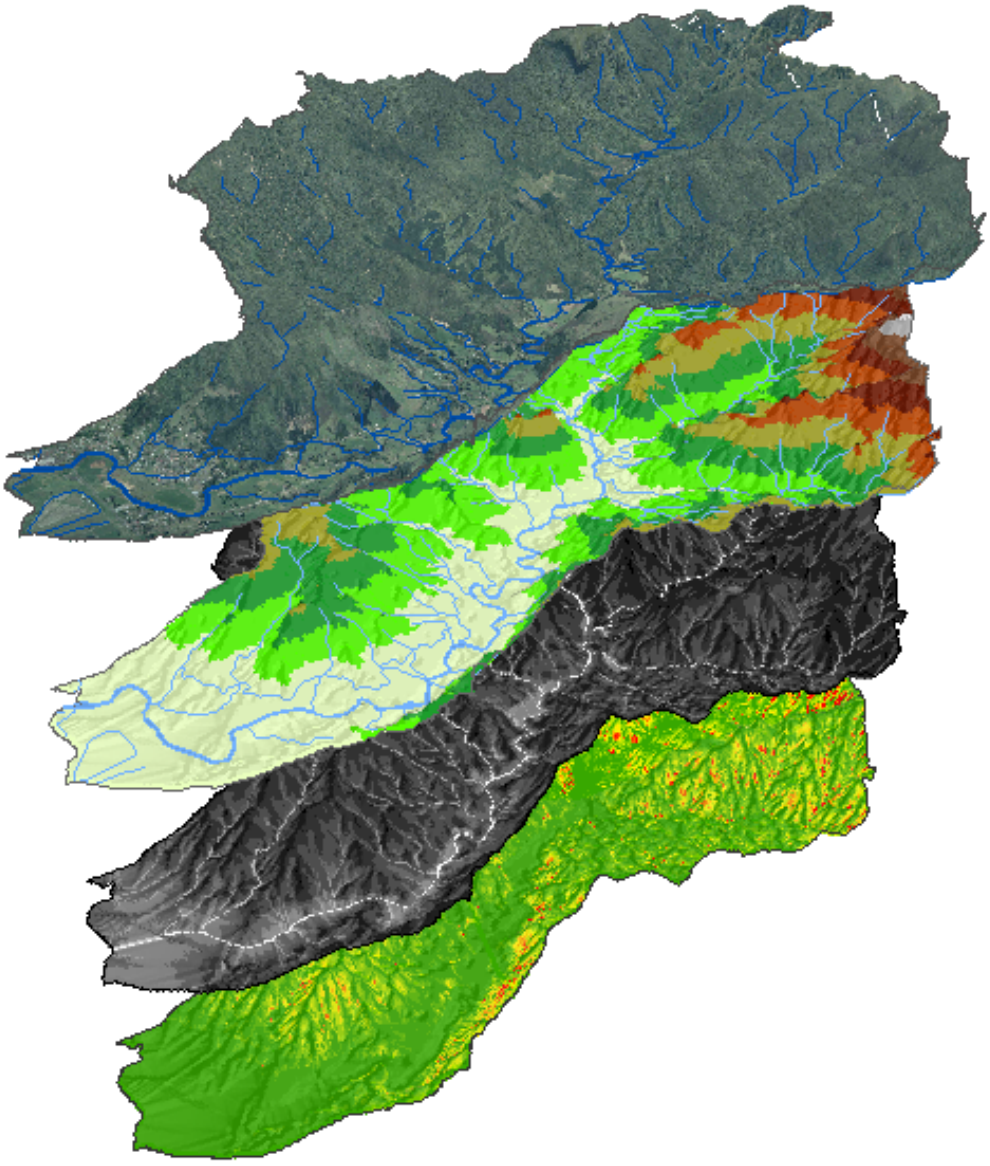
Webster, R., Oliver, M.A. 2007. Geostatistics for Environmental Sciences. Second edition, Wiley and Sons, 317 pp.

Wilson, J.P., Gallant, J.C. 2000. Terrain Analysis: Principles and Applications. New York, John Wiley and Sons.

Woollons, R.C., Skinner, M.F., Richardson, M.F., Rijske, W.C., 2002. Utility of "A" horizon soil characteristics to separate pedological groupings, and their influence with climatic and topographic variables on *Pinus radiata* height growth. New Zealand Journal of Forestry Science 32, 195-207.

"… after climbing a great hill, one only finds that there are many more hills to climb. I have taken a moment here to rest, to steal a view of the glorious vista that surrounds me. To look back on the distance I have come. But I can rest only for a moment, for with freedom come responsibilities, and I dare not linger, for my long walk is not yet ended."

Nelson Mandela - *Long Walk to Freedom*



Three dimensional graphic of vegetation (orthophotograph with waterways, upper graphic), elevation (with waterways), topographic wetness index, and the length and slope factor (lowest graphic) illustrating these properties across the Coromandel macro-catchment. All graphics were draped over a 25 m resolution DEM displayed with a 1.5 vertical exaggeration. Hillshading was used to highlight topographical features excluding the vegetation graphic.

CHAPTER TWO

DEVELOPING NATIONAL-SCALE TERRAIN ATTRIBUTES FOR NEW ZEALAND (TANZ)

D.J. Palmer[a*], B.K. Höck[b], D.J. Lowe[a], A.G. Dunningham[b], T.W. Payn[b]

[a]*Department of Earth and Ocean Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand 3240*

[b]*Scion, Private Bag 3020, Rotorua, New Zealand*

[*]Corresponding author. Tel.: +64-7-8562889 *E-mail address*: djp8@waikato.ac.nz

Published: Scion Bulletin, 2008

ABSTRACT

Primary and secondary terrain attributes provide information essential for environmental modelling in a wide range of disciplines. Although topographical attributes are frequently developed and used in New Zealand, currently these attributes remain piecemeal and site specific. In this paper we define the parameters and protocols used in developing national-scale terrain attributes for New Zealand (TANZ). By 'national scale' we mean digital maps or surfaces extending over the entire national territory of New Zealand. The software TAPES-G (Terrain Analysis Programs for the Environmental Sciences - Grid) was used to develop primary terrain attributes at a 25 m resolution. The GIS platform ArcInfo, in conjunction with Arc Macro Language (AML), was used to extract source DEM data for large catchments across New Zealand using a macro-catchment coverage vector surface. Macro-catchments serve two purposes: (1) to ensure array size parameters for TAPES-G are not exceeded; and (2) the macro-catchment parameter acts as a boundary for TAPES-G processing. AMLs automate the development of terrain attributes and the subsequent merging of raster data into a series of national topographic attributes. In developing TANZ we set out to create high-quality surfaces useful for productivity modelling and spatial interpolation of *Pinus radiata* data across New Zealand. TANZ surfaces will be additionally useful in other modelling environments including hydrology, geomorphology, pedology, and ecology with potential applications in resource management, environmental modelling, GIS, pedometrics, and geostatistics. This bulletin provides comprehensive technical information covering the protocols and processes used to develop TANZ.

CONTENTS

1. INTRODUCTION

1.1 Project rationale

The shape of the Earth's surface influences surface water flow and the removal and transfer of sediments, solutes and nutrients. Landform also influences the amount of solar energy reaching the Earth's surface, with impacts on the distribution and nature of organisms (Blaszcynski, 1997). Therefore, the development of digital terrain-based attributes provides information useful in many disciplines involving environmental modelling.

With the increase of high speed computer processors and the increased digital storage capacity of modern computer systems, along with access to finer resolution data sets, national-scale modelling using large datasets is now possible. Automated digital terrain analysis in conjunction with geographical information systems (GIS) have greatly facilitated the capture and characterisation of properties and processes of the land, increasing the understanding of hydrological, geomorphological, pedological, and ecological processes (Moore et al., 1991). Although high-quality gridded (raster) surfaces covering New Zealand have been derived for climatic and other environmental properties (Leathwick et al., 2002, 2003a, 2003b), comprehensive terrain attributes at a national scale, other than primary attributes such as elevation and slope, have not yet been developed.

Notwithstanding the ambiguities surrounding the definition of the word 'scale' in a mapping context (Bian, 1997; Hewitt and Lilburne, 2003; Lilburne et al., 2004), this document defines 'national scale' as digital terrain maps or surfaces extending over the entire national territory of New Zealand, which is in keeping with the term 'continental scale' when modelling or mapping across continents (Henderson et al., 2001; Wynn et al., 2006). In this paper we set out to define the protocols and programming processes involved in developing fine-resolution terrain attributes at a national scale: Terrain Attributes for New Zealand (TANZ).

1.2 New Zealand geomorphology

The three main islands of the New Zealand archipelago, with a total area of ~260,000 km^2 (Table 1), span 13 degrees of the temperate mid-latitudes of the South Pacific Ocean (Figure 1) and are bisected by an active, obliquely converging boundary between the Australian and Pacific lithospheric plates (Kamp, 1992; Pettinga, 2001). Consequently, tectonism and volcanism, together with climate, are major controls on the evolution of the present-day landscape. Because tectonic processes have operated in concert with extreme climate fluctuations during the Quaternary, and because of generally high precipitation rates (including frequent rainstorms) and very fast rates of denudation (landsurface lowering), the resulting landscape is characterised by strong relief marked by diversity, complexity, and youthfulness (Molloy and Christie, 1998; Newnham et al., 1999; Fitzsimons, 2001; Leathwick et al., 2003a, 2003b).

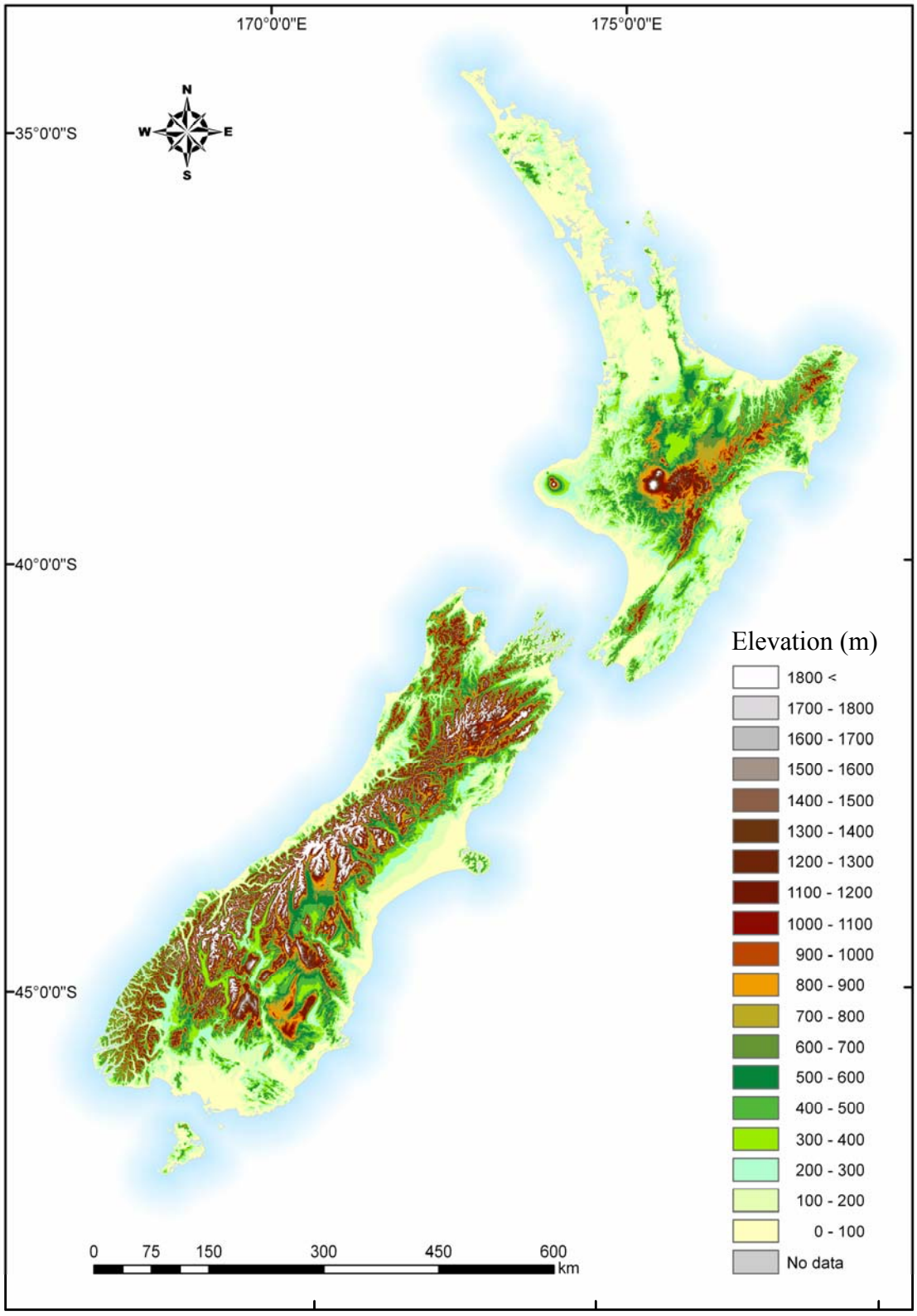


Figure 1: Digital elevation map of New Zealand (elevation in metres above sea-level) (data derived from Barringer et al., 2002).

The two main islands are divided into eastern and western sections by a central mountain axis that extends from Fiordland in the South Island to East Cape in the North Island. The Southern Alps form a chain of mountains in the South Island high enough to support glaciers and to form an orographic barrier to prevailing westerlies. The highest peak is Mt Cook (known also as Aoraki or Aorangi) at 3754 m. The central North Island is characterised by a variety of volcanic features and products including large calderas (e.g. Lake Taupo), stratovolcanoes (e.g. Mts Ruapehu, 2797 m; Taranaki, 2518 m), lava domes, and extensive plateaux of welded to nonwelded pyroclastic flow and tephra-fall deposits (Neall, 2001; Lowe and Palmer, 2005; Smith et al., 2006). More than 66% of the land in New Zealand is above 200 m elevation: about 25% is steep to very steep (slopes > 26°), about 20% is hilly (slopes 16-25°), and about 55% is flat, undulating, and rolling (slopes 0-15°) (see Table 1 for details).

1.3 Project outline

We describe the technical details of the methods used to create primary and secondary terrain attributes for New Zealand. First, we present a general overview of digital elevation modelling (Section 2). Then we discuss methods and problems associated with terrain modelling (Section 3). Next, we describe and discuss the methods used in developing TANZ (Section 4). Then we present the TAPES-G model (Section 5) and discuss algorithms used in the development of the attributes (Sections 6-8). Finally, we

provide examples illustrating each of the terrain attributes created for New Zealand (Section 9), and summarise our main findings and conclusions (Section 10).

More specifically, Section 2 assesses methods used to spatially capture elevation data and we discuss some of the problems associated with data source and quality, and with choosing an appropriate spatial scale and digital elevation model (DEM) resolution. The representation of terrain and terrain attributes are reviewed in Section 3 with discussion on how these are influenced by topography. In this section we also discuss the 'macro-catchment concept' developed to automate DEM-data extraction across New Zealand. The TAPES-G installation and technical requirements, and input parameters and data pre-processing requirements, are discussed in Section 4. Sections 5 and 6 examine the technical side of terrain-attribute development using the TAPES-G model. In Section 8 we discuss the creation of national-scale raster surfaces and how individual macro-catchments were merged into national-scale surfaces. The resultant outputs of terrain attributes created for New Zealand are presented in Section 9, with examples illustrating a representative macro-catchment (on Coromandel Peninsula, eastern North Island) displayed two dimensionally. A subcatchment within the larger macro-catchment was the focus of more detailed, cell-by-cell analysis through the use of three-dimensional graphics. These graphics were designed to illustrate the spatial patterns occurring within a landscape using primary and secondary terrain attributes.

2. DIGITAL ELEVATION MODELLING

2.1 The representation of elevation

A digital elevation model (DEM) has been described as "an ordered array of numbers that represents the spatial distribution of elevations above some arbitrary datum in a landscape" (Moore et al., 1991) – in other words, the digital representation of altitude. Typically, digital elevation data are collected from various sources including satellite imagery, aerial photographs, or from ground surveys such as global positioning systems (GPS). All these types of elevation data are used to capture and store information for terrain modelling (Wilson and Gallant, 2000). Primary terrain attributes are often derived directly from a DEM, whereas secondary terrain attributes are calculated from the primary terrain attributes. The DEM forms the basis or building blocks for all terrain attributes discussed in this paper, and therefore a brief discussion covering issues of DEM development and quality is included.

2.2 GIS data structures for modelling elevation

There are three main methods to structure digital elevation data:

1) point elevation data on a regular grid;

2) point elevation data in a triangulated irregular network;

3) digital contour lines.

Figure 2 illustrates these approaches. Regular grids (Figure 2a) have digital information arranged in rows and columns with each grid point representing the elevation at that location. The main advantage of regular grids is found in the simple data structure, making data processing, storage and visualisation relatively straightforward. Regular DEMs, however, are sometimes criticised because of their inability to handle sharp or abrupt changes in elevation readily. A smaller grid size can reduce this concern, but this increases computation times and storage requirements. Issues relating to grid size are discussed in greater detail in the next section. Another issue highlighted by Wilson and Gallant (2000) is the undesirable zigzagging pattern of upslope flow paths formed when calculated from grids, though there are processes to assist with this, as discussed later in the paper.

A second structure used to store elevation data is the triangulated irregular networks (TINs) (Figure 2b). The basis of TINs form a series of points containing x, y, and z coordinates preferably representing peaks, ridges, and breaks in slope. Networks are constructed based on triangular elements, where each triangle consists of a plane joining three adjacent points, or facets with vertices at the sample points (Moore et al., 1991; Wilson and Gallant, 2000). TINs have the advantage of modelling sharp features such as ridges or peaks well (Wilson and Gallant, 2000), and realistically reflect surface roughness by appropriate sample point intensity. Furthermore, TINs have the advantage of requiring less storage space when compared with the regular grid equivalent. On the negative side, calculating terrain attributes is not straightforward because of the nature of

TINs and the difficulty in determining the upslope connection of a facet (Moore et al., 1993b).

Finally, a contour-based network structure (Figure 2c) consists of numerous small and irregular shaped polygons bounded by adjacent contour lines and streamlines (lines drawn orthogonal to the contour lines). This structure is popular in hydrological applications because the flow of water can be reduced to a series of coupled one-dimensional equations in areas of complex terrain, hence reducing redundancy (Moore and Foster, 1990; Moore et al., 1991).

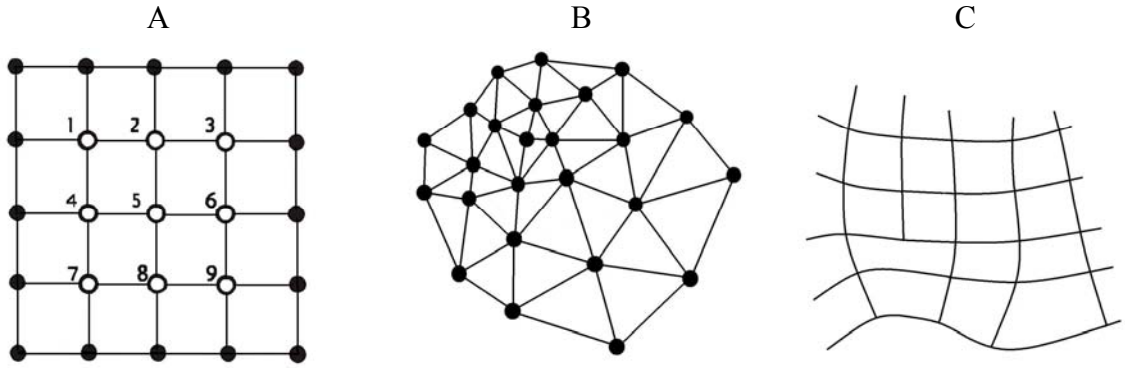


Figure 2: Methods used in structuring elevation data: (A) rectangular grids, (B) triangulated irregular networks, and C) network based on contour data (after Moore et al., 1991).

Our study used regular gridded data (Figure 2a), because (1) high quality DEM raster data are readily available for New Zealand, (2) DEM raster data can easily be converted to regular grid data, and (3) the terrain model we had selected, TAPES-G, requires input data to be formatted to regular grid point data. By using an adequate, relatively fine-grid resolution, the previously mentioned problem associated with abrupt changes in elevation is, to some extent, resolved. The visually and analytically undesirable zigzag patterns

referred to above can be minimised through the implementation of different algorithms (see Section 5.1.4).

2.3 Spatial scales and DEM resolution

The accuracy of a raster DEM is strongly affected by the spatial scale and quality of the source data. Even the most sophisticated algorithms for surface fitting will produce substandard models should the source data be inadequate (Wilson and Gallant, 1998).

When deriving a raster DEM from contour data, the relationship between the cell-size and the spatial density of the contour lines is important. There is no benefit gained by selecting a cell-size that is smaller than the smallest distance between the contour lines. Furthermore, care must be taken when interpolating a DEM at an excessively fine resolution relative to its source data to avoid the creation of artefacts. To illustrate this relationship, if 20 m contour lines (i.e. every 20 m change in elevation) are used to derive 5 m, 25 m and 50 m resolution DEMs (i.e. with a 5 m, 25 m and 50 m horizontal resolution, respectively), these DEMs will have different levels of accuracy locally. In a generally flat area, for example the Canterbury Plains, the 50 m DEM will be just as accurate as the 5 and 25 m DEMs. In very steep terrain, however, the accuracy of the 50 m DEM is likely to be low by comparison with the other DEMs. However, a DEM with a resolution less than the smallest distance between the contour lines will not enhance accuracy – the accuracy will remain equivalent to that of the original spatial scale.

Numerous papers discuss issues relating to grid resolution and its sensitivity on resultant terrain attributes (Florinsky, 1998; Wilson et al., 2000; Thompson et al., 2001; Claessens, 2005). Table 2 provides information that relates relative scale and DEM resolution to appropriate model applications.

Table 1: Land surface areas in km^2 and percent for classes of elevation above sea level and slope in degrees for the North Island, South Island, and New Zealand derived from a 25-m DEM (Barringer et al., 2002). Adjacent surrounding offshore islands were included, whereas lakes and major water bodies were excluded from the analysis using *Land Information New Zealand* (LINZ) 1:50,000 topographic data (NZTopo).

		North Island		South Island		New Zealand	
		km^2	%	km^2	%	km^2	%
Elevation (m)	Class						
	0 - 100	31215	27.7	20071	13.4	51287	19.6
	100 - 200	20492	18.2	15746	10.5	36238	13.8
	200 - 300	16447	14.6	13606	9.1	30052	11.5
	300 - 400	12314	10.9	12321	8.2	24635	9.4
	400 - 500	9134	8.1	11282	7.6	20415	7.8
	500 - 600	7200	6.4	10365	6.9	17566	6.7
	600 <	16039	14.2	66005	44.2	82044	31.3
Slope (degree) *							
Flat to undulating	0 - 3	36104	32.0	40190	26.9	76293	29.4
Undulating	3 - 7	12486	11.1	11400	7.6	23885	9.3
Rolling	7 - 15	22345	19.8	20233	13.5	42578	16.7
Hilly	15 - 25	22343	19.8	28875	19.3	51218	19.6
Steep	25 - 34	12081	10.7	22082	14.8	34164	12.7
Very steep	34 <	7482	6.6	26616	17.8	34099	12.2
Total		112841		149396		262237	

* Slope classes based on Leathwick et al. (2003b)

Table 2: Spatial scales and general applications for digital elevation models within New Zealand (adapted from Hutchinson and Gallant, 2000).

Scale	DEM resolution	Model applications
Locally fine scale	1-5 m	Spatial modelling of land-based operations such as forest harvesting, roading and logging planning
Fine toposcale	5-50 m	Spatial analysis of soil properties Spatial analysis of soil survey data Topographic aspect influences on solar radiation, evaporation, and vegetation patterns
Coarse toposcale	50-200 m	Subcatchment analysis for hydrological modelling Assessment of biodiversity
Mesoscale	200 m-5 km	Elevation dependant representations of climate (temperature, precipitation, solar radiation) Determination of national scale drainage divisions
Macroscale	5 km-500 km	Major orographic barriers for general circulation models

Note: Coarse resolution DEMs are often derived by averaging finer scale DEM grids

2.4 DEM for New Zealand: source and quality

The source data for the raster DEM used in this paper were 20 m contours, spot heights, lake shoreline elevations and coastline information from *Land Information New Zealand* (LINZ) 1:50,000 vector topographic data (NZTopo). Technical information relating to NZTopo (select the "technical information about NZTopo") is available at the following website:

<http://www.linz.govt.nz/core/topography/topographicdata/topodtabase/index.html>.

For New Zealand, a 25 m DEM is a readily available resolution, derived from the LINZ source data. The 25 m resolution DEM used in this study fits into the fine toposcale category. There were two main reasons for choosing this scale for modelling our terrain attributes at the national level.

(1) Our future modelling applications include spatial analysis and predictive modelling of soil properties and soil survey data which require a resolution finer than 50 m.
(2) Even the current capabilities of the latest high-speed computer processors restrict or limit the volume of data that can be processed over any given period of time. Furthermore, storage requirements (workspace, swap space, and storage space) of data can restrict modelling capabilities.

Overall, the choice in using a 25 m DEM to model terrain attributes was a trade off between computer processor and storage capabilities and the need for fine-resolution data when potentially undertaking simulations for erosion models or predicting the spatial variability of a soil property on a relatively flat land surface. For our modelling purposes, future applications of these derived terrain attributes will include prediction of soil properties to growth models and productivity surfaces for *Pinus radiata* using spatial interpolation techniques such as geostatistics and regression analysis. These surfaces, along with their associated prediction error, will also feed directly into other future modelling applications.

Landcare Research developed a national 25 m DEM using LINZ data (Figure 1). This DEM was created using the procedure described by Barringer et al. (2002), which is considered to have an advanced interpolation method when compared with the previous methods used in national-extent 25 m DEM interpolation. A validation process was also undertaken for this floating point DEM using differential GPS data collected from 25 m postings on the Port Hills near Christchurch. In addition, a larger DEM validation project was undertaken in the lower Waitaki/Kurow area using filtered LIDAR data. For details refer to Barringer et al. (2002).

Sinks are frequently encountered artefacts in some surface-fitting methods. They are potentially problematic because they represent a terminal point in hydrological flow paths. Spurious sinks require correction and removal to create a hydrologically sound surface representative of the real land surface. Valid sinks in New Zealand include those in karst landscapes, and depressional wetlands and lakes where sinks occur naturally. When smoothing data to remove spurious depressions, consideration has to be given to the effect that such smoothing has on the overall accuracy of data. The disadvantage of altering surface elevations is that the uncertainty in the modelling environment is increased. Identification of real-world sinks was possible through the LINZ topodata contour depressions and sinkhole vector datasets. These are the same series of data used in Landcare Research's development of the national-extent 25 m floating point DEM and thereby have an excellent spatial correspondence with the DEM. Locations of natural sinks and their surrounding topography were identified for retaining while all spurious

unwanted sinks were removed using standard sink-removal techniques. For detailed discussion covering sink retention or removal refer to Section 4.4.

3. DIGITAL MODELLING OF TERRAIN

3.1 Digital terrain modelling

Elevation is a measure of height above an arbitrary datum in the landscape (Moore et al., 1991), and in this modelling environment each 25 m by 25 m cell represents the number of metres above sea level. In the development of TANZ the TAPES-G programme used a 25 m floating point DEM for modelling all primary attributes (Section 5.1). The DEM has a New Zealand map grid projection with a New Zealand geodetic datum NZGD1949. Elevation is commonly used in modelling environments because it directly impacts on temperatures as measured by lapse rates and indirectly through the influence of slope, aspect and overshadowing on solar radiation regimes. Elevation also influences precipitation through orographic effects. For example, the Southern Alps in the South Island form an orographic barrier with western regions intercepting and receiving much of the rainfall during prevailing westerly winds leaving eastern regions comparatively dry. Local changes in elevation define slope and aspect, which in turn characterise the movement of water across the Earth's surface, impacting on erosion, deposition and other processes (Moore and Wilson, 1992; Krysanova et al., 2000), pedogenesis, topsoil thickness, and other soil properties including root volume and nutrient status (Moore et al., 1993a; Gessler et al., 2000). Therefore the ability to describe topography is

fundamental to the prediction of surface flow and velocity, and other related characteristics including soil moisture regimes and stream flows (Moore et al., 1988; Sulebak et al., 2000). Other examples include papers by, Dikau (1989), Ventura and Irvin (2000), and Burrough et al. (2001), which used topographic-based attributes to delineate contiguous spatial units into elements such as ridgelines, shoulder slopes, backslopes, footslopes, and valley bottoms, emphasising how mapping landforms can be an effective method for capturing information that relates topography to landform processes.

Digital terrain models (DTMs) are models providing digital representation of the shape in the landscape, which is defined by local changes in elevation. Typical examples include slope, curvature, aspect, and water flow direction. DTMs are widely used in modelling soil erosion (e.g. De Roo, 1998; Rodda et al., 2001), hydrological modelling (Moore et al., 1991; Romano and Palladino, 2002), soil survey and land resource assessment (McKenzie, et al., 2000; Thompson et al., 2001), environmental modelling (Mummery et al., 1999; Mackey et al., 2000), and for predicting soil and regolith properties (Gessler et al., 1995; Florinsky et al., 2002). Because elevation is an attribute of a terrain, the DTM model includes the relevant DEM of the area. It is common for attributes such as slope to be derived from the DEM, as we have done in this paper.

3.2 The macro-catchment concept

Computer limitations required the New Zealand DEM to be divided into smaller, more manageable areas. Such subdivision, however, can create edge effects along the boundaries of terrain data where the edge of the DEM does not coincide with an actual catchment boundary, or when discontinuities are introduced into a continuous surface. The former problem in particular has an impact on the quality of modelled terrain attributes such as upslope contributing area (Gallant and Wilson, 2000).

Macro-catchments were developed to cope with these and other computer limitations noted above (Section 2.4), as well as avoiding the introduction of edge effects. Much of the research surrounding terrain attributes focuses on small to medium-size catchment scales when using fine-resolution gridded data (Table 3). The macro-catchment concept was designed to divide New Zealand into a series of large, naturally draining areas that consist of numerous catchments and sub-catchments representing the natural drainage patterns for that location (Figure 3) while not exceeding maximum processing capabilities. The maximum size of the macro-catchments was chosen so that the computer processor and storage system was able to cope with the magnitude of each data set (the number of rows and columns of each grid).

Table 3: Research studies using terrain attributes, their DEM resolution, catchment size, total relief, and main focus.

Research paper	Main focus	DEM (m)	Total relief (m)	Size of area studied (km^2)
Florinsky et al. (2002)	Soil property prediction	15	6	0.66
Franklin et al. (2000)	Predictive vegetation mapping	30	1700	~ 80
Fried, et al. (2000)	Storm-water discharge investigation	10	-	17
Gessler et al. (2000)	Modelling soil-landscape and ecosystem properties	2, 4, 6, 8, & 10	33	0.02
Gessler et al. (1995)	Prediction of soil attributes	20	-	100
Krysanova et al. (2000)	Soil-moisture deficit and potential soil loss	1000	1161	128,000.00
Mackey et al. (2000)	Modelling forest ecosystems	20	200	900
McKenzie et al. (2000)	Predictive soil survey and pedological applications	25	-	5000
Mummery et al. (1999)	Environmental forestry modelling	25	60	0.83
Thompson et al. (2001)	Soil-landscape modelling	10, & 20	20	0.13
Wheatley et al. (2000)	Vegetation mapping	30	-	~ 900
Wilson et al. (2000)	Sensitivity analysis of grid resolution and flow-routing algorithms on terrain attributes	30, 100, 200	1495	105
This study	TANZ: National-scale modelling of terrain attributes for New Zealand	25	3754	270,000 *

* Approximate total area of North Island and South Island of New Zealand

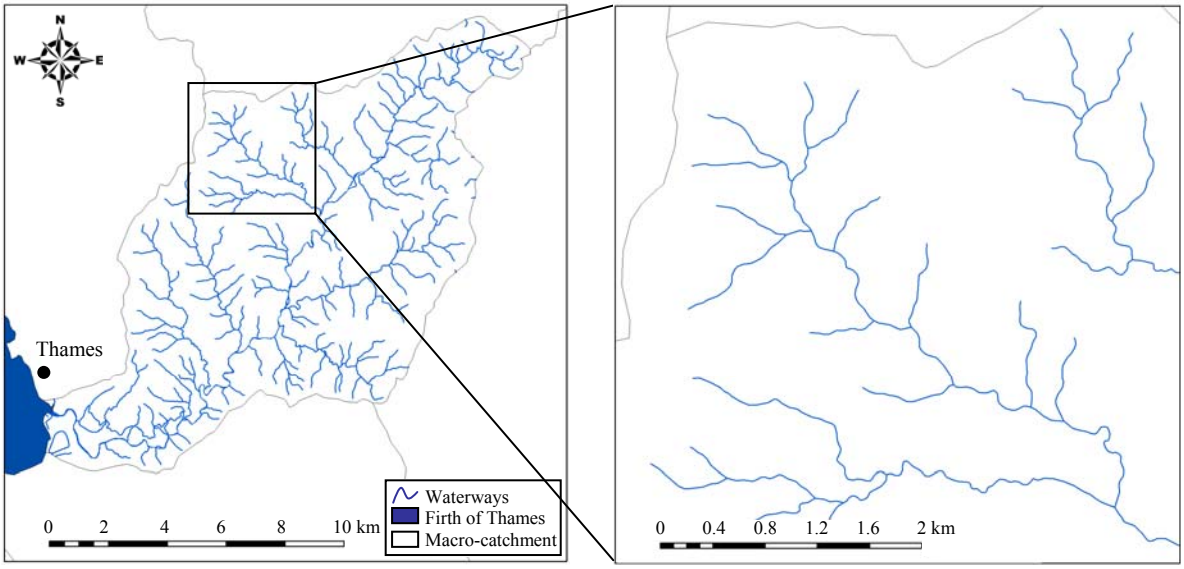


Figure 3: Maps showing river and stream network within a Coromandel Peninsula macro-catchment (left), and one of its many constituent sub-catchments. Location of macro-catchment is shown in Fig. 4.

To develop the macro-catchment boundaries, an existing NIWA vector catchment boundary map (Ude Shankar, NIWA, Christchurch, pers. comm., May 2003) was upgraded and adapted. Firstly, boundaries between adjacent catchments were checked that they corresponded with the local highest DEM values, i.e. that they matched the DEM-based catchment boundaries. Because the DEM was to be the basis for the terrain analysis, where necessary, catchment boundaries were modified to match the DEM-indicated extents. Along the coastline, the catchment boundaries were updated to the latest coastline from the NZTopo dataset. Finally, larger catchments were divided into smaller catchments to ensure that the maximum preset size limits in the modelling process were not exceeded. Small island catchments were sometimes grouped with larger mainland macro-catchments.

This maximum size of macro-catchment was set to 100 km by 100 km for reasons described in Section 4.2. To exclude the analysis of numerous smaller coastal islands, which were considered unsuitable for forestry purposes, all coastal islands smaller than 100 ha were removed from the analysis. The resultant vector map for the North Island with surrounding (large) islands contains 392 macro-catchments (Figure 4), and 245 macro-catchments for the South Island and its surrounding islands (Figure 5).

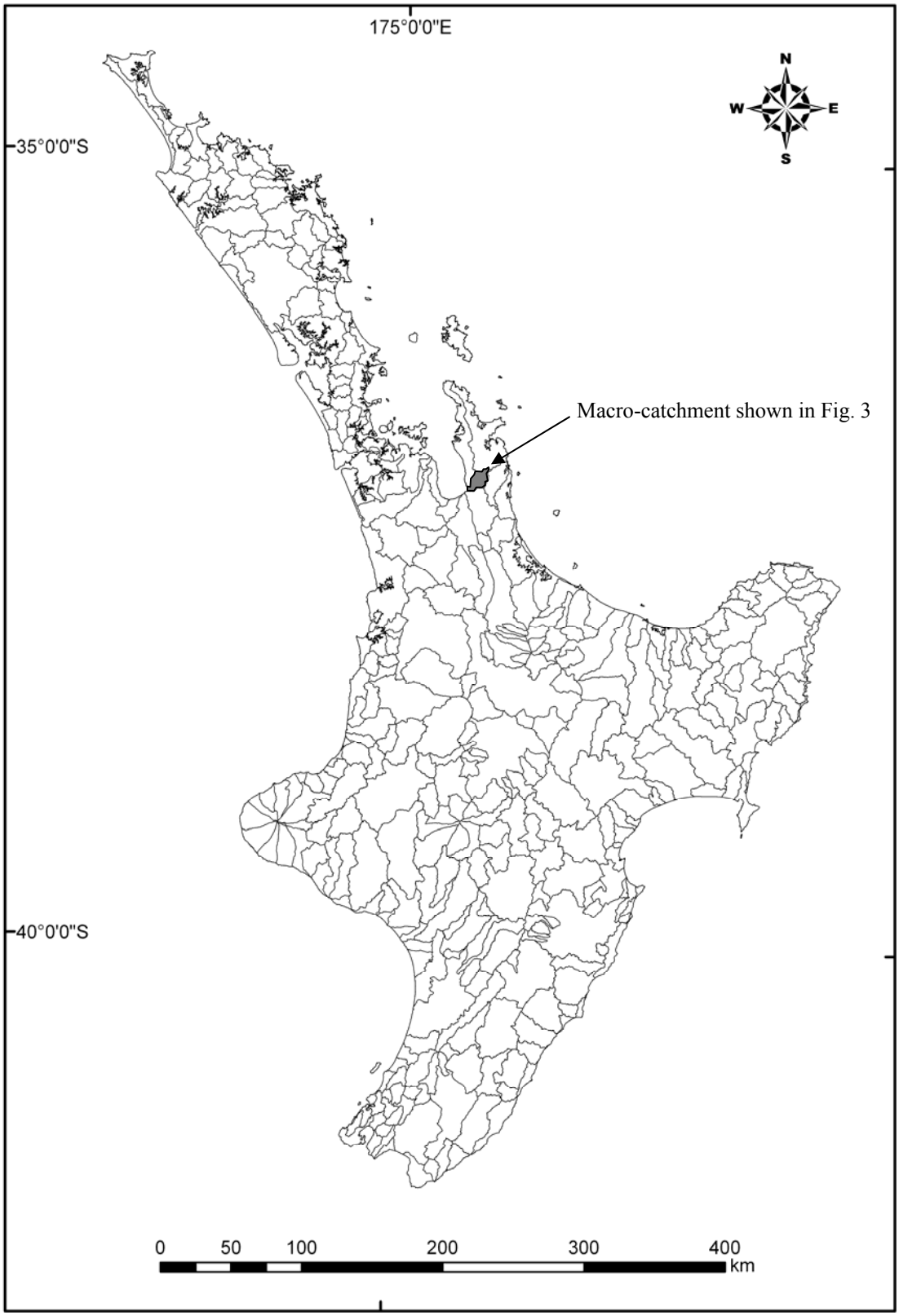


Figure 4: Locations of 392 North Island macro-catchments.

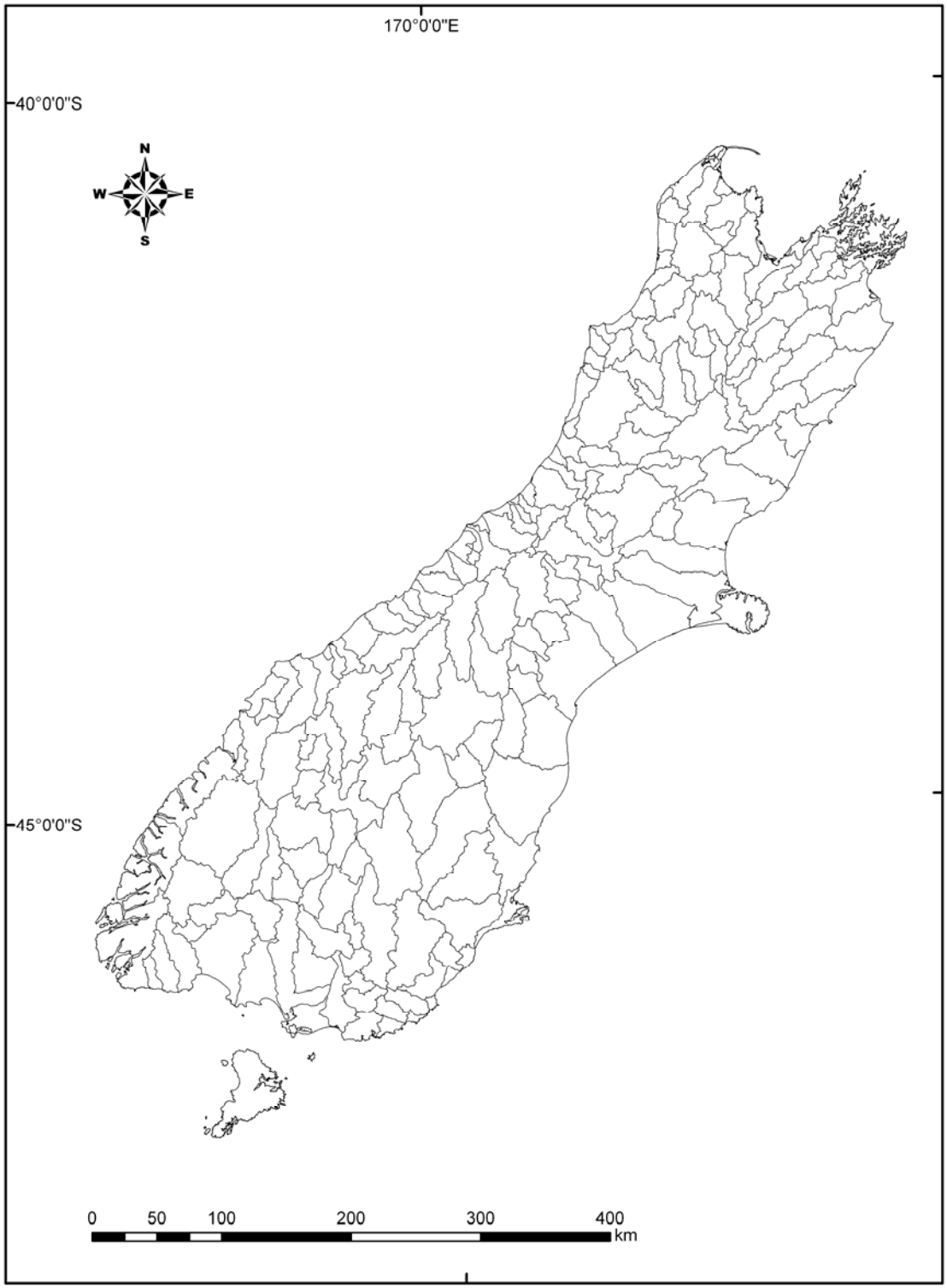


Figure 5: Locations of 245 South Island macro-catchments.

3.2.1 Macro-catchment delineation

With macro-catchment boundaries defined, a script was written in Arc Macro Language (AML) to extract each catchment polygon. The script uses the macro-catchment's unique polygon identification number to extract individual catchment polygons into individual data layers (coverages). Each polygon boundary was then used to create two raster masks (Figure 6). The first mask is ten grid cells (i.e. 250 m) larger than the macro-catchment. This cell size allowed the terrain-based modelling to process up to and over the catchment boundary in order to avoid any of the discontinuities mentioned earlier. Secondly, a mask equivalent to the macro-catchment boundary was created to ensure masking of individual catchment data back to the correct size in order to merge resulting terrain attribute data into a single comprehensive national-scale surface again. The issues of the merging process are discussed later in detail in Section 8.1.

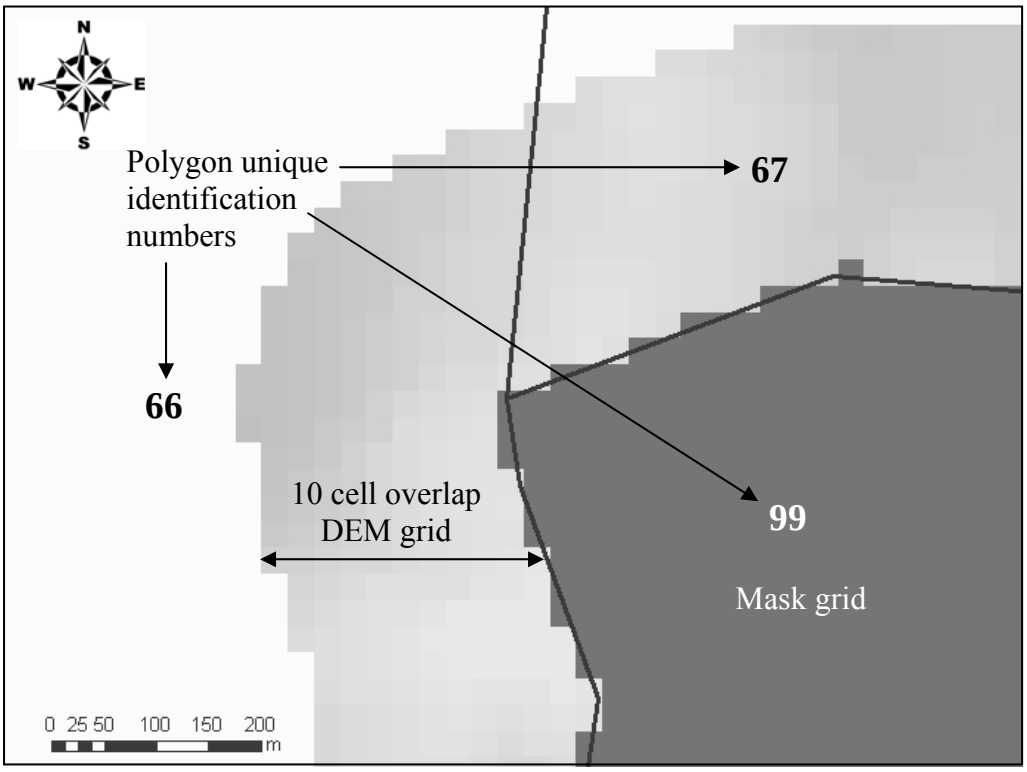


Figure 6: Graphic showing the boundaries between three macro-catchments (black line) and their associated unique identification numbers, the 10-cell overlap dem grid (light grey scale) to ensure that there are no discontinuities at the boundaries of the macro-catchment, and the mask grid (dark grey) designed to mask the catchment back to original size for optimal merging between each of the macro-catchments.

3.2.2 Automation of macro-catchment DEM extraction and pre-processing

Using the GIS platform ArcInfo, AMLs were written to automatically extract elevation data for individual macro-catchments using the extended macro-catchment mask. The automated extraction process was necessary to reduce user error and to ensure that our computer system could process data for not only the full extent of the macro-catchments, but also without producing any discontinuities at any part of the process.

The primary AMLs (do_catnorth.aml and do_catsouth.aml) contain iterative loops calling each macro-catchment's unique identification number incrementally from the base number until the respective final number for each of the North and South islands is reached (1 to 392 for the North Island and 1 to 246 for the South Island).

The primary AML calls a secondary AML named catnorth.aml or catsouth.aml (Figure 7). These routines identify the macro-catchment being processed in the polygon coverage through its unique identification number. The selected polygon is buffered by 9 m, converted into a raster layer and stored as a mask for later use when merging output grids (Section 8.2). A 9 m buffer ensures a one cell overlay or an exact match with adjacent macro-catchment boundaries in the polygon-to-grid conversion. The same polygon is buffered by 250 m and converted to raster, and is then used for extracting the appropriate subset of the 25 m resolution New Zealand raster DEM. Finally, the routine removes all values less than zero, replacing them with a small value greater than zero – see Section 4.5 for rationale. The output grid is given the number 1000 plus its unique identification number for the North Island grids or respectively 2000 plus the unique identification number for the South Island grids. The grids are now ready for conversion to any specific input format required for modelling.

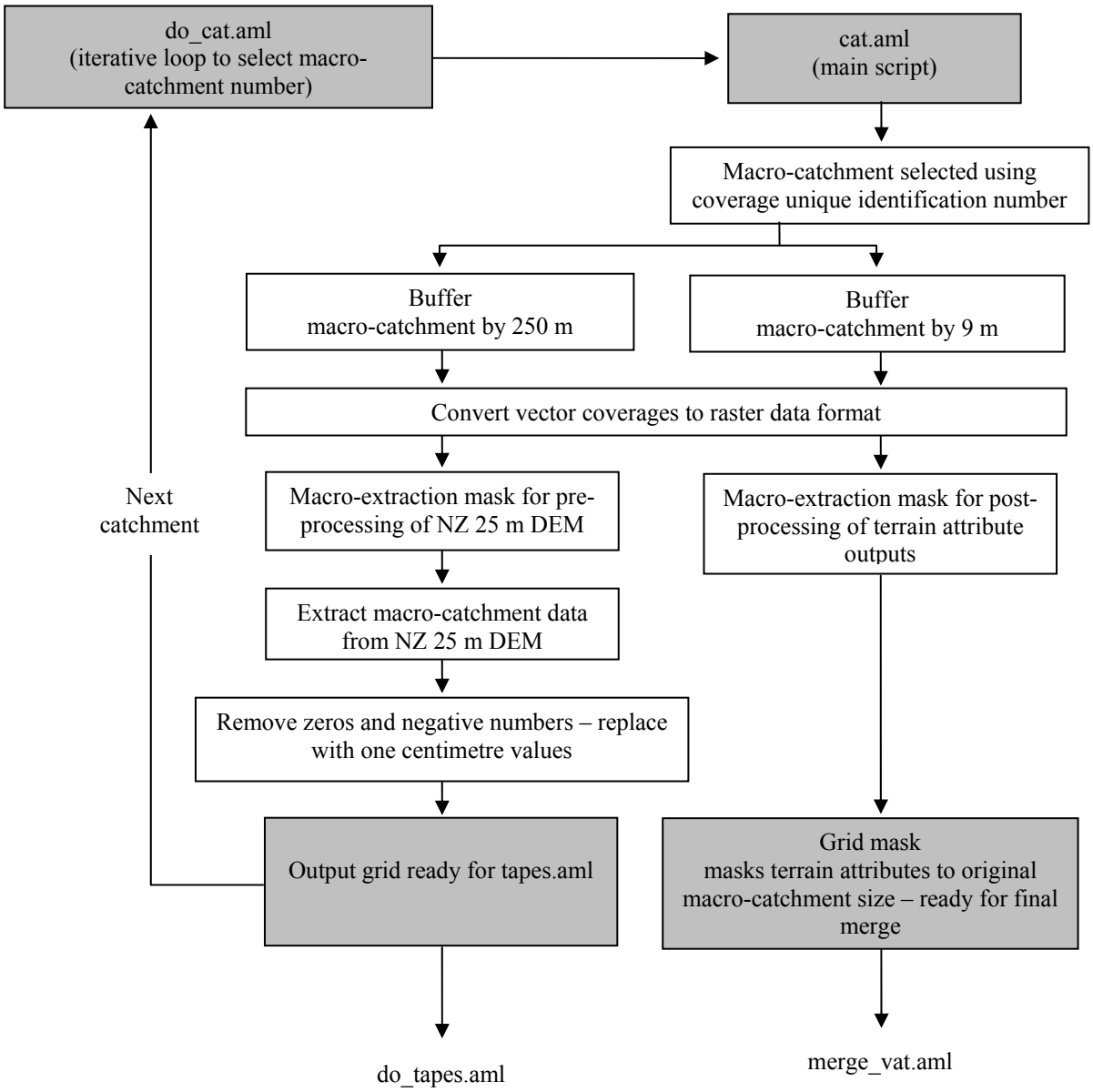


Figure 7: Flow chart illustrating main steps in the macro-catchment extraction from the New Zealand DEM.

4. TAPES-G: THE MODEL AND DATA REQUIREMENTS

4.1 TAPES-G and the modelling environment

Terrain Analysis Programs for the Environmental Sciences, TAPES, is a series of programmes designed to model terrain attributes. The TAPES programmes were written in Fortran-77 and C for Unix operating systems. When compiling code, array sizes can be set and easily adjusted to suit the requirements. The TAPES-G programme 'grid' requires up to 80 bytes of memory per grid cell. Thus a grid containing one million cells uses around eighty megabytes of virtual memory. A standard grid size often used is 1000 rows by 1000 columns. An option for reducing virtual memory requirements by up to half is to compile the programme without the DEMON flow accumulation algorithm if this is not required. Refer to the TAPES programmes for details (TAPES-G, 1995).

4.2 Installation of TAPES-G

The software TAPES-G, developed by Ian Moore, was downloaded from the Australian National University (ANU) website (http://cres.anu.edu.au) onto a Unix Silicon Graphics Workstation. Note that ANU no longer retains the TAPES-G software and further enquires should be made to John.Gallant@csiro.au.

The array size parameters LJ and LI within tapes_g.f were reset to accommodate the maximum size input DEM macro-catchment within our dataset. LJ refers to the number of columns, or width, and LI is the number of rows, or height, which was set to 4000 by 4000 cells, respectively. These settings result in a maximum extent of 100 km by 100 km (4000 rows by 4000 columns at a 25 m resolution).

In order to maximise processing speeds, TAPES-G was compiled without the DEMON module by editing the "makefile" lines: DEMON = demon.o and #DEMON = nodemon.o and changing them to #DEMON = demon.o and DEMON = nodemon.o. The command "make all" compiled the TAPES-G programme to fit the new parameters.

4.3 TAPES-G input parameters

The input parameters used in the TAPES-G programme remain constant for each macro-catchment (Table 4). Many of these parameters are straightforward, requiring little explanation. The mesh size overlay was set to 100 m, but has no influence on specific catchment area calculations and is retained as a background display from the TAPES-G programme. The critical slope-area used in the channel initiation was set to 5000 m^2, indicating that any flow path contributing area greater than 5000 m^2 would be considered part of the stream network or concentrated flow, whereas flow paths less than 5000 m^2 were considered unconcentrated flow. The difference between concentrated and

unconcentrated flow does not influence attribute calculations, however, and is only used in colour differentiation of concentrated versus unconcentrated flow (TAPES-G, 1995).

Table 4: TAPES-G input parameters.

TAPES-G inputs	Using FD8 flow-routing algorithm
Standard input/output	yes
Traditional attribute output or new, standard output	T, traditional output
Grid cell size	25
DEM file name	*file name*
Output file name	*file name*
Output file name in binary format	no
Dummy records for no data cells	yes
DEM input as	(3) file of real x, y, and z values
X - Y data entered as	(2) absolute length (real world) units
Restrict calculations within a catchment boundary	no
Mesh size	100
Artifical flow	no
Critical area-slope	5000
Value of exponent, R	2
Create depressionless DEM	yes
Drainage direction calculated	(1) D8 algorithm
Catchment area computation	(2) multiple drainage direction methods
Maximum cross grading area (m^2)	25,000
Output depressionless DEM	no
First record of output file	(1) north west
Output slope	(2) finite difference algorithm

Specific catchment area was calculated using a maximum cross-grading area threshold of 25,000 m^2. The rationale behind this threshold is given in Section 5.1.5. This threshold means that upland areas above defined channels (< 25,000 m^2) used the FD8 (multiple-flow direction method) algorithm allowing flow to be distributed to multiple nearest-neighbouring nodes, whereas below points of presumed channel initiation (> 25,000 m^2) the D8 algorithm (single-flow direction method) was activated. Finally, the finite difference slope method was used to calculate slope.

4.4 Sink identification: retain or remove?

In the real world, surface and near-surface water flows across the landscape, contributing towards stream and river flow, before finally discharging into the sea. However, there are exceptions in real landscapes where naturally closed basins with no outlet occur, commonly termed sinks. Digital representations of elevation, depending on the accuracy and quality of the data from which they were derived, may have sinks representing either the natural landscape, or spurious artefacts resulting from the surface generation process. Generally, we refer to sink features in a digital surface as pits if they comprise a single cell, and depressions if they consist of a group of cells. Examples are found in depressional wetlands, lakes, and karst environments. Sinks become problematic and undesirable during hydrological modelling because flow direction becomes undefined when a cell or group of cells is lower than all of the surrounding cells. Therefore, differentiating actual sinks from spurious sinks is an important DEM pre-processing

issue. Spurious sinks require eliminating (Jensen and Domingue, 1988), while actual sinks need to be maintained.

The identification of natural sinks across New Zealand was possible using an extract of the 20 m contour data and the sink point data. More specifically, depression contours are defined as "an imaginary line that connects points of equal height value… used to depict depressions in the Earth surface", and sinks as "a hole or funnel shaped cavity made in the earth by the action of water on the soil, rock or underlying strata (LINZ, 2007)", used to identify depressions too small to be shown by contours (pers. comm. LINZ, 2007). For details refer to: <http://www.linz.govt.nz>.

A national surface of all the sinks was created from these datasets. Depressions may have single or multiple contours; all were extracted from the 20 m contour dataset. Smaller depressions with areas less than 5000 m^2 were selected and converted to a single raster cell. The reason for applying a threshold was to avoid the loss of data during the polygon-to-raster conversion process where polygons smaller than a single cell are not recognised and thereby not created. The centre of the lowest location within each of the larger polygons was determined and converted to a single raster cell. Cell locations were checked and manually adjusted based on the following criteria:

(1) Where no stream, waterway, or lake is associated with a sink, identify the deepest cell within the sink (retains the sink's surrounding topography).

(2) Where a lake is within the bounds of a sink, choose the cell closest to the upper level of the sink within the lake, but below the sink's pour point.

(3) Where a stream or waterway passes continuously through a sink (sink is assumed to represent a depressional wetland environment), choose the cell closest to the upper levels of the sink, but below the pour point.

(4) Where two or more streams or waterways intersect within a sink, decide on the direction of water flow and act appropriately by identifying cells closest to stream or waterway intersections.

In flat areas where flow is not easily predicted, and which could potentially bypass the cell chosen from the above selection criteria, the flat sink area was identified and lowered into the surface by 25 cm, ensuring flow was retained within the sink.

LINZ sink point data were also converted to a single raster cell. The three datasets were then merged into one surface suitable for masking sinks across the national extent. This sink mask was used in the data pre-processing to identify and thereby retain real-world sinks across New Zealand. All other spurious sinks were subsequently removed in the modelling process.

4.5 DEM source data pre-processing

Prior to developing terrain attributes with the TAPES-G programme we needed to clarify the difference between grid cells and grid points. Our source data DEM contains grid

cells that are orientated in a north to south and east to west direction, referenced by a single coordinate pair representing the southwest (bottom left corner) of each cell. For our modelling purposes each of these cells has a fixed area of 25 m by 25 m or 625 m^2. The data structure for TAPES-G is different, requiring grid points, which are mathematical points containing no area information. Therefore, pre-processing of grid cells to grid points is required before executing the TAPES-G programme, as described below.

Other pre-processing requirements include dealing with zero elevation and no-data values. For example, terrain at sea level is represented by cell values of zero metres elevation, whereas a cell value of no data (in our DEMs, a cell value of -9999) indicates either that there is not enough reliable information at that spatial location, or that it is outside the area of interest. Conversely, in the TAPES-G environment, an elevation of zero indicates a missing value or point outside the border of the analysis area, equivalent to our no-data value. This issue requires conversion of zero elevations (which mainly occur at coastal locations on New Zealand's DEM) to a positive integer value and changes of no-data values to zero elevations. Hence zero elevation values were replaced with 0.01 metre above sea level (a.s.l.) for the 25 m floating point DEM.

Finally, TAPES-G cannot manage areas of depressions containing valid negative numbers – for example, at a below sea-level mine at Huntly in the North Island of New Zealand. These major depressions were dealt with concurrently at the time of zero elevation removal. A conditional statement is used in the AML routine that states: if the

input grid is less than or equal to zero then give it the value of 0.01 m, but if it is not less than or equal to zero then the value remains unchanged.

4.6 Automation of grid conversion and pre-processing

Grid cell conversion to grid-point data required for the TAPES-G programme was undertaken using purpose-written AMLs. The primary AML (do_tapes.aml) contains an iterative loop that calls each grid starting from a number predetermined in the macro-catchment AML (e.g. DEM1001, DEM1002, …. DEM1392), then adds one number with each iteration until reaching the final DEM number (Figure 8). The primary AML calls a secondary AML named tapes.aml. This script takes the macro-catchment extracted DEMs from the initial routine and converts the nodata values in each DEM to zero. The DEM grid cells are converted to point data (a point coverage), and then the x and y spatial coordinates are added for the points. Converting the point data from coverage to text format required the execution of a final read-and-write routine to remove header information and commas from the x and y coordinates and the z values, in order to provide the ".txt" files ready for input into the TAPES-G programme. We recognise that options exist that could speed up data transfer for different environments, but this approach suited our combination of computers.

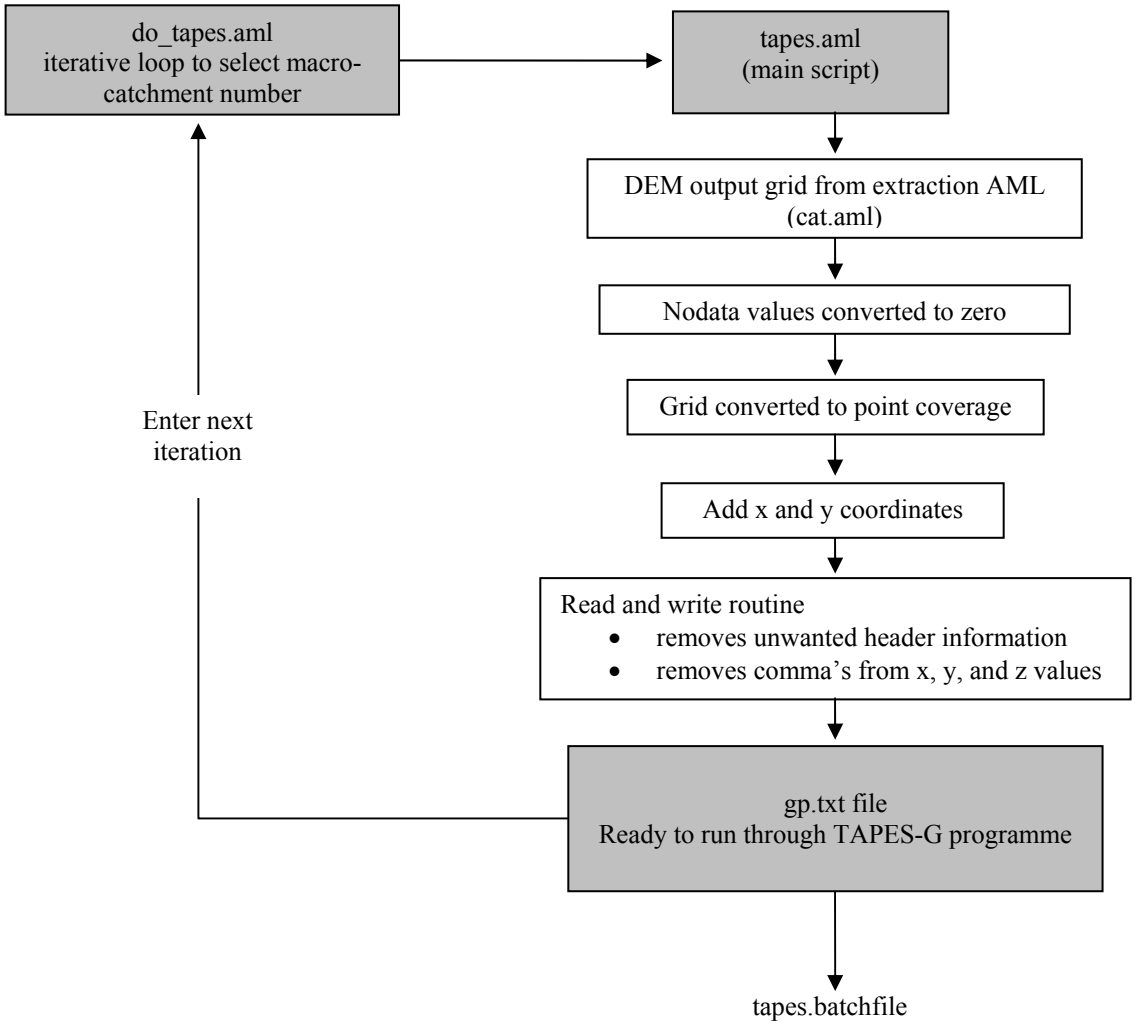


Figure 8: Flow chart illustrating main steps undertaken in the conversion of grid cell data to grid point.

5. TAPES-G: TERRAIN ATTRIBUTE COMPUTATIONS

5.1 Computation of primary terrain attributes

Primary topographic attributes derived using TAPES-G software are computed from directional derivatives of the topographic surface. They "measure the rate at which elevation changes in response to changes in location" (Gallant and Wilson, 2000), and include slope, aspect, plan and profile curvature, and upslope contributing area (Table 5).

The majority of the topographic attributes in Table 5 can be calculated locally from the derivatives of the topographic surface using centred finite differences (Gallant and Wilson, 1996). The first-order derivatives show the rate of change in elevation with a distance along the x or y axes (slope in those directions) as shown in Equations 1 and 2. Equations 3 and 4 are second-order derivatives that denote the rate of change of the first derivative in the x and y directions (curvature in those directions). Equation 5 describes the rate of change of the x derivative in the y direction and is a second derivative. Equations 6 and 7 are combinations of terms that are used in subsequent equations. Figure 9 illustrates the numbering convention of the nine points entering the finite difference equations, with h referring to the grid spacing of the DEM. At the edge of a DEM, or adjacent to no-data regions, forward and backward difference schemes are used (Gallant and Wilson, 1996).

Table 5: TANZ outputs and TAPES-G derived primary terrain attributes.

TANZ output [a]	Data type [b]	Correction factor [a] (divide by) [d]	TAPES-G attributes [c]	Units [c]	Definition [c]
Inherent in all raster outputs	fp	-	*x, y*	Metres	x and y coordinates as determined from DEM
fdir	int	1	Flow direction	None	Computed using D8 algorithm
elev	int	100	*z*	Metres	Elevation as read from DEM
drain	fp	-	Contributing area	Number of cells	Area draining from each cell
fwid	fp	-	Flow width	Multiple of cell width	Width associated with flow leaving cell
slopp	fp	-	Slope	Percentage	Slope in steepest downslope direction
slopr	fp	-	-	Radians	Slope in steepest downslope direction (derived from percentage slope)
slopd	fp	-	-	Degrees	Slope in steepest downslope direction (derived from percentage slope)
asp	int	1	Aspect	Degrees clockwise from north	Direction of steepest downslope slope
profc	fp	-	Profile curvature	Radians per 100 metres	Curvature of surface in direction of steepest descent
planc	fp	-	Plan curvature	Radians per 100 metres	Curvature of contour drawn through grid point
tanc	fp	-	Tangent curvature	Radians per 100 metres	Plan curvature multiplied by sine of slope angle
elres	int	100	Elevation residual	Metres	Difference between original DEM and depressionless DEM
fplen	fp	-	Flow-path length	Metres	Longest flow path from the catchment divide or edge of DEM to the cell
dasds	fp	-	*d(As)/ds*	None	Rate of change of specific catchment area along flow path

[a] Based on this work
[b] int (integer), fp (floating point)
[c] Based on Gallant and Wilson (2000)
[d] Divide surface by this number to convert data to correct decimal place

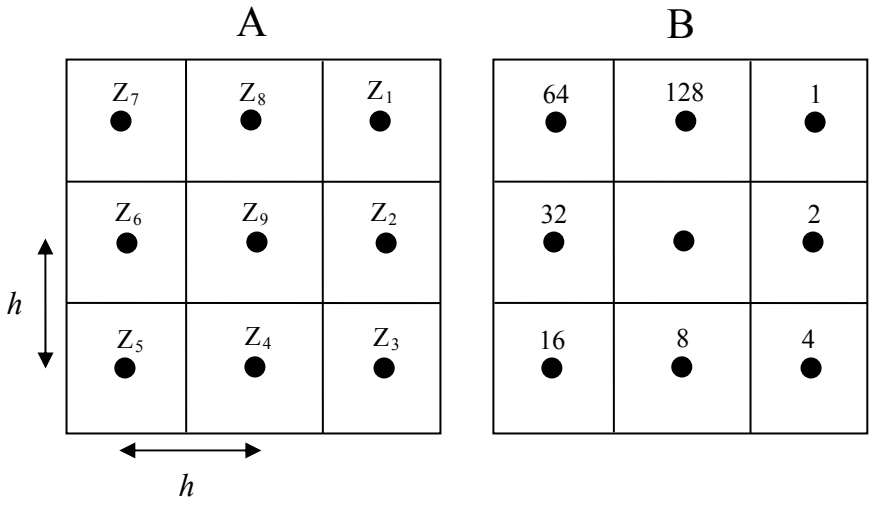


Figure 9: Diagram illustrating a 3 x 3 window from a gridded DEM (A) with the node numbering convention where *h* refers to the grid spacing, and (B) numbering convention for flow direction (after Gallant and Wilson, 2000).

$$z_x = \frac{\partial z}{\partial x} \approx \frac{z_2 - z_6}{2h} \tag{1}$$

$$z_y = \frac{\partial z}{\partial y} \approx \frac{z_8 - z_4}{2h} \tag{2}$$

$$z_{xx} = \frac{\partial^2 z}{\partial x^2} \approx \frac{z_2 - 2z_9 + z_6}{h^2} \tag{3}$$

$$z_{yy} = \frac{\partial^2 z}{\partial y^2} \approx \frac{z_8 - 2z_9 +}{h^2} \tag{4}$$

$$z_{xy} = \frac{\partial^2 z}{\partial x \partial y} \approx \frac{-z_7 + z_1 + z_5 - z_3}{4h^2} \tag{5}$$

$$p = z_x^2 + z_y^2 \tag{6}$$

$$q = p + 1 \tag{7}$$

5.1.1 Slope

Slope refers to the rate at which elevation changes in the direction of the steepest descent (Gallant and Wilson, 2000). TAPES-G provides two possible methods for calculating slope, the D8 and the finite-differences algorithm. The D8 method calculates slope using the steepest descent to one of its eight nearest neighbours and is written as:

$$S_{D8} = \max_{i = 1,8} \frac{z_9 - z_i}{h\Phi(i)} \tag{8}$$

where $\Phi(i) = 1$ for cardinal (N, S, E, and W) neighbours (i = 2, 4, 6, and 8) and $\Phi(i) = \sqrt{2}$ for diagonal neighbours (NE, SE, SW, and NW) to account for increased distance to those cells (Gallant and Wilson, 2000). An explanation of node numbering and neighbouring cells is given in Figure 9a.

The finite-difference algorithm uses four cardinal directions and takes the form:

$$S_{FD} = \sqrt{p} \tag{9}$$

The finite-difference method was chosen and implemented because it was considered to calculate slope to a greater degree of accuracy (Gallant and Wilson, 2000).

The TAPES-G percentage slope, S was converted to slope angle (in both radians and degrees) to provide alternatives for various slope modelling applications using an AML script. Slope in radians was calculated as:

$$\text{radians} = \arctan (S / 100)$$

Slope in degrees was calculated as:

$$\text{degrees} = \text{radians} * 180 / \pi$$

5.1.2 Aspect and primary flow direction

The terrain attribute aspect, ψ is considered reflective of a site's climatic environment. Generally speaking, New Zealand sites with northerly aspects exhibit warmer, drier environments when compared with southerly aspects. Aspect relates to the direction of steepest descent for each grid cell or point in a catchment and is the direction water would flow across a surface at that point (Moore, 1996). The TAPES-G programme measures aspect in degrees in a clockwise direction from north (Gallant and Wilson, 2000). Aspect in degrees is calculated using the finite-difference algorithm and takes the form:

$$\psi_{FD} = 180 - \arctan \left(\frac{z_y}{z_x} \right) + 90 \left(\frac{z_x}{|z_x|} \right) \qquad (10)$$

Note that as slope becomes very small, aspect becomes meaningless and difficult to define. Therefore when slope is below a predefined minimum value, aspect is considered to be undefined, or the slope has no aspect. Currently TAPES-G has a minimum slope value of zero and these cells are given a no-data value indicating undefined aspect.

In a modelling context, aspect can be difficult to define mathematically because of the circular nature of the values, for example 1 and 365 are similar aspects, but are mathematically very different. Therefore in a modelling situation, aspect is typically not used and instead a solar radiation index that accounts for the collective effects of slope, aspect and shadowing is favoured.

Primary flow direction, *FLOWD*, can be considered an approximate surrogate for aspect in that it defines the direction to the nearest neighbour with the maximum gradient, exactly as found in Equation 8 (Gallant and Wilson, 1996; 2000). Primary flow direction is shown in Equation 11.

$$FLOWD = 2^{j-1} \qquad \text{where } j = \underset{i=1,8}{\arg\max} \frac{z_9 - z_i}{h\Phi(i)} \tag{11}$$

i.e. where j is determined by the i that gives the largest slope value and therefore the direction of steepest descent. The approximate aspect corresponding to this flow direction is $\psi_{D8} = 45j$ (Gallant and Wilson, 1996; 2000). *FLOWD* is encoded using a binary notation (Figure 9 b), allowing identification of flow to multiple nearest neighbours (although TAPES-G does not use this capability) (Gallant and Wilson, 1996, 2000).

Also note that if the central node, z_9 elevation is lower than the surrounding z_1 to z_8 nodes (Figure 9a), then the node is considered a sink, or in a flat area, and has undefined flow direction. This situation highlights the importance of identifying actual sinks from spurious sinks and acting appropriately (See Section 4.4 for details). In TAPES-G the "create a depressionless DEM" can be implemented to remove unwanted sinks using Jenson and Domingue's (1988) algorithm to assign flow directions in depressions to ensure that the surface will drain correctly.

5.1.3 Curvature

The two most commonly calculated curvatures are plan (contour) curvature (K_c), which refers to the rate of change of aspect along a contour (Equation 12), and profile curvature (K_p), the rate of change of slope down a flow line (Equation 13) (Gallant and Wilson, 2000). Profile curvature is important when requiring the capture of information relating to flow velocity and sediment transport. Plan curvature captures topographic convergence and divergence across a landscape. Tangential curvature (K_t) is plan curvature multiplied by the sine of the slope angle. Plan, profile and tangential curvature are used frequently in modelling flow characteristics. Curvatures can also be used to delineate geomorphic units as described by Dikau (1989), where plan curvature can be used to distinguish between ridges, valleys, and hillslope. In contrast, profile curvature is used to differentiate between upper (convex) slopes and lower (concave) slopes. Figure 10 illustrates plan and profile curvature and the influence they can have on surface and subsurface water movement. Curvatures are calculated from second derivatives (rate of change of a first

derivative, for example slope or aspect) (Gallant and Wilson, 2000). Mitasova and Hofierka (1993) suggested that tangential curvature provides an improved representation of flow convergence and divergence compared with plan curvature because it does not have large values when the slope is small (Equation 14).

For plan and profile curvatures, the output values reflect the degree of curvature – gentle curvature is represented by small values whereas strong curvature has larger values. TAPES-G units for curvature are in radians per 100 metres and are calculated from:

$$K_c = \frac{z_{xx}z_y^2 - 2z_{xy}z_xz_y + z_{yy}z_x^2}{p^{3/2}} \quad (12)$$

$$K_p = \frac{z_{xx}z_x^2 + 2z_{xy}z_xz_y + z_{yy}z_y^2}{pq^{3/2}} \quad (13)$$

$$K_t = \frac{z_{xx}z_y^2 - 2z_{xy}z_xz_y + z_{yy}z_x^2}{pq^{1/2}} \quad (14)$$

In the TANZ surfaces, profile curvatures with negative values represent a convex profile, which typically occur on upper slopes, but positive values indicate a concave profile usually found on lower slopes (see Section 9.2 for examples). Plan curvature has negative values for diverging flow (on ridges) and positive values for converging flow in valleys.

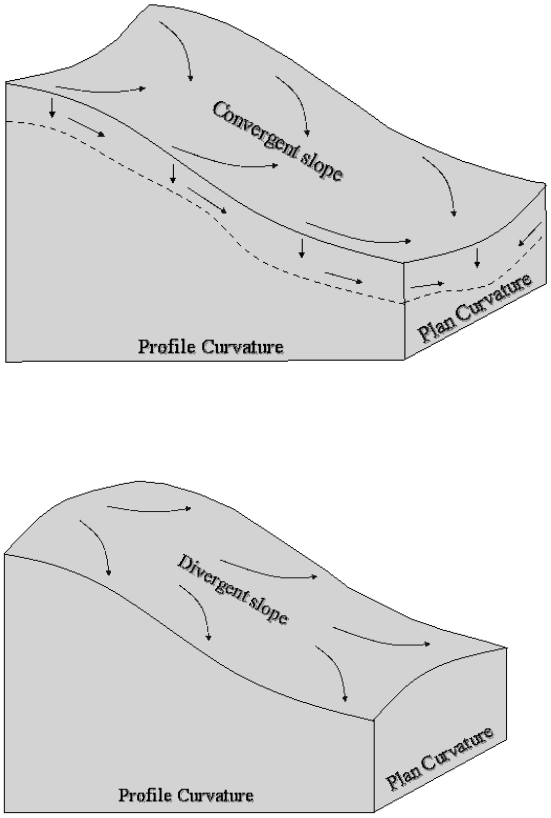


Figure 10: Schematic diagrams illustrating plan and profile curvatures and their influence on surface and subsurface water movement associated with convergent and divergent hillslopes (adapted from Pennock et al. 1987, after Hall and Olsen, 1991, p.15).

5.1.4 Upslope contributing area

Upslope contributing area, A, is estimated using flow direction(s) from a given node and is the area above a given length of contour that contributes to flow across the contour, as described in Figure 11 (Gallant and Wilson, 2000). Upslope contributing area is also often referred to as the drainage area or catchment area. Closely related is specific catchment area, A_s, which is the drainage area per unit width orthogonal to a flowline

(Gallant and Wilson, 1996). Upslope contributing area and specific catchment area outputs from TAPES-G can be represented either in the number of contributing cells or in square metres.

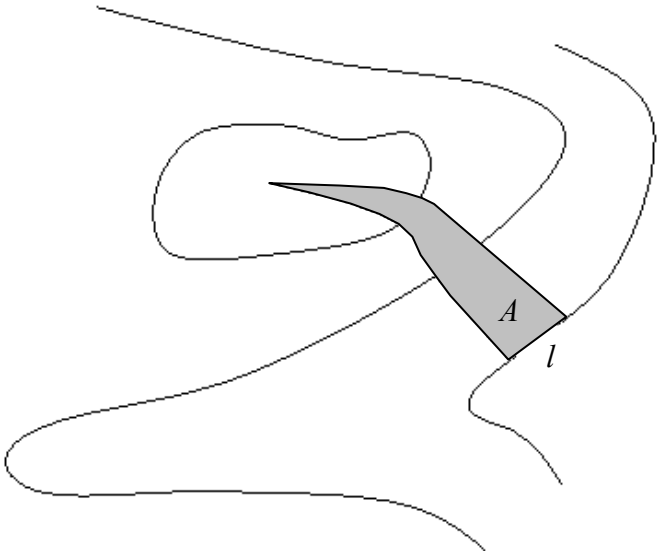


Figure 11: Upslope contributing area *A* is the area of land upslope of a length of contour *l*, whereas specific catchment area *A*s is *A*/*l* (after Gallant and Wilson, 2000).

The TAPES-G programme provides four different algorithms for the calculation of contributing area: the single-flow-direction, D8 method (O'Callaghan and Mark 1984); the randomised single-flow-direction, Rho8 method (Fairfield and Leymarie, 1991); the multiple-flow-direction methods, FD8 and FRho8 (Moore et al., 1993b); and the DEMON stream-tube method (Lea, 1992; Costa-Cabral and Burges, 1994). A flow routing algorithm simulates the distribution of flow from a cell to one or more cells down

slope. The choice of algorithm is important because it influences calculations for upslope contributing area, specific catchment area, stream power index and other terrain attributes.

The D8 algorithm developed by O'Callaghan and Mark (1984) is referred to as the deterministic eight-node method because it allows flow to pass from the processing cell to only one of its eight neighbours in the direction of steepest descent, as defined in Equation 11 (*FLOWD*). A limitation to the D8 method is that it cannot model flow dispersion and tends to "produce flow in parallel lines along preferred directions that will agree with the aspect only when the aspect is a multiple of 45° (Gallant and Wilson, 2000)." For example, on a surface where aspects range from 0 and 22.5° , the D8 algorithm will predict a constant flow direction *FLOWD* = 128 (north) (Gallant and Wilson 1996, 2000). Even with this limitation the D8 algorithm remains a popular method for calculating upslope contributing area.

The randomised single-flow-direction, Rho8 method, developed by Fairfield and Leymarie (1991), offers a random or stochastic element to the D8 method in an attempt to disperse flow paths and avoid parallel flowlines. In Equation 11, Φ (i) is replaced with ($2–r$) for the diagonal neighbours (i = 1, 3, 5, or 7 in Figure 9a), where r is a uniformly distributed random variable between 0 and 1 (Gallant and Wilson, 1996). A decision was made not to implement this method in our national-scale modeling because, like the D8 algorithm, Rho8 cannot model flow dispersion and uses a stochastic solution to address

parallel flow paths. As the stochastic nature of the algorithm makes the results non-repeatable, we therefore considered the approach inappropriate for our purposes.

The FD8 and FRho8 algorithms provide alternatives to the D8 and Rho8 algorithms by allowing flow dispersion to be captured (Moore et al., 1993b). The FD8 and FRho8 methods enable flow to be distributed to multiple nearest neighbouring nodes in elevated areas above defined channels recognising flow dispersion. The portion of flow or upslope contributing area assigned to each downslope neighbour, F_i is determined on a slope-weighted basis as suggested by Freeman (1991) and Quinn et al., (1991), and is defined as:

$$F_i = \frac{\max(0, S_i^v)}{\sum_{i=1}^{8} \max(0, S_i^v)} \tag{15}$$

where S_i is the slope from the central node to the neighbour i (the same term used in Equation 8) and v is a positive constant predefined as 1.1 in the TAPES-G programme. The value of 1.1 arises because Freeman (1991) found this value produced the highest degree of accuracy for artificial conical surfaces (Gallant and Wilson 1996).

A fourth and final option is the DEMON method which provides a completely different approach to modelling flow accumulation and dispersion. It is similar to the stream-tube method developed by Moore and Grayson (1991) used with contour-based DEMs (Gallant and Wilson, 1996; 2000). The DEMON module refers to Digital Elevation Model Networks, developed by Costa-Cabral and Burges (1994) based on a method

proposed by Lea (1992). Gallant and Wilson (1996; 2000) stated that, "in DEMON, flow is generated at each pixel (source pixel) and is followed down a stream tube until the edge of the DEM or a pit is encountered." There are two issues of concern if applying the DEMON method for modelling flow at the national extent. Firstly, the DEMON method for modelling flow is not yet fully tested against other available methods, and secondly, it is a relatively memory-hungry programme compared with the other previously mentioned modelling methods. Therefore, based on these restrictions, a decision was made to compile the TAPES-G programme without the DEMON module.

For our project we used both the D8 and FD8 algorithms using a maximum cross-grading area of 25,000 m^2 (see Section 5.1.5). The greatest advantage of TAPES-G-derived TANZ surfaces is that these two methods are used to develop upslope contributing area and their associated data. This pair of algorithms allows flow to be distributed to multiple neighbouring cells in upland areas above defined stream channels, and hence both flow convergence and divergence may be modelled or simulated. The FD8 algorithm is employed above areas of presumed channel initiation, allowing flow divergence to be captured. This ability to model both divergent and convergent flow within the landscape provides an improvement to hydrological digital modelling not seen before, hence justifying the implementation of these algorithms into the TANZ surfaces.

5.1.5 Rationale behind maximum cross-grading area threshold

Initially, relevant literature discussing appropriate maximum cross-grading area thresholds for switching between the D8 and FD8 algorithms was consulted. Parameter values used for the maximum cross-grading area thresholds were shown to vary greatly depending on the type of landscape modelled. Gallant and Wilson (2000) discussed the improvement in calculation of contributing area when implementing the FD8 algorithm in upslope areas. In valleys where floodplains and riparian areas are represented, however, the D8 algorithm is useful because stream lines are usually well defined. Therefore, setting of the maximum cross-grading area threshold in the TAPES-G programme becomes important. The rationale is that finer textured landscapes with high valley densities require smaller cross-grading areas compared with coarse landscapes with lower valley densities, which require larger cross-grading area thresholds (Gallant and Wilson, 2000). An example would be a highly dissected region (fine textured) such as the Southern Alps of New Zealand's South Island compared with the gently undulating to strongly rolling older landscapes (coarse textured) of Northland in the North Island. These examples show that optimal cross-grading threshold depends greatly on the landscape texture or valley density. Gallant and Wilson (2000) also suggested that 10 hectares or 100,000 m^2 is a reasonable starting position from which to work.

When comparing New Zealand's landscape pattern marked with strong relief, diversity, complexity and general youthfulness with those in which overseas studies are based, special consideration must be given to this heterogeneous landscape structure. For

example, the Coromandel macro-catchment (Figure 17) illustrates the extremes in landscape structure and texture. It typifies many New Zealand catchments beginning at sea-level and within a short distance becoming diverse and structurally complex. This complexity becomes problematic when justifying a generic maximum cross-grading area threshold for national-scale modelling across an entire country. Nationally, about 39% of New Zealand topography is in the flat to undulating category (slopes 0-7°), whereas 36% is rolling to hilly (slopes 8-25°) and about 25% is steep to very steep terrain (slopes > 25°) (Table 2). We investigated maximum cross-grading area thresholds of 10,000, 15,000, 20,000, 25,000, and 50,000 m^2 in areas of relatively fine, medium, and coarse textured landscapes. Generally, we found strong visual differences because the parallel flow lines created by the D8 algorithm were replaced by the FD8 multiple flow algorithm with increasing cross-grading area. However, increasing cross-grading area increases computational time. Taking into account thresholds used in the literature and the complex diverse textural differences found across New Zealand, a maximum cross-grading area of 25,000 m^2 was set. We recognise, however, that some macro-catchments consisting of coarse textured landscapes, for example the Canterbury Plains, may have some improvement in upslope contributing area were the threshold to be increased. Indeed, Fried et al. (2000) commented that "flow-path determination by the various methods is especially challenging in areas of low relief." Also of interest are the comments of Wilson et al. (2000), who stated that the cross-grading area has less impact on resultant terrain attributes than either the DEM-source data and resolution or the flow routing algorithm driving the calculations.

5.1.6 Specific catchment area, flow width, and maximum flow path-length

The terrain attribute, flow width, in conjunction with upslope contributing area, is used to calculate specific catchment area. Specific catchment area is defined as the ratio of the contributing area to the contour length or flow width. Therefore the importance of flow width is found in the development of other terrain attributes. The smallest possible flow width is the cell width, which in TANZ is 25 m. Increasing flow width is dependent on the flow algorithm used (D8 and/or FD8) and whether flow is passing in a cardinal or diagonal direction from each cell. In contrast, the maximum flow path length is referred to as the maximum length of all flow paths from the catchment boundary to a given point in the DEM (Gallant and Wilson, 2000). Closely related is the rate of change of specific catchment area along a flow path (dA_s/ds), which is computed by the TAPES-G programme as the specific catchment area of flow leaving a cell less the average of the specific catchment areas entering the cell, divided by the flow length across the cell (Gallant and Wilson, 2000). Moore (1996) suggested that this terrain attribute has some applications in erosion models.

The flow-width attribute, w, and upslope contributing area, A, are used to calculate specific catchment area A_s:

$$A_s = \frac{A}{w} \tag{16}$$

When using the D8 and Rho8 methods in TAPES-G, flow width is computed as equal to the cell width in cardinal directions, and diagonal cell width for flow in diagonal directions:

$$w = \begin{cases} h & i = 2, 4, 6, 8 \\ \sqrt{2}\,h & i = 1, 3, 5, 7 \end{cases} \tag{17}$$

TAPES-G originally used the D8 method to model FD8/FRho8 flow width, but was updated to account for the amount of dispersion of flow by first calculating the effective number of flow directions from the weights:

$$n_{\text{flow}} = \frac{1}{\sum_{i=1}^{8} F_i^2} \tag{18}$$

This estimates n_{flow} from a single direction through to flow dispersed in eight directions equally. Flow width is then estimated as:

$$w = \begin{cases} h & n_{\text{flow}} < 3 \\ h\left[\, 1 + 3\left(\frac{n_{\text{flow}} - 3}{5}\right) \,\right] & n_{\text{flow}} \geq 3 \end{cases} \tag{19}$$

When n_{flow} is small in convergent and planar areas the flow width is h. However, in areas where n_{flow} is large, flow width increases to a maximum of $4h$. “This method takes no

account of whether flow is in a cardinal or diagonal direction" (Gallant and Wilson, 2000).

TAPES-G also calculates maximum flow-path length, which is defined as the "maximum length of all flow paths from the catchment boundary to a given point in the DEM" (Gallant and Wilson, 2000). It is calculated using the D8 or Rho8 methods previously described by accumulating flow distances across the cells.

5.2 Automating the TAPES-G process

Scripts were written to automate the process of running the TAPES-G programme for each macro-catchment, removing all possible manual user input errors and reducing the time requirements. Effectively this meant that the TAPES-G programme could run through consecutive macro-catchments until completion with little technical input.

The main steps of the automated process are illustrated in Figure 12. For each macro-catchment, the appropriate command file containing the TAPES-G input parameters (see Table 4 for details) is selected. These parameters, together with the .txt file of elevations previously created by the tapes.aml for each macro-catchment, form the input to the primary terrain attribute computations. The resultant ascii grid file contains the x and y spatial coordinates along with a series of values representing the thirteen primary terrain

attributes. The result for each macro-catchment was copied to a suitable storage space capable of containing the large volumes of data.

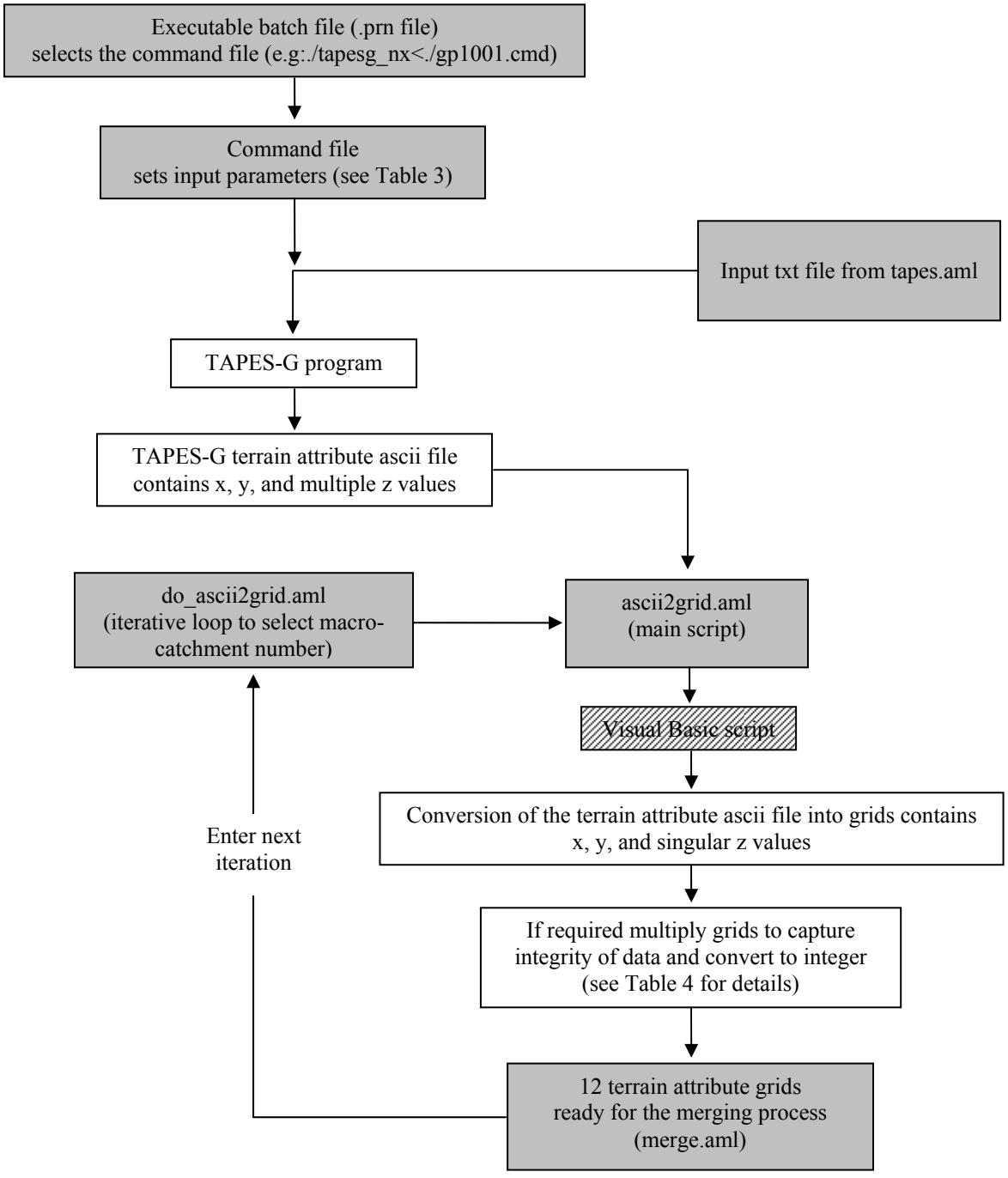


Figure 12: Flow chart illustrating main steps involved in the automation and running of the TAPES-G programme and the extraction of resultant output data.

Extracting the terrain attribute data from the ASCII text files into grids was undertaken by additional AMLs. The primary AML (do_ ascii2grid.aml) contains an iterative loop that calls each text file consecutively, representing each macro-catchment. This primary AML runs a secondary AML (ascii2grid.aml) where a VB script is called and executed. The VB script extracts the x and y spatial coordinates together with one of the thirteen primary terrain attributes within the TAPES-G derived text file. The primary terrain attribute data are given appropriate names (see Table 5) and are converted to a raster format. The output of this process is twelve grids.

6. SECONDARY TERRAIN ATTRIBUTES

6.1 Development of secondary terrain attributes

Secondary topographic attributes were derived from two or more primary attributes (see Table 5) that characterise specific processes occurring within a landscape. In this work we derived a topographic wetness index (TWI), stream power index (SPI), the length and slope factor (LS) for the revised universal soil loss equation (RUSLE), and shortwave radiation surfaces (Table 6). Note that the shortwave radiation model required an additional latitude surface, not strictly considered a primary terrain attribute.

The topographic wetness index (TWI), commonly called the compound topographic index, is extensively used to "describe the effects of topography on the location and size of saturated source areas of runoff generation" (Wilson and Gallant, 2000). The TWI is

useful in capturing information related to the distribution and abundance of water, depth of water table, evapotranspiration, susceptibility of soils to erosion, and soil horizon thickness. These features in turn influence the spatial patterns of native vegetation and pedogenic processes. TWI is commonly used in hydrological, geomorphological, pedological, and ecological applications. Jones (1986; 1987) discussed some of the pitfalls when using the TWI in inappropriate ways. Quinn et al. (1995) described how the steady-state TWI could be used effectively in the TOPMODEL hydrologic modelling structure as well as highlighting the various problems associated with use and application of TWI models.

Stream power index (SPI) is considered a "measure of erosive power of flowing water based on the assumption that discharge (q) is proportional to specific catchment area (A_s)" (Wilson and Gallant, 2000). At locations of "profile convexity and tangential concavity (flow acceleration and convergence zones)" net erosion occurs, while at locations of "profile concavity (zones of decreasing flow velocity)", net deposition occurs (Wilson and Gallant, 2000). SPI has been extensively used in erosion studies, sediment transport models, and in geomorphological studies to measure the erosive power of flowing water (Moore et al., 1991). Therefore, SPI influences properties such as the potential erosive power of overland flows, plant cover type and distribution, soil-horizon thickness, and organic matter content.

The length-slope factor (LS factor) developed here for TANZ was derived from primary terrain attributes created in TAPES-G using the equations of McCool et al. (1987; 1989).

The RUSLE is a well-researched, empirically driven, model for sheet and rill erosion comprising rainfall-runoff erosivity, soil erodibility, plant cover management, supporting practices and the slope-length and slope-steepness factors. Typically, the slope-length factor and the slope-steepness factor are combined together as a function of both slope and length of the landscape. In general, as the length of the slope increases, cumulative runoff increases. Furthermore, as the slope of the landscape increases, higher runoff velocities contribute to increased erosion. Griffin et al. (1988) and Moore and Wilson (1994) suggested that the LS factor values could be multiplied by 1.6 to estimate soil-loss values for individual grid cells.

These secondary topographic attributes were calculated using AML scripts. The TWI takes the form:

$$\mathrm{TWI} = \ln\left(\frac{A_s}{\tan\beta}\right) \tag{20}$$

where A_s is the specific catchment area ($m^2\ m^{-1}$), and β is the slope angle in radians. This form of the topographic wetness index assumes steady state conditions prevail (Moore et al. 1991, 1993b; Moore and Wilson, 1992).

Table 6: TANZ secondary terrain attributes and correction factors.

TANZ output [a]	Data type [b]	Correction factor [a] (divide by) [d]	Attributes	Description
twi	fp	-	Topographic wetness index [(c)]	Assumes steady-state conditions and uniform soil properties (i.e., transmissivity is constant throughout the catchment and equal to unity). This pair of equations predicts zones of saturation where A_s is large (typically in converging segments of landscapes), β is small (at base of concave slopes where slope gradient is reduced), and T_i is small (on shallow soils). These conditions are usually encountered along drainage paths and in zones of water concentration in landscapes [(c)].
spi	fp	-	Stream power index [(c)]	Measure of erosive power of flowing water based on assumption that discharge (q) is proportional to specific catchment area (A_s). Predicts net erosion in areas of profile convexity and tangential concavity (flow acceleration and convergence zones) and net deposition in areas of profile concavity (zones of decreasing flow velocity) [(c)].
ls	fp	-	Length and slope factor (RUSLE) [(e, f)]	This equation estimates rill and sheet erosion based on the revised universal soil loss equation [(e, f)].
srad	int	100	Clear sky shortwave radiation [(g)]	Estimation of clear sky shortwave radiation that accounts for influences of surrounding overshadowing terrain [(g)].

[a] Based on this work
[b] int (integer), fp (floating point)
[c] Based on Gallant and Wilson (2000)
[d] Divide surface by this number to convert data to correct decimal place
[e] Based on McCool et al. (1987)
[f] Based on McCool et al. (1989)
[g] Based on Kumar et al. (1997)

The SPI uses the same primary attributes as the TWI and takes the form:

$$\text{SPI} = A_s \tan \beta \tag{21}$$

The LS factor uses the equations by McCool et al. (1987, 1989), and is also referred to by Wilson et al. (2000). It takes the form:

$$L = \left(\frac{\lambda}{22.13} \right)^m \tag{22}$$

$$S = 10.8 \sin \beta + 0.03 \ (\tan \beta < 0.09) \tag{23}$$

$$S = 16.8 \sin \beta - 0.5 \ (\tan \beta \geq 0.09) \tag{24}$$

where L is the slope-length factor,

λ refers to the flow-path length (in metres) ($\lambda = 4$ m when $\lambda \leq 4$ m),

$m = F/(1 + F)$, where $F = (\sin \beta / 0.0896) / [3(\sin \beta)^{0.8} + 0.56]$ in erosional areas, and F = 0 in depositional areas,

while S is the slope steepness factor,

and β is the slope angle (McCool et al. 1987, 1989).

"The LS factor becomes unity for the case where the flow-path length is 22.13 m and the slope is 9% (Wilson et al., 2000)."

7. SOLAR RADIATION

7.1 Solar-radiation modelling

Incoming solar radiation (insolation) is essential to most biological and physical processes (Pinde and Rich, 2002) and is a major source of energy for the Earth with direct shortwave radiation contributing substantially to global radiation and the energy balance. When modelling environmental and biological systems, properties such as aspect and slope are useful in providing information relating to soil temperature, soil moisture, physical and biological processes and consequently site productivity. As described in Section 5.1.2, aspect is mathematically problematic because northern aspects can be expressed by different values. This is true also of areas with no aspect because most aspect models (including TAPES-G derived surfaces) require slope to be zero for no aspect to occur (Wilson and Gallant, 2000). Because of these issues, another approach was sought. An alternative approach is the shortwave radiation model of Kumar et al. (1997). This model encompasses slope and aspect adequately yet also accounts for the influence of shaded terrain. It has modelling simplicity and flexibility, as well as minimal computation time and storage requirements. Hence the shortwave radiation model of Kumar et al. (1997) was used to model solar radiation across New Zealand at a 25 m-resolution utilising DEM and latitude surfaces.

7.2 Solar-radiation model implementation

For our modelling requirements we created a core AML routine that assigned model parameters (DEM input, grid output, latitude grid, model start day, model finishing day, and modelling time interval) to be passed through to the main shortwave-radiation model of Kumar et al. (1997). A latitude surface of New Zealand was created (Section 7.4) and the latitude values were converted to a raster grid containing New Zealand map grid coordinates compatible with the other input data and modelled surfaces. The core AML functions in a similar manner to other AMLs created for TANZ (do_amls) by selecting one macro-catchment at a time for processing (Figure 13). Initially, each region of interest is extracted using a macro-catchment boundary mask (for modelling rationale see Section 3.2) and the centroid or central cell in each macro-catchment grid is determined. The latitude value for this centroid cell is passed onto the main shortwave radiation AML of Kumar et al. (1997).

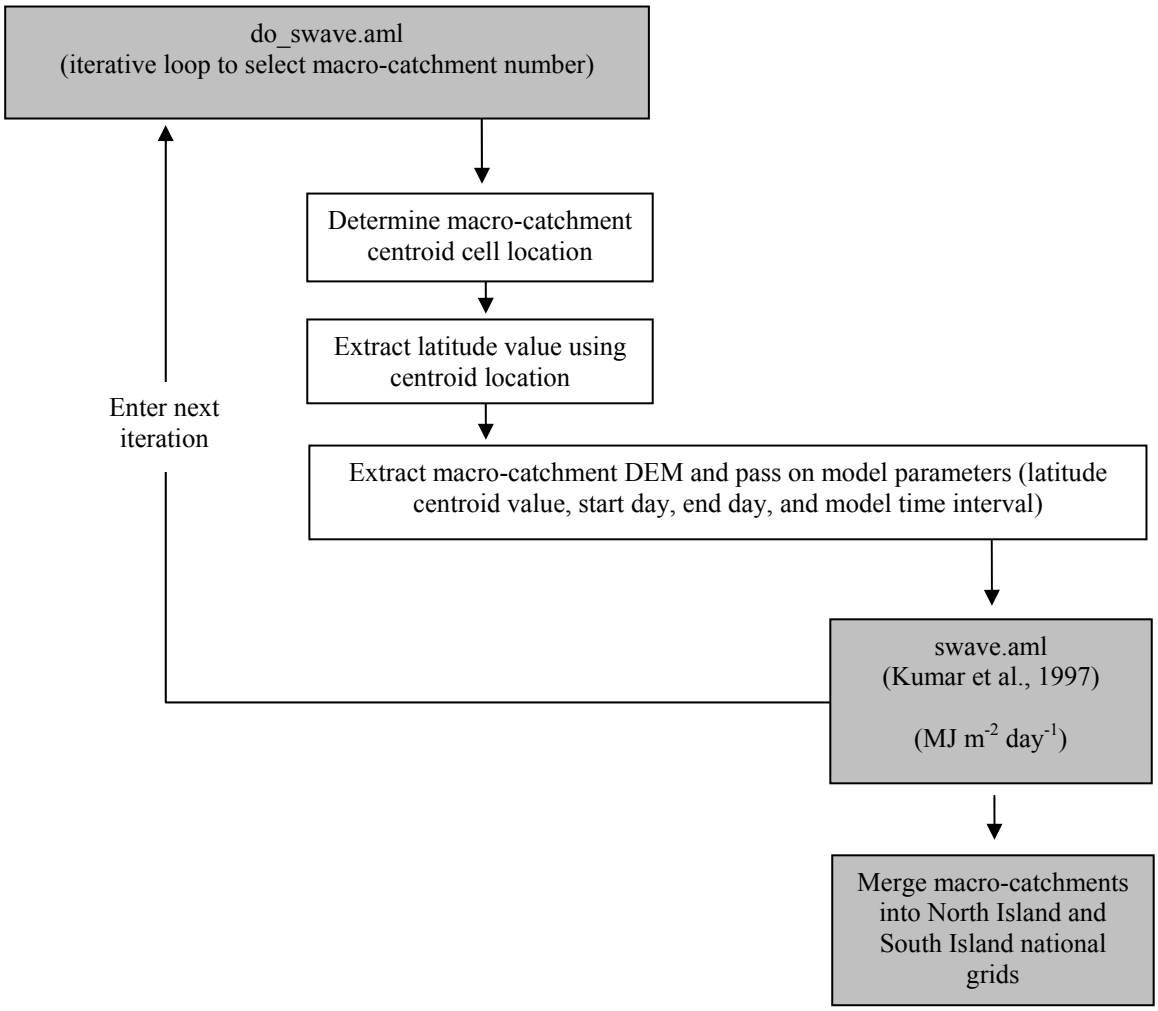


Figure 13: Flow chart illustrating main steps in modelling shortwave radiation adjusted for topographic overshadowing.

7.3 Solar-radiation model rationale

Sensitivity analysis was undertaken on one macro-catchment to determine the effect of time and latitude interval resolution on modelling accuracy and model computational time. A macro-catchment in the Coromandel Peninsula on eastern North Island was selected (Figure 3). Eight time intervals were trialled, namely, 1, 5, 10, 15, 30, 60, 90, and 120 minute intervals were modelled and descriptive statistics derived. Statistics were also collected for modelling the Coromandel macro-catchment using the northern, central, and most southern latitude values. Table 7 provides descriptive statistics highlighting (1) model sensitivity to change in time interval between calculations, and (2) model sensitivity to change in latitude.

When choosing parameter values, consideration was given to possible future applications of these data and a trade-off was made between processing time requirements and modelling accuracy. As a result two processing time intervals were chosen. Firstly, 120 minutes was used to create shortwave-radiation grids for each month of the year with the resultant surfaces averaged to provide a mean annual shortwave-radiation surface for New Zealand. Secondly, a time interval of 60 minutes was chosen to calculate average monthly shortwave radiation for the months of January (mid summer) and July (mid winter). Furthermore, implementing the centroid latitude value was considered adequate for our modelling purposes because sensitivity analysis did not indicate significant changes to model outcomes from the changes in the latitude from the north to the south of the Coromandel macro-catchment. A further advantage of using the centroid latitude is

that centroid selection can be automated, removing the possibility of human induced errors.

Table 7: Descriptive statistics illustrating model sensitivity to a) decreasing modelling time interval, and b) change in latitude values from north to south across the Coromandel macro-catchment at 120 and 60 minute time intervals, and the influence these had on insolation (monthly mean data: MJ m^{-2} day^{-1}).

Parameter Sensitivity	Time interval (mins)	Latitude (decimal degrees)	Mean	Median	Range	Standard deviation
a) Time interval	120	-37.093	18.53	19.26	24.96	3.5361
	90	-37.093	18.50	19.20	24.91	3.5348
	60	-37.093	18.48	19.16	2488	3.5311
	45	-37.093	18.48	19.16	24.88	3.5315
	30	-37.093	18.48	19.16	24.88	3.5317
	15	-37.093	18.48	19.15	24.88	3.5303
	10	-37.093	18.48	19.16	24.88	3.5306
	5	-37.093	18.48	19.15	24.87	3.5307
	1	-37.093	18.47	19.15	24.87	3.5306
b) Latitude	120	-37.019	18.56	19.29	24.96	3.5329
	120	-37.093	18.53	19.26	24.96	3.5361
	120	-37.167	18.51	19.24	24.95	3.5395
	60	-37.019	18.51	19.19	24.89	3.5279
	60	-37.093	18.48	19.16	24.88	3.5311
	60	-37.167	18.45	19.15	24.87	3.5342

Sensitivity analysis was also undertaken across several macro-catchments with parallel relief running in north to south and east to west directions to assess the possible impact of differing catchment buffers from the original buffer on model outcomes. The buffer resolutions tested were 250, 500, 1000, 2500, and 5000 m. Results indicated that there was no advantage in expanding the boundary buffer beyond 250 m. This outcome was not surprising considering catchment boundaries are at the highest point in the landscape and are generally not overshadowed by terrain. However, there may be some residual impact

from larger landscapes, for example the Southern Alps on the Canterbury Plains. However, capturing this influence would defeat the purpose of dividing New Zealand into macro-catchments of appropriate size to match with computer requirements. Indeed, ever-increasing the catchment sizes will greatly increase computational time. Therefore, the standard 250 m macro-catchment buffer was considered appropriate for TANZ modelling.

7.4 Latitude surface development

A major component of the shortwave solar radiation model of Kumar et al. (1997) is the latitude value in the area of interest. For this purpose, a latitude value surface was created for New Zealand by dividing the North Island into five and the South Island into six regions. The spatial extent of these regions was chosen to allow resultant text data to fit within a standard ~ 64,000 row Excel spreadsheet at 1000 m grid-point resolution. An AML was written to convert shape files to ArcInfo point coverages and to form topology prior to conversion to grids at a 1000 m-resolution. Finally, x, y, and z coordinates were added to each grid cell location before writing these data to a .txt file (containing < 64,000 rows of data). Each resultant .txt file was extracted to spreadsheet and edited to contain x and y spatial location coordinates in the New Zealand map grid system.

Using the LINZ website, each of the 11 spreadsheet files was manually copied and pasted to:

http://www.linz.govt.nz/rcs/linz/pub/web/root/core/SurveySystem/GeodeticInfo/CoordinateConversions/coordinateconversions/index, where the New Zealand map grid x and y coordinates were converted to the latitude and longitude coordinate system (World Geodetic System). The x, y, and new latitude values were saved before converting back to point coverages and finally to 1000 m-resolution grid surfaces containing x and y New Zealand map grid projection and World Geodetic System latitude z values. All grids were merged into North Island and South Island surfaces and converted to the required resolution.

8. CREATION OF NATIONAL SURFACES

8.1 Merging data into national-scale coverages

The merge command in ArcInfo was used to bring all the macro-catchments together into national-scale coverages. Initially, the merging AML takes each macro-catchment and masks the attributes to the original macro-catchment input coverage size. For masking details refer to Section 3.2.1 and Figure 6 (mask grid). The original macro-catchments are merged together into groups of macro-catchments adjacent to one another because ArcInfo limits merging to fifty concurrent grids at any one time. Figures 14 and 15 illustrate the groupings of macro-catchments across New Zealand. These groups consist of between 26 and 48 macro-catchments. Finally, the grids are masked to the original coastline.

Figure 14: Map illustrating merged groupings of macro-catchments for the North Island.

Figure 15: Map illustrating merged groupings of macro-catchments for the South Island.

9. TERRAIN ATTRIBUTES OF NEW ZEALAND (TANZ)

The following section shows examples of primary and secondary terrain attribute outputs developed for New Zealand. Graphics for a selected Coromandel macro-catchment, displayed in both the second and third dimensions, provide realistic digital representations from TANZ data (Figures 16 and 17). The two dimensional maps from the Coromandel macro-catchment were clipped using the macro-catchment polygon coverage discussed in Section 3.2. Further, a subcatchment within the macro-catchment was selected and displayed three dimensionally using ArcGIS 3D scene, emphasising and illustrating the spatial patterns associated with each terrain attribute occurring within that area.

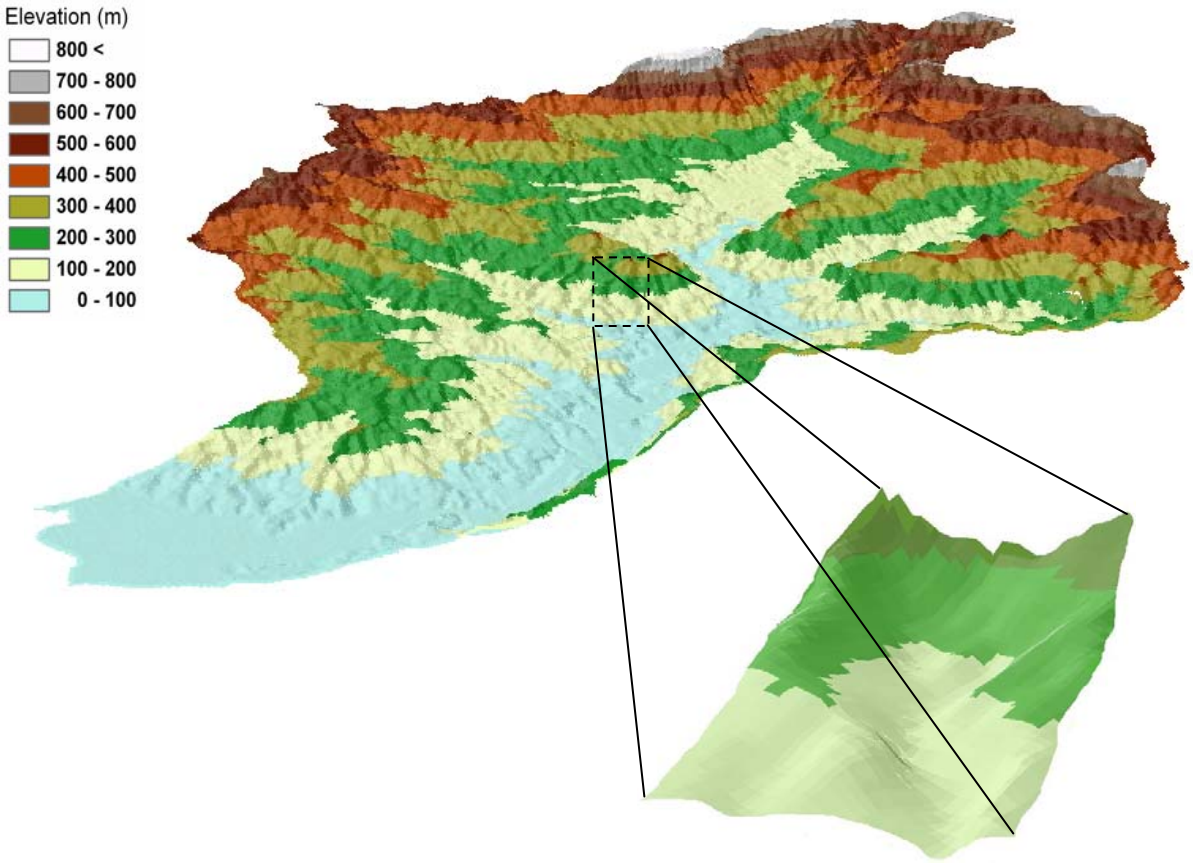


Figure 16: Three dimensional DEMs of the Coromandel macro-catchment (upper) and a subcatchment (lower) displayed with a 1.5 vertical exaggeration and hillshade to highlight topographical features (see Fig. 17 for representations of scale and directional bearing not available in the third dimension). Note the expanded box is an approximate subcatchment location with the expanded graphic displayed from an angle different from that of the macro-catchment.

9.1 Elevation, slope, aspect, and flow direction

Figure 17 illustrates the strong relief occurring across the Coromandel macro-catchment, lying in a northeast to southwest direction. This catchment is approximately 15 km in width and rises from sea-level at the Firth of Thames (southwest corner) with a broad flattish alluvial fan up to an elevated and strongly dissected catchment boundary, peaking at more than 800 m above sea-level.

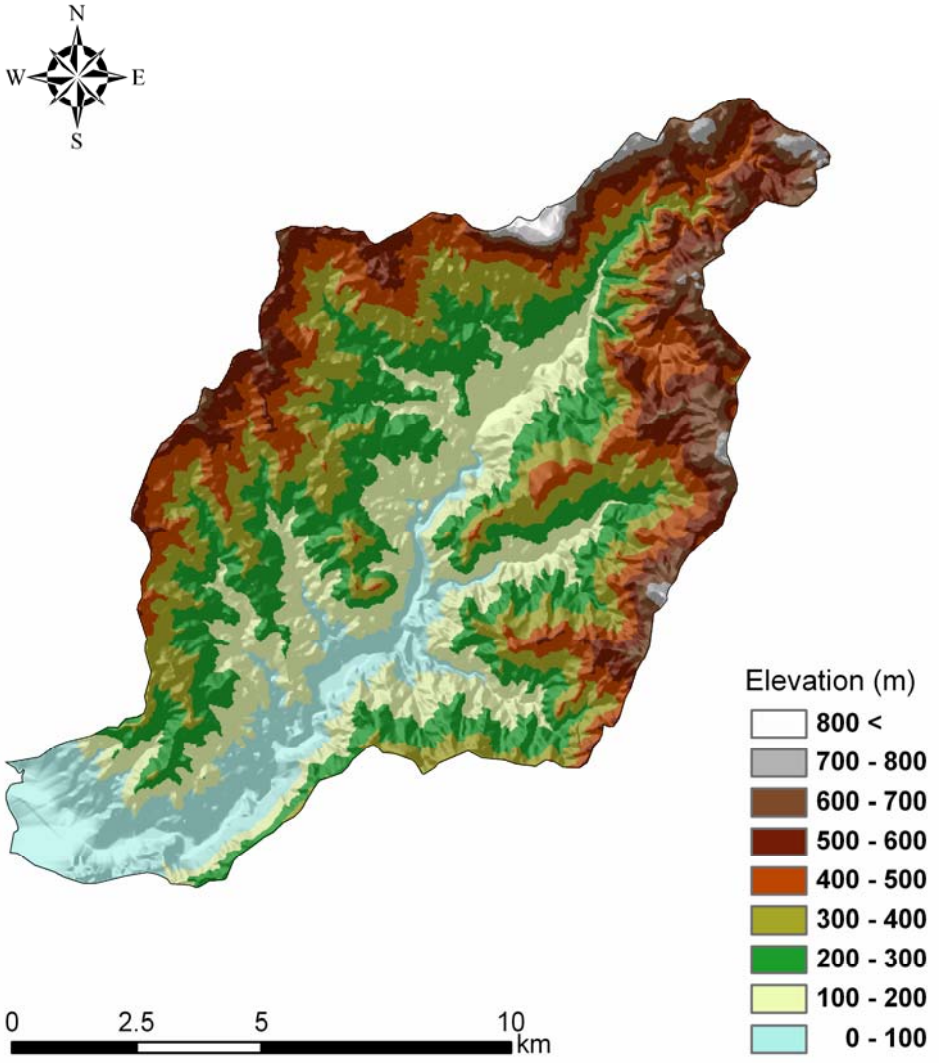


Figure 17: DEM with shaded relief of the Coromandel macro-catchment (hillshade inclusive).

Slope angle was classified into seven categories according to Young's (1972) steepness criteria and is shown in Figure 18. The rationale behind this application is that the Coromandel macro-catchment, like many catchments in New Zealand, covers a large range of slopes from flattish surfaces with little or no slope on alluvial deposits and river plains, through to very steep erosional regions in the upper parts of the catchment. Therefore, slope was divided into the seven classes providing easy visual assessment. Slope in degree angle is a common method for representing landscapes in the New Zealand context visually. Figure 18 indicates that a large portion of the macro-catchment has slopes < 18 degrees. The lower-lying stream channel and alluvial plain are predominately less than 2 degrees, whereas valley tributaries often fall within the 5 to 10 degree range. A small portion of the upper catchment (northeastern end) contains slopes ≥ 45 degrees. Figure 19 provides a three dimensional view of the Coromandel subcatchment, including hillshade, showing slope angles in degrees. This subcatchment graphic indicates that the majority of lower slope angles at this location are in valley bottoms or on hill crests or summit slopes of spurs on the low-lying rolling terrain. Steeper slopes occur on the elevated mountainous terrain.

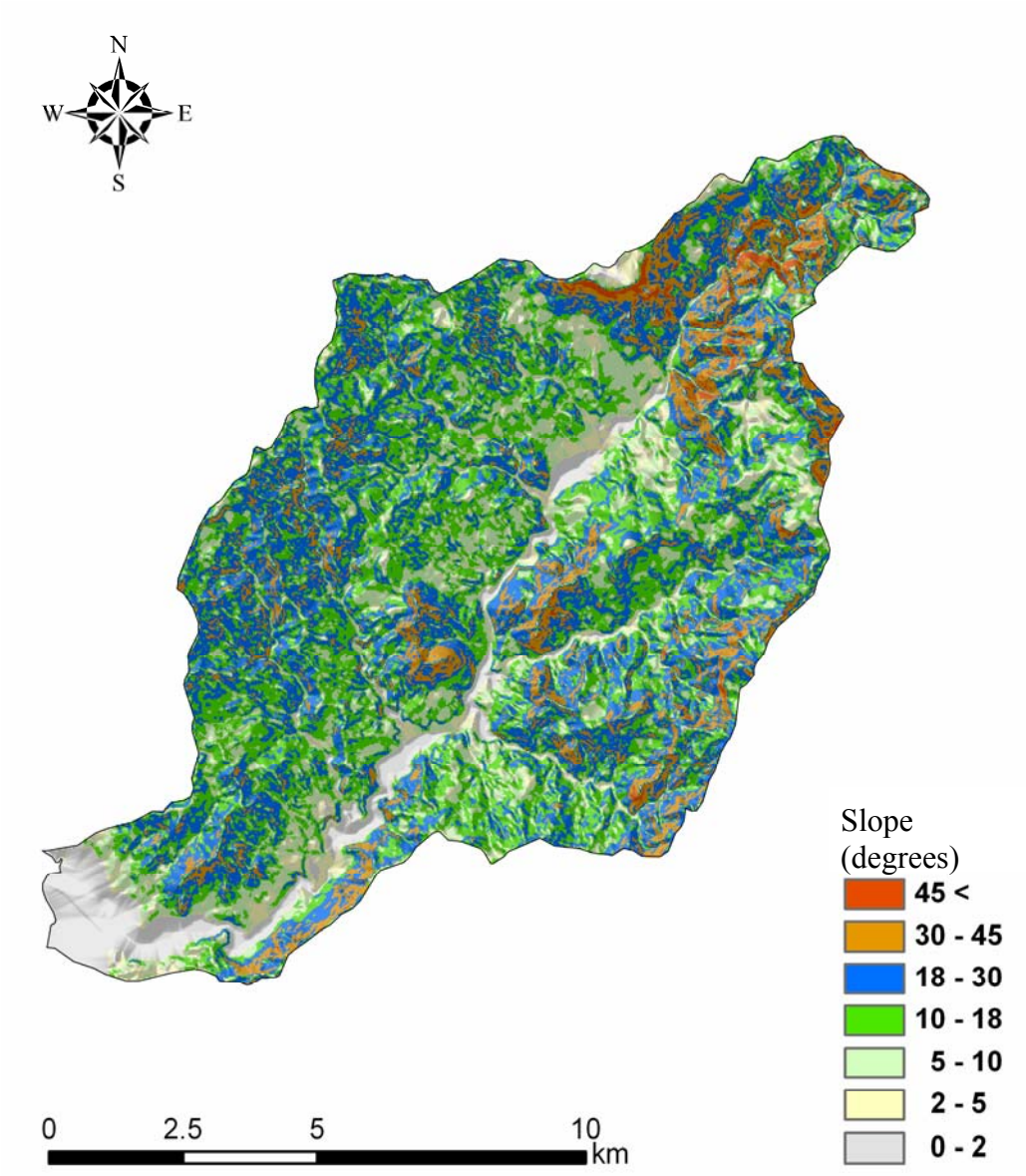


Figure 18: Slope angle (degrees) of the Coromandel macro-catchment (hillshade inclusive).

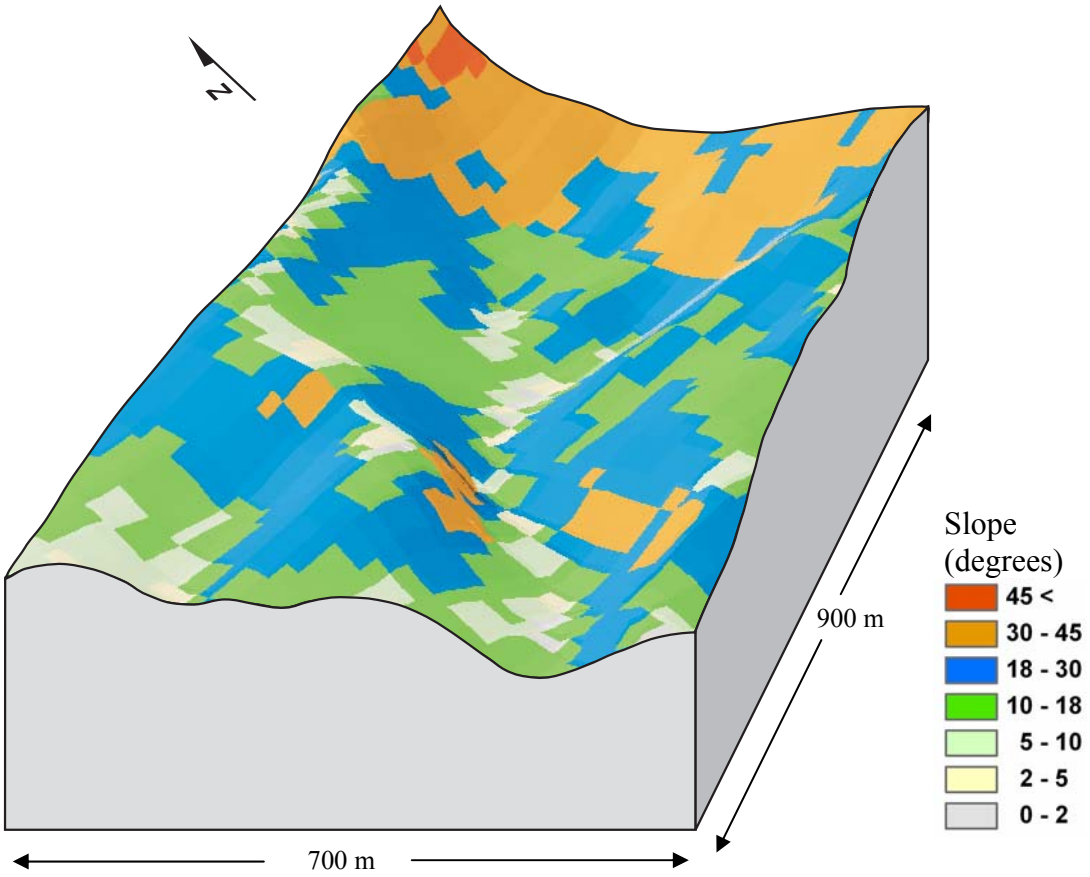


Figure 19: Three-dimensional view the Coromandel subcatchment showing slope angle in degrees (hillshade inclusive).

The TANZ aspect output consists of values in a clockwise direction from 0 to 360 degree with no aspect represented by a no-data value. The Coromandel macro-catchment was reclassified into north (0-45° and 315-360°), east (45-135°), south (135-225°), and west facing aspects (225-315°) to provide a map for easy visual assessment. Figure 20 illustrates these aspects with the north western part of the catchment dominated by southern and eastern aspects. This is conversely true for the south eastern parts of the catchment where northern and eastern aspects dominate. Areas of no aspect are mostly

found on the valley floors or to a lesser extent on ridge lines. The three dimensional graphic of the Coromandel subcatchment (Figure 21) indicates that the majority of this landscape is facing either a southerly or westerly direction with the occasional cell facing north, east, or with no aspect. Note that 'no aspect' indicates the representative cell falls within a flat area.

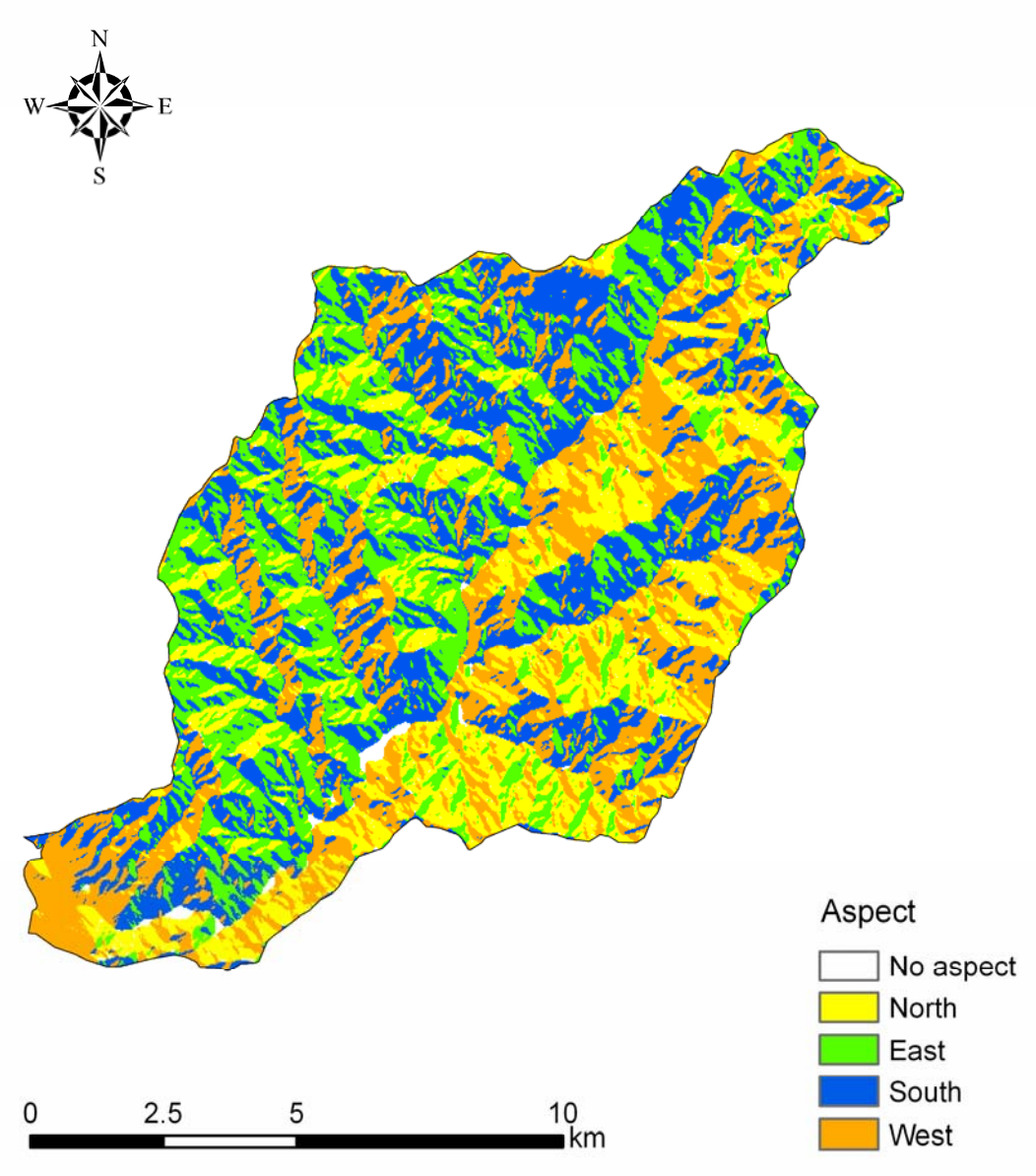


Figure 20: Graphic illustrating aspect for the Coromandel macro-catchment.

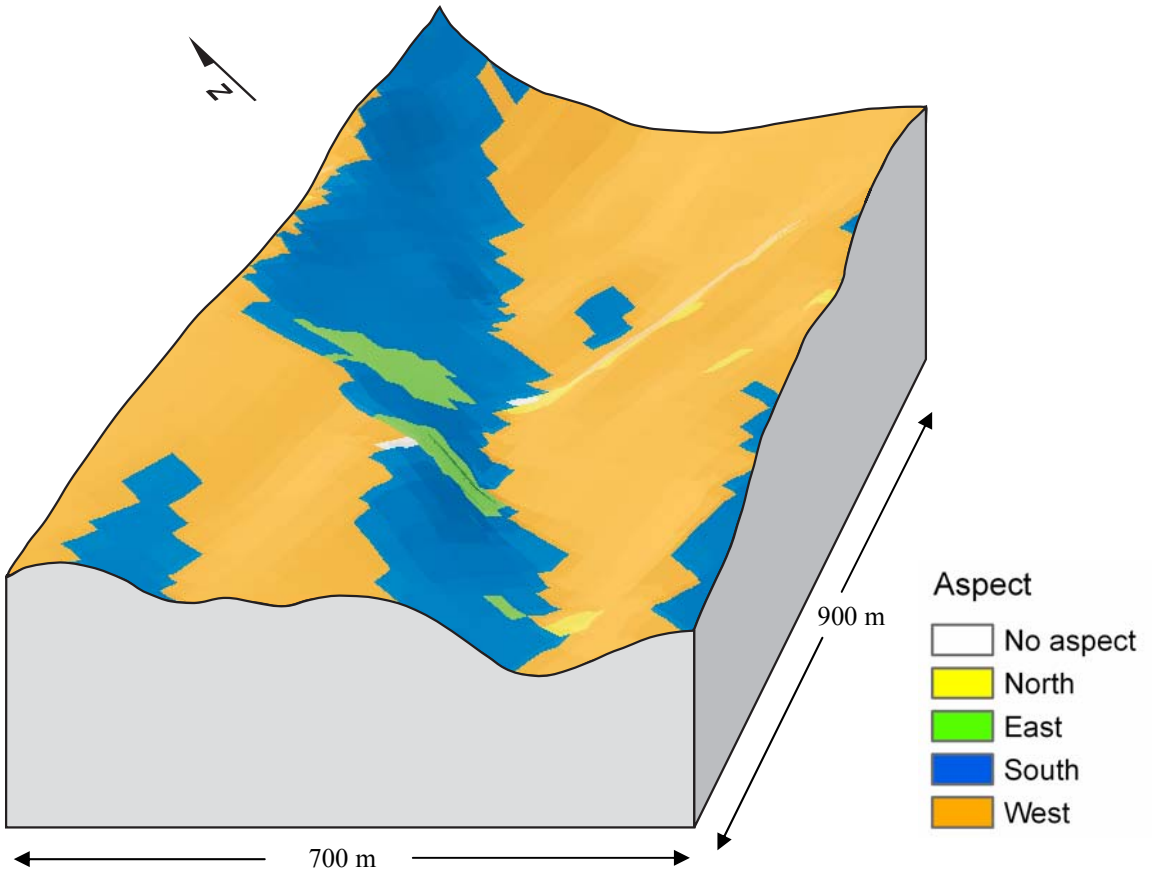


Figure 21: Three dimensional view of the Coromandel subcatchment showing aspect (hillshade inclusive).

Primary flow direction was reclassified into eight primary directions consisting of north (0-22.5° and 337.5-360°), northeast (22.5-67.5°), east (67.5-112.5°), southeast (112.5-157.5°), south (157.5-202.5°), southwest (202.5-247.5°), west (247.5-292.5°), northwest (292.5-337.5°) and one class delineating no primary flow direction (no-data) to provide map visualisation. Figure 22 illustrates these flow directions for the Coromandel macro-catchment. Because the macro-catchment lies in a southwest (low elevation) to northeast direction (elevated elevation), the north-western side of the catchment consists mainly of

southern and eastern flow directions. The opposite is true of the south-eastern side of the catchment where flow directions are mostly in northern or western flow directions. The three dimensional graphic of the Coromandel subcatchment (Figure 23) indicates that the majority of landscape flow direction is southerly (SW, S, or SE). However, west and east flow directions exist with the occasional cell facing northward. Note that 'no flow direction' indicates the cells represent a relatively flat area.

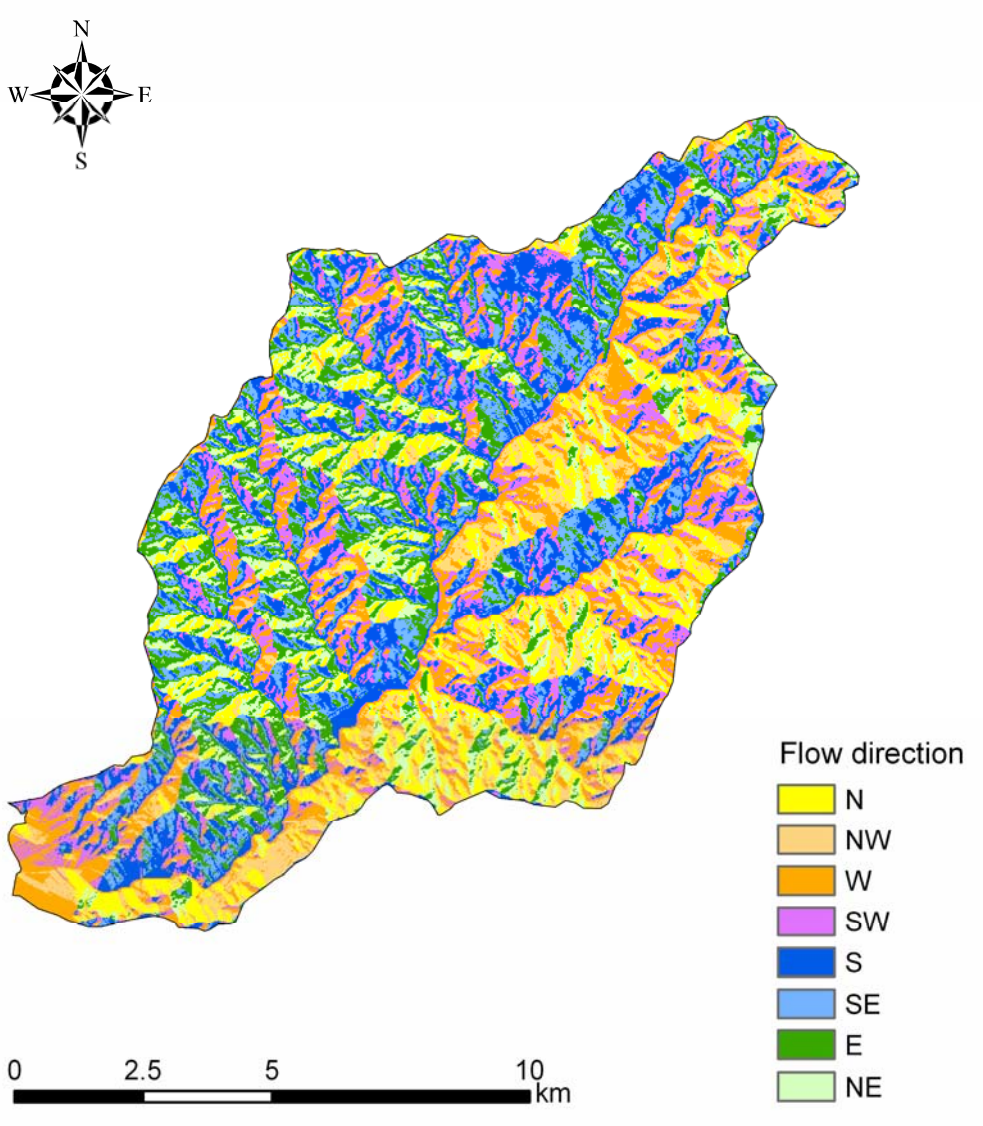


Figure 22: Graphic illustrating primary flow directions for the Coromandel macro-catchment (hillshade inclusive).

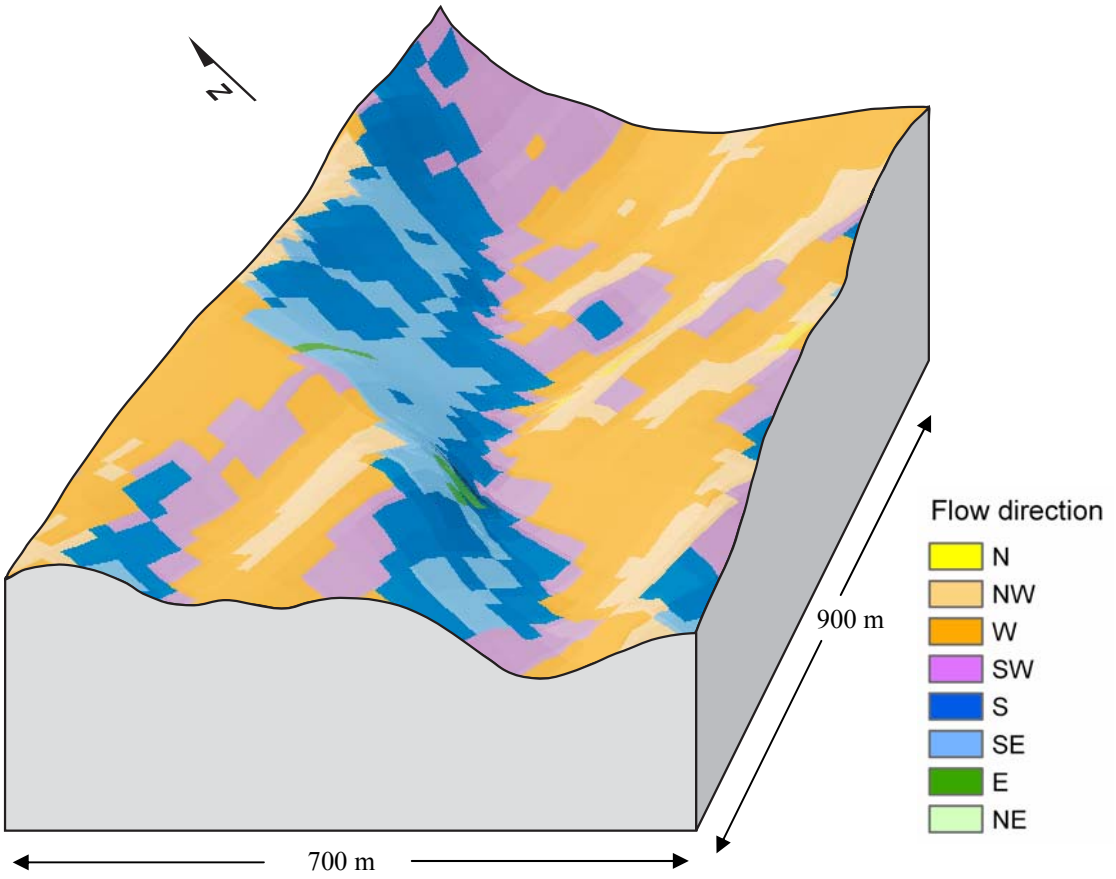


Figure 23: Three-dimensional view of the Coromandel subcatchment showing primary flow directions (hillshade inclusive).

9.2 Curvature

Figures 24, 25, and 26 describe plan, profile and tangential curvature, respectively, for the Coromandel macro-catchment in units of radians per 100 m. The TANZ curvature surfaces provide information describing the tightness of curvature, the inverse of radius of curvature with small curvature corresponding to a large radius of curvature. At the scale of detail and map extent given for the Coromandel macro-catchment, curvature is difficult to assess visually other than in general terms. The three-dimensional graphics shown in Figure 27 highlight the various curvature elements occurring within this subcatchment. Note the similarity in spatial patterns occurring between plan and tangential curvatures because tangential curvature is developed by multiplying plan curvature by the sine of the slope angle (Figures 27a and 27c). The plan and tangential curvature graphics highlight that negative numbers (brown colours) typically occur on convex ridge crests, whereas positive numbers (blue colours) are often found on the concave valley floors. The profile curvature graphic (Figure 27b) illustrates different spatial patterns because it represents the curvature of the surface in the direction of steepest descent.

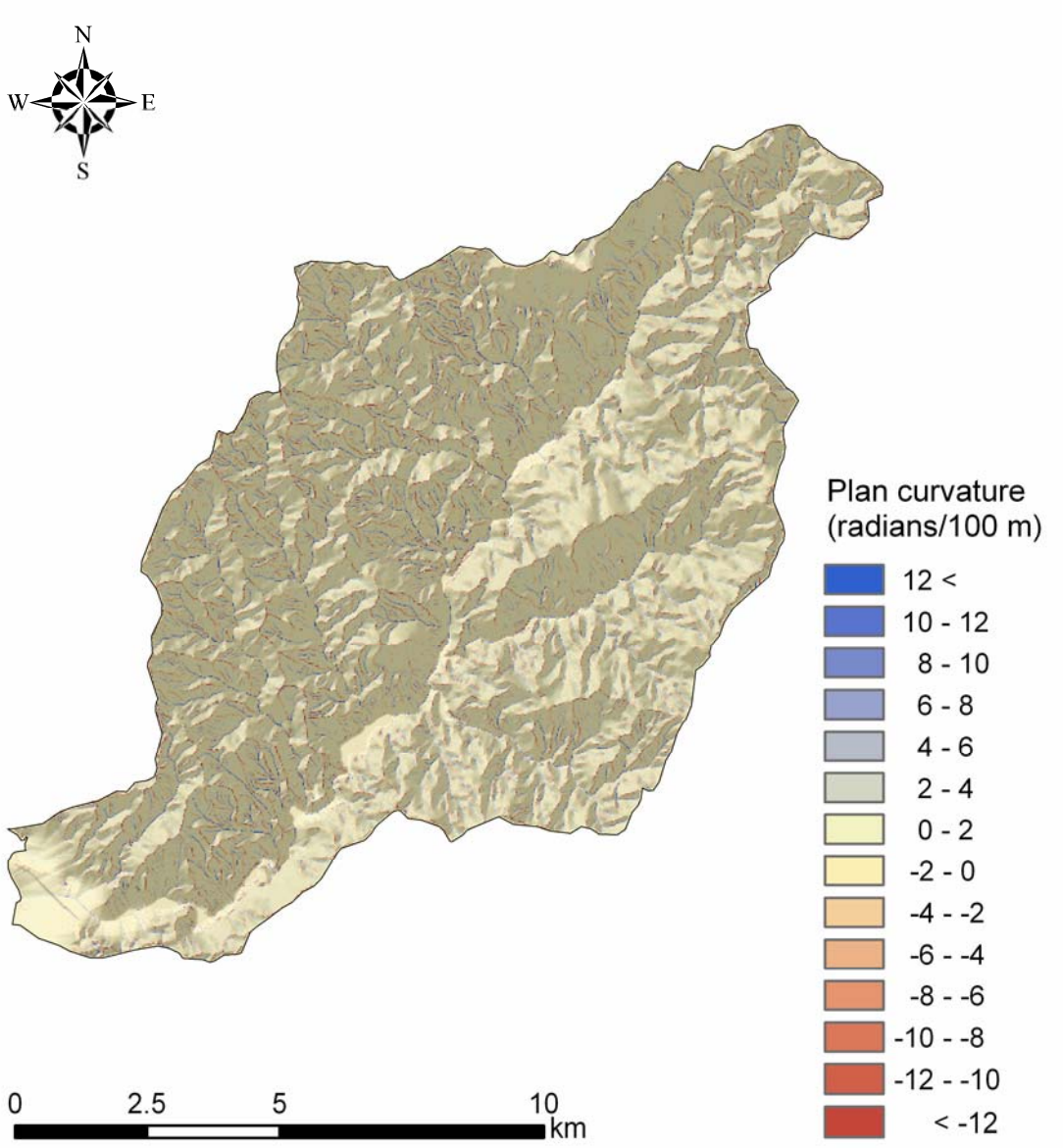


Figure 24: Graphic illustrating plan curvature for the Coromandel macro-catchment (hillshade inclusive).

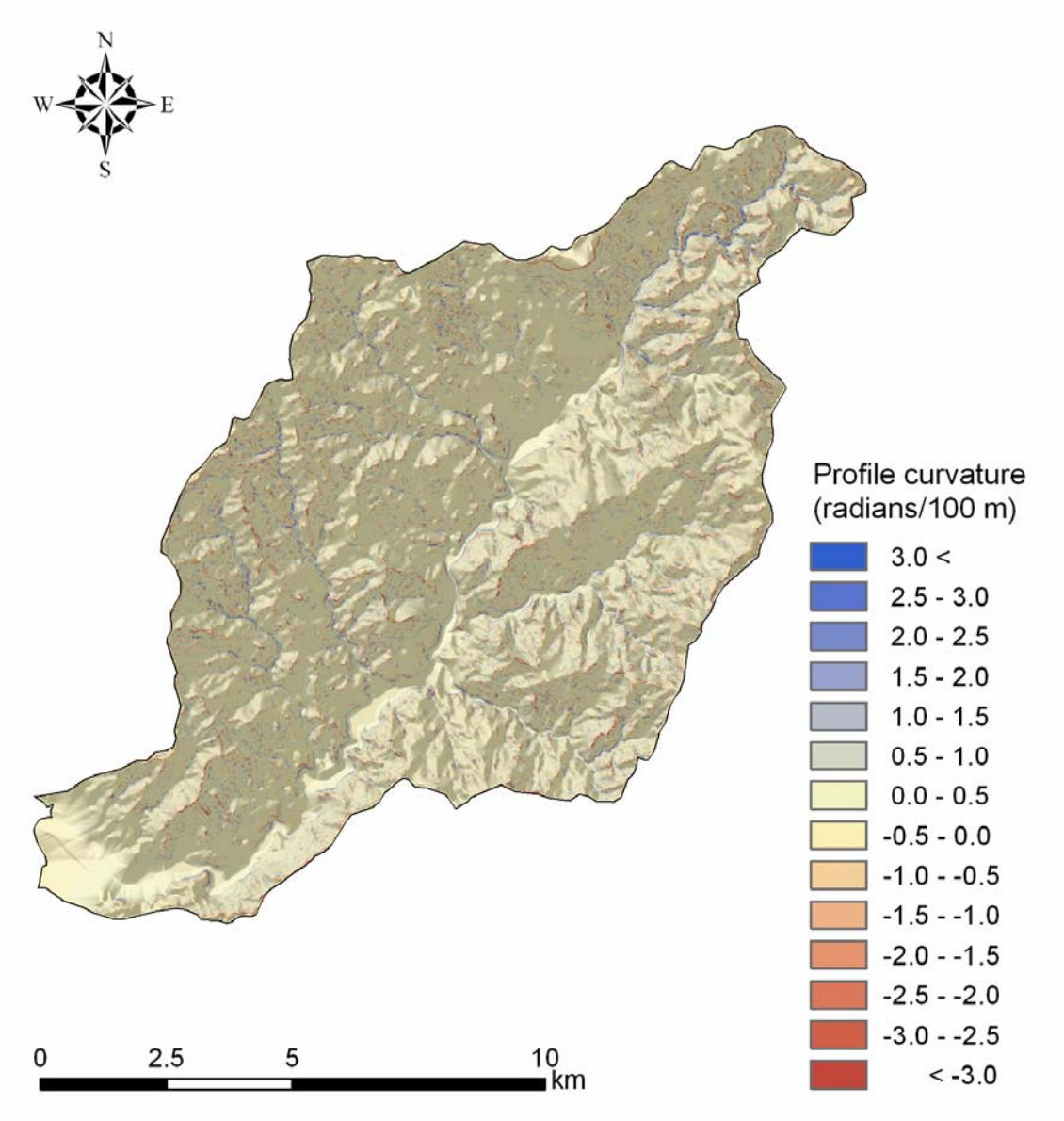


Figure 25: Graphic illustrating profile curvature for the Coromandel macro-catchment (hillshade inclusive).

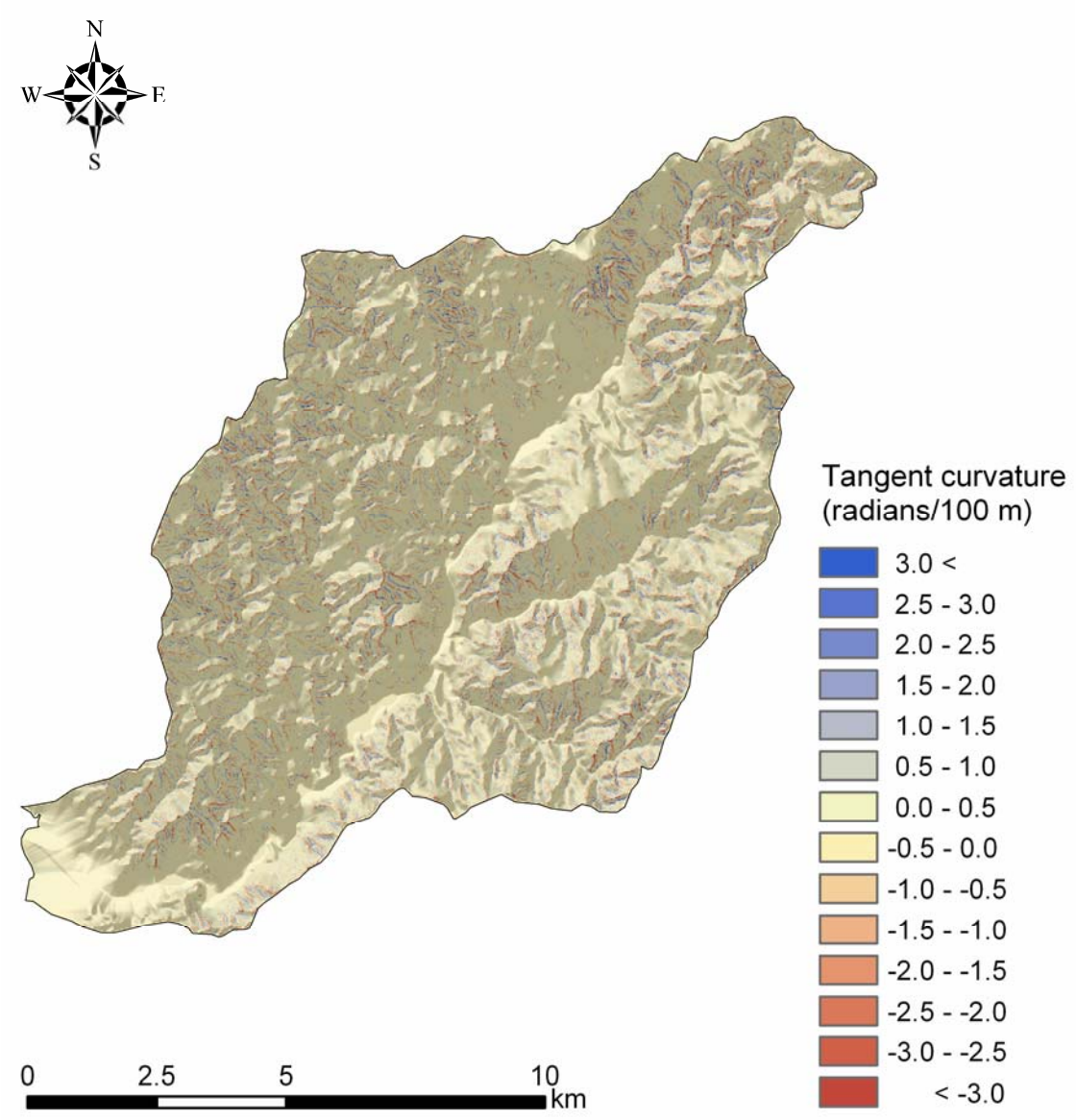


Figure 26: Graphic illustrating tangential curvature for the Coromandel macro-catchment (hillshade inclusive).

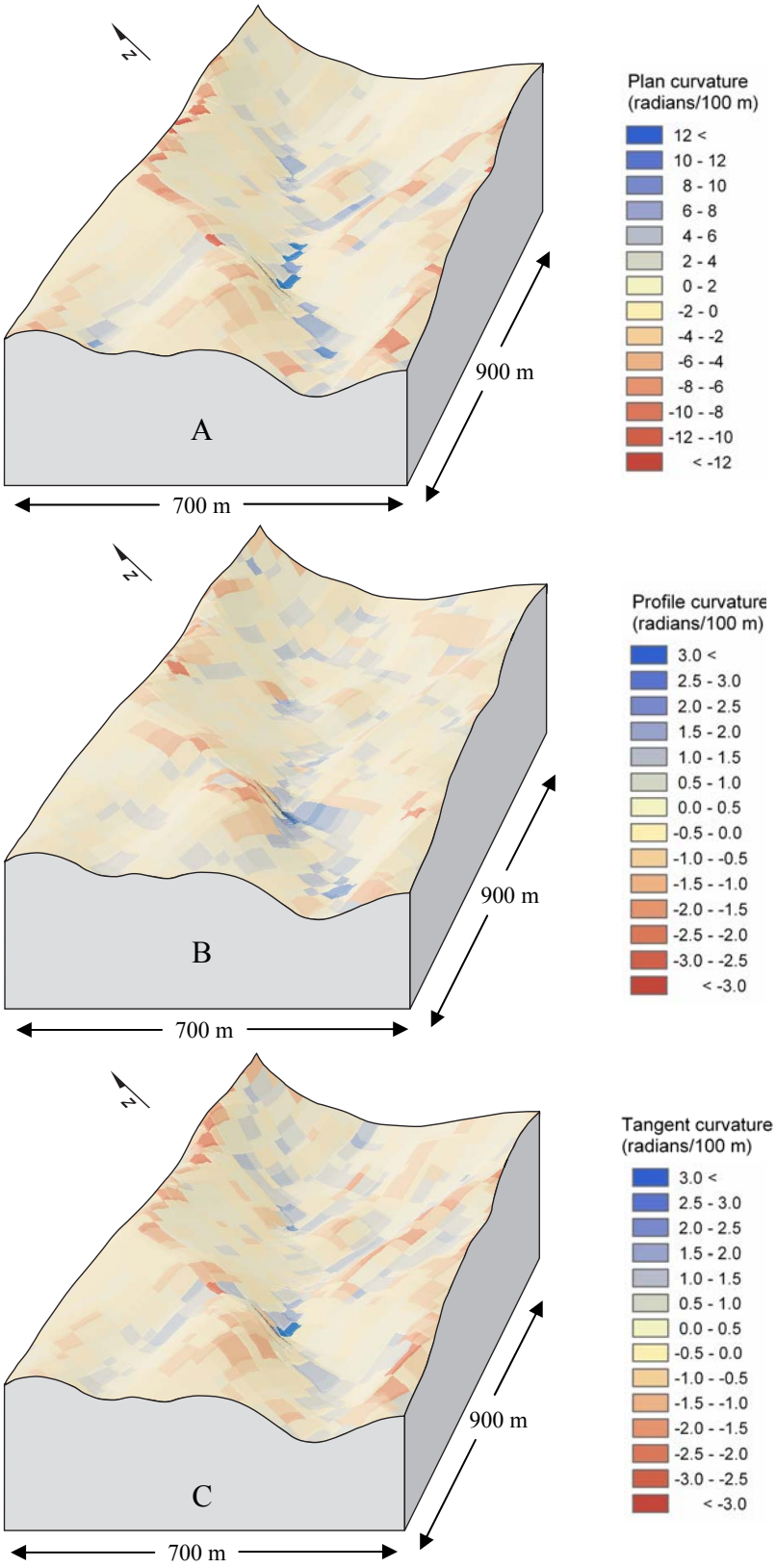


Figure 27: Three dimensional views of the Coromandel subcatchment showing (A) plan curvature, (B) profile curvature, and (C) tangent curvature (hillshade inclusive).

9.3 Upslope contributing area

The upslope contributing area graphics of the Coromandel macro-catchment and subcatchment (Figures 28 and 29) were converted to hectares, allowing visualisation of a standard, comprehensive representative unit (note: TANZ data are in numbers of contributing cells). Upslope contributing area (Figure 28) is difficult to visually assess other than in general terms because of the map extent and scale of detail given for the Coromandel macro-catchment. Nevertheless, the three-dimensional graphic shown in Figure 29 highlights the areas with greater contributing area. Overall, the spatial patterns highlight the increasing upslope contributing area to down-slope cells that are further from the catchment boundaries.

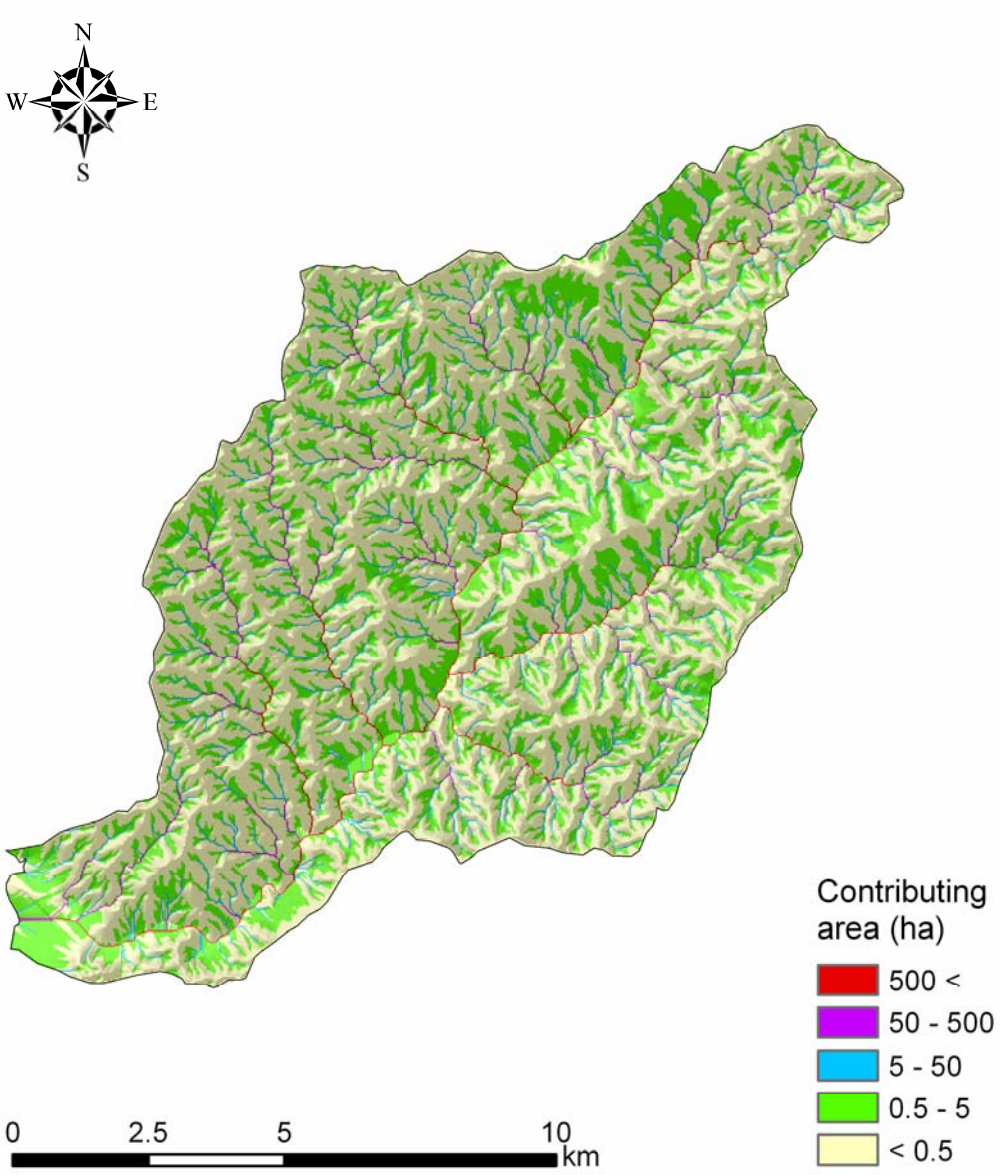


Figure 28: Graphic illustrating upslope contributing area for the Coromandel macro-catchment (hillshade inclusive).

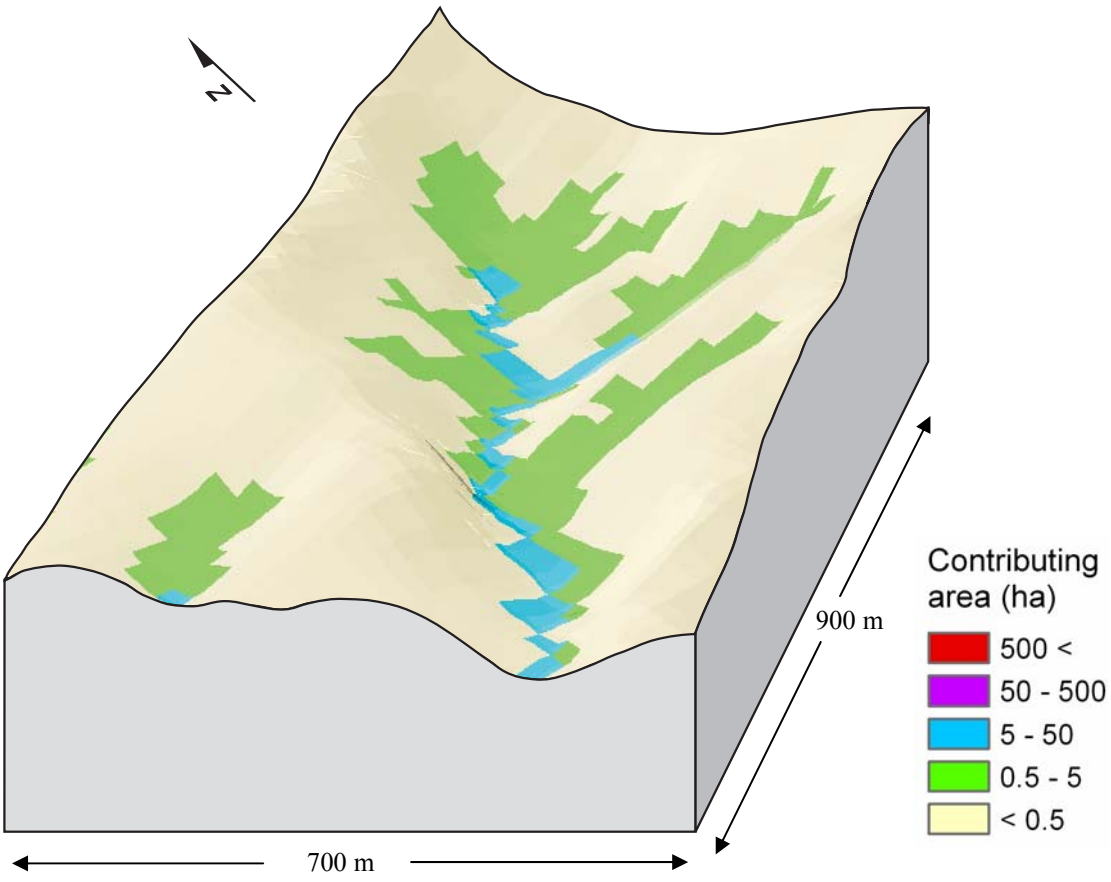


Figure 29: Three dimensional view of the Coromandel subcatchment showing upslope contributing area (hillshade not included in graphic).

9.4 Rate of change of specific catchment area along the flow path

The rate of change of specific catchment area along a flow path (dA_s/ds) indicates a strong spatial pattern with increasing values along defined channels (Figures 30 and 31). In general terms areas closer to catchment boundaries have lower values with values increasing down slope.

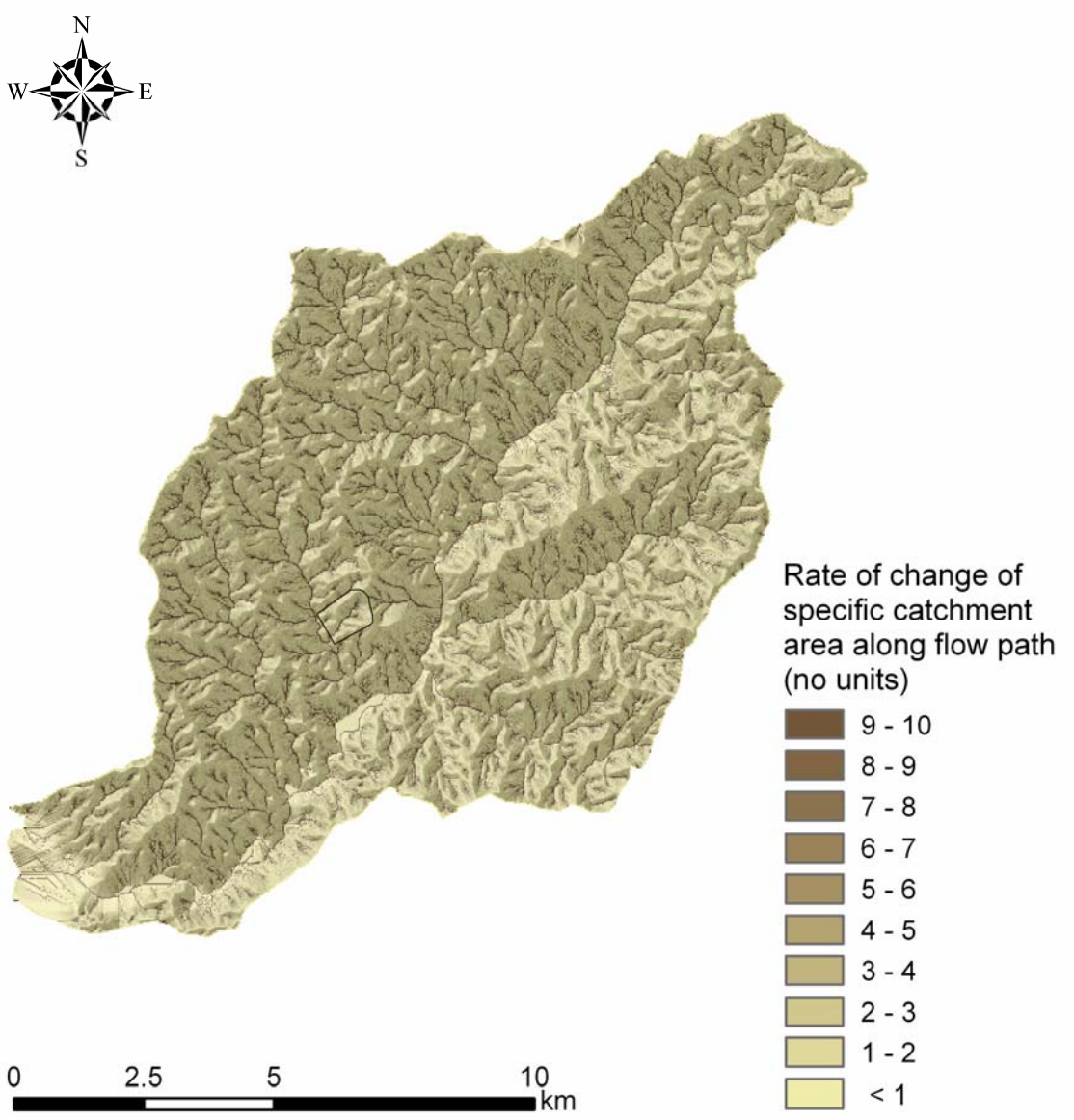


Figure 30: Graphic illustrating the rate of change of specific catchment area along the flow path (*dA*s/*ds*) for the Coromandel macro-catchment (hillshade inclusive).

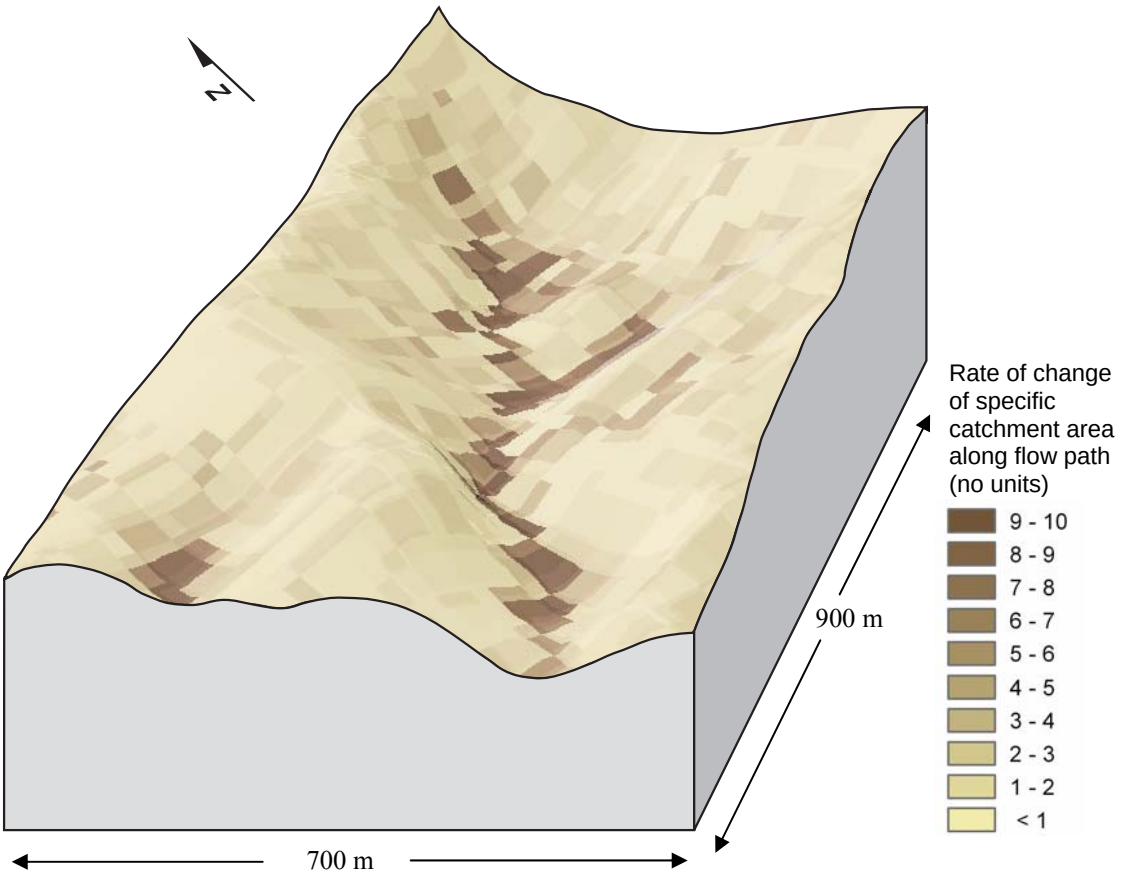


Figure 31: Three-dimensional view of the Coromandel subcatchment showing the rate of change of specific catchment area along the flow path (*d*A*s*/*ds*) (hillshade inclusive).

9.5 Flow width and maximum flow path length

The terrain attribute, flow width in conjunction with upslope contributing area, is used to calculate specific catchment area. Flow widths are illustrated in Figures 32 and 33. It is interesting to note that in general terms the lower lying to flat areas have lower flow width values, whereas the ridgelines tend to have increased flow widths. Maximum flow

path lengths in Figures 34 and 35 show increasing values along the defined channels further from the catchment boundaries.

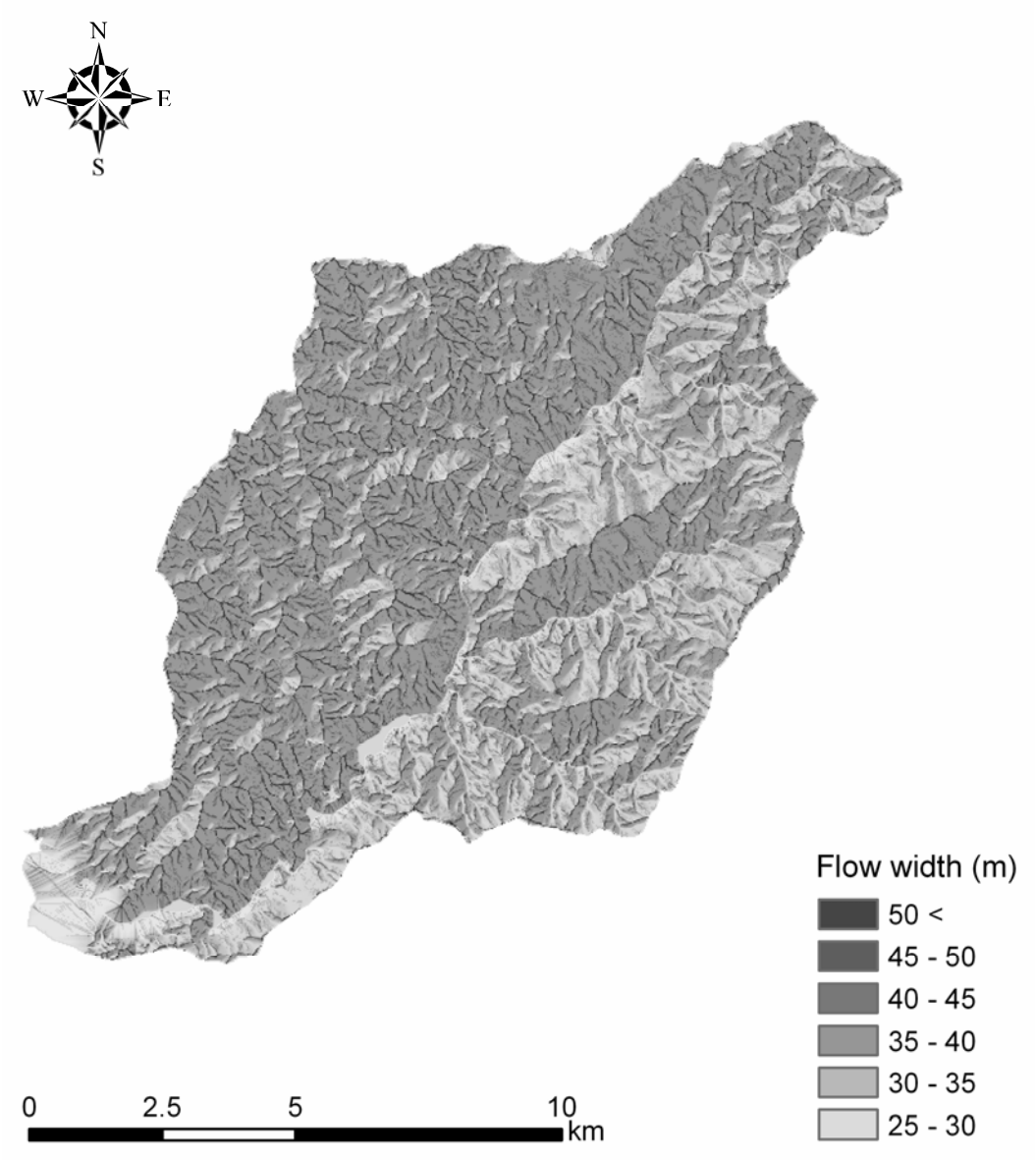


Figure 32: Graphic illustrating the flow width for the Coromandel macro-catchment (hillshade inclusive).

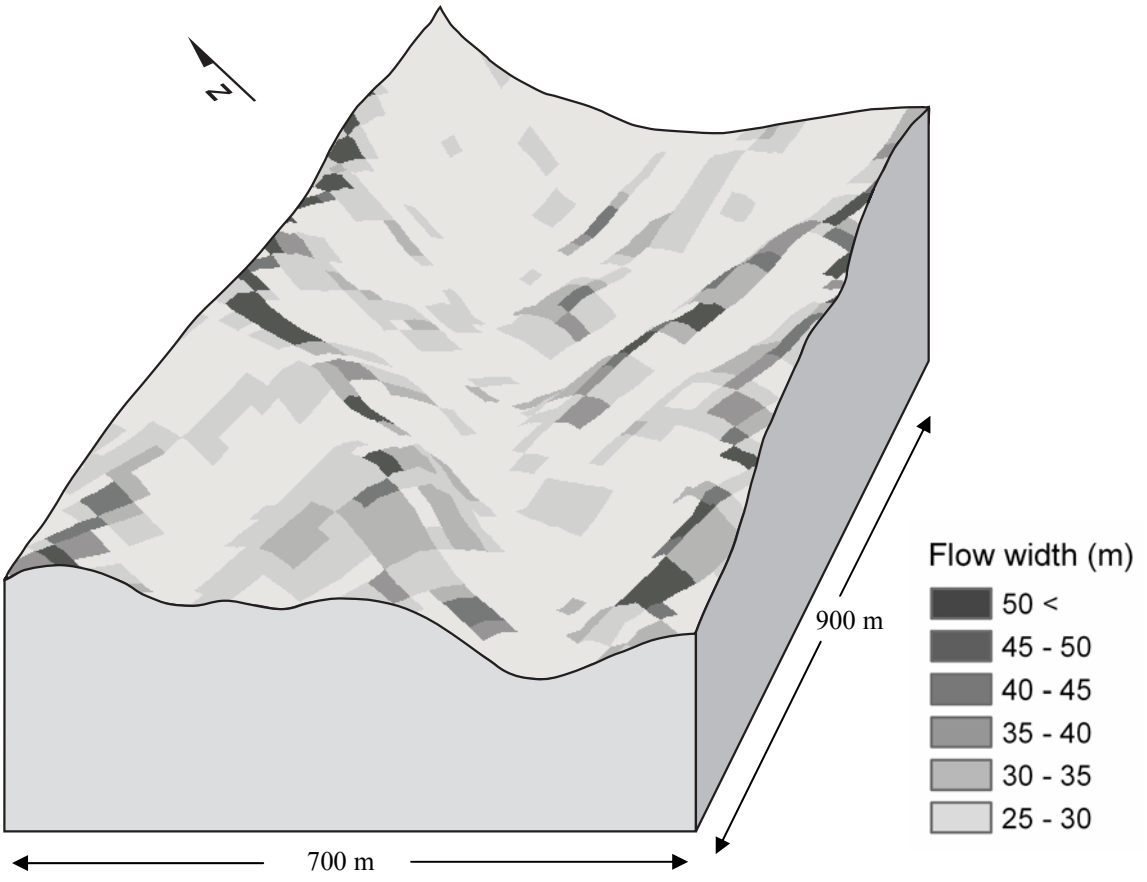


Figure 33: Three-dimensional view of the Coromandel subcatchment showing the flow width.

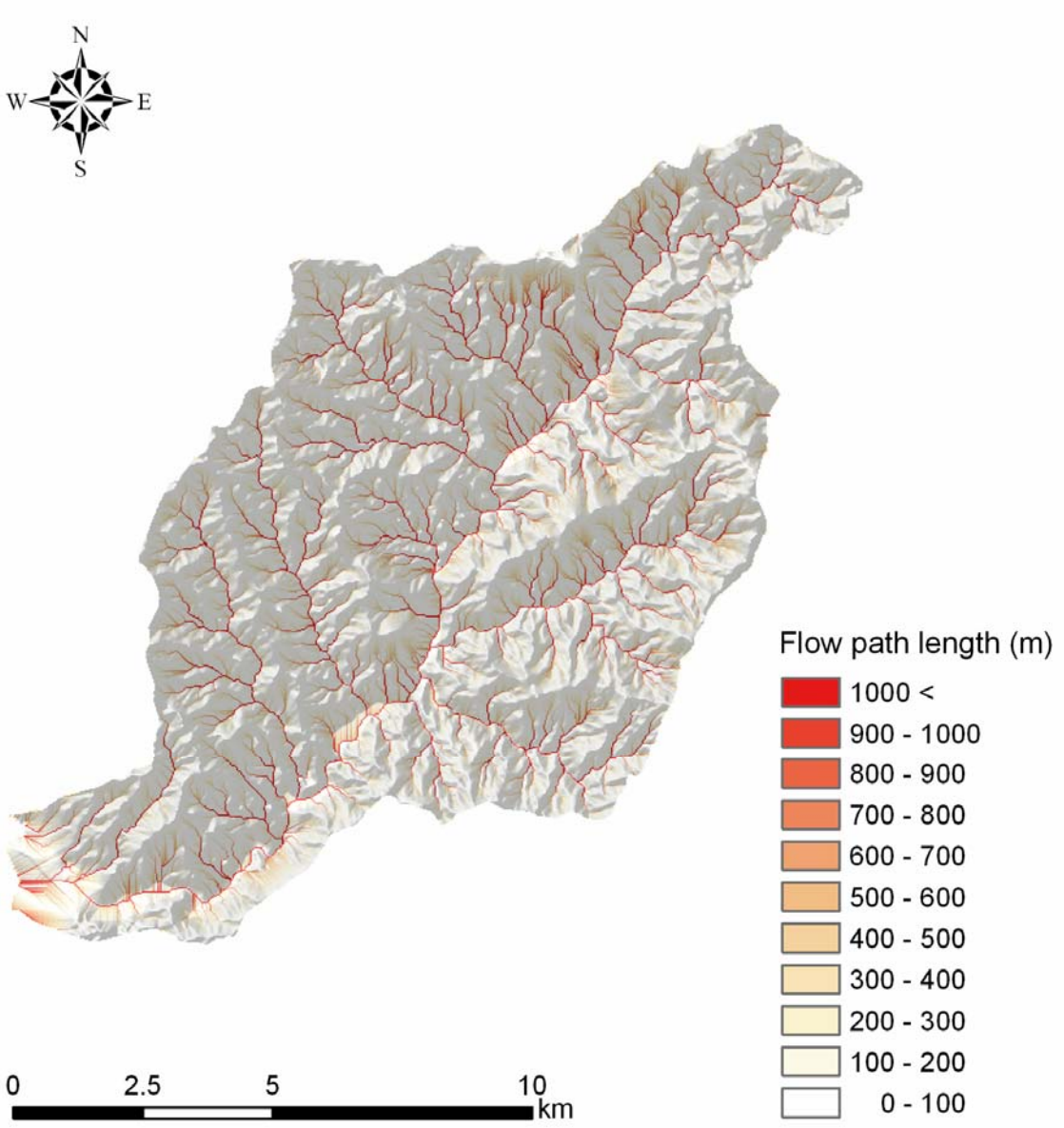


Figure 34: Graphic illustrating the flow-path length for the Coromandel macro-catchment (hillshade inclusive).

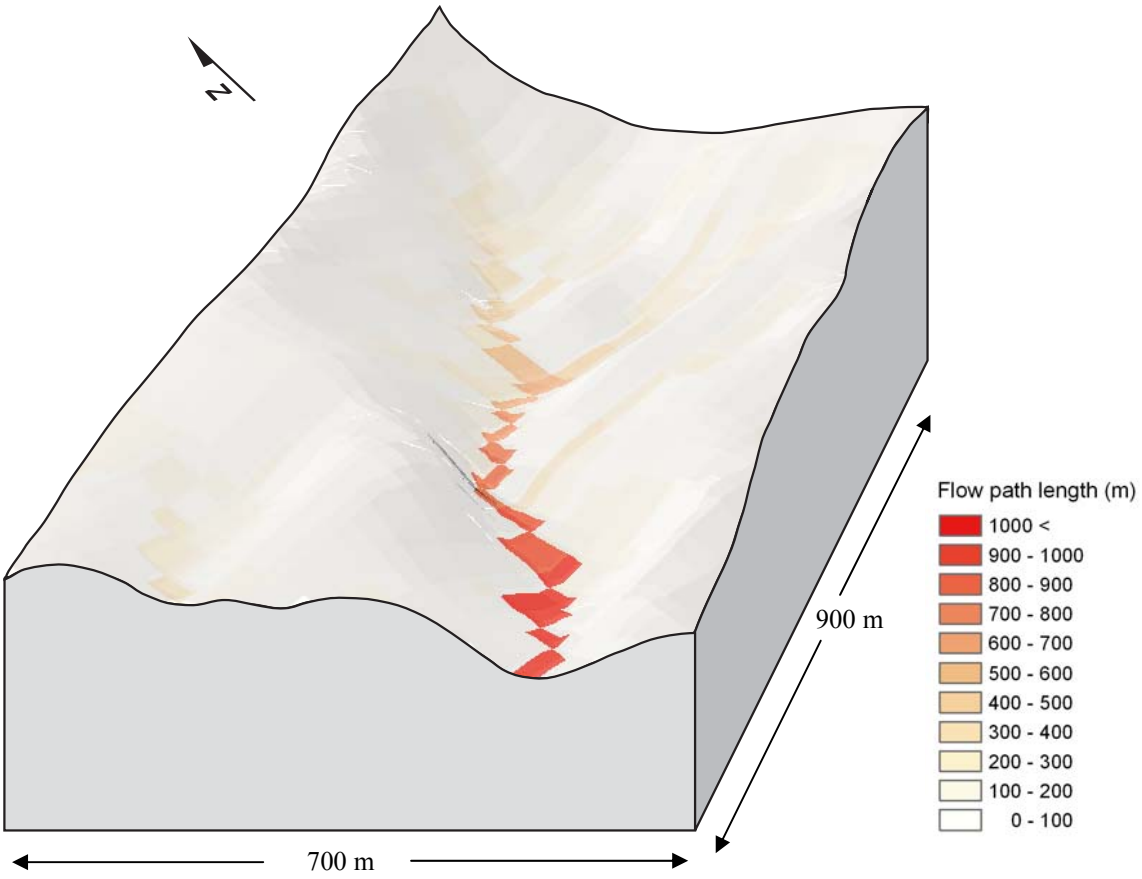


Figure 35: Three-dimensional view of the Coromandel subcatchment showing the flow-path length (hillshade inclusive).

9.6 Topographic wetness index, stream power index, the length-slope factor, and shortwave radiation index

Figures 36 and 37 illustrate the TWI graphically with values from zero to > 20 represented. A value of zero indicates a location where saturation is low, and a value of 20 or more indicates locations were saturation will be high. Generally, high saturation values occur in areas where the specific catchment area (A_s) is large (typically convergent areas in the landscape) and slope angle (β) is relatively small. Conversely, low saturation values are found in areas where specific catchment area is small and slope gradient is relatively elevated.

Stream power index (SPI) is considered a measure of the erosive power of overland flow (Moore et al., 1993a). Moreover, depending on the sediment budget to flowing water within a catchment, the SPI provides an indication of the ability to transport sediment. For example, deposition may occur along a flow path because either slope decreases (reduced velocity) or flow disperses. Figures 38 and 39 illustrate that at locations with profile convexity and tangential concavity (flow acceleration and convergence zones), net erosion occurs, whereas at locations of profile concavity (zones of decreasing flow velocity), net deposition occurs (Wilson and Gallant, 2000).

The length-slope factor (LS factor) developed for TANZ describes the combined effects of slope length (i.e. flow length) and slope gradient, indicating the propensity for erosion to take place, assuming all other universal soil loss equation factors (soil erodibility,

cover and management, rainfall and runoff, and support practice) remain constant. For example, as the length of the slope increases, cumulative runoff increases. Increasing landscape slope and subsequently higher runoff velocities contribute towards increased soil particle transportation (Figures 40 and 41).

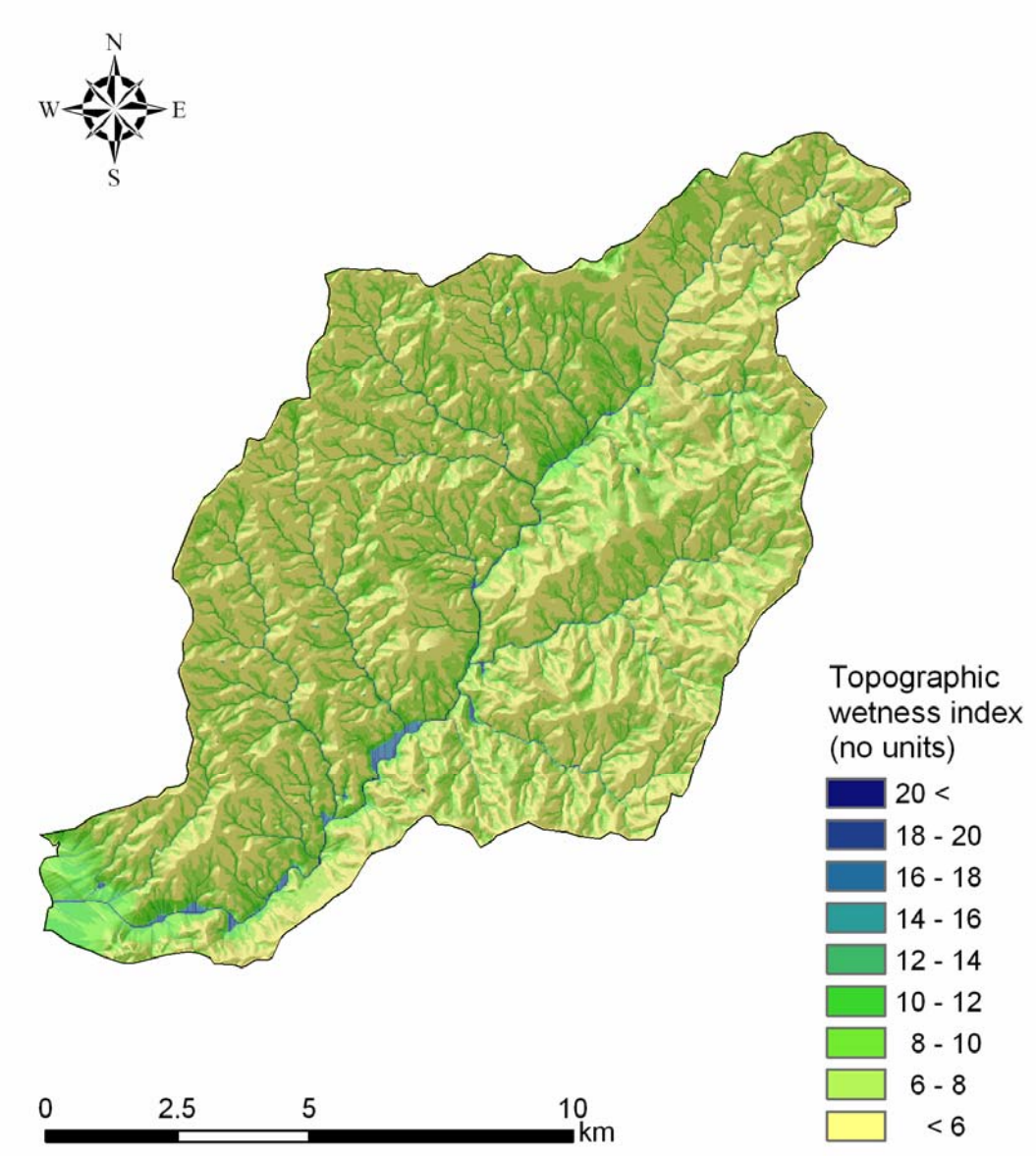


Figure 36: Graphic illustrating the topographic wetness index for the Coromandel macro-catchment (hillshade inclusive).

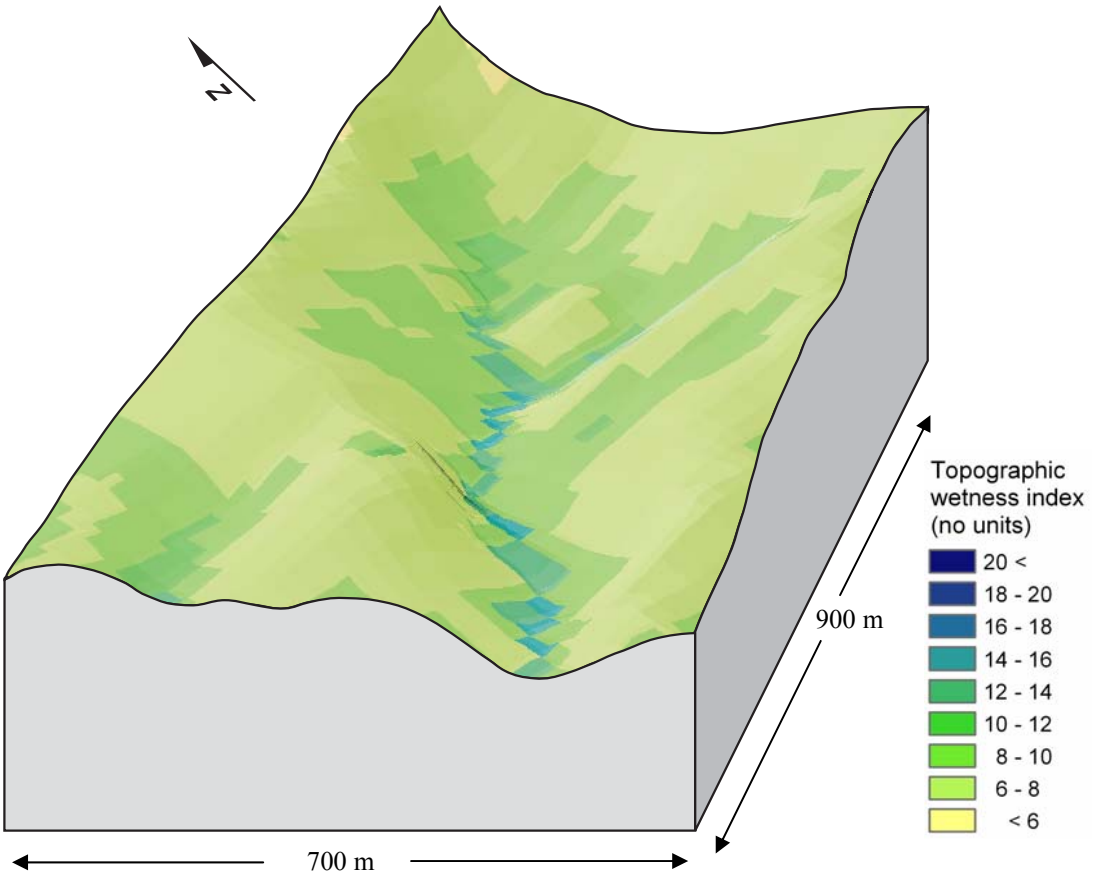


Figure 37: Three-dimensional view of the Coromandel subcatchment showing the topographic wetness index (hillshade inclusive).

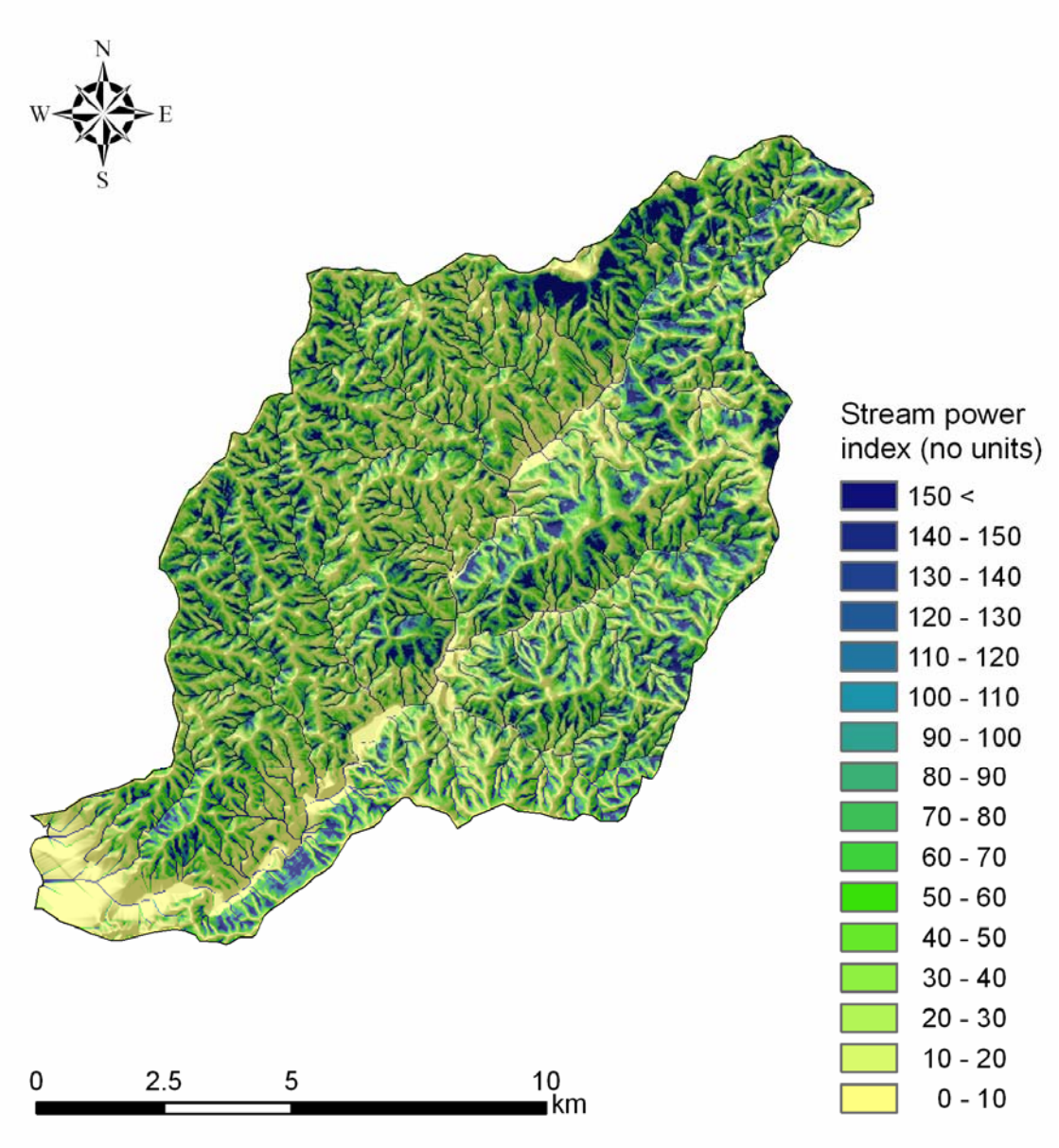


Figure 38: Graphic illustrating the stream power index for the Coromandel macro-catchment (hillshade inclusive).

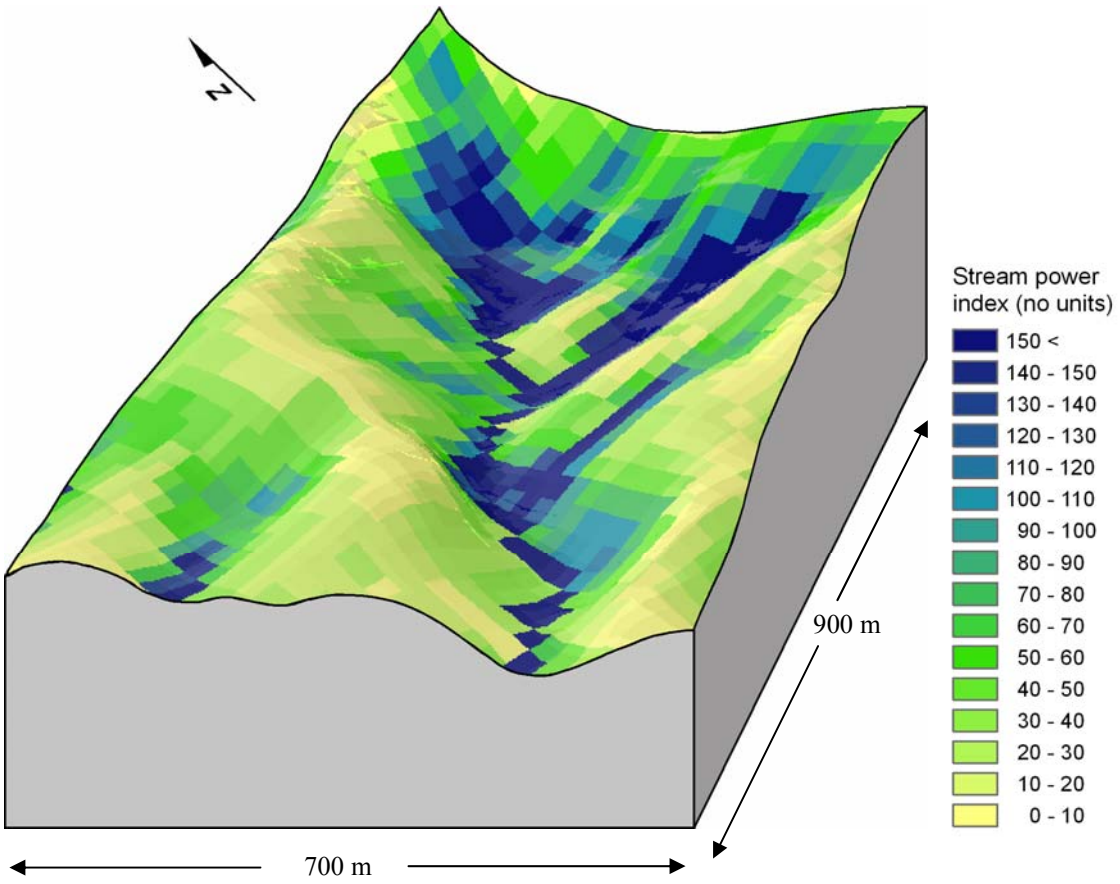


Figure 39: Three-dimensional view of the Coromandel subcatchment showing the stream power index (hillshade inclusive).

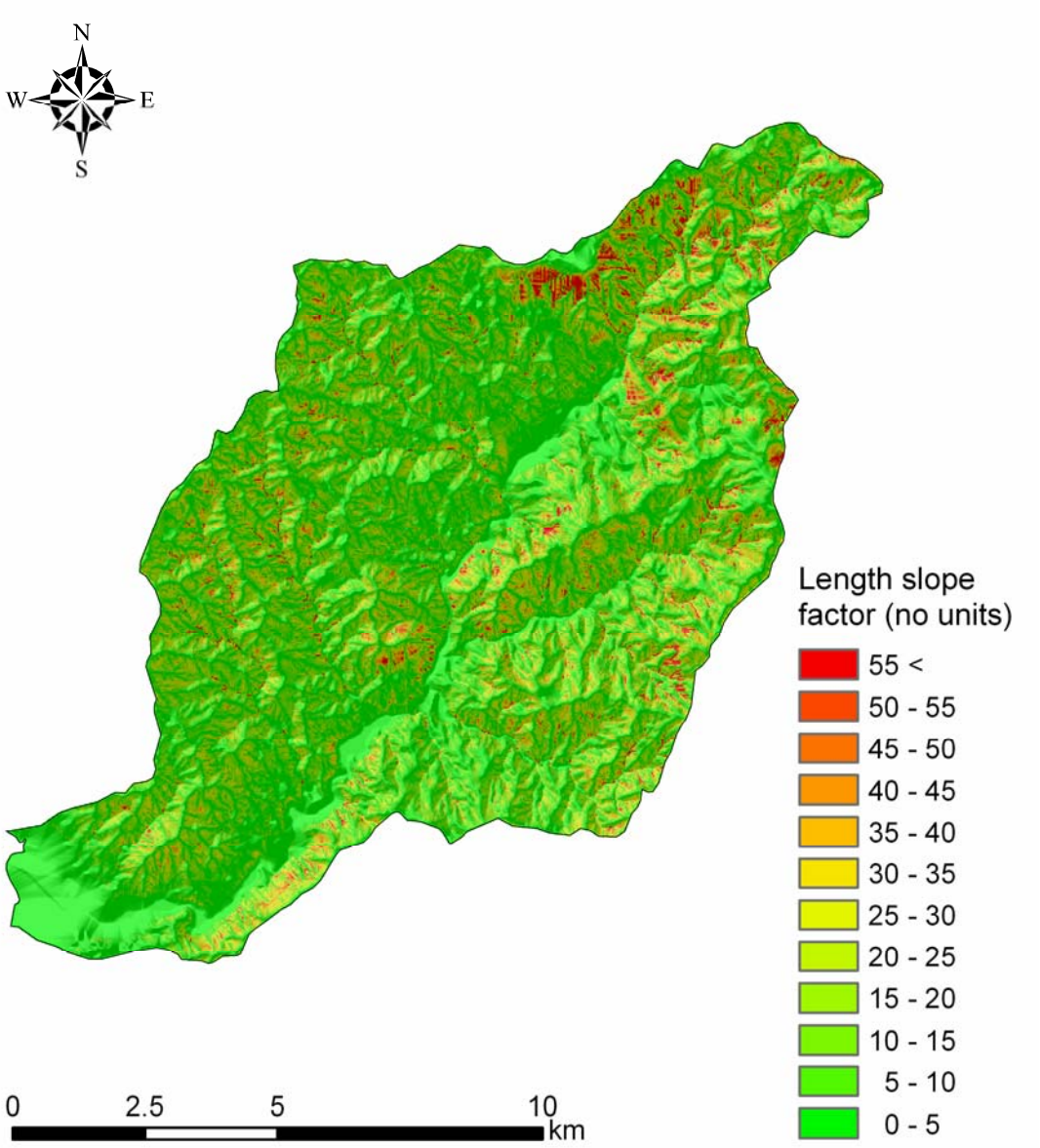


Figure 40: Graphic illustrating the length-slope factor for the Coromandel macro-catchment (hillshade inclusive).

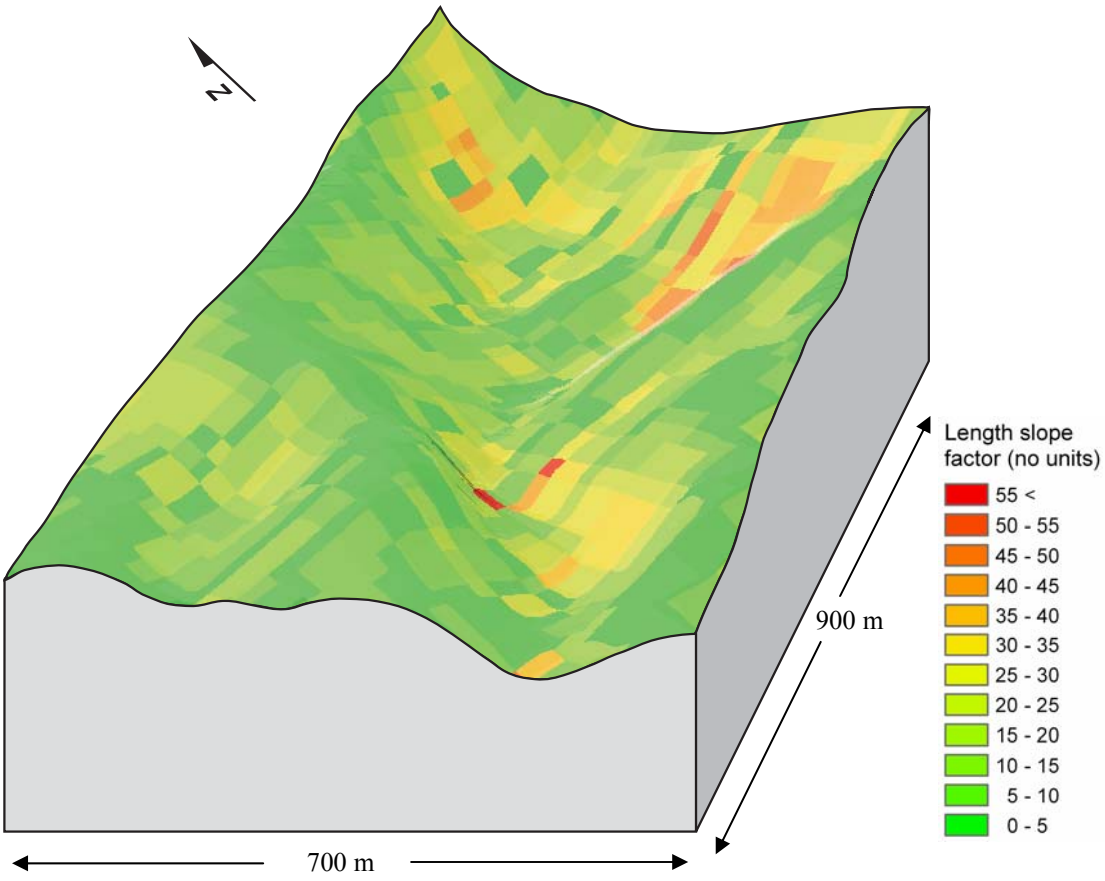


Figure 41: Three-dimensional view of the Coromandel subcatchment showing the length-slope factor (hillshade inclusive).

The TANZ solar radiation graphic (Figure 42) illustrates the mean annual shortwave radiation for New Zealand with values across the range 0 to 30 MJ m^{-2} day^{-1}. Topographically dissected regions such as the mountainous ranges of the Southern Alps clearly show the influence of shadowing terrain with lower index values. Conversely, regions such as the adjacent Canterbury Plains, with relatively flat topography and minimal influence from overshadowing, have some of the highest index values.

The graphics in Figures 43 and 44 use the Coromandel macro-catchment to demonstrate the modelled solar radiation for the months of January, April, July, and October. The graphic shows that January (summer) has the highest Coromandel macro-catchment values. It falls away for the month of April (autumn), decreases further for the month of July (winter), and then increases in October (spring). Also noteworthy at this scale is the influence of terrain overshadowing or reducing insolation on southern aspects. The influence of terrain on the solar radiation modelled values is highlighted at the finer scale within the three-dimensional graphics (Figure 45): cells with higher terrain in the north reduce incoming light to the macro-catchment when the Sun is lower in the sky.

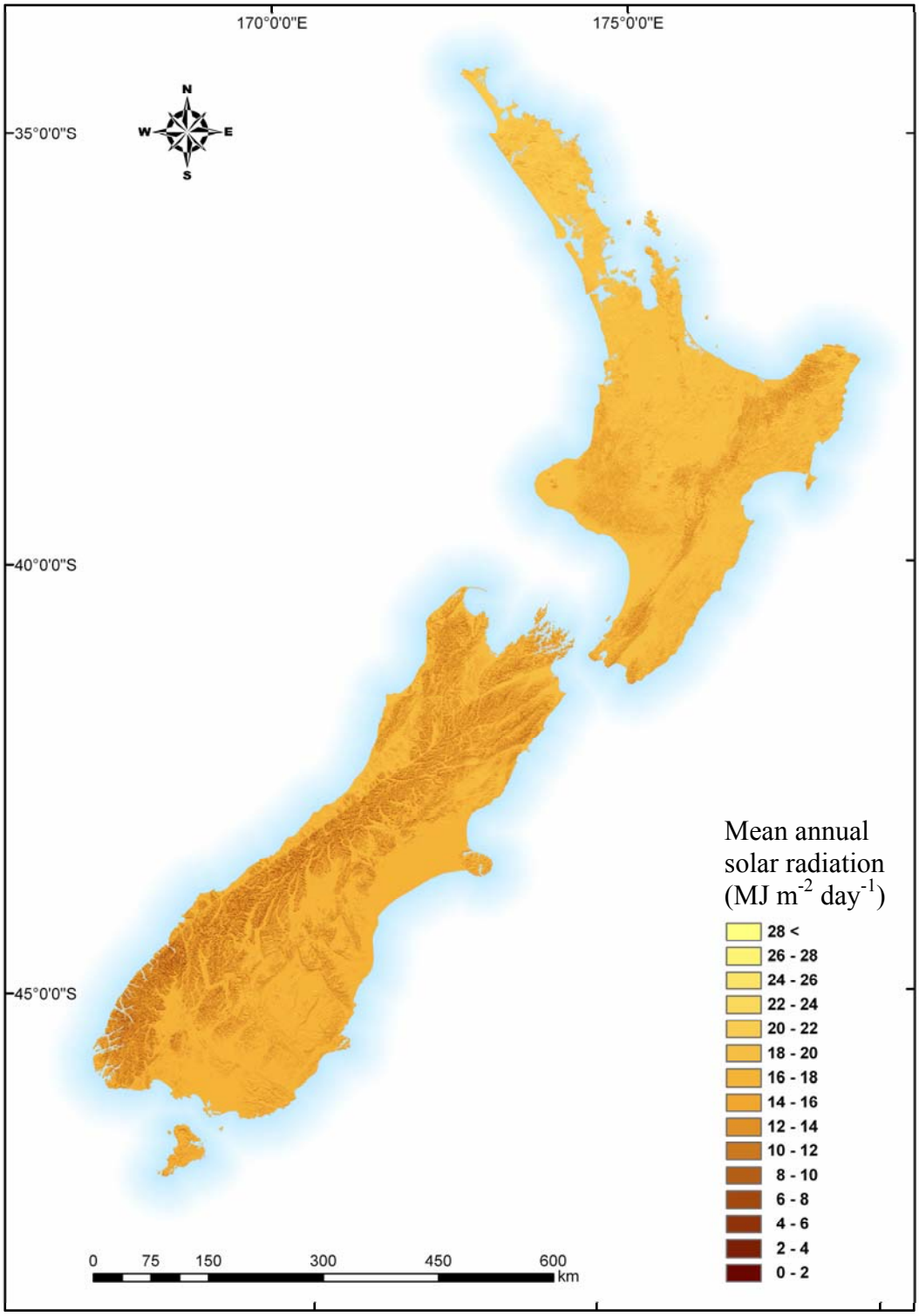


Figure 42: Average annual shortwave radiation for New Zealand (MJ m^{-2} day^{-1}).

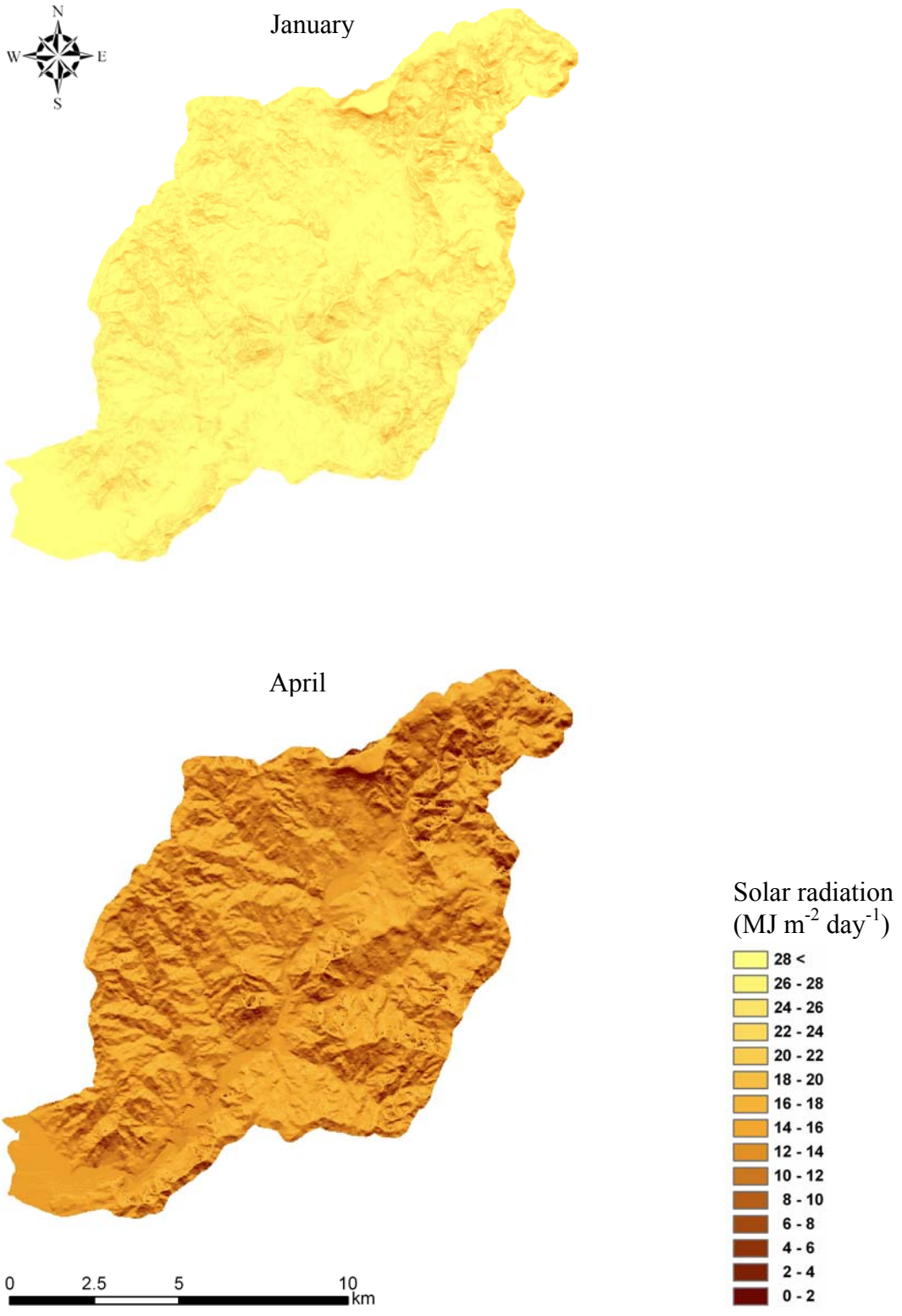


Figure 43: Shortwave radiation (MJ m^{-2} day^{-1}) for January and April for the Coromandel macro-catchment.

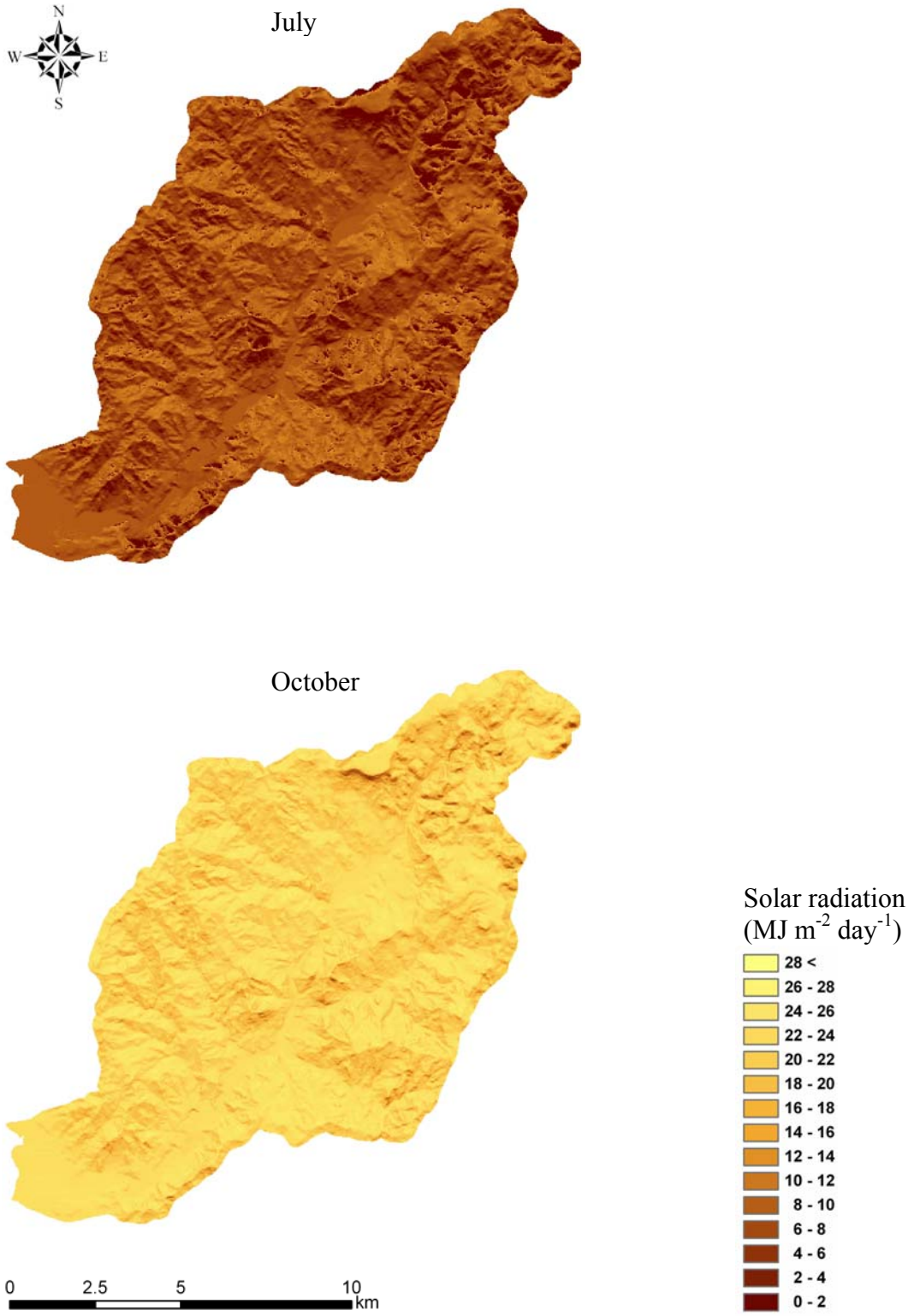


Figure 44: Shortwave radiation (MJ m^{-2} day^{-1}) for July and October for the Coromandel macro-catchment.

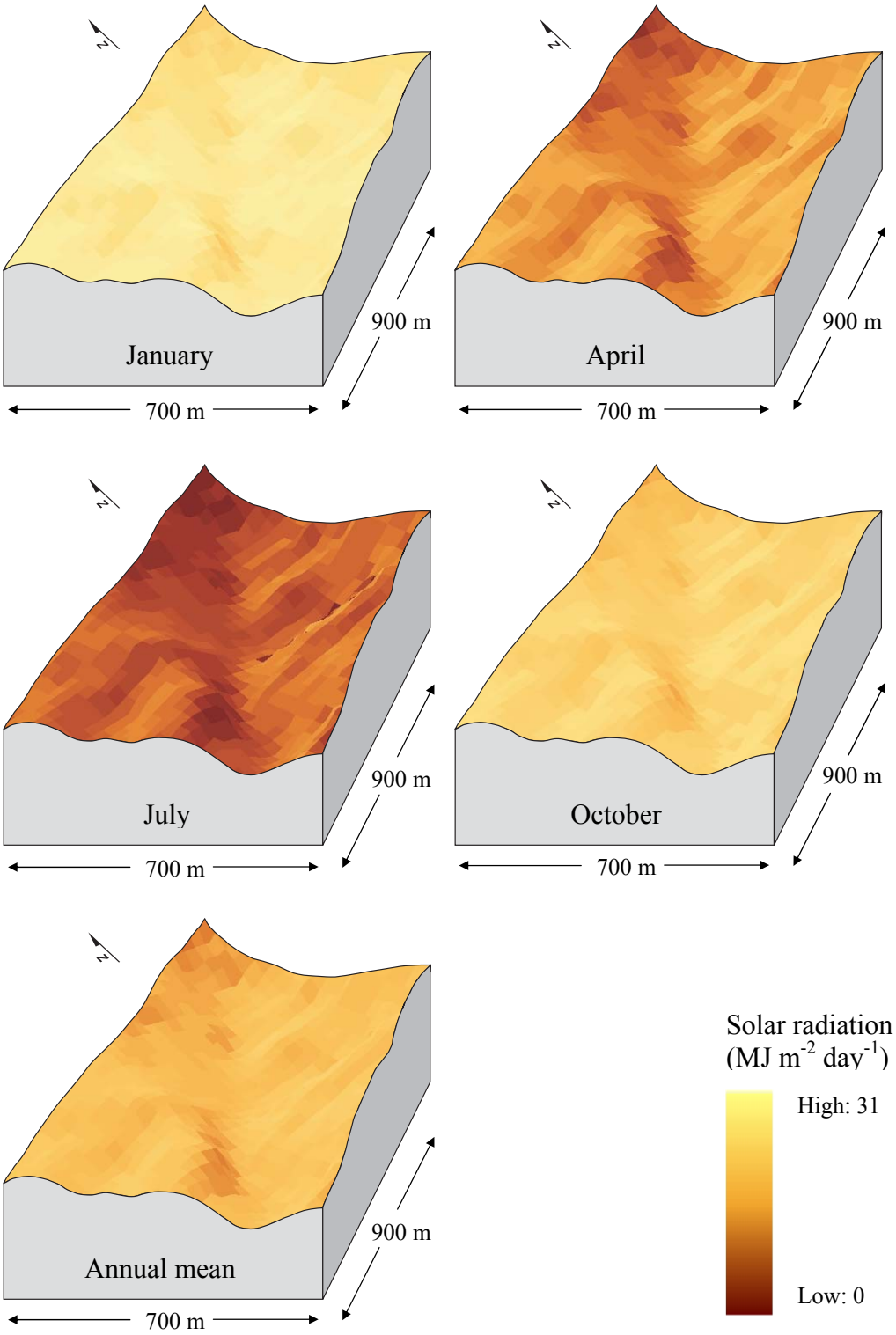


Figure 45: Three-dimensional views of the Coromandel subcatchment showing the shortwave radiation (MJ m^{-2} day^{-1}) for the months of January, April, July and October.

10. SUMMARY AND CONCLUSIONS

This bulletin defines the protocols and programming processes involved in developing national-scale primary and secondary terrain attributes (TANZ), directly derived from a New Zealand 25 m-resolution DEM. New Zealand was divided into a series of macro-catchments to enable passing DEM data through the TAPES-G programme for the development of TANZ. The resulting primary and secondary terrain attributes were merged together, providing gridded data for the North Island and South Island. They include slope, aspect, contributing area, plan and profile curvature, topographic wetness index, and the stream power index. The length-slope factor, 12 monthly and one annual shortwave radiation index surfaces were also derived using primary terrain attributes.

Typically, terrain attributes are created for site-specific projects, remaining piecemeal with the ownership firmly held within the organisation that created the surfaces. A further data dissemination barrier remains because of the high cost of surface development. The creation of terrain attribute surfaces like other environmental surfaces requires large development resources, including both human (time and technical skill) and computing resources. Should these types of data be made available, those organisations that might purchase these surfaces do so with careful consideration because of the major financial outlay.

The advance our research brings is not only the development of comprehensive national-scale terrain attributes across New Zealand, but also the development of surfaces with the

highest possible degree of accuracy. This accuracy was achieved by implementing the most advanced, published and validated algorithms in the development of TANZ. More specifically, we developed processes and protocols for managing large datasets by dividing the input data into smaller, manageable macro-catchments. Macro-catchments were of optimal size to allow the computer to manage and process these data at a resolution of 25 m. Another reason for the development of these naturally draining macro-catchments compared to a simple division of New Zealand into a series of rectangular polygons was to minimise edge effects. Edge effects occur where the edge of the DEM does not coincide with the boundary of the catchment. Edge effects impact on contributing area values and to a lesser degree on shading. Our terrain attribute development protocol negated this problem by calculating beyond the boundary edge, thus ensuring correct contributing area calculations were performed. The unwanted surplus data beyond the boundary were removed before finally merging individual macro-catchments into national-scale surfaces.

The final consideration for developing national scale terrain attributes was the ability of TAPES-G to model both divergent and convergent flow within a landscape by switching between algorithms during modelling. The choice of algorithms is important because it impacts on the calculations of some terrain attributes including upslope contributing area, specific catchment area, and stream power index. Our project used the D8 and FD8 algorithms with a maximum cross-grading area of 25,000 m^2. In upland areas the FD8 method allows flow to be distributed to multiple neighbouring cells allowing flow

convergence to be modelled, whereas the D8 method enables flow divergence to be captured in valley bottoms.

This work was undertaken towards fulfilment of a Ph.D. Our overriding aim in developing TANZ has been to create high-quality national-scale surfaces useful in further development and spatial interpolation of *Pinus radiata* productivity data across New Zealand. The TANZ will also be useful in other modelling environments including hydrology, geomorphology, pedology, and ecology. Further applications potentially include resource management, environmental modelling, GIS, pedometrics, and geostatistics. This bulletin provides comprehensive technical information on the processes and protocols used in developing TANZ. For more information regarding TANZ development or access to the derived surfaces please contact the authors or Scion.

11. ACKNOWLEDGEMENTS

We gratefully thank John Gallant for his comments and very helpful suggestions that assisted us in overcoming many of the hurdles associated with this project. We also thank John Wilson for his comments and suggestions in the development of TANZ. David Palmer wishes to acknowledge and thank the Foundation for Research Science and Technology (FRST) for funding through a Top Achiever Scholarship, without which this project would not have been possible. Barbara Höck also thanks FRST for their funding

(Protecting and enhancing the environment through forestry CO4X0304). We are grateful to comments from two reviewers, and the editor, that improved this bulletin.

12. REFERENCES

Barringer, J.R.F., Pairman, D., McNeill, S.J. 2002, Development of a high-resolution digital elevation model for New Zealand. Landcare Research Contract Report LC0102/170 (unpublished).

Blaszcynski, J.S. 1997. Landform characterization with Geographic Information Systems. Photogrammetric Engineering and Remote Sensing, 63: 183-191.

Bian, L. 1997. Multiscale nature of spatial data in scaling up environmental models. In: Quattrochi, D.A., Goodchild, M.F. (eds), Scale in remote sensing and GIS. Boca Raton, Florida, USA, Lewis Publishers: pp. 13-26.

Burrough, P.A., Wilson, J.P., van Gaans, P.F.M., Hansen, A.J. 2001. Fuzzy k-means classification of digital elevation models as an aid to forest mapping in the Greater Yellowstone Area, USA. Landscape Ecology, 16: 523-46.

Claessens, L. 2005. Modelling Landslide Dynamics in Forested Landscapes: Addressing landscape evolution, landslide soil redistribution and vegetation patterns in the Waitakere Ranges, west Auckland, New Zealand. PhD thesis, Wageningen University, Netherlands, 143pp.

Costa-Cabral, M.C., Burges, S.J. 1994. Digital elevation model networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal areas. Water Resources Research, 30: 1681-1692.

De Roo, A.P.J. 1998. Modelling runoff and sediment transport in catchments using GIS. Hydrological Processes, 12: 905-922.

Dikau, R. 1989. The application of a digital relief model to landform analysis. In: Raper, J.F. (eds), Three dimensional Applications in Geographical Information Systems. London, Taylor and Francis: pp. 51-77.

Fairfield, J., Leymarie, P. 1991. Drainage networks from grid digital elevation models. Water Resources Research, 27: 709-717.

Fitzsimons, S. 2001. Denudation, weathering, and slope development. In: Sturman, A., Spronken-Smith, R. (eds), The Physical Environment – a New Zealand Perspective. Oxford University Press, Melbourne: pp. 240-253.

Florinsky, I.V. 1998. Accuracy of local topographic variables derived from digital elevation models. International Journal of Geographical Information Science, 12: 47-61.

Florinsky, I.V., Eilers, R.G., Manning, G.R., Fuller, L.G. 2002. Prediction of soil properties by digital terrain modelling. Environmental Modelling and Software, 17: 295-311.

Franklin, J., McCullough, P, Gray, C. 2000. Terrain variables used for predictive mapping of vegetation communities in Southern California. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 331-353.

Freeman, G.T. 1991. Calculating catchment area with divergent flow based on a regular grid. Computers and Geosciences, 17: 413-422.

Fried, J.S., Brown, D.G., Zweifler, M.O., Gold, M.A. 2000. Mapping contributing areas for stormwater discharge to streams using terrain analysis. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 183-203.

Gallant, J.C., Wilson, J.P. 1996. TAPES-G: A grid-based terrain analysis program for the environmental sciences. Computers and Geosciences, 22: 713-722.

Gallant, J.C., Wilson, J.P. 2000. Primary topographic attributes. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 51-85.

Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L., Holmes, K. 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Science Society of America Journal, 64: 2046-2056.

Gessler, P.E., Moore, I.D., McKenzie, N.J., Ryan, P.J. 1995. Soil-landscape modelling and spatial prediction of soil attributes. International Journal of Geographical Information Systems, 9: 421-432.

Griffin, M.L., Beasley, D.B., Fletcher, J.J., Foster, G.R. 1988. Estimating soil loss on topographically non-uniform field and farm units. Journal of Soil and Water Conservation, 43: 326-331.

Hall, G.F., Olson, C.G. 1991. Predicting variability of soils from landscape models. In: Mausbach, M.J., Wilding, L.P. (eds), Spatial Variabilities of Soils and Landforms. SSSA Special Publication Number 28. Soil Science Society of America, Madison, pp. 9-24.

Henderson, B., Bui, E., Moran, C., Simon, D., Carlile, P. 2001. ASRIS: Continental-scale soil property predictions from point data. CSIRO Land and Water Technical Report 28/01.

Hewitt, A.E., Lilburne, L.R. 2003. Effect of scale on the information content of soil maps. New Zealand Soil News, 51: 78-81.

Hutchinson, M.F., Gallant, J.C. 2000. Digital elevation models and representation of terrain shape. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 29-50.

Jensen, S.K., Domingue, J.O. 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. Photogrammetric Engineering and Remote Sensing, 54: 241-244.

Jones, J.A.A.1986. Some limitations to the a/s index for predicting basin-wide patterns of soil water drainage. Zeitschrift für Geomorphologie, 60: 7-20.

Jones, J.A.A.1987. The initiation of natural drainage networks. Progress in Physical Geography, 11: 205-245.

Kamp, P.J.J. 1992. Tectonic architecture of New Zealand. In: Soons, J.M., Selby, M.J. (eds), Landforms of New Zealand, Auckland, 2nd edition. Longman Paul: pp.1-30.

Krysanova, V., Muller, D., Cramer, W., Becker, A. 2000. Spatial Analysis of Soil-Moisture Deficit and Potential Soil Loss in the Elbe River Basin. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp.163-181.

Kumar, L., Skidmore, A.K., Knowles, E. 1997. Modelling topographic variation in solar radiation in a GIS environment. International Journal of Geographical Information Science, 11: 475-497.

Lea, N.J. 1992. An aspect-driven kinematic routing algorithm. In: Parsons, A.J., Abrahams, A.D. (eds), Overland Flow: Hydraulics and Erosion Mechanics. London: UCL Press.

Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K. 2003a. LENZ Technical Guide. David Bateman Ltd, Auckland, New Zealand, 237pp.

Leathwick, J.R., Stephens, R.T.T., Wilson, G.W. 2002. New Zealand Climate Surfaces. Landcare Research Contract Report 9798/126, Hamilton, New Zealand.

Leathwick, J., Wilson, G., Rutlege, D., Wardle, P., Morgan, F., Johnston, K., McLeod, M., Kirkpatrick, R. 2003b. Land Environments of New Zealand. David Bateman Ltd, Auckland, New Zealand, 183pp.

Lilburne, L.R., Webb, T.H., Benwell, G.L. 2004. The Scale Matcher: a procedure for assessing scale compatibility of spatial data and models. International Journal of Geographical Information Science, 18: 257-279.

LINZ, 2007. http://www.linz.govt.nz.

Lowe, D.J. Palmer, D.J. 2005. Andisols of New Zealand and Australia. Journal of Integrated Field Science, 2: 39-65.

Mackey, B.G., Mullen, I.C., Baldwin, K.A., Gallant, J.C., Sims, R.A., McKenney, D.W. 2000. Towards a spatial model of Boreal forest ecosystem: The role of digital terrain analysis. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 391-422.

McCool, D.K., Brown, L.C., Foster, G.R., Mutchler, C.K., Meyer, L.D. 1987. Revised slope steepness factor for the Universal Soil Loss Equation. Transactions of the American Society of Agricultural Engineers, 30: 1387-1396.

McCool, D.K., Foster, G.R., Mutchler, C.K. 1989. Revised slope length factor for the Universal Soil Loss Equation. Transactions of the American Society of Agricultural Engineers, 32: 1571-1576.

McKenzie, N.J., Gessler, P.E., Ryan, P.J., O'Connell, D.A. 2000. The role of terrain analysis in soil mapping. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 245-265.

Mitasova, H., Hofierka, J. 1993. Interpolation by regularized spline with tension, II: application to terrain modelling and surface geometry analysis. Mathematical Geology, 25: 657-669.

Molloy, L., Christie, Q. 1998. The Living Mantle: Soils in the New Zealand Landscape, 2nd edition. New Zealand Society of Soil Science, Lincoln University, Canterbury, 253 pp.

Moore, I.D. 1996. Hydrologic modelling and GIS. In: Goodchild, M.F., Steyaert, L.T., Parks, B.O., Crane, M.P., Johnston, C.A., Maidment, D.R., Glendinning, S. (eds), GIS and Environmental Modelling: Progress and Research Issuses. Fort Collins, CO: GIS World Books: 143-148.

Moore, I.D., Foster, G.R. 1990. Hydraulics and overland flow. In: Anderson, M.G., Burt, T.P. (eds), Process Studies in Hillslope Hydrology. New York, Wiley: pp. 215-154.

Moore, I.D., Grayson, R.B. 1991. Terrain-based catchment partitioning and runoff prediction using vector elevation data. Water Resources Research, 27: 1177-1191.

Moore, I.D., Wilson, J.P. 1992: Length-slope factors for the Revised Universal Soil Loss Equation: simplified method of estimation. Journal of Soil and Water Conservation, 47: 423-428.

Moore, I.D., Wilson, J.P. 1994: Reply to comments by Foster on "Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation." Journal of Soil and Water Conservation, 49: 174-180.

Moore, I.D., Burch, G.J., MacKenzie, D.H. 1988. Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. Transactions of the American Society of Agricultural Engineers, 31: 1098-1107.

Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A. 1993a. Soil attribute prediction using terrain analysis. Soil Science Society of America Journal, 57: 443-452.

Moore, I.D., Grayson, R.B., Ladson, A.R. 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological Processes, 5: 3-30.

Moore, I.D., Lewis, A., Gallant, J.C. 1993b. Terrain attributes: estimation methods and scale effects. In: Jakeman, A.J., Beck, M.B., McAleer, M.J. (eds), Modelling Change in Environmental Systems. New York: 189-214.

Mummery, D., Battaglia, M., Beadle, C.L., Turnbull, C.R.A., McLeod, R. 1999. An application of terrain and environmental modelling in a large-scale forestry experiment. Forest Ecology and Management, 118: 149-159.

Neall, V.E. 2001. Volcanic landforms. In: Sturman, A., Spronken-Smith, R. (eds), The Physical Environment – a New Zealand Perspective. Oxford University Press, Melbourne: pp. 39-60.

Newnham, R.M., Lowe, D.J., Williams, P.W. 1999. Quaternary environmental change in New Zealand: a review. Progress in Physical Geography, 23: 567-610.

O'Callaghan, J.F., Mark, D.M. 1984. The extraction of drainage networks from digital elevation data. Computer Vision, Graphics and Image Processing, 28: 323-344.

Pennock, D.J., Zebarth, B.J., de Jong, E. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. Geoderma, 40: 297-315.

Pettinga, J. 2001. Rock formation and Earth-building processes. In: Sturman, A., Spronken-Smith, R. (eds), The Physical Environment – a New Zealand Perspective. Oxford University Press, Melbourne: pp. 12-38.

Pinde, F., Rich, P. 2002. A geometric solar radiation model with applications in agriculture and forestry. Computers and Electronics in Agriculture, 37: 25-35.

Quinn, P.F., Beven, K.J., Chevallier, P., Planchon, O. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrological Processes, 5: 59-79.

Quinn, P.F., Beven, K.J., Lamb, R. 1995. The ln(a/tan β) index: how to calculate it and how to use it within the TOPMODEL framework. Hydrological Processes, 9: 161-182.

Rodda.H., Stroud, MJ., Shankar, U., Thorrold, B.S. 2001. A GIS based approach to modeling the effects of land-use change on soil erosion in New Zealand. Soil Use and Management, 17: 30-40.

Romano, N., Palladino, M. 2002. Prediction of soil water retention using soil physical data and terrain attributes. Journal of Hydrology, 265: 56-75.

Smith, R.T., Lowe, D.J., Wright, I.C. 2006. Volcanoes. Te Ara – The Encyclopaedia of New Zealand [online]. Updated 9th June, 2006. New Zealand Ministry for Culture and Heritage, Wellington.
URL: http//www.TeAra.govt.nz/EarthSeaAndSky/NaturalHazardsAndDisasters/Volcanoes/en.

Sulebak, J.R., Tallaksen, L.M., Erichsen, B. 2000. Estimation of areal soil moisture by use of terrain data. Geografiska Annaler Series A: Physical Geography, 82: 89-105.

TAPES-G, 1995. Manuscript describing methods for implementing the TAPES-G program. Terrain Analysis Programmes for the Environmental Sciences – Grid. Documentation version 1.4.

Thompson, J.A., Bell, J.C., Butler, C.A. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma, 100: 67-89.

Ventura, S.J., Irvin, B.J. 2000. Automated Landform Classification Methods for Soil-Landscape Studies. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 267-294.

Wheatley, J.M., Wilson, J.P., Redmond, L.R., Ma, Z., DiBenedetto, J. 2000. Automated land cover mapping using Landsat thematic mapper images and topographic attributes. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 355-389.

Wilson, J.P., Gallant, J.C. 1998. Terrain-based Approaches to Environmental Resource Evaluation. In: Lane, S.N., Richards, K.S., Chandler, J.H. (eds) Landform Monitoring, Modelling, and Analysis. New York, John Wiley and Sons: pp. 219-240.

Wilson, J.P., Gallant, J.C. 2000. Digital Terrain Analysis. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 1-28.

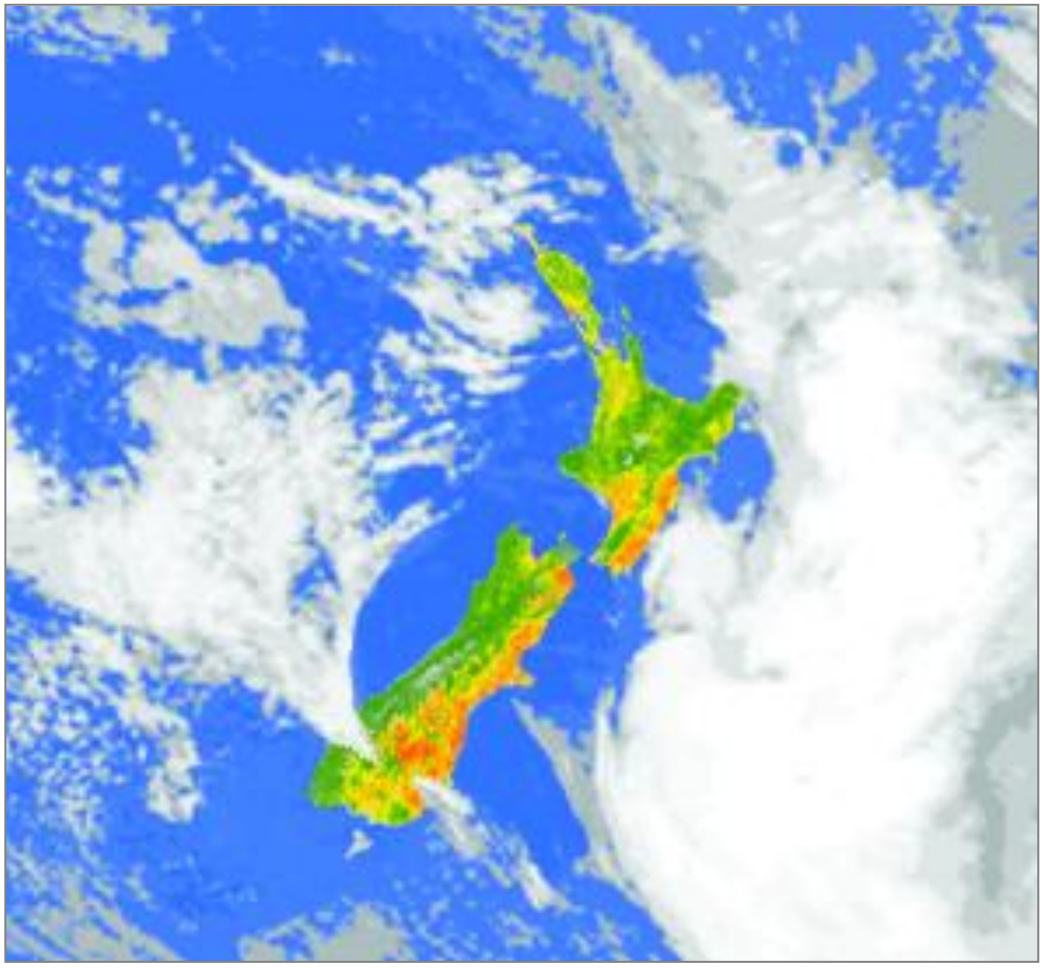
Wilson, J.P., Repetto, P.L., Snyder, R.D. 2000. Effect of Data Source, Grid Resolution, and Flow-Routing Method on Computed Topographic Attributes. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 133-162.

Wynn, J.G., Bird, M.I., Vellen, L., Grand-Clement, E., Carter, J., Berry S.L. 2006. Continental-scale measurement of the soil organic carbon pool with climatic, edaphic, and biotic controls. Global Biogeochemical Cycles, Vol. 20, No. 1, GB1007.

Young, A. 1972. Slopes. Oliver and Boyd, Edinburgh.

"…what is the use of a book," thought Alice, "without pictures or conversations?"

Lewis Carroll, *Alice in Wonderland*



Satellite perspective of New Zealand's cloud cover over a map depicting the mean fraction of available root-zone water storage (W_f) for February. W_f data derived from this study superimposed over an adjusted free online satellite cloud cover image from NIWA.

CHAPTER THREE

A DYNAMIC FRAMEWORK FOR SPATIAL MODELLING *PINUS RADIATA* SOIL WATER BALANCE (SWATBAL) ACROSS NEW ZEALAND

D.J. Palmer[a*], M.S. Watt[b], B.K. Höck[b], D.J. Lowe[a], T.W. Payn[b]

[a]*Department of Earth and Ocean Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand 3240*

[b]*Ensis,Private Bag 3020, Rotorua, New Zealand*

[*]Corresponding author. Tel.: +64-7-8562889 *E-mail address*: djp8@waikato.ac.nz

Submitted as a Scion Bulletin

ABSTRACT

Available root-zone water storage is a key environmental property useful in modelling many physical systems including crop production, hydrological processes, ground water pollution, and biodiversity. We have developed a dynamic spatial soil-water balance model, SWatBal, using Arc Macro Language (AML) and the GIS platform ArcInfo™ by adapting a point-location model specific to *Pinus radiata* in New Zealand. Spatial estimates of available root-zone soil water content across New Zealand were modelled at 100 m-resolution using the best available soil and climate data. SWatBal uses monthly climatic data distributed on a daily basis to simulate naturally occurring stochastic rainfall events, rather than an assumed constant monthly rainfall, making this model unique by improving its temporal scale. We discuss the development of SWatBal, protocols, input parameters, and the implementation of spatial soil and climate data for New Zealand. SWatBal is designed to ensure the model can be readily updated whenever new and increasingly accurate data become available. Although SWatBal surfaces will prove valuable for many disciplines, the rationale driving our model development is further research and investigation into spatial relationships between available root-zone water storage and *P. radiata* crop productivity.

Keywords: soil water balance, spatial modelling, national scale modelling, SWatBal, GIS, New Zealand.

List of symbols

Symbol	Description	Units
c	Fractional stone content of the soil	$m^3\ m^{-3}$
D	Air saturation deficit	kPa
D_{0t}	Sensitivity of *Pinus radiata* g_{st} to D	kPa
E_g	Evaporation from the soil surface	mm
E_t	Transpiration from the *Pinus radiata* canopy	mm
E_{tw}	Evaporation from the wet *Pinus radiata* canopy	mm
F	Drainage from the root zone	mm
f_{wt}	Fraction of rainfall intercepted by the tree canopy	-
G_g	Available energy beneath the tree canopy	$MJ\ m^2\ day^{-1}$
g_{st}	*Pinus radiata* stomatal conductance	$mmol\ m^2\ s^{-1}$
g_{stmax}	*Pinus radiata* maximum stomatal conductance	$mmol\ m^2\ s^{-1}$
L_t	*Pinus radiata* leaf area index	$m^2\ m^{-2}$
P	Precipitation	mm
PAW	Profile total available water	mm
PRD	Potential rooting depth	m
RAV	Reasoned and allocated virtual rainfall data	mm
s	Slope of the relationship between saturated vapour pressure and temperature at a given air temperature	$kPa\ {}^oC^{-1}$
W	Root-zone water storage	mm
W_a	Available root-zone water storage	mm
W_f	Fraction of available root-zone water storage	-
W_m	Maximum available root-zone water storage (W_{max}-W_{min})	mm
W_{max}	Maximum root-zone water storage	mm
W_{min}	Minimum root-zone water storage	mm
W_t	Value of W when E_t, E_g decline	-
W_v	Available volumetric root-zone water content	$m^3\ m^{-3}$
γ	Psychrometric constant	$kPa\ {}^oC^{-1}$

CONTENTS

1. INTRODUCTION

1.1 Project rationale

In New Zealand *Pinus radiata* D. Don is the most widely distributed commercial forestry crop covering an estimated 1.6 million hectares (NZFOA, 2005). The primary environmental factors required for *P. radiata* biomass productivity include water, solar energy (light and heat), gaseous exchange and nutrient uptake from soil, and structural support. For the most part, improved management practices can mitigate limiting environmental factors, but adequate soil moisture on free draining and dryland sites remains limiting (Watt et al., 2003b). Numerous studies (Arneth et al., 1998a, 1998b; McMurtrie et al., 1990; Richardson et al., 2002; Watt et al., 2003a, 2003b) have used process-based models to assess *P. radiata* growth at locations exhibiting seasonal water deficits throughout New Zealand.

Historically, modelling more than one location in a forest stand was considered cost prohibitive because of the large data collection requirements needed for model input parameters. Now with the introduction of high speed computer processors and large digital storage capabilities, modelling forest stands, catchments, and whole climatic regions has become possible. Although the degree of accuracy and spatial resolution with which modelling is undertaken remains dependant on available data, improvements to model simulations will continue to increase into the future with the release of new and better-quality data. The ability to simulate soil water balance across large regions opens up a whole new area of research that explores the influence soil water deficit and surplus

plays on tree productivity at a regional or national extent. In this paper we set out to define the protocols and programming processes involved in developing a spatial soil water balance model for *P. radiata* across New Zealand (SWatBal).

1.2 New Zealand's climatic environment

New Zealand's three main islands cover a total area of ~270,000 km^2 (Leathwick, 1998), spanning 13 degrees of temperate mid-latitudes in the South Pacific Ocean (Figure 1). New Zealand's highest mountain peak, Mt Cook (Aoraki) is situated in the Southern Alps standing 3754 m above sea level. In the central North Island Mount Ruapehu reaches 2797 m, closely followed by Taranaki to the west at 2518 m. More than 75% of land in New Zealand is above 200 m elevation, with ~50% consisting of alpine or steepland, ~20% hill country, and ~30% flat to rolling landscape (Palmer et al., 2008). Northern New Zealand is considered subtropical compared with cooler temperate climates found to the south. To further complicate New Zealand's climate, the country is divided by mountain chains extending from Fiordland in the south to East Cape in the north. These mountainous regions are characterised by strong relief (Palmer et al., 2008; Leathwick et al., 2003) and severe alpine climatic conditions influence New Zealand's climate in an east to west direction. A good example of this climatic variability can be seen with the Southern Alps effectively dividing the South Island into two regions with the West Coast recognised as the wettest region of New Zealand and in contrast the eastern regions renowned for their markedly drier climatic conditions. Another climatic phenomena

influencing South Island climate is the aptly named "nor' wester" (fohn winds) that occasionally brings wind to the Canterbury plains east of the mountain ranges (Met. Service, 2006; NIWA, 2006).

Annual rainfall extremes can range from as little as 300 mm in Central Otago to greater than 14,000 mm in Fiordland/West Coast. Setting extremes aside, New Zealand rainfall averages, for the most part fall between 600 and 1500 mm a year, however, substantial areas across both the North and South Islands can exceed 2500 mm a year (Figure 2). For a large part of New Zealand, rainfall is spread relatively evenly throughout the year. However, in general rainfall dominates during winter months with nearly twice the rainfall of summer for northern regions. The predominance of winter rainfall diminishes southward with summer rainfall prevailing in the lower half of the South Island. Annual temperatures decrease gradually at southern latitudes, ranging from ~16°C in the north to ~10°C in the south (Met. Service, 2006; NIWA, 2006). The warmest months of the year are January and February with July tending to be the coldest month. For the greater part, variations in seasonal temperatures tend to be relatively small across New Zealand. However, extremes do occur with an estimated decrease in temperature of ~0.7°C for every 100 m gain in elevation. Snow is seldom found in coastal regions of the North Island and western regions of the South Island. Occasionally the South Island will experience short term snow cover, but usually snow is only found at higher elevations. In winter and spring months frosts are common throughout almost all of New Zealand, especially during periods of clear skies and still night air (Met. Service, 2006; NIWA, 2006).

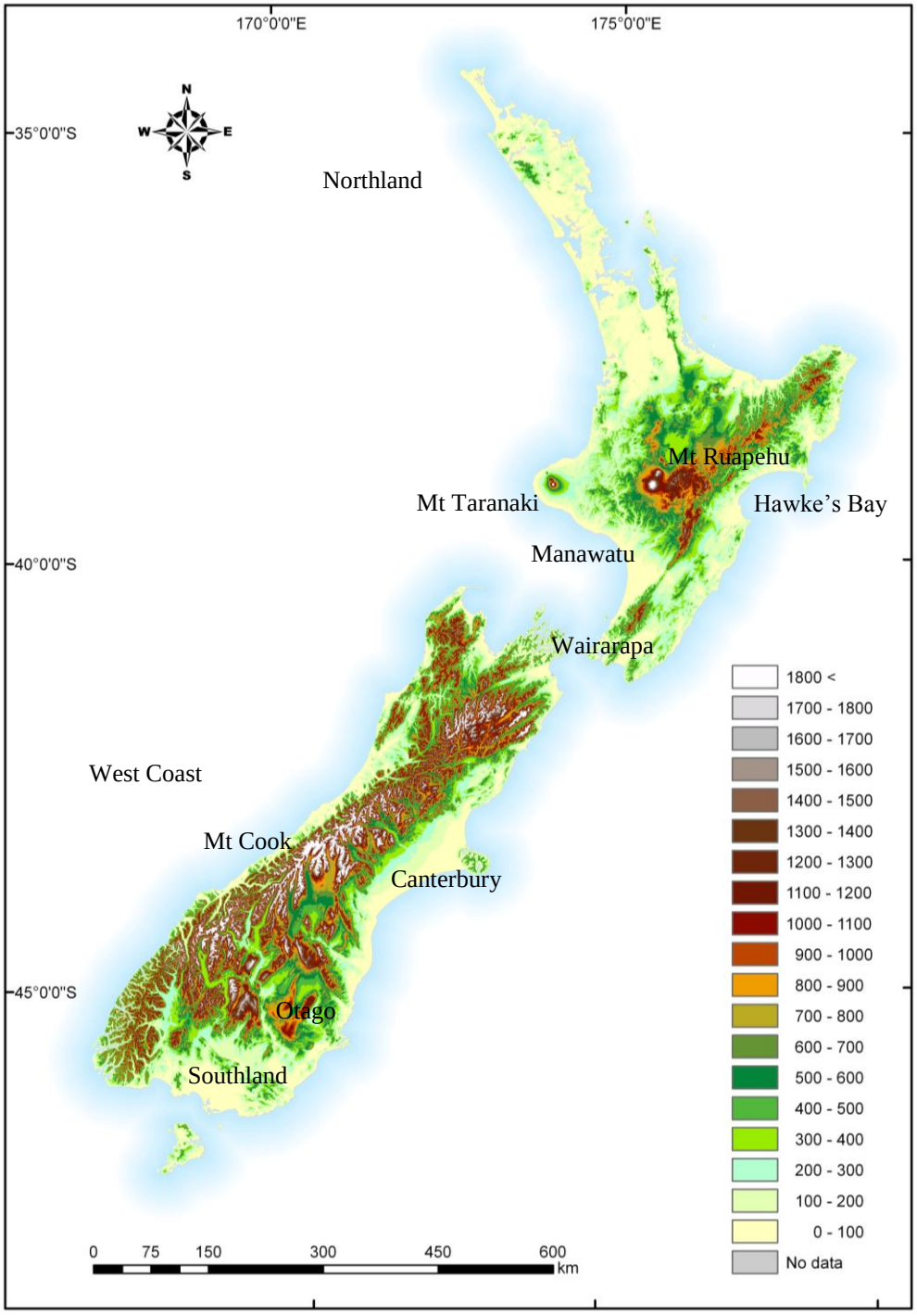


Figure 1: Digital elevation model (DEM) at a 25 m resolution illustrating New Zealand's topography and relief and geographic location on the Earth's surface (after Palmer et al., 2008).

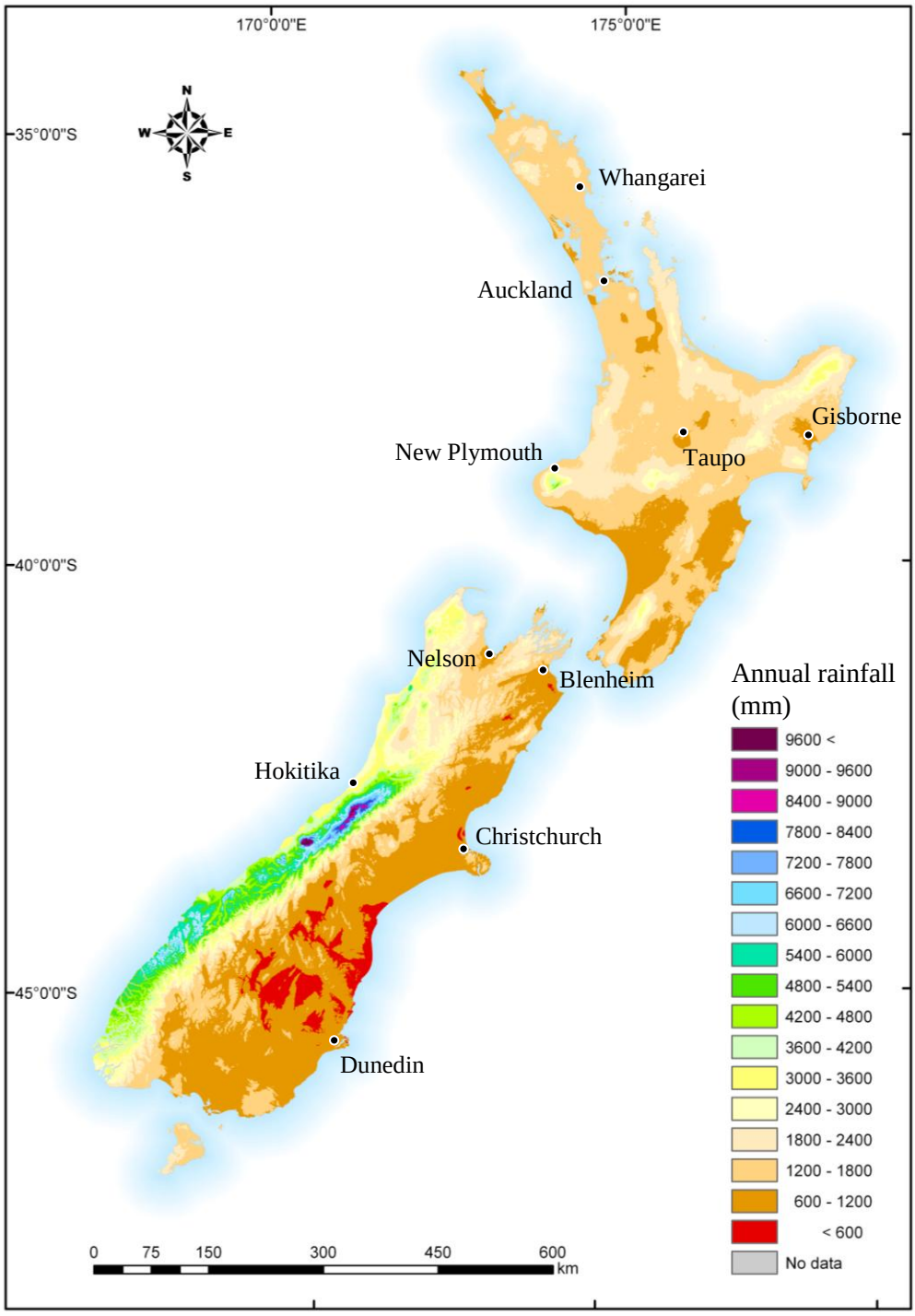


Figure 2: Annual rainfall for New Zealand at a 100 m resolution. Note data derived by summing 12 mean monthly rainfall surfaces (Leathwick et al., 2002b). Place names mentioned in the text are also shown on the figure.

1.3 Project outline

This manuscript describes the various protocols and methods used to create a nation-wide soil water balance model for New Zealand in a series of steps. First, we present a general overview of existing soil water balance models, including point location spreadsheet models and spatially driven models, before covering issues relating to data scale, resolution, and accuracy (Section 2). In Section 3 we discuss the underlying SWatBal spatial data input requirements. More specifically, Section 3.1 describes data source and preparation for each input surface prior to modelling, Section 3.2 covers climate surface development, Section 3.3 the soil component surface creation, and Section 3.4 latitude surface development. The advance made by allocation of Landcare Research's monthly rainfall data and their daily distribution using NIWA's Virtual data is discussed in Section 3.5. The evaporation and transpiration surfaces are discussed in Section 3.6. In Section 4 we provide the technical details used in developing the Arc Macro Language routine (AML) driving SWatBal. Section 5, provides examples of the modelled SWatBal surfaces, illustrating mean monthly fraction of available root-zone water storage, mean monthly available root-zone water storage (mm), and monthly mean drainage (mm). Section 6 discusses the sensitivity of SWatBal to changes in input parameters and variables, and also validates SWatBal against two published soil water balance studies. In Section 7 we discuss SWatBal advancements and model limitations. Finally we summarise the main points and discuss SWatBal future directions (Section 8).

2. MODELLING SOIL WATER

2.1 Point location models: estimation of available root-zone water storage

Available root-zone water storage for *P. radiata* has been modelled extensively across New Zealand for a variety of differing soils and climatic conditions. Usually the purpose of these studies is to examine water deficit as a main constraint to tree growth and development at these locations. Previous research on dryland sites has investigated how weed competition influences water availability and growth in juvenile (Watt et al., 2003b) and mid-rotation (Richardson et al., 2002) stands of *P. radiata*. Modelled estimates of root-zone water storage have also been used to investigate how water deficit influences carbon sequestration in an eight-year old *P. radiata* crop in north Canterbury (Arneth et al., 1998a, 1998b). Studies in which root-zone water storage has been modelled in *P. radiata* stands include an examination of rainfall interception by a seven-year-old *P. radiata* crop (Kelliher et al., 1992), modelling a small catchment containing 13-year-old *P. radiata* in south Waikato (Whitehead and Kelliher, 1991), and investigation of *P. radiata* growth from a nine to 12-year-old crop on Pinaki sands north west of Auckland (McMurtrie et al., 1990). These studies used process-based models to partition water balance into components of evaporation, transpiration, and root-zone water storage, thereby modelling water surplus or deficit at daily time steps for a specific point location. Generally, the daily water balance model used to calculate root-zone water storage (W) on the ith day is written as,

$$W_i = W_{i-1} + P_i - E_{ti} - E_{twi} - E_{gi} - F_i \quad (1)$$

where P_i is rainfall, E_{ti} transpiration from the tree canopy, E_{twi} evaporation from the wet tree canopy, E_{gi} evaporation from the soil surface, and F_i drainage from the root-zone. Model input parameters can include rainfall, temperature, solar radiation, vapour pressure deficit, leaf area index, latitude coordinate (Goudriaan and van Laar, 1994, day length calculation), minimum and maximum root-zone water content, root-zone stone content, and effective rooting depth. For modelling details see Whitehead et al. (2001) and Watt et al. (2003b). Climatic properties for each location are measured using on site weather station data. However, where data are unavailable because of project constraints or system malfunctions, climate data estimates are sometimes made from local neighbouring meteorological stations (Met. station). Soil data are collected in the field to make estimations of the maximum and minimum root-zone water storage, effective rooting depth, and stone fraction. These published models are extensively validated for estimating root-zone water content and allow fine-tuning adjustments specific to the study environment within their spatial extent (i.e. point locations). The drawback of point location models is that data accuracy quickly diminishes from the site of interest. Therefore, extrapolating soil water balance across a forest stand, compartment or a catchment remains problematic. Indeed, point location soil water balance models provide little or no information relating to spatial variability across the landscape. To conclude, point location soil water balance models and their applications are appropriate for studying at a specific location of interest, but where information across a greater extent is needed, a new approach is required. With the advent of modern high speed computer processors and increased digital storage capacity we can now develop spatial soil water

balance models using the wealth of existing knowledge contained in published point location models.

2.2 Spatially driven models for estimating soil water

Spatially driven soil water balance models offer advancements for researchers providing estimates of available root zone water content across many spatial extents including forest stand, compartment, catchment, or landscape level analysis together with temporal scales from hourly, daily, monthly and annual time scales. For the researcher such development opens new opportunities not previously available, allowing comparisons of moisture content whether within the same forest stand or at distal locations.

The introduction of GIS has played an instrumental role in the estimation of spatial patterns for soil water content. An early New Zealand example for the use of GIS technology can be seen in Barringer et al. (1995). Spatial estimates of water deficit zones were made across a ~250,000 hectare site in Canterbury illustrating spatial patterns for 30-year average rainfall, dry and wet rainfall years, and for El Niño and La Niña climate circulation patterns. This study highlighted the ability of GIS to integrate data from a variety of sources and resolutions, emphasising the potential role GIS technology can play in agroclimate studies. More recently, Barringer and Lilburne (1999) presented a GIS driven regional-scale soil water balance for the Canterbury Plains. Leathwick (1995) investigated the relationships between New Zealand tree species distribution,

environmental properties and climate using MOIST, which calculates the ratio of rainfall over potential evapotranspiration for months with rainfall deficits. MOIST uses the method of Thornthwaite and Mather (1954), chosen in the absence of reliable soil moisture information, thereby bypassing the need for soil water storage information. Predicting changes to the distribution of New Zealand's forests in the event of global warming (Leathwick et al., 1996) was undertaken using a more sophisticated model in the estimation of daily soil root zone moisture content. This model commences on the first day of July when New Zealand soils are considered to be at or close to field capacity. BIOCLIM derived data (Mitchell, 1991) was used to estimate monthly rainfall after adjustment was made for extreme outlier rainfall events. In an attempt to estimate rainfall across each month realistically, the average number of rain days in a month was calculated, average monthly rainfall evenly apportioned and allocated to each rain day across the month. The soil root zone water deficit approach was also used to investigate the distribution of *Northofagus* species across New Zealand (Leathwick, 1998). This model was modified following Whitehead, Leathwick, and Walcroft (2001) to account for the diurnal trend in canopy transpiration and soil evaporation (Goudriaan and van Laar, 1994). Whitehead et al. (2001) investigated carbon sequestration of New Zealand's indigenous forests with the updated model used in Leathwick and Whitehead (2001) to further investigate the distribution of New Zealand's tree species in relation to soil and atmospheric water deficits. Environmental variables were progressively refined for the prediction of New Zealand's forest patterns (Leathwick, 2001; Leathwick et al., 2002a, 2003). A numerical classification system was used to define areas with similar character using spatial layers inclusive of soil root zone water deficit and other environmental

properties related to New Zealand's climate and landforms. The water balance models described here provide valuable information and a foundation on which to build and improve future models that estimate soil moisture for New Zealand environments. For our purposes, spatial estimates of monthly root zone moisture content were developed for New Zealand to produce spatial digital surfaces that will assist in explaining *P. radiata* productivity across New Zealand.

2.3 Issues surrounding scale, resolution, data accuracy and errors

The SWatBal model uses an assemblage of environmental and climatic surfaces representing daily rainfall, solar radiation, temperature, vapour pressure deficit, latitude, rooting depth and pedon water storage. The national climate surfaces exist with a known degree of accuracy and reliability, not requiring further processing prior to model execution (Leathwick et al., 2002b), but the remaining surfaces required pre-processing or creation from existing data. The soil surfaces exist as vector digital maps requiring conversion to raster format, whereas 365 surfaces representing daily rainfall and a latitude grid were purpose-built for the SWatBal model. This section considers the development of the input surfaces for the SWatBal model setting out and discussing issues related to scale, resolution, data accuracy and their associated errors.

Scaling up or scaling down (Payn et al., 2000) are terms often used to describe the aggregation or disaggregation of spatial data to coarser or finer scales, respectively.

Aggregation of spatial surfaces to accommodate modelling at coarser scales is common practice, which can significantly affect model outcomes and consequently decision making processes. Therefore an in-depth understanding of the effects of data aggregation are fundamental to the SWatBal modelling process.

Scale can be defined in many ways often leading to ambiguity and confusion. Nevertheless, the single term 'scale' has emerged over the years to mean 'level of detail' (Goodchild, 2001). Bian (1997) summarised scale into the common connotations of cartographic scale, geographic scale, resolution, and operational scale. Cartographic scale, commonly referred to as map scale, describes a larger scale as having greater detail. For example, a soil map of 1: 5000 (1 cm on a map represents 5000 cm on the Earth's surface, i.e. 50 m) will have greater attribute detail, providing higher differentiation of soil units compared with a 1: 50000 scale map (1 cm on a map represents 50,000 cm on the Earth's surface, e.g. 500 m) capable of delineating only soil associations. This is analogous to comparing a map delineating soil properties to the subgroup level (fine detail) of the New Zealand Soil Classification System (NZSC) (Hewitt, 1998) (e.g. Immature Orthic Pumice Soils) compared with the coarser, less detailed Order level of the NZSC (e.g Pumice Soils), comprised of numerous soil sub classes lumped into one homogenous group often termed a soil complex or soil associations (Taylor and Pohlen, 1970; Singer and Munns, 1991). The connotation geographic scale, often termed spatial extent, refers to the area captured by a map. For example, a map of the Waikato region compared with a map of New Zealand having a smaller to larger geographic scale or spatial extent, respectively. Resolution refers to the

"size of the smallest distinguishable part of a spatial data set" (Bian, 1997). For example the resolution of a raster dataset, containing grid cells or picture elements (pixels) with a predetermined resolution, such as New Zealand's national digital elevation model (Barringer et al., 2002) or the terrain attribute surfaces developed for New Zealand (TANZ) (Palmer et al., 2008), have a resolution of 25 m compared with Landcare Research's climate surfaces (Leathwick et al., 2002b) with a 100 m resolution. The final connotation, operational scale, describes the scale over which the phenomenon operates or varies and can encompass both spatial and temporal variation. For example, soil water balance across regions with extreme rainfall events operate at larger scales both spatially and temporally than locations with little seasonal variation in rainfall. Throughout this manuscript unless otherwise stated or implied by context, scale will refer to the level of geographic detail used to describe the Earth's surface.

For SWatBal modelling it was important to choose spatial datasets at cartographic scales and resolutions appropriate to the property of interest, capturing sufficient detail to adequately explain attribute variation (operational scale), yet not overburdening computer processing time and storage constraints with unnecessary detail (Palmer et al., 2008). The accuracy of spatial data, commonly referred to as surfaces or layers, is influenced by spatial scale and source data quality and even the most sophisticated algorithms used in surface fitting will produce substandard outcomes should the input source data be inadequate (Wilson and Gallant, 1998). Spatial surfaces are seldom, if ever, error-free containing inherited inaccuracy from initial sampling methods including measurement, positional accuracy, observation density, classification interpretation and protocols

(Heuvelink, 1998), whereas post sampling may contain inherent map production, generalisation, format conversion, and numerical errors. The common practice of digitising paper maps will not only replicate existing errors contained within the spatial data, but may also compound errors during the reproduction process itself. For detailed discussion covering the many different sources of error consult Burrough (1986); Burrough and McDonnell (1998); and Zang and Goodchild, (2002). Another consideration when creating a spatial surface is statistical error (Burrough, 1986) where sharp, crisp boundaries representing a homogenous polygon unit provides little if any statistical information about the variation occurring within each unit. In reality, geological, pedological, and environmental data are seldom single homogenous units and are better represented by a raster format enabling boundaries to change gradually over space or time. In earth sciences not all data are derived directly from paper maps or digital images with digital surfaces frequently interpolated from point data observations (Heuvelink, 1998) where the prevailing error propagation is found in the measurement and interpolation processes. Geostatistical spatial interpolation techniques such as kriging have the added advantage of quantifying these errors (Journel and Huijbregts, 1978; Webster and Oliver, 1990; Cressie, 1991; Goovaerts, 1997, 1999; Burrough and McDonnell, 1998; Pawlowsky-Glahn and Olea, 2004).

When deriving raster data from vector point or line datasets, grid cell resolution in relation to underlying vector data becomes important. Consider the Canterbury Plains where elevation changes little over relatively large distances compared with the mountainous Southern Alps where elevation varies over markedly short distances. When

interpolating a DEM representing elevation for steep terrain (Southern Alps), then finer resolutions of 5 m or 25 m improve the capture of landscape short-range variability. However, a 50 m DEM may adequately capture and represent the spatial variability occurring across the Canterbury Plains because of its relatively flat landform, with little detail gain with finer DEM resolutions, which would only result in higher computer computational time and storage costs (Palmer et al., 2008). In relative terms a 50 m resolution raster pixel represents 2500 m^2 on the Earth's surface compared with a 5 m cell representing 25 m^2, 100 times greater detail. Other considerations are the spatial and temporal density (operational scale) of the underlying vector data and the purpose for which the resultant surface will be used. In general, a high resolution raster will capture greater detail compared with a low resolution, but there is a lower and upper limit (floor and ceiling) with which detail can be captured and improved. Consider spatially interpolating a DEM from a vector contour map representing 20 m elevations. A cell-size resolution of 5 m will provide too much detail with little, if any gain, compared with a 100 m resolution DEM, which exceeds the spatial density, thus over simplifying the representation of underlying data. To further complicate surface creation, the temporal aspect or resolution associated with climatic properties needs addressing. Ask the question, "For what purpose and at what temporal scale should the data be represented?". These examples illustrate the importance of using data that adequately capture and illustrate the property of interest's spatial and temporal variability while appropriately representing the property's spatial density.

3. UNDERLYING SWATBAL SPATIAL SURFACE DEVELOPMENT

3.1 SWatBal modelling rationale

SWatBal uses a variety of primary environmental surfaces including rainfall, solar radiation, temperature, vapour pressure deficit, soil moisture content, effective rooting depth, and latitude. Climate surfaces with recognized methodologies and a known degree of accuracy were available for SWatBal modelling (Leathwick et al., 2002b). However, surfaces representing the soil components required further development before use, and latitudinal and daily rainfall surfaces required building from primary data.

For our application we decided to model on a daily basis providing temporal detail not available when modelling across longer time intervals. To model water balance at an hourly interval, thereby providing greater temporal detail, although possible, provides little detail gain, when the purpose of our modelling is comparing average monthly soil water balance across New Zealand. A standard 100 m cell-size resolution was used for all processing data. There were three main reasons in choosing this resolution when modelling soil water balance at a national extent.

1) In general the spatial variability of climate data is adequately accounted for at 100 m resolution and existing climate data of known source and quality is at 100 m resolution.
2) Even the latest capabilities of high-speed computer processors restrict or limit data volume processed over any given period of time (Palmer et al, 2008).

3) Storage requirements, even with gigabytes available, may still constrain overall modelling capabilities.

Overall, using a 100 m cell-size resolution was a trade off between computer processing and storage capabilities, the spatial extent of underlying data and its associated errors, compared with the need for increasingly detailed outcomes when creating a soil root zone water balance model covering the extent of New Zealand.

3.2 Climate surfaces

Climate surfaces utilised in the SWatBal model are monthly rainfall, solar radiation and temperature, all of which were created by Landcare Research (Leathwick et al., 2002b). New Zealand Meteorological Service climate normals for years 1951 to 1980 were extracted by Landcare Research in the development of temperature and rainfall surfaces, while other climate properties utilised all available records up to 1980. Surface fitting implemented ANUSPLIN software using thin-plate splines (Leathwick et al., 2002b) for interpolation optimisation of irregularly spaced climate data networks (Hutchinson and Gessler, 1994). For detailed information relating to underlying climate data, its surface fitting, and data accuracy refer to Leathwick et al. (2002b).

3.3 Soil component surfaces

Existing Landcare Research fundamental soil layers (FSLs) were used in the development of potential rooting depth and root-zone water storage. The FSLs profile total available water (*PAW*) and potential rooting depth (*PRD*) were used to estimate SWatBal available root zone water content across New Zealand. The FSLs are a collection of soil physical and chemical properties developed into GIS surfaces by linking soil information stored in New Zealand's National Soil Database (NSD) with New Zealand's Land Resource Information (LRI) polygons through their associated soilform classifications. The soilform Clayden and Webb, 1994 is the fourth category of the NZSC (Hewitt, 1998). Each FSL contains a series of fields representing the minimum, middle, maximum values, categorical and class descriptor fields, and fields that estimate data variability and lineage for each polygon unit. Newsome et al. (2000) commented that the middle value, termed 'mid', was calculated as the middle value in the range from minimum to maximum, and because these ranges may straddle class boundaries the 'mid' value will not necessarily equate to the expected mid-point of a class. For detailed information and discussion related to class boundaries and associated methodologies see Webb and Wilson (1995). Although the two FSLs used to represent the soil component of the SWatBal model, *PRD* and *PAW* have greatly improved since the conception of the original LRI maps developed in the 1970s, the accuracy associated with these layers remains difficult to quantify compared with the climate surfaces. The two estimates associated with data accuracy for the FSLs are data variation and lineage, which are associated with each polygon unit contained within the FSLs. Variation provides an

estimation of data reliability based on how many classes the data may straddle (Table 1). Lineage indicates whether the data were derived from the named soil unit, or relationships with other soils, and whether the estimate is considered reliable (Table 2). Spatial variation for the reliability and lineage for the FSL classes are shown in Figures 3 and 4.

Table 1: Variation distribution for the fundamental soil layers (FSLs) as defined by Newsome et al. (2000).

Class	Data position	Class variability description
0	Occurs mostly within the nominated class	Middle of the nominated class is a good approximation for a numerical value
1	Straddles the class above and below	Mean is the middle of the nominated class
1-	Straddles this class and the class below	Mean is taken at the class boundary
1+	Straddles this class and the class above	Mean is taken at the class boundary
2	Straddles two classes above and below	Mean is the middle of the nominated class

Table 2: Lineage distribution for the fundamental soil layers (FSLs) as defined by Newsome et al. (2000).

Class	Data lineage description	Reliability estimate
m	Estimated from analysis or measurements of the named soil	Estimate considered reliable
p	Measured from profile	Estimate considered reliable
r	Estimated from relationships with other soils	Estimate considered reliable
u	Estimated from relationships with other soils	Unknown level of accuracy
uf	Estimated from General Soil Survey data (scale 1: 253,440)	In general, quality of estimate considered less reliable than class ‘u’ above

[a] Class ‘p’ is not described in Newsome et al. (2002), but it occurs as a class in the digital data. Peter Newsome pers. comm. (2006) provided the class description stated here.

In general, class zero dominates the *PAW* and *PRD* variation estimate surfaces shown in Figures 3 and 4 suggesting values occurring within this class are considered a good approximation of their values. Classes 1, 1+, and 1- also account for a large portion of their layers, suggestive that value estimates straddle their current class and the class above and below (class 1), the current class and the class above (class 1+), and the current class and class below (1-). Class 2 accounts for only small areas across New Zealand, indicating that estimates from data straddling two classes above and below are minimal. The *PAW* layer (Figure 3) indicates that lineage estimates are reliable for the central North Island and Waikato, Hawke's Bay, New Plymouth, parts of Southland, and extensively across the Canterbury Plains. Areas of concern are the Wairapapa region. *PRD* estimates of reliability are generally good with the exception of parts of East Cape, Wairapapa, Central Otago, and Fiordland (Figure 4).

PAW values (Table 3) were derived for each soil profile to the maximum *PRD* depth (Newsome et al., 2000). *PAW* was estimated from the volumetric water content difference between -10 kPa and -1500 kPa across all root utilisable horizons (Webb and Wilson, 1995). Methods were developed from the work of Gradwell and Birrell (1979); Wilson and Giltrap (1982); and Griffiths (1985) and currently remain the best available estimates of profile total available water at a national extent across New Zealand. Stone content was also accounted for when deriving the *PAW* surface (Trevor Webb pers. comm. (2007).

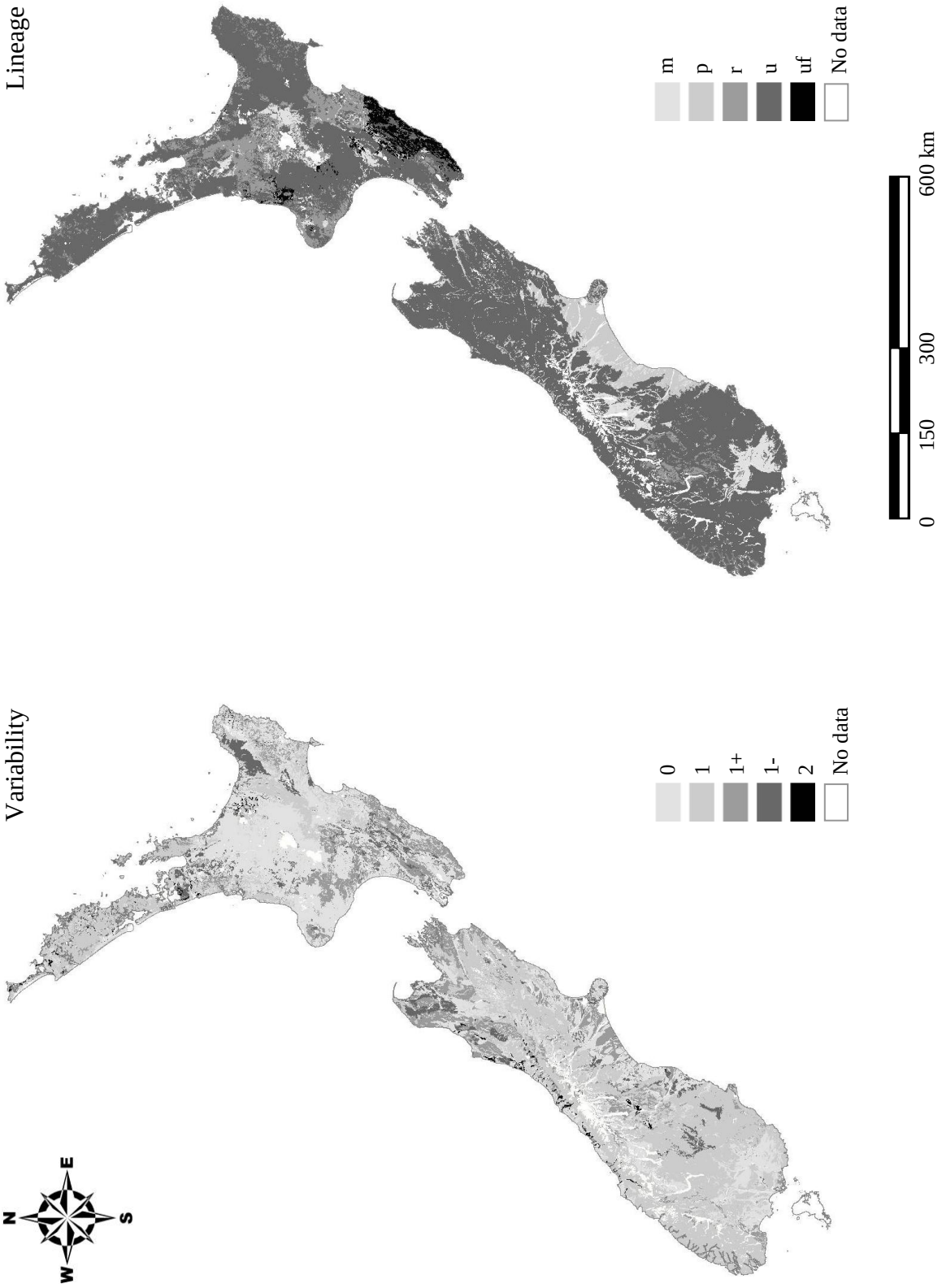


Figure 3: Profile total available water (*PAW*) estimates of variability, and lineage as recorded in the fundamental soil layers (refer to Tables 1 and 2 for legend explanations).

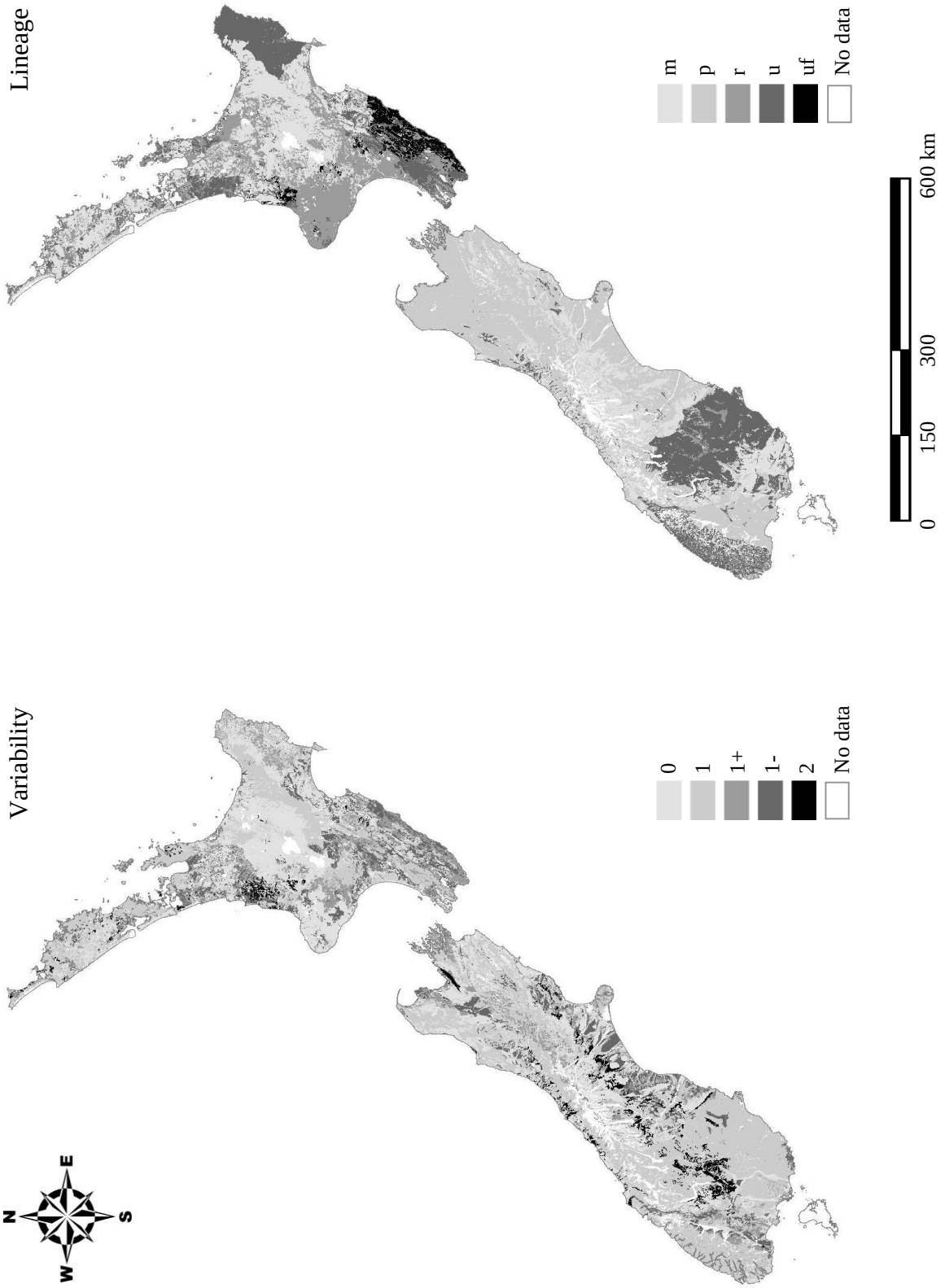


Figure 4: Potential rooting depth (*PRD*) estimates of variability, and lineage as recorded in the fundamental soil layers (refer to Tables 1 and 2 for legend explanations).

Table 3: Definition of profile total available water classes (after Newsome et al., 2000)

PAW class	PAW min (mm)	PAW max (mm)	Description
1	250	350	Very high
2	150	249	High
3	90	149	Moderately high
4	60	89	Moderate
5	30	59	Low
6	0	29	Very low

Griffiths (1985) stated that the potential rooting depth (*PRD*) is the "depth of soil that can provide a suitable medium for root development, retain available water in desirable quantities, and supply nutrients". More specifically, the *PRD* identifies the depth (m) to a soil layer that may impede root extension, identified by penetration resistance, poor aeration or very low water capacity (Wilde et al., 2000) (Table 4). The *PRD* is not plant-species specific and the question of whether or not the entire *PRD* is fully utilised will depend on the type of plant species and its stage of development. For comprehensive discussions covering *PRD* development methods see Webb and Wilson (1995) and Griffiths (1985). Currently the *PRD* is the best available national extent dataset, providing an estimation of rooting depth at which the *P. radiata* species can occupy a soil volume throughout New Zealand.

Table 4: Definition of potential rooting depth classes (after Newsome et al., 2000)

PRD class	PRD min (m)	PRD max (m)	Description
1	1.2	1.5	Very deep
2	0.9	1.19	Deep
3	0.6	0.89	Moderately deep
4	0.45	0.59	Slightly deep
5	0.25	0.44	Shallow
6	0.15	0.24	Very shallow

Preprocessing of the FSLs prior to running the SWatBal model required conversion of the polygon feature classes representing PAW and PRD to a gridded raster format. A purpose-written Arc Macro Language (AML) routine was executed in the GIS platform ArcInfo™ to predefine the surface extent, cell resolution, and snapping environment prior to converting "mid" values contained in each coverage to a raster format.

On completion of the FSL pre-processing, an available volumetric root-zone water content (W_v) surface was derived from the *PAW* and *PRD* surfaces. The reason for developing a W_v surface prior to SWatBal model execution was to reduce "on the fly" model calculations. The W_v surface was developed using the equation

$$W_v = PAW / ((PRD \times 1000) \times (1 - \text{stone})) \quad (2)$$

Note that stone content was set to zero, but remains an input variable should future models require it.

3.4 Latitude surfaces

Day length calculations for the SWatBal model require a continuous national surface representing latitude across New Zealand. The NZMG projection was used to align 100 m resolution grids with the other SWatBal input variables while z values represent latitude at each location were in decimal degree. Refer to Palmer et al. (2008) for development details.

3.5 Development of daily rainfall surfaces

The SWatBal model climate property with the greatest sensitivity to change in temporal trends is precipitation. Rudimentary soil water balance models may distribute rainfall (monthly daily averages) evenly across each day in a month, effectively constraining model simulation, impacting on drainage rates and the natural daily variability that occurs.

The allocation of rainfall data on a daily basis could be approached from a number of directions. For example, climate data from one representative year could be used in the modelling process. This approach remains problematic because annual climatic heterogeneity makes it difficult to determine a year in which rainfall is representative for

all New Zealand locations[1]. For example, Figure 5 highlights the difference between years with total rainfall at Burnham for June 1988 and 1989, as 43.6 and 62.4 mm, respectively. Another approach is calculating the daily rainfall from the same Julian days averaged across long-term rainfall data. However, this approach may prove detrimental to short-term rainfall distribution events that typically clump together as seen in Figure 6. A simplistic approach is the allocation of monthly mean rainfall evenly across each day of the month leaving constant non-stop rainfall (Figure 6A). An advancement on this approach is to derive the average number of days it rains in a month from long-term rainfall data, distribute the rain days evenly across the month and allocate rainfall evenly to each rain day (Figure 6B). A further step might be the allocation of rainfall amount (mm) based on a stochastic mathematical function (Figure 6C). Once again, these approaches may produce an unrealistic rainfall distribution pattern that is not repeatable.

To improve model simulations, 365 rainfall surfaces were created. These surfaces were called 'reasoned and allocated virtual data' (*RAV*). Two datasets were used to create daily rainfall for New Zealand, Landcare Research's monthly mean rainfall surfaces (Leathwick et al., 2002b) and NIWA's virtual climate data (Tait et al., 2006).

[1] For example the Insurance Council of New Zealand (www.icnz.org.nz) shows the most recent year with no significant claims for storm and flood events in New Zealand is in 1989.

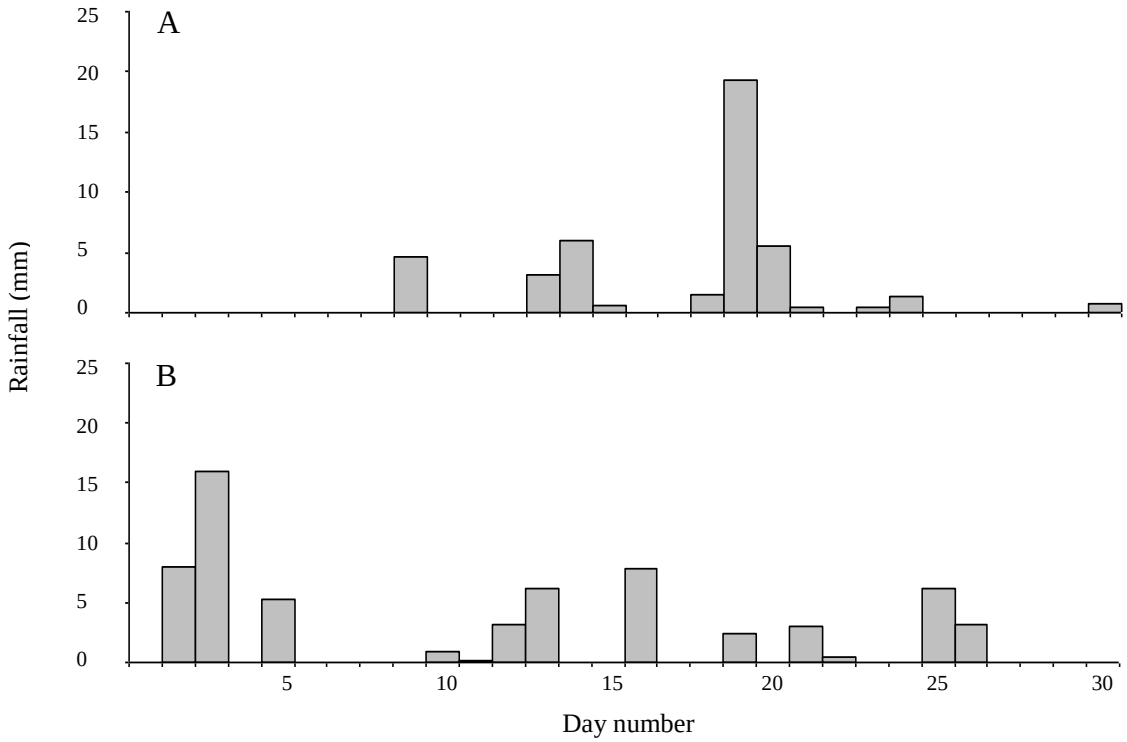


Figure 5: Rainfall distributions at Burnham (Canterbury), illustrating rainfall events clumped together and demonstrating the different distribution patterns across the month of June for the years 1988 (A), and 1989 (B) (Met. Service data).

The criteria used to select months representing 'normal' rainfall from the ~40 years of virtual rainfall data were,

1) months closest to average number of rain days, and

2) months closest to the sum of rainfall

compared to Landcare Research's monthly mean rainfall values. The portion of rainfall within each rain day from the selected month was allocated from Landcare Research's monthly mean rainfall.

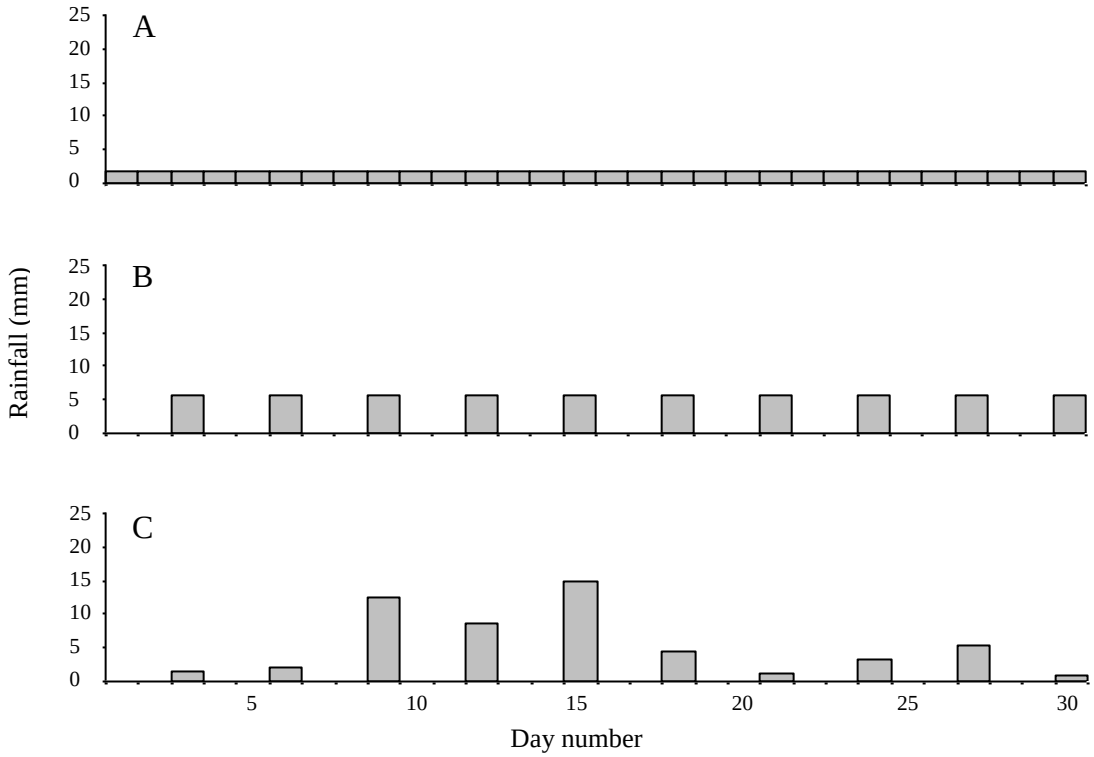


Figure 6: Hypothetical rainfall distributions: A) distribution of monthly mean (55.2 mm) across all days in month (1.84 mm daily); B) distribution of monthly mean across the number of days it rains (10 rain days) as determined from long-term data; and C) random allocation of rainfall to graph B rain day distribution.

The virtual climate data consist of 11521 files with actual and interpolated climate data for the period 1962 to 2004 with each file representing a point location across New Zealand on a 5 km grid. The rainfall surfaces built for the SWatBal model used a combination of monthly mean rainfall and virtual climate data for the selection criteria and rainfall allocation. A Visual Basic (VB) programme was developed to extract and allocate rainfall to months most representative of normal rainfall distribution from the ~40 years of virtual climate data and rainfall from the monthly mean rainfall surface. This allocation was achieved through a series of iterative steps for each location represented by a file containing ~40 years of rainfall data (Figure 7). Initially extreme rainfall events

occurring within a month at each location across the ~40 years of data were removed prior to determining the average number of days it rained (> 1 mm) in each month across the ~40 years of virtual climate data. Extreme rainfall events were excluded from further analysis by summing the three highest daily rainfall events in each month and calculating whether this exceeded 80 % of the monthly mean rainfall. A monthly selection criteria was used to determine for each month across the ~40 years of virtual rainfall data the month that is closest to the number of days it rains and the total monthly average rainfall based on the Landcare Research monthly mean rainfall surface. In total, 12 months of data was selected from the ~40 years virtual rainfall data representing rainfall distribution at each location. The percentage of rain allocated to each rain day is calculated before allocating and distributing Landcare Research's monthly mean rainfall to each day it rains. Table 5 shows the number of months selected closest to the number of days it rained (> 1 mm) compared with the total volume of rainfall. Although nearly all the months selected are within one rain day of the monthly average over the ~40 years, there remain some months with higher than average total monthly rainfall. However, the higher than average total monthly rainfall was not considered a concern because the portion of rain occurring on each rain day was determined before allocating the mean monthly rainfall data (Landcare Research climate surfaces). The final output provides 365 days of rainfall data (12 representative months were selected from long-term data) using monthly virtual data representing a realistic rainfall distribution across each month. The interim vector dataset contains the spatial components for each of the x and y coordinates representing 5 km locations across New Zealand and 365 fields providing representative rainfall occurring on each Julian day.

Table 5: Table with the number of months selected to represent the amount and distribution of monthly rainfall based on the closest fit with the number of days it rains and the total rainfall within a month (note: locations may have multiple representative months for each rainfall class).

Rain days in month [1]	Rainfall (mm) 0 – 50	50 – 100	100 – 200	200 – 300	> 300
0	133365	3816	877	97	22
1	70	3	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	2	2	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0

[1] Number of rain days with > 1 mm

The interim vector dataset required further processing to a raster format with an appropriate spatial scale suitable for the SWatBal model. An AML routine was developed to transfer these data into 365 grids representing 'typical' rainfall for each day of the year. The vector point data representing each day was initially converted to raster format with a 200 m cell resolution before using the *eucallocation* command to populate the surface (Figure 8). This routine calculates for each cell the zone of the closest source cell and can be seen in Figure 9. This method was chosen because data remain as discrete rainfall zones populated outward from each location. The rationale for this approach is that one month was selected to represent a 'typical' months rainfall from the ~ 40 years of data, but adjacent locations may have very different distributions and hence different daily rainfall for their selected month potentially from a different year. For example, on June the 1st one site may have no rain, while the adjacent site has 50 mm of rainfall for the same day. Consequently, because data were selected on representativeness of the number of rain days and total rainfall in a month, spatial interpolation between locations was considered inappropriate.

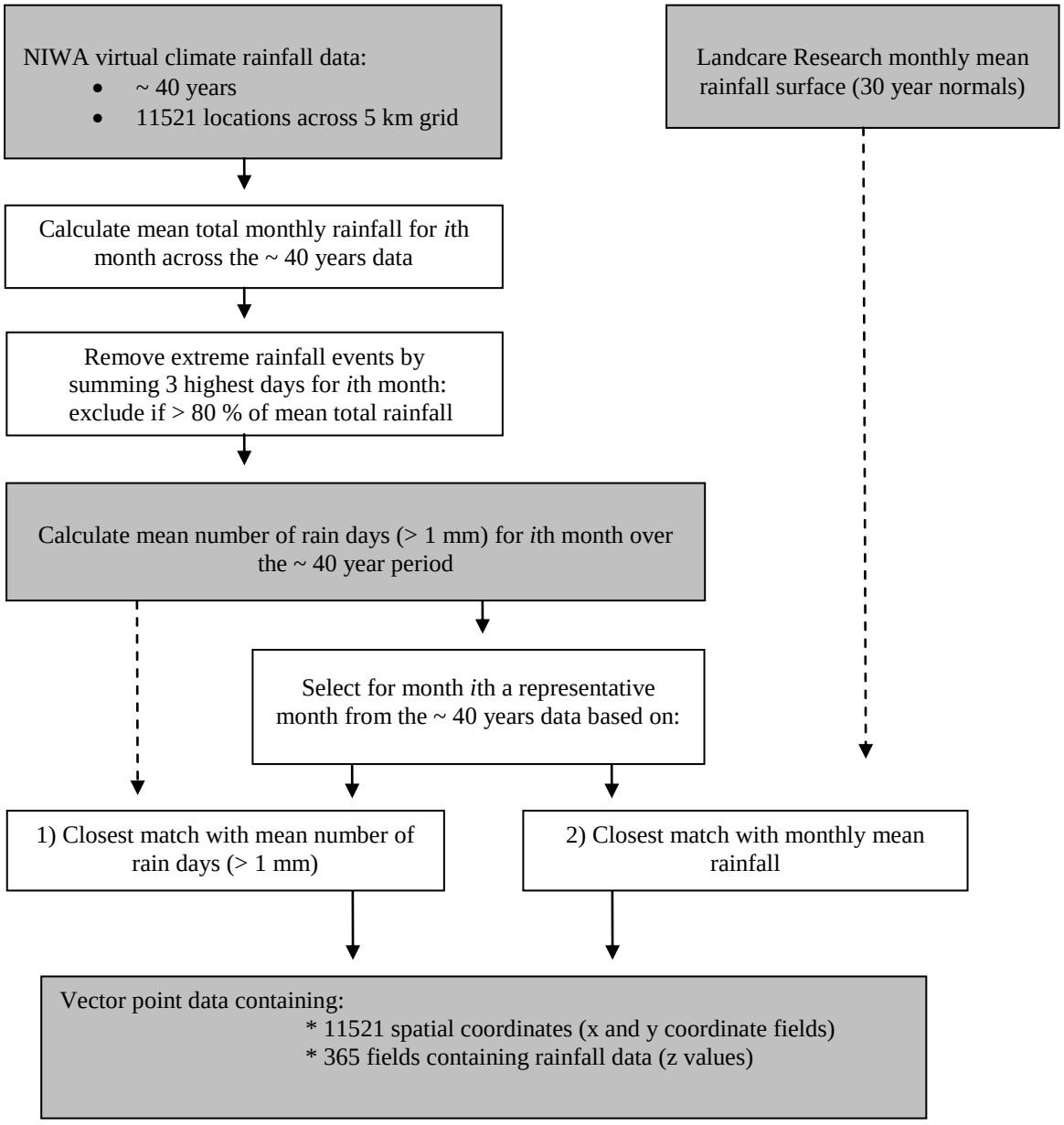


Figure 7: Flow chart illustrating processes and protocols in developing *RAV* data using a purpose built Visual Basic script that executes an iterative process for each of the 11521 locations across New Zealand.

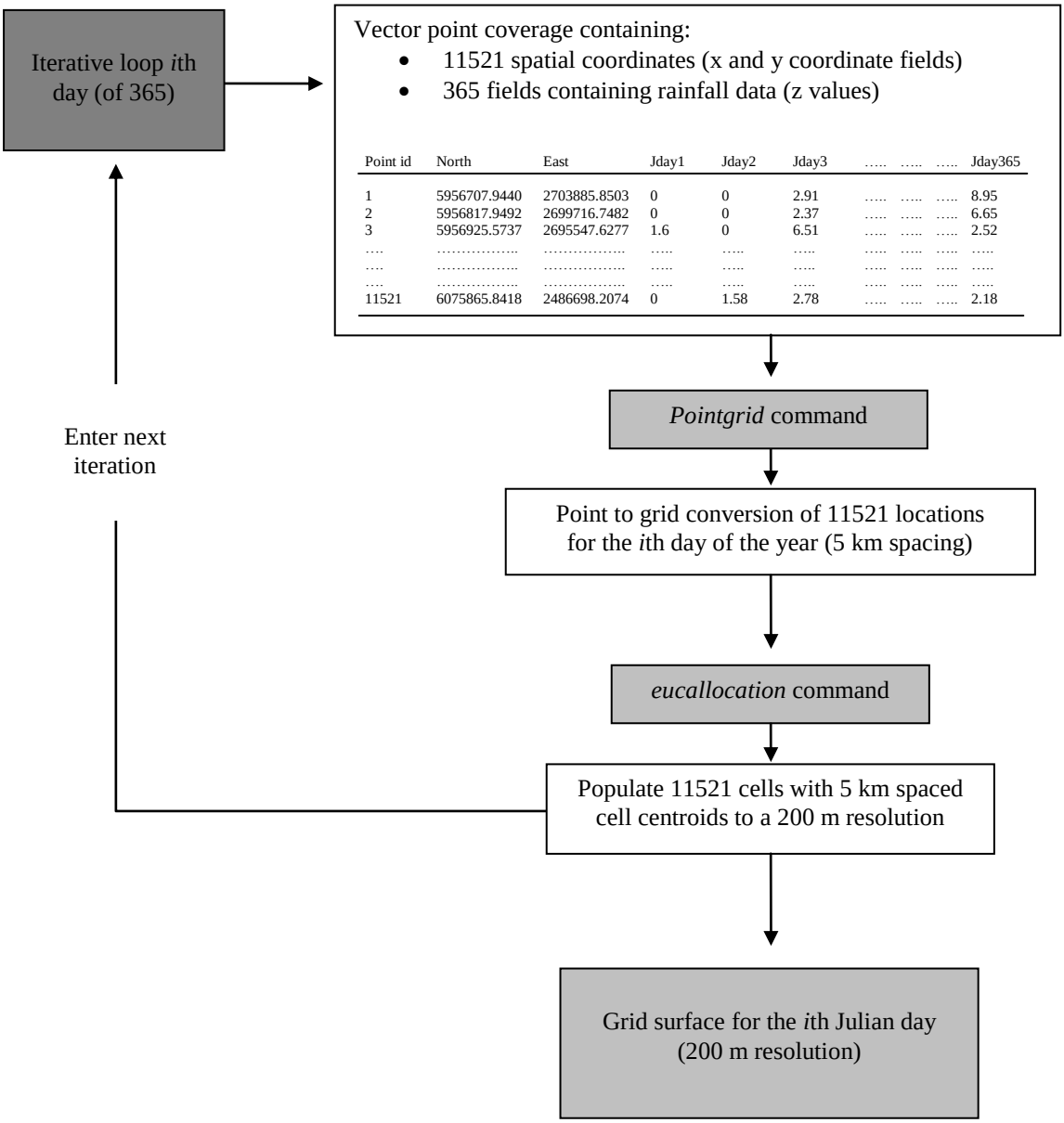


Figure 8: Flow chart showing the AML processes used in the creation of raster surfaces called reasoned allocated virtual surfaces (*RAV*) representing 365 days of rainfall.

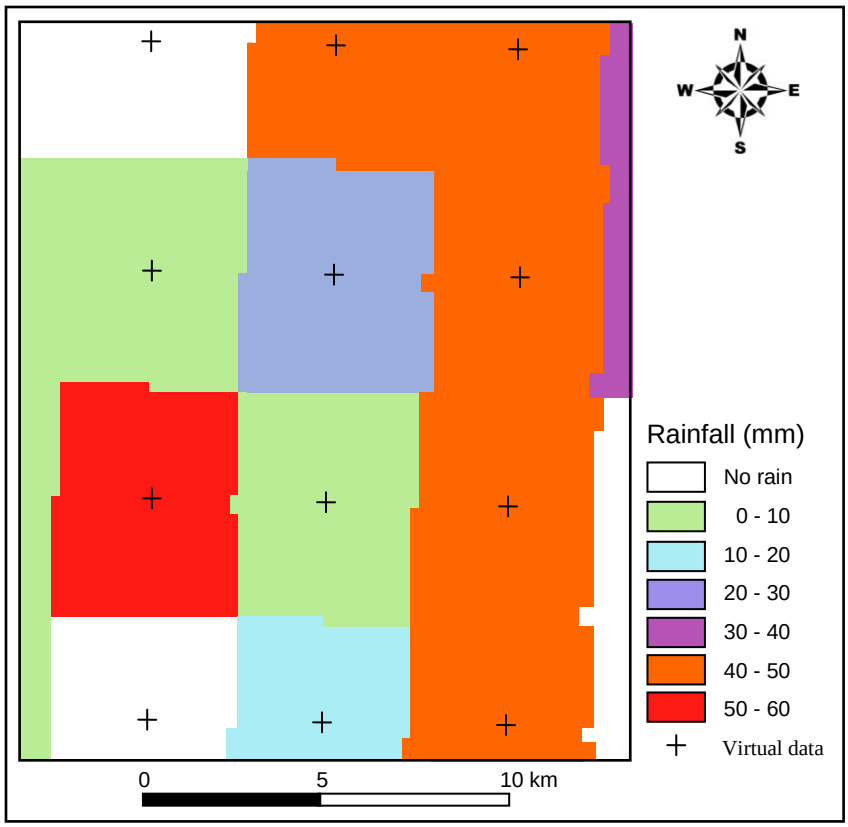


Figure 9: Graphic illustrating the populating of vector rainfall data at 5 km centres to a 200 m resolution raster surface – reasoned allocated virtual data for Julian day number 90. Note the jigsaw pattern produced from exact allocation of data out from each location.

3.6 Evaporation and transpiration surfaces

During the execution of the SWatBal model understorey evaporation and transpiration are estimated. Calculating evaporation and transpiration requires the input data vapour pressure deficit (D), transmitted radiation (G_g), psychrometric constant (γ) and (s), which is the slope of the relationship between saturated vapour pressure and temperature at a given air temperature. To simplify modelling processes and to reduce computational time, stand alone surfaces were developed for these properties using standard formulae.

During the calculation of evaporation and transpiration the SWatBal model calls each of the monthly surfaces, D, G_g, γ, and s as required and assumes a constant L_t value of 5. Refer to Section 6 for discussion surrounding the sensitivity of SWatBal simulations to change in input parameters and variables.

4. MODEL DESCRIPTION, PARAMETERISATION AND EXECUTION

4.1 SWatBal model description

A daily water balance model (SWatBal) was used to estimate root-zone water storage (W_i) on the ith day as,

$$W_i = W_{i-1} + P_i - E_{ti} - E_{twi} - E_{gi} - F_i \quad (3)$$

where P_i is daily rainfall, E_{ti} transpiration from the tree canopy, E_{twi} evaporation of intercepted rainfall from the tree canopy, E_{gi} evaporation from the soil surface, and F_i drainage from the root-zone (Whitehead et al., 2001; Watt et al., 2003b). Root-zone water storage ranges from a maximum (W_{max}) to a minimum, (W_{min}) value, and maximum available root-zone water storage, W_m, is defined as $W_{max} - W_{min}$. As W_m was available for the spatial modelling, W_{min} in all modelling was set to zero. Modelled values of W_i give the available root-zone water storage and when expressed as a fraction of W_m (=(W_i/W_m)*100), define the fraction of available root-zone water storage, W_f.

Transpiration from the tree canopy was calculated using a simple diffusion equation,

$$E_t = D\, g_{st}\, L_t \qquad (4)$$

where D refers to the air saturation deficit, g_{st} the average stomatal conductance for the canopy, and L_t is the half surface leaf area index. In response to increasing air saturation deficit, stomatal conductance decreases, which was modelled using the function described by Lohammer et al. (1980) as,

$$g_{st} = \alpha\, g_{stmax}\, (1+ (D - D_{smin})/ D_{0t})^{-1} \qquad (5)$$

where g_{stmax} is the maximum stomatal conductance, and D_{0t} is the sensitivity of g_{st} to D, when $D > D_{smin}$ (the value of D below which g_{st} is constant) and α will be defined later. Evaporation from the wet tree canopy, E_{twi}, ($=f_{wt}\, P$) was calculated as a constant fraction, f_{wt}, of rainfall.

Evaporation from the soil, E_g, was calculated from the available energy beneath the tree canopies, G_g, as,

$$E_g = \alpha\, \tau s G_g / [\lambda\, (s + \gamma)] \qquad (6)$$

where the term $s\, G_g/[\lambda(s + \gamma)]$ is the equilibrium rate of evaporation and the coefficient τ describes the degree of coupling of the soil surface with the air above the canopy

(Kelliher et al., 1990), s is the slope of the relationship found between saturated vapour pressure and the temperature at any given air temperature, λ the latent heat of vaporisation, and γ the psychrometric constant. Using Beers Law, G_g was determined from (e^{-kLt} G_a) where G_a represents the available energy above the canopy (assumed to be 70% of shortwave irradiance) and k is the light extinction coefficient (assumed to be 0.5 for a spherical leaf angle distribution).

The coefficient α was used to reduce evaporation from the soil and transpiration as soil water storage declined. The value of the coefficient α was set to 1 at maximum values of W (W_m), and was not reduced until W declined to a threshold value (W_t). As W_i progressively declined below this threshold α was linearly reduced from 1 to 0 at minimum values of W (=0). The threshold values (W_t) for reducing evaporation from the soil and transpiration, taken from Watt et al. (2003b), were assumed to be, respectively 60% and 55% of root-zone water storage, W_f (defined above).

Estimated evaporation from the soil and transpiration were reduced to 75% of their potential rates on days when rain fell. Drainage from the root zone was assumed to be zero when $W_i \leq W_m$ and equal to rainfall reaching the soil when $W_i > W_m$.

Meteorological data requirements for the SWatBal model are discussed in Section 3.2. The model also requires input parameters including g_{stmax}, D_{0t}, W_m, f_{wt}, W_t, and L_t (Table 6). Values obtained from the literature for these parameters include 0.1 mmol^{-2} s^{-1} and 0.6 kPa, respectively, for g_{stmax}, and D_{0t} (Whitehead et al., 1996; Richardson et al., 2002), 0.2

for f_{wt} (Kelliher et al., 1992), and W_t equal to 0.6 and 0.55 of W_m for E_g and g_{st}, respectively (Watt et al., 2003b).

Table 6: Values used for parameters in the SWatBal model.

Parameter	Value	Units
g_{stmax}	0.1	mol m^{-2} s^{-1}
D_{0t}	0.6	kPa
W_t for E_g	0.60 W_m	Unitless
W_t for E_t	0.55 W_m	Unitless
f_{wt}	0.2	Unitless
PRD	Extracted from *PRD* surface [1]	mm
W_m	Extracted from *PAW* surface [1]	m^3 m^{-3}

See text for full explanation of these symbols
[1] Refer to section 3.3

4.2 Arc Macro Language routine driving SWatBal

Arc Macro Language (AML) is used to undertake the daily soil water balance calculations. This iterative process uses a primary AML that 'calls' and executes three associated AMLs, ext.aml, watt.aml, and ratio.aml (Figure 10). Initially the do_watt AML sets the constant model input parameters including L_t, directory paths, and the start and end Julian day numbers of each month before extracting the annual data for the region of interest. The annual data include *PAW*, W_v, *PRD*, and latitude, and are used in all calculations across all months of the year.

SWatBal is designed to calculate soil water balance over the potential rooting depth of *P. radiata* for a catchment or subcatchment, of any size, as required by the user. The SWatBal program is an iterative process calculating soil water balance for every day in a year. With each iteration SWatBal automatically extracts the required data allowing large regions (e.g. New Zealand) to be divided into grids of appropriate size aligned with the researcher's computer computational power and digital storage capacity. On completion of the modelling process, overlapping grids can be merged into one continuous surface representing the soil water balance for the forested region. The ext.aml extracts monthly data for the *i*th month including D, G_g, γ and s. The same AML extracts rainfall data for every day in the *i*th month from the *RAV* surfaces.

The watt.aml takes all previously extracted data to calculate a soil water balance. This process begins on the first day of July, where soils are assumed to be fully recharged, by developing the output surfaces W_a and drainage for every day in *i*th month. The final AML-ratio is called at the end of each month with the purpose of determining the monthly mean W_a and drainage. Finally, W_f is calculated for each month by W_i/W_m.

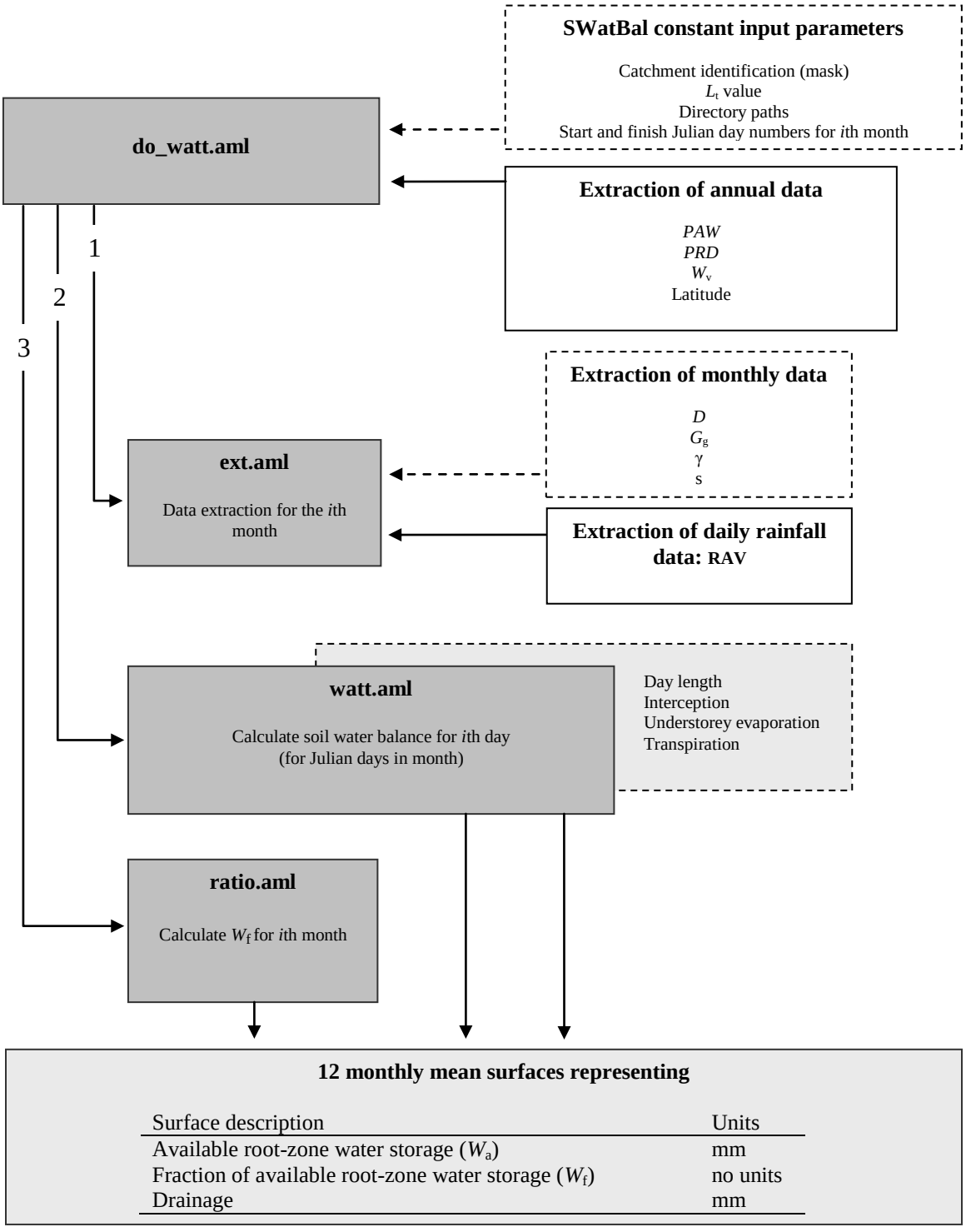


Figure 10: Flow chart illustrating the data extraction and soil water balance calculation processes involved in the SWatBal model.

5. SOIL WATER BALANCE (SWATBAL) SURFACES FOR NEW ZEALAND

The following section contains maps illustrating monthly mean W_a, W_f and drainage at a 100 m-resolution across New Zealand. There are 12 monthly and one annual map for each W_a, W_f and drainage attribute, providing a total of 39 surfaces. To display and discuss all maps in this section would provide unnecessary detail. Consequently, only the months of January and July are shown here, providing the reader with an overview of the likely extreme months (wettest and driest) to highlight their differences. For detailed maps refer to appendices.

5.1 Spatial and temporal variation in rainfall and soil water storage

Pinus radiata productivity remains an important issue for forestry managers in New Zealand. Increased, or indeed diminished, growth of *P. radiata* is directly related to a site's ability to provide resources to a forestry crop (Boomsma and Hunter, 1990). These resources are numerous; however, it is generally accepted that intercepted solar radiation, and the soil's ability to retain and supply adequate nutrients and water to a demanding crop profoundly influence forest productivity. Although an adequate supply of water is fundamental to all plants, measures of rainfall alone provide useful, but limited, information when attempting to understand the relationship between rainfall and crop productivity. The next logical step was the development of a model that accounts for the

spatial variation of soil water balance enabling comparisons of sites across New Zealand. Furthermore, SWatBal delineates sites with soil water surpluses, typically maximising crop productivity, and deficits that may reduce crop productivity.

The central graphic in Figure 11 illustrates spatial variation of the annual fraction of available root-zone water content (W_f) in an east to west direction across the South Island of New Zealand. This soil water balance cross section was draped over a DEM to provide the reader with a sense of topography, highlighting the western regions with higher moisture contents throughout the year (80-100), while eastern regions are in deficit for extended periods of the year (< 60). The two graph inserts highlight on a month-by-month basis mean monthly W_f and rainfall occurring at arbitrarily chosen locations for the Hokitika and Canterbury regions. The Hokitika graph shows a relatively consistent soil water balance with all months close to full soil water recharge. Rainfall for this location is also high with values from 220 to 322 mm for July and December, respectively. In contrast, the Canterbury graph shows a location with strong seasonal variation with summer months in deficit, while reaching close to full soil water recharge during July with values of 27 to 95. Rainfall for the Canterbury location is less than 70 mm a month, with the lowest value of 48 mm occurring in September. The overall spatial variation typifies the west to east moisture patterns of New Zealand's South Island with the Southern Alps intercepting and receiving much of the rainfall during prevailing westerly winds leaving the eastern regions comparatively dry.

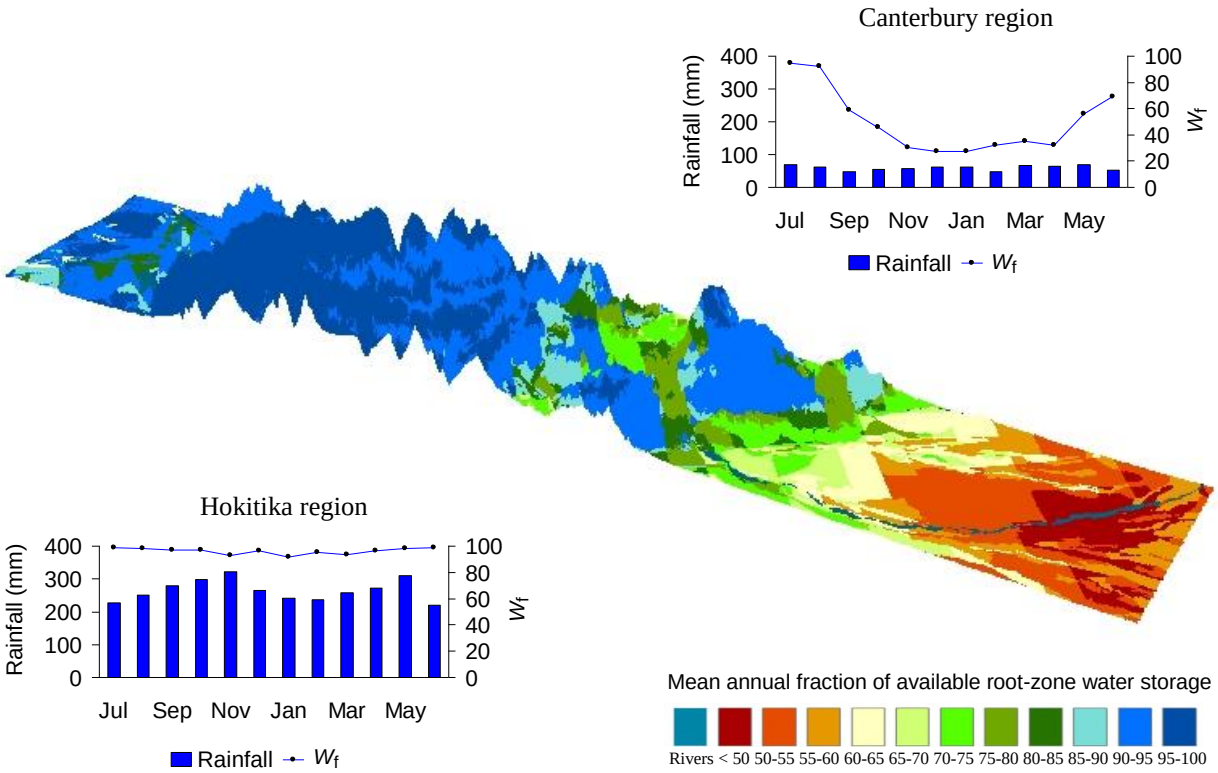


Figure 11: Cross section of the South Island from the West Coast to the Canterbury Plains in the east illustrating the mean annual fraction of available root-zone water storage (W_f). Graphs highlight the total monthly rainfall (bars) and mean monthly fraction of available root-zone water storage (lines) for the Hokitika and Canterbury regions. Total rainfall was derived from *RAV* data and mean monthly W_f was calculated using SWatBal.

5.2 Mean monthly fraction of available root-zone water storage

Mean monthly W_f was calculated as W_a / W_m, expressed as mm per mm (units cancel out). Determining mean monthly W_f enables the comparison of sites across New Zealand not possible with mean monthly W_a alone. The W_f value is the fraction of available root-zone water storage relative to the maximum root-zone water storage. The spatial variation across New Zealand for the mean W_f can be seen in Figure 12 for January and July. Soil water deficits are common features across many regions of New Zealand during January. Regions with deficit values less than 30 are Central Otago, South Canterbury, the Canterbury Plains, North Canterbury, Blenheim, Wairarapa, Hawke's Bay, Gisborne, and parts of Northland (Figure 12). Much of the east coast of the South Island and southern North Island, Nelson, the Manawatu and Northland regions suffer deficits to some extent for the month of January. In contrast, the majority of New Zealand reaches full soil water recharge during July with the exception of central Otago (Figure 12). It is evident that ~50 % of New Zealand never experiences soil water deficits even during January. The implications of summer soil water deficits at the national extent may be realised in diminished crop productivity (see Figures 1 and 2 for place and region locations).

The SWatBal model will be useful to study in many fields including crop productivity, hydrology, leaching and groundwater contamination, erosion, plant ecology and biodiversity. Furthermore, SWatBal allows the comparison of moisture content across many spatial extents including forest stand, compartment, catchment, or landscape level analysis. SWatBal model development allows user flexibility and versatility, and with the

advent of new and improved data the model can be rerun. Additionally, should the practitioner require surface creation at different temporal scales the model is flexible enough to handle daily, weekly, fortnightly, or monthly time scales. This flexibility is also true for different spatial scales where a forest researcher has captured high quality data for a specific forest stand, compartment, catchment or sub-catchment, allowing re-execution of the SWatBal model at finer resolutions.

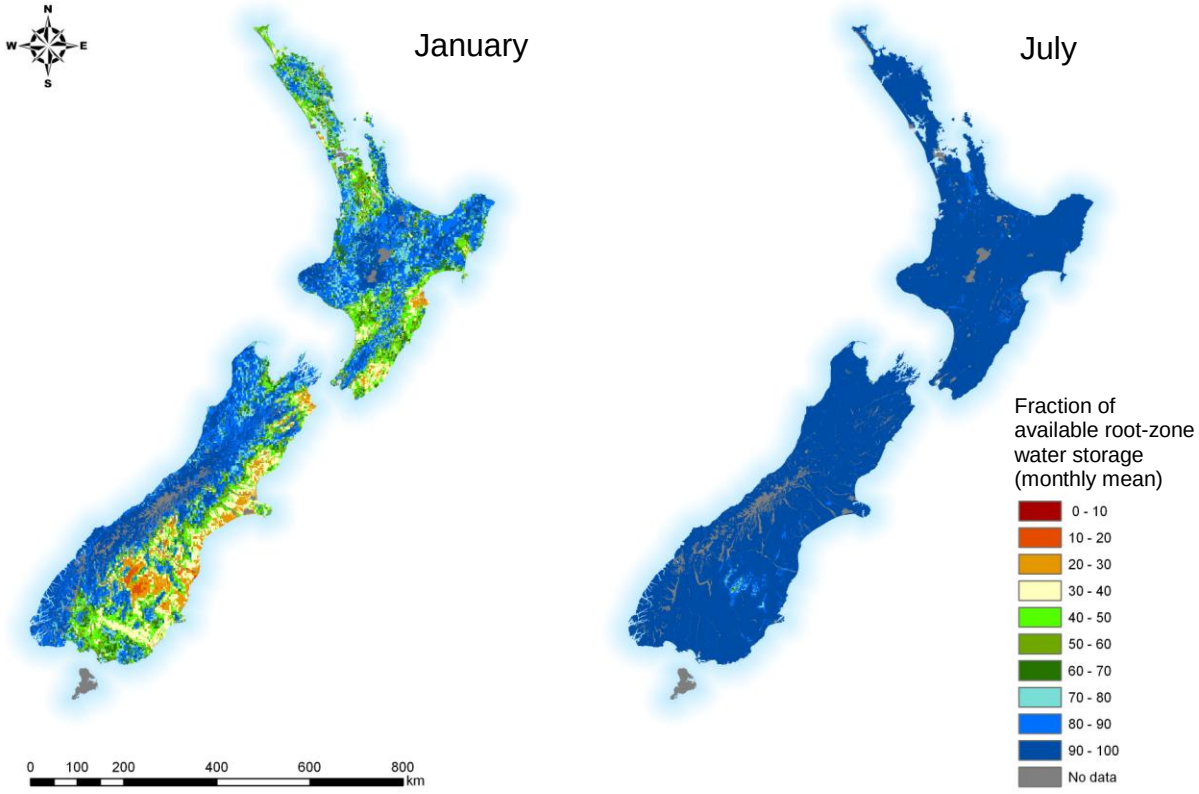


Figure 12: Geographic variation in the mean fraction of available root-zone water storage (W_f) for the months of January and July across New Zealand.

Finally, SWatBal surfaces provide visualisation of spatial patterns whether at the national extent (Figure 12) or more detailed view of any location across the country (Figure 13 - Napier), essentially by "zooming in" at an expanding spatial scale until reaching the limit of the operational scale. Potentially forest managers can view a close up of the forest of interest or assess the feasibility of forest development at different locations.

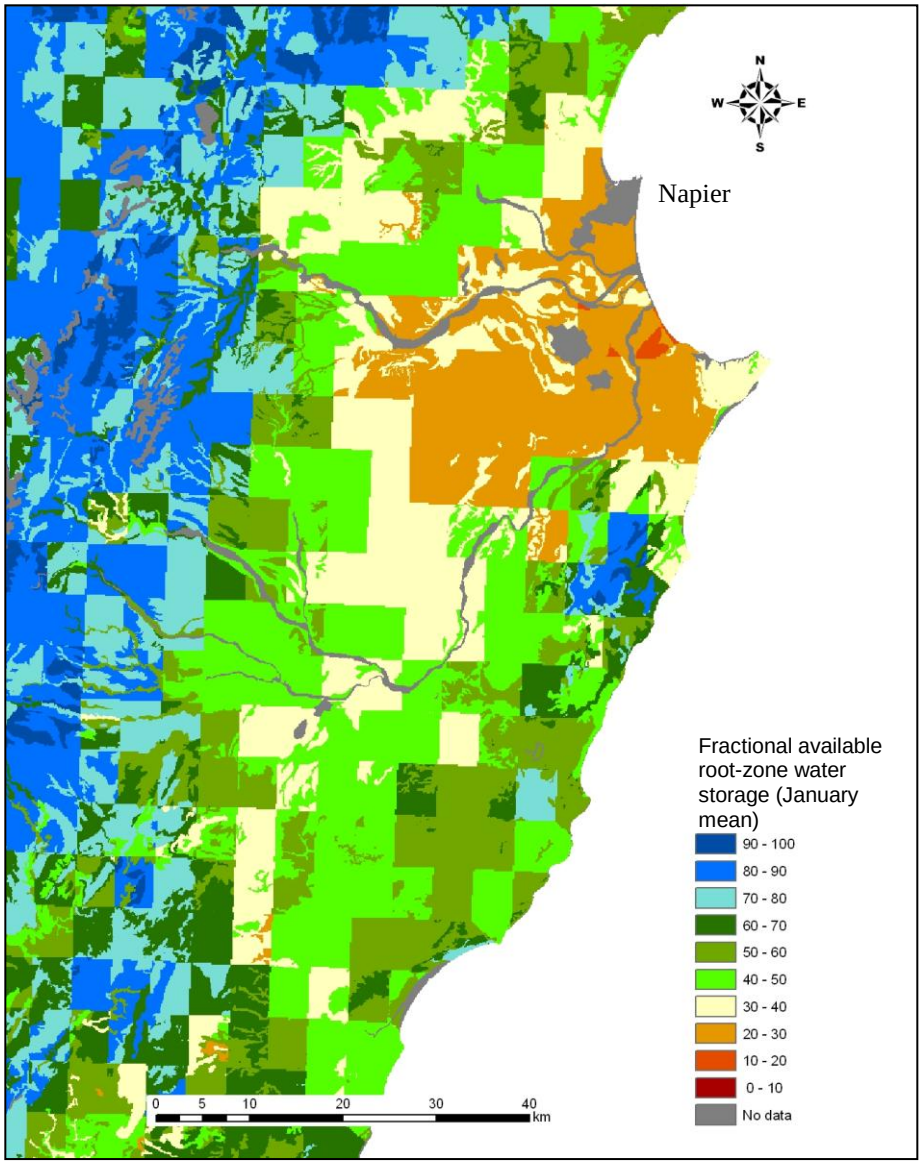


Figure 13: Geographic variation in the mean fraction of available root-zone water storage (W_f) during January for part of the Hawke's Bay region. The map is drawn at a scale of 1:400,000.

5.3 Mean monthly available root-zone water storage

Available root-zone water storage (W_a) provides a different perspective from W_f with mean monthly W_a representing the monthly mean available water (mm) within the "effective rooting depth". Figure 14 highlights the spatial variation of mean available root-zone water storage (W_a) across New Zealand for the months of January and July. Soils with the greatest W_a for the months of January and July are in the Taranaki and Taupo regions with greater than ~225 mm mean available root-zone water storage. In contrast, the Otago region retains the lowest W_a values of less than ~25 and ~50 mm mean available root-zone water storage for January and July, respectively (Figure 14). Assessing mean monthly W_a is less intuitive than W_f data because direct comparisons can only be made at specific locations. For example, the Balmoral series of North Canterbury typically has a W_{max} of ~ 37 mm available water over an effective rooting depth of ~0.3 m (Arneth et al., 1998a). In comparison the Haupapa series in the central North Island typically has a W_{max} of ~380 mm of water available within an effective rooting depth of ~1 m (Kelliher et al., 1992).

Although W_a remains less intuitive than the derivative W_f, it can provide information useful in water balance calculations when applying the latest NIWA virtual climate data, including rainfall requirements to fully recharge soil water and climate change simulations. Furthermore, W_a surfaces can be used to assist in identifying locations suited to the production of specific tree species. For example, *Cupressus lusitanica* is becoming a popular alternative to *P. radiata* because of its high productivity and versatile uses,

including exterior joinery and panelling (Haslett, 1986). Although the *C. lusitanica* species is considered to demand greater site resources that its *P. radiata* counterpart, it also requires well drained soils of at least moderate fertility (Clifton, 1994).

Suitability maps could also be beneficial to the commonly grown Douglas fir species (*Pseudotsuga menziesii*), which are currently grown on cooler, drier sites (Fahey et al., 2001). As work progresses in the area of defining tree species growth response in relation to volumetric water content, models like CLIMEX (Sutherst et al., 1999) could be used to develop and model optimal tree species distribution for New Zealand. An applied example of CLIMEX and its potential to model species distribution with climate change can be seen in Kriticos et al. (2003).

Tree-species suitability maps based on crop water demand could be produced identifying and delineating regions for resource allocation and planning for foresters, resource managers, policy and resource planners including local, regional and central Government. Other possible uses for SWatBal include forest carbon sequestration modelling, climate change modelling, and numerous other applications including erosion, leaching, and ecological studies.

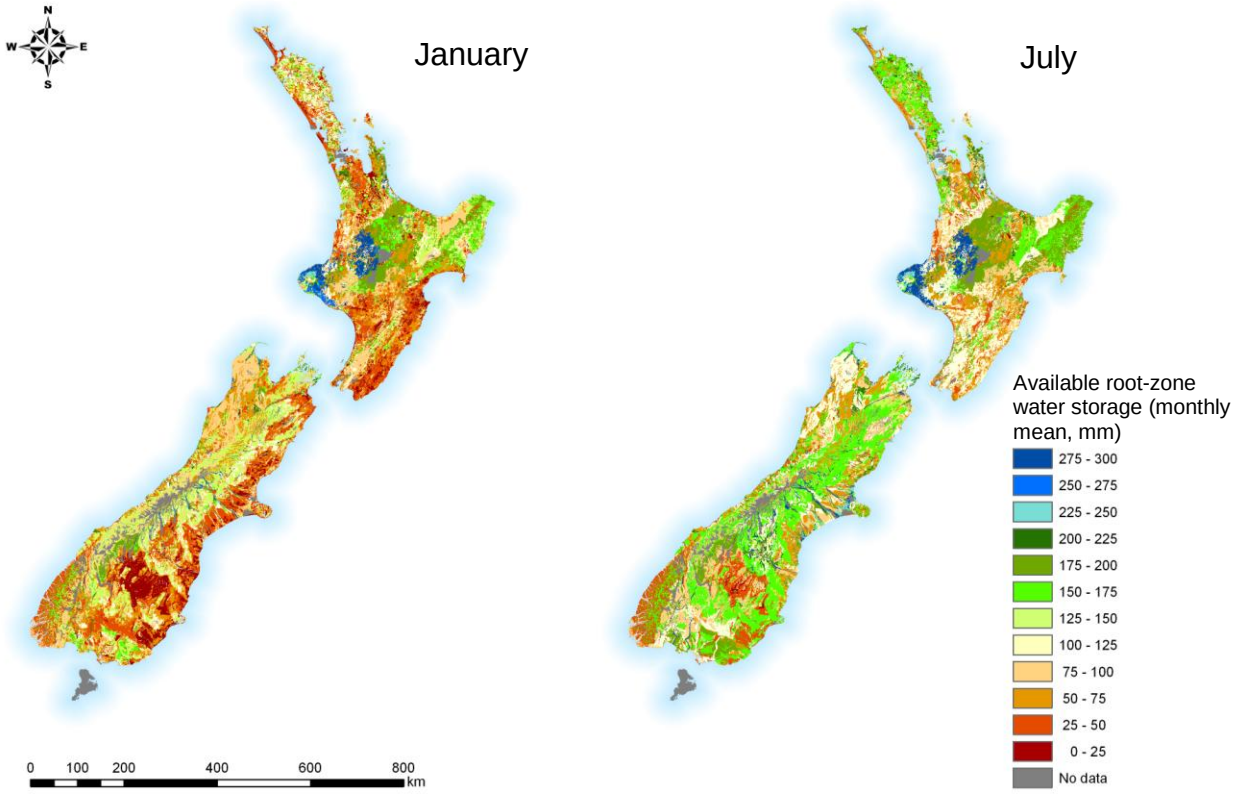


Figure 14: Geographic variation in the mean available root-zone water storage (W_a) for the months of January and July across New Zealand.

5.4 Mean monthly drainage

The drainage surfaces produced during the SWatBal model simulations provides information useful to many disciplines, especially in hydrologically related fields. Figure 15 illustrates the mean monthly drainage in mm across the extent of New Zealand for January and July. As would be expected drainage during the month of January is minimal for most locations across New Zealand. High drainage rates predominate along the West Coast of the South Island and the adjacent mountainous interior. However, for the

remainder of the South Island during January, drainage rates are minimal with less than a monthly average of 2 mm. The North Island drainage values for January are minimal with < 2 mm mean drainage values for most locations. For July, mean drainage values remain minimal with < 2 mm throughout the Canterbury, Otago and Southland regions. The lower slopes of the Southern Alps, the Port Hills and the northern parts of the South Island increase marginally to the 2 to 4 mm class. In contrast, North Island average drainage rates for July increase in all regions (2 to 10 mm) with the exceptions of the Hawke's Bay and Manawatu regions, which remain unchanged.

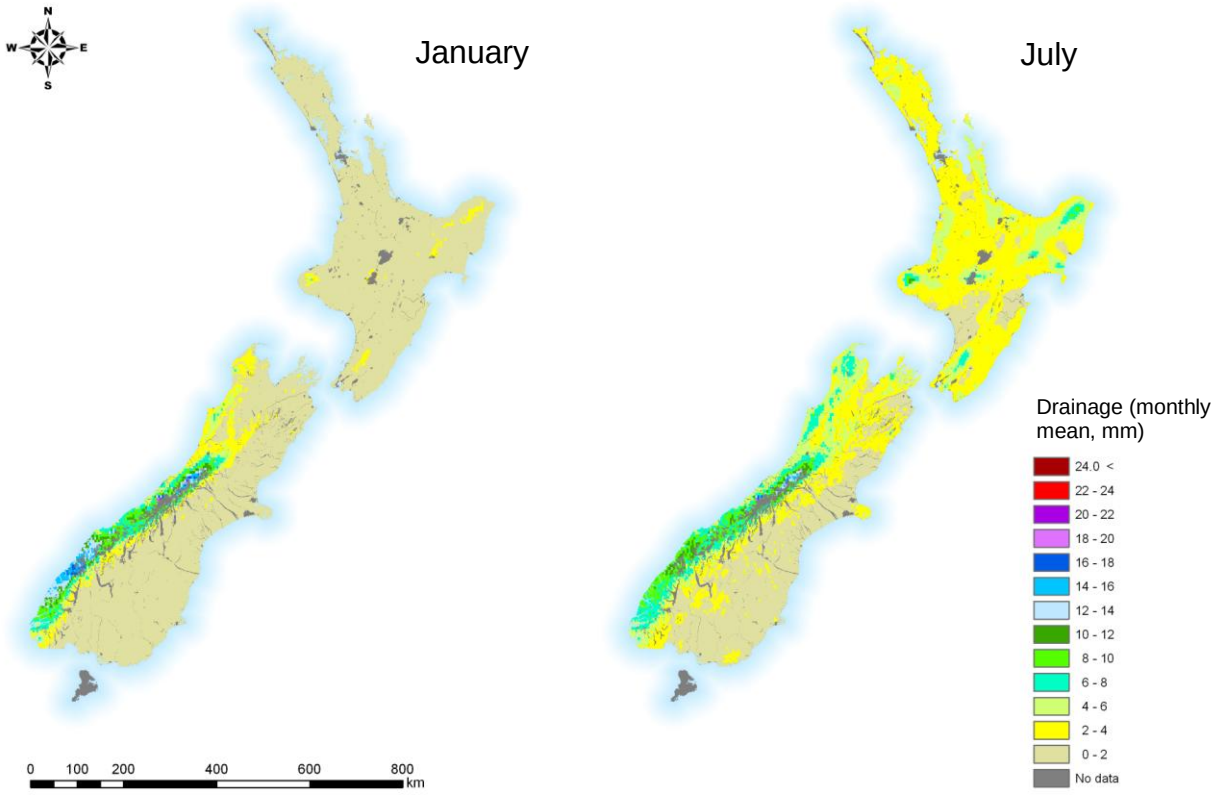


Figure 15: Geographic variation in the mean drainage for the months of January and July across New Zealand.

The SWatBal model calculates drainage on a daily basis, which is effectively the water that is surplus to requirements for that period. Much of this drainage occurs during peak rainfall events leading to these periodic events becoming 'averaged out' during monthly calculations. Consequently, mean monthly drainage rates may have lost some of their detail required for other modelling processes. Other inherent impacts for the drainage model surfaces are the selection of daily *RAV* rainfall data. During *RAV* data preprocessing extreme rainfall events were removed from the NIWA virtual rainfall data. Not to detract from the overall importance and visual impact of the mean monthly drainage surfaces, research related to hydrological terrain based studies would benefit from rerunning. SWatBal could use appropriate temporal scales, with either empirically measured field data where available or, where these are not available, by using NIWA's interpolated virtual rainfall surfaces for the period of interest. Another approach to reduce the impact of monthly averaging is the use of cumulative drainage over a specified period. Terrain based flow modelling would also require partitioning drainage into the various components, for example runoff verses infiltration rates requiring information related to physical properties like slope, aspect, curvature, and hydraulic conductivity. Considering the availability and accuracy for many of these datasets, this type of modelling may remain better suited to the catchment and sub-catchment extent where modelling at finer resolutions is appropriate.

6. SWATBAL VALIDATION AND SENSITIVITY ANALYSIS

The purpose of this section was to (i) validate the SWatBal model simulations against measured soil water storage data, from published studies on water balance within *P. radiata* stands, and (ii) evaluate the sensitivity of the SWatBal model to changes in biological and physical input parameters and variables.

6.1 Methods

6.1.1 General description of the Burnham and Haupapa sites

Two sites were chosen for validation because of their contrasting soil water storage. The Burnham forest site was located on the Canterbury Plains (43^{o} 36'S, 172^{o} 18'E) and characterised by summer soil water deficits with an average rainfall of around 590 mm per year (Figure 16), a stony silt loam soil texture with low soil water storage, and an udic moisture regime in "Soil Taxonomy" (Soil Survey Staff, 1999). At Burnham, the soil belongs to the Lismore series (Brown Soils in New Zealand Soil Classification (NZSC: Hewitt, 1998), or Inceptisols in "Soil Taxonomy"). Data used for the Burnham study (Richardson et al., 2002) covered a three year period from a stand age of 9 years, at which time stand density was 610 stems ha^{-1}.

The Haupapa site (Kelliher et al., 1992) is located on the volcanic plateau in the central North Island (38^{o} 33'S, 176^{o} 25'E) (Figure 16). The soil at this site is a free draining pumiceous loamy sand to sand of the Taupo series (Pumice Soils in NZSC, or Vitrands in

"Soil Taxonomy"). Typical site characteristics include a relatively high mean annual rainfall of around 1260 mm, evenly distributed across the year, very high soil water storage, and an udic moisture regime (Soil Survey Staff, 1999). A strong feature of the Haupapa study is that the soil water storage is in surplus across the majority of the year. At this site measurements of soil water storage were taken over the course of a year in a 7-year-old plantation, at which time the stand density was 450 trees ha^{-1}.

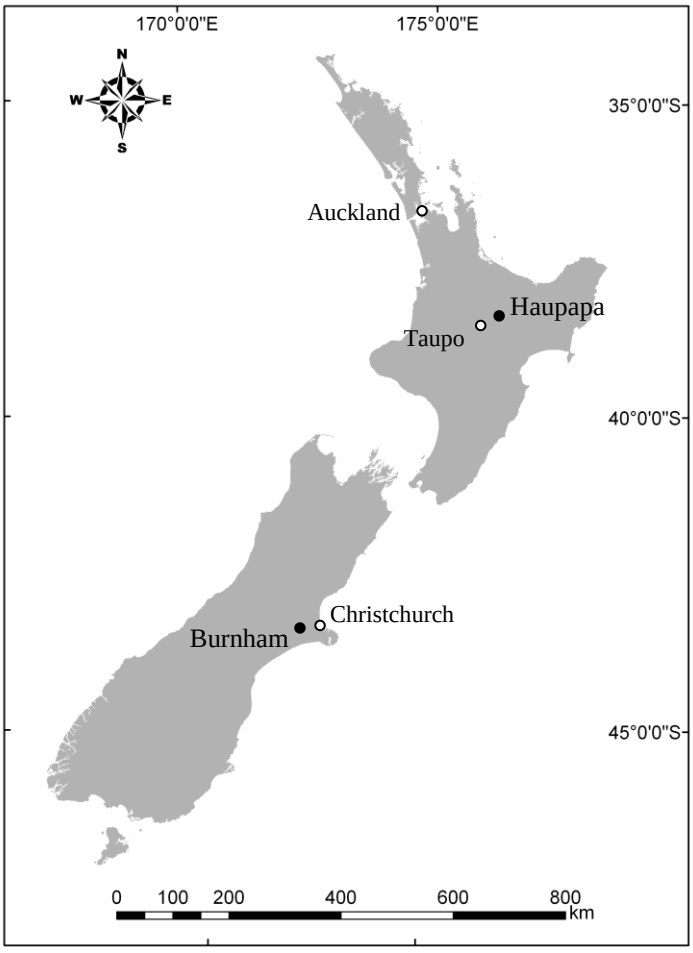


Figure 16: Map showing the Burnham and Haupapa forest plantation locations.

6.1.2 Experimental design

At the Burnham and Haupapa sites, measurements of available root-zone water storage (W_a) were available for their study periods (Kelliher et al., 1992; Richardson et al., 2002). W_a site measurements were converted to fraction of available root-zone water storage (W_f) as defined in Section 4.1 to allow comparisons with modelled data (filled circles in Figure 17 and 18). Three SWatBal models were calculated to determine W_f at the Burnham and Haupapa sites using the input variables and additional parameters air saturation deficit (D), rainfall (P), transmitted radiation (G_g), maximum available root-zone water storage (W_m), and potential rooting depth (PRD) from (1) field measurements used in their respective published studies, (2) field measurements from their respective published studies, but rainfall substituted with *RAV* data (see Section 3.5 for RAV development details), and (3) SWatBal data extracted from climate surfaces and fundamental soil layers (FSLs) representative of their site soil attributes. Leaf area index (L_t) values used in models (1) and (2) were from Richardson et al. (2002) and Kelliher et al. (1992) for the Burnham and Haupapa sites, respectively. The third SWatBal model assumed a constant L_t value of 5 with other input parameters fully described in Section 4.1. Sensitivity analysis used SWatBal model (3) as the control and adjusted input values and parameters independently by ± 25 and 50 % to assess the impact on modelled W_f values (deviation from the control model).

6.2 Results

6.2.1 Validation of the SWatBal model

Using actual meteorological data there is close correspondence between measured and modelled values at the Burnham site (Fig. 17A). Substitution of Burnham rainfall data with *RAV* rainfall data does not unduly diminish correspondence between modelled and measured W_f (Figure 17B). When spatially interpolated long-term climate surfaces and FSLs are used as inputs to the SWatBal model, a reasonable level of correspondence between modelled and measured W_f is observed across the three-year study period (Figure 17C).

The SWatBal model simulations using meteorological data for the Haupapa site show a close correspondence between measured and modelled values (Figure 18A). Substitution of the meteorological rainfall data with *RAV* rainfall data shows a good correspondence between modelled and measured W_f across the winter and spring periods, but abnormalities occur across the summer and autumn periods (Figure 18B). When spatially interpolated long-term climate surfaces and FSLs are used as inputs to the SWatBal model, the level of correspondence between modelled and measured W_f remains relatively good across all months with the exceptions of January and March (Figure 18C). Overall, the SWatBal model simulations using spatially interpolated long-term climate surfaces and FSLs (Figure 18C) performed better than with RAV rainfall data alone (Figure 18B).

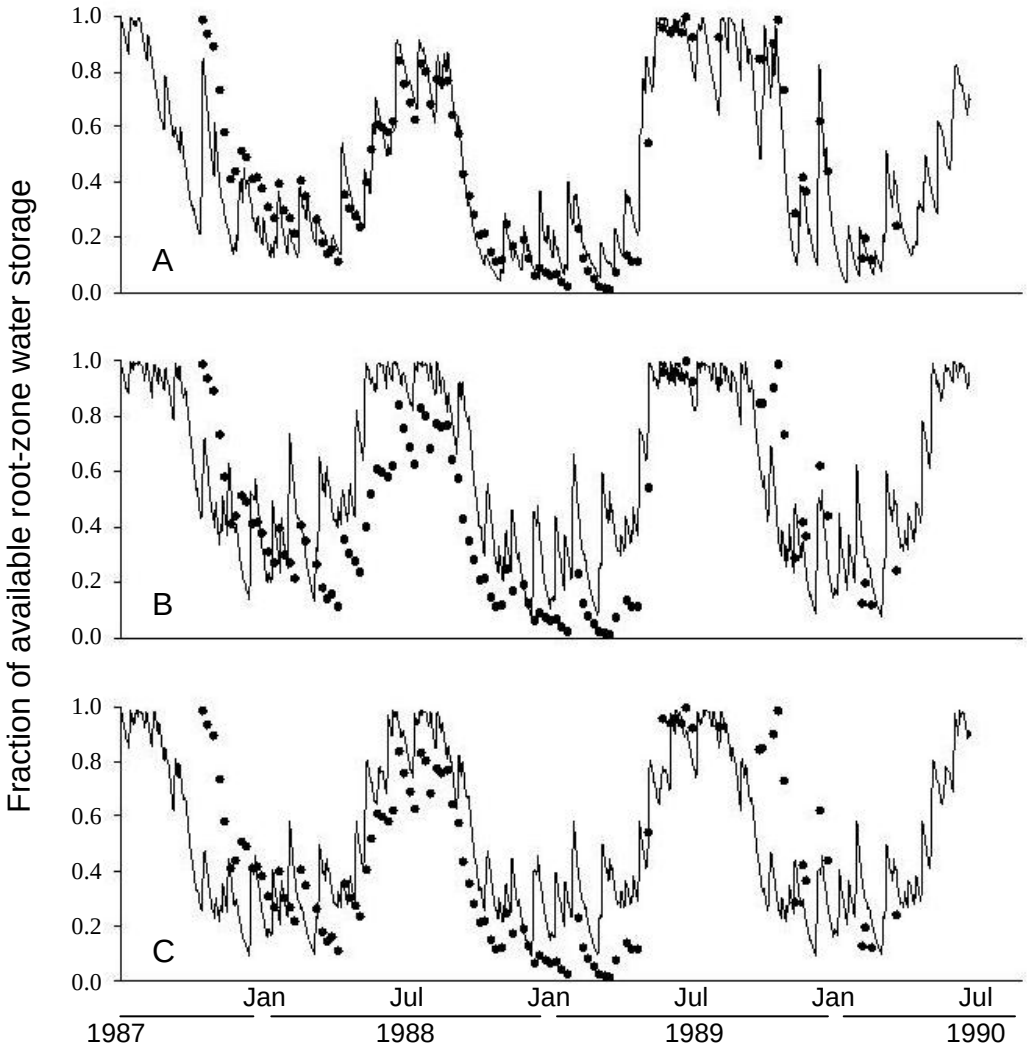


Figure 17: Modelled (line) and measured (filled circles) fraction of available root-zone water storage (W_f) for a dryland site at Burnham for three years from 1987 to 1990 using (A) field collected data, (B) field collected data but rainfall substituted with *RAV* data, and (C) SWatBal data.

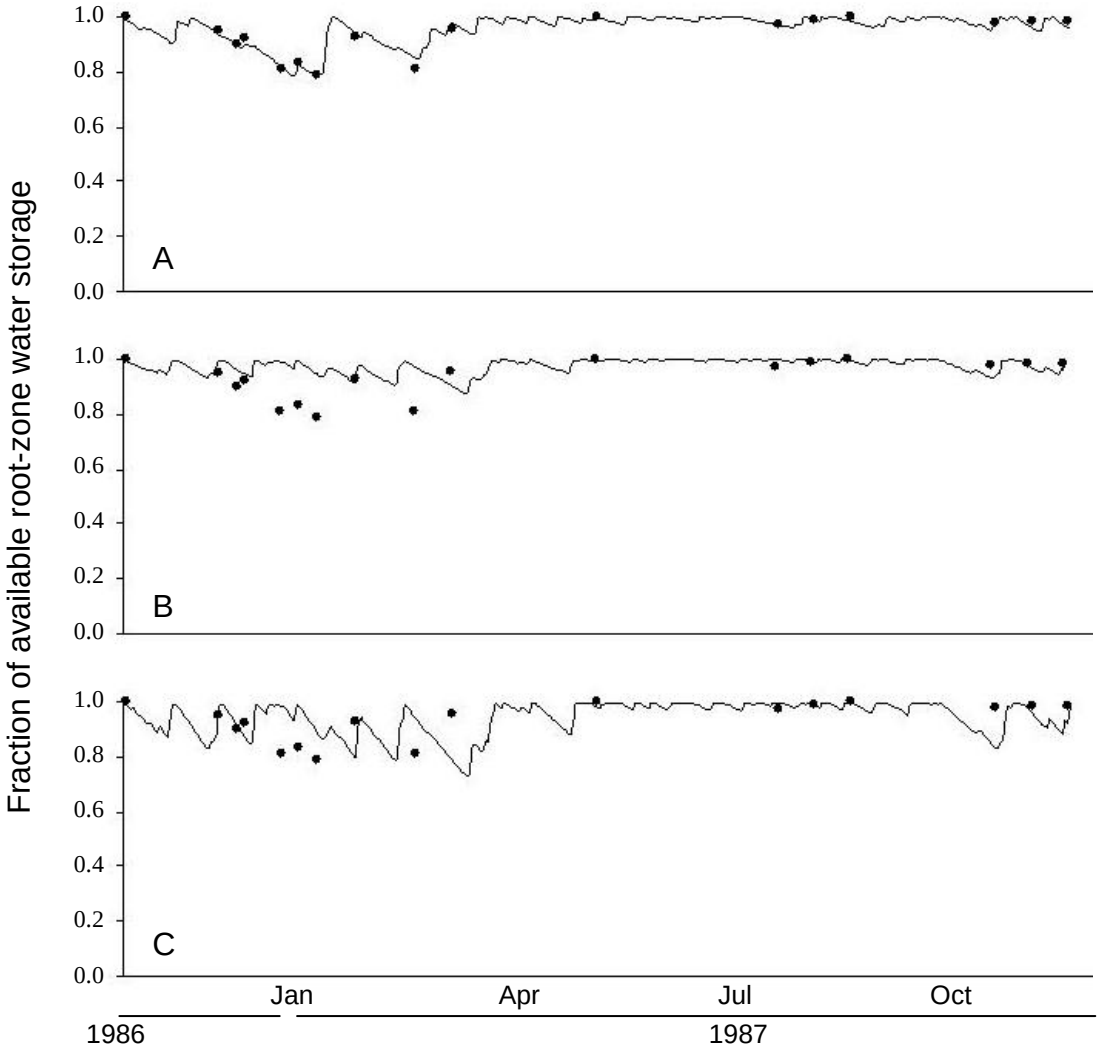


Figure 18: Modelled (line) and measured (filled circles) fraction of available root-zone water storage (W_f) at Haupapa across one year for the 1986 to 1987 period using (A) field collected data, (B) field collected data but rainfall substituted with *RAV* data, and (C) SWatBal data.

6.2.2 Sensitivity of SWatBal to changes in model parameters and variables

All major contributing model parameters and variables were considered in the sensitivity analysis of SWatBal W_f predictions. Sensitivity analysis was undertaken by adjusting precipitation (*P)*, *P. radiata* leaf area index (L_t), maximum available root-zone water storage (W_m), maximum stomatal conductance (g_{stmax}), air saturation deficit (D), and the sensitivity of *P. radiata* g_{st} to D (D_{ot}) by ± 25 and 50%.

For the Burnham site, modelled W_f showed the greatest sensitivity to changes in P, g_{stmax}, L_t, and to a lesser extent D (Table 7). When rainfall was increased by 25 %, their was an overall increase in W_f by 9 %. However, decreasing P by 25 and 50 % produced a greater negative response in W_f, reducing values by 18 and 29 %, respectively. The SWatBal model was also highly sensitive to changes in g_{stmax} with 25 and 50 % increases in this parameter significantly reducing estimates of W_f by 12 and 20 %, respectively. Conversely, 25 and 50 % decreases in g_{stmax} increased W_f, respectively, by 11 and 30 %. The SWatBal model was less sensitive to changes in L_t with 25 and 50 % increases only reducing W_f by 9 and 19 %, respectively. When D was changed by (±) 25 %, an inverse response of ~ 5 % occurred to W_f. Further increases to D showed little change to W_f however, a 50 % decrease provided a 16 % increase to W_f. On an annual basis W_f is relatively insensitive to changes in W_m and D_{ot}.

Table 7: Deviation from the control for modelled annual fraction of available root-zone water storage (W_f) when the input variables are altered (± 25 and 50%) at the Burnham and Haupapa sites.

	Burnham				Haupapa			
Variable[1]	+ 50%	+ 25%	- 25%	- 50%	+ 50%	+ 25%	- 25%	- 50%
P	0.17	0.09	-0.18	-0.29	0.02	0.01	-0.08	-0.32
L_t	-0.19	-0.09	0.05	0.08	-0.10	-0.04	0.01	0.01
W_m	-0.01	-0.01	-0.02	-0.03	-0.04	0.00	-0.02	-0.07
g_{stmax}	-0.20	-0.12	0.11	0.30	-0.16	-0.06	0.02	0.04
D	-0.07	-0.04	0.05	0.16	-0.01	-0.01	0.01	0.03
D_{ot}	-0.02	-0.01	0.02	0.05	-0.02	-0.01	0.01	0.02

[1] P refers to precipitation, L_t *Pinus radiata* leaf area index, W_m maximum available root-zone water storage, g_{stmax} maximum stomatal conductance, D air saturation deficit, and D_{ot} sensitivity of *P. radiata* g_{st} to D.

For the most part, W_f at the Haupapa site was only sensitive to extreme changes in input variables and parameters. W_f was only responsive to a 25 and 50 % decrease in P with decreases in W_f of 8 and 32 %, respectively. The SWatBal model also showed some sensitivity to changes in g_{stmax} with 25 and 50 % increases reducing estimates of W_f by 6 and 16 %, respectively. However, it took a 50 % increase in L_t to produce a 10 % decrease in W_f. For W_m it took a 50 % decrease for a 7 % decrease in W_f. On an annual basis W_f was found to be relatively insensitive to changes in D and D_{ot}.

6.3 Discussion

Although it was only possible to use two validation sites, the overall utility of the SWatBal model is clear. Firstly, when the SWatBal model was tested using data from the original Burnham and Haupapa studies, W_f predictions showed good agreement with the site measured data, demonstrating that the SWatBal model performed well at these

locations (Figures 17A, and 18A). This level of agreement indicates that the model is correctly formulated and that key parameter values are robust. Secondly, when the SWatBal model uses spatially interpolated long-term climate and soil physical data (FSLs), modelled values continue to closely correspond to measured values, although the strength of the relationship is slightly reduced (Figures 17C, and 18C). This close correspondence suggests that spatially interpolated long-term climate data provide a useful approximation of actual climate data, providing confidence in the general applicability of the model.

Considering SWatBal simulations are derived from long-term climate averages and the comparison was made with field measurements considered inherently variable, SWatBal model simulations of W_f are considered good. Richardson et al. (2002) noted that rainfall during the first two years of the Burnham study was lower (~470 mm) than the annual average. This lower than average rainfall could for the most part explain the model overestimation of W_f using long term average measured rainfall in Figures 17B and C.

Overall, W_f was most sensitive to changes in input variables and parameters at the Burnham site where summer water deficits are common, compared with the Haupapa site where seasonal deficits are minimal. Rainfall is normally considered highly variable and sensitivity analysis showed that even a 25% change in rainfall will impact on modelled W_f at the Burnham site. However, decreasing rainfall by 50 % was required before the Haupapa site showed any substantial change.

W_f also showed high sensitivity to changes in g_{stmax} at the Burnham site. However, a substantial increase in g_{stmax} was required before any major response to W_f at the Haupapa site became apparent. Published g_{stmax} values for the *P. radiata* species are generally considered to vary between 0.25 mol m^{-2} s^{-1} (Watt et al., 2003) and 0.1 mol m^{-2} s^{-1} (Richardson et al., 2002), for juvenile through to mature *P. radiata* stands. However, there are exceptions. For a detailed account of published g_{stmax} values and relevant information refer to Table 2 in Whitehead et al. (1996). In a *P. radiata* stand, g_{stmax} values are thought to rapidly decline as trees mature. The SWatBal model uses a g_{stmax} parameter of 0.1 mol m^{-2} s^{-1} with the assumption that the stand is approaching canopy closure and maturity. Overall, the sensitivity analysis shows W_f values at the dryland Burnham site as highly sensitive to changes in g_{stmax}. These findings highlight the impact of g_{stmax} values on model simulations, placing emphasis on choosing an appropriate g_{stmax} value.

Published *P. radiata* one-sided L_t values typically range between 2 m^2 m^{-2} (McMurtrie et al., 1990; Richardson et al., 2002) and 7 m^2 m^{-2} (Arneth et al., 1998b). Overestimated L_t values will lead to underestimated predictions of W_f whereas underestimated L_t values will result in over prediction of W_f. The sensitivity analysis suggests W_f is not unduly influenced by changes in L_t values. For example, a 25 % change in L_t shows a response of less than 9 % for W_f at the Burnham site and only 4 % at the Haupapa site. These responses indicate that substantial changes (± 50 %) are required in L_t values before W_f simulations become greatly influenced. This relatively low sensitivity of changes to L_t suggests the development of a nation-wide surface may provide little improvement to the estimation of W_f.

6.4 Validation and sensitivity analysis summary and conclusion

When the SWatBal model was tested using data from the Burnham and Haupapa studies, good agreement was found between modelled and measured W_f. This agreement demonstrated that the SWatBal model was correctly formulated and that key parameter values were robust. Using spatially interpolated long-term climate and soil physical data, SWatBal model values continue to closely correspond with site measured values, although the strength of the relationship is slightly reduced. This good correspondence provides confidence in the general applicability of the SWatBal model and suggests that the use of spatially interpolated long-term climate data provides a useful approximation of actual climate data.

The sensitivity analysis highlighted a number of other points. At both Burnham and Haupapa, changes to P, and g_{stmax}, had the greatest impact on model predictions. Overall, changes to rainfall had the greatest impact on W_f, especially at the Burnham dryland site. The high sensitivity of soil water storage to changes in rainfall highlights the importance of modelling with accurate and representative rainfall data. This finding justifies the effort taken in this study to develop *RAV* data, which simulates rainfall that what would normally occur on a monthly basis using long-term rainfall averages.

Change to the input parameter g_{stmax} was found to be highly sensitive to SWatBal modelled W_f, especially at the Burnham dryland site. Of all the parameters and variables tested, uncertainty surrounds the application of one standard generic g_{stmax} value. These

findings suggest that the SWatBal model would benefit from further research to determine whether a generic value could be used for g_{stmax} and, if not, what alternative approach could be taken.

The sensitivity analysis found that W_f was not unduly influenced by changes to L_t. This lack of influence suggests that the development of a nation-wide L_t surface may provide little improvement toward SWatBal model simulations.

7. SWATBAL ADVANCEMENTS AND LIMITATIONS

7.1 The advancement SWatBal provides

SWatBal model was developed from a well recognised and published point-location model initially developed and refined over many years for estimating the soil water balance specific to the *Pinus radiata* species under New Zealand conditions. Furthermore, the SWatBal model utilises readily available GIS software enabling model execution by general practitioners, removing the need for specialist skills of programmers. The major advance SWatBal brings is a fully operational spatial model capable of determining soil moisture surplus or deficit across large regions. Indeed, for this paper water balance was simulated monthly for the whole of New Zealand. A further advance is that SWatBal uses a raster masking system, selecting regions of appropriate size, compatible with the operator's computer processor. This feature allows easy model execution for smaller forest stands, subcatchments or catchments by creating a raster

mask of the area of interest and passing this through the programme. Currently, input data remain at a relatively coarse scale, but as new and improved datasets become available the model can easily be rerun improving model outcomes. If a specific forest stand of interest should require modelling in closer detail, then the collection of detailed data for that area and the creation of raster surfaces with an appropriate level of detail for rerunning in the SWatBal model can also be undertaken. The SWatBal model is easily run using its transparent and repeatable methods.

The *RAV* daily rainfall surfaces represent realistic rainfall distribution across each month based on ~40 years of historic data providing an advance not realised in previous New Zealand models. This advancement of the data improves the SWatBal model simulations for drainage and rainfall, especially for sites with relatively homogenous rainfall across each year. In the future the approach used to create *RAV* surfaces could be applied to other associated input properties leading to further improvements.

7.2 SWatBal limitations

SWatBal was designed to provide insight into forest soil moisture regimes occurring across New Zealand on a monthly basis and to explore the relationships between soil moisture and *P. radiata* productivity. From this perspective our modelling approach falls into the coarse scaled or less detailed (macro) category of scientific research. Therefore, future use and applications of these data must be at appropriate scales. In a modelling

environment the output data are directly related to the scale and accuracy of input data. The underlying SWatBal data remain the best available and therefore provide the best national extent outputs at this point in time. However, improvement to input surfaces, especially those related to the soil storage component, could advance future model outcomes. *RAV* daily rainfall data provide a method for improving input data not used before in New Zealand, which could potentially be used to improve other climate-related properties. Leaf area index (L_t) was a property of concern because it was held constant across New Zealand, which under normal conditions would be influenced by soil nutrient status, previous land use and other associated environmental properties. In a point-location model this parameter can be estimated or measured in the field. Data at the national extent were not available to estimate L_t. Research into creating a national L_t surface, although problematic, could greatly enhance future modelling applications. The limitations highlighted here make the point that although SWatBal adequately achieves the outcomes it was designed for, the underlying model structure is robust enough to enable upgrades into the future, an essential requirement in this fast moving era of spatial data development.

8. CONCLUSIONS

The overriding purpose for developing the SWatBal monthly surfaces was to explore the spatial relationships between soil moisture and *P. radiata* productivity across New Zealand. When creating datasets or modelling at the national extent there remain

numerous issues related to data availability, resolution and quality. While issues of data availability and measures of their error continue to be problematic in New Zealand, the SWatBal model must work within current dataset capabilities until data upgrades become available for existing nationwide surfaces. Consideration of these issues was central to the SWatBal advancement with extensive emphasis given to literature reviews, expanding the knowledge base, and learning from current *P. radiata* soil water balance models. This paper discusses each input parameter in turn reviewing their capabilities, describing the advancement SWatBal brings and highlights possible future developments.

The SWatBal model was tested using data from two contrasting sites, Burnham and Haupapa. The validation found good agreement between modelled and measured $W_{f,}$ demonstrating that SWatBal was correctly formulated and that key parameter values were robust. Although the sensitivity analysis and validation was only undertaken at these two sites, this agreement provides confidence in the general applicability of the SWatBal model. It also shows that spatially interpolated long-term climate data provides a useful approximation of actual climate data.

From the conceptual stages of this project it was recognised that some available climatic properties were not ideal for soil water balance modelling. However, from the onset SWatBal was designed not as "the last word", but as a dynamic framework capable of model updates as new and improved data become available. Central to SWatBal is its flexibility, accommodating soil water balance simulations at a variety of user defined resolutions, geographic extents, and temporal or spatial scales. Estimates of daily soil

moisture were developed at a 100 m resolution across New Zealand producing a series of 12 monthly surfaces and one annual surface for mean fraction of available root-zone water storage (W_f), mean available root-zone water storage (W_a), and mean drainage.

Over and above SWatBal's design flexibility, its major advance for soil water balance modelling is the development and implementation of 'reasoned and allocated virtual data' (*RAV*). To date, spatially driven soil moisture balance models developed for New Zealand assume constant rainfall distribution across each day in a month. The more advanced models determine the number of days it rains on average for each month, distributing the rain days evenly across each month, then allocating constant rainfall to the days it rains. Although these approaches are an improvement beyond distributing monthly rainfall constantly to each day in a month, the next advancement in simulating 'naturally occurring' rainfall events was the development of *RAV* that contains stochastic or random elements.

Although the rationale driving model development was focused and specific, this does not diminish SWatBal's possibilities in other applications including hydrological processes, ground water pollution, and biological studies. Examples of such applications include the implementation of these data into forestry carbon sequestration models, investigation into the impacts of climate change on forest productivity, or the construction of suitability maps for water-deficit-tolerant forest species across the national extent. In conclusion, the SWatBal surfaces not only allow direct comparisons of the soil water status across New Zealand but also enable visualisation of spatial patterns

of weather across the national extent, catchment, sub-catchment, forest stand or compartment level of detail.

9. ACKNOWLEDGEMENTS

We gratefully acknowledge and thank David Whitehead for his helpful and valuable comments in the development of SWatBal. We are grateful to Trevor Webb for his technical assistance and advice with Landcare Research's datasets. David Palmer wishes to thank and acknowledge the Foundation for Research Science and Technology (FRST) and the University of Waikato (Doctoral Scholarship) for their funding. Barbara Höck thanks FRST for funding (Protecting and enhancing the environment through forestry CO4X0304). We also acknowledge the helpful comments provided by two anonymous referees.

10. REFERENCES

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998a. Assessment of annual carbon exchange in a water-stressed *Pinus radiata* plantation: an analysis based on eddy covariance measurements and an integrated biophysical model. Global Change Biology 5, 531-545.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998b. Fluxes of carbon and water in a *Pinus radiata* forest subject to soil water deficit. Australian Journal of Plant Physiology 25, 557-570.

Barringer, J., Lilburne, L. 1999. Scale issues in developing regional-scale soil water balance surfaces. In: Whigham, P.A., (Eds.), Proceedings of the Eleventh Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand, pp. 231-240.

Barringer, J.R.F., McNeill, S.J., Pairman, D. 2002. Progress on assessing the accuracy of a high-resolution digital elevation model for New Zealand. In: Hunter, G., Lowell, K. (Eds), Proceedings of the 5th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 10-12 July, Melbourne, Australia, pp. 187-195.

Barringer, J.R.F., Porteous, A., Salinger, M.J., Trangmar, B.B. 1995. Estimating spatial patterns of wilting point deficit using a water balance model and a geographic information system. Journal of Hydrology 34, 42-59.

Bian, L. 1997. Multiscale nature of spatial data in scaling up environmental models. In: Quattrochi, D.A., Goodchild, M.F., (Eds), Scale in remote sensing and GIS, Lewis Publishers, Boca Raton, pp. 13-26.

Boomsma, D.B., Hunter, I.R., 1990. Effects of water, nutrients and their interaction on tree growth, and plantation forest management practices in Australasia: a review. Forest Ecology and Management 30, 455-476.

Burrough, P.A. 1986. Principles of Geographic Information Systems for Land Resource Assessment. Oxford University Press, New York.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographic Information Systems. Oxford University Press, New York.

Clayden, B., Webb, T.H. 1994. Criteria for defining the soilform – the fourth category of the New Zealand Soil Classification. Landcare Research Science Series No. 3, Lincoln, 36 pp.

Clifton, N.C., 1994. New Zealand Timbers. GP Publications, Wellington, 170 pp.

Cressie, N. 1991. Statistics for Spatial Data. New York, Wiley.

Fahey, B., Watson, A., Payne, J. 2001. Water loss from plantations of Douglas-fir and radiata pine on the Canterbury Plains, South Island, New Zealand. Journal of Hydrology 40, 77-96

Goodchild, M.F. 2001. Models of scale and scales of modelling. In: Tate, N.J., Atkinson, P.M, (Eds.), Modelling scale in geographical information science. Wiley and Sons, England, pp. 3-10.

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, USA, 483 pp.

Goovaerts, P. 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma, 89: 1-45.

Goudriaan, J., van Laar, H.H. 1994. "Modelling Crop Growth". Kluwer, Amsterdam, Netherlands.

Gradwell, M.W., Birrell, K.S. 1979. Soil Bureau laboratory methods. Part C: Methods for physical analysis of soils. New Zealand Soil Bureau Scientific Report 10C.

Griffiths, E. 1985. Interpretation of soil morphology for assessing moisture movement and storage. New Zealand Soil Bureau Scientific Report 74, 20 pp.

Haslett, A.N. 1986. Properties and utilisation of exotic specialty timbers grown in New Zealand. Part III: Cypresses. FRI Bulletin No. 119, Forest Research Institute, Rotorua, New Zealand, 12 pp.

Heuvelink, G.B.M. 1998. Error Propagation in Environmental Modelling with GIS. Taylor and Francis, London, UK, 127 pp.

Hewitt, A.E. 1998. New Zealand Soil Classification (Second edition). Landcare Research Science Series No. 1. 133 pp.

Hutchinson, M.F., Gessler, P.E. 1994. Splines – more than just a smooth interpolator. Geoderma 62, 45-67.

Journel, A.G., Huijbregts, Ch.J. 1978. Mining Geostatistics, Academic Press, London, 600 pp.

Kelliher, F.M., Whitehead, D., McAneney, K.J., Judd, M.J. 1990. Partitioning evapotranspiration into tree and understorey components in two young *Pinus radiata* D. Don stands. Agricultural and Forest Meteorology 50, 211-227.

Kelliher F.M., Whitehead, D., Pollock, D.S. 1992. Rainfall interception by trees and slash in a young *Pinus radiata* D. Don stand. Journal of Hydrology, 131, 187-204.

Kriticos, D. J., Sutherst, R. W., Brown, J. R., Adkins, S. W., Maywald, G. F. 2003. Climate change and the potential distribution of an invasive alien plant: *Acacia nilotica* ssp. *indica* in Australia. Journal of Applied Ecology, 40,111-124

Leathwick, J.R 1995. Climatic relationships of some New Zealand forest tree species. Journal of Vegetation Science 6, 237-248.

Leathwick, J.R. 1998. Are New Zealand's *Northofagus* species in equilibrium with their environment? Journal of Vegetation Science 9, 719-732.

Leathwick, J.R. 2001. New Zealand's potential forest pattern as predicted from current species-environment relationships. New Zealand Journal of Botany 39, 447-464.

Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K. 2002a. Land Environments of New Zealand: A technical guide. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 237 pp.

Leathwick, J.R., Whitehead, D. 2001. Soil and atmospheric water deficits and the distribution of New Zealand's indigenous tree species. Functional Ecology 15, 233-242.

Leathwick, J.R. Whitehead, D. McLeod, M. 1996. Predicting changes in the composition of New Zealand's indigenous forests in response to global warming: a modelling approach. Environmental Software 11, 81-90.

Leathwick, J., Wilson, G., Rutledge, D., Wardle, P., Morgan, F., Johnston, K., McLeod, M., Kirkpatrick, R. 2003. Land Environments of New Zealand. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 184 pp.

Leathwick, J.R., Wilson, G., Stephens, R.T.T. 2002b. Climate surfaces for New Zealand. Landcare Research contract report: LC9798/126. Landcare Research, Hamilton, New Zealand, 22 pp.

Lohammer, T., Larsson, S., Linder, S., Falk, S.O. 1980. FAST – Simulation models of gaseous exchange in Scots pine. Ecological Bulletin 32, 505-523.

McMurtrie, R.E., Rook, D.A., Kelliher, F.M. 1990. Modelling the yield of *Pinus radiata* on a site limited by water and nitrogen. Forest Ecology and Management, 30, 381-413.

Met. Service, 2006. About MetService: MetService history, viewed 2006. http://www.metservice.co.nz/default/index.php?alias=aboutmetservice.

Mitchell, N.D., 1991. The derivation of climate surfaces for New Zealand, and their application to the bioclimatic analysis of the distribution of kauri (*Agathis australis*). Journal of the Royal Society of New Zealand, 21, 13-24.

Newsome, P.F.J., Wilde, R.H., Willoughby, E.J. 2000. Land resource information system spatial data layers. Landcare Research, Palmerston North, New Zealand, 84 pp.

NIWA, 2006. Overview of New Zealand climate, viewed 2006. http://www.niwascience.co.nz/edu/resources/climate/overview/.

NZFOA, New Zealand Forest Owners Association, 2005. New Zealand Forest Industry Facts and Figures 2004/2005. New Zealand Forest Owners Association, P.O. Box 1208, Wellington.

Palmer, D.J., Höck, B.K., Dunningham, A.G., Lowe, D.J., Payn, T.W. 2008. Developing national-scale terrain attributes for New Zealand (TANZ). Scion Bulletin.

Pawlowsky-Glahn, V., Olea, R.A. 2004. Geostatistical analysis of compositional data. Oxford University Press, New York, USA, 181 pp.

Payn, T.W., Skinner, M.F., Hill, R.B., Thorn, A.J., Scott, J., Downs, S., Chapman, H. 2000. Scaling up or down: the use of foliage and soil information for optimising the phosphate nutrition of radiata pine. Forest Ecology and Management, 138, 79-89.

Richardson, B., Whitehead, D., McCracken, I.J. 2002. Root-zone water storage and growth of *Pinus radiata* in the presence of a broom understorey. New Zealand Journal of Forestry Science, 32, 208-220.

Singer, M.J., Munns, D.N. 1991. Soils an introduction (Second edition). Macmillan Publishing Company, New York, USA, 473 pp.

Soil Survey Staff, 1999. Soil Taxonomy, 2nd ed. USDA Natural Resources Conservation Service Agriculture Handbook 436. 869 pp.

Sutherst, R.W., Maywald, G.F., Yonow, T., Stevens, P.M. 1999. CLIMEX; predicting the effects of climate on plants and animals. User guide. CSIRO Publishing, Melbourne, Australia.

Tait, A., Henderson, R., Turner, R., Zheng, X. 2006. Spatial interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26, 2097-2115.

Taylor, N.H., Pohlen, I.J. 1970, Soil survey method: A New Zealand handbook for the field study of soils. New Zealand Soil Bureau Bulletin 25, 242 pp.

Thornthwaite, C.W., Mather, J.R. 1954. Climate in relation to crops. Meteorological Monographs, 2, 1-10.

Watt, M.S., Clinton, P.W., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C. 2003a. Above-ground biomass accumulation and nitrogen fixation of broom (*Cytisus scoparius* L.) growing with juvenile *Pinus radiata* on a dryland site. Forest Ecology and Management, 184, 93-104.

Watt, M.S., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C. 2003b. Modelling the influence of weed competition on the growth of young *Pinus radiata* at a dryland site. Forest Ecology and Management, 178, 271-286.

Webster, R., Oliver, M.A. 1990. Statistical methods in soil and land resource survey. Oxford University Press, Oxford. 316 pp.

Webb, T.H., Wilson, A.D. 1995. A manual of land characteristics for evaluation of rural land. Landcare Research Science Series 10. Lincoln, New Zealand, Manaaki Whenua Press, 32 pp.

Whitehead, D., Kelliher, F.M. 1991. Modeling the water balance of a small *Pinus radiata* catchment. Tree Physiology, 9, 17-33.

Whitehead, D., Leathwick, J.R., Walcroft, A.S. 2001. Modeling annual carbon uptake for the indigenous forests of New Zealand. Forest Science, 47, 9-19.

Whitehead, D., Livingston, N.J., Kelliher, F.M., Hogan, K.P., Pepin, S., McSeveny, T.M., Byers, J.N. 1996. Response of transpiration and photosynthesis to a transient change in illuminated foliage area for a *Pinus radiata* D. Don tree. Plant Cell and Environment, 19, 949-957.

Wilde, R.H., Willoughby, E.J., Hewitt, A.E. 2000. Data manual for the national soils database spatial extension. Landcare Research, Palmerston North, New Zealand.

Wilson, A.D., Giltrap, D.J. 1982. Prediction and mapping of soil water retention properties. New Zealand Soil Bureau District Office Report, WN7. 15 pp.

Wilson, J.P., Gallant, J.C. 1998. Terrain-based approaches to environmental resource evaluation. In: Lane, S.N., Richards, K.S., Chandler, J.H. (eds.), Landform monitoring, modelling, and analysis. Wiley & Sons, New York, pp. 219-240.

Zang, J., Goodchild, M.F. 2002. Uncertainty in Geographical Information. Taylor and Francis, London, 266 pp.

Landscape view across Lake Tekapo towards the Two Thumb Range with the Church of the Good Shepherd to the right of the frame (photograph by Adam Vonk)

Not chaos-like together crushed and bruised,
But, as the world, harmoniously confused:
Where order in variety we see,
And where, though all things differ, all agree.

A. Pope, *Windsor Forest*

CHAPTER FOUR

COMPARING THE PREDICTION OF *PINUS RADIATA* PRODUCTIVITY USING A SOIL WATER BALANCE MODEL (SWATBAL) DERIVED FROM SITE-MEASURED CLIMATE DATA AND LONG-TERM AVERAGED CLIMATE SPATIAL DATA

D.J. Palmer[a*], M.S. Watt[b], B.K. Höck[c], D.J. Lowe[a], T.W. Payn[c]

[a]*Department of Earth and Ocean Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand 3240*

[b] *Scion, PO Box 29237, Christchurch, New Zealand*

[c] *Scion, Private Bag 3020, Rotorua, New Zealand*

[*]Corresponding author. Tel.: +64-7-8562889 *E-mail address*: djp8@waikato.ac.nz

To be submitted to: Canadian Journal of Forest Research

ABSTRACT

Available root-zone water storage (W_a), over wide environmental and edaphic gradients throughout New Zealand, was determined at 31 sites using two approaches: (1) a soil water balance model using actual climate data (ACD), and (2) SWatBal, a spatial soil water balance approach that uses readily available spatially interpolated long-term (SILT) climate surfaces. Using these estimates of W_a the objective of this study was to compare the predictive power of W_a determined using ACD, and SILT climatic data for modelling productivity of *Pinus radiata* .

Mean annual volume increment (MAI_v) for *P. radiata* D. Don was determined using small, highly stocked plots designed to compress the crop rotation length. MAI_v observations covered a substantial range from 7 to 49 m^3 ha^{-1} yr^{-1}. Mean annual air temperature, T_a, was the environmental variable most strongly related to MAI_v, exhibiting a significant and positive linear relationship which accounted for 42% of the variance in MAI_v. When environmental variables were included in combination the best predictive model accounted for 61% of the variance in MAI_v using actual T_a and W_a determined from either ACD or SILT climate surfaces. This combination of variables accounted for approximately 15% more of the variance in MAI_v than T_a and P. Although we recognise the need for the continued collection of some low cost site data, the utility of the SWatBal model and the use of SILT climate data simplify soil water balance estimations considerably.

Keywords: soil water balance, SWatBal, forest productivity, radiata pine, New Zealand.

1. INTRODUCTION

Pinus radiata D. Don is the most widely planted forestry crop in New Zealand (NZFOA, 2005) and the Southern Hemisphere (Lewis and Ferguson 1993). Given its importance, the identification of key environmental determinants of productivity for this species is likely to be of considerable use in determining site suitability and possible future impacts under climate change. Reductions in tree productivity are often attributed to limiting environmental factors including temperature, water, light and nutrients (Nambiar and Sands, 1993). To date, however, our knowledge and understanding of the properties influencing *P. radiata* productivity have been limited and the exact form of these relationships remains unclear.

Estimation and modelling *P. radiata* site productivity across New Zealand's national extent began in the early 1970s and although the focus has changed from productivity to sustainability, this pool of knowledge continues to slowly accrue. Jackson and Gifford (1974) published New Zealand's first comprehensive national study investigating the effects of climate and other environmental variables on *P. radiata* productivity across 132 nationwide sites. They found that 66% of the periodic volume increment could be attributed to mean annual and seasonal distribution of rainfall, the seasonal departures

from optimum temperature, total nitrogen and available phosphorus, and effective soil depth.

A decade later, Hunter and Gibson (1984) investigated *P. radiata* site index at 299 North Island sites. A model was constructed that predicted site index from climate data including annual rainfall, and departures from the optimal average annual temperature, in conjunction with nutrients, topsoil depth, soil penetrability, and soil pH. Woollons et al. (2002) revisited the Hunter and Gibson (1984) data using climate variables extracted from the BIOCLIM database (Wahba and Wendelberger, 1980; Hutchinson, 1984; Nix, 1986) and pooled soil class information (New Zealand Soil Classification, soil order level; Hewitt, 1998). Main predictors of site index were A-horizon depth, average wind velocity, mean annual rainfall, and altitude.

More recently a national plot series, in which trees were grown at high stand density over a four year period, was used to identify the main soil properties influencing productivity and sustainability of *P. radiata* and *Cupressus lusitanica* (Watt et al., 2008). While this study focussed on key soil properties influencing sustainability, multiple regression showed that temperature and mean annual root-zone water storage were the key climatic determinants of volume for *P. radiata*. This research clearly identified soil water balance as a key determinant of *P. radiata* productivity, showing this variable to have a stronger relationship with productivity than the more frequently used variable of mean annual rainfall.

Although spatial estimates of water balance have been calculated for New Zealand (Barringer at al. 1995; Barringer and Lilburne, 1999), these models are not applicable to plantation forestry systems. The collection of environmental and climate data and the estimation of soil water balance across broader scales, at more than one site, remains logistically difficult and often cost prohibitive. Logistics aside, the time and effort required to set up site-specific meteorological stations, the parameterisation and measurement of climatic variables, and the accurate calculation of the soil water storage component is not without high cost. Although not ideal, it is not surprising that some studies opt for data collected from the nearest climate station (McMurtrie et al., 1990, Richardson et al., 2002), often many kilometres from the study location because of instrument failure or incomplete dataset collection.

A contrast to site-specific soil water balance with its associated logistical and cost concerns is the spatial approach utilising long-term climate data. Spatial modelling using GIS platforms has come into its own because of the affordability and availability of high-speed computer processor technology and large digital storage capacity (Palmer et al., 2008). These advantages mitigate many of the historic issues associated with measurement and collection costs. Soil water balances potentially can be modelled spatially across large regions using SWatBal (Chapter 3), a model that utilises data extracted from spatially interpolated long-term (SILT) climate surfaces. The main advantage of SWatBal is the reduced modelling costs and the relatively quick model development and processing period. There is the added advantage that modelling is undertaken with normalised climate data (Leathwick et al., 2002; Chapter 3). However,

there remains a trade-off between SWatBal cost effectiveness and model simulation efficacy. Although SWatBal's modelled predictions are largely dependant on input surface accuracy and precision, these datasets are continually improving.

To identify key environmental determinants at the national extent requires datasets that cover the wide environmental gradient across New Zealand. This type of diversity is covered by a national network of 31 plots grown over a period of four years at high stand density to identify key determinants of productivity (Watt et al., 2005; 2008). By establishing these plots at high density, the experimental period over which the stand develops is compressed by increasing intra-tree competition and the rate at which leaf area develops and the green crown recedes. In these plots, canopy closure is typically attained at most sites within two years of planting. This technique has significant advantages over conventional experimental methods because time frames over which measurements can be taken are shortened and relatively large numbers of plots can be installed, thus permitting sampling over a wide range of environments. Using site specific measurements of climate Watt et al. (2008) found W_a and mean annual temperature to be the key determinants of productivity across this dataset.

Using this national network of plots the objective of this study was to determine the key climatic drivers of *Pinus radiata* productivity within this dataset using SILT climate surfaces compared with those determined using actual climatic data. A specific aim was to compare the utility of W_a derived from SILT surfaces with than obtained from actual climatic data in modelling productivity across New Zealand.

2.0 METHODS

2.1 Site locations

Site quality plots (Figure 1) were selected to represent the range in soil properties on which plantation forests are currently grown in New Zealand. Soils were grouped in soil orders of the New Zealand Soil Classification system (NZSC) (Hewitt, 1998), which recognises 15 soil orders. Plantation forests within New Zealand (Thompson et al., 2003) occur on nine orders, namely Allophanic, Brown, Oxidic, Pallic, Podzol, Pumice, Raw, Recent and Ultic orders (Hewitt, 1995; Rijkse and Hewitt, 1995; Kirkpatrick, 1999). The research trial was established across all these orders with the exception of the Oxidic Soils, which represent only a small proportion (~0.7%) of the area on which the forest estate is planted (R. Simcock, pers. comm. 2008). The number of sites established on each soil order was weighted to be representative of the corresponding plantation area on which the order occurs.

Sites were further screened using climate surfaces (Figure 1) to ensure that selected areas represented the extensive range in meteorological conditions found throughout New Zealand forest plantations. Long-term climate data obtained from thin-plate spline surfaces (Hutchinson and Gessler, 1994) fitted to meteorological station data (Leathwick and Stephens, 1998) were used to determine climatic conditions at each selected site. When compared with long-term average values for all plantation forests (Anon., 1983), the 31 selected sites almost encompass the total range in annual rainfall, mean annual air temperature, relative humidity and solar radiation (see Watt et al., 2005).

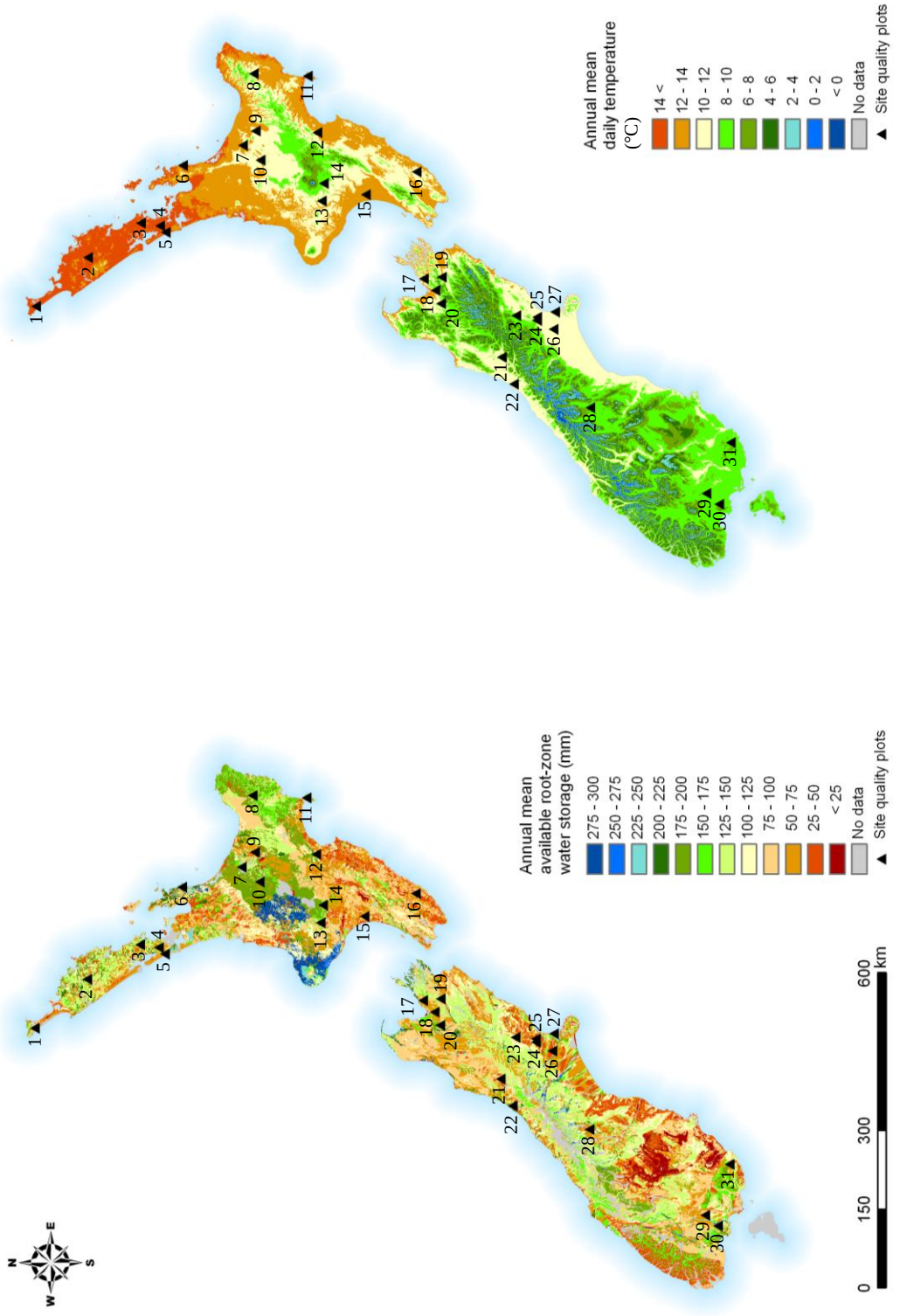


Figure 1. Maps illustrating the geographic variation in annual mean available root-zone water storage (W_a) (left) and annual mean daily temperature (T_a) (right) for New Zealand. Site quality plot locations are shown by triangular symbols and their associated numbers. The W_a surface was derived from the SWatBal model (Chapter 3) and the T_a surface was calculated by averaging Landcare Research's 12 monthly mean daily temperature surfaces (Leathwick et al., 2002).

2.2 Site quality plot experimental design

Site quality plots were located at each of the 31 selected locations with a series of eight plots installed using a factorial design with the following three factors: species (*P. radiata* and *C. lusitanica* Mill.), nutrient status (no fertiliser and nutrients supplied in excess of crop requirements), and disturbance (low and high disturbance). Plots were located in close proximity to one another and were of similar slope and aspect. In this study, measurements and modelling were restricted to the fertilised *P. radiata* subplots to reduce the effect of fertility limitations on growth so that environmental determinants of growth could be more readily identified.

Each plot was small in size (3 m × 3 m), containing nine trees spaced at 0.5 m × 0.5 m (equivalent to 40,000 stems ha^{-1}) surrounded by a two-row buffer to reduce edge effects. Regular applications of herbicide ensured weed-free conditions. All sites were planted with seed stock of the same origin (growth and form factor of 19 for *P. radiata* (Vincent and Dunstan, 1989)) sourced from the Scion nursery in Rotorua. The site quality plots were installed over a three-year period from 2000 to 2002.

In fertilised plots, broadcast application of fertiliser was undertaken three times a year during the first year (winter, spring and summer) and thereafter during spring at annual intervals. Over the four years of the study, the elemental quantity of nitrogen, phosphorus, potassium, sulphur, magnesium and calcium applied to each fertilised plot was 690, 200, 558, 160, 40 and 160 kg ha^{-1}, respectively. At sites identified by forest

history or soil analysis (at planting time) as being deficient in boron, magnesium and copper, additional applications of these nutrients to fertilised plots were made during summer in sufficient quantities to ensure these nutrients were not limiting.

2.3 Measurement collection and extraction of climate data

Ground-line diameter (D) and height (h) of the nine measurement trees were measured annually in all plots. Stem volume (V) was determined from tree height and ground-line diameter using the following equation previously found to be applicable to stands covering a range of ages and stockings (P.W. Beets, unpublished data):

$$V = (\pi(D/2)^2\, h)\, 0.25 \qquad (1)$$

The mean annual increment in volume (MAI_v) was determined as $(V_2 - V_1)/(T_2 - T_1)$, where V_2 and T_2 represent volume and time at harvest, while V_1 and T_1 represent volume and time immediately after planting.

At each site, photosynthetically active radiation (I), air temperature and relative humidity were measured on sensors installed on a 3-m high tower located adjacent to the experimental plots. A tipping bucket rain gauge, positioned on top of the tower, was used to measure above-canopy rainfall. Vapour pressure deficit (D) was determined from relative humidity and temperature measurements using standard formulae.

The SILT climate data required for the determination of soil water balance using the SWatBal model is a combination of long-term monthly averaged climate surfaces (Leathwick et al., 2002) and data derived from them, and data derived from 40 years of daily rainfall interpolated into national virtual rainfall surfaces (Tait et al., 2006). The monthly climate data were extracted from 100 m climate raster surfaces for each of the 31 plot locations. Specifically, solar radiation and temperature values were extracted from long-term monthly climate surfaces. Rainfall values were extracted from reasoned and allocated virtual data (RAV) surfaces (Chapter 3). Briefly, the RAV surfaces comprise 365 surfaces with rainfall for each day of the year. Rainfall for a day was determined by selecting a monthly rainfall distribution from the virtual rainfall surfaces with extreme events removed so that total rainfall for the month was the closest to that of the long-term monthly averages. If necessary, daily rainfall was proportionally adjusted so that total monthly rainfall exactly matched the long term average. Standard formulae were used to determine *D* from relative humidity and temperature values extracted from the monthly climate surfaces.

2.4 Determination of root-zone water storage

A daily water balance model was used to calculate root-zone water storage (*W*) on the *i*th day as,

$$W_i = W_{i-1} + P_i - E_{ti} - E_{twi} - E_{gi} - F_i \quad (1)$$

where P_i is rainfall, E_{ti} transpiration from the tree canopy, E_{twi} evaporation of intercepted rainfall from the tree canopy, E_{gi} evaporation from the soil surface, and F_i drainage from the root-zone (Whitehead et al., 2001; Watt et al., 2003). Root-zone water storage ranges from a maximum (W_{max}) to a minimum (W_{min}) value, and maximum available root-zone water storage, W_m, is defined as W_{max} - W_{min}. As W_{min} was set to zero at all sites, modelled values of W_i are analogous to available root-zone water storage values, which will be hereafter referred to as W_a.

Transpiration from the tree canopy was calculated using a simple diffusion equation,

$$E_t = D\, g_{st}\, L_t \qquad (2)$$

where D is vapour pressure deficit, g_{st} the average stomatal conductance for the canopy, and L_t is the half-surface leaf area index. The decline in stomatal conductance, which occurs in response to increasing D, was modelled using the function described by Lohammer et al. (1980) as,

$$g_{st} = \alpha\, g_{stmax}\, (1+ (D - D_{smin})/ D_{0t})^{-1} \qquad (3)$$

where g_{stmax} is the maximum stomatal conductance, and D_{0t} is the sensitivity of g_{st} to D when $D > D_{smin}$ (the value of D below which g_{st} is constant). The coefficient α will be defined later. Evaporation from the wet tree canopy, E_{twi}, ($=f_{wt}\, P$), was calculated as a constant fraction, f_{wt}, of rainfall.

Evaporation from the soil, E_g, was calculated from the available energy beneath the tree canopies, G_g, as,

$$E_g = \alpha \tau s G_g / [\lambda (s + \gamma)] \quad (4)$$

where the term $s\ G_g/[\lambda(s + \gamma)]$ is the equilibrium rate of evaporation and the coefficient τ describes the degree of coupling of the soil surface with the air above the canopy (Kelliher et al., 1990), s is the slope of the relationship found between saturated vapour pressure and the temperature at any given air temperature, λ the latent heat of vaporisation, and γ the psychrometric constant. Using Beers Law, G_g was determined from ($e^{-kLt}\ G_a$) where G_a represents the available energy above the canopy (assumed to be 70% of shortwave irradiance) and k is the light extinction coefficient (assumed to be 0.5 for a spherical leaf angle distribution).

The coefficient α was used to reduce evaporation from the soil and transpiration as soil water storage declined. The value of α was set to 1 at maximum values of W (W_m), and was not reduced until W declined to a threshold value (W_t). As W_i progressively declined below this threshold, α was linearly reduced from 1 to 0 at minimum values of W_a (=0).

Estimated evaporation from the soil and transpiration were reduced to 75% of their potential rates on days when rain occurred. Drainage from the root zone was assumed to be zero when $W_i \leq W_m$ and equal to rainfall reaching the soil when $W_i > W_m$.

In addition to meteorological data, the model requires parameter values for g_{stmax}, D_{0t}, f_{wt}, and W_t, the site specific variables W_m, and seasonal estimates of leaf area index (L_t). Parameter values obtained from the literature include 0.1 mol^{-2} s^{-1} and 0.6 kPa, respectively, for g_{stmax}, and D_{0t} (Whitehead et al., 1996; Richardson et al., 2002), 0.2 for f_{wt} (Kelliher et al., 1992), and W_t equal to 0.6 and 0.55 of W_m for E_g and g_{st}, respectively (Watt et al., 2003). Site-specific estimates of maximum available root-zone water storage (W_m) were determined at harvest following the methods fully described in Watt et al. (2008). Leaf area index (L_t) was directly measured every four months at each site using a canopy analyser (LAI-2000, Li-Cor Inc., Lincoln, NE, USA). Measurements were taken at four locations in each plot and averaged to obtain a plot level estimate. Daily estimates of L_t were obtained by fitting a spline function to these plot level estimates.

Two soil water balance models, a standard site specific model and a spatial GIS model (SWatBal), were used to calculate W_a at daily time steps for each of the 31 site quality plots. The standard site specific model used actual climatic data (ACD) collected over the study period. The SWatBal model uses the same input data as the standard site specific model, but substituted ACD for data extracted from SILT climate surfaces used in the standard SWatBal application (Section 2.3). SWatBal uses the GIS platform ArcInfo™ and its associated Arc Macro Language (AML) within the 'GRID' module to calculate W_a across each month, aligned with long-term normalised monthly surfaces (see Chapter 3 for modelling details and application). The SWatBal-derived daily surfaces include W_i and drainage, averaged to monthly mean values.

2.5 Data analyses

All analyses were undertaken in SAS (SAS Institute, 2000). Using appropriate functional forms, the strength of bivariate relationships were examined between (i) the same climatic variables obtained from ACD and SILT climatic surfaces and (ii) mean annual volume increment (MAI_v) and climatic variables obtained from both data sources. Multiple regression models of MAI_v were constructed using a combination of climatic variables for data obtained from both ACD and SILT climatic surfaces. Using the general linear modelling procedure (PROC GLM) variables were introduced sequentially into the model starting with the variable which exhibited the strongest correlation, until further additions were not significant at P=0.05. Although the forward stepping procedure was used to assist in variable selection, final selection was undertaken manually one variable at a time to enable the identification of any non-linear relationships. For the final models, redundancy between variables was assessed using the variance inflation factor, with values less than 10 indicating that multicollinearity is within acceptable bounds (Der and Everitt, 2001). The relationship between measured and predicted MAI_v was plotted for both models to assess model bias.

3.0 RESULTS

3.1 Site level variation

Site level variation in the meteorological data shows a substantial range across the 31 sites (Table 1). Variation was most pronounced for W_a with ACD values ranging from 10 to 120 mm, and from 10 to 86 mm for W_a determined from SILT climatic surfaces. Averaged across the two datasets average annual rainfall (P) covered a six-fold range, while mean (T_a) and minimum (T_{min}) average annual air temperature ranged two-fold and four-fold, respectively. Site variation in I and D was somewhat lower, ranging two-fold for both variables in the ACD data, and 28 to 49%, for I and D, respectively, in the SILT dataset.

Table 1. Site level variation in average annual climate variables measured from the 31 sites over the study (ACD) and from spatially interpolated long-term (SILT) data. For each variable, the mean and range, in brackets, are shown for the 31 sites. Also shown is the coefficient of determination and significance level between the ACD and SILT data. Asterisks represent significance at $P<0.001$ (***) and $P<0.01$ (**).

Symbol	Variable	Measurement period		R^2
		ACD	SILT	
W_a	Available root-zone water storage (mm)	49 (10-120)	42 (10-86)	0.91***
P	Rainfall (mm year^{-1})	1349 (458-3035)	1408 (635-2990)	0.89***
T_{min}	Minimum air temperature (^oC)	7.6 (3.6-12.3)	6.9 (3.3-12.0)	0.86***
T_a	Mean air temperature (^oC)	13.4 (8.5-17.0)	11.6 (8.4-15.7)	0.76***
D	Vapour pressure deficit (kPa)	0.51 (0.31-0.72)	0.54 (0.43-0.64)	0.37**
I	Photosynthetically active radiation (MJ PAR m^{-2} d^{-1})	6.0 (4.2-7.5)	7.2 (6.0-7.7)	0.24**

Site level variation in tree dimensions and site characteristics affecting water balance were considerable (Table 2). Both ground-line diameter, and height at age four, ranged two-fold across all sites, and MAI_v ranged seven-fold from 7 to 49 m^3 ha^{-1} yr^{-1}. Average leaf area index ranged from 0.9 to 4.4 m^2 m^{-2}, while rooting depth ranged from 0.21 to 0.96 m. Maximum available root-zone water storage (W_{max}) covered a substantial range from 30 mm to 200 mm.

Table 2. Site level variation in stand characteristics and water balance input variables. Variable symbols are defined in the text and Table 1.

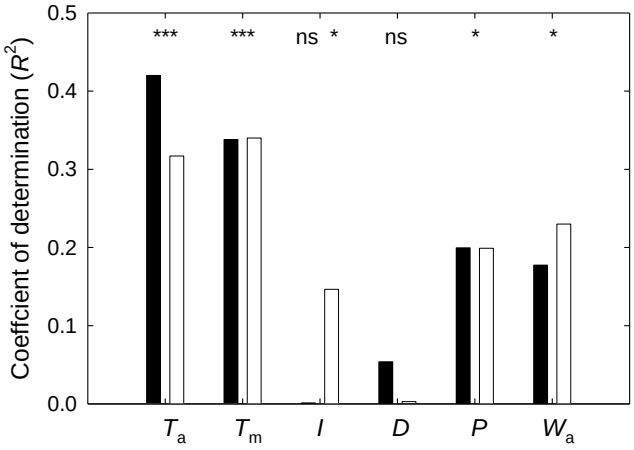
Variable	Mean and range
Stand characteristics	
Ground line diameter, age 4 (mm)	50 (34-75)
Height, age 4 (m)	4.0 (2.2-5.3)
Total stem volume, age 4 (m^3 ha^{-1})	87 (28-199)
MAI_v (m^3 ha^{-1} yr^{-1})	23 (7-49)
Water balance input variables	
Average leaf area index (m^2 m^2)	3.1 (0.9-4.4)
Root depth (m)	0.53 (0.21-0.96)
W_m (mm)	120 (30-200)

3.2 Bivariate correlations

The variables W_a, P, T_{min}, and T_a exhibited strong significant relationships ($R^2 > 0.76$) between ACD and SILT datasets (Table 1). Although correlations between datasets for I and D were significant, they were relatively weak (Table 1). Mean annual volume

increment exhibited significant (P<0.001) positive linear relationships with T_a and T_{min} from both ACD and SILT datasets, with all relationships having coefficients of determination greater than 0.31 (Figure 2). Annual average root-zone water storage was significantly related to MAI_v in both datasets with coefficients of determination ranging from 0.17 (ACD) to 0.23 (SILT climatic surfaces). Average annual rainfall exhibited significant bivariate correlations, using data from both sources, similar in strength to that of W_a. Average annual rainfall was modelled as a downward opening parabola with a maximum MAI_v occurring at ~2000 mm. With the exception of SILT derived *I*, no other variables exhibited significant relationships with MAI_v (Figure 2).

Figure 2. Coefficient of determination between volume mean annual increment (MAI_v) and the climatic variables from actual climate data (ACD) obtained during the study (filled bars) and data extracted from spatially interpolated long-term (SILT) climate surfaces (open bars). Asterisks at the top of the graph represent significance between mean annual volume increment (MAI_v) and the environmental variables at P<0.001 (***) and P<0.05 (*); ns= non-significant at P=0.05. A single significance level is given for each variable when significance categories do not differ between the two groups. T_a refers to mean air temperature, T_m minimum air temperature, *I* photosynthetically active radiation, *D* vapour pressure deficit, *P* rainfall, and W_a available root-zone water storage.



3.3 Multiple regression results

The best multiple regression models formulated for MAI_v from ACD included T_a and W_a, which collectively explained 61% of the variance (Table 3). The coefficient of determination remained unchanged when T_a from ACD was combined with W_a determined from SILT climatic surfaces. Using only SILT climatic surfaces, the best multiple regression model included both T_a and W_a, which collectively explained 46% of the variance in the data. Multiple regression models for both ACD and SILT climatic data included T_a and W_a as positive linear terms (Figure 3). The rate of increase in MAI_v with increases in T_a and W_a was very similar for both ACD and SILT climatic data (Figure 3).

Table 3. Summary of statistics for the final predictive models of volume mean annual increment (MAI_v) using actual climate data (ACD) obtained during the study and data extracted from spatially interpolated long-term (SILT) climate surfaces. Variable symbols are defined in Table 1.

Variables	Variable source	
	ACD	SILT
T_a (^oC)	0.42***	0.32**
W_a (mm)	0.19***	0.14*
Total R^2	0.61	0.46

The partial R^2 of each significant variable is shown with the significance level of the variable included in the model. Asterisks represent significant at $P<0.001$ (***), $P<0.01$ (**) and $P<0.05$ (*).

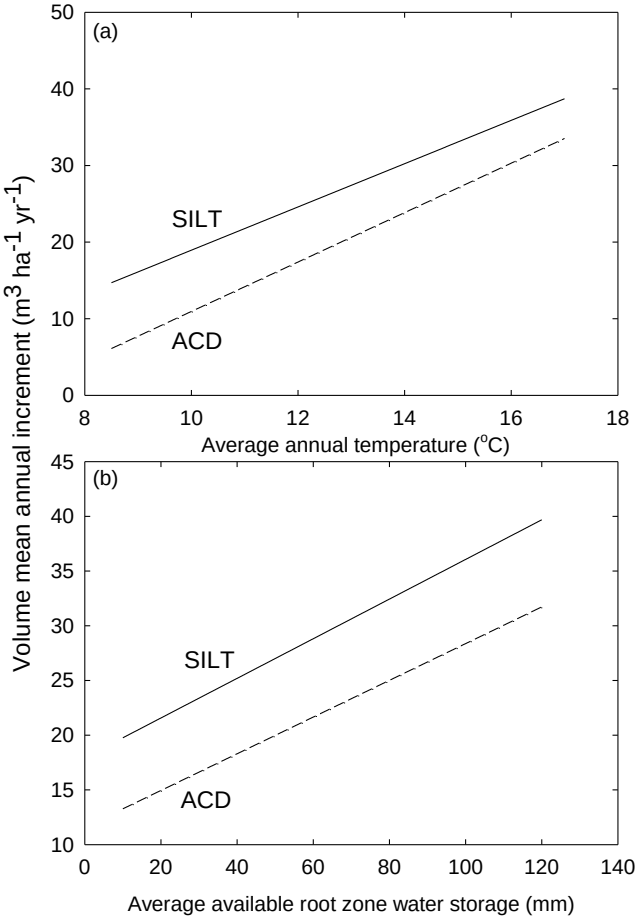


Figure 3. Response of volume mean annual increment (MAI_v) to (a) average annual temperature and (b) average available root-zone water storage for data obtained from actual climate data (ACD) and data extracted from spatially interpolated long-term (SILT) climate surfaces. For each response curve the variable not shown was held constant in the model at average values.

For both data sources bias and multicollinearity for the final models, which included T_a and W_a, were acceptable. Figure 4 shows a plot of the relationship between measured and predicted MAI_v, developed using ACD and SILT data, which exhibits little apparent model bias for both relationships. The variance inflation factor was less than 1.01 for all variables which indicated that multicollinearity between independent variables was within acceptable limits (less than 10).

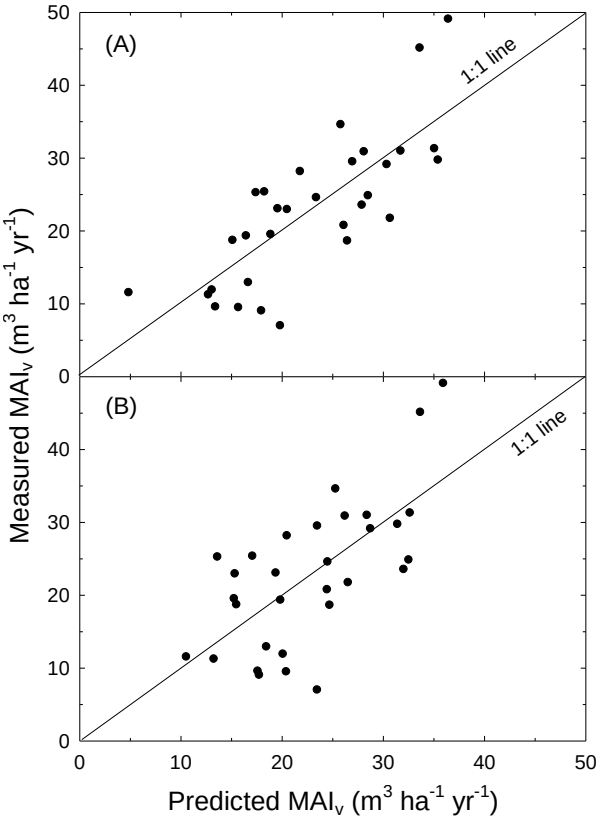


Figure 4. Plots showing the relationship between predicted and measured mean annual volume increment (MAI_v) for (A) actual climate data and (B) spatially interpolated long-term climate data. The solid line shows a 1:1 relationship.

Using both data sources, P was not found to be a significant predictor of MAI_v when included in combination with T_a. Multiple regression models, which included P and T_a, had coefficients of determination of 0.42 for ACD and 0.36 for the SILT climatic surfaces. Use of W_a, rather than P in the multiple regression model, accounted for an additional 19 % and 10 % of the variance, respectively, for ACD and SILT climatic surfaces.

4.0 DISCUSSION

Soil water balance has long been recognised as a major determinant of *P. radiata* growth at specific locations exhibiting seasonal water deficits (McMurtrie et al., 1990; Arneth et al., 1998a, 1998b; Richardson et al., 2002; Watt et al., 2003). More recently research has shown that root-zone water storage and mean annual temperature are the key climatic determinants of productivity across the national extent for *P. radiata* (Watt et al., 2008). Results from this study confirm these findings and clearly show the utility of GIS derived climatic data in predicting root-zone storage values, using the recently developed spatial model SWatBal (Chapter 3). This approach provides a cost effective means of modelling root-zone water storage that could potentially have a number of applications including the development of models to predict species productivity, both under current and future climate.

In our study, the application of nutrients in excess of crop requirements enabled determinants of MAI_v to be identified more readily across the environmental range by mitigating the confounding effects of limiting soil nutrient status. Hunter and Gibson (1984) commented that assessing optimal temperature ranges for productivity across New Zealand was problematic "since all the areas with temperatures warmer than the apparent optimum lie in the endemically nutrient deficient areas of the North Auckland Peninsula." They suggested an optimum average annual temperature for growth being 12°C, for stands that were mostly unfertilised. Our study, however, shows that *P. radiata* productivity (MAI_v) increases linearly with increasing T_a across the New Zealand

temperature range (Figure 3). Response curves indicate that MAI_v values for ACD at 8 °C to be ~13 m^3 ha^{-1} yr^{-1}, increasing to ~37 m^3 ha^{-1} yr^{-1} at 16 °C, suggesting optimum air temperature values were not reached in northern regions of New Zealand.

The development of new and improved national-extent environmental and climate surfaces for New Zealand will continue. For example, the National Institute of Water and Atmospheric Research (NIWA) continuously collects and updates data from climate stations across New Zealand (Tait et al., 2006). The advantages of regularly updated climate data are two-fold. The latest information can be used either to calculate new long-term climate averages – for example updated RAV data (Chapter 3) – or, should modelling require specific, harmonised periods of time, then data representing these periods can be extracted and used in modelling water balance.

Although this study was undertaken with high stand densities over a shorter growth period, when compared to mature forest stands, research suggests that trees grown at extremely high stand densities follow similar developmental patterns to operational stands (Watt et al., 2005). Amateis et al. (2003a, b) and Sharma et al. (2003) suggested that diameter and height patterns of loblolly pine (*Pinus taeda L.*) grown over a four-year period have similar development patterns to those of operational stands grown over 16 years when tree spacings are reduced to 1/16 of the operational stand. Watt et al. (2005, 2008) supported these findings, suggesting that growth patterns closely follow those of stands growing at operational stockings of 300 stems ha^{-1} over a 30-year period. These

results suggest that operational stands will command similar demand for resources when compared with trees grown at these high stocking densities.

As the conclusions drawn from this study are based on data derived from *P. radiata* plots grown at high stand densities, the utility of these data require some validation. The purpose of this approach was to compress the experimental period over which the stand grows and develops by increasing intra-tree competition and the rate at which the leaf area develops and the base of the green crown recedes. While literature suggests the efficacy of this experimental design is robust, further research into forest stands grown at operational spacings over the long term and across natural fertility gradients is required.

5.0 CONCLUSION

Actual mean annual temperature, T_a, was the variable most strongly related to mean annual volume increment (MAI_v), accounting for 42% of the variance in MAI_v. Modelling clearly demonstrated that using available root-zone water storage (W_a), in combination with T_a, accounted for 61% of the variance in *P. radiata* MAI_v. The main implication of this research is to endorse the move towards relatively low cost, long-term normalised spatial climate GIS surfaces in preference to the high cost of measuring and collecting data at each individual study location across New Zealand. Even though we recognise the collection of some essential environment site data will continue, the utility of the SWatBal model and the use of SILT climate data simplifies soil water balance

estimations considerably. While the results are compelling, further research ascertaining whether these findings are applicable to mature stands grown at conventional stand densities across natural fertility gradients is needed. If findings are confirmed, the role of soil water balance could become a more affordable and important component of modelling *P. radiata* MAI_v across New Zealand.

6.0 ACKNOWLEDGEMENTS

We gratefully acknowledge and thank David Whitehead for his helpful and valuable comments in the development of SWatBal. We thank Graham Coker for establishing the trial series used in this study and subsequent collection and collation of measurements used in the analysis. We are grateful to Trevor Webb for his technical assistance and advice with Landcare Research's datasets. David Palmer wishes to thank and acknowledge the Foundation for Research Science and Technology (FRST) and the University of Waikato (Doctoral Scholarship) for their funding. Barbara Höck thanks FRST for funding (Protecting and enhancing the environment through forestry CO4X0304). We are also indebted to the numerous forestry companies and private owners for supporting the installation of trial sites. We also acknowledge the helpful comments provided by two anonymous referees.

7.0 REFERENCES

Amateis, R.L., Sharma, M., Burkhart, H.E., 2003a. Scaling growth relationships from seedling plots using similarity analysis. Forest Science 49, 188-195.

Amateis, R.L., Sharma, M., Burkhart, H.E., 2003b. Using miniature-scale plantations as experimental tools for assessing sustainability issues. Canadian Journal of Forest Research 33, 450-454.

Anon., 1983. Summaries of Climatological Observation to 1980. New Zealand Meteorological Service Miscellaneous Publication 177, Wellington, New Zealand.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998a. Assessment of annual carbon exchange in a water-stressed *Pinus radiata* plantation: an analysis based on eddy covariance measurements and an integrated biophysical model. Global Change Biology 5, 531-545.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998b. Fluxes of carbon and water in a *Pinus radiata* forest subject to soil water deficit. Australian Journal of Plant Physiology 25, 557-570.

Barringer, J.R.F., Porteous, A., Salinger, M.J., Trangmar, B.B., 1995. Estimating spatial patterns of wilting point deficit using a water balance model and a geographic information system. Journal of Hydrology 34, 42-59.

Barringer, J., Lilburne, L., 1999. Scale issues in developing regional-scale soil water balance surfaces. In: Whigham, P.A., (Ed.), Proceedings of the Eleventh Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand, pp. 231-240.

Der, G., Everitt, B.S., 2001. A handbook of statistical analyses using SAS, 2nd ed. CRC Press.

Gradwell, M.W., 1972. Methods for physical analysis of soils. New Zealand Soil Bureau Scientific Report No. 10C. Wellington, DSIR.

Hewitt, A.E., 1995. Soil map of the South Island, New Zealand Soil Classification 1:100,000 scale. Manaaki Whenua Press, Lincoln, NZ.

Hewitt, A.E., 1998. New Zealand Soil Classification, 2nd ed. Landcare Research Science Series 1, 133 pp.

Hunter I.R., Gibson, A.R., 1984. Predicting *Pinus radiata* site index from environmental variables. New Zealand Journal of Forestry Science 14, 53-64.

Hutchinson, M.F., 1984. A summary of some surface fitting and contouring programs for noisy data. Consulting Report No. ACT 84/6. CSIRO Division of Mathematics and Statistics. CSIRO. Canberra.

Hutchinson, M.F., Gessler, P.E., 1994. Splines – more than just a smooth interpolator. Geoderma 62, 45-67.

Jackson, D.S., Gifford, H.H., 1974. Environmental variables influencing the increment of radiata pine (1) periodic volume increment. New Zealand Journal of Forestry Science 4, 3-26.

Kelliher, F.M., Whitehead, D., McAneney, K.J., Judd, M.J., 1990. Partitioning evapotranspiration into tree and understorey components in two young *Pinus radiata* D. Don stands. Agricultural and Forest Meteorology 50, 211-227.

Kelliher F.M., Whitehead, D., Pollock, D.S., 1992. Rainfall interception by trees and slash in a young *Pinus radiata* D. Don stand. Journal of Hydrology 131, 187-204.

Kirkpatrick, R., 1999. Plate 11 (Soils) in: Bateman Contemporary Atlas New Zealand. David Bateman, Auckland.

Leathwick, J.R., Stephens, R.T.T., 1998. Climate surfaces for New Zealand. Landcare Research contract report: LC9798/126. Landcare Research, Hamilton, New Zealand, 19 pp.

Leathwick, J.R., Wilson, G., Stephens, R.T.T., 2002. Climate surfaces for New Zealand. Landcare Research contract report: LC9798/126. Landcare Research, Hamilton, New Zealand, 22 pp.

Lewis, N.B, Ferguson I.S., 1993. Management of radiata pine. Inkata Press, Melbourne. 404 pp.

Lohammer, T., Larsson, S., Linder, S., Falk, S.O., 1980. FAST – Simulation models of gaseous exchange in Scots pine. Ecological Bulletin 32, 505-523.

McMurtrie, R.E., Rook, D.A., Kelliher, F.M., 1990. Modelling the yield of *Pinus radiata* on a site limited by water and nitrogen. Forest Ecology and Management 30, 381-413.

Nambiar, E.K.S., Sands, R., 1993. Competition for water and nutrients in forests. Canadian Journal of Forest Research 23, 1955-1968.

Nix, H.A., 1986. A biogeographical analysis of Australian Elapid snakes. In: Longmore, R. (Ed.) "Atlas of Elapid snakes of Australia". Australian Government Publishing Service.

NZFOA, New Zealand Forest Owners Association, 2005. New Zealand Forest Industry Facts and Figures 2004/2005. New Zealand Forest Owners Association, P.O. Box 1208, Wellington.

Palmer, D.J., Höck, B.K., Dunningham, A.G., Lowe, D.J., Payn, T.W., 2008. Developing national-scale terrain attributes for New Zealand (TANZ). Scion Bulletin (in press).

Richardson, B., Whitehead, D., McCracken, I.J., 2002. Root-zone water storage and growth of *Pinus radiata* in the presence of a broom understorey. New Zealand Journal of Forestry Science 32, 208-220.

Rijkse, W.C., Hewitt, A.E., 1995. Soil map of the North Island, New Zealand Soil Classification 1:100,000 scale. Manaaki Whenua Press, Lincoln, NZ.

SAS-Institute-Inc 2000. SAS/STAT User's Guide: Version 8. Volumes 1, 2 and 3. SAS Institute Inc., Cary, North Carolina. pp. 3884.

Sharma, M., Amateis, R.L., Burkhart, H.E. 2003. Forest stand dynamics and similarity theory. Ecological Modelling, 167, 165-180.

Tait, A., Henderson, R., Turner, R., Zheng, X., 2006. Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26, 2097-2115.

Thompson, S., Grüner, I., Gapare, N., 2003. New Zealand Land Cover Database Version 2. Illustrated Guide to Target Classes. Report for the Ministry for the Environment.

Vincent, T.G., Dunstan, J.S., 1989. Register of commercial seedlots issued by the New Zealand Forest Service. Ministry of Forestry, Rotorua. Forest Research Institute Bulletin No. 144. Forest Research Institute, Rotorua, New Zealand, 155 pp.

Wahba, G., Wendelberger, J., 1980. Some new mathematical methods for variational objective analysis using splines and cross validation. Monthly Weather Review, 108, 1122-1143.

Watt, M.S., Coker, G., Clinton, P.W., Davis, M.R., Parfitt, R., Simcock, R., Garrett, L., Payn, T., Richardson, B., Dunningham, A., 2005. Defining sustainability of plantation forests through identification of site quality indicators influencing productivity – A national view for New Zealand. Forest Ecology and Management 216, 51-63.

Watt, M.S., Davis, M.R., Clinton, P.W., Coker, G., Ross, C., Dando, J., Parfitt, R.L., Simcock, R., 2008. Identification of key soil indicators influencing plantation productivity and sustainability across a national trial series in New Zealand. Forest Ecology and Management (in press).

Watt, M.S., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C., 2003. Modelling the influence of weed competition on the growth of young *Pinus radiata* at a dryland site. Forest Ecology and Management 178, 271-286.

Whitehead, D., Leathwick, J.R., Walcroft, A.S., 2001. Modeling annual carbon uptake for the indigenous forests of New Zealand. Forest Science 47, 9-19.

Whitehead, D., Livingston, N.J., Kelliher, F.M., Hogan, K.P., Pepin, S., McSeveny, T.M., Byers, J.N., 1996. Response of transpiration and photosynthesis to a transient change in illuminated foliage area for a *Pinus radiata* D. Don tree. Plant Cell and Environment 19, 949-957.

Woollons, R.C., Skinner, M.F., Richardson, M.F., Rijske, W.C., 2002. Utility of "A" horizon soil characteristics to separate pedological groupings, and their influence with climatic and topographic variables on *Pinus radiata* height growth. New Zealand Journal of Forestry Science 32, 195-207.

View from the Kepler track towards the South Fiord of Lake Te Anau illustrating the upper vegetation limit at higher elevations (photograph by Dr Kyle Bland)

Slight not what's near, though aiming for what is far.

Euripides, *Rhesus*

PREDICTING THE SPATIAL DISTRIBUTION OF *PINUS RADIATA* PRODUCTIVITY IN NEW ZEALAND USING INTERPOLATED SURFACES AND ANCILLARY MAPS

D.J. Palmer[a*], M.O. Kimberley[b], B.K. Höck[b], M.S. Watt[c], D.J. Lowe[a], T.W. Payn[b]

[a]*Department of Earth and Ocean Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand 3240*

[b] *Scion, Private Bag 3020, Rotorua, New Zealand*

[c] *Scion, PO Box 29237, Christchurch, New Zealand*

[*]Corresponding author. Tel.: +64-7-8562889 *E-mail address*: djp8@waikato.ac.nz

Prepared for submission to: Agricultural and Forest Meteorology

ABSTRACT

The aim of this study was to develop maps of known certainty that describe the spatial variability of *P. radiata* productivity across New Zealand's national extent. A model for predicting *Pinus radiata* volume mean annual increment (referred to as the 300 Index) and mean top height at age twenty (referred to as Site Index) for New Zealand using regression kriging (RK) was developed. Partial least squares (PLS) regression was used to predict forest productivity from 1146 permanent sample plots (PSP) using variables extracted from GIS surfaces covering the national extent of New Zealand. These spatial datasets included climate, landuse, terrain, and environmental surfaces. The best PLS models explained 58% and 67% of the variance for the 300 Index and Site Index, respectively. Ordinary kriging was used to spatially interpolate residuals from the PLS models before summation with the PLS-modelled surfaces to improve predictions. A cross validation procedure withheld 618 out of 1764 sites from the modelling process for assessment of measured forest productivity against model predictions. Results show RK substantially improved PLS regression predictions, increasing R^2 values by 12 % and 13 % for Site Index and 300 Index, respectively. Root mean square error (RMSE) and goodness of precision (G) cross validation statistics confirmed these findings, with RK showing significant improvement for both the 300 Index and Site Index validation predictions. The main contributors of *P. radiata* productivity were assessed by comparing variable groupings averaged across New Zealand environmental regions (LENZ). Findings highlight the importance of air temperature, water status (mainly water deficit), and profile variable groupings for both models.

Keywords: forest productivity, 300 Index, Site Index, spatial modelling, national scale modelling, partial least squares, regression kriging, New Zealand.

1. INTRODUCTION

The most widely planted commercial forestry crop in New Zealand (NZFOA, 2005) and the Southern Hemisphere (Lewis and Ferguson, 1993) is *Pinus radiata* D. Don. Given its extensive distribution and substantial contribution to the New Zealand economy, the identification of highly productive sites is important. Beyond the economic significance of wood and wood-fibre, and with the new era of carbon budgeting and sequestration, less productive sites requiring less capital outlay may become increasingly desirable. There is also increasing demand for renewable energy sources including biofuel from wood fibre. Thus, considering the economic and environmental importance of the *P. radiata* species, mapping productivity and site quality for New Zealand using observed data across the national extent will provide information essential in determining national inventory and site suitability. Therefore, the development of appropriate empirically based national models and their predictive maps will provide new and useful information for forestry investors, managers, planners, and scientists.

In New Zealand the most widely used measure of site quality for *P. radiata* stands is Site Index, defined as the mean top height at age twenty years (Goulding, 2005). However, stocking trials indicate that Site Index remains only weakly related to stand basal area

growth, especially in older stands with higher stocking (Kimberley et al., 2005). Therefore, when assessing forest productivity, Site Index provides only a partial measure of stand productivity. Consequently, the 300 Index was developed as a measure of productivity for New Zealand, and is defined as "the stem volume mean annual increment at age thirty years for a defined reference regime of 300 stems ha^{-1}" (Kimberley et al., 2005).

Since the early 1970s many studies covering wide spatial extents have looked at drivers of *P. radiata* forest productivity and site quality in locations including New Zealand (Jackson and Gilford, 1974; Hunter and Gibson, 1984; Woollons et al., 2002; Watt et al., 2005, 2008), Chile (Schlatter and Gerding, 1984), Spain (Sánchez-Rodríguez et al., 2002; Romanyà and Vallejo, 2004), and Australasia (Czarnowski et al., 1971). Varying degrees of success were achieved with up to 72% of the variance in *P. radiata* forest productivity explained by Watt et al. (2005). For many of the earlier studies the two variables air temperature and rainfall were considered to be the greatest contributors toward *P. radiata* site quality and productivity. Numerous New Zealand studies (e.g. Arneth et al., 1998a, 1998b; McMurtrie et al., 1990; Richardson et al., 2002; Watt et al., 2003) have shown *P. radiata* growth to be correlated with seasonal water deficits at single point locations. However, little research has investigated the utility of soil water balance as a predictor of *P. radiata* productivity across the national extent of New Zealand.

For New Zealand, the majority of these previous studies have used observed field data as independent explanatory variables. These types of models were designed to assess the

relationships between site properties and forest productivity or site quality. Because of their reliance on observational field data, often from relatively small datasets, thematic spatial representation was not possible. One exception, considered complementary to the Hunter and Gibson (1984) study, was the site ranking index developed by Eyles (1986). The indices generated by Eyles (1986) for *P. radiata* were associated with the New Zealand land use capability units (NWASCA, 1975-79) in collaboration with expert knowledge.

Since the 1990s, GIS platforms have become increasingly important because of their ability to process and store large volumes of spatial data (Palmer et al., 2008). Although the role of GIS in managing spatial data and analysing forestry spatial trends is clear (Höck et al., 1994; Payn et al., 1999; Palmer et al., 2005), data measurement and collection for projects other than localised detailed studies are often cost prohibitive. In addition, even though the availability and accuracy of New Zealand climate data (Leathwick et al., 2002b; Mitchell, 1991; Tait et al., 2006; Chapter 3) and terrain attributes (Palmer et al., 2008) are relatively high, the precision and accuracy of some New Zealand soil-related surfaces remain unclear. In Australia, a new approach utilised the national soils point database (Henderson et al., 2005) to fill gaps in soil survey information over large spatial extents (Bui and Moran, 2003) using techniques including expert knowledge and rule-based induction (Bui and Moran, 2001). In New Zealand, the fundamental soil layers (Newsome et al., 2000) currently provide the best nationally available soils information dataset; S-map, currently being developed (Landcare Research, 2008), will eventually provide digital maps for a raft of soil physical and

chemical properties across New Zealand. Although the accuracy and precision of spatially interpolated data vary greatly, potentially in many situations these data could be used effectively, thus reducing many of the costs associated with the measurement and collection of field data. Indeed, the use of primary and secondary terrain attributes is becoming increasingly popular for studies of both restricted and wide spatial extents.

Historically, many of the site productivity models were developed using multiple linear regression analysis. Multiple linear regression is considered an appropriate modelling technique where (i) independent variables are few in number, (ii) such variables are not significantly redundant (collinear), and (iii) the variables have well understood relationships to the responses. Cross validated partial least squares (PLS) regression may provide an alternative modelling technique where the above first and second conditions may have been breached. Tegelmark (1998) illustrated the use of PLS technique for modelling the regeneration of Scots pine in Sweden. It may be possible to gain further improvements to multiple linear regression model predictions through the application of regression kriging techniques as described by Odeh et al. (1995).

Using 300 Index and Site index values derived from a PSP database, the objectives of this study were to (i) develop productivity models by correlation with data derived from national extent ancillary maps and interpolated surfaces using the PLS approach, (ii) extend the PLS model through regression kriging (RK, kriged PLS residuals), (iii) undertake an independent quantitative validation of methods described in (i) and (ii) to determine which of the two approaches provided the best estimate of Site Index and 300

Index across the national extent, (iv) determine the drivers of *P. radiata* productivity across environmental regions where possible, and (v) develop surfaces from these models useful in the delineation of spatial variation of *P. radiata* 300 Index and Site Index across both North and South islands of New Zealand.

2.0 METHODS

2.1 Permanent sample plot data and preliminary screening

Scion Research collects and stores permanent sample plot (PSP) data on behalf of many private forestry owners. Extractions of PSP data were undertaken for sites across New Zealand where permission for use was given. The PSP data were further scrutinised for sites that could adversely influence the integrity of the dataset. Exclusions included removal of Nelder (spacing), oversowing (legume N), disturbance (forest floor removal), and fertiliser (P, N, and K) site trials. However, in many of these excluded trials, it was possible to identify control plots which were retained. Other exclusions included the removal of data from stands planted prior to 1975 and for stands less than seven years in age. Based on preliminary screening of the PSP data, it was found that stands established in the 1930s had 300 Index values ~25% lower than stands established since 1975. Furthermore, experience suggested that site productivity indices based on very young trees (less than 7 years age) were unreliable. PSPs with plot history data inadequate for estimating the 300 Index were also excluded. The national dataset was also vetted for

duplicates and all sites were averaged to a 100-m grid that aligned with model development surfaces.

The productivity indices were calculated for the 1764 PSP locations across New Zealand using the procedure discussed by Kimberley et al. (2005). To calculate Site Index, a national height/age model (an equation for predicting height for any age and Site Index) was used. By inverting the equation, it is possible to obtain Site Index as a function of age and mean top height. In our study, the measurement closest to age 20 years was used for each PSP. Estimation of the 300 Index, which is a measure of stem volume productivity, is more complex because, unlike height, stem volume is strongly influenced by stocking density and, to a lesser extent, thinning and pruning history. To calculate the 300 Index, a plot measurement consisting of the Basal Area, Mean Top Height and Stocking Density at a known age (again, the measurement closest to age 20 years was used for each PSP in this study), along with stand history information detailing the initial stocking, timing and extent of thinnings, and timing and height of prunings, is required. The 300 Index estimation procedure utilised the 300 Index Model, an empirical growth model which is sensitive to all the above inputs, and which was calibrated to a site by the 300 Index, effectively a local site productivity parameter. An iterative procedure was used to determine the 300 Index parameter value that will predict the plot measurement. All variables were projected to a New Zealand map grid projection with a New Zealand geodetic datum, NZGD1949.

2.2 GIS data extraction and pre-processing

Properties of interest that could potentially serve as independent variables in the modelling of *P. radiata* productivity were prepared. Pre-processing included the conversion of scaled integer data to floating point and the use of variable transformations to convert each dataset into an approximately normal Gaussian distribution. These transformations were achieved using Box-Cox power transformations (Box and Cox, 1964). Datasets explored included primary and secondary terrain attributes (Palmer et al., 2008), monthly soil water balance (Chapter 3), climate (Mitchell, 1991; Leathwick et al., 2002b; Tait 2006), fundamental soil layers (FSL) and land resource information (LRI) (Newsome et al., 2000), vegetative cover (Newsome, 1987), foliar nutrition (Hunter et al., 1991), and other environmental surfaces (Leathwick et al., 2002a; 2003). From these datasets, variables were extracted for each of the 1764 PSP locations. Descriptions of the variables including units, classifications and transformations, are given in Tables 1 and 2. A brief synopsis follows.

Terrain attributes used in modelling productivity included the compound topographic wetness index (TWI), the length and slope factors (LS), and plan curvature (PlanC). TWI is used extensively to "describe the effects of topography on the location and size of saturated source areas of runoff generation" (Wilson and Gallant, 2000). The TWI dataset was further divided into four classes from unsaturated through to saturated locations. The LS factor is a component of the universal soil loss equation and represents the slope-length and slope-steepness, calculated using the equations of McCool et al. (1987; 1989).

The LS factor used here was separated into seven equal classes representing a continuum from low (1) through to high (7). PlanC captures topographic convergence and divergence of flow across a landscape and can be used to differentiate between valleys and ridges. Negative values represent converging flow in valleys whereas positive values represent diverging flow on ridge lines. Refer to Table 1 for classification details.

The water status variables used in modelling *P. radiata* productivity comprised rainfall and water balance. The development of mean annual available root-zone water storage (W_a), summer mean fraction of available root-zone water storage (W_f), and drainage for spring and summer for New Zealand were described in Chapter 3. Both W_a and W_f datasets were also reclassified into five classes (see Tables 1 and 2 for details). The annual water deficit and monthly water balance ratio were developed by Leathwick et al. (2002a; 2003). The only rainfall dataset used in the productivity modelling was the summer average which was developed by averaging the Leathwick et al. (2002b) surfaces for the summer rainfall months.

Climate variables used in modelling productivity included seasonal averages for temperature, numbers of days ground frost, vapour pressure deficit, solar radiation, and wind speed. All seasonal climate data were calculated by averaging the Leathwick et al. (2002b) surfaces across their respective summer, autumn, and winter periods. The mean daily minimum temperature for June (Tmin Jun) was taken from Leathwick et al. (2002a; 2003), while the mean daily temperature for summer (Tdaily sum) was calculated from

Leathwick et al. (2002b) minimum and maximum temperature datasets using ((2(Tmax))+Tmin)/3.

Acid soluble phosphorus (AcidP) provided an estimation of parent material phosphorus levels and has five classes from very low (class 1) through to very high (class 5) (Leathwick et al., 2003). The variable phosphate retention (Prent) is an estimation of the soil's capacity to fix phosphorus in the upper 20 cm of the profile, whereas soil pH is an estimation of the acidity or alkalinity of the soil (Newsome et al. (2000).

The landuse variables used in modelling *P. radiata* productivity were land use classification (NZLUC) and vegetation cover (VCMNZ). NZLUC (Newsome et al., 2000) are digital layers developed from NWASCA (1975-79) maps provide an estimate of the general versatility of land for productive use. Our study reclassified these data into high (classes 1, 2, and 3), medium (classes 4 and 5), and low versatility (classes 6, 7, and 8). The vegetation cover map for New Zealand was developed over the 1981 to 1985 period (Newsome et al., 1987). VCMNZ was developed by categorising the vegetation cover map into (1) cropland and grassland, (2) grassland-scrub, grassland forest, forest-scrub, and forest, and (3) exotic forest classes.

The general grouping termed profile comprised soil data including the New Zealand Soil Classification (Hewitt, 1998) for soil order, the topsoil gravel content, and profile readily available water content, were all derived from Newsome et al. (2000). The average particle size and induration of soil parent material were from Leathwick et al. (2002a;

2003). Soil parent material based on common lithology was developed by us using maps from New Zealand Soil Bureau (1973) and data from Newsome et al. (2000). Further details are given in Tables 1 and 2.

Table 1: Independent variables and general groupings for the 300 Index partial least squares model. For transformed variables, power coefficients are given in parentheses.

Variable name	Variable description	Units	General groupings
PM[1]	Soil parent material based on common lithology[a]	-	Profile
NZSC Order[2]	New Zealand Soil Classification	-	Profile
GRAV[3]	Topsoil gravel content	%	Profile
PRAW[3]	Profile readily available water	mm	Profile
Induration[4]	Induration of parent material	-	Profile
Psize[4]	Particle size	mm	Profile
Deficit[4]	Annual water deficit	mm	Water status
Wbal ratio[4]	Monthly water balance ratio ($^{-0.5}$)	-	Water status
Rain sum[5]	Mean monthly rainfall for summer ($^{-0.5}$)	mm	Water status
W_a avg[6]	Mean annual available root-zone water storage ($^{0.5}$)	mm	Water status
Drain spr[6]	Mean drainage for spring[b]	mm	Water status
Drain win[6]	Mean drainage for winter	mm	Water status
W_a avg class[6]	Mean annual fraction of available root-zone water storage[c]	-	Water status
pH[3]	pH	-	Nutrient status
AcidP[4]	Subsoil acid soluble phosphorus	Mg/100g	Nutrient status
VCMNZ[7]	Vegetation cover map for New Zealand[d]	-	Landuse
NZLUC[3]	New Zealand Land Use Classification[e]	-	Landuse
PlanC[8]	Plan curvature[f]	-	Terrain attributes
TWI[8]	Topographic wetness index[g]	-	Terrain attributes
DGF aut[5]	Mean number of days ground frost for autumn	°C	Air temperature
DGF sum[5]	Mean number of days ground frost for summer	°C	Air temperature
Tmin win[5]	Mean daily minimum temperature for winter	°C	Air temperature
VPD aut[5]	Mean daily 9 am vapour pressure deficit for autumn ($^{1.25}$)	kPa	VPD
VPD oct[4]	Mean daily 9 am vapour pressure deficit for October	kPa	VPD
VPD win[5]	Mean daily 9 am vapour pressure deficit for winter	kPa	VPD
Srad sum[5]	Mean daily solar radiation for summer	MJ m^{-2} day^{-1}	Solar radiation

[1] Refer to Newsome et al. (2000) and New Zealand Soil Bureau (1973)
[2] Refer to Newsome et al. (2000) for spatial details and Hewitt (1998) for class descriptions
[3] Refer to Newsome et al. (2000)
[4] Refer to Leathwick et al. (2002a; 2003)
[5] Refer to Leathwick et al. (2002b)
[6] Refer to Chapter 3
[7] Refer to Newsome (1987)
[8] Refer to Palmer et al. (2008)
[a] Dataset derived from LRI[2] based on parent materials of common lithology (1) dune sand, (2) loess, (3) volcanic rocks: loose (4) volcanic rocks: extremely weak to weak (5) volcanic rocks: weak to extremely strong (6) soft sediments (7) sedimentary rocks: moderate to extremely strong (< Tertiary) (8) older sedimentary rocks: (pretretiary and schist) (9) intrusive rocks: including gneiss (10) sedimentary rocks: very compact (stiff) to weak (< Tertiary)
[b] Model includes a quadratic form of the dataset
[c] W_a avg was reclassified into five classes (1) 0 to 60, (2) 60 to 70, (3) 70 to 80, (4) 80 to 90, (5) 90 to 100 mm
[d] Vegetation cover map was reclassified into 3 classes (1) improved pasture, (2) exotic forests, (3) other classes
[e] NZLUC was reclassified into three classes (1) 1, 2, and 3, (2) 4 and 5, (3) 6, 7, and 8 (classes grouped)
[f] PlanC was reclassified into five classes (1) < -29, (2) -30 to -14, (3) -15 to 15, (4) 16 to 30, (5) 31 < (radians 100 m^{-1})
[g] TWI was reclassified into four classes (1) < 5, (2) 5 to 10, (3) 10 to 15, (4) 15 <

Table 2: Independent variables and general groupings for the Site Index partial least squares model. For transformed variables, power coefficients are given in parentheses.

Variable Name	Variable description	Units	General groupings
DGF sum[1]	Mean number of days ground frost for summer[a]	°C	Air temperature
Tmax sum[1]	Mean daily maximum temperature for summer[a]	°C	Air temperature
Tavg sum[1]	Mean daily average temperature for summer	°C	Air temperature
Tmin jun[5]	Mean daily minimum temperature for June[a]	°C	Air temperature
Tdaily sum[1]	Mean daily average temperature for summer[b]	°C	Air temperature
PM[2]	Soil parent material based on common lithology[c]	-	Profile
NZSC Order[3]	New Zealand Soil Classification[d]	-	Profile
GRAV[4]	Topsoil gravel content	%	Profile
Wind sum[1]	Mean daily wind speed for summer ($^{-0.25}$)	km hr^{-1}	Wind
Deficit[5]	Annual water deficit	mm	Water status
W_f sum[6]	Mean monthly fraction of available root-zone water storage for summer	mm	Water status
W_f sum class[6]	Mean monthly fraction of available root-zone water storage for summer[e]	-	Water status
AcidP[5]	Subsoil acid soluble phosphorus	Mg/100g	Nutrient status
Prent[4]	Phosphorus retention in upper profile	%	Nutrient status
LS[7]	Length and slope factor[f] ($^{\text{natural log}}$)	-	Terrain attributes
TWI[7]	Topographic wetness index[g]	-	Terrain attributes

[1] Refer to Leathwick et al. (2002b)
[2] Refer to Newsome et al. (2000) and New Zealand Soil Bureau (1973)
[3] Refer to Newsome et al. (2000) for spatial details and Hewitt (1998) for class descriptions
[4] Refer to Newsome et al. (2000)
[5] Refer to Leathwick et al. (2002a; 2003)
[6] Refer to Chapter 3
[7] Refer to Palmer et al. (2008)
[a] Model includes a quadratic form of the dataset
[b] Dataset was calculated from ((2*tmax[1])+tmin[1])/3 across the summer period
[c] Dataset derived from LRI[2] based on parent materials of common lithology (1) dune sand, (2) loess, (3) volcanic rocks: loose (4) volcanic rocks: extremely weak to weak (5) volcanic rocks: weak to extremely strong (6) soft sediments (7) sedimentary rocks: moderate to extremely strong (< Tertiary) (8) older sedimentary rocks: (pretretiary and schist) (9) intrusive rocks: including gneiss (10) sedimentary rocks: very compact (stiff) to weak (< Tertiary)
[d] NZLUC was reclassified into three classes (1) 1, 2, and 3, (2) 4 and 5, (3) 6, 7, and 8 (classes grouped)
[e] W_f sum was reclassified into five classes (1) 0 to 40, (2) 40 to 50, (3) 50 to 60, (4) 60 to 80, (5) 80 to 100
[f] LS factor was reclassified into seven classes (1) < 5, (2) 5 to 10, (3) 10 to 15, (4) 15 to 20, (5) 20 to 25, (6) 25 to 30, (7) 30 <
[g] TWI was reclassified into four classes (1) < 5, (2) 5 to 10, (3) 10 to 15, (4) 15 <

2.3 The partial least squares modelling technique

Stand data were analysed using the multivariate method PLS regression originally developed by Herman Wold (1966). The SAS (Version 9) PLS procedure was used, with the observations separated into training (n = 1146) and test (n = 618) datasets for cross validation purposes. Figure 1 shows the PSP locations across New Zealand for both the training and randomly selected test sites. Note that the test dataset locations were withheld from all further interpolation procedures and were used for the determination of model prediction bias and precision.

The overall goal of PLS is to predict a vector or matrix Y from a matrix X of explanatory variables, using the common structure underlying Y and X (Abdi, 2003). PLS is a regression technique for the identification of underlying orthogonal factors, which are linear combinations of the explanatory variables X (these factors are also known as latent variables) which best model the response or Y variable or variables. The assumption in PLS is that directions in the predictor space that are well sampled should provide better prediction for new observations when the predictors are highly correlated. Prior to calculation, predictors and responses are centred and scaled to have a mean of zero and a standard deviation of one, giving all variables the same relative importance. The number of factors chosen is usually based on the minimum predicted residual sum of squares (PRESS). We applied the van der Voet (1994) test choosing the least number of extracted factors whose residuals were not significantly greater than those of the model with minimum error (minimum PRESS) with a significant level of 0.05.

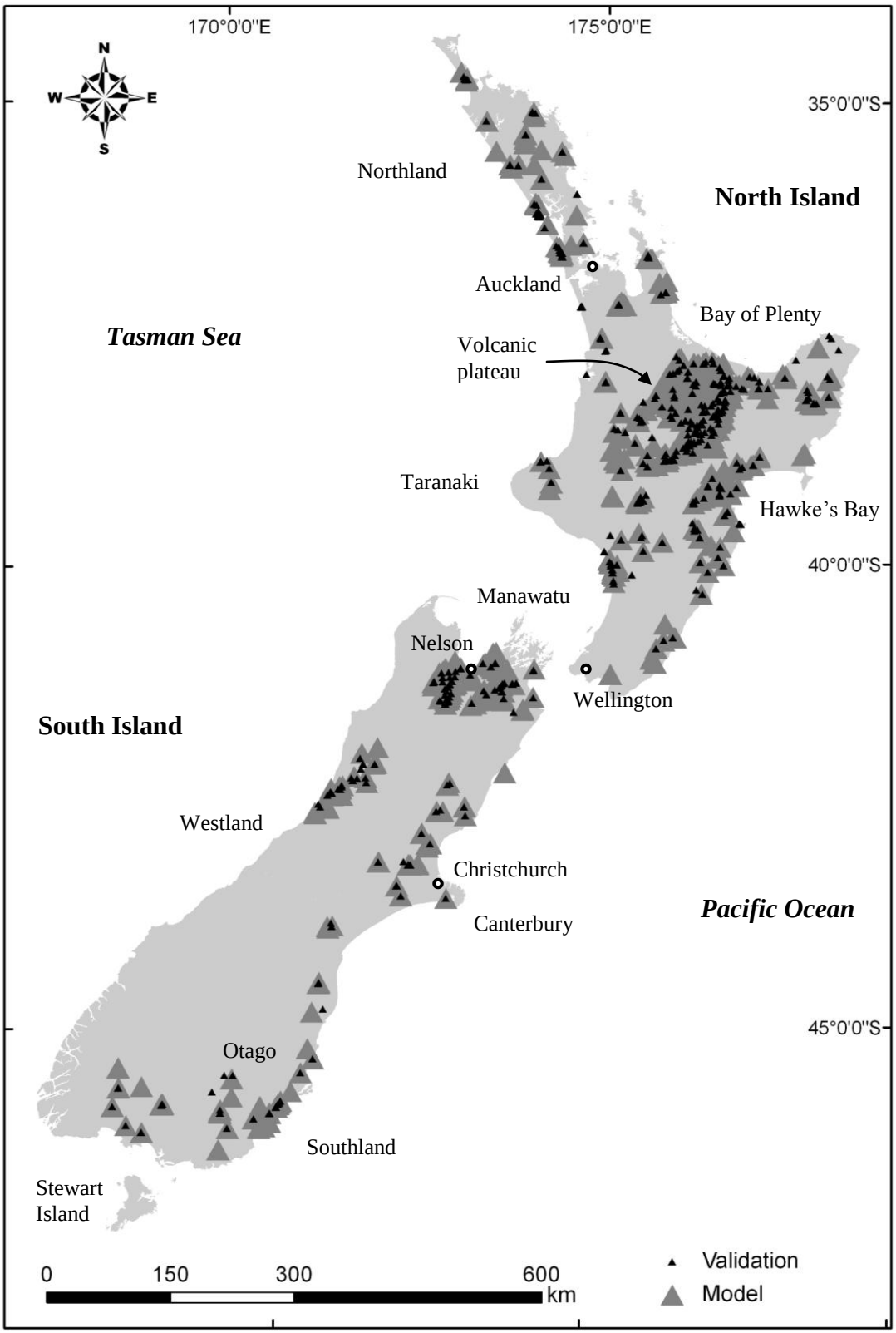


Figure 1: Distribution of PSP sites used to model *Pinus radiata* productivity and the withheld sites randomly selected for model validation. Place names mentioned in the text are also shown on the figure.

Predictors providing only small contributions towards the model were eliminated using a combination of two methods. Regression coefficients for the standardised data were screened for predictors with minimal contribution to the response prediction. Another statistic, variable importance for projection (VIP) (Wold, 1994), represents the value of each predictor in fitting the PLS model for both predictors and response. Therefore, if a predictor has a relatively small coefficient (in absolute value) and a small VIP, then it is a prime candidate for removal. Using the above protocol, independent variables were reduced to a minimum while maximising the explanation of the variance for the 300 Index and Site Index.

2.4 Development methods for assessing environmental drivers of productivity

Although the PLS modelling technique is robust when facing data noise, missing data, and collinearity among independents, the interpretation of the loadings of the independent latent variables is not straightforward (Tobias, 1997). Furthermore, PLS is not necessarily ideal in delineating underlying relationships between variables. To provide an overview, and as a progression towards understanding the importance of variables contributing towards productivity, the following analysis was performed using land environments of New Zealand (LENZ – level 1) (Figure 2).

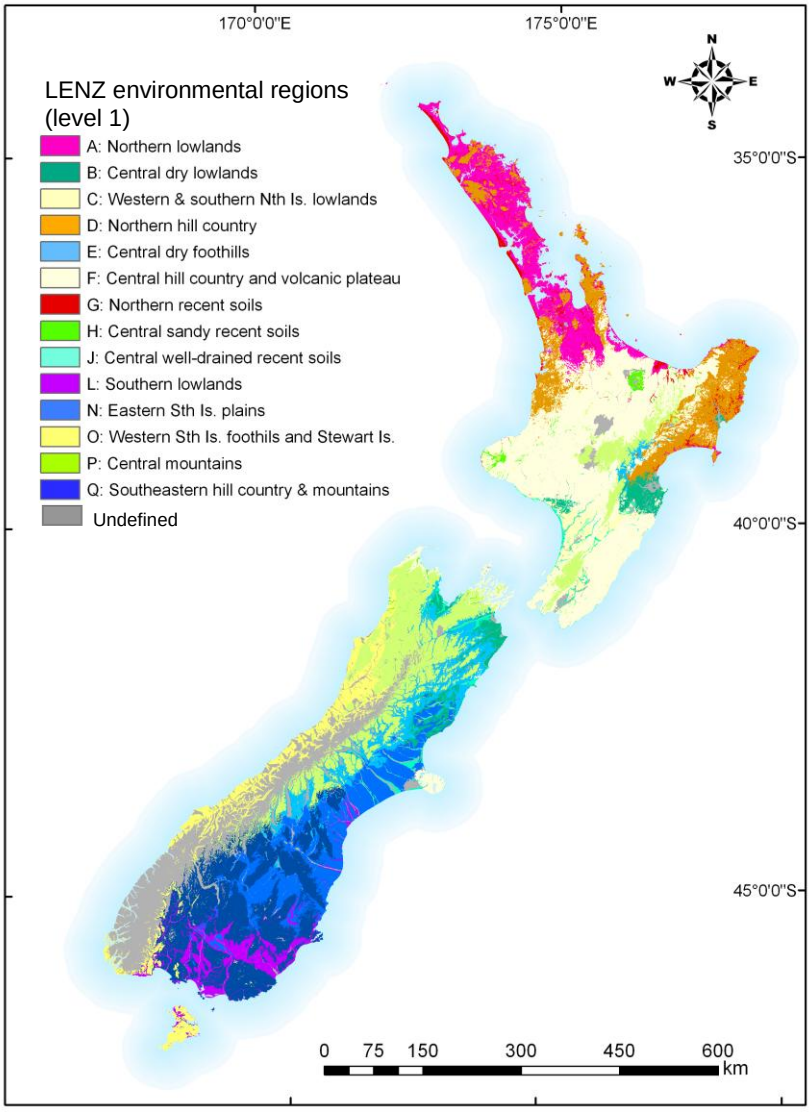


Figure 2: Map illustrating the fourteen most general classes (level 1) of Land Environments for New Zealand (LENZ) (Leathwick et al., 2002a; 2003) used to delineate and group biophysical and environmental variables used in the modelling of the 300 Index and Site Index.

To help ensure that our investigation did not get too unwieldy, the independent variables were first given a group according to their common properties (general groupings, Tables 1 and 2), and then assigned the appropriate LENZ class (level 1) (where LENZ is New Zealand's environmental classification system with 14 classes recognised in Level 1 (Figure 2) (Leathwick et al., 2002a; 2003). Secondly, the contribution of each independent variable was calculated by multiplying its value at each PSP location by its

coefficient within the PLS model, i.e., we assumed all other variables and the intercept are zero. The model contribution of the grouping was then determined by (i) averaging the individual contributions of each of the independent variables in each grouping (this is the mean for each grouping for all of New Zealand), and (ii) averaging the individual contributions of each of the independent variables for each LENZ region. The deviation from the mean within each LENZ class is then the difference between (i) and (ii).

2.5 Geostatistical prediction techniques

The regression kriging technique (Odeh et al., 1995) was used to improve predictions from the PLS model. Regression kriging used here requires the spatial interpolation of the PLS residuals using ordinary kriging before summation of the residual surface with the PLS predictions. PLS residuals refer to the difference between predicted values from the PLS model and the measured (observed) values. The 300 Index and Site Index predictions (z^*), made using regression kriging, were calculated at unvisited sites (S_o) using:

$$z^*(S_o) = zpr^* + \varepsilon^* \quad (1)$$

where zpr^* is the 300 Index or Site Index PLS predictions and ε^* is the kriged PLS residuals.

Experimental (semi-)variograms were developed for the PLS residuals from the training datasets and theoretical curves fitted to represent the data. Ordinary kriging estimated the weighted average of the 300 Index and Site Index at the unvisited sites, $z^*(S_o)$, and takes the form:

$$z^*(S_o) = \sum_{i=1}^{n} \lambda_i z(S_i) \qquad (2)$$

where λ_i are the weights given to each of the n development sites and $z(S_i)$ represents the 300 Index or Site Index PLS residual values measured at the development sites. An unbiased estimation was achieved because the weights sum to unity (Odeh et al., 1994; Schloeder et al., 2001). SAS (Version 9) was used in the development of all kriged surfaces.

2.6 Evaluation of prediction techniques

Performances of both the prediction technique PLS and regression kriging were evaluated by determining how close the predicted values (z^*) were to the observed values (z) at the (l) test (validation) sites (S_j). The bias and precision of the prediction techniques were assessed using the mean error (ME) and the root-mean-square error (RMSE), respectively, as defined by Voltz and Webster (1990) and Triantafilis et al. (2001). The ME was calculated using:

$$\mathrm{ME} = \frac{1}{l}\sum\left[z(s_j) - z^*(s_j)\right] \tag{3}$$

ME values should be closer to zero for unbiased predictions (Odeh et al., 1994; Triantafilis et al., 2001). Additionally, a positive ME statistic indicates a prediction underestimation, whereas a negative ME suggests the prediction technique is an over-estimation (Triantafilis et al., 2001). Bias was also assessed by plotting residuals against predicted values for all models constructed.

Root mean square error was defined as:

$$\mathrm{RMSE} = \left\{\frac{1}{l}\sum_{j=1}^{l}\left[z(s_j) - z^*(s_j)\right]^2\right\}^{0.5} \tag{4}$$

RMSE is a measure of prediction accuracy, and is the mean square root of sum of squared prediction error (Odeh et al., 1995). Lower RMSE values indicate that predictions are increasing in precision (Odeh et al., 1994; Triantafilis et al., 2001).

Goodness-of-prediction (G) (Agterberg, 1984; Schloeder et al., 2001) was used as a relative measure of precision. For example, the G value provides a measure of how precise the prediction technique was relative to the precision of the sample mean. G was calculated using:

$$G = \left(1 - \left\{ \sum_{j=1}^{l} \left[z(s_j) - z^*(s_j) \right]^2 / \sum_{j=1}^{l} \left[z(s_j) - \bar{z} \right]^2 \right\} \right) \times 100 \quad (5)$$

where $\bar{z}$ is the sample mean (Schloeder et al., 2001). A positive G value indicates that the prediction technique performed better than the sample mean, whereas a negative G value indicates that the prediction technique had a poorer performance than the sample mean. A perfect prediction (with a RMSE of zero) would have a G statistic of 100%.

3.0 RESULTS

3.1 Partial least squares modelling

The number of factors chosen for both the 300 Index and Site Index PLS models was four in both cases, based on the minimum predicted residual sum of squares (PRESS). Table 3 lists the amount of variation accounted for by each of the four factors chosen, both individually and cumulatively. The total variance explained using four factor models was 58 and 67% for the 300 Index and Site Index, respectively.

Table 3: Percent variation accounted for using partial least squares for 300 Index and Site Index. Number of factors reduced to 4 and the smallest number with $p>0.05$.

	300 Index		Site Index	
Extracted factors	Dependent variables		Dependent variables	
	Current	Total	Current	Total
1	40.53	40.53	57.36	57.36
2	10.89	51.42	6.77	64.13
3	4.37	55.79	1.67	65.80
4	2.11	57.90	1.23	67.02

3.2 Regression kriging modelling

Exponential experimental (semi-)variogram models were chosen for both the 300 Index and Site Index data (Table 4). The 300 Index model has a small to moderate nugget, and a sill at around 7 km. Site Index has a smaller nugget and also attains a sill over a range of about 7 km, but the sill has half the height of that for the 300 Index (Figure 3).

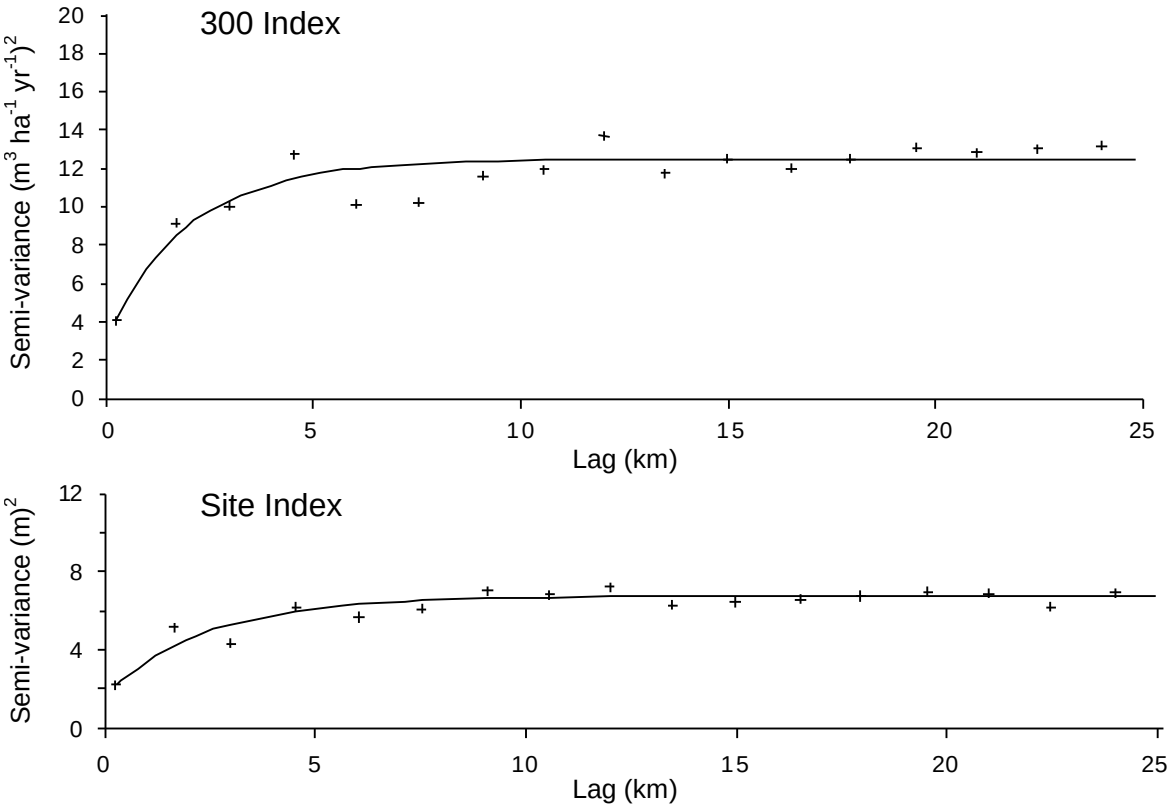


Figure 3: Semi-variograms of the PLS residuals for the 300 Index (upper) and Site Index (lower).

Table 4: Coefficients of the models fitted to the semi-variograms for the PLS model residuals for the 300 Index and Site Index.

Variable	Units	Model	Nugget	Sill	Range (m)	Lag (m)
300 Index	(m^3 ha^{-1} yr^{-1})	Exponential	3.0	9.5	2000	1500
Site Index	(m)	Exponential	1.8	5.0	2400	1500

3.3 *Pinus radiata* productivity surfaces for the 300 Index and Site Index models

Maps illustrating the geographic variation in the 300 Index and Site Index for New Zealand, modelled using the PLS technique, are shown in Figures 4 and 5, respectively. In general, both North Island productivity surfaces show higher values than their southern counterparts and values decrease at higher elevations. As discussed below, based on expert knowledge, generally, the modelled surfaces behaved as expected.

Land surface coverage in square kilometres and as a percentage of New Zealand's total land area developed from the PLS productivity surfaces can be seen in Table 5. These data confirm the North Island's higher 300 Index values, and the majority of land surface areas occur between 24 to 36 m^3 ha^{-1} yr^{-1}. In contrast, in the South Island, 20 to 32 m^3 ha^{-1} yr^{-1} values dominate. Site Index displays a similar trend with the 24 to 36 m classes and 20 to 28 m classes dominating in the North and South Islands, respectively.

Table 5: Land surface area in square kilometres and percentages for classes in the 300 Index and Site Index surfaces for North Island, South Island, and for New Zealand. 300 Index and Site Index surfaces were modelled using partial least squares regression. Where modelled, adjacent surrounding offshore islands were included, whereas lakes, major water bodies, and regions with mean annual temperature below 8.25 °C were excluded from analysis.

		North Island		South Island		New Zealand	
300 Index	(m^3 ha^{-1} yr^{-1})	km^2†	%*	km^2†	%*	km^2†	%*
	< 12	0	0.0	447	0.3	447	0.3
	12 - 16	22	0.0	1960	1.1	1982	1.1
	16 - 20	108	0.1	9012	5.1	9119	5.1
	20 - 24	2329	1.3	21767	12.2	24097	13.5
	24 - 28	15143	8.5	29907	16.8	45050	25.3
	28 - 32	47622	26.8	13076	7.4	60698	34.1
	32 - 36	31340	17.6	725	0.4	32065	18.0
	36 <	4392	2.5	1	0.0	4392	2.5
	Total	100955	56.8	76894	43.2	177850	100.0
Site Index	(m)						
	< 16	21	0.0	2405	1.4	2426	1.4
	16 - 20	107	0.1	8980	5.1	9087	5.1
	20 - 24	2318	1.3	21653	12.2	23971	13.5
	24 - 28	15092	8.5	29780	16.8	44871	25.3
	28 - 32	47471	26.8	12932	7.3	60403	34.1
	32 - 36	31206	17.6	719	0.4	31925	18.0
	36 <	4348	2.5	1	0.0	4349	2.5
	Total	100562	56.8	76470	43.2	177032	100.0

† km^2 of model surface
* Percentage of 300 Index and Site Index land surface areas were calculated using land surface area total derived from Palmer et al. (2008) as a base measurement.

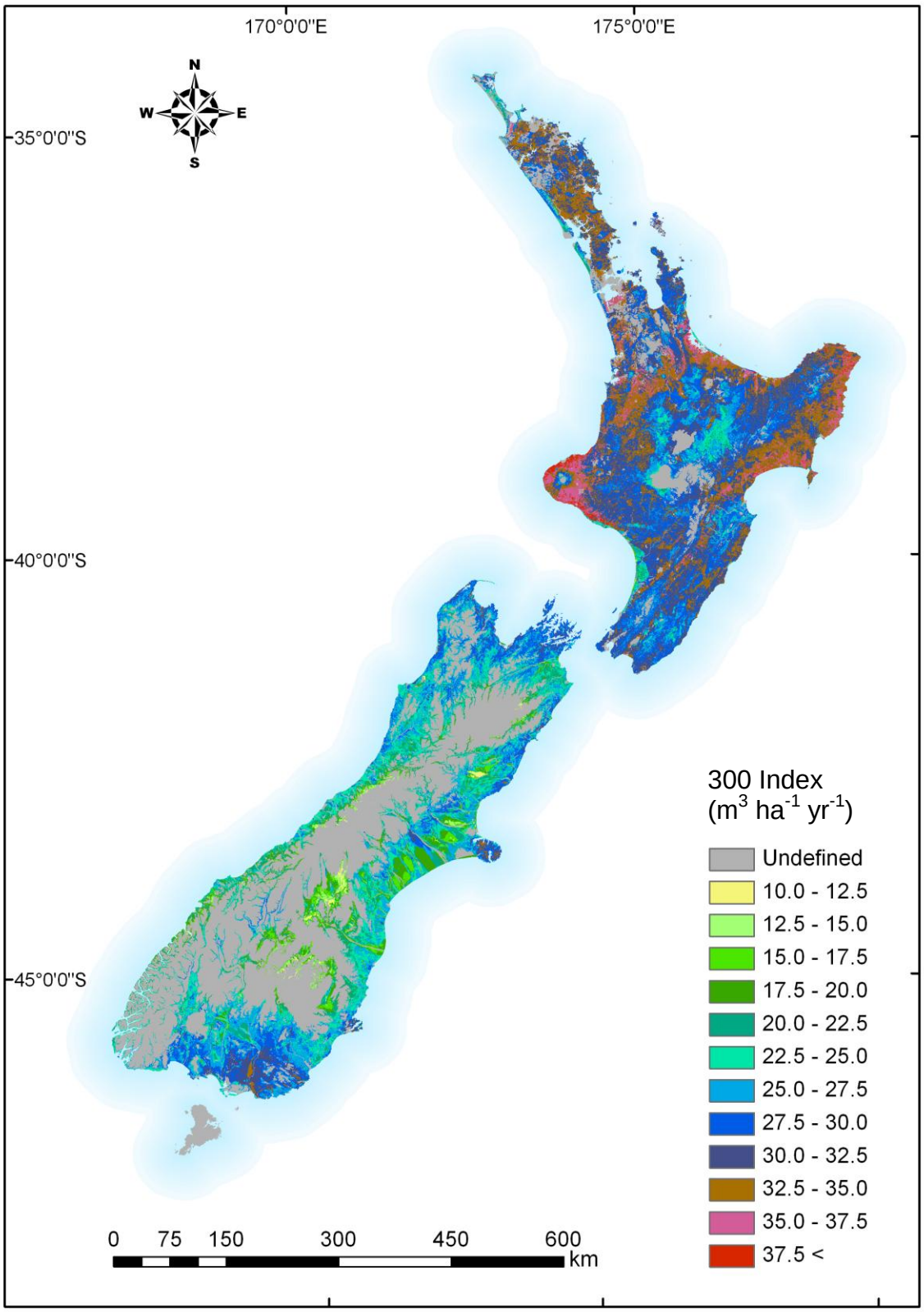


Figure 4: Distribution of the geographic variation of the 300 Index across New Zealand developed using partial least squares.

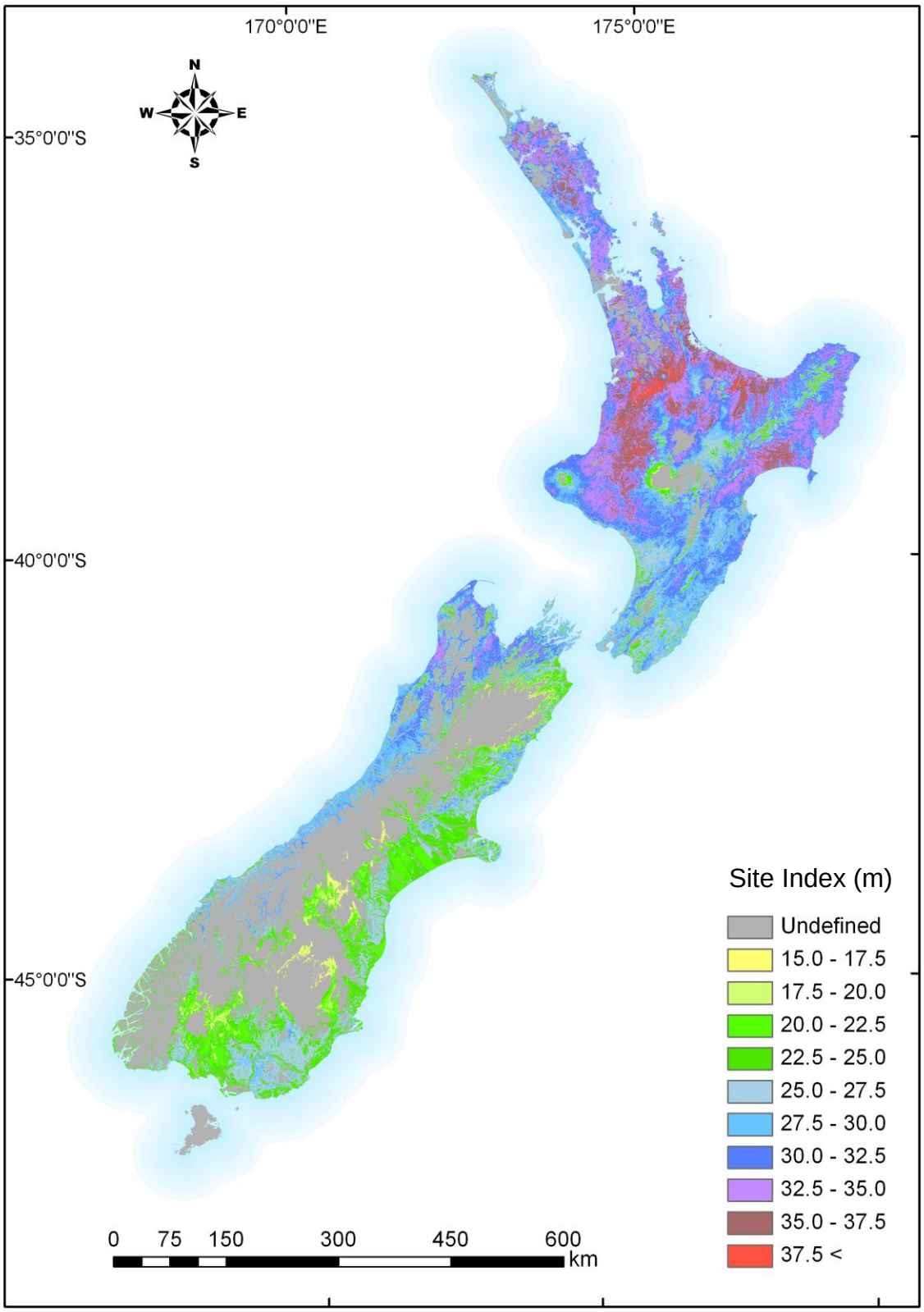


Figure 5: Distribution of the geographic variation of the Site Index across New Zealand developed using partial least squares regression.

3.4 Spatial prediction of productivity surfaces

The best PLS model formulated explained 58% of the variance in the 300 Index data. The Site Index model exceeded this, explaining 67% of the variance. Plots of residual and actual values against predictions exhibited little apparent bias for either 300 Index or Site Index (Figure 6).

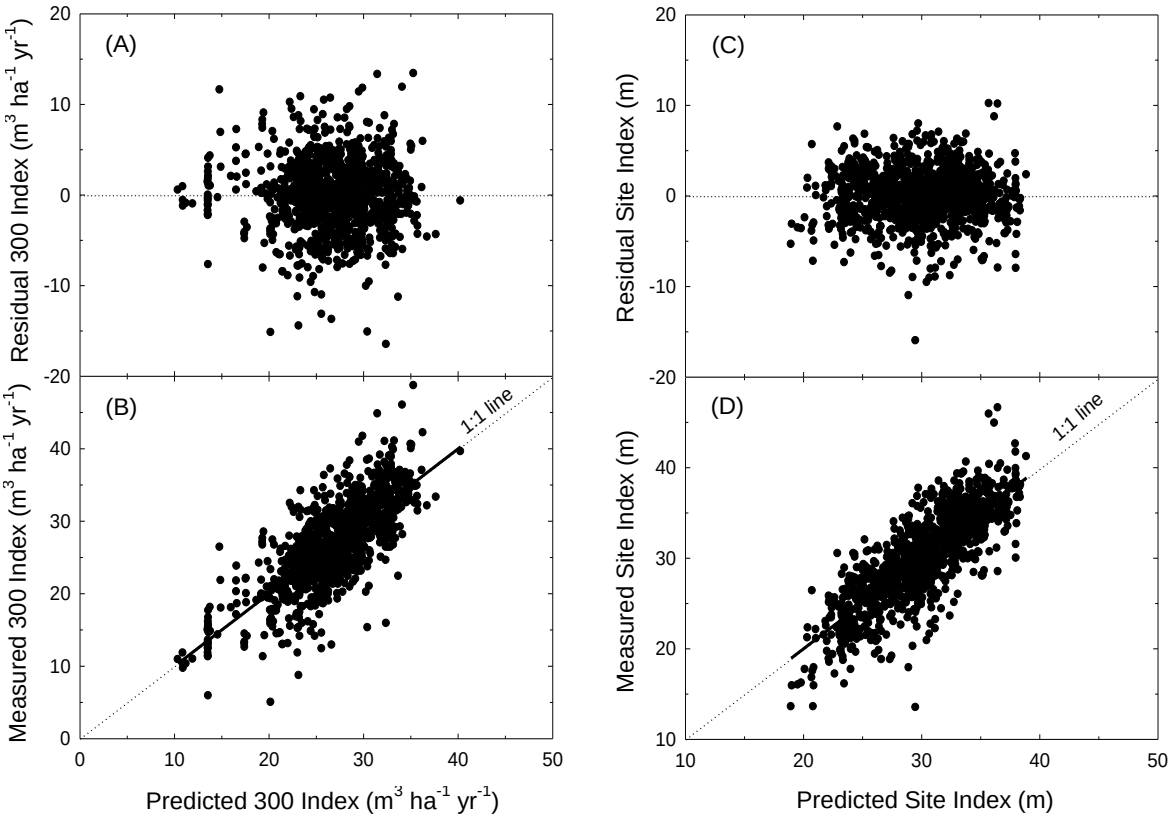


Figure 6. Plots of partial least squares regression for the model dataset showing the relationship between (A) predicted and residual 300 Index and (B) predicted and measured 300 Index, (C) predicted and residual Site Index and (D) predicted and measured Site Index. In (B) and (D) the 1:1 line is shown as a dotted line. Also shown as a thick line is a line of best fit (n = 1146).

3.5 Validation of productivity surfaces

Validation results show that RK substantially improved PLS regression predictions, increasing the R^2 values by 12 % and 13 % for Site Index and 300 Index, respectively (Table 6). The root mean square error (RMSE) and goodness-of-precision (G) cross validation statistics confirmed these findings with RK showing significant improvement for both the 300 Index and Site Index validation predictions. Furthermore, plots for the measured against predicted validation values for the 300 Index and Site Index also confirmed the substantial improvements by the RK technique (Figure 7).

Table 6: Validation statistics for the coefficient of determination (R^2), mean error (ME), root-mean-square error (RMSE), and goodness-of-prediction (G) statistics for 300 Index and Site Index prediction techniques.

Prediction techniques[1]	300 Index				Site Index			
	R^2	ME	RMSE	G	R^2	ME	RMSE	G
PLS	0.48	-0.39	4.22	48.03	0.59	-0.16	3.13	58.63
RK	0.61	-0.45	3.65	61.11	0.70	-0.10	2.65	70.35

[1] PLS refers to partial least squares regression, and RK is regression kriging.

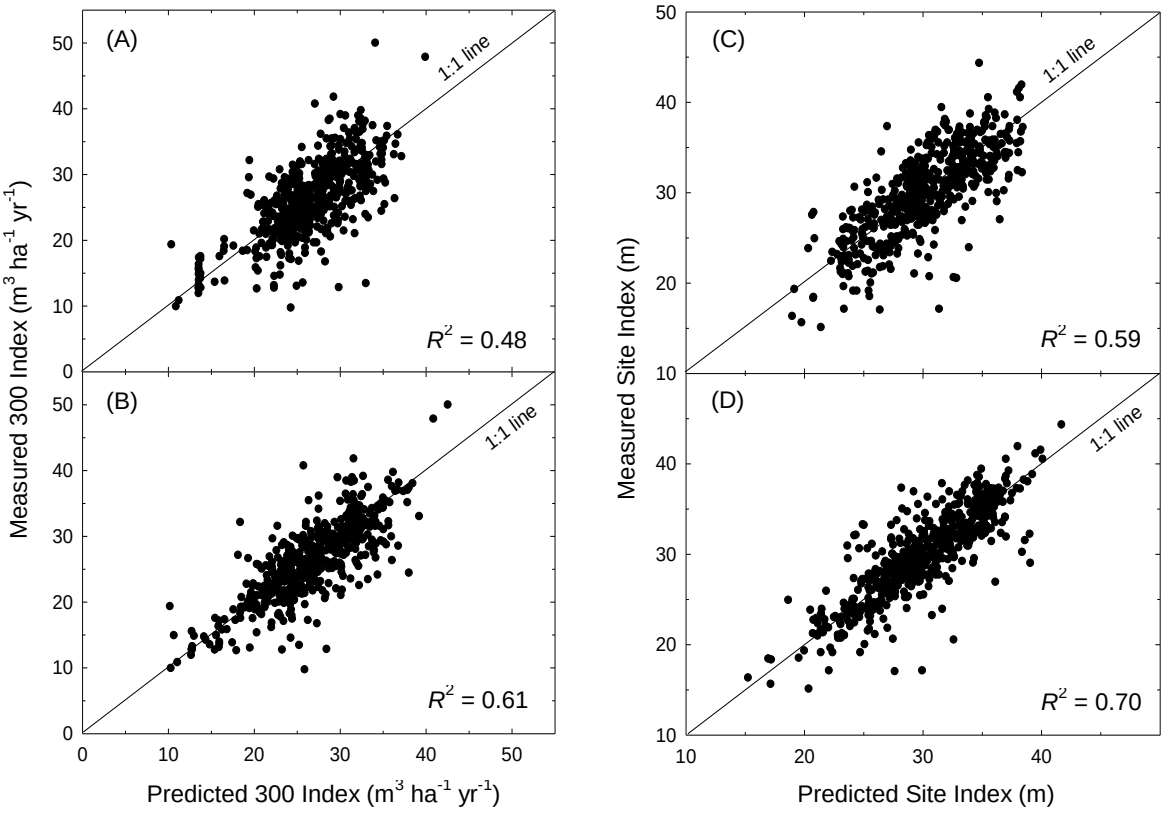


Figure 7. Validation plots of predicted verses measured for the 300 Index values using (A) partial least squares, (B) regression kriging, and for Site Index using (C) partial least squares, (D) regression kriging (n = 552).

3.6 Assessment of environmental drivers of productivity

Mean values for the measured and modelled (PLS) data grouped by LENZ environmental regions were closely correlated and showed similar spatial trends for both the 300 Index and Site Index data. To summarise, northern hill country (D), western and southern North Island lowlands (C), and central sandy recent regions (H) displayed the highest mean values for the 300 Index (Table 7). Conversely, the lowest 300 Index mean values were

derived for the eastern South Island plains (N) and the central well-drained recent soils (J). The ranking of mean values followed a broadly similar pattern for Site Index. Regions with the highest Site Index mean values were the western and southern North Island lowlands (C), followed by the central sandy recent soils (H). The lowest Site Index mean values were found in the eastern South Island plains (N) and southeastern hill country and mountain regions (Q).

The deviations from the mean for each general grouping within each environmental region are shown in Figures 8 and 9. For 300 Index, profile and water status groupings showed the greatest deviations from their means. Profile groupings with the highest positive deviations are the northern hill country (D) and the central hill country and volcanic plateau (F), whereas central well-drained recent soils (J) and the northern recent soils (G) have the greatest negative deviations from their respective means (Figure 8). For water status groupings, central sandy recent soils (H), central mountains (P), and the northern hill country (D) have the highest positive deviations, whereas the groupings eastern South Island plains (N) and the central well-drained recent soils (J) have the greatest negative deviations from their respective means.

Table 7: Mean values of measured and predicted (PLS fitted regression data) grouped by level 1 of the LENZ environmental regions for comparison (Leathwick et al., 2003). Brackets provide regional productivity rankings from lowest to highest, 1 to 14, respectively. Lowest and highest three regions are in **bold** typeface.

LENZ[1]	Description	Sample number	Mean 300 Index		Mean Site Index	
			Measured	Predicted	Measured	Predicted
A	Northern lowlands	57	27.3 (9)	28.5 (11)	31.9 (10)	31.8 (10)
B	Central dry lowlands	97	26.3 (8)	26.5 (9)	28.6 (6)	29.1 (7)
C	Western and southern North Island lowlands	12	29.5 (**12**)	29.8 (**13**)	34.4 (**14**)	33.8 (**14**)
D	Northern hill country	95	32.9 (**14**)	32.3 (**14**)	33.8 (**12**)	33.0 (**12**)
E	Central dry foothills	71	25.2 (7)	25.1 (7)	29.0 (9)	29.1 (8)
F	Central hill country and volcanic plateau	454	28.2 (10)	28.3 (10)	32.0 (11)	32.1 (11)
G	Northern recent soils	40	22.0 (**3**)	23.7 (4)	28.8 (8)	29.1 (6)
H	Central sandy recent soils	34	30.6 (**13**)	29.4 (**12**)	33.8 (**13**)	33.2 (**13**)
J	Central well-drained recent soils	28	19.7 (**2**)	18.0 (**2**)	25.0 (**3**)	24.5 (**3**)
L	Southern lowlands	24	28.5 (11)	25.5 (8)	27.1 (4)	25.0 (4)
N	Eastern South Island plains	37	17.3 (**1**)	16.8 (**1**)	22.7 (**1**)	23.7 (**1**)
O	Western South Island foothills and Stewart Island	27	22.2 (4)	22.6 (**3**)	27.6 (5)	28.8 (5)
P	Central mountains	59	23.8 (5)	24.2 (5)	28.8 (7)	29.5 (9)
Q	Southeastern hill country and mountains	104	24.6 (6)	24.8 (6)	24.4 (**2**)	24.2 (**2**)

[1]Refer to Leathwick et al. (2003) for class details

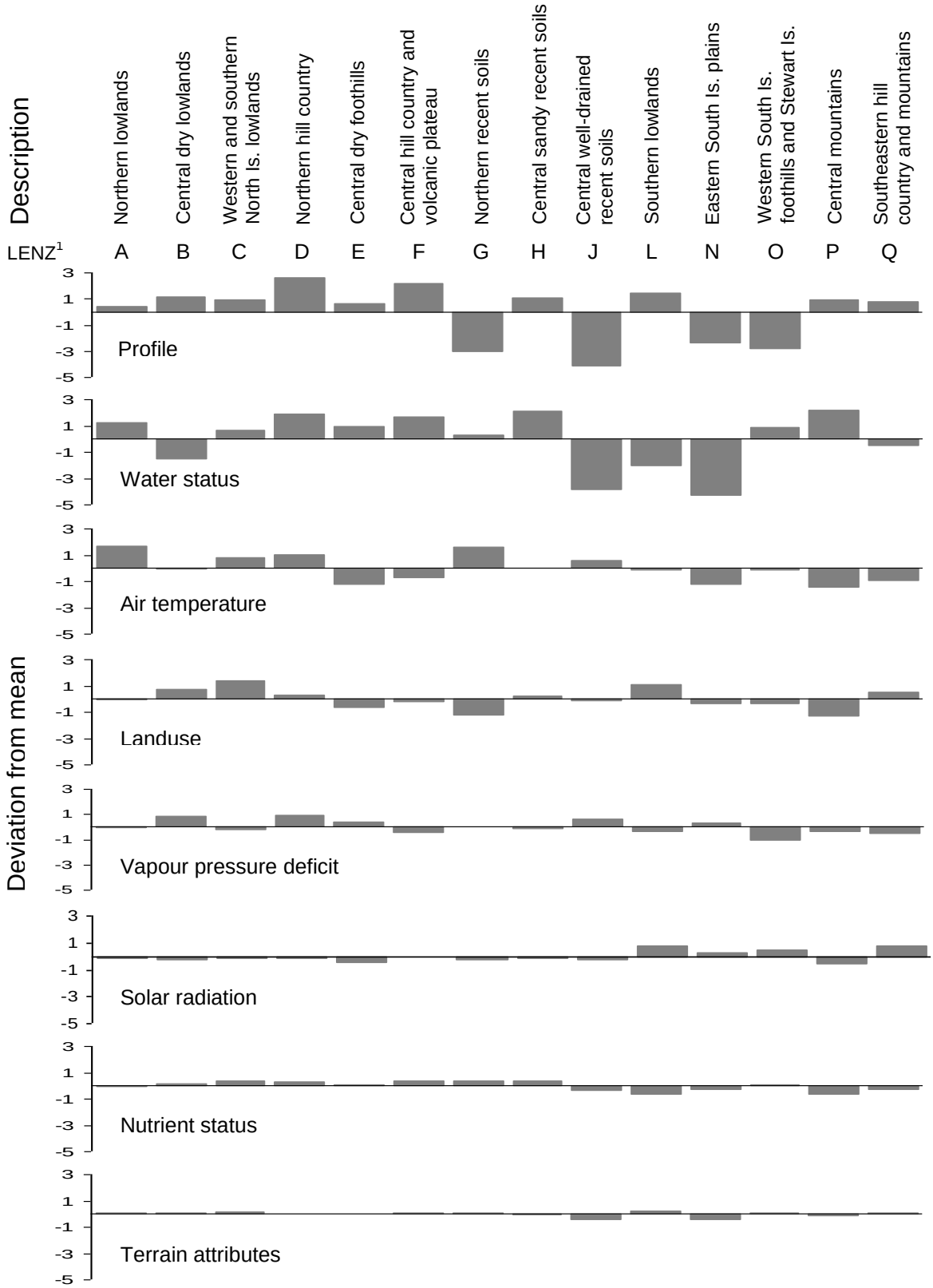


Figure 8: 300 Index deviation from the mean within regression coefficient groupings of profile, water status, air temperature, landuse, solar radiation, nutrient status, and terrain related variables separated into [1]LENZ (level 1) environmental regions for comparison (Leathwick et al., 2003).

Site Index mean values with the most pronounced deviations were air temperature, water status, and profile general groupings (Figure 9). Air temperature has a generally positive impact on northern regions, whereas southern regions remain generally negative. More specifically, northern lowlands (A), western and southern North Island lowlands (C), and the central sandy recent soils (H) have the greatest positive deviations in comparison to the southeastern hill country and mountains (Q) and the southern lowlands (L) that show the greatest negative deviations from their respective mean air temperatures. For water status, the western South Island foothills and Stewart Island (O), the central sandy recent soils (H), and the central mountain regions (P) have the greatest positive deviations, whereas the central well-drained recent soils (J), eastern South Island plains (N), and the central dry lowlands (B) have the greatest negative deviations from means. Profile groupings with the greatest positive deviations were the central hill country and volcanic plateau (F), central sandy recent soils (H), and the northern hill country regions (D). Conversely, regions with the strongest negative deviations from their means were the central well-drained recent soils (J), northern recent soils (G), and the eastern South Island plains (N).

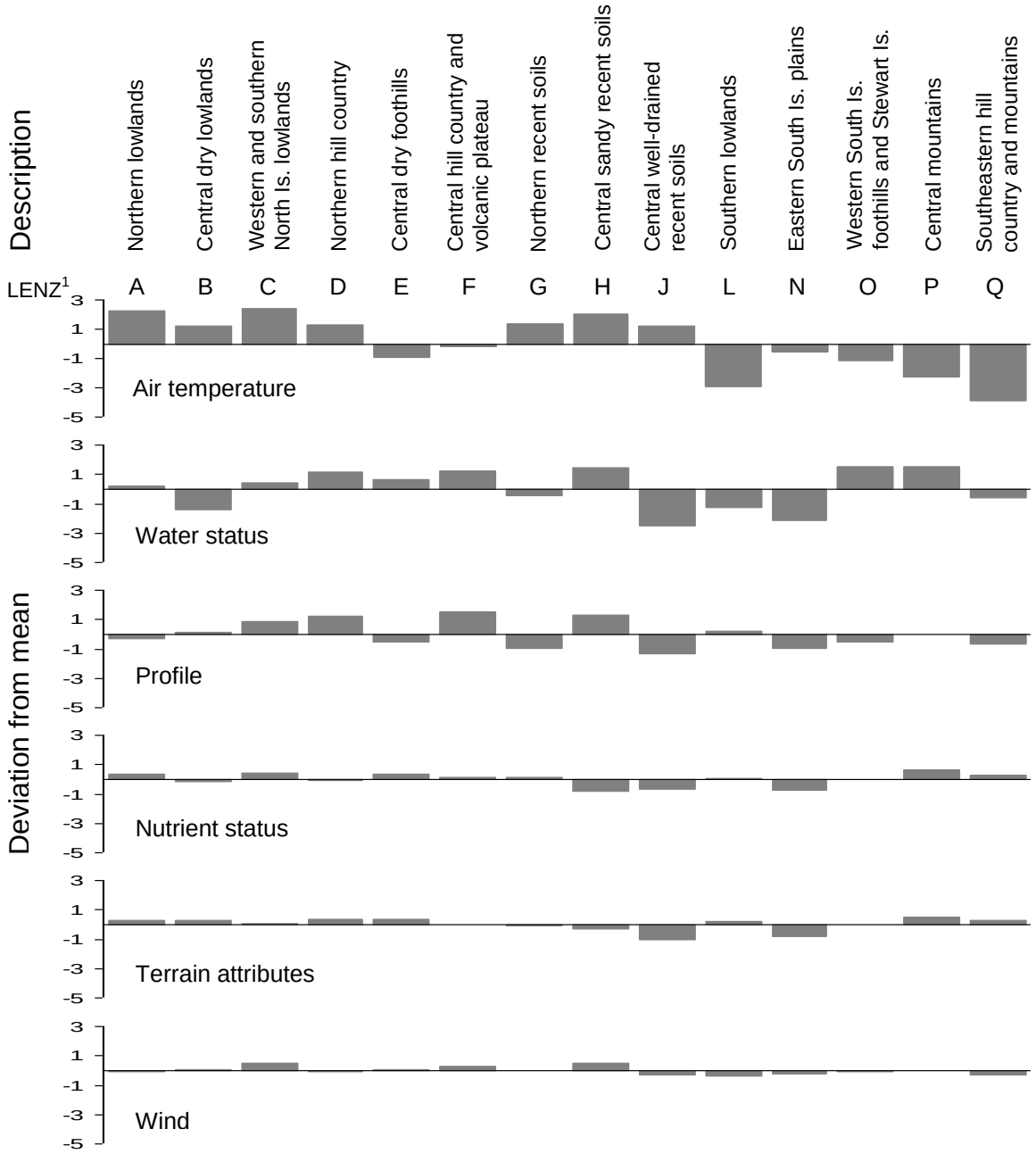


Figure 9: Site index deviation from the mean within regression coefficient groupings of air temperature, water status, profile, nutrient status, terrain, and wind related variables separated into [1]LENZ (level 1) environmental regions for comparison (Leathwick et al., 2003).

4.0 DISCUSSION

4.1 Spatial variability of the 300 Index and Site Index surfaces

The 300 Index and Site Index models and maps developed here are a major step forward in the visualisation and assessment of forest productivity. The maps are empirically based and clearly define model and map limitations while providing detailed information of spatial distribution across the national extent of New Zealand.

Generally, PSP locations are well represented across New Zealand with the exception of the South Island mountainous interior (Figure 1). To achieve best modelling results from our irregularly spaced datasets, regression transfer models (Burrough and McDonnell, 1998) provided predictions for the 300 Index and Site Index using PLS. Overall, expert knowledge suggests that modelled values for both the 300 Index (Figure 4) and Site Index (Figure 5) followed expected trends with the respective coefficients of determination providing estimates of model prediction outcomes (Figure 6). From a visual assessment of the 300 Index map, the region of highest *P. radiata* productivity is Taranaki on the North Island's west coast (Figure 4). Conversely, some of the lowest 300 Index values occur along the South Island's Canterbury Plains. In general, mapped trends indicate higher productivity along warmer coastal regions, with decreasing values along the cooler mountainous interior. Some exceptions are the lower elevation coastal regions and sandy soils (e.g., Woodhill forest, northwest of Auckland), and the south-western plains of the Rangitikei and Manawatu rivers, where lower productivity prevailed. A possible explanation for such productivity is the limiting nutrient and moisture status of

associated parent materials (for example, McMurtrie et al., 1990). Relatively high 300 Index values occur in cool Southland sites, but Site Index values are comparatively low. These differences strongly suggest that temperature is important for *P. radiata* plantation height (Site Index) but not as important for plantation volume (300 Index).

The impact of temperature on Site Index is very evident with values decreasing dramatically along the cooler mountainous interior of the North and South islands (Figure 5). Generally, Site Index values decrease with lower latitudes. Lower Site Index values in the very north of the North Island could be explained by the impact of summer and annual water deficits (see SwatBal, Chapter 3). Conversely, elevated Site Index values for the western margins of the South Island could be explained by adequate water storage. North Island Site Index values follow a broadly similar spatial trend to 300 Index with the exception of lower values around the Taranaki region. These elevated 300 Index values could be associated with the improved pasture and high landuse classes (Table 1) for the Taranaki region that feature strongly in the 'landuse' groupings (Figure 8). In contrast, Site Index tree height values for Southland are relatively low compared with volume suggesting these trees are shorter but still produce reasonably high volumes.

As with most models there were areas inconsistent with expectations – for example, the 300 Index map locations where mean annual temperature values are below 8.25 °C (grey mask, Figure 4). Expert knowledge suggests productivity values should reduce to about 12 to 15 m^3 ha^{-1} yr^{-1} at around 900 m elevations (refer to Chapter 4). Under this premise, the central ranges across the North Island had higher than expected values. For the South

Island, expectations generally held true. On closer examination of PSP locations at higher elevations, North Island observations remained relatively high, which could explain the model predictions. Consequently, a lower limit mask of 8.25 °C mean annual temperature was placed on the index maps. This threshold was chosen because temperature variables are major model contributors, and few PSP observations occur below this temperature. Nevertheless, the mask should not limit the utility of the model because plantations are unlikely to be established below this temperature. Without research focussing on and exploring these cooler regions and assessing growth of the *P. radiata* species in these regions, productivity predictions below this temperature regime will remain uncertain.

4.2 Generalised assessment of model contributors

For this study, PLS was chosen for its ability to manage multiple independents even when predictors displayed multicollinearity. The trade off is that PLS remains less than satisfactory in its ability to filter out variables of minor causal importance (Tobias, 1997), and greater difficulty is experienced when interpreting independent variable loadings. As a meaningful assessment of model contributors, model coefficients were used to calculate the influence of independent variables by grouping them into common themes (Tables 1 and 2), and averaging their respective predictions within the LENZ environmental regions (Figure 2).

For the 300 Index, the grouping ‘profile’ showed the highest degree of deviation from environmental regions with variables that included soil class, parent materials and induration, particle size and profile readily available water depth (Table 1). The environmental regions – northern recent soils (G), central well drained recent soils (J), eastern South Island plains (N), and western South Island foothills and Stewart Island (O) – illustrated that the group of variables termed ‘profile’ had a strong negative impact on forest productivity (Figure 8). Notably, these regions included Woodhill Forest, coastal regions of the Manawatu, and the Canterbury Plains, which previously we noted as having lower levels of productivity. Conversely, the environmental regions where the grouping ‘profile’ had the greatest positive deviations from the mean were the northern (D) and central hill country and the volcanic plateau (F). More specific determinations of independent variable contribution within the ‘profile’ general grouping were not possible.

The grouping ‘water status’ for the 300 Index, as expected, highlighted drier regions as having negative deviations which included the eastern South Island plains (N), the central well-drained recent soils (J), and the central dry lowlands (B). These regions are considered highly correlated with seasonal soil water deficits displayed in SWatBal (Chapter 3). The opposite is generally true for environmental regions with positive deviations from mean water status. These regions with wetter climates include central sandy recent soils (H), and the northern hill country (D).

The grouping for ‘air temperature’ also performed as expected with regions with low air temperatures and subject to frosts having a negative influence on productivity, whereas

warmer regions tended to be increasingly positive. The central dry foothills (E), the eastern South Island plains (N), the central mountains (P), and to a lesser extent the central hill country and volcanic plateau (F) regions showed strong negative deviations from means. Conversely, the northern lowlands (A), and the northern recent soils (G) had positive deviations.

Site Index variables and their respective groupings were considerably fewer in number when compared with the 300 Index groupings. The grouping 'air temperature' stands out as having the greatest influence on Site Index values. 'Water status' and to a lesser extent 'profile' variables also play an influential role in some environmental regions. The northern regions, with the exception of the cooler central dry foothills (E), and the hill country and volcanic plateau (F), showed favourable conditions for height growth, whereas southern environmental regions were negative.

Water status, primarily consisting of water deficit variables, illustrated the negative impact on height growth at dryland regions including the central well-drained recent soils (J), the eastern South Island plains (N), and to a lesser extent the central dry lowlands (B). These regions all correspond with locations that displayed annual and seasonal water deficits as determined by SWatBal (Chapter 3).

The Site Index 'profile' groupings for the central hill country and volcanic plateau (F), the central sandy recent soils (H), and the northern hill country (D) have the greatest

positive deviations from the mean. Regions with the greatest negative deviations were the central well-drained recent soils (J) and the northern recent soils (G).

4.3 Cross validation

Map development with global interpolation and the estimation of long range variation in forest productivity were undertaken using PLS regression. This technique allows the exploration of the functional relationships between spatial variables and the property of interest (forest productivity), sometimes referred to as transfer functions (Bouma and Bregt, 1989). The technique has advantages when sample observations are numerous, but spatial representation is sometimes sparse, and irregular, as in this study. The main advantages over other interpolation techniques are as follows: (1) the coefficient of determination (R^2) provides an estimation of the prediction certainty (variance explained); and (2) the technique is not reliant on the separation distance between observations. Indeed, multiple linear regression models (MLR), including PLS, implicitly disregard spatial covariation as well as lacking sensitivity to measurement error (Odeh, 1995). These attributes give PLS models the advantage of using local environmental and landform information not always available with other interpolation techniques. The variance explained for the 300 Index and Site Index maps developed here were considered high when compared with other New Zealand *P. radiata* productivity studies. The coefficient of determination (R^2) indicated that 58% and 67% of the variance was explained for the 300 Index and Site Index, respectively.

Ordinary kriging of the regression residuals from the PLS models showed a substantial improvement in the coefficient of determination (R^2), mean error (ME), root-mean-square error (RMSE), and goodness-of-prediction (G) statistics for both Site Index and the 300 Index (Table 6). The only exception was the ME statistic for the 300 Index which suggested that PLS predictions were less biased than regression kriging predictions. Negative ME statistics also suggested that the model predictions for Site Index and 300 Index were to some degree over estimated. However, overall it was found that the RK technique substantially improved model predictions at the validation locations assessed.

For forestry management, consideration should be given to both the PLS and regression kriging outcomes. Clearly, the range over which the (semi-)variogram operates in relation to data observation density and spatial locations is important. Figure 3 shows an exponential model fitted to the 300 Index regression residuals, having a small to moderate nugget effect, and sill attainment at around 7 km. Site Index also reached its sill over a range of about 7 km, but with a sill half the height of that for the 300 Index. This relationship is also true of the nugget, which is half that for the 300 Index. These results imply that there is no spatial dependence between observed data points beyond ~7 km and that all estimates of variances of differences will be invariant with sample separation distance (Burrough and McDonnell, 1998).

For the forester with a single observation representing a forest stand, many kilometres from surrounding sample observations, kriging of the PLS regression residuals will be of little benefit to predictions. Conversely, observations within the bounds of the (semi-

)variogram model will provide improvement to the 300 Index or Site Index predictions. Where productivity observations are relatively dense such as in the central North Island plateau and the Nelson regions, RK is the preferred prediction technique. However, the RK technique will provide little advantage for forest managers or scientists interested in mapped locations where observational data are sparse. In this situation the PLS model will provide an optimal prediction performance without the added cost of RK.

Our study explained less of the variance in productivity compared with that determined by Watt et al. (2005) where model development was from a limited dataset (n = 33) and *P. radiata* volumes were estimated from high stand densities grown over short periods (~4 years). The measured variables phosphorus (P) and nitrogen (N) used in the Watt et al. (2005) study were not directly available for this project, but remained as indirect or soft information datasets. For example, soil pH, subsoil acid soluble P, vegetation cover (VCMNZ), and land use class (NZLUC), were the only variables likely to be related to soil P and N. Furthermore, the strongest linkage was indirectly through Class One of the VCMNZ and NZLUC (Table 2) datasets indicating intensive farming practices at these locations (and thus likely high N and P fertiliser applications). The PLS models used foliar N and P from the nutritional atlas (Hunter et al., 1991). However, incomplete national coverage meant abandoning these explanatory variables. It is clear that future productivity models will benefit from the use of national extent soil and foliar N and P datasets.

For Site Index the model was considerably stronger with 67% of the variance explained. Site Index maps (Figure 5) from this study were compared with a site index ranking map (Eyles, 1986), and some visual similarities are evident. For example, a decreasing trend from north to south is evident with most North Island locations exhibiting values above 25, whereas South Island values decrease to below 25. The main differences between the models remain in the spatial detail, and the map by Eyles (1986) does not account for cooler regions, other than in more generic terms.

The main advance of our work is not only that models were developed using observed data from currently growing plantation stands, but also in the production of maps defining the spatial variability of the 300 Index and Site Index across all of New Zealand. The models and maps developed here provide a significant step forward in the visualisation and assessment of forest productivity. As intended, our study was empirically based and clearly defined model and map limitations through model validation, while providing spatial distribution information across New Zealand.

5.0 SUMMARY AND CONCLUSIONS

PLS regression was used to develop models of mean annual increment volume (300 Index) and mean top height at age 20 (Site Index) by PLS regression using data derived from national extent ancillary maps and interpolated surfaces. Forestry maps were developed using the best models, and explained 58% and 67% of the variance for the 300 Index and Site Index, respectively. Where possible, drivers of *P. radiata* productivity were assessed across New Zealand environmental regions (derived from LENZ), highlighting the importance of air temperature, soil water deficits and profile properties. Model predictions were improved through RK in a process where regression residuals were kriged before summation with the PLS modelled surfaces. Observed data withheld for the test set validation procedure when developing the PLS models were used for assessing prediction bias and precision using cross validation procedures. Overall, RK provided a substantial improvement over PLS in prediction and precision for both surfaces. To conclude, we successfully developed models and maps to provide forest managers and scientists with tools for exploring the spatial relationships of *P. radiata* productivity across New Zealand.

6.0 ACKNOWLEDGEMENTS

We gratefully acknowledge and thank Carolyn Anderson and Mark Dean for their assistance in obtaining permissions from the forestry companies involved and for the PSP data extraction. We are also indebted to numerous forestry companies and private owners for supporting this research. David Palmer wishes to thank and acknowledge the Foundation for Research Science and Technology (FRST) and the University of Waikato (Doctoral Scholarship) for their funding. Barbara Höck thanks FRST for funding (Protecting and enhancing the environment through forestry CO4X0304). We also acknowledge the helpful comments provided by anonymous referees.

7.0 REFERENCES

Abdi, H., 2003. Partial least squares (PLS) regression. In: Lewis-Beck, M., Bryman, A., Futing, T., (Eds.), Encyclopaedia of social sciences research methods. Thousand Oaks, Sage.

Agterberg, F.P., 1984. Trend surface analysis. *In*: G.L. Gaile and C.J. Willmott (Eds.), Spatial Statistics and Models. Reidel, Dordrecht, Netherlands, pp. 147-171.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998a. Assessment of annual carbon exchange in a water-stressed *Pinus radiata* plantation: an analysis based on eddy covariance measurements and an integrated biophysical model. Global Change Biology 5, 531-545.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998b. Fluxes of carbon and water in a *Pinus radiata* forest subject to soil water deficit. Australian Journal of Plant Physiology 25, 557-570.

Bouma, J., Bregt, A.K., (Eds), 1989. Land quantities in space and time. Proceedings of a symposium organized by the International Society of Soil Science (ISSS), Wageningen, the Netherlands 22-26 August 1988. Wageningen: Pudoc. 352 pp.

Box G.E.P., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistics Society, B-26, 211-252.

Bui, E.N., Moran C.J., 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. Geoderma 103, 79-94.

Bui, E.N., Moran C.J., 2003. A strategy to fill gaps in soil survey over large spatial extents: an example from the Murray-Darling basin of Australia. Geoderma 111, 21-44.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographic Information Systems. Oxford University Press, Oxford, 333 pp.

Czarnowski, M.S., Humpreys, F.R., Gentle, S.W., 1971. Quantitative expression of Site-Index in terms of certain soil and climate characteristics of *Pinus radiata* D. Don plantations in Australia and New Zealand. Ekologia Polska 19, 295-309.

Eyles, G.O., 1986. *Pinus radiata* site index rankings for New Zealand. New Zealand Forestry 31, 19-22.

Gotway, C.A., Ferguson, R.B., Hergert, G.W., Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Science Society of America Journal, 60: 1237-1247.

Goulding, C.J., 2005. "Measurement of Trees"; Section 6.5 of the NZIF Forestry Handbook, 4th Edition, Mike Colley (Ed.), NZIF. 318 pp.

Henderson, B., Bui, E., Moran, C., Simon, D., 2005. Australia-wide predictions of soil properties using decision trees. Geoderma, 124: 383-398.

Hewitt, A.E. 1998. New Zealand Soil Classification (Second edition). Landcare Research Science Series No. 1. 133 pp.

Höck, B.K., Payn, T.W., Shirley, J.W., 1994. Using a geographic information system and geostatistics to estimate site index of *Pinus radiata* for Kaingaroa Forest, New Zealand. New Zealand Journal of Forestry Science 23, 264-277.

Hunter I.R., Gibson, A.R., 1984. Predicting *Pinus radiata* site index from environmental variables. New Zealand Journal of Forestry Science 14, 53-64.

Hunter, I.R., Rodgers, B.E., Dunningham, A., Prince, J.M., Thorn, A.J., 1991. An Atlas of Radiata Pine Nutrition in New Zealand. Forest Research Bulletin No. 165, Forest Research Institute, Rotorua, New Zealand, 24 pp.

Jackson, D.S., Gifford, H.H., 1974. Environmental variables influencing the increment of radiata pine (1) periodic volume increment. New Zealand Journal of Forestry Science 4**,** 3-26.

Kimberley, M.O., West, G., Dean, M., Knowles, L., 2005. Site Productivity: The 300 Index - a volume productivity index for radiata pine. New Zealand Journal of Forestry 50, 13-18.

Landcare Research, 2008. S-map project: A cooperative soil information system for New Zealand. Landcare Research, http://www.landcareresearch.co.nz.

Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K., 2002a. Land Environments of New Zealand: A technical guide. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 237 pp.

Leathwick, J., Wilson, G., Rutledge, D., Wardle, P., Morgan, F., Johnston, K., McLeod, M., Kirkpatrick, R., 2003. Land Environments of New Zealand. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 184 pp.

Leathwick, J.R., Wilson, G., Stephens, R.T.T., 2002b. Climate surfaces for New Zealand. Landcare Research contract report: LC9798/126. Landcare Research, Hamilton, New Zealand, 22 pp.

Lewis, N.B, Ferguson I.S., 1993. Management of radiata pine. Inkata Press, Melbourne. 404 pp.

McCool, D.K.; Brown, L.C.; Foster, G.R.; Mutchler, C.K.; Meyer, L.D. 1987. Revised slope steepness factor for the Universal Soil Loss Equation. Transactions of the American Society of Agricultural Engineers, 30: 1387-1396.

McCool, D.K.; Foster, G.R.; Mutchler, C.K. 1989. Revised slope length factor for the Universal Soil Loss Equation. Transactions of the American Society of Agricultural Engineers, 32: 1571-1576.

McMurtrie, R.E., Rook, D.A., Kelliher, F.M. 1990. Modelling the yield of *Pinus radiata* on a site limited by water and nitrogen. Forest Ecology and Management, 30, 381-413.

Mitchell, N.D., 1991. The derivation of climate surfaces for New Zealand, and their application to the bioclimatic analysis of the distribution of kauri (*Agathis australis*). Journal of the Royal Society of New Zealand 21, 13-24.

National Water and Soil Conservation Authority, 1975-70. New Zealand Land Resource Inventory Worksheets, 1:63,360. National Water and Soil Conservation Organisation, Wellington.

Newsome, P.F.J., 1987. The Vegetative Cover of New Zealand. Soil Conservation Centre, Aokautere, Ministry of Works and Development. Water and Soil Miscellaneous Publication No 112, Wellington, New Zealand.

Newsome, P.F.J., Wilde, R.H., Willoughby, E.J., 2000. Land resource information system spatial data layers. Landcare Research, Palmerston North, New Zealand, 84 pp.

New Zealand Soil Bureau, 1973. Map of parent rocks of New Zealand soils. 1: 1,000,000. 2 sheets. New Zealand Soil Survey Report 5.

NZFOA, New Zealand Forest Owners Association, 2005. New Zealand Forest Industry Facts and Figures 2004/2005. New Zealand Forest Owners Association, P.O. Box 1208, Wellington.

Odeh I.O.A., McBratney, A.B., Chittleborough, D.J., 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63, 197-214.

Odeh I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215-226.

Palmer, D.J., Höck, B.K., Dunningham, A.G., Lowe, D.J., Payn, T.W., 2008. Developing national-scale terrain attributes for New Zealand (TANZ). Scion Research Bulletin, Rotorua, New Zealand (in press).

Palmer, D.J., Lowe, D.J., Payn, T.W., Höck, B.K., McLay C.D.A., Kimberley, M.O., 2005. Soil and foliar phosphorus as indicators of sustainability for *Pinus radiata* plantation forestry in New Zealand. Forest Ecology and Management 220, 140-154.

Payn, T.W., Hill, R.B., Höck, B.K., Skinner, M.F., Thorn, A.J., Rijkse, W.C., 1999. Potential for the use of GIS and spatial analysis techniques as tools for monitoring changes in forest productivity and nutrition, a New Zealand example. Forest Ecology and Management 122, 187-196.

Richardson, B., Whitehead, D., McCracken, I.J. 2002. Root-zone water storage and growth of *Pinus radiata* in the presence of a broom understorey. New Zealand Journal of Forestry Science, 32, 208-220.

Romanyà, J., Vallejo, V.R., 2004. Productivity of *Pinus radiata* plantations in Spain in response to climate and soil. Forest Ecology and Management 195, 177-189.

Sánchez-Rodríguez, F., Rodríguez-Soalleiro, R., Español, E., López, C.A., Merino, A., 2002. Influence of edaphic factors and tree nutritive status on the productivity of *Pinus radiata* D. Don plantations in northwest Spain. Forest Ecology and Management 171, 181-189.

Schlatter, J.E., Gerding, V.R., 1984. Important site factors for *Pinus radiata* growth in Chile. In: Grey, D.C., Schonau, A.P.G., Schutz, C.J. (Eds.), Proceedings, IUFRO Sypmosium on Site and Productivity of Fast Growing Plantations. Pretoria and Pietermaritzburg, South Africa, 30 April–11 May. 1984. ISBN 0 621 08513 8. pp. 541–550.

Schloeder, C.A, Zimmerman, N.E., Jacobs, M.J., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Science Society of America Journal, 65: 470-479.

Tait, A., Henderson, R., Turner, R., Zheng, X. 2006. Spatial interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26, 2097-2115.

Tegelmark, D.O., 1998. Site factors as multivariate predictors of the success of natural regeneration in Scots pine forests. Forest Ecology Management 109, 231-239.

Tobias, R.D., 1997. An introduction to partial least squares regression. Cary, SAS Institute, 7 pp. http://ftp.sas.com/techsup/download/technote/ts509.pdf.

Triantafilis, J., Odeh, I.O., McBratney, A.B., 2001. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. Soil Science Society of America Journal 65, 869-878.

van der Voet, H., 1994. Comparing the predictive accuracy of models using a simple randomization test. Chemometrics and Intelligent Laboratory Systems 25, 313-323.

Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science, 41: 473-490.

Watt, M.S., Coker, G., Clinton, P.W., Davis, M., Parfitt, R., Simcock, R., Garrett, L., Payn, T., Richardson, B., Dunningham, A., 2005. Defining sustainability of plantation forests through identification of site quality indicators influencing productivity - a national view for New Zealand. Forest Ecology and Management 216, 51-63.

Watt, M.S., Davis, M.R., Clinton, P.W., Coker, G., Ross, C., Dando, J., Parfitt, R.L., Simcock, R., 2008. Identification of key soil indicators influencing plantation productivity and sustainability across a national trial series in New Zealand. Forest Ecology and Management (in press).

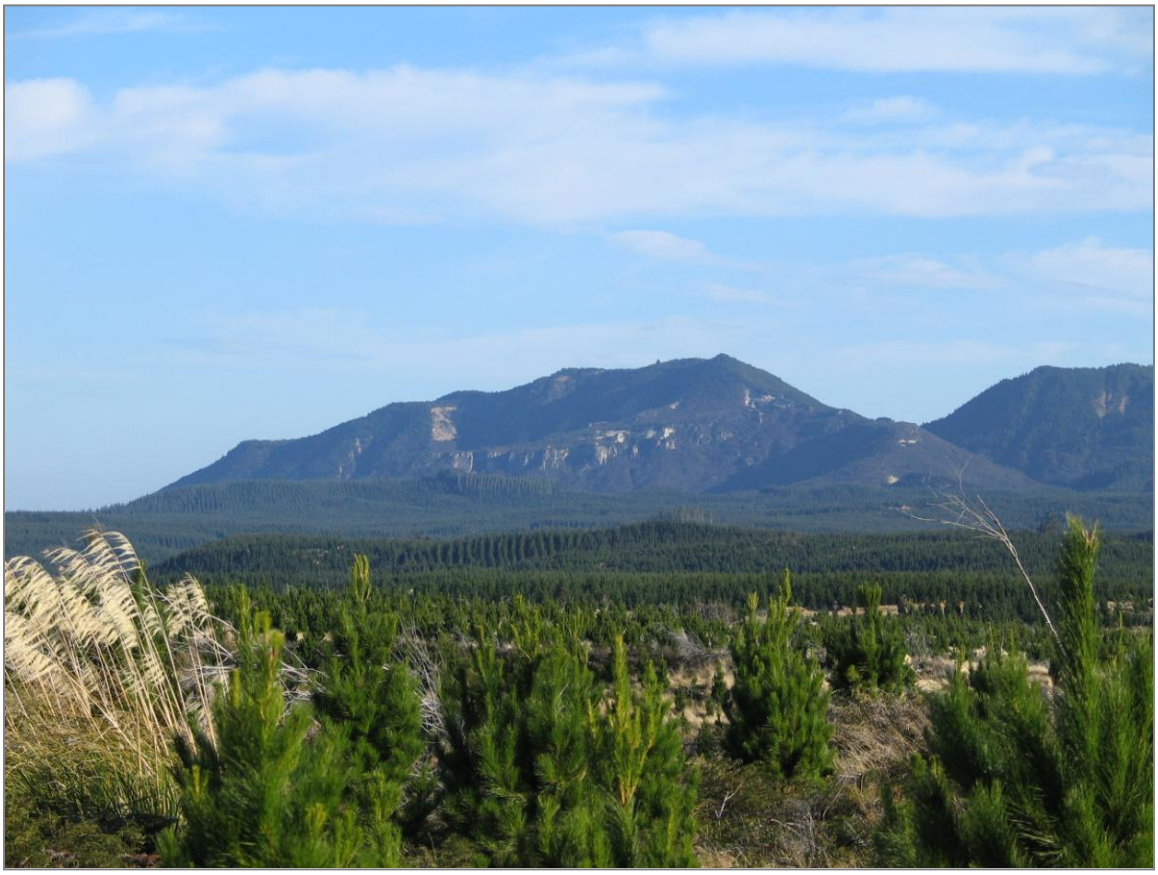
Watt, M.S., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C., 2003. Modelling the influence of weed competition on the growth of young *Pinus radiata* at a dryland site. Forest Ecology and Management, 178, 271-286.

Wilson, J.P., Gallant, J.C. 2000. Digital Terrain Analysis. In: Wilson, J.P., Gallant, J.C. (eds), Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: pp. 1-28.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares. In P. R. Krishnaiah (Eds.), Multivariate Analysis. New York, Academic Press, pp. 391-420.

Wold, S., 1994. PLS for Multivariate Linear Modeling, QSAR: Chemometric Methods in Molecular Design. Methods and Principles in Medicinal Chemistry, (Eds.), H. van de Waterbeemd, Weinheim, Germany, Verlag-Chemie.

Woollons, R.C., Skinner, M.F., Richardson, M.F., Rijske, W.C., 2002. Utility of "A" horizon soil characteristics to separate pedological groupings, and their influence with climatic and topographic variables on *Pinus radiata* height growth. New Zealand Journal of Forestry Science 32, 195-207.



Pinus radiata plantations across the Castle Rocks plateau in Hawkes' Bay. View towards the Kaweka Forest with Mount Miroroa and Kohinga centre and right, respectively (photograph by Dr Kyle Bland)

Truth lies within a little and certain compass, but error is immense.

Henry St John, *Reflections on Exile*

CHAPTER SIX

A COMPARISON OF SPATIAL PREDICTION TECHNIQUES FOR DEVELOPING *PINUS RADIATA* PRODUCTIVITY SURFACES ACROSS NEW ZEALAND

D.J. Palmer[a*], B.K. Höck[b], M.O. Kimberley[b], D.J. Lowe[a], T.W. Payn[b]

[a]*Department of Earth and Ocean Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand 3240*

[b] *Scion, Private Bag 3020, Rotorua, New Zealand*

[*]Corresponding author. Tel.: +64-7-8562889 *E-mail address*: djp8@waikato.ac.nz

Prepared for submission to: New Zealand Journal of Forestry

ABSTRACT

Spatial interpolation is frequently used to predict values at unvisited sites enabling the spatial variation and patterns of a property across a landscape to be quantified. Inverse distance weighting (IDW), ordinary kriging (OK), regression kriging (RK), and partial least squares (PLS) regression are interpolation techniques typically used where the region of interest's spatial extent is relatively small and observations are numerous and regularly spaced. In the current era of data 'mining' and utilisation of sparse data, the above criteria are not always fully met, increasing model uncertainties. Furthermore, regression modelling and kriging techniques require good judgement, experience, and expertise by the practitioner compared with IDW with its more rudimentary approach. In this study we compared spatial predictions derived from IDW, PLS, RK, and OK for *Pinus radiata* volume mean annual increment (referred to as 300 Index) and mean top height at age twenty (referred to as Site Index) across New Zealand using cross validation techniques. Our results showed that for the 300 Index measure of productivity, the OK technique consistently gave the greatest accuracy and most unbiased predictions. Nevertheless, RK and IDW techniques provided similar results. For Site Index predictions, RK performed with the greatest precision followed by OK, IDW, and PLS. However, OK still retained the most unbiased predictions. RK predictions across large spatial extents have the added advantage of combining statistical and geostatistical information, which together provide insight not available without the regression model component. Maps illustrating the spatial variation of *P. radiata* forest productivity are provided.

Keywords: plantation forestry, regression kriging, ordinary kriging, inverse distance weighting, spatial modelling, national scale modelling, New Zealand.

1. INTRODUCTION

The forestry sector has become New Zealand's third largest export earner contributing approximately NZ$ 2.38 billion to the economy annually. Plantation forests comprise 6 % of New Zealand's total land area of which *Pinus radiata* D. Don contributes ~91 % to the entire national plantation estate (NZFOA, 2005). Beyond economics, the sequestration of carbon by forest plantations and New Zealand's commitment to Kyoto protocol (UNFCCC, 1998) are also important. The development of renewable energy (biofuel) from wood fibre is also on the increase. Further, New Zealand is a signatory to the Montreal Process (Anon, 1995), advocating a commitment to maintaining the health, vitality and productive capacity of its forest ecosystems. Given the economic and environmental importance of the *P. radiata* species, mapping the spatial distribution of productivity will be of considerable use when determining site suitability and possible future impacts under climate change.

Modelling *P. radiata* productivity and developing spatially representative maps are not without problems, especially at the national extent. In New Zealand, generalised additive models (GAM) have been used in the development of land environments of New Zealand (LENZ) by defining areas with similar character using spatial layers including root zone

water deficit and other environmental properties related to climate and landforms (Leathwick et al., 2002a; 2003). The role of spatial models and the appropriateness of modelling using GAM, generalised linear models, regression trees, and neutral networks were discussed by Austin (1995; 2002).

Multiple linear regression techniques have been used in a number of studies to determine the major environmental contributors for New Zealand *P. radiata* productivity and site quality (Jackson and Gilford, 1974; Hunter and Gibson, 1984; Woollons et al., 2002; Watt et al., 2005, 2008). An early map describing site index rankings was developed by Eyles (1986) where site index values for *P. radiata* were associated with New Zealand land use capability units (NWASCA, 1975-79) in collaboration with expert knowledge. However, the thematic spatial distribution of *P. radiata* productivity using empirically derived models at the national extent has not been developed.

The two main estimates of *P. radiata* productivity and site quality for New Zealand are (1) the 300 Index, defined as the stem volume mean annual increment at age thirty years with a reference regime of 300 stems ha^{-1} (Kimberley et al., 2005); and (2) the Site Index, defined as the mean top height at age twenty (Goulding, 2005). Spatial modelling using regression techniques and easy-to-measure data (Burrough and McDonnell, 1998) shows potential for developing national extent productivity models and maps. Multiple linear regression using variables that are relatively cheap to measure is often generically referred to as a transfer function (Bouma and Bregt, 1989). Burrough and McDonnell (1998) defined a transfer function as a numerical model for computing new attribute

values from existing data using regression models or algorithms, i.e. mathematical relationships. New Zealand variables that could be used to develop statistical models include modelled terrain attributes (Palmer et al., 2008), interpolated climate surfaces (Leathwick et al., 2002b; Tait et al., 2006; RAV data, Chapter 3), modelled available root-zone water content (Chapters 3 and 4), and existing ancillary maps representing environmental properties (Newsome et al., 2000; Leathwick et al., 2002a; 2003). Issues of redundancy (multicollinearity) between easy-to-measure variables can be problematic when using multiple linear regression models. Partial least squares (PLS) regression (Wold, 1966) provides an alternate that circumvents issues of redundancy. It may also be possible to improve multiple linear regression predictions by summation of the model with the model's kriged residuals, commonly referred to as regression kriging (Odeh et al., 1995).

Various geostatistical techniques have the potential for interpolating forest productivity observations across the national extent of New Zealand. Numerous publications are dedicated to the discipline of geostatistics (e.g., Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989; Webster and Oliver, 1990, 2007; Cressie, 1991; Goovaerts, 1997, Diggle and Ribeiro, 2007). Examples of kriging for forestry across extensive areas can be seen for site index (Höck et al., 1993), total tree height (Samra et al,. 1989), timber volume (Homgren and Thuresson, 1997; Gunnarsson et al., 1998), and height/diameter models (Nikos, et al., 2004). Kriging techniques are diverse, with ordinary kriging, co-kriging, and regression kriging all having potential for interpolating forest productivity.

Although the advantages of applying regionalised variable theory to "natural phenomena that are too irregular at the scale of interest to be modelled analytically" are well known (Burrough and McDonnell, 1998), developing maps with kriging techniques does require practitioner experience, expertise and effort. A relatively simple and easy-to-apply interpolation technique, inverse distance weighting, is hence also worth consideration. Our study compares PLS, ordinary kriging (OK), regression kriging (RK), and inverse distance weighting (IDW) interpolation techniques for the prediction of *P. radiata* productivity across New Zealand from irregularly spaced and at times sparse sampling plot observations.

2.0 METHODS

2.1 Data acquisition

Permanent sampling plot data (PSP) are collected and stored on a national database held by Scion on behalf of many New Zealand forest owners. With owner permissions, PSP data were extracted from this database. PSP data that would impact dataset integrity were removed, specifically Nelder (spacing), oversowing (legume N), disturbance (forest floor removal), and fertiliser (P, N, and K) site-trial data. Within these excluded trials, control plots were retained. The 300 Index and Site Index values were calculated from the PSP data using the procedure of Kimberley et al. (2005), as discussed below. A further preliminary screening found that PSP stands established from the 1930s have 300 Index values ~25% lower than those of stands established since 1975. We also

determined that site productivity indices derived from trees younger than seven years of age were unreliable. These preliminary findings resulted in PSP data from stands prior to 1975, and stands younger than seven years old, being excluded from further analysis. PSPs with plot history data inadequate for estimating the 300 Index were also excluded. The final productivity variables available to this research were from 1764 PSP locations across New Zealand. The final productivity variables were averaged to a 100 m grid aligned with the variables used in the modelling of forest productivity. All variables were projected to a New Zealand map grid projection with a New Zealand geodetic datum NZGD1949.

To calculate Site Index, a national height/age model (an equation for predicting height for any age and Site Index) was used. By inverting the equation, it is possible to obtain Site Index as a function of age and mean top height. In our study, the measurement closest to age 20 years was used for each PSP. Estimation of the 300 Index, which is a measure of stem volume productivity, is more complex because, unlike height, stem volume is strongly influenced by stocking density and, to a lesser extent, thinning and pruning history. To calculate the 300 Index, a plot measurement consisting of the Basal Area, Mean Top Height and Stocking Density at a known age (the measurement closest to age 20 years was used for each PSP in this study), along with stand history information detailing the initial stocking, timing and extent of thinnings, and timing and height of prunings, are required. The 300 Index estimation procedure utilises the 300 Index Model, an empirical growth model which is sensitive to all the above inputs, and which is calibrated to a site by the 300 Index, effectively a local site productivity parameter. An

iterative procedure is used to determine the 300 Index parameter value that will predict the plot measurement.

2.2 Data extraction and pre-processing

At the 1764 *P. radiata* productivity sites, data were extracted from biophysical GIS surfaces, which included primary and secondary terrain attributes (Palmer et al., 2008), monthly and annual soil water balance (Chapter 3), monthly and annual climate variables (Mitchell, 1991; Leathwick et al., 2002b), fundamental soil layers (FSL) and land resource information (Newsome et al., 2000), vegetative cover (Newsome, 1987), foliar nutrition (Hunter et al., 1991), and biophysical surfaces (Leathwick et al., 2002a; 2003) for New Zealand. Pre-processing of this GIS data included the conversion of scaled integer data to floating point and the use of variable transformations to convert each dataset into an approximately normal Gaussian distribution. These transformations were achieved using Box-Cox power transformations (Box and Cox, 1964). SAS (SAS Institute, 2000) was used to randomly split the data into training (n = 1146) and test (n = 618) datasets (Figure 1). The test dataset was withheld from further modelling procedures and retained for validation purposes only.

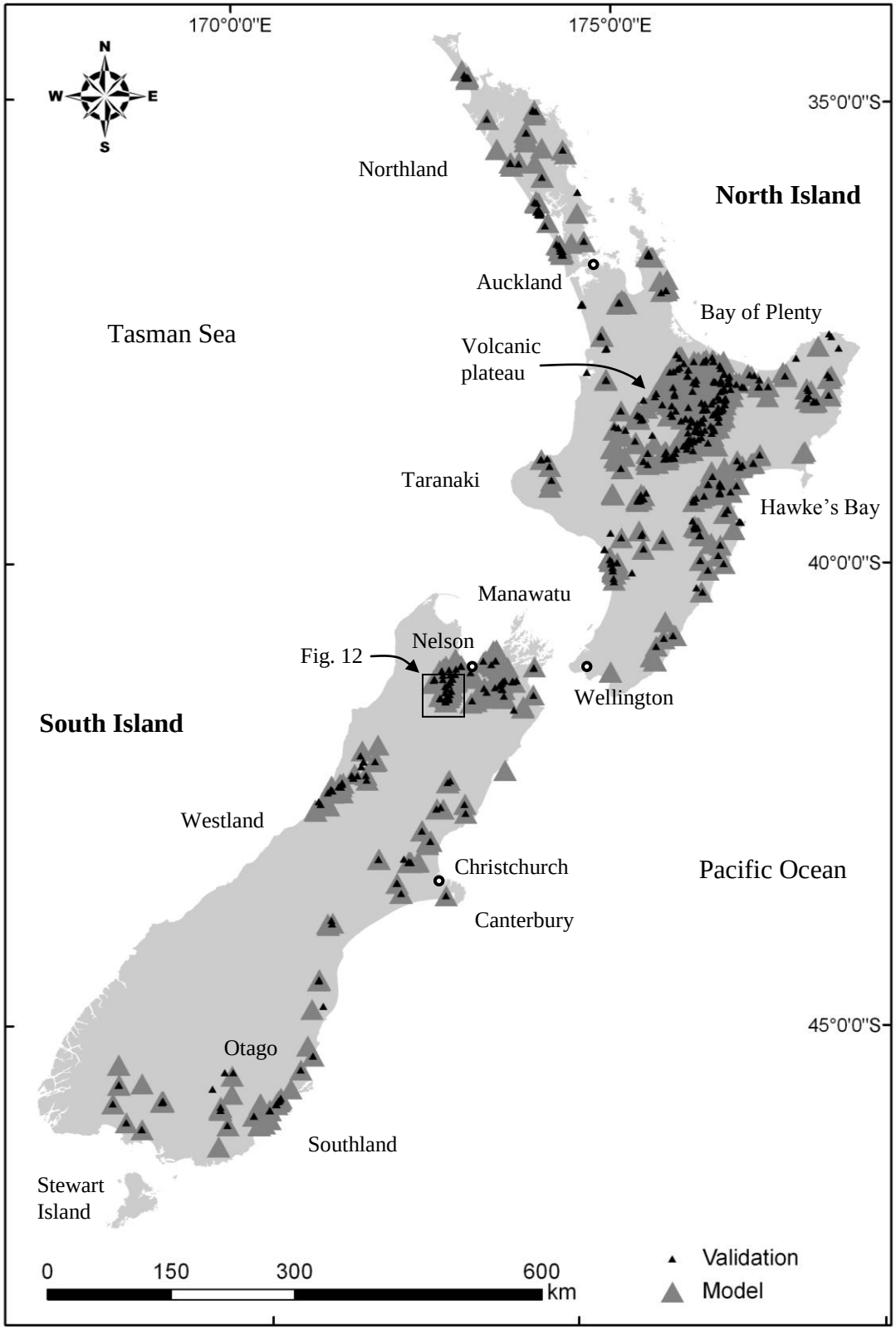


Figure 1: Distribution of PSP sites used to model *Pinus radiata* productivity and the withheld sites randomly selected for model validation. Place names mentioned in the text are also shown.

2.3 Partial least squares regression surface development

The PLS procedure originally developed by Herman Wold (1966) was used to develop models for *P. radiata* 300 Index and Site Index for New Zealand. The test set validation method allows the choosing of successive orthogonal factors that minimise the covariance between each *X*-score (also referred to as extracted factors) and the corresponding *Y*-score. The van der Voet (1994) test was used to choose the least number of extracted factors whose residuals are not significantly greater than those of the model with minimum error (minimum PRESS) with a significance level of 0.05.

Those predictors only providing small contributions towards the model were identified using two methods. Firstly, data were screened for regression coefficients for the standardised data with minimal contribution to the response prediction. Secondly, the variable importance for projection (VIP) statistic (Wold, 1994) was used to identify the value of each predictor in fitting the PLS model for both predictors and response. If a predictor has a relatively small coefficient (in absolute value) as well as a small VIP, then the variable is a prime candidate for removal. Independent variables were reduced to a minimum using the above protocol while maximising the explanation of the variance for the 300 Index and Site Index. For details regarding independent variables used in the modelling procedure and variation accounted by each extracted factor, refer to Tables 1, 2, and 3 in Chapter 5. The SAS (Version 9) PLS procedure was used for model development.

2.4 Regression kriging and ordinary kriging

Regression kriging (Odeh et al., 1995) is a spatial interpolation technique that combines regression of the dependent variables (300 Index and Site Index) with ordinary kriging of the regression residuals. PLS residuals refer to the difference between predicted values from the PLS model and the measured (observed) values. Predictions (z^*) made using regression kriging were calculated at unvisited sites (S_o) using:

$$z^*(S_o) = zpr^* + \varepsilon^* \quad (1)$$

where zpr^* represents the 300 Index or Site Index PLS predictions and ε^* represents the kriged PLS residuals. This is a hybrid modelling technique because it combines statistical and geostatistical techniques.

Ordinary kriging estimates the weighted average of the 300 Index and Site Index at the prediction locations (unvisited sites), $z^*(S_o)$, and takes the form:

$$z^*(S_o) = \sum_{i=1}^{n} \lambda_i z(S_i) \quad (2)$$

where λ_i are the weights given to each of the n development sites and $z(S_i)$ are the 300 Index or Site Index for either observed values or PLS residual values from the training dataset. Using the training dataset, experimental semi-variograms were developed for the

300 Index and Site Index using the observed variables and residuals from the PLS models, and curves were fitted to represent the data. SAS (Version 9) was used in the development of all kriged surfaces.

2.5 Inverse distance weighting model development

IDW is regarded as one of the standard interpolation procedures found in many GIS packages (Burrough and McDonnell, 1998; Longley et al., 2005) and it is often used as the default surface generation method for attribute sample locations (Lu and Wong, 2007). IDW estimates the unknown value $\hat{y}(S_0)$ at location S_0 in the following manner:

$$\hat{y}(S_0) = \sum_{i=1}^{n} \lambda_i y(S_i) \tag{3}$$

where n is the number of measured sample points surrounding the prediction location used in the prediction, λ_i are the weights assigned to each measured point, and $y(S_i)$ is the observed value at the location S_i. Weights are determined by:

$$\lambda_i = d_{io}^{-p} / \sum_{i=1}^{n} d_{io}^{-p} \qquad \sum_{i=1}^{n} \lambda_i = 1, \tag{4}$$

As the distance increases, the weight is reduced by a factor of p, and d_{io} is the distance between the prediction location, S_0, and each of the measured locations, S_i. Therefore, the

power parameter *p* influences the weighting of the observed location's value on the prediction location's value. Effectively this means that as the distance increases between the observed sample locations and the prediction location, the weight the observed data point will have on the prediction decreases exponentially. The optimal *p* value was determined by minimising the root-mean-square prediction error (RMSPE) through standard cross-validation techniques available in ArcGIS™ using geostatistical analyst (Johnston et al., 2001). For our modelling purposes, *p* was set to 1.3.

2.6 Validation and assessment of the prediction techniques

Performance of the techniques PLS, RK, OK, and IDW were evaluated by determining how close the predicted values (z^*) were to the observed values (z) at the (l) test (validation) locations (S_j).

The mean error (ME) was used to assess prediction bias, whereas the root-mean-square-error (RMSE) was used as a measure of prediction precision as defined by Voltz and Webster (1990) and Triantafilis et al. (2001). The ME was determined using:

$$\mathrm{ME} = \frac{1}{l}\sum\left[z(s_j) - z^*(s_j)\right] \quad (5)$$

ME values should be closer to zero for unbiased predictions (Odeh et al., 1994; Triantafilis et al., 2001). A negative ME value indicates that the prediction technique is

over-estimating, conversely a positive ME value suggests an underestimation of prediction (Triantafilis et al., 2001). RMSE was calculated in the following manner:

$$\mathrm{RMSE} = \left\{ \frac{1}{l} \sum_{j=1}^{l} \left[z(s_j) - z^*(s_j) \right]^2 \right\}^{0.5} \tag{6}$$

Predictions are considered increasingly precise with lower RMSE values (Odeh et al., 1994; Triantafilis et al., 2001).

A relative measure of precision was also determined, known as goodness-of-prediction (G) (Agterberg, 1984; Schloeder et al., 2001). The G value is a measure of the prediction technique's precision relative to the precision of the sample mean. G was determined by:

$$\mathrm{G} = \left(1 - \left\{ \sum_{j=1}^{l} \left[z(s_j) - z^*(s_j) \right]^2 / \sum_{j=1}^{l} \left[z(s_j) - \bar{z} \right]^2 \right\} \right) \times 100 \tag{7}$$

Where $\bar{z}$ refers to the sample mean (Schloeder et al., 2001). A positive G value indicates that the prediction technique performed better than the sample mean, whereas a negative G value indicates that the prediction technique had a poorer performance than the sample mean. A perfect prediction (with a RMSE of zero) would have a G value of 100% (Gotway et al., 1996).

ME and RMSE under some conditions are not always considered to be consistent and reliable assessments of bias and precision (Odeh et al., 1994; Triantafilis et al., 2001). The mean rank (MR) and standard deviation of rank (SDR) were used as an additional evaluation of the prediction techniques. The ranking of each prediction technique was determined for each validation data location (r_{ij}) on the basis of the squared error of the prediction (Odeh et al., 1994; Triantafilis et al., 2001). As a result, lowest ranked prediction techniques have the smallest squared error, and highest ranked prediction techniques have the greatest squared error. The MR of the *i*th prediction technique was determined using:

$$\mathrm{MR} = \frac{1}{n}\left[\sum_{i=1}^{n} r_{ij}\right] \tag{8}$$

The SDR of the *i*th prediction technique was determined in the following manner:

$$\mathrm{SDR} = \left[\frac{1}{n-1}\sum_{j=1}^{n}\left(r_{ij} - \mathrm{MR}\right)^{2}\right]^{0.5} \tag{9}$$

Prediction techniques with the lowest MR and SDR are considered to be the most accurate and consistent, respectively.

2.7 Map prediction and neighbourhood search parameters

Defining a search neighbourhood eliminates the influence of distant observations to prediction locations. Therefore, all map prediction neighbourhood parameters were set to a minimum of two points and a maximum of 15 points with a search radius of 80 km for OK and IDW techniques. The RK technique was an exception to this where the search radius was set to 25 km to reflect the shorter range over which spatial dependence was operating.

3.0 RESULTS

3.1 Regression kriging modelling

For both the RK and OK techniques, exponential curves were fitted to the experimental semi-variogram models (Table 1). For RK the 'kriged residuals' for the 300 Index had a small to moderate nugget, with a sill reached at around 7 km. Site Index has a smaller nugget, but also attains its sill over a range of about 7 km, but with a sill about half that for the 300 Index (Figure 2).

Table 1: Coefficients of the models fitted to the semi-variograms for the 300 Index and Site Index observed data and the PLS model residuals.

Variable	Units	Model	Nugget	Sill	Range (m)	Lag (m)
Observed						
300 Index	(m^3 ha^{-1} yr^{-1})	Exponential	5.0	29.0	27000	1500
Site Index	(m)	Exponential	2.5	16.0	20000	1500
Residuals						
300 Index	(m^3 ha^{-1} yr^{-1})	Exponential	3.0	9.5	2000	1500
Site Index	(m)	Exponential	1.8	5.0	2400	1500

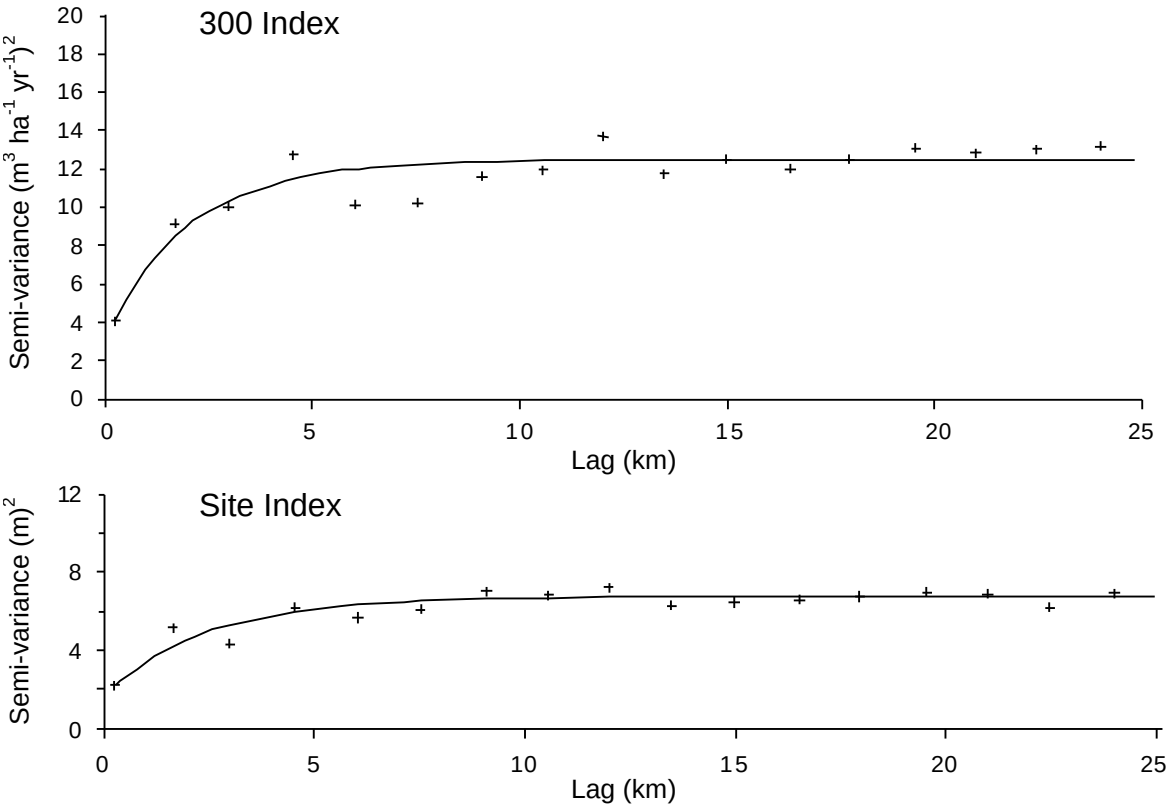


Figure 2: Semi-variograms of the partial least squares residuals for the 300 Index (upper) and Site Index (lower).

The OK semi-variogram for the 300 Index display a moderate nugget, with a sill attained at around 80 km. Site Index also reaches its sill over a range of about 80 km, but with a nugget and sill about half that for the 300 Index (Figure 3). Parameters of the curves fitted to the experimental semi-variogram models are in Table 1.

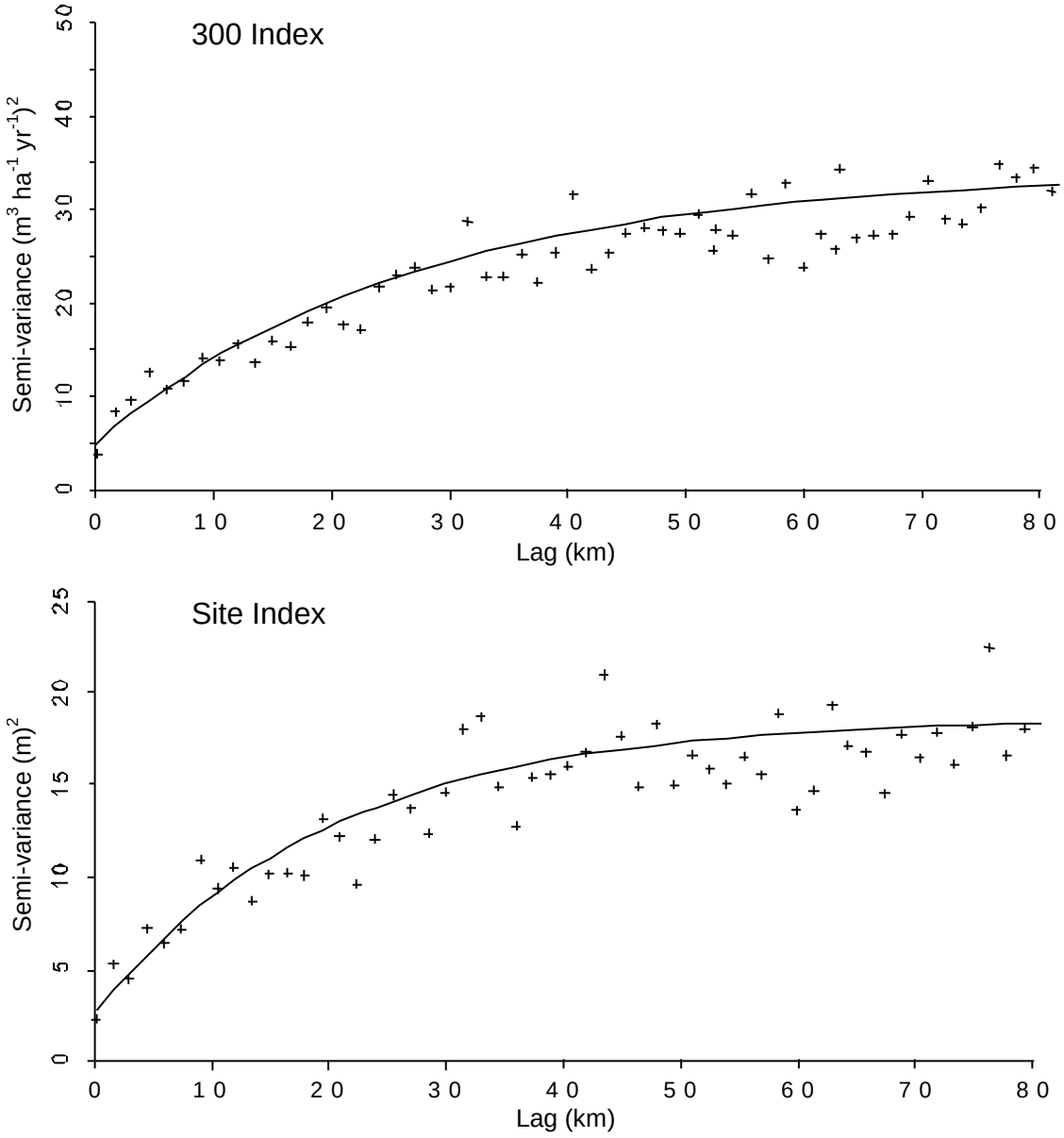


Figure 3: Semi-variograms for the 300 Index (upper) and Site Index (lower) observed data.

3.2 *Pinus radiata* productivity surfaces for the 300 Index and Site Index

In general, the reference dataset (REF) has the lowest mean values while having the greatest range for 300 Index and Site Index values when compared with PLS, IDW, RK, and OK prediction techniques (Table 2). The RK technique for the 300 Index on average over predicted by 0.45 m^3 ha^{-1} yr^{-1}, whereas the PLS technique only marginally over predicts Site Index by 0.16 m above the observed data mean.

Table 2: 300 Index and Site Index summary statistics for the observed data (REF), partial least squares (PLS), inverse distance weighting (IDW), regression kriging (RK), and ordinary kriging (OK) models predicted at validation sites. For each observed and predicted variable the sample mean, minimum, maximum, and standard deviation are shown.

		REF	PLS	IDW	RK	OK
300 Index	(m^3 ha^{-1} yr^{-1})					
	Mean	26.34	26.73	26.73	26.79	26.62
	Min	9.70	10.39	10.26	10.22	10.77
	Max	49.95	39.95	45.32	42.57	42.74
	Standard deviation	5.85	4.74	5.22	5.23	5.15
Site Index	(m)					
	Mean	30.01	30.17	30.08	30.11	30.06
	Min	15.07	18.99	15.08	15.27	15.63
	Max	44.30	38.45	42.54	41.71	42.53
	Standard deviation	4.87	4.00	4.40	4.47	4.33

Maps of the 300 Index are shown in Figures 4 and 5, and Figures 6 and 7 show Site Index across New Zealand. Generally, for the PLS and RK prediction techniques (Figures 5 and 7), estimates of forest productivity are elevated at northern latitudes and decline towards southern regions. Also, forest productivity values tend to decrease in the cooler mountainous interior regions. However, from a visual analysis of the map, OK and IDW prediction techniques (Figures 5 and 7) do not always follow these trends.

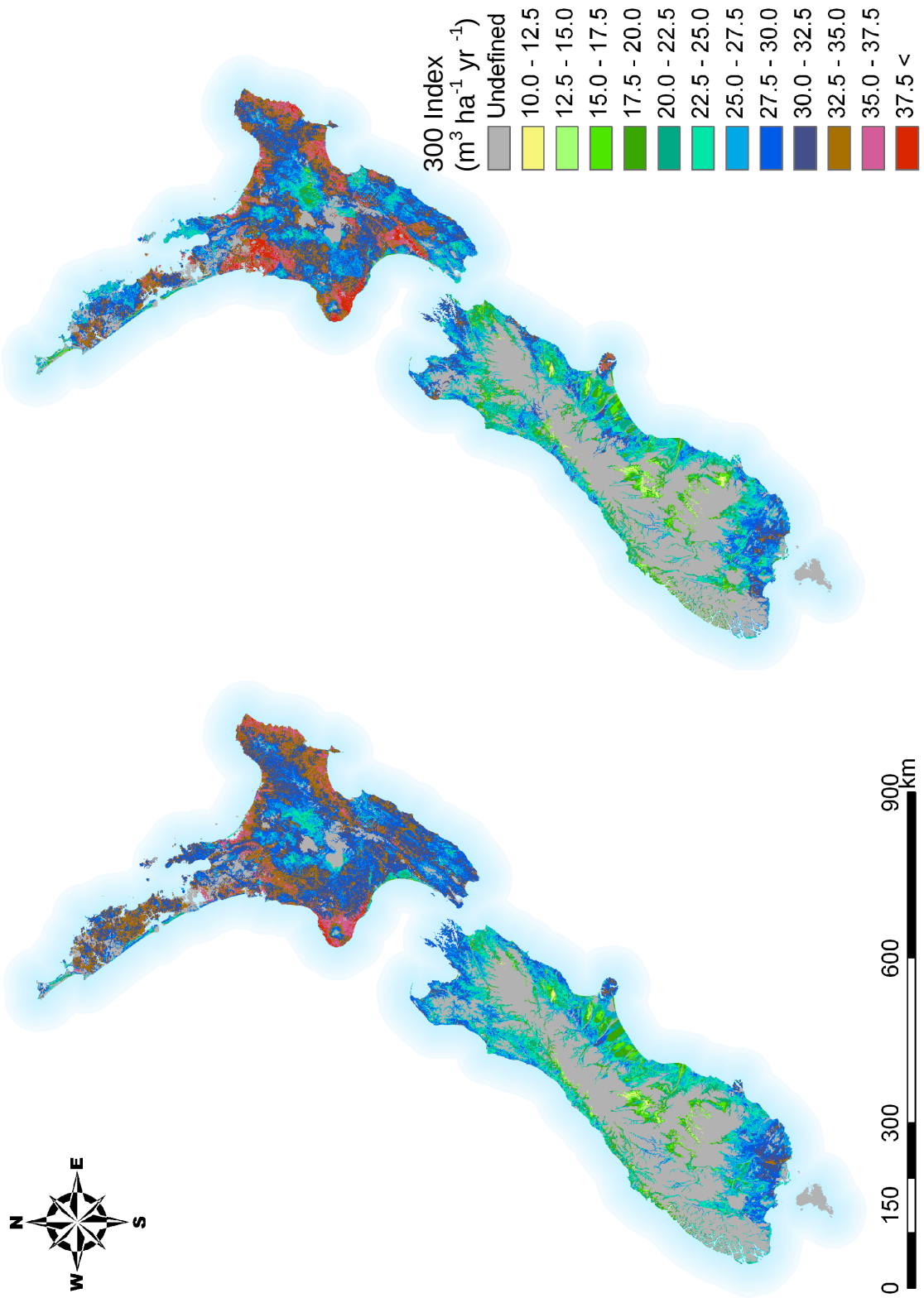


Figure 4: Map of the 300 Index values across New Zealand developed using the partial least squares (left) and regression kriging (right) prediction techniques.

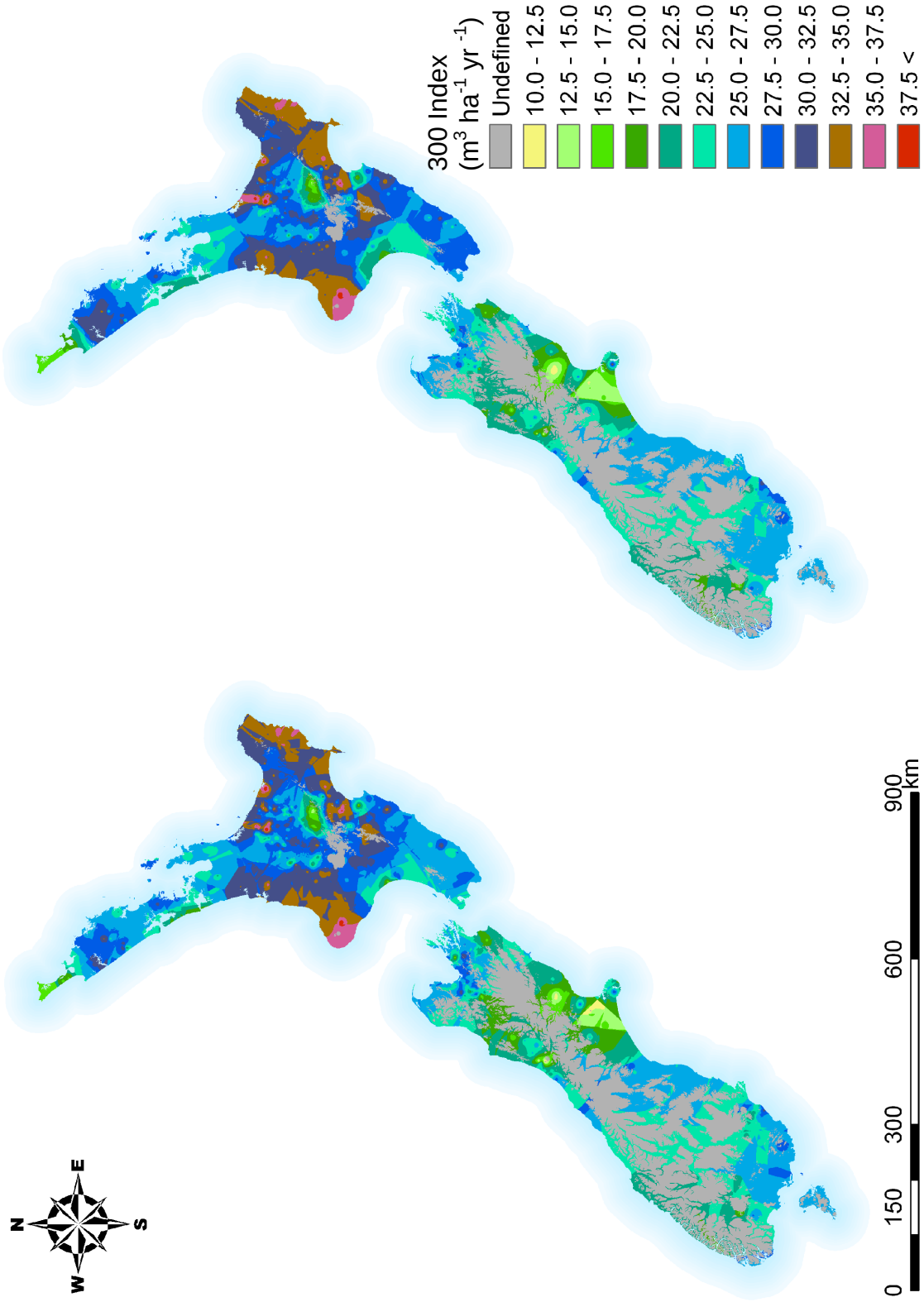


Figure 5: Map of the 300 Index values across New Zealand developed using ordinary kriging (left) and inverse distance weighting (right) prediction techniques.

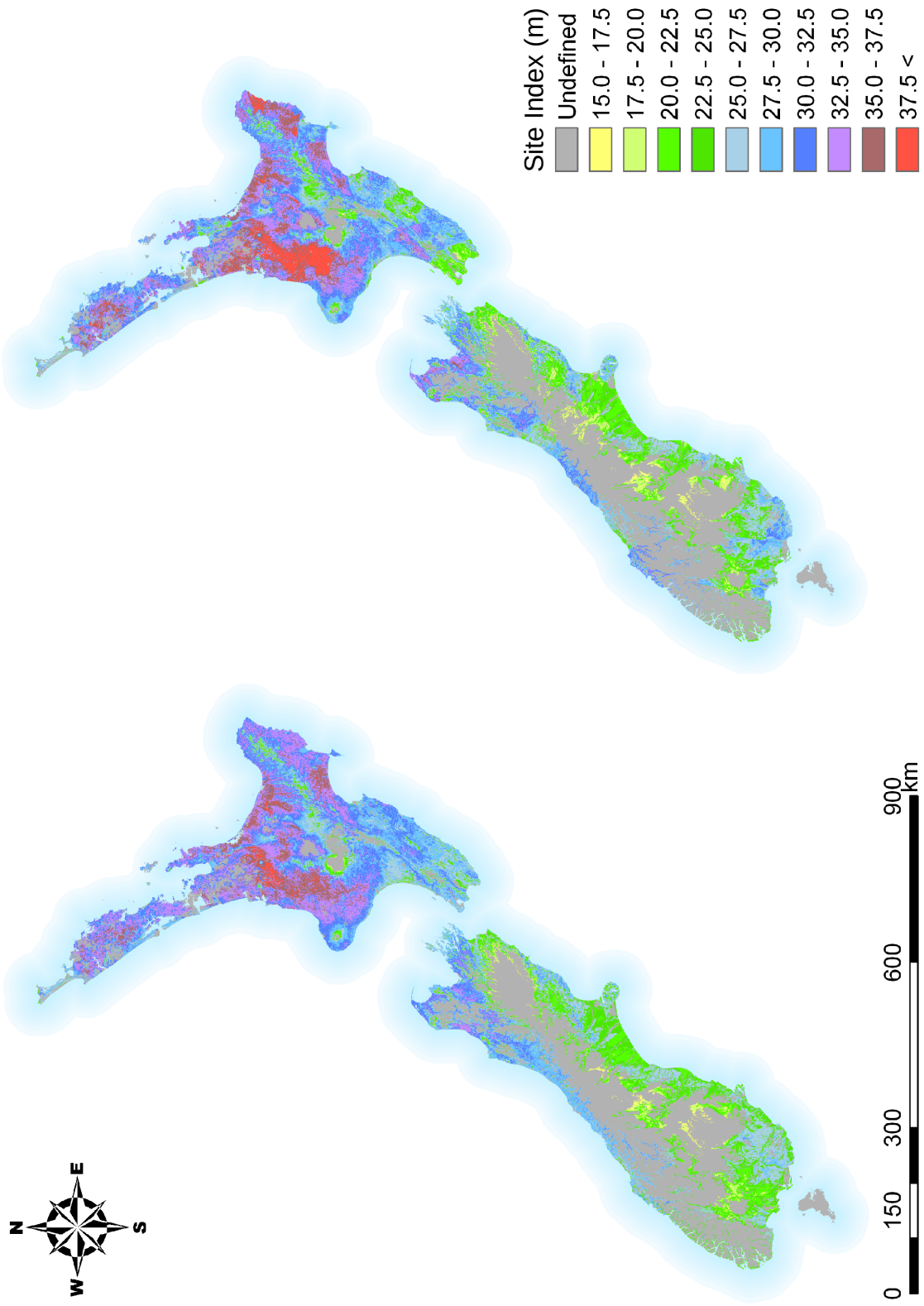


Figure 6: Map of Site Index values across New Zealand developed using the partial least squares (left) and regression kriging (right) prediction techniques.

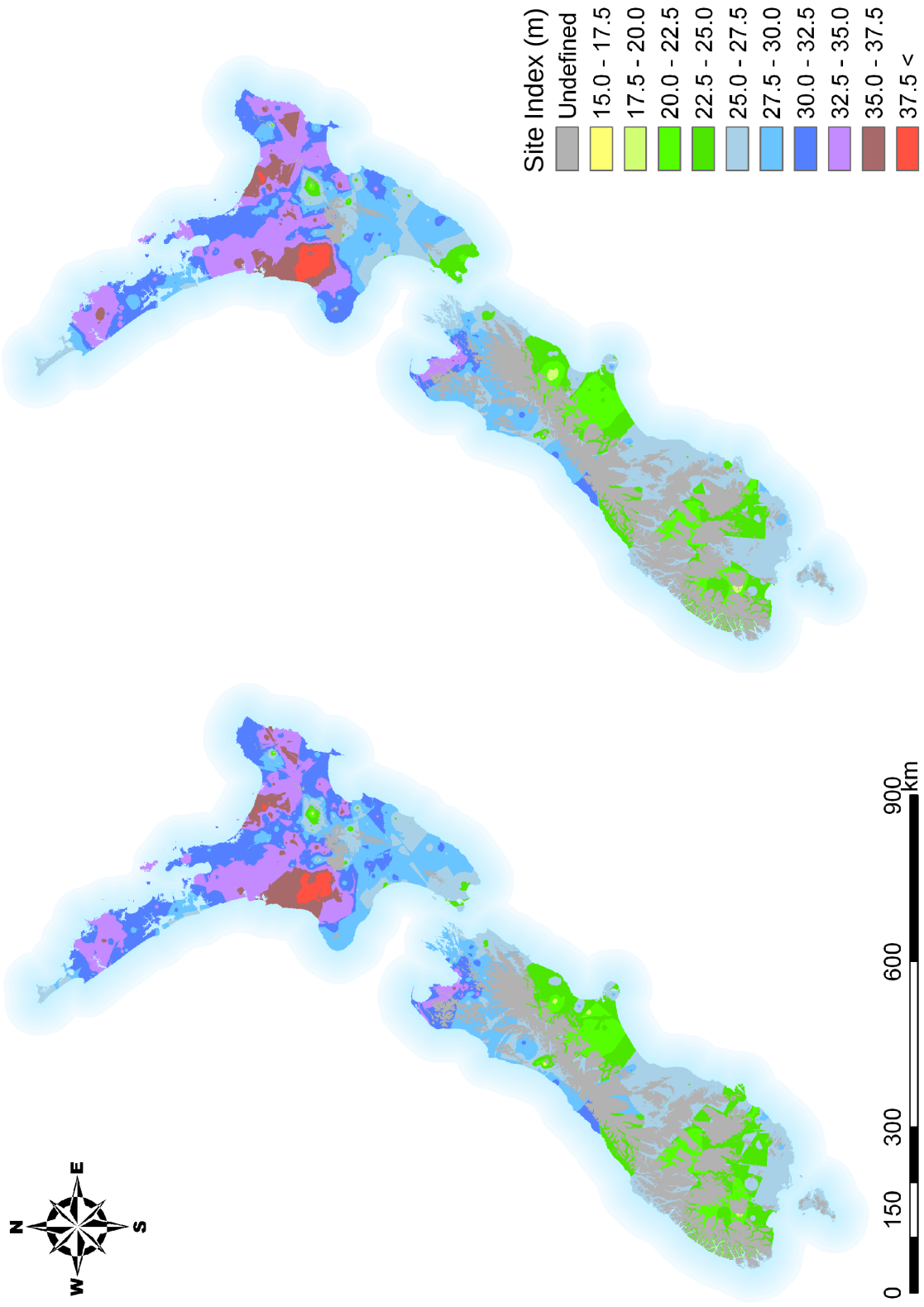


Figure 7: Map of Site Index values across New Zealand developed using ordinary kriging (left) and inverse distance weighting (right) prediction techniques.

3.3 Spatial prediction of the productivity surfaces

The best predictive PLS models developed explain 58% and 67% of the variance in the 300 Index and Site Index, respectively. Plots of residual and actual values against predictions exhibit little apparent bias for either 300 Index or Site Index data (Figure 8).

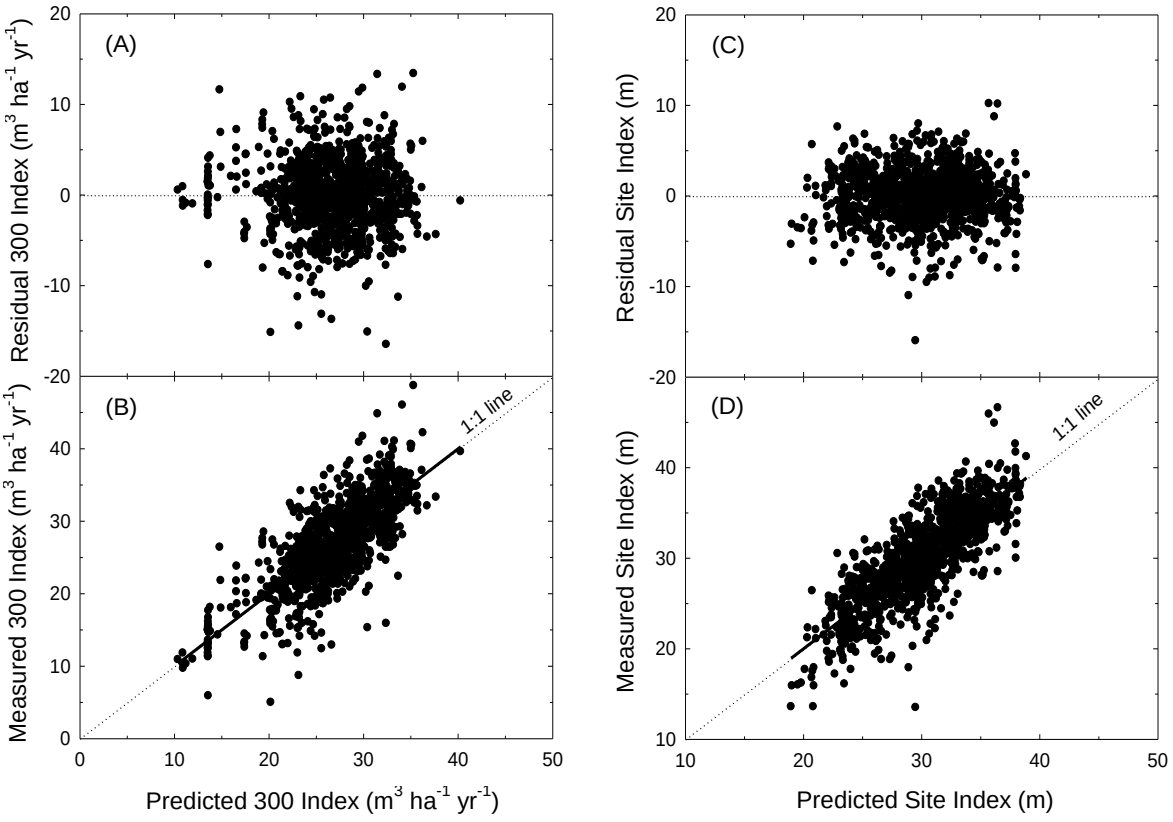


Figure 8. Plots of partial least squares regression showing the relationship between (A) predicted and residual 300 Index and (B) predicted and measured 300 Index, (C) predicted and residual Site Index and (D) predicted and measured Site Index. In (B and D) the 1:1 line is shown as a dotted line. Also shown as a thick line is a linear line of best fit (n = 1146) (figures reproduced from Chapter 5, Figure 6).

3.4 Validation of productivity surfaces

Validation results show a substantial improvement in the PLS regression predictions through RK, with coefficient of determination (R^2) values increasing by 11 % and 13 % for Site Index and 300 Index, respectively. Furthermore, cross validation procedures show substantial improvements to the PLS models through RK, with decreases in RMSE and increases in G values for both the 300 Index and Site Index models (Table 3). RK had the best RMSE and G prediction statistics for Site Index, whereas OK provided the best overall predictions at validation locations for the 300 Index.

Overall, the RK prediction technique was ranked to have the lowest squared error (MR), whereas OK was the most consistent (SDR) predictor (Figure 9). To generalise, the interpolation techniques RK, OK, and IDW performed at a similar standard followed by PLS and then the data reference mean (REF).

Table 3: Validation statistics for the correlation of determination (R^2), mean error (ME), root-mean-square error (RMSE), and goodness-of-prediction (G) statistics for 300 Index and Site Index prediction techniques.

Prediction techniques[1]	300 Index				Site Index			
	R^2	ME	RMSE	G	R^2	ME	RMSE	G
PLS	0.48	-0.39	4.22	48.03	0.59	-0.16	3.13	58.63
RK	0.61	-0.45	3.65	61.11	0.70	-0.10	2.65	70.35
OK	0.61	-0.28	3.63	61.46	0.69	-0.05	2.70	69.26
IDW	0.60	-0.39	3.68	60.47	0.69	-0.07	2.71	68.80
REF	-	0.00	5.85	0.00	-	0.00	4.86	0.00

[1] PLS = partial least squares regression, RK = regression kriging, OK = ordinary kriging, IDW = inverse distance weighting, REF = reference technique.

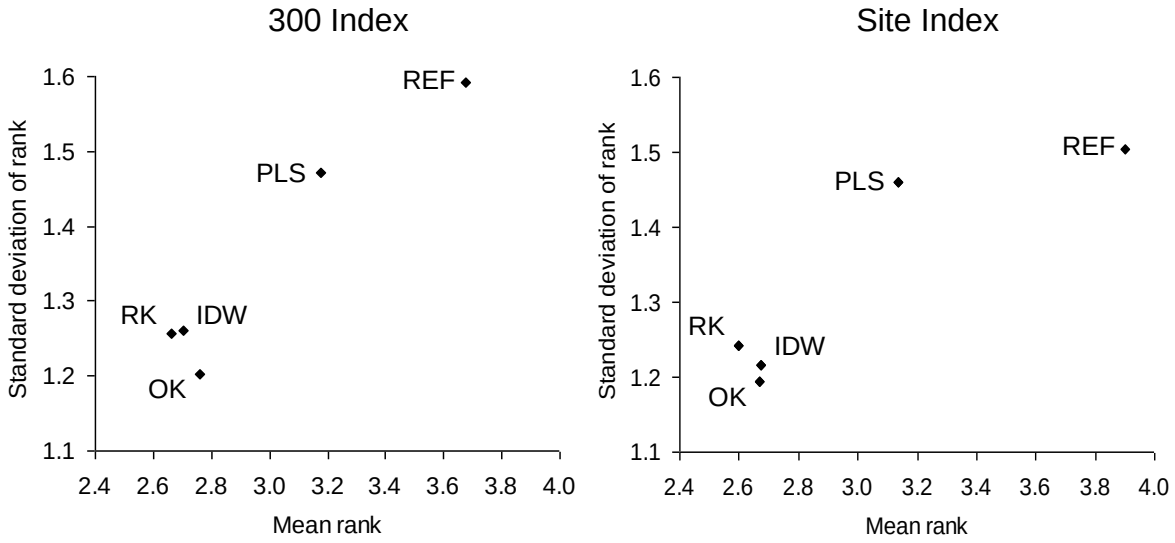


Figure 9: Validation statistics for the accuracy and consistence of the prediction techniques for PLS = partial least squares regression, RK = regression kriging, OK = ordinary kriging, IDW = inverse distance weighting, REF = reference technique using mean rank (MR), and standard deviation of rank (SDR).

Substantiation of the validation statistics can be seen in Figures 10 and 11. RK provided the best predictions for the 300 Index (R^2 = 0.61), with OK and IDW prediction techniques following closely. A similar result was found at the validation locations for Site Index predictions with R^2 = 0.70, 0.69, and 0.69 for RK, IDW, and OK, respectively.

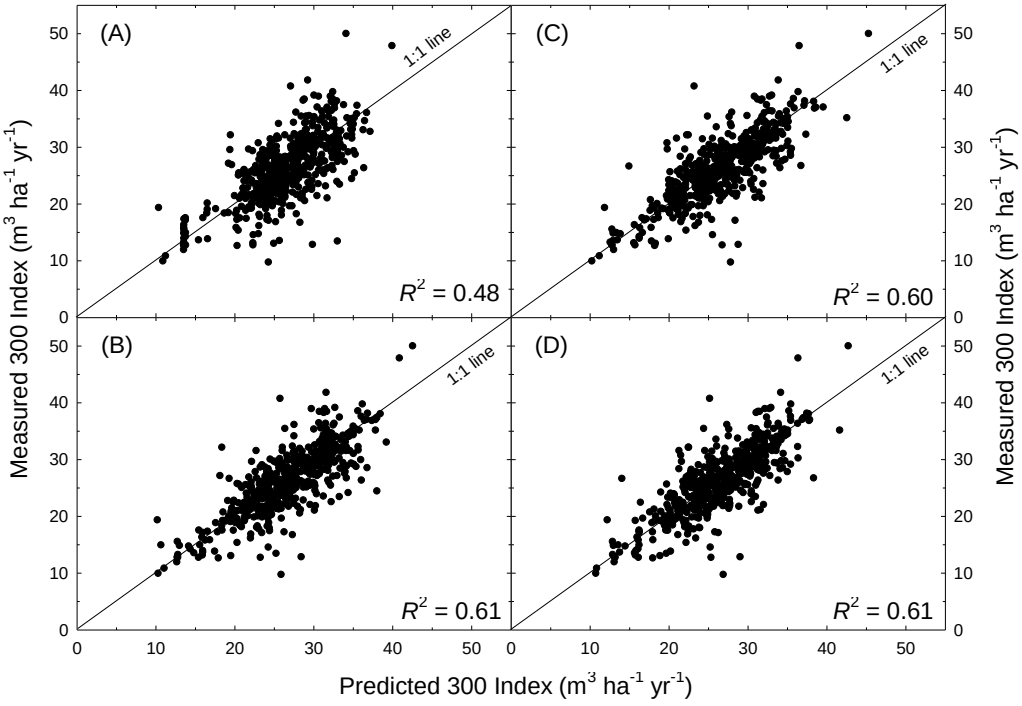


Figure 10. Validation plots of predicted verses measured for the 300 Index values using (A) partial least squares, (B) regression kriging, (C) inverse distance weighting and (D), ordinary kriging (n = 552).

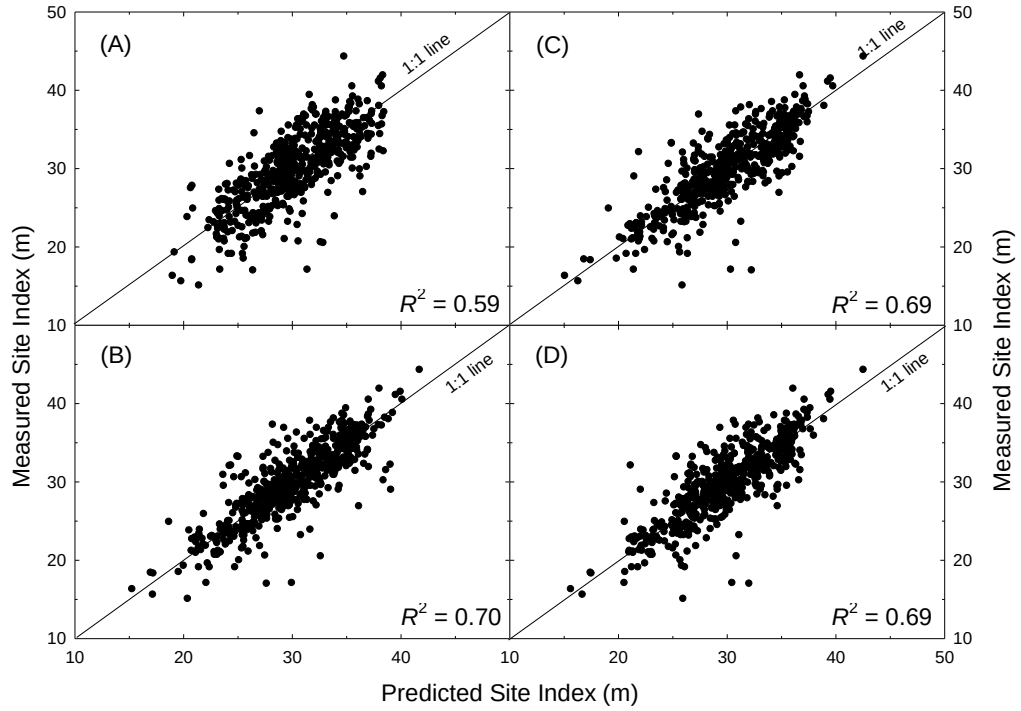


Figure 11. Validation plots of predicted verses measured for the Site Index values using (A) partial least squares, (B) regression kriging, (C) inverse distance weighting and (D), ordinary kriging (n = 552).

4.0 DISCUSSION

4.1 Visual assessment of the 300 Index and Site Index maps

The 300 Index and Site Index maps displayed here are a major step forward in the visualisation and assessment of forest productivity. Maps are empirically based and provide detailed information of the spatial distribution across the national extent of New Zealand.

When developing OK, IDW, and to a lesser degree RK models, the spatial location of the property of interest and its proximity to other observations has an influence on mapping predictions. Spatial location also influences the degree of certainty associated with each prediction. Ideally, sample observations should be represented at an operational scale that adequately captures the spatial variability across the region of interest. In this context we refer to operational scale as defined by Bian (1997) to mean the scale over which the phenomenon operates or varies spatially. In the real world, however, capturing all spatial variability is seldom ever achieved and we resort to techniques to improve predictions under these circumstances (Schloeder et al., 2001). Certainly, this project is no exception with an irregularly spaced and sometimes sparse sample dataset (see Figure 1). For this reason IDW and OK techniques for the 300 Index and Site Index (Figures 5 and 7) have resulted in maps that miss some important features, making actual productivity values at these locations unclear. For example, expert knowledge tells us that Site Index heights should reduce to about 15 at around 900 m elevation. However, the cooler mountainous

interiors of central North and South islands where productivity values normally taper off are generally absent in IDW and OK predictions.

The impact of sparse sample observational density is illustrated in Figures 12C and D where elevated mountainous regions along the western and south eastern graphic margins remain undefined (undefined grey areas shown Figures 12A and B). At the small map scale shown (1: 200,000 as depicted on A4 page), the detailed spatial patterns of the Nelson region (Figure 1) become evident, highlighting the differences between PLS, RK, OK and IDW maps. The PLS map (Figure 12A) provides the greatest spatial detail because it implicitly disregards spatial covariation and is not reliant on observational density. The RK model consists of a PLS regression component and a residuals component (kriged PLS residuals) and can be termed a hybrid model. The RK map, therefore, shows similarity to its related PLS counterpart (Figure 12B). It is interesting to note that central locations in the PLS map (Figure 12A) decline by one class in the RK map from 30-32.5 to 27.5-30 m class (Figure 12B). Conversely, the RK technique elevates some spatial predictions in response to higher observed site values at those locations. These extreme Site Index values can be seen at OK and IDW locations, represented by pink colours (Figure 12C and D).

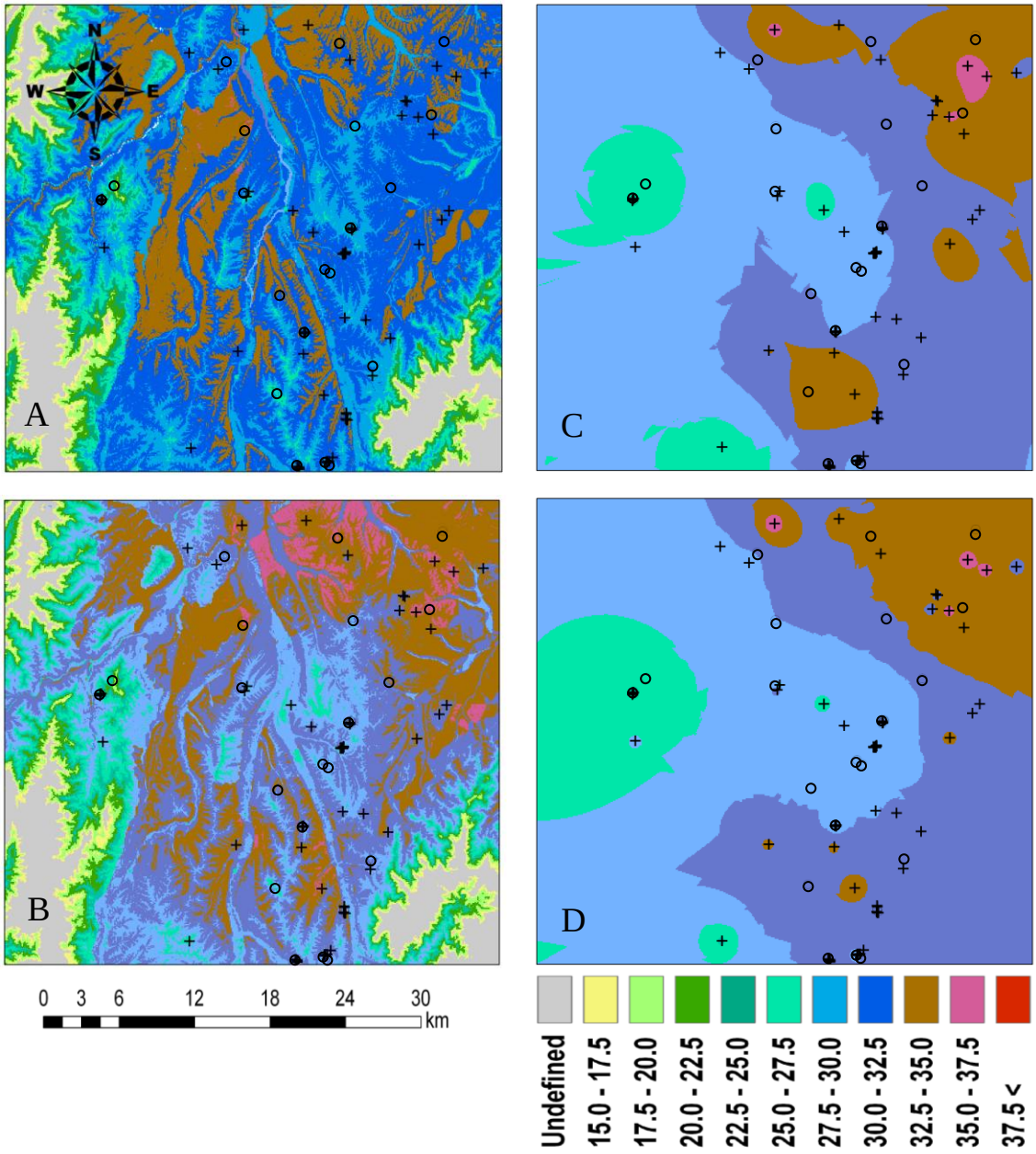


Figure 12: Comparison of (A) partial least squares, (B) regression kriging, (C) ordinary kriging, and (D) inverse distance weighting interpolated maps and the impact on spatial variation when productivity observations are dense (centre of the graphic) and sparse (elevated grey areas) for Site Index (m) in the Nelson region (Figure 1). Crosses are observed data locations whereas circles are validation locations withheld from the modelling processes.

Kriging techniques in many situations are considered substantially superior compared with IDW techniques (Zimmerman et al., 1999). However, the overall accuracy of a spatial map will depend on many factors including sample observation density and configuration, search neighbourhood size, shape, and orientation (e.g. Gotway et al., 1996). For our applied study, sample configuration and density cannot be controlled. However, analysing the semi-variogram provides some insight into neighbourhood size and shape. OK semi-variograms show productivity observations separated by distances less than ~80 km to be autocorrelated, whereas locations beyond this range are not (Figure 3). For the kriged residuals (RK models), autocorrelation continues to a range of less than 7 km (Figure 2). Although not provided here, an advantage the kriging approach offers is the quantitative measure of the error associated with each prediction. Such a measure is not available with the IDW technique (Bregt et al., 1992, Gotway, et al., 1996). For the forest manager who wishes to consider the reliability of the spatial prediction at the compartment or stand level, the spatial dependence over which productivity varies is important. Furthermore, assessment of the error associated with each prediction becomes crucial.

Interestingly, cross validation statistics for the 300 Index and Site Index show RK, OK, and IDW generally performed to a similar level of precision and bias, but outperformed the PLS technique. Considering the spatial prediction limitations surrounding sparse, scattered data, it is perhaps not surprising that RK, OK, and IDW have similar prediction validation statistics. In reality, validation locations, although randomly selected, remain within current operational forest stands and by default generally lie within close

proximity to other productivity observations. This proximity relationship means that the validation statistics presented here are at best only applicable to currently forested regions with relatively high sample observation densities.

In the assessment of PLS, RK, OK, and IDW techniques at the New Zealand spatial extent (e.g., at a map scale 1: 6,000,000), the visual differences are substantial (Figures 4-7). However, without 'zooming' into the spatial extent of an individual region, such discussion remains beyond the scope of this chapter. When assessing the validity of mapping techniques, it is essential to first determine the purpose of the maps. For example, if the map is to provide an overview of *P. radiata* productivity across New Zealand (e.g., 1: 6,000,000 map scale), then IDW and OK techniques provide useful generalisations. Although cross validation statistics indicate that OK provided slightly better predictions, the IDW prediction technique requires substantially less effort to implement. However, some issues to consider when developing maps from IDW are the use of appropriate power functions (Gotway et al., 1996, Kravchenko, and Bullock, 1999) and the "bulls eye" pattern that sometimes dominate maps (Gotway et al., 1996). This "bulls eye" pattern can be seen to some degree in Figure 5.

Another option not used here would be interpolation using cokriging. Cokriging could potentially use temperature and soil water balance datasets known to be highly correlated with productivity (Chapter 5) to overcome some of the issues arising from the irregularly spaced and sometimes sparse productivity observations. However, if surface development

requires a map with greater quantitative rigor, then techniques implementing transfer functionality should be considered.

4.2 Partial least squares and regression kriging for mapping the 300 Index and Site Index

Multiple linear regression (MLR) techniques, as noted earlier, have been commonly used to model *P. radiata* productivity for New Zealand (e.g., Watt et al., 2005, 2008). The PLS regression technique used here explores the functional relationship between spatially variable attributes. Evaluating long range variation can be undertaken using global regression (Burrough and McDonnell, 1998). The 300 Index and Site Index maps developed using PLS have a coefficient of determination (R^2) explaining 58% and 67% of the variance, respectively. The advantage that global regression has over other techniques is two fold: (1) PLS provides an estimation of the certainty of prediction outcomes; and (2) PLS is not reliant on the separation distance between observations. MLR models implicitly disregard spatial covariation in addition to their lack of sensitivity to measurement error (Odeh, 1995). Thus MLR models have the advantage through the utilisation of local environmental and landform information. For maps developed from PLS this advantage means better spatial representation of productivity estimates when data are numerous but sparse (Figure 1).

The RK technique used here is the summation of the PLS model and ordinary kriging of the PLS model residuals. This combined approach has the advantages of the regression prediction technique previously discussed, while OK of the regression residuals incorporates regression and measurement uncertainties into the kriging environment (Odeh et al., 1995). The findings from our study show that RK has similar prediction precision and bias to those of OK and IDW prediction techniques (Table 3). These results are reaffirmed with RK having the lowest mean rank, while OK and IDW have the lowest standard deviation of rank (Figure 8). If our assessment of prediction performance were based purely on cross validated data from forested regions, RK would be the model of choice. However, from a visual assessment and perspective of expert knowledge, PLS provides information at remote spatial locations that IDW and OK do not provide. Additionally, OK of the residual component in RK can only improve predictions within the range of spatial dependence in conjunction with supporting observations. Therefore regions including the volcanic plateau of the central North Island and the Nelson area, where greater productivity observational densities occur, prediction certainty will increase. For these reasons, RK is the preferred prediction technique where productivity observations are relatively dense. However, the RK technique provides little advantage for forest managers or scientists interested in mapped locations where observational data are sparse. In this situation the PLS model is recommended.

5.0 CONCLUSION

PLS and RK interpolation techniques have shown substantial improvements in visualising and assessing *P. radiata* productivity across New Zealand. These improvements are attributed to the multiple linear regression model that implicitly disregards spatial covariation. Contrary to the visual assessment, cross validation findings suggest OK, IDW, and RK techniques predicted at similar levels of precision and bias. However, OK and IDW failed to predict important features, including the reduced productivity at elevated mountainous regions because of low sample observational representation. Therefore, the PLS models were considered to be the best predictors of forest productivity at locations where observational data were sparse. Cross validation techniques were limited in the practical assessment of productivity prediction because validation sites were restricted to locations currently forested and thereby increasing the chance of higher density observations. RK was considered an improvement on PLS predictions, but only where observations were abundant. Conversely, the PLS map evidently has greater certainty across regions where productivity observations were sparse.

6.0 ACKNOWLEDGEMENTS

We gratefully thank Carolyn Anderson and Mark Dean for their assistance in obtaining permission to use and extract PSP data. We are also indebted to the numerous forestry companies and private owners for supporting this research. David Palmer wishes to thank and acknowledge the Foundation for Research Science and Technology (FRST) and the University of Waikato (Doctoral Scholarship) for their funding. Barbara Höck thanks FRST for funding (Protecting and enhancing the environment through forestry CO4X0304).

7.0 REFERENCES

Agterberg, F.P., 1984. Trend surface analysis. *In*: G.L. Gaile and C.J. Willmott (Eds.), Spatial Statistics and Models. Reidel, Dordrecht, Netherlands, pp. 147-171.

Anon, 1995. Criteria and indicators for the conversion and sustainable management of temperate and boreal forests - the Montreal Process. Canadian Forest Service, Natural Resources Canada, Hull, Quebec, Canada K1A 1GD, 27 pp.

Austin, M.P., Meyers, J.A., Belbin, L., Doherty, M.D., 1995. Simulated data case study, Sub-project 5, Modelling of landscape patterns and processes using biological data. Division of Wildlife and Ecology, Commonwealth Scientific and Industrial Research Organisation, 99 pp.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157, 101-118.

Bian, L. 1997. Multiscale nature of spatial data in scaling up environmental models. In: Quattrochi, D.A., Goodchild, M.F. (Eds), Scale in remote sensing and GIS, Lewis Publishers, Boca Raton, pp. 13-26.

Bouma, J., Bregt, A.K., (Eds), 1989. Land quantities in space and time. Proceedings of a symposium organized by the International Society of Soil Science (ISSS), Wageningen, the Netherlands 22-26 August 1988. Wageningen: Pudoc. 352 pp.

Box G.E.P., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistics Society, B-26, 211-252.

Bregt, A.K., Gesink, H.J., Alkasuma, 1992. Mapping the conditional probability of soil variables. Geoderma 53, 15-29.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographic Information Systems. Oxford University Press, Oxford, 333 pp.

Cressie, N. 1991. Statistics for Spatial Data. New York, Wiley.

Diggle, P. J., Ribeiro Jr., P.J. 2007. Model-based Geostatistics. New York, Springer, 228 pp.

Eyles, G.O., 1986. *Pinus radiata* site index rankings for New Zealand. New Zealand Forestry 31, 19-22.

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, 483 pp.

Gotway, C.A., Ferguson, R.B., Hergert, G.W., Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Science Society of America Journal, 60: 1237-1247.

Goulding, C.J., 2005. "Measurement of Trees"; Section 6.5 of the NZIF Forestry Handbook, 4th Edition, Mike Colley (Ed.), NZIF. 318 pp.

Gunnarsson, F., Holm, S., Holmgren, P., Thuresson, T., 1998. On the potential of kriging for forest management planning. Scandinavian Journal of Forest Research 13, 237-245.

Höck, B.K., Payn, T.W., Shirley, J.W., 1993. Using a geographic information system and geostatistics to estimate site index of *Pinus radiata* for Kaingaroa Forest, New Zealand. New Zealand Journal of Forestry Science 23, 264-277.

Holmgren, P., Thuresson, T., 1997. Applying objectively estimated and spatially continuous forest parameters in tactical planning to obtain dynamic treatment units. Forest Science, 43, 317-326.

Hunter I.R., Gibson, A.R., 1984. Predicting *Pinus radiata* site index from environmental variables. New Zealand Journal of Forestry Science 14, 53-64.

Hunter, I.R., Rodgers, B.E., Dunningham, A., Prince, J.M., Thorn, A.J., 1991. An Atlas of Radiata Pine Nutrition in New Zealand. Forest Research Bulletin No. 165, Forest Research Institute, Rotorua, New Zealand, 24 pp.

Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press.

Jackson, D.S., Gifford, H.H., 1974. Environmental variables influencing the increment of radiata pine (1) periodic volume increment. New Zealand Journal of Forestry Science 4**,** 3-26.

Johnston, K., Ver Hoef, J. M., Krivoruchko, K., 2001. Using ArcGIS Geostatistical Analyst. ESRI Press, Redlands, CA, 300 pp.

Journel, A.G., Huijbregts, Ch.J. 1978. Mining Geostatistics, Academic Press, London, 600 pp.

Kimberley, M.O., West, G., Dean, M., Knowles, L., 2005. Site Productivity: The 300 Index – a volume productivity index for radiata pine. New Zealand Journal of Forestry 50, 13-18.

Kravchenko, A., Bullock, D.G., 1999. A comparative study of interpolation methods for mapping soil properties. Agronomy Journal, 91: 393-400.

Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K., 2002a. Land Environments of New Zealand: A technical guide. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 237 pp.

Leathwick, J., Wilson, G., Rutledge, D., Wardle, P., Morgan, F., Johnston, K., McLeod, M., Kirkpatrick, R., 2003. Land Environments of New Zealand. Ministry for the Environment, Wellington, and Manaaki Whenua Landcare Research, Hamilton, 184 pp.

Leathwick, J.R., Wilson, G., Stephens, R.T.T., 2002b. Climate surfaces for New Zealand. Landcare Research contract report: LC9798/126. Landcare Research, Hamilton, New Zealand, 22 pp.

Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W. 2005. Geographic information systems and science. Second edition, John Wiley and Sons, Chichester, 517 pp.

Lu, G.Y., Wong, D.W. 2007. An adaptive inverse distance weighting spatial interpolation technique. Computers and Geosciences, doi:10.1016/j.cageo.2007.07.010.

Mitchell, N.D., 1991. The derivation of climate surfaces for New Zealand, and their application to the bioclimatic analysis of the distribution of kauri (*Agathis australis*). Journal of the Royal Society of New Zealand 21, 13-24.

National Water and Soil Conservation Authority, 1975-70. New Zealand Land Resource Inventory Worksheets, 1: 63,360. National Water and Soil Conservation Organisation, Wellington.

Newsome, P.F.J., 1987. The Vegetative Cover of New Zealand. Soil Conservation Centre, Aokautere, Ministry of Works and Development. Water and Soil Miscellaneous Publication No. 112, Wellington, New Zealand.

Newsome, P.F.J., Wilde, R.H., Willoughby, E.J., 2000. Land resource information system spatial data layers. Landcare Research, Palmerston North, New Zealand, 84 pp.

Nikos, N., Calama, R., Montero, G., Gil, L., 2004. Geostatistical prediction of height/diameter models. Forest Ecology and Management, 195, 221-235.

NZFOA, New Zealand Forest Owners Association, 2005. New Zealand Forest Industry Facts and Figures 2004/2005. New Zealand Forest Owners Association, P.O. Box 1208, Wellington.

Odeh I.O.A., McBratney, A.B., Chittleborough, D.J., 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63, 197-214.

Odeh I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215-226.

Palmer, D.J., Höck, B.K., Dunningham, A.G., Lowe, D.J., Payn, T.W., 2008. Developing national-scale terrain attributes for New Zealand (TANZ). Scion Research Bulletin, Rotorua, New Zealand.

Samra, J.S., Gill, H.S., Bhatia, V.K., 1989. Spatial stochastic modelling of growth and forest resource evaluation. Forest Science 35, 663-676.

SAS-Institute-Inc., 2000. SAS/STAT User's Guide: Version 8. Volumes 1, 2 and 3. SAS Institute Inc., Cary, North Carolina. pp. 3884.

Schloeder, C.A, Zimmerman, N.E., Jacobs, M.J., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Science Society of America Journal, 65: 470-479.

Tait, A., Henderson, R., Turner, R., Zheng, X. 2006. Spatial interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26, 2097-2115.

Triantafilis, J., Odeh, I.O., McBratney, A.B., 2001. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. Soil Science Society of America Journal 65, 869-878.

UNFCCC, 1998, Kyoto Protocol to the United Nations Framework Convention for Climate Change, UN, Bonn.

van der Voet, H., 1994. Comparing the predictive accuracy of models using a simple randomization test. Chemometrics and Intelligent Laboratory Systems 25, 313-323.

Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science, 41: 473-490.

Watt, M.S., Coker, G., Clinton, P.W., Davis, M.R., Parfitt, R., Simcock, R., Garrett, L., Payn, T., Richardson, B., Dunningham, A., 2005. Defining sustainability of plantation forests through identification of site quality indicators influencing productivity - A national view for New Zealand. Forest Ecology and Management 216, 51-63.

Watt, M.S., Davis, M.R., Clinton, P.W., Coker, G., Ross, C., Dando, J., Parfitt, R.L., Simcock, R., 2008. Identification of key soil indicators influencing plantation productivity and sustainability across a national trial series in New Zealand. Forest Ecology and Management (in press).

Webster, R., Oliver, M.A. 1990. Statistical methods in soil and land resource survey. Oxford University Press, Oxford. 316 pp.

Webster, R., Oliver, M.A. 2007. Geostatistics for environmental sciences. Second edition, Wiley and Sons, 317 pp.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares. In P. R. Krishnaiah (Eds.), Multivariate Analysis. New York, Academic Press, pp. 391-420.

Wold, S., 1994. PLS for Multivariate Linear Modeling, QSAR: Chemometric Methods in Molecular Design. Methods and Principles in Medicinal Chemistry, (Eds.), H. van de Waterbeemd, Weinheim, Germany, Verlag-Chemie.

Woollons, R.C., Skinner, M.F., Richardson, M.F., Rijske, W.C., 2002. Utility of "A" horizon soil characteristics to separate pedological groupings, and their influence with climatic and topographic variables on *Pinus radiata* height growth. New Zealand Journal of Forestry Science 32, 195-207.

Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P., 1999. An experimental comparison of ordinary and universal kriging and inverse distance weighting. Mathematical Geology 31, 375-390.

Fox Glacier viewed from the Lake Matheson Road (photograph by Dr Kyle Bland)

Say, from whence you owe this strange intelligence?

William Shakespeare, *Macbeth*

CHAPTER SEVEN

THESIS DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

1.0 BACKGROUND

The interest in mapping forest productivity has dramatically increased since the inception of this project. Initially, the forest industry focussed on defining its productive capacity and sustainability, but more recently there has been a shift towards the issue of carbon sequestration. Nevertheless, the rationale behind this project and the development of models and associated maps is a deeper understanding of the spatial distribution of *Pinus radiata* productivity across the national extent. In New Zealand, two main estimates of *P. radiata* productivity are used. The first is height, defined as the mean top height at age twenty (Goulding, 2005) (Site Index), and the second is volume, defined as the stem volume mean annual increment at age thirty years with a reference regime of 300 stems ha^{-1} (Kimberley et al., 2005) (300 Index). Previous models for New Zealand (Jackson and Gilford, 1974; Hunter and Gibson, 1984; Woollons et al., 2002; Watt et al., 2005, 2008) provided insight into the main environmental contributors towards *P. radiata* productivity. Despite this, maps were never developed because the 'hard data' used in developing the models were site specific. The modelling and mapping in my study took a different approach, utilising GIS layers, interpolated surfaces, and ancillary maps often referred to as 'soft information' (Burrough and McDonnell, 1998). Information related to

the soil nutrient and soil moisture status is not fully represented in New Zealand national datasets. Therefore, terrain attributes for New Zealand (TANZ) and soil water balance (SWatBal) were developed as essential supporting information for the modelling and mapping of *P. radiata* forest productivity across New Zealand.

Chapters 2 and 3 of this thesis described the development of a model of terrain attributes for New Zealand (TANZ) and the development of a soil water balance model (SWatBal). Conclusions are therefore largely technical in nature, providing details of model advancements as well as visualisations of mapped surfaces. Chapters three to six discussed the development of *P. radiata* 300 Index and Site Index models utilising data developed in chapters 2 and 3 in conjunction with other nationally available GIS layers, surfaces, and ancillary maps.

2.0 OVERVIEW AND SYNTHESIS OF RESULTS

2.1 Development of terrain attributes for New Zealand (TANZ)

The main results from Chapter 2 were the development of high quality primary and secondary terrain attributes at a 25-m cell resolution across New Zealand. The macro-catchment concept was designed to divide New Zealand into a series of large, naturally draining areas that consist of numerous catchments and sub-catchments representing the natural drainage patterns for that region. The maximum size of the macro-catchments was chosen so that the computer processor and storage system was able to manage the

magnitude of each data set. The primary terrain attributes developed were elevation (inclusive of spurious sink identification and removal) (Figure 1), flow direction, upslope contributing area, flow width, slope as a percentage, radians, and degrees, aspect, curvature (profile, plan, and tangent), flow path length, and rate of change of specific catchment area along flow path. Secondary terrain attributes developed included topographic wetness index, stream power index, the length and slope factors from the revised universal soil loss equation, and 12 monthly and one annual clear-sky shortwave radiation surfaces (topographically adjusted).

The TANZ model provides an advance not only in developing comprehensive national-scale terrain attributes across New Zealand, but also in developing surfaces with the highest possible degree of accuracy. The advancements in the modelling process included sink identification and correction, the removal of edge effects around macro-catchment boundaries, and modelling of flow convergence and divergence. The removal of spurious sinks, while retaining natural sinks and depressional features that impede overland flow, was undertaken by firstly identifying naturally occurring sinks across New Zealand and then eliminating spurious sinks using the technique of Jensen and Domingue (1988). Edge effects were eliminated by extending terrain attributes 10 cells beyond the macro-catchment boundary. On completion of the terrain modelling process, the surplus cells that may contain inconsistencies around the edges were discarded. A further advance in the development of terrain attributes was the recognition of flow divergence and convergence in the modelling process. In upland areas, the FD8 algorithm allows flow to

be distributed to multiple neighbouring cells, thus allowing flow convergence to be modelled. In contrast, the D8 method enables flow divergence to be captured in valleys.

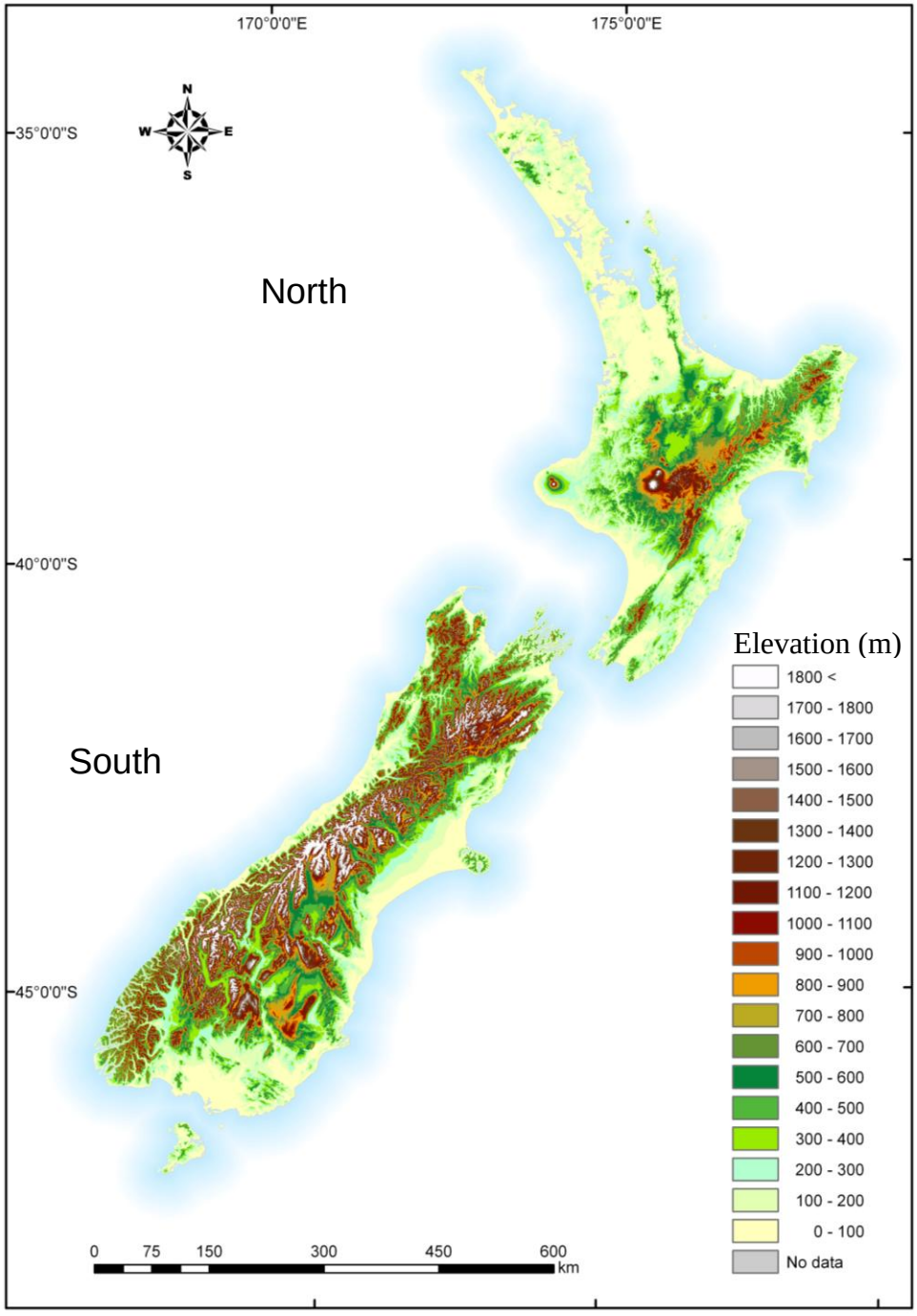


Figure 1: Spatial distribution of elevation across New Zealand (elevation in metres above sea-level) (data derived from Barringer et al., 2002).

2.2 Development of a soil water balance model for New Zealand (SWatBal)

Chapter three described how monthly soil water balance surfaces were developed across New Zealand to assess the spatial relationships between soil moisture and *P. radiata* productivity. Soil water balance has long been recognised as a major contributor to forest productivity and much research has been undertaken at locations exhibiting seasonal water deficits (Arneth et al., 1998a, 1998b; McMurtrie et al., 1990; Richardson et al., 2002; Watt et al., 2003, 2008). The soil water balance model developed was designed with a dynamic framework capable of upgrading as new and improved data become available. Furthermore, SWatBal is flexible, accommodating soil water balance simulations at a variety of user defined resolutions, geographic extents, and temporal or spatial scales as required. The model include 12 monthly surfaces and one annual surface for mean fraction of available root-zone water storage (W_f), mean available root-zone water storage (W_a), and mean drainage. Figure 2 provides an example of the mean annual fraction of available root-zone water storage (W_f) across New Zealand. A critical attribute of SWatBal's development was the implementation of 'reasoned and allocated virtual data' (*RAV*). Prior to SWatBal, New Zealand spatial soil water balance models assumed constant rainfall distribution across each day in a month. *RAV* monthly rainfall distribution was developed by choosing a representative month from ~ 40 years of rainfall records (Tait et al., 2006).

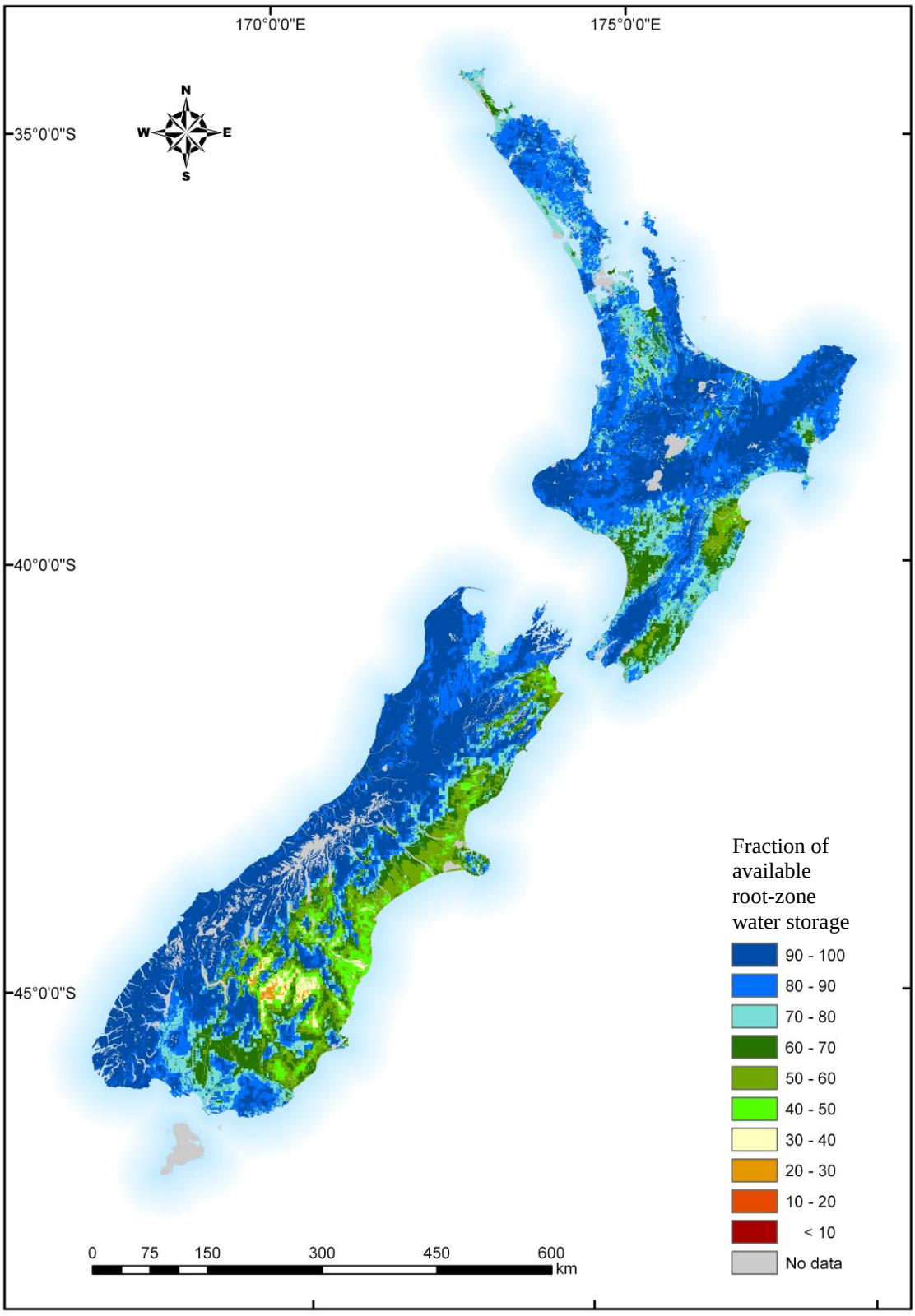


Figure 2: Spatial distribution of annual mean fraction of available root-zone water storage (W_f) across New Zealand.

The selection criteria were (1) the number of days it rains in a month, and (2) the total monthly rainfall for the representative month closest to the 40 year average. Overall, SWatBal modelled surfaces not only allow direct comparisons of the soil moisture status across New Zealand but also allow the visualisation of spatial patterns at the required level of detail for users.

2.3 Application of soil water balance in modelling forest productivity

In Chapter 4 it was found that a substantial improvement could be made to model predictions of *Pinus radiata* mean annual increment volume (MAI_v) through the use of available root-zone water content W_a (SWatBal, Chapter 3). W_a, in combination with annual temperature (T_a), accounted for 61% of the variance in *P. Radiata*, MAI_v. The main implication of this chapter is to endorse a move towards relatively low cost, long-term normalised spatial climate data as an alternative to measuring and collecting site specific climate data, which are high cost. Although it is recognised that some essential environmental site specific data will still need to be collected, the utility of SWatBal and its associated data simplifies soil water balance calculations considerably whilst retaining similar accuracy.

2.4 Predicting the spatial distribution of *Pinus radiata* productivity across New Zealand using interpolated surfaces and ancillary maps

Chapter five described the use of data developed in chapters 2 and 3, in conjunction with interpolated surfaces and ancillary maps, for the development of *P. radiata* productivity surfaces for New Zealand. The main purpose of this work was to develop models and maps for forest managers and scientists and thus tools for exploring the spatial relationships of *P. radiata* productivity across New Zealand. Partial least squares (PLS) regression was used in the development of mean annual increment volume (300 Index) and mean top height at age 20 (Site Index) by correlation with data derived from national extent ancillary maps and interpolated surfaces. The best forestry maps developed explained 58% and 67% of the variance for the 300 Index and Site Index, respectively. A cross validation procedure withheld 618 out of 1764 sites from the modelling process to enable a comparison between measured and modelled forest productivity. The findings showed that regression kriging (RK) substantially improved PLS regression predictions, increasing R^2 values by 12 % and 13 % for Site Index and 300 Index, respectively (Figure 3). The root mean square error (RMSE) and goodness of precision (G) cross validation statistics confirmed these results, with RK showing significant improvement for both the 300 Index and Site Index validation predictions. Where possible, contributors of *P. radiata* productivity were assessed across New Zealand environmental regions (derived from Land Environments of New Zealand, LENZ). These contributions highlighted the importance of air temperature, rainfall, soil water status and profile properties in both productivity models.

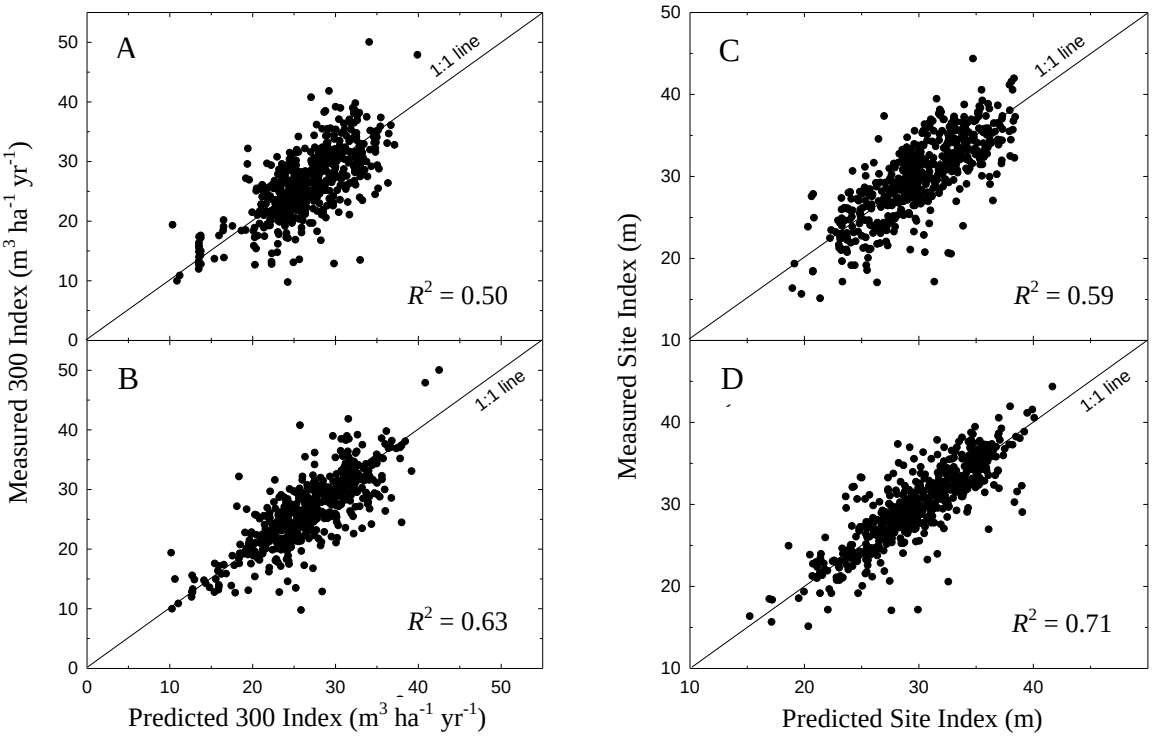


Figure 3. Relationships between predicted and measured 300 Index for (A) PLS and (B) RK, and Site Index for (C) PLS and (D) RK models. The 1:1 line is shown (n = 552).

2.5 Comparing the interpolation prediction techniques for *Pinus radiata* productivity across New Zealand

The development of the 300 Index and Site Index maps for New Zealand using multiple linear regression in association with national extent ancillary maps and interpolated surfaces is not without substantial cost. Therefore a comparison of the interpolation techniques used in chapter 5 (PLS and RK) with the comparatively cheaper and easy-to-apply modelling techniques, ordinary kriging (OK), and inverse distance weighting (IDW), is warranted. Visually, the PLS and RK interpolation techniques provided maps

with the greatest spatial detail. This is because multiple linear regression models implicitly disregard spatial covariation, with the mapped predictions being derived from statistical relationships developed from underlying datasets related to variables (for example, terrain and climate). Contrary to the visually detailed PLS maps, the cross validation statistics showed that OK, IDW, and RK outperformed the PLS prediction technique. In addition, the cross validation statistics indicated that OK, IDW, and RK predicted at similar levels of precision and bias. However, OK and IDW maps failed to predict some obvious features. For example, declining levels of productivity at elevated cooler regions where sample observations were sparse (Figure 4) were not predicted. Sample observations of *Pinus radiata* productivity in this study were limited to locations currently forested and therefore observations were biased towards higher sample observation densities. Consequently, at distal locations OK and IDW are unlikely to perform well because of their reliance on observations at the optimal operational scale. The main implications of these comparisons are that PLS provide detailed information globally across all spatial locations. In contrast, RK provided improved predictions at locations where observational densities remain relatively high. Therefore the combined effect (summation) of PLS and its kriged residuals provided predictions equivalent to, if not better than, those derived using OK or IDW.

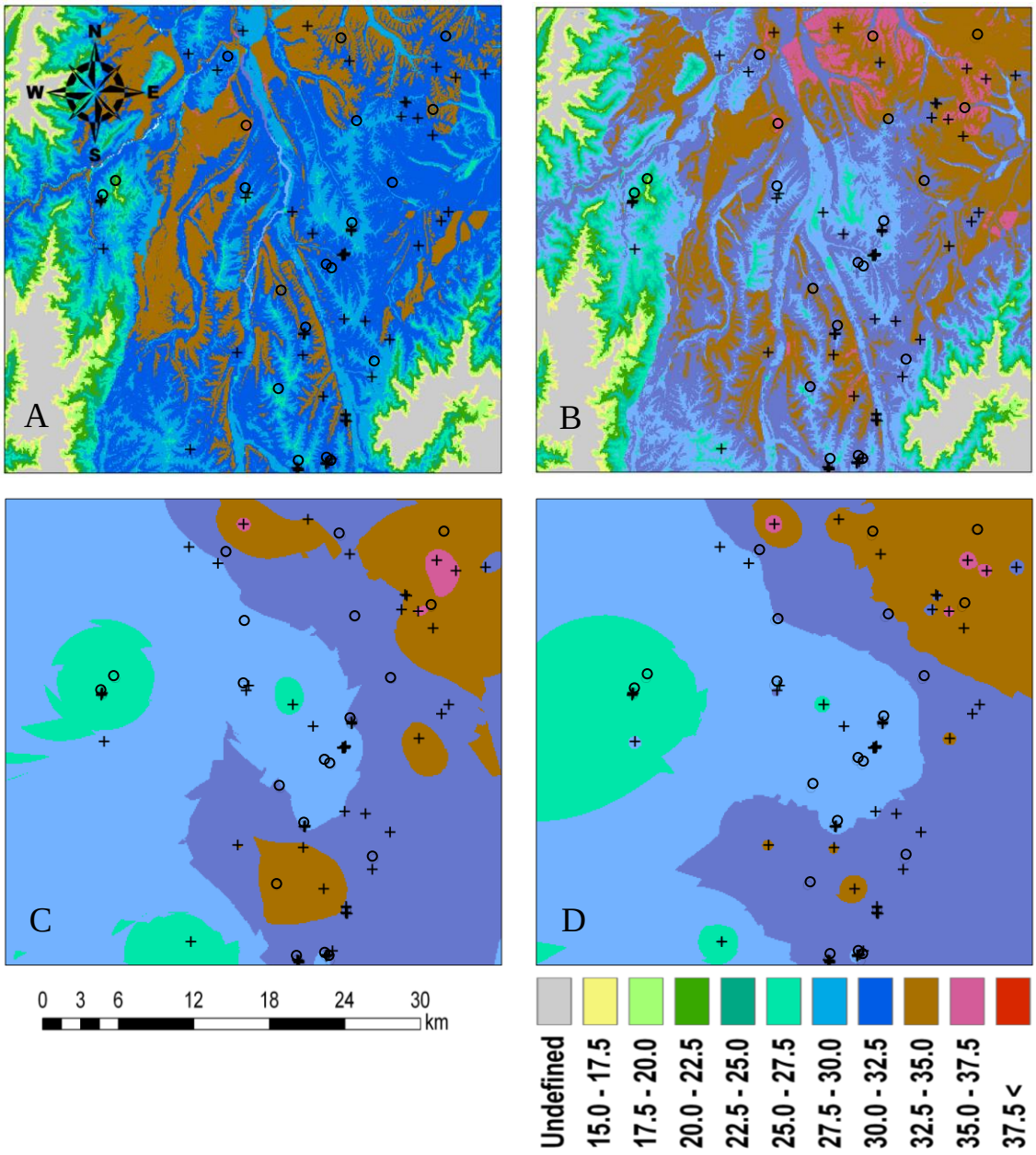


Figure 4: Comparison of (A) partial least squares, (B) regression kriging, (C) ordinary kriging, and (D) inverse distance weighting interpolation maps and the impact on spatial variation when productivity observations are dense (central graphic) and sparse (elevated grey areas) for Site Index in the Nelson area. Crosses are observed data locations, whereas circles are validation locations withheld from the modelling processes.

2.6 Summary

Maps describing the spatial variability across New Zealand's national extent for *Pinus radiata* productivity using ancillary data have never been produced before. In this thesis, high quality national datasets were developed for terrain attributes and soil water balance to bolster available climate and environmental datasets. Multiple linear regression modelling was used in association with these datasets to develop empirically based models for measures of *P. radiata* productivity, namely 300 Index and Site Index values. Using cross validation procedures, the model precision and bias (in conjunction with knowledge of variance explained for each model) defined the certainty applying to the models and maps. Furthermore, it was found that the implementation of regression kriging improved predictions at validation locations. The implications of this research are to use PLS maps where productivity observations are remote or sparse. Conversely, RK models are considered better predictors of productivity at the compartment level where higher density observations occur.

3.0 CONCLUSIONS

The main conclusions of the research relating to the specific objectives given in Chapter 1 (Section 1.3) are as follows.

1. Primary and secondary terrain attributes were developed for New Zealand at a 25-m grid cell resolution for the purpose of assessing the relationships between terrain attributes and *P. radiata* productivity. These terrain attributes are currently the most comprehensive and highest quality digital surfaces available for New Zealand because algorithms used in their development modelled flow divergence and convergence while accounting for naturally occurring sinks.

2. SWatBal monthly soil water balance surfaces were developed for the exploration of the spatial relationships between *P. radiata* and soil moisture. A major advance of the SWatBal model is the development and use of 'reasoned and allocated virtual data' rainfall data, RAV, where rainfall is distributed across each day of a month, simulating naturally occurring rainfall events across New Zealand. The SWatBal model also allows for updates as new and improved data become available. Another flexible feature of the SWatBal model is the ability of the user to define the geographic extent, temporal or spatial scale, and the resolution of a specific project.

3. It was clearly demonstrated that the variable available root-zone water storage modelled using SWatBal in association with field measured temperature (average annual) provided better predictions of *P. radiata* mean annual volume increment (MAI_v) than the variable rainfall used historically. The development of SWatBal allows a move towards relatively low cost, long-term normalised spatial climate GIS surfaces in preference to the high cost of measuring and collecting data at each individual study location.

4. Partial least squares (PLS) regression was used to explain 58% and 67% of the variance in 300 Index and Site Index, respectively. The regression kriging technique (RK) was used to further improve model predictions through ordinary kriging (OK) of PLS regression residuals and summation with PLS predictions. Cross validation using 618 out of the 1764 sites showed a substantial improvement in prediction precision and bias for Site Index. However, RK only provided improvement to prediction precision for the 300 Index, not prediction bias.

5. Statistical models were used to develop national extent *P. radiata* 300 Index and Site Index maps using PLS regression models and RK models. Cross validation sites were considered biased toward locations currently forested, increasing the chance of high density of observations. Although RK generally provided improved predictions at validation sites, and because of concerns about observation density, the PLS models were considered better global estimates

(national spatial extent) of forest productivity at locations where observational data are sparse. However, at the forest compartment level where the observational density of data may increase, RK models should be considered for improved predictions of the 300 Index and Site Index surfaces.

6. The spatial interpolation techniques inverse distance weighting (IDW), ordinary kriging (OK), partial least squares regression (PLS), and regression kriging (RK), were compared using cross validation for the prediction of the 300 Index and Site Index across New Zealand. Findings showed that OK, IDW, and RK techniques provided predictions at similar levels of precision. However, OK and IDW failed to show important features because of sparse observational densities (as previously stated). Therefore, where observations are sparse and scattered, PLS regression are the most useful assessment tools. If the purpose of the map derived from these interpolation methods is simply a national overview of general forest productivity trends then IDW provides the greatest cost advantage. Conversely, PLS and RK, depending on the level of interest (spatial extent), become important tools were the assessment requires greater certainty.

4.0 REFERENCES

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998a. Assessment of annual carbon exchange in a water-stressed *Pinus radiata* plantation: an analysis based on eddy covariance measurements and an integrated biophysical model. Global Change Biology 5, 531-545.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998b. Fluxes of carbon and water in a *Pinus radiata* forest subject to soil water deficit. Australian Journal of Plant Physiology 25, 557-570.

Barringer, J.R.F., Pairman, D., McNeill, S.J. 2002, Development of a high-resolution digital elevation model for New Zealand. Landcare Research Contract Report LC0102/170 (unpublished).

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographic Information Systems. Oxford University Press, Oxford, 333 pp.

Goulding, C.J., 2005. "Measurement of Trees"; Section 6.5 of the NZIF Forestry Handbook, 4th Edition, Mike Colley (Ed.), NZIF. 318 pp.

Hunter I.R., Gibson, A.R., 1984. Predicting *Pinus radiata* site index from environmental variables. New Zealand Journal of Forestry Science 14, 53-64.

Jackson, D.S., Gifford, H.H., 1974. Environmental variables influencing the increment of radiata pine (1) periodic volume increment. New Zealand Journal of Forestry Science 4**,** 3-26.

Jensen, S.K., Domingue, J.O. 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. Photogrammetric Engineering and Remote Sensing 54, 241-244.

Kimberley, M.O., West, G., Dean, M., Knowles, L., 2005. Site Productivity: The 300 Index - a volume productivity index for radiata pine. New Zealand Journal of Forestry 50, 13-18.

McMurtrie, R.E., Rook, D.A., Kelliher, F.M., 1990. Modelling the yield of *Pinus radiata* on a site limited by water and nitrogen. Forest Ecology and Management 30, 381-413.

Richardson, B., Whitehead, D., McCracken, I.J., 2002. Root-zone water storage and growth of *Pinus radiata* in the presence of a broom understorey. New Zealand Journal of Forestry Science 32, 208-220.

Tait, A., Henderson, R., Turner, R., Zheng, X. 2006. Spatial interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26, 2097-2115.

Watt, M.S., Coker, G., Clinton, P.W., Davis, M.R., Parfitt, R., Simcock, R., Garrett, L., Payn, T., Richardson, B., Dunningham, A., 2005. Defining sustainability of plantation forests through identification of site quality indicators influencing productivity - A national view for New Zealand. Forest Ecology and Management 216, 51-63.

Watt, M.S., Davis, M.R., Clinton, P.W., Coker, G., Ross, C., Dando, J., Parfitt, R.L., Simcock, R., 2008. Identification of key soil indicators influencing plantation productivity and sustainability across a national trial series in New Zealand. Forest Ecology and Management (in press).

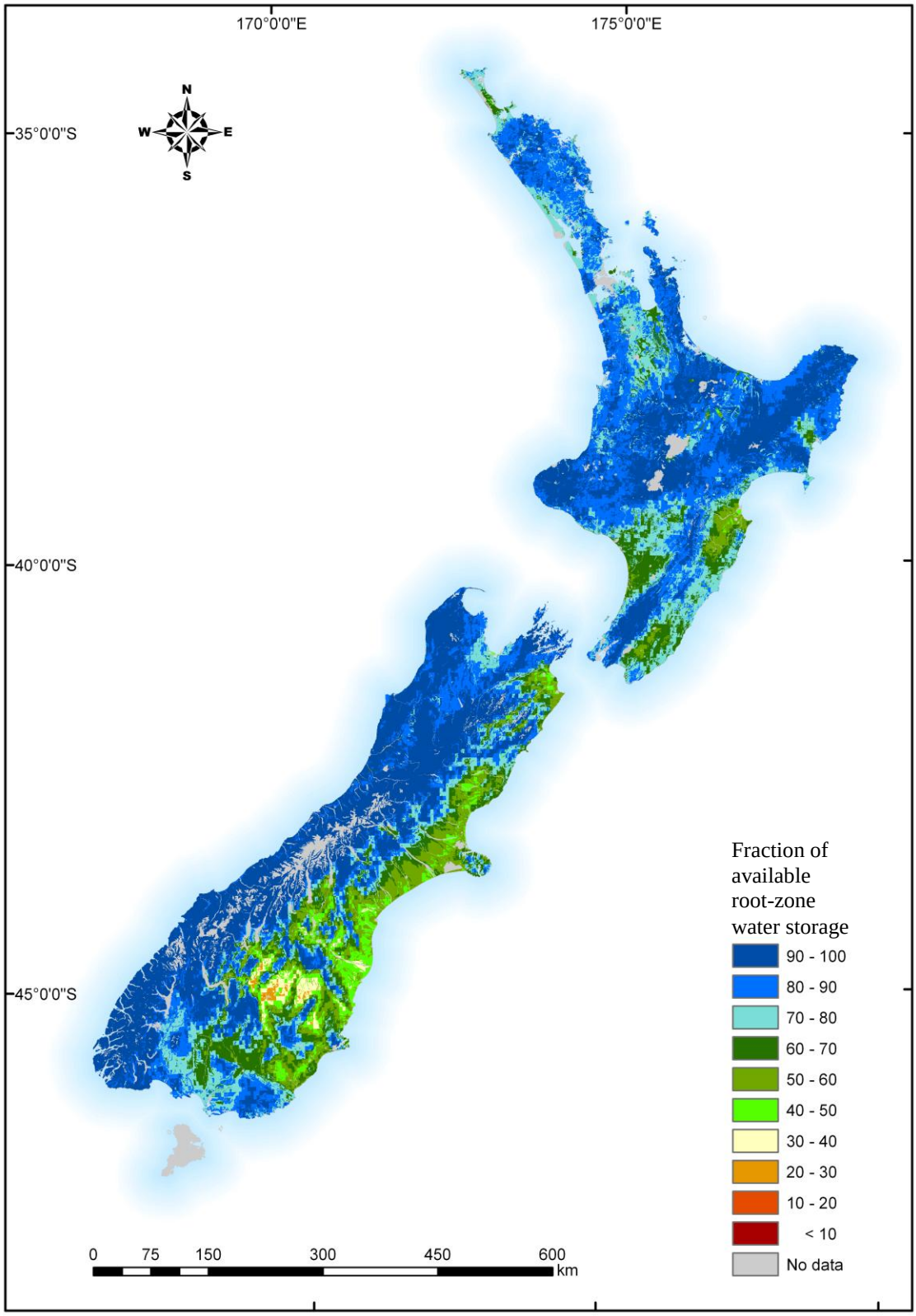
Watt, M.S., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C., 2003. Modelling the influence of weed competition on the growth of young *Pinus radiata* at a dryland site. Forest Ecology and Management 178, 271-286.

Woollons, R.C., Skinner, M.F., Richardson, M.F., Rijske, W.C., 2002. Utility of "A" horizon soil characteristics to separate pedological groupings, and their influence with climatic and topographic variables on *Pinus radiata* height growth. New Zealand Journal of Forestry Science 32, 195-207.

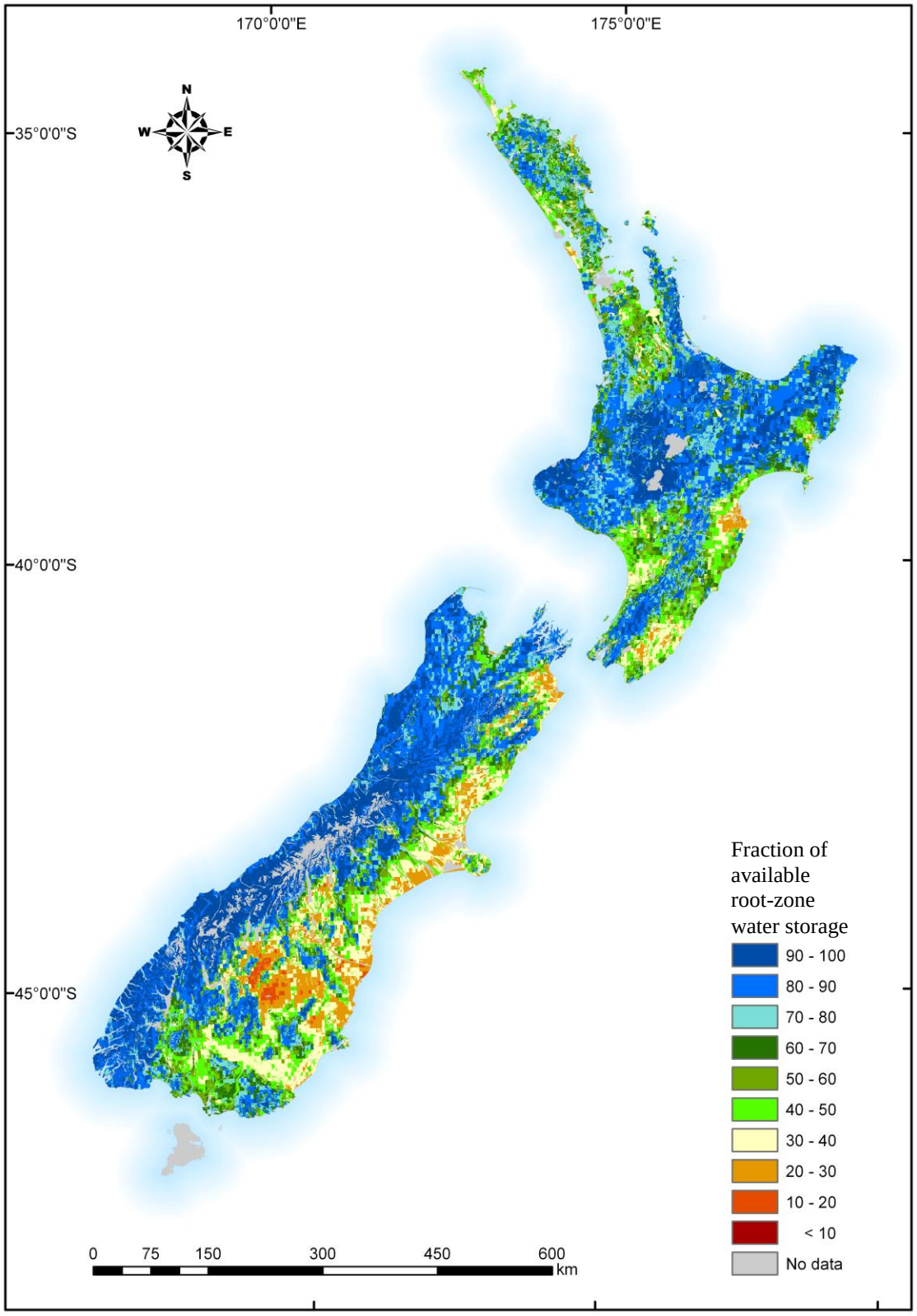
Mohaka River near State Highway 5, looking towards the Crohane Forest, western Hawke's Bay
(photograph by Dr Kyle Bland)

APPENDICES

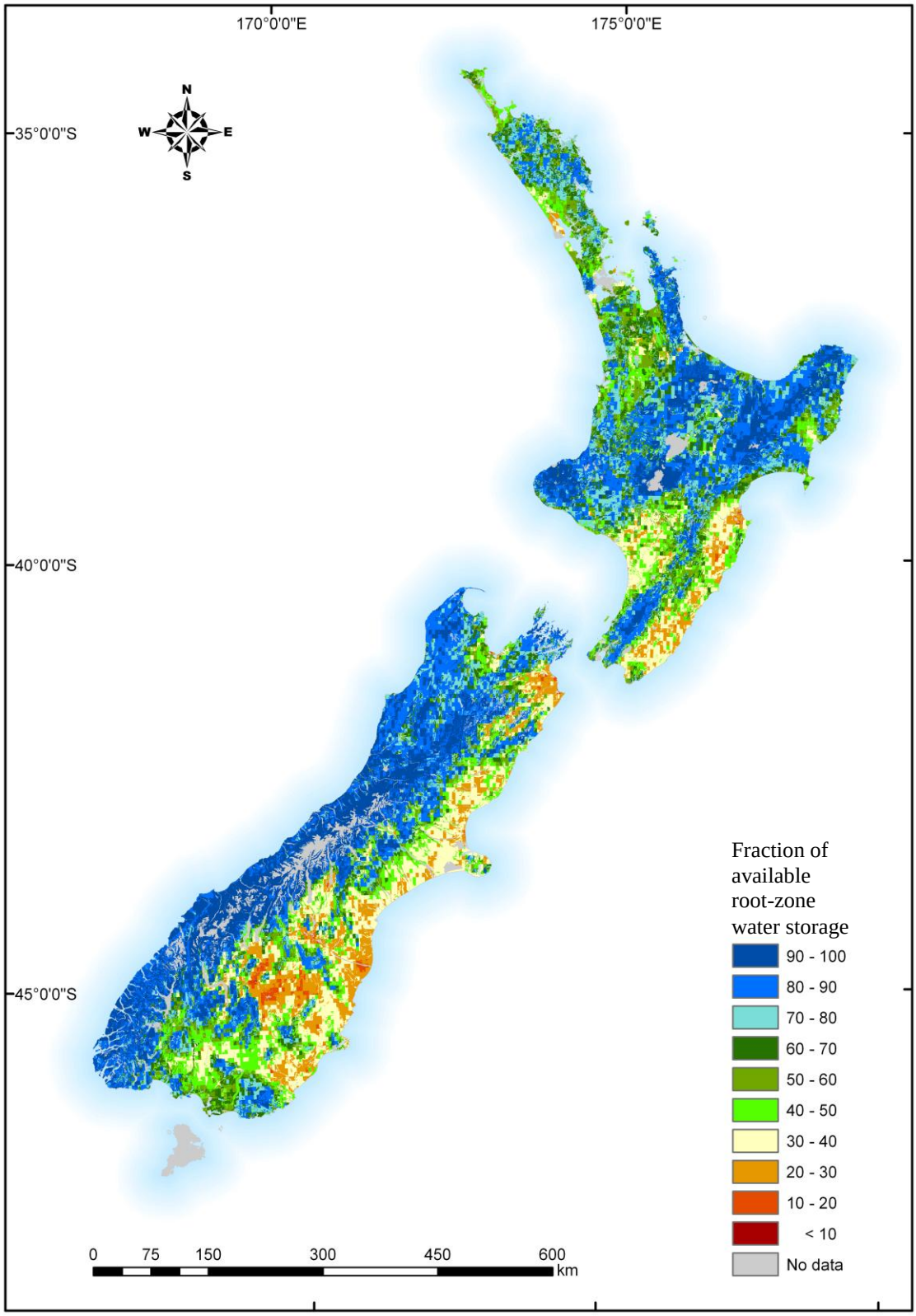
Appendix Ia: Annual mean fraction of available root-zone water storage (W_f)



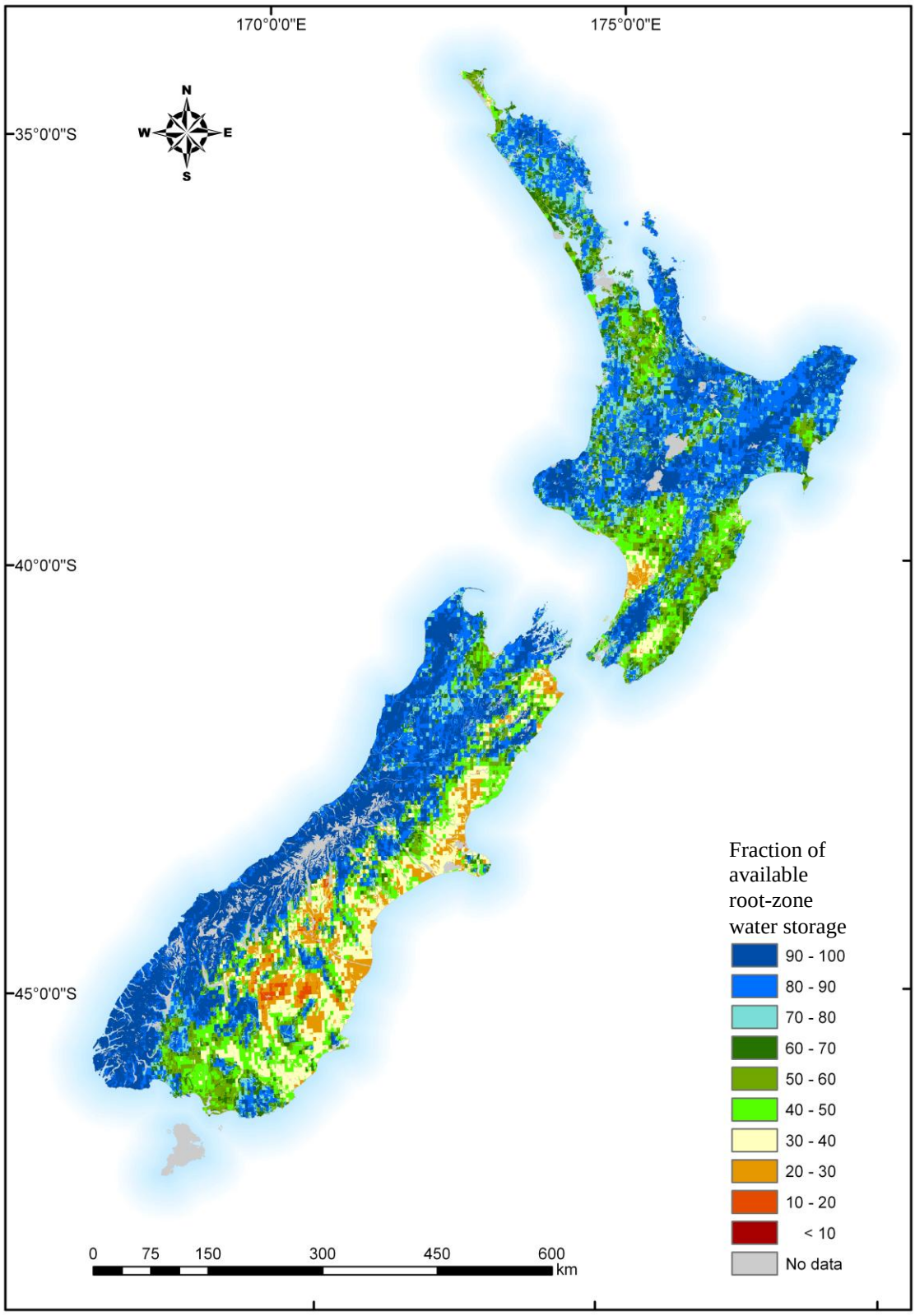
Appendix Ib: January monthly mean fraction of available root-zone water storage (W_f)



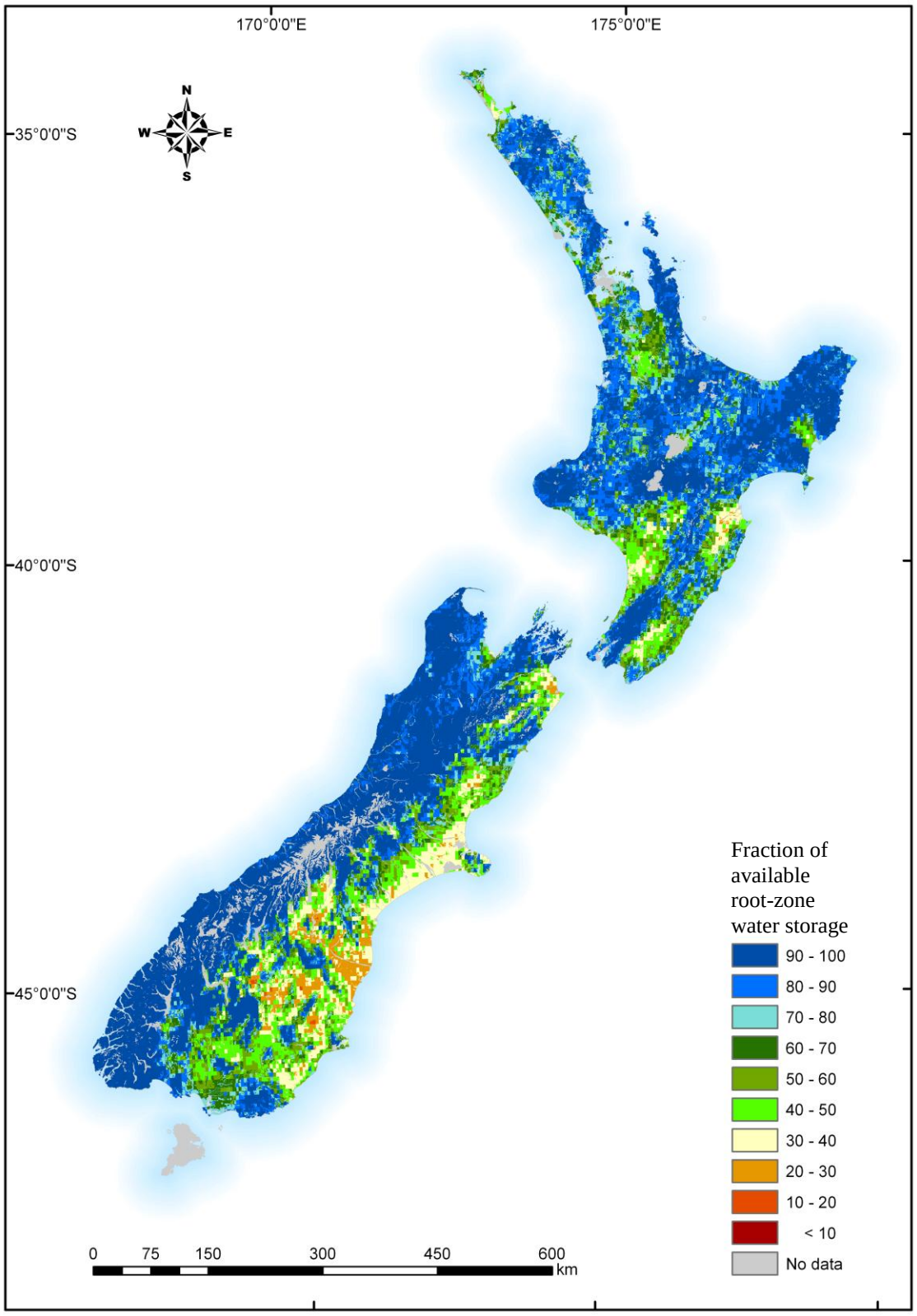
Appendix Ic: February monthly mean fraction of available root-zone water storage (W_f)



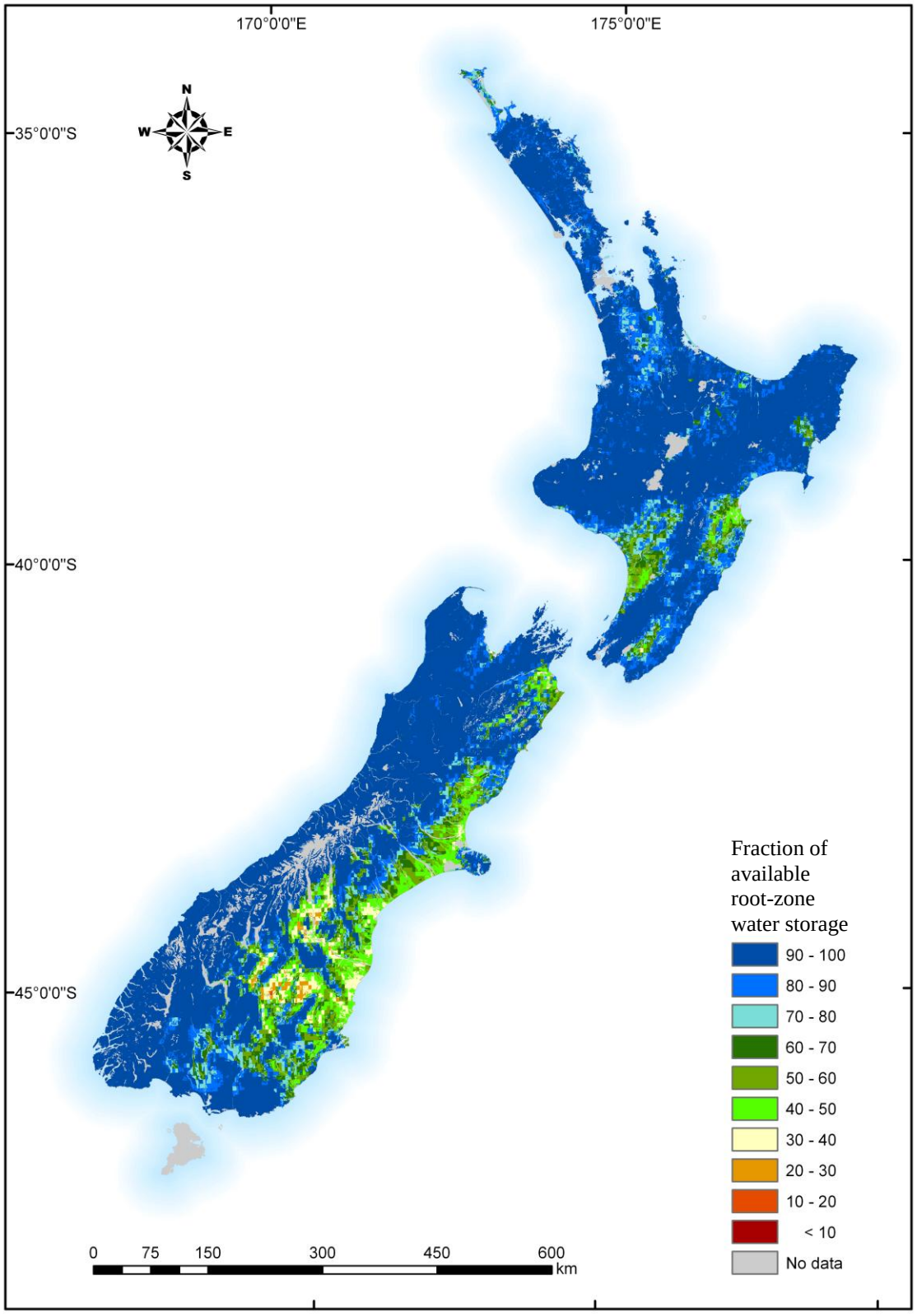
Appendix Id: March monthly mean fraction of available root-zone water storage (W_f)



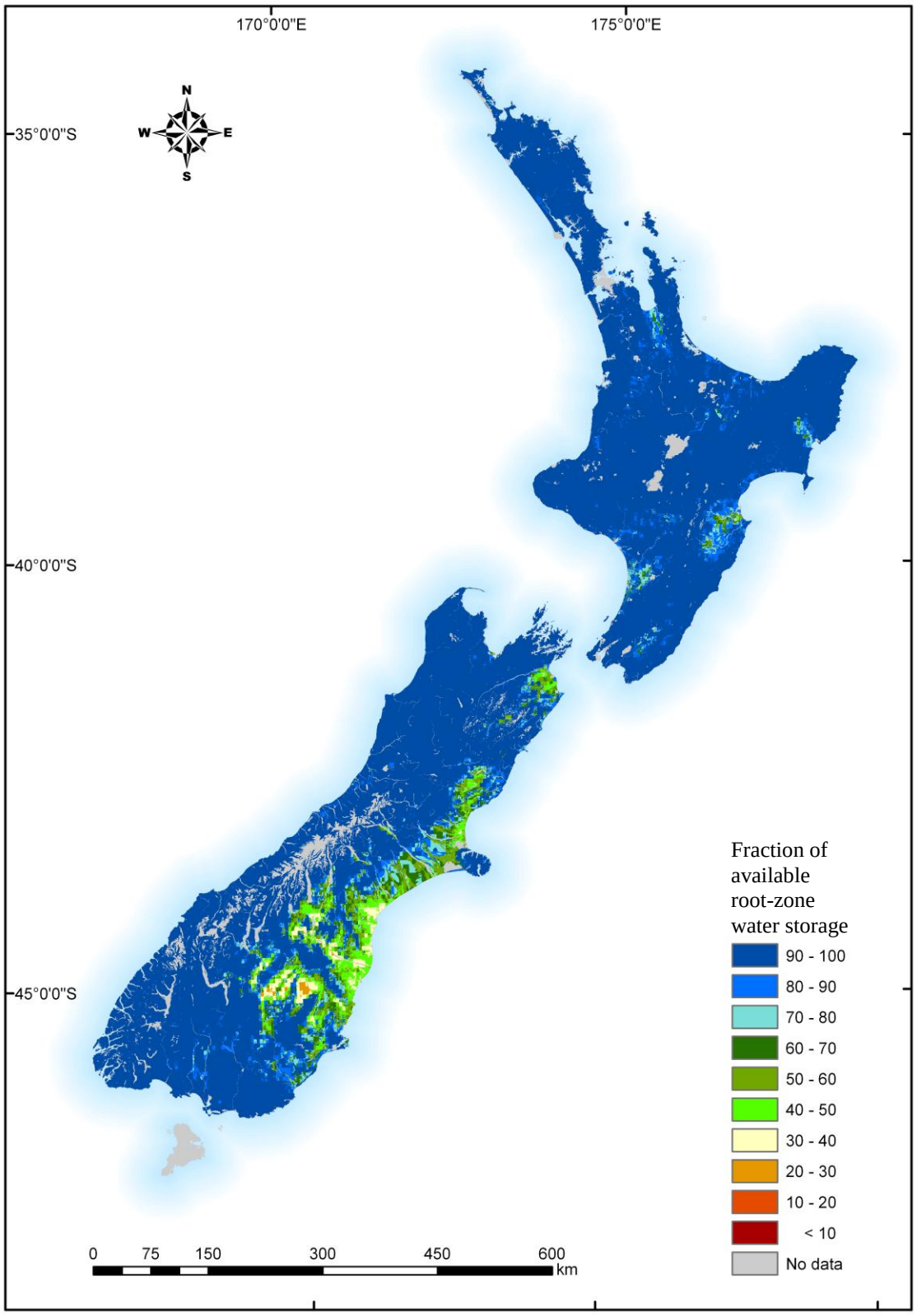
Appendix Ie: April monthly mean fraction of available root-zone water storage (W_f)



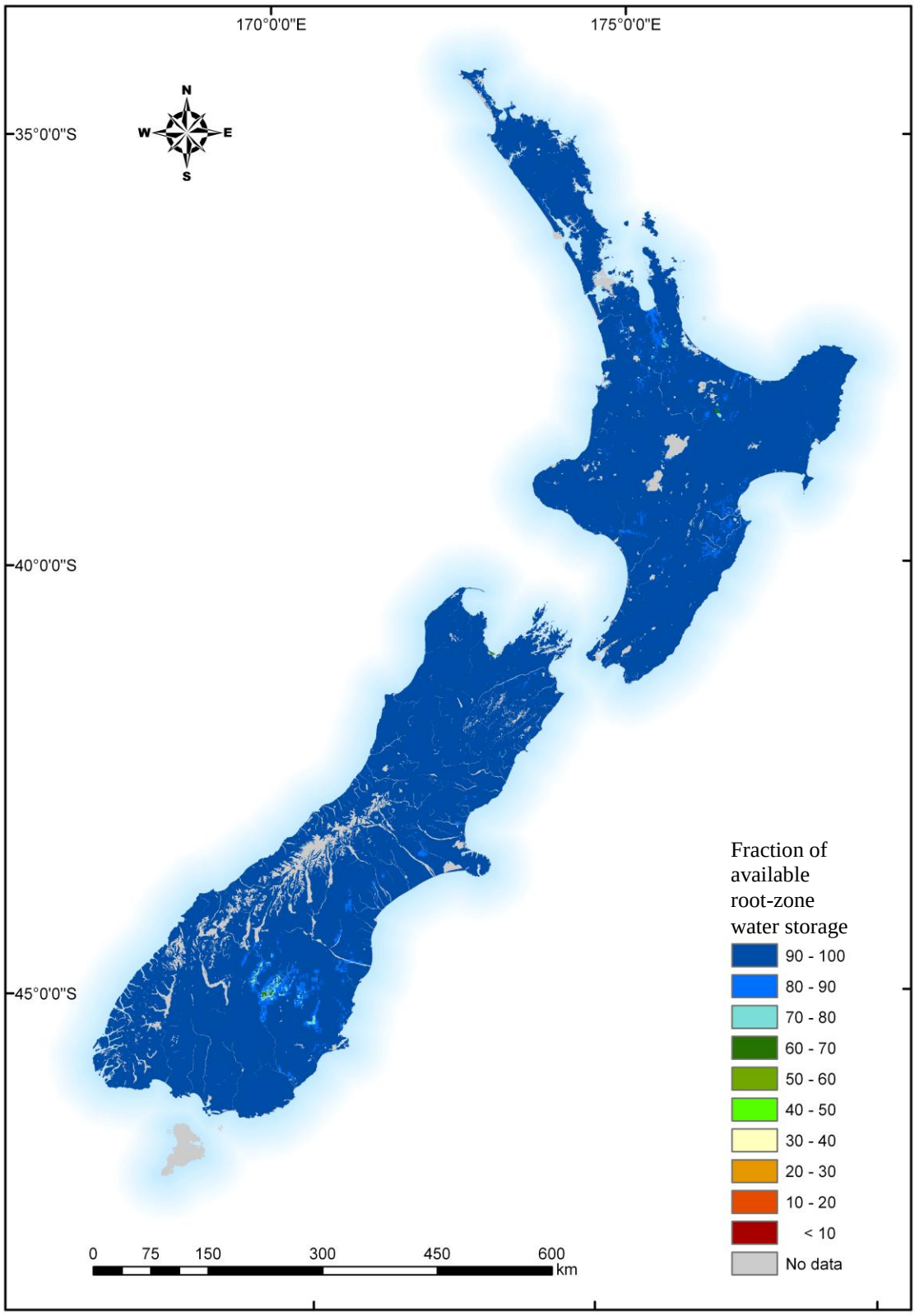
Appendix If: May monthly mean fraction of available root-zone water storage (W_f)



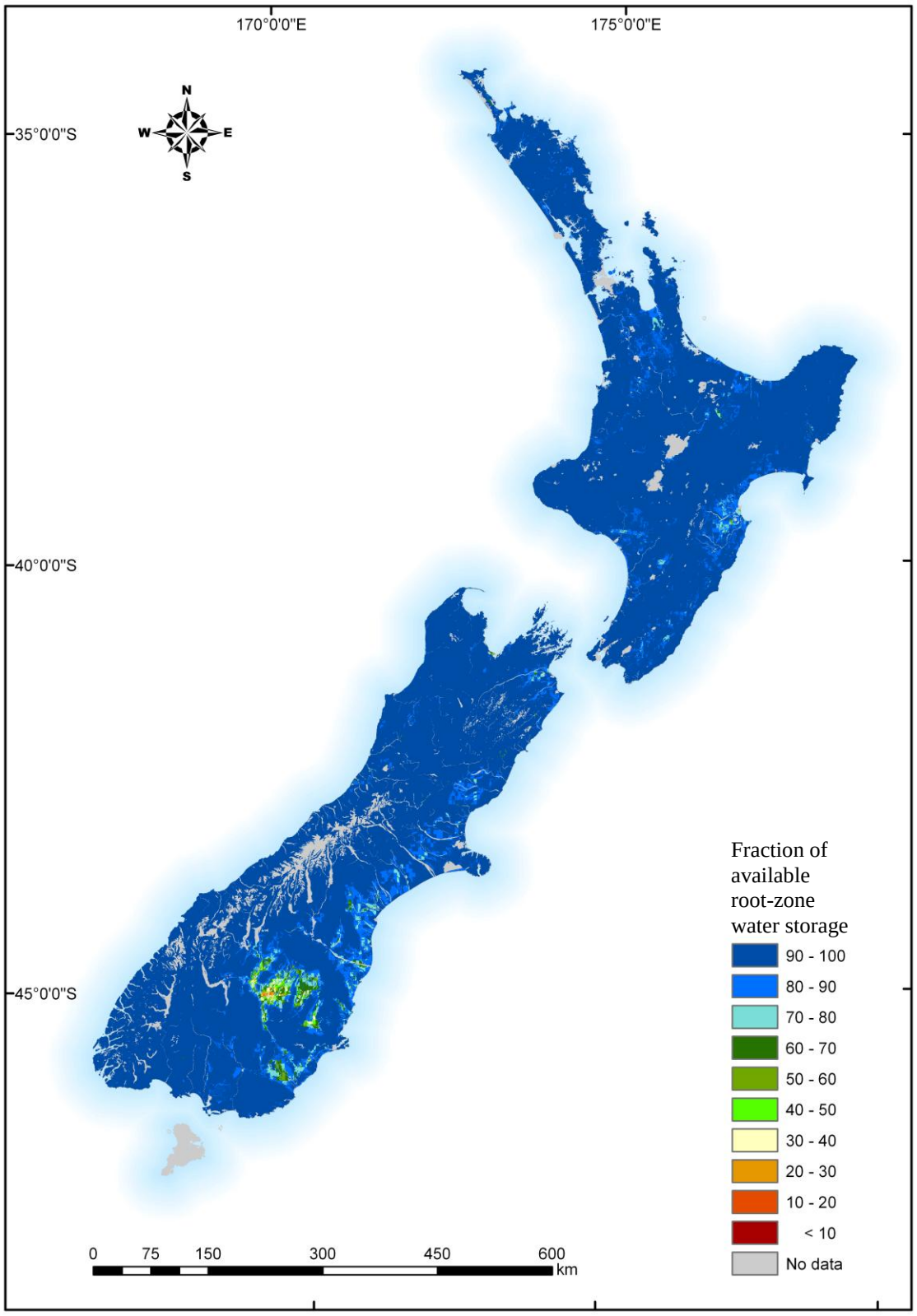
Appendix Ig: June monthly mean fraction of available root-zone water storage (W_f)



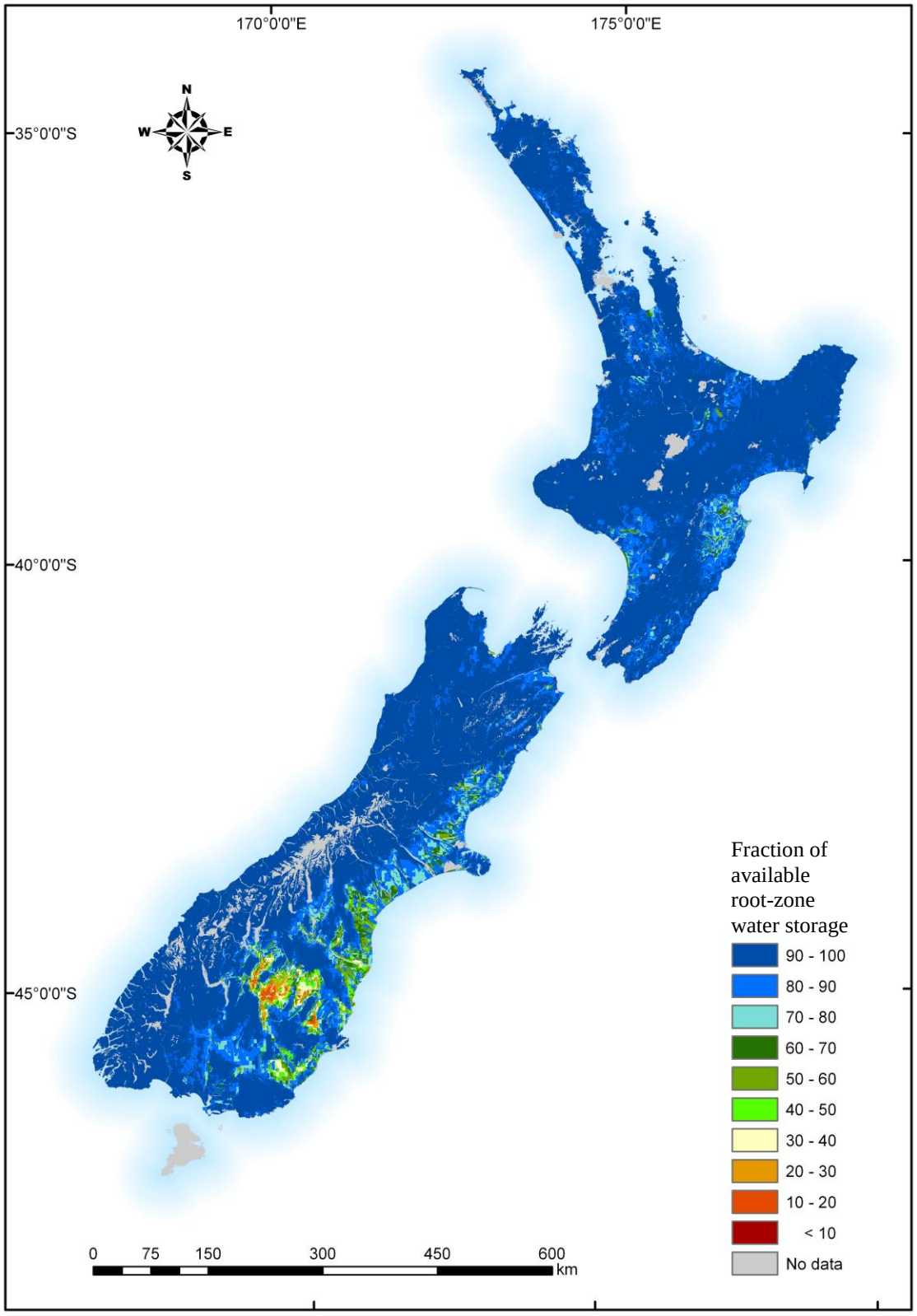
Appendix Ih: July monthly mean fraction of available root-zone water storage (W_f)



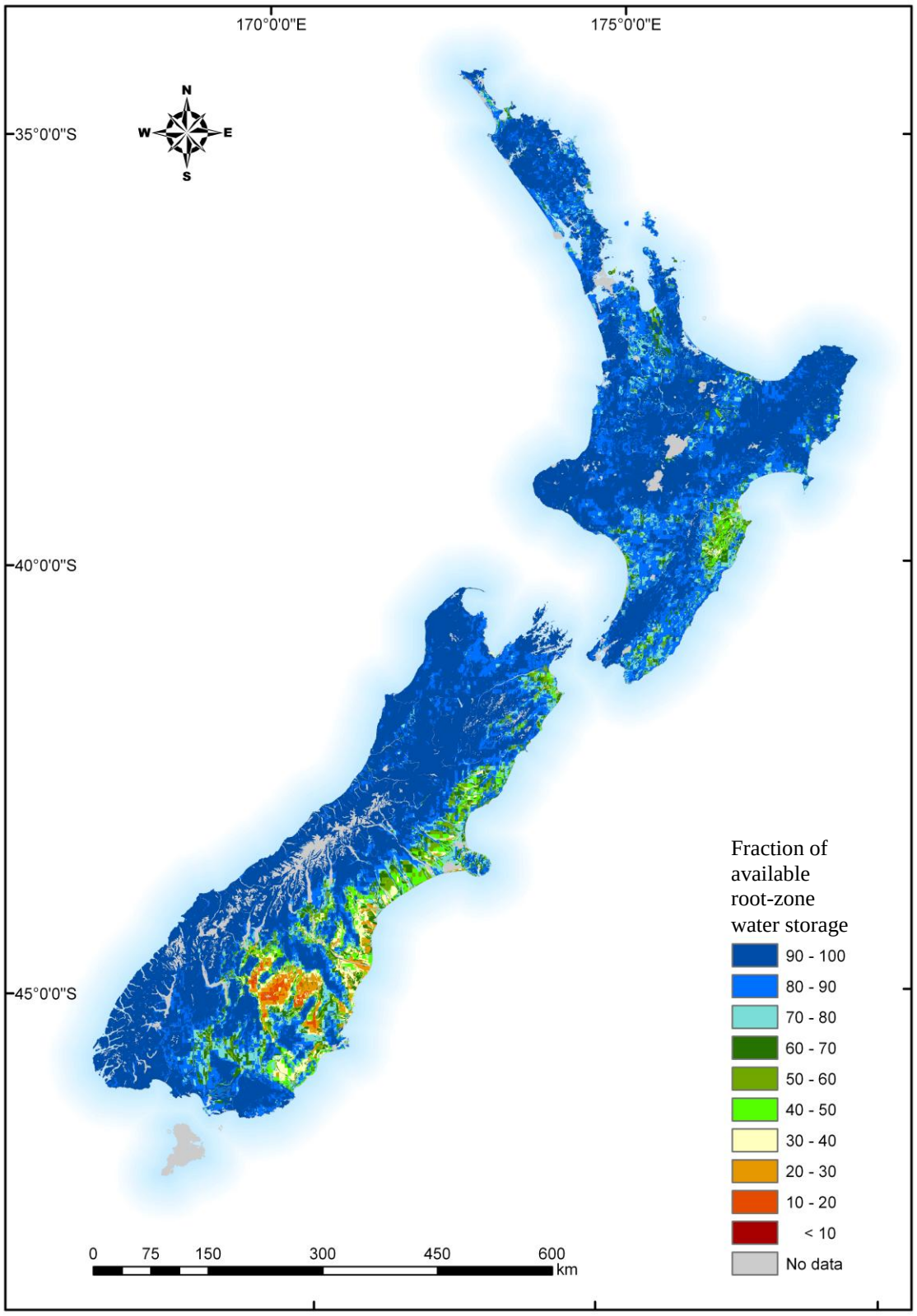
Appendix III: August monthly mean fraction of available root-zone water storage (W_f)



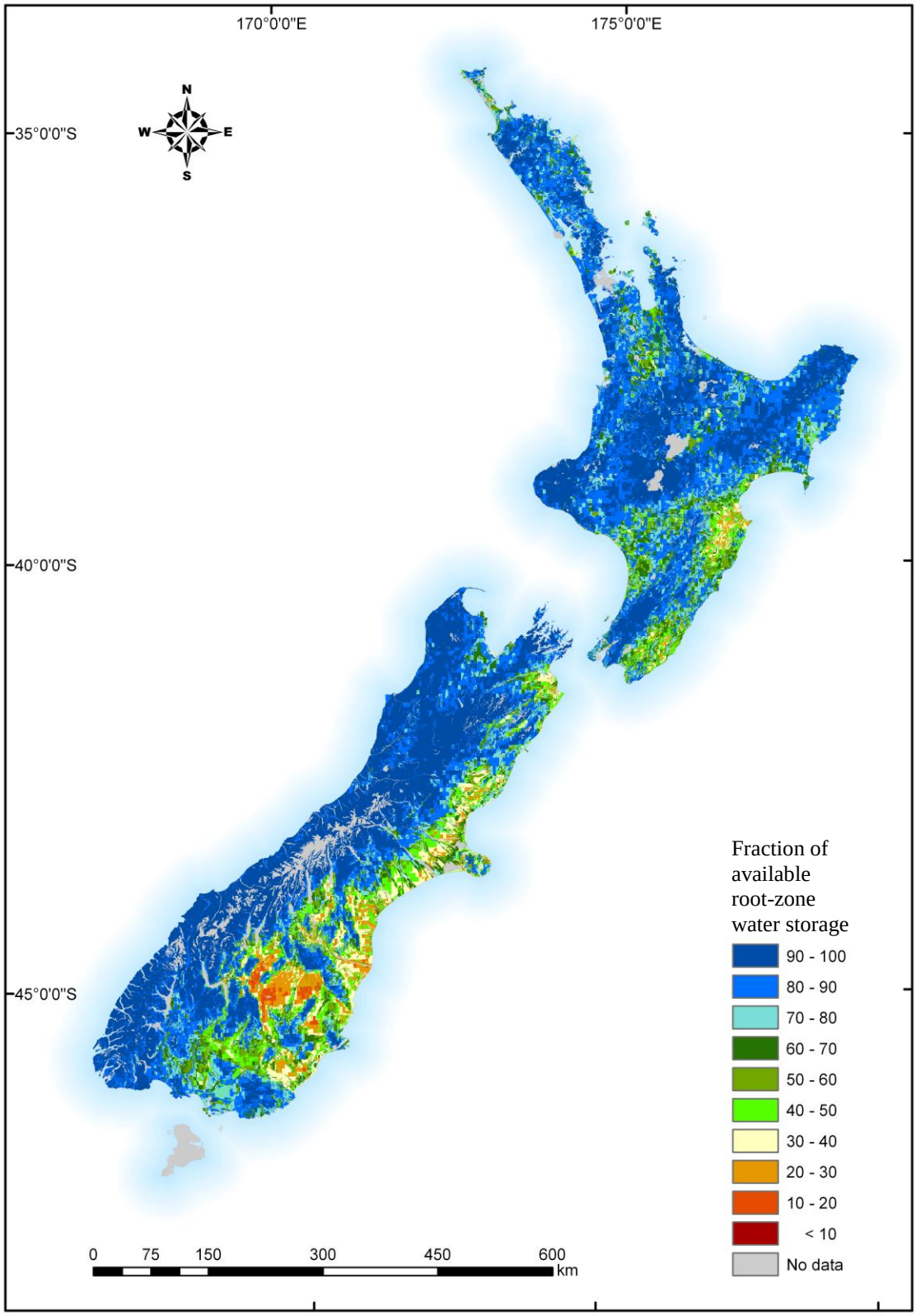
Appendix Ij: September monthly mean fraction of available root-zone water storage (W_f)



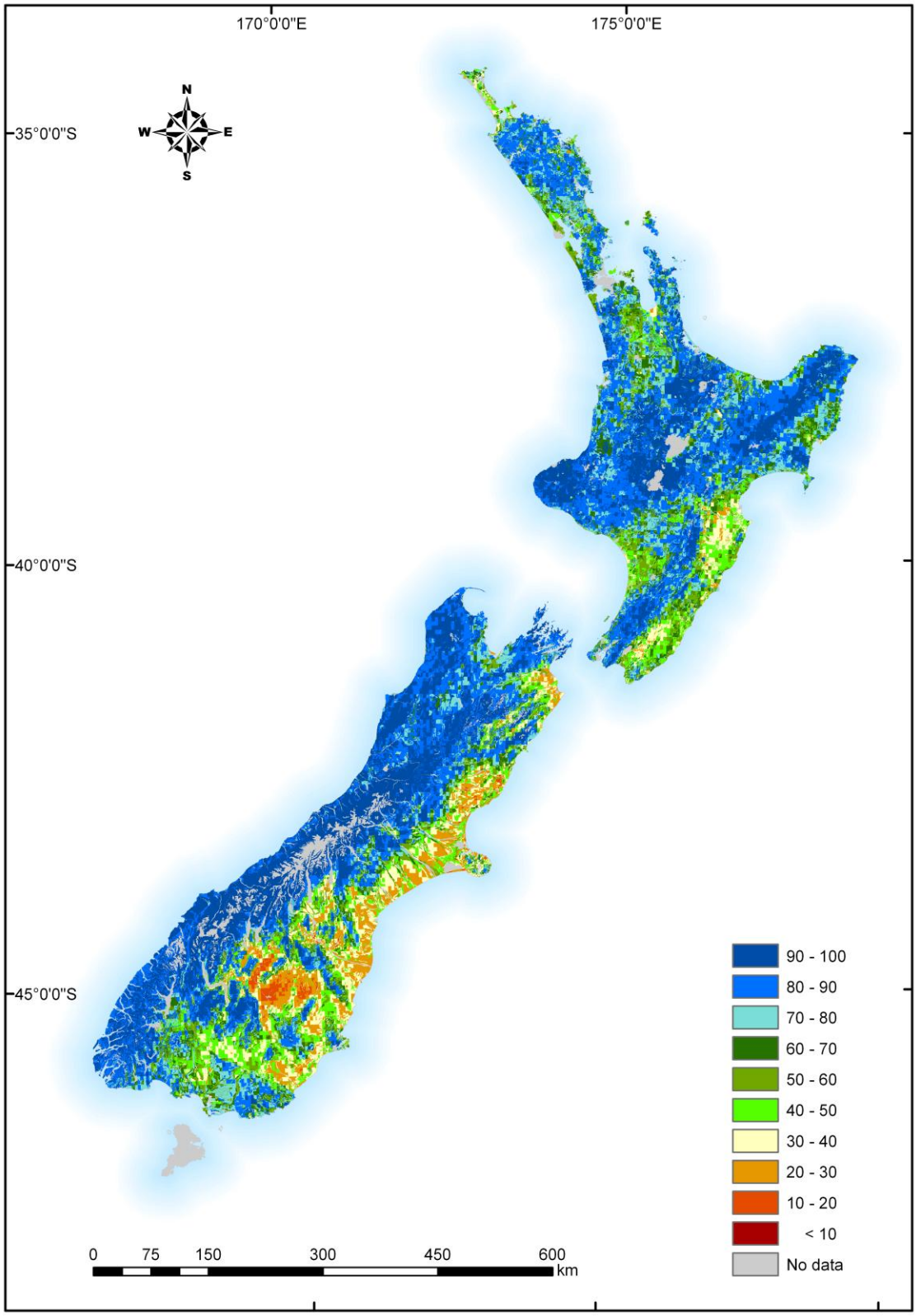
Appendix Ik: October monthly mean fraction of available root-zone water storage (W_f)



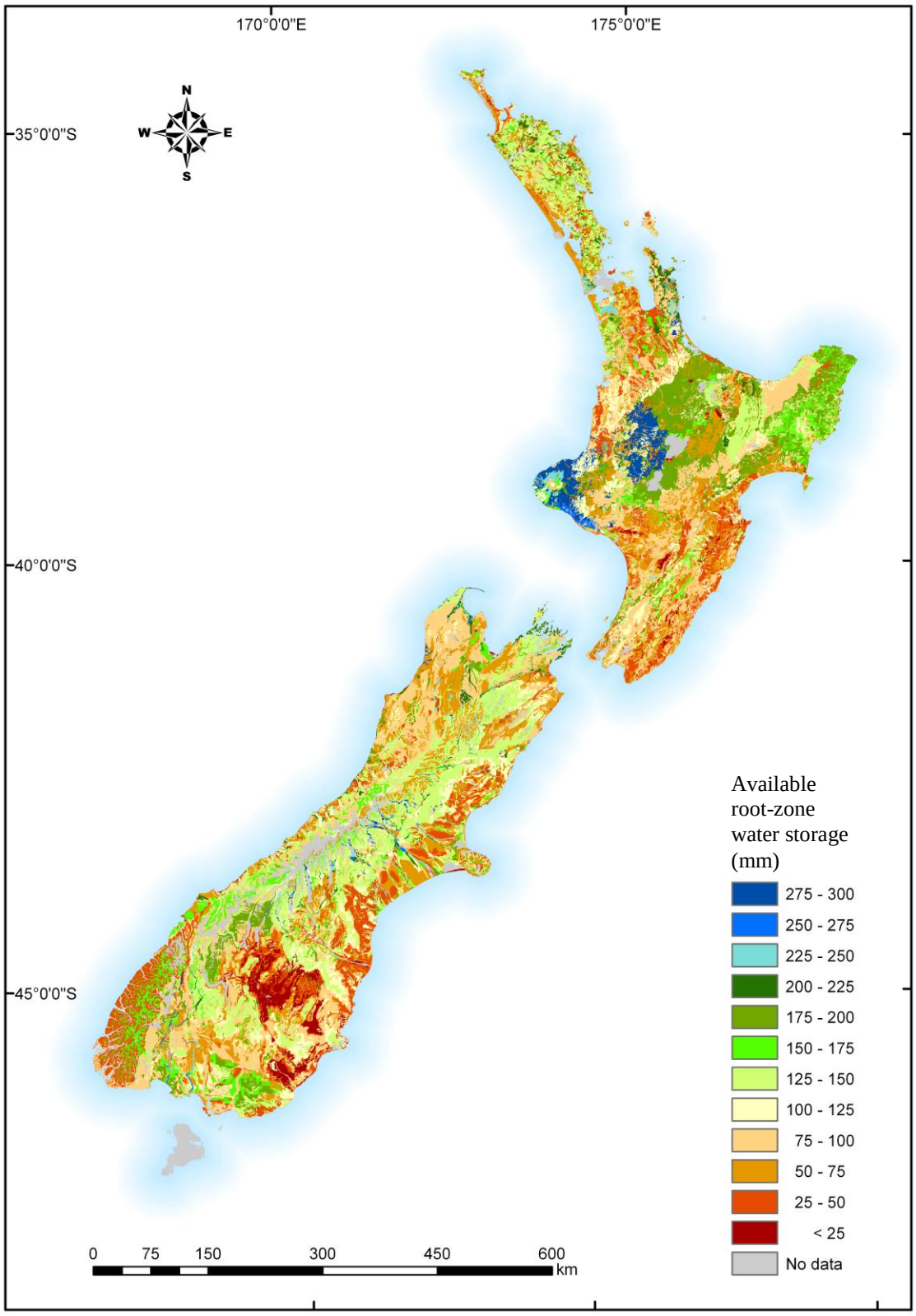
Appendix Il: November monthly mean fraction of available root-zone water storage (W_f)



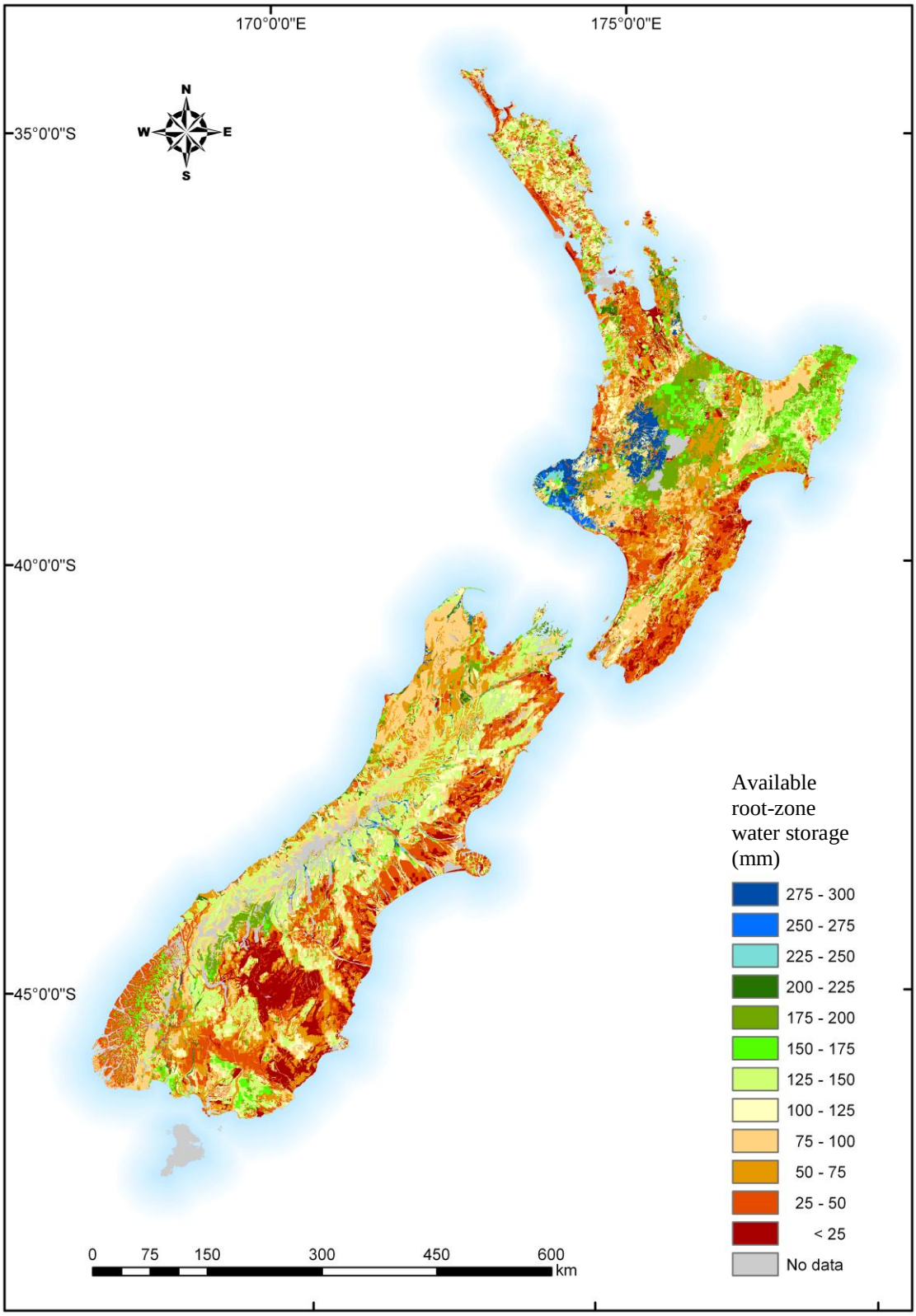
Appendix Im: December monthly mean fraction of available root-zone water storage (W_f)



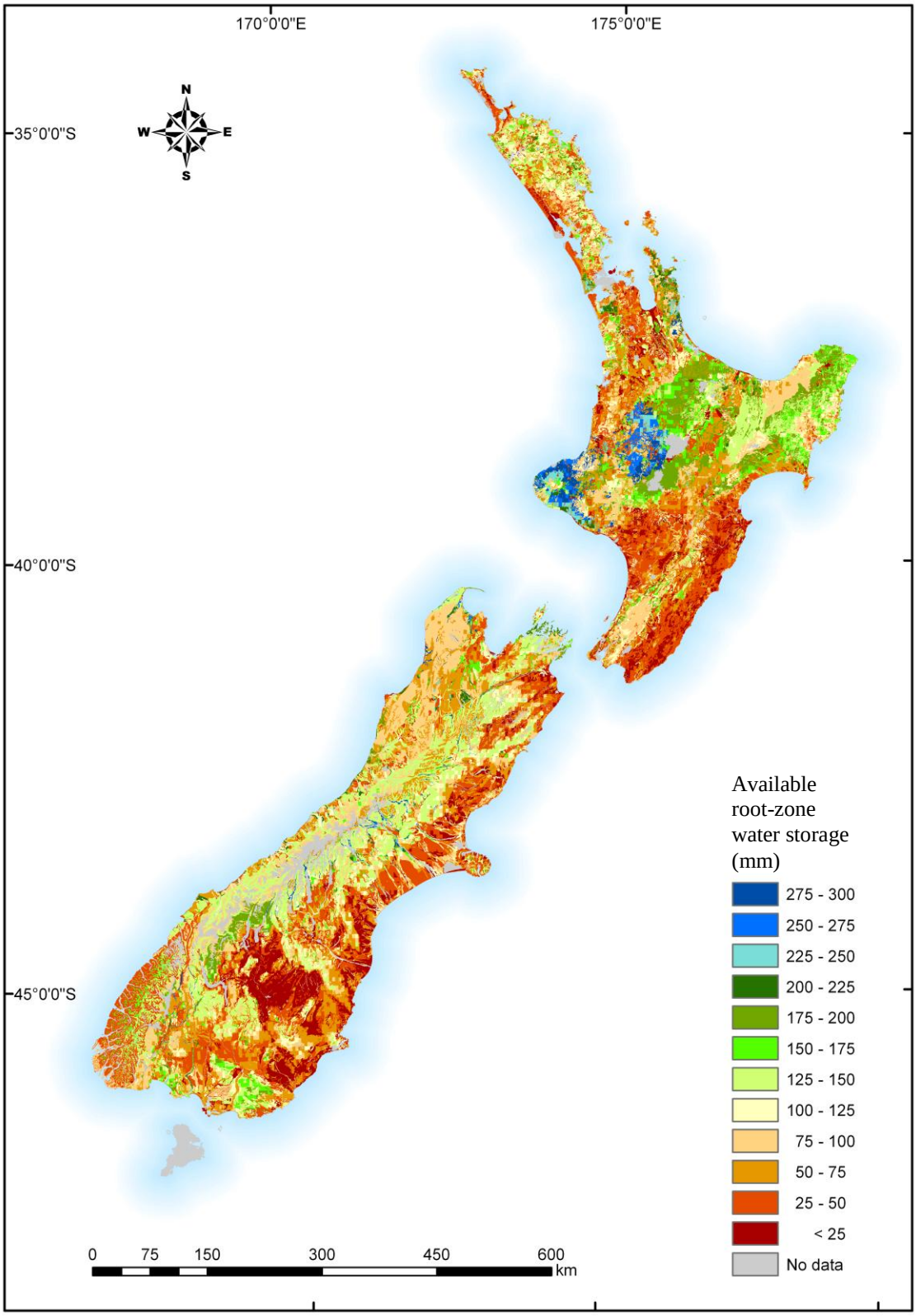
Appendix IIa: Annual mean available root-zone water storage (W_a)



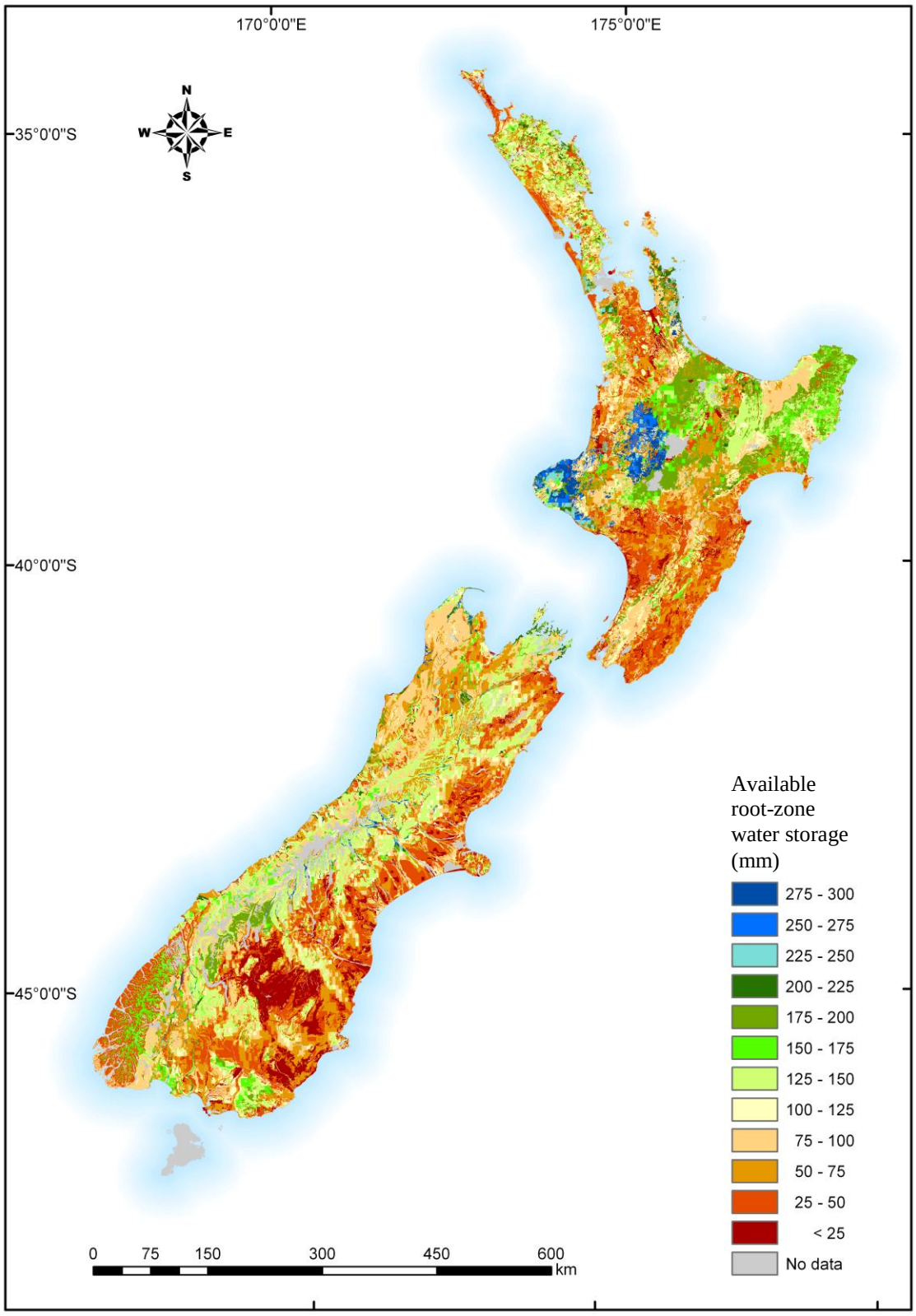
Appendix IIb: January monthly mean available root-zone water storage (W_a)



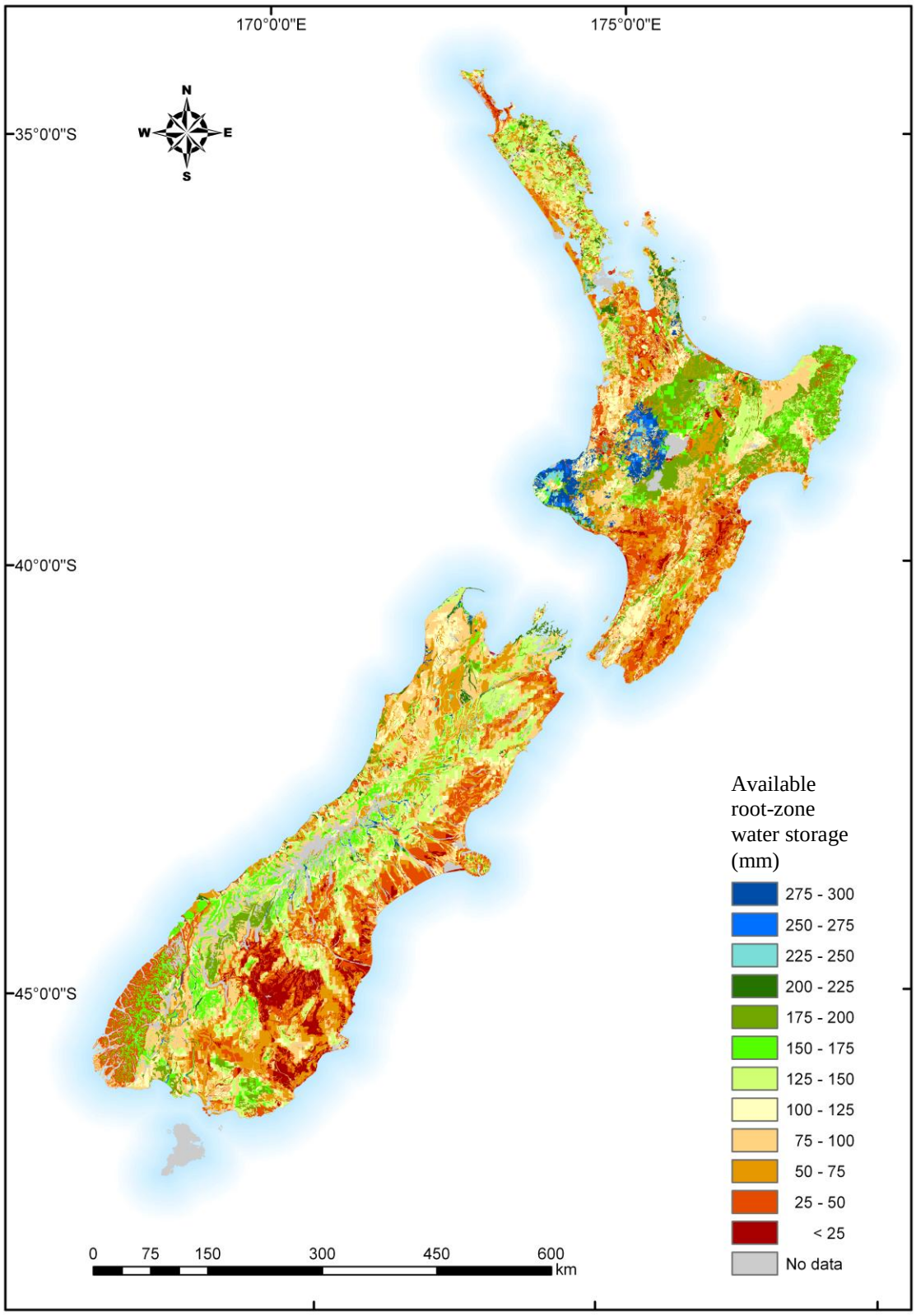
Appendix IIc: February monthly mean available root-zone water storage (W_a)



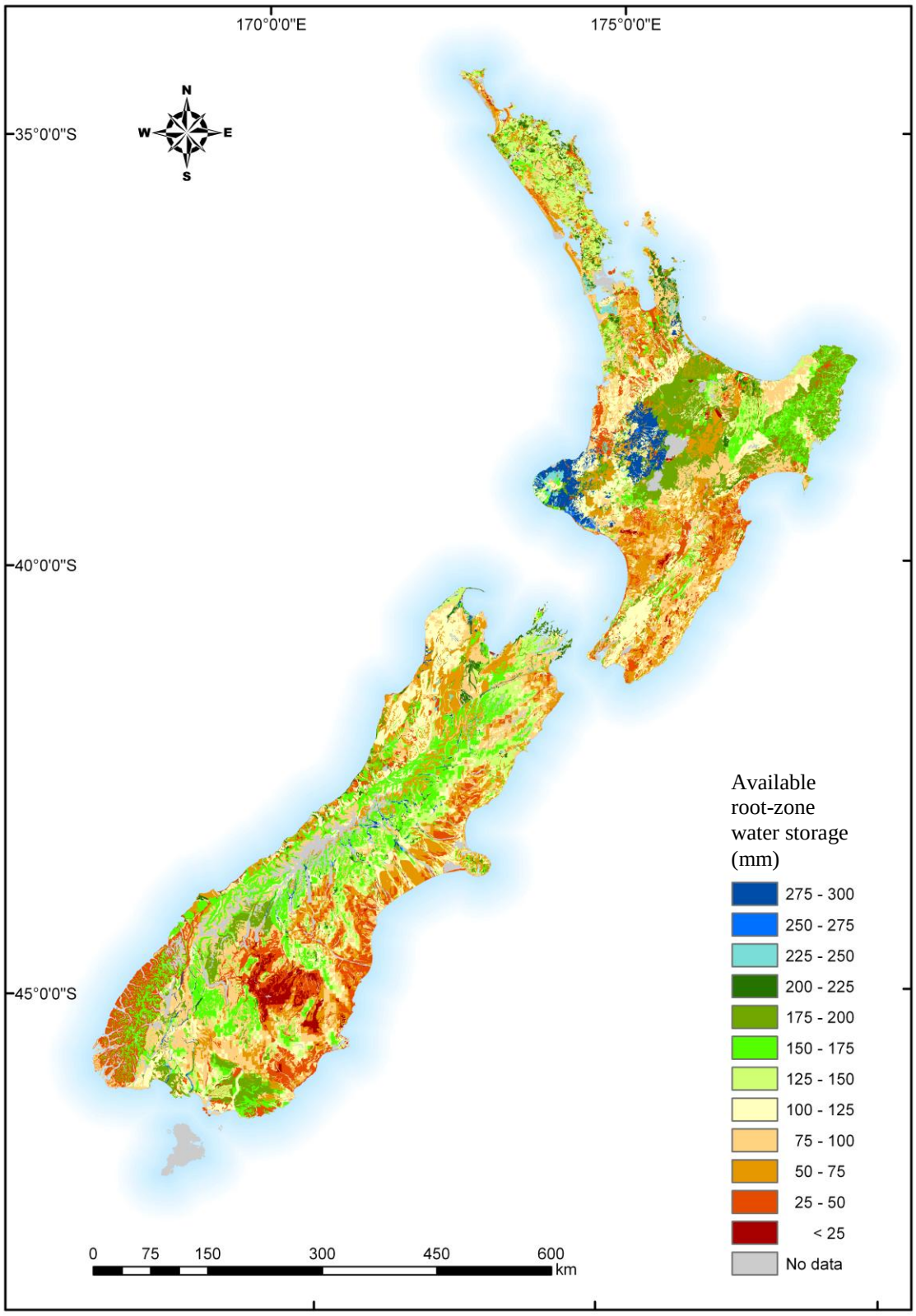
Appendix IId: March monthly mean available root-zone water storage (W_a)



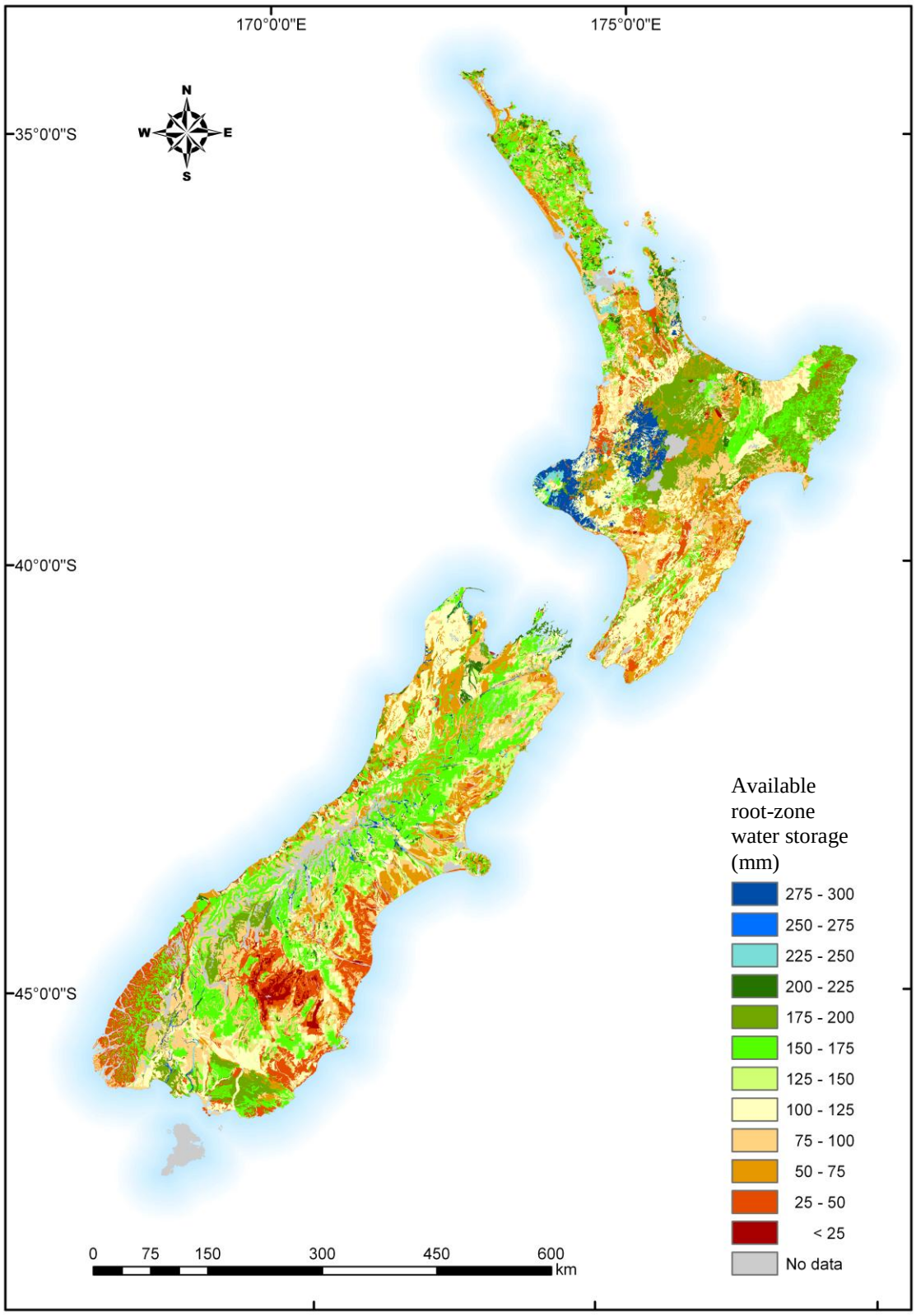
Appendix IIe: April monthly mean available root-zone water storage (W_a)



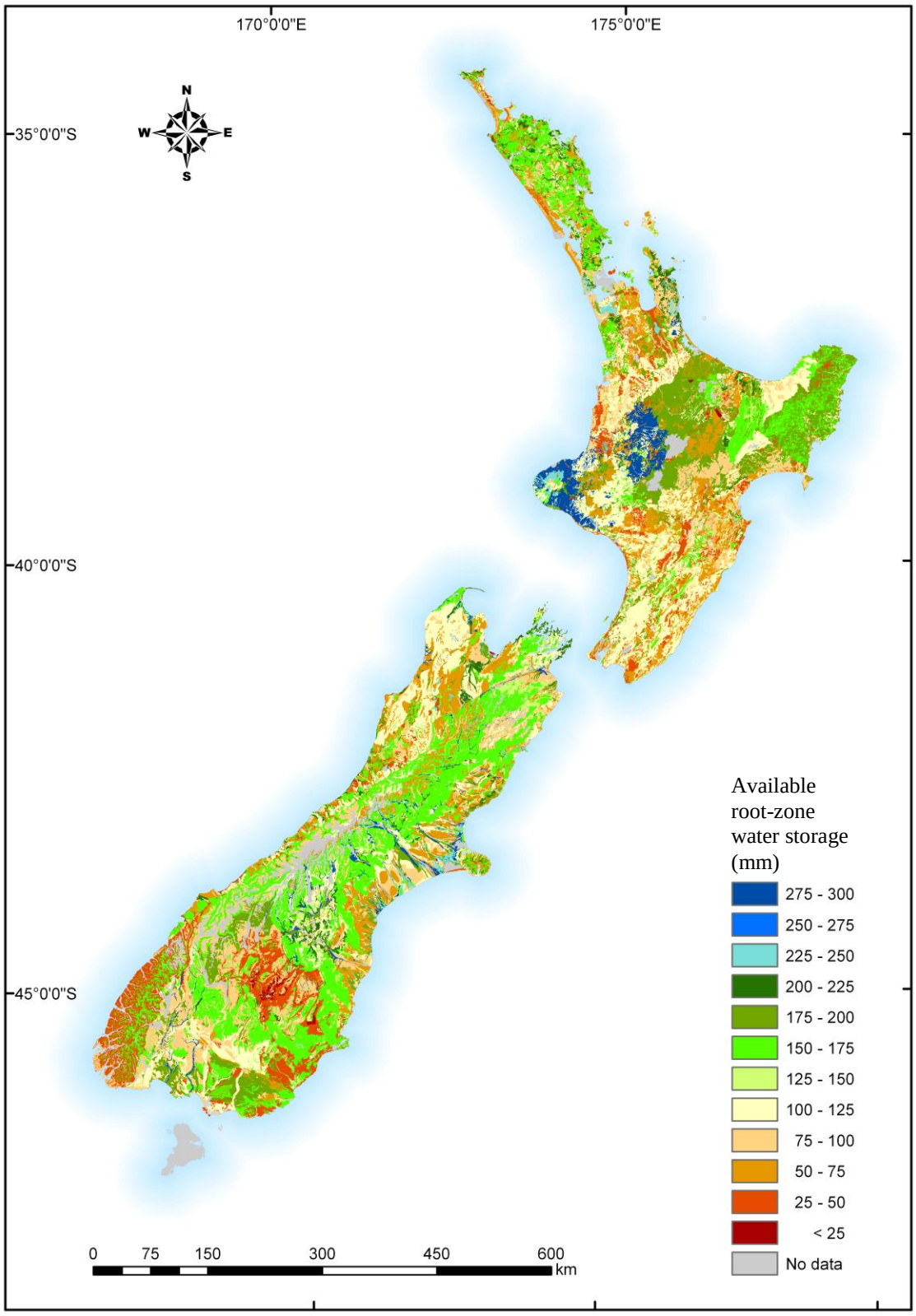
Appendix IIf: May monthly mean available root-zone water storage (W_a)



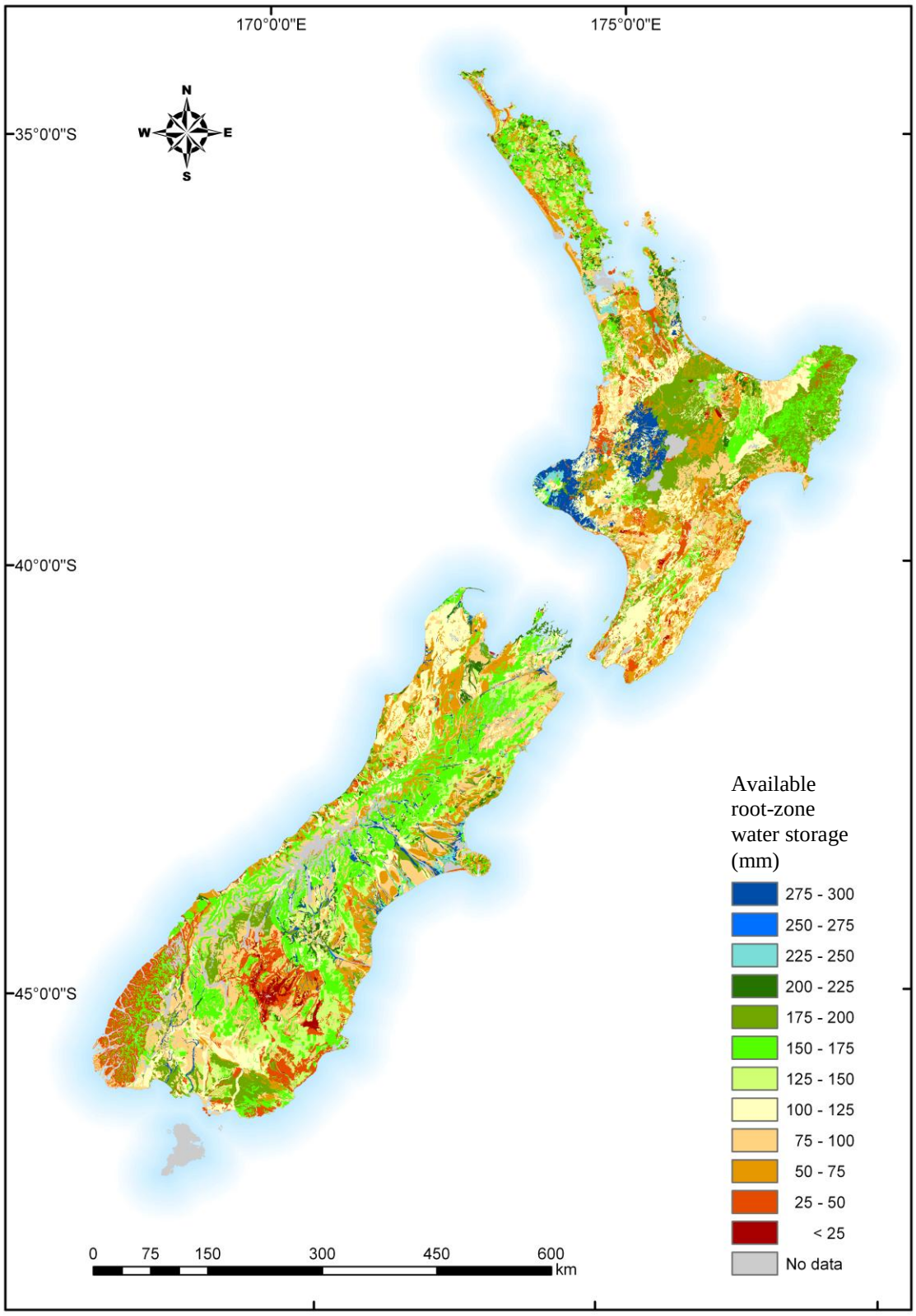
Appendix IIg: June monthly mean available root-zone water storage (W_a)



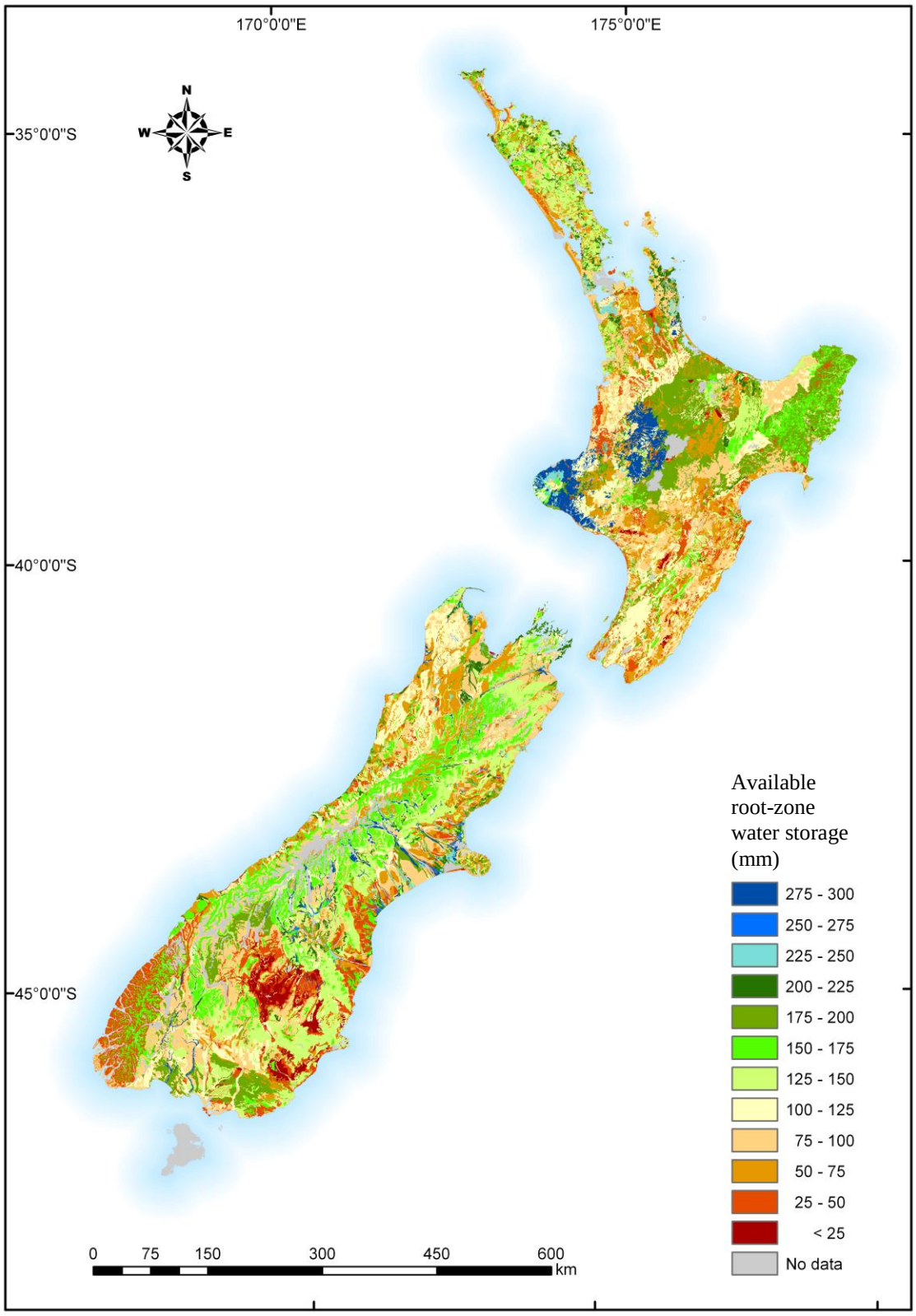
Appendix IIh: July monthly mean available root-zone water storage (W_a)



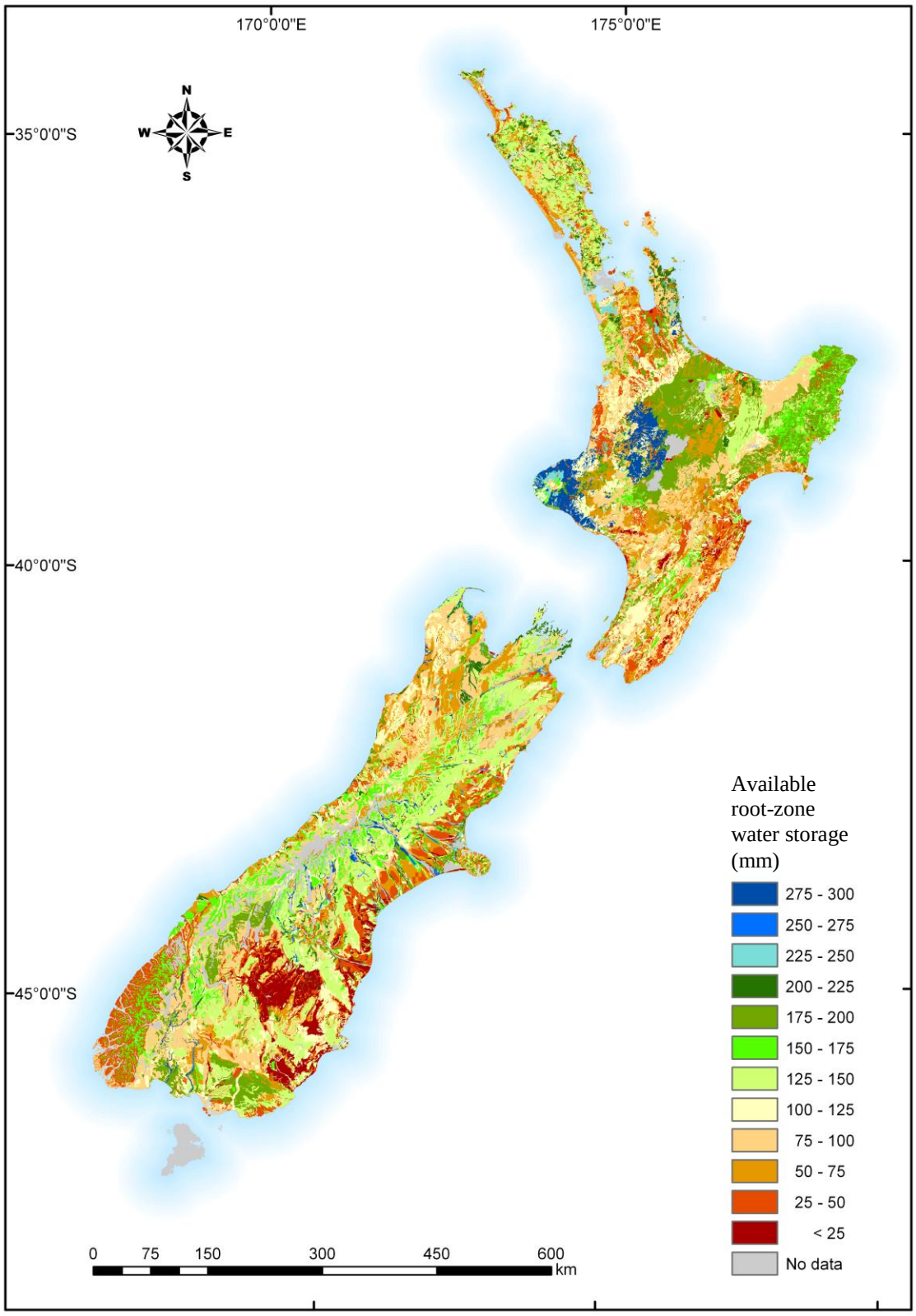
Appendix IIi: August monthly mean available root-zone water storage (W_a)



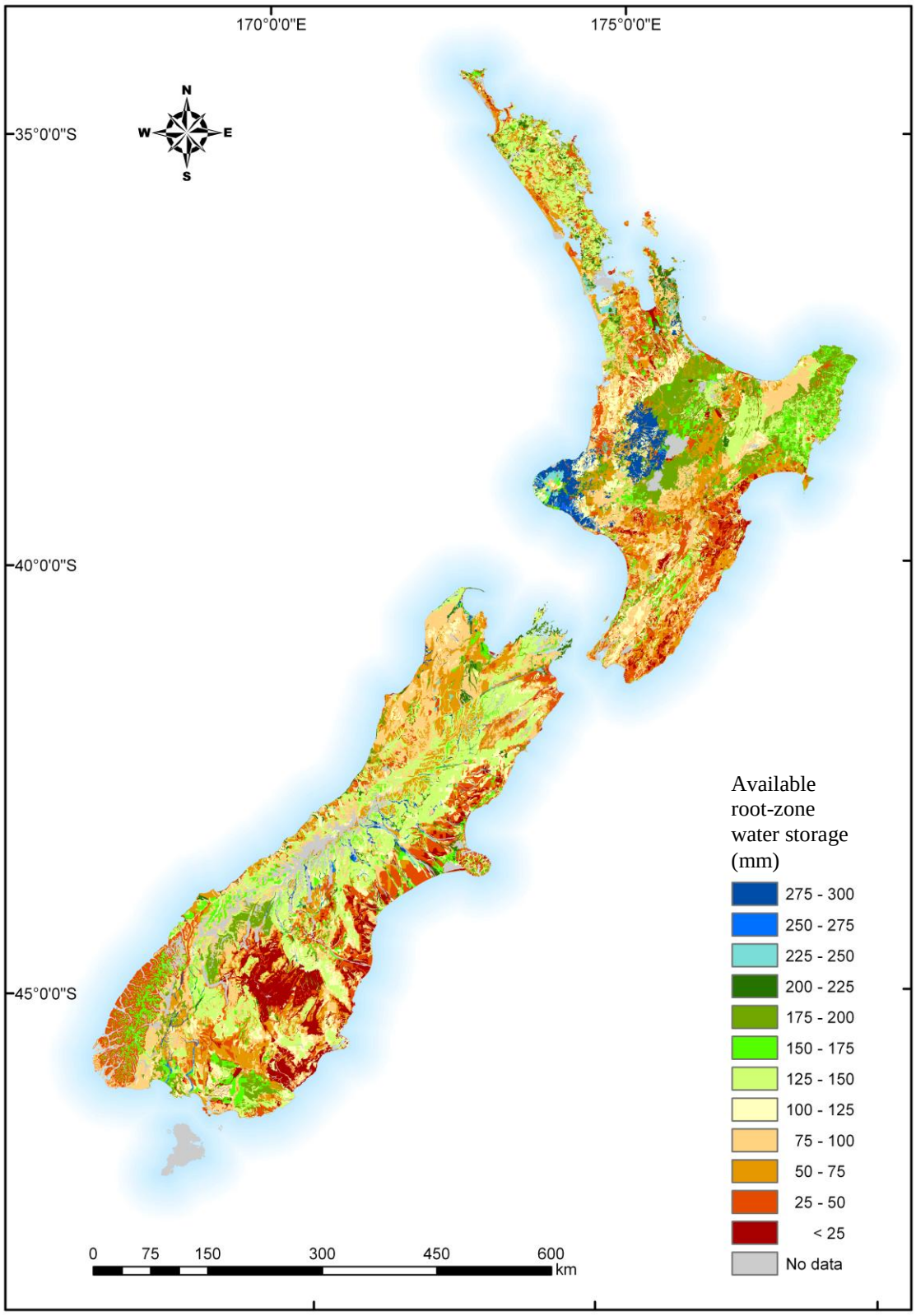
Appendix IIj: September monthly mean available root-zone water storage (W_a)



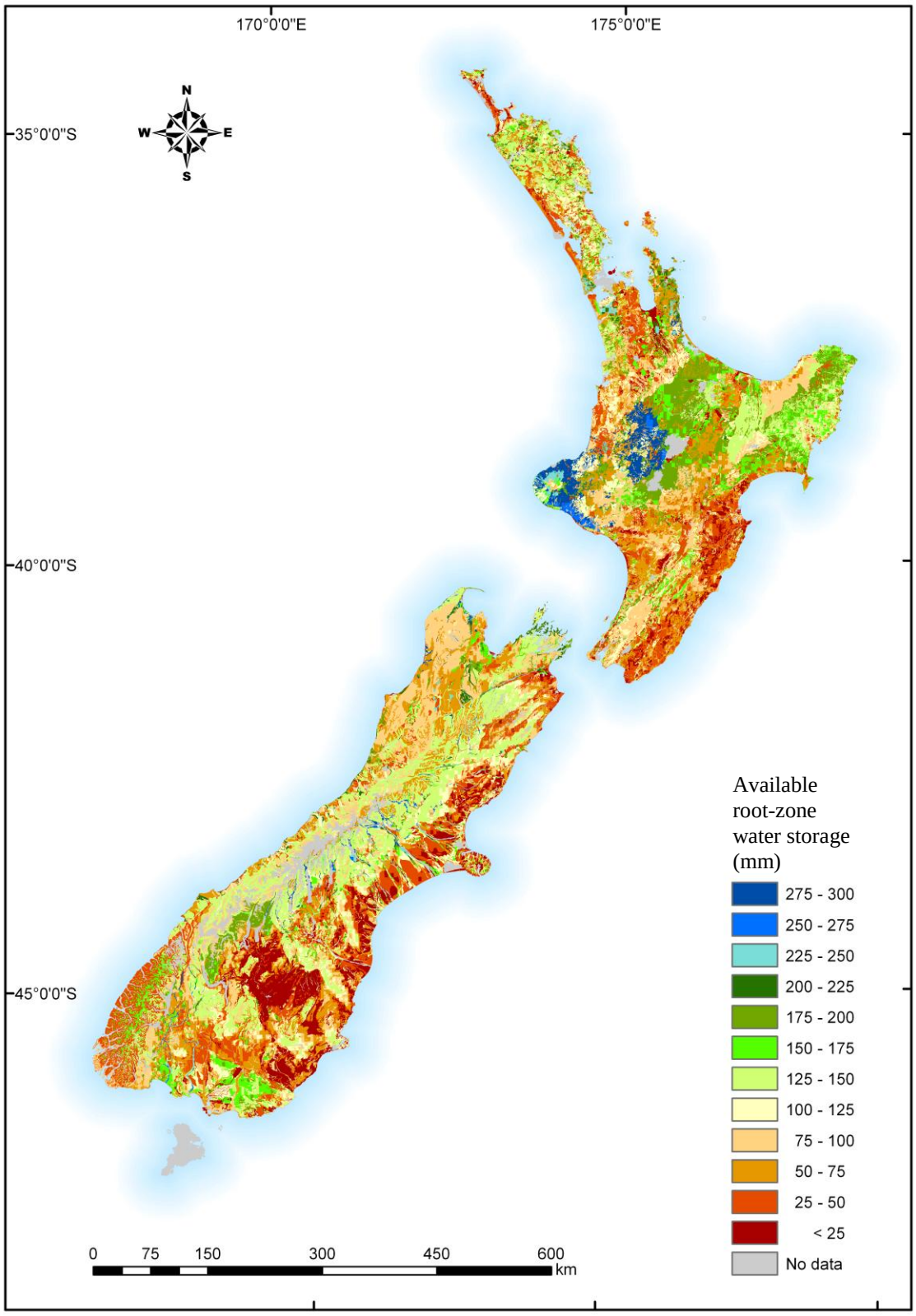
Appendix IIk: October monthly mean available root-zone water storage (W_a)



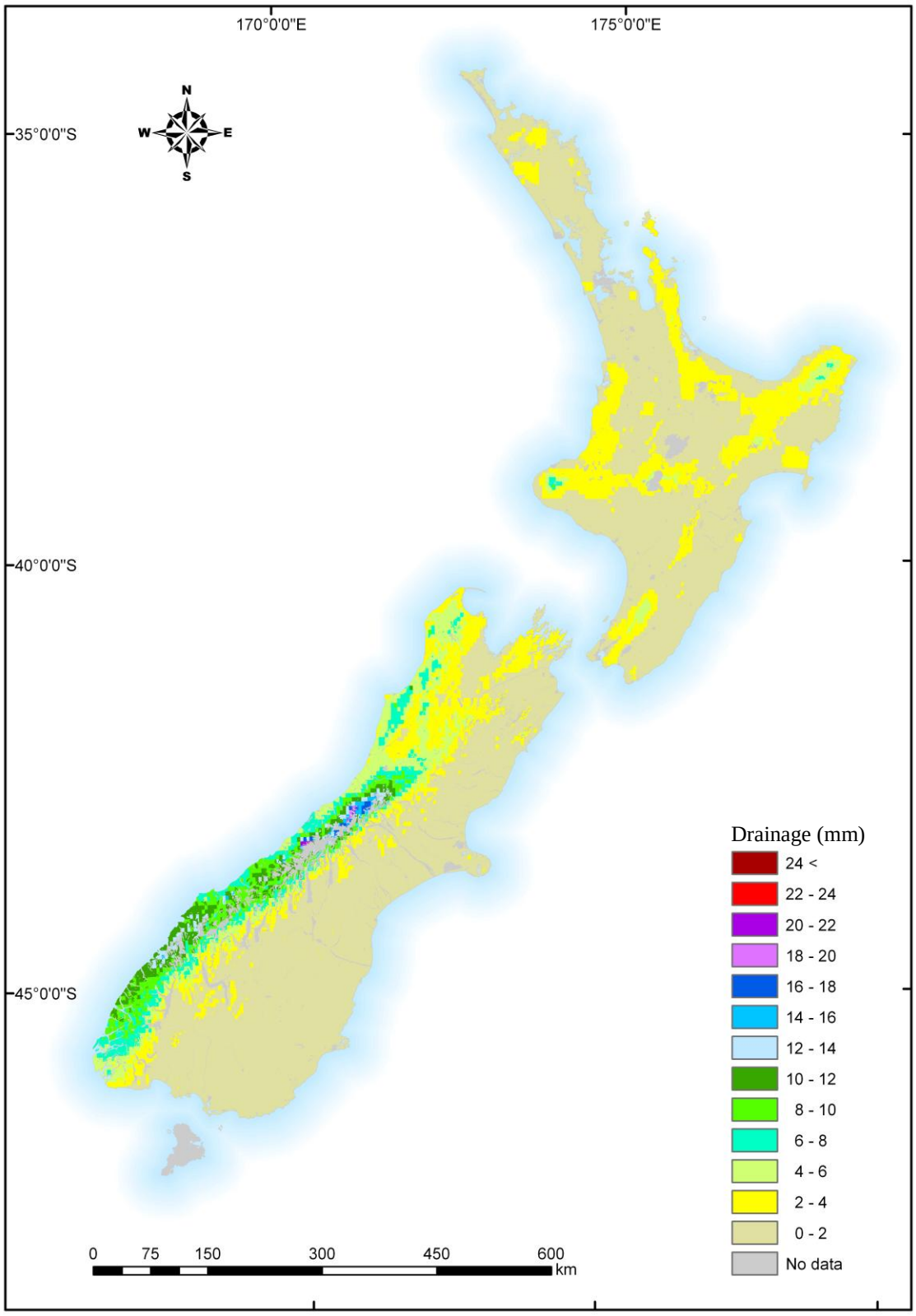
Appendix IIl: November monthly mean available root-zone water storage (W_a)



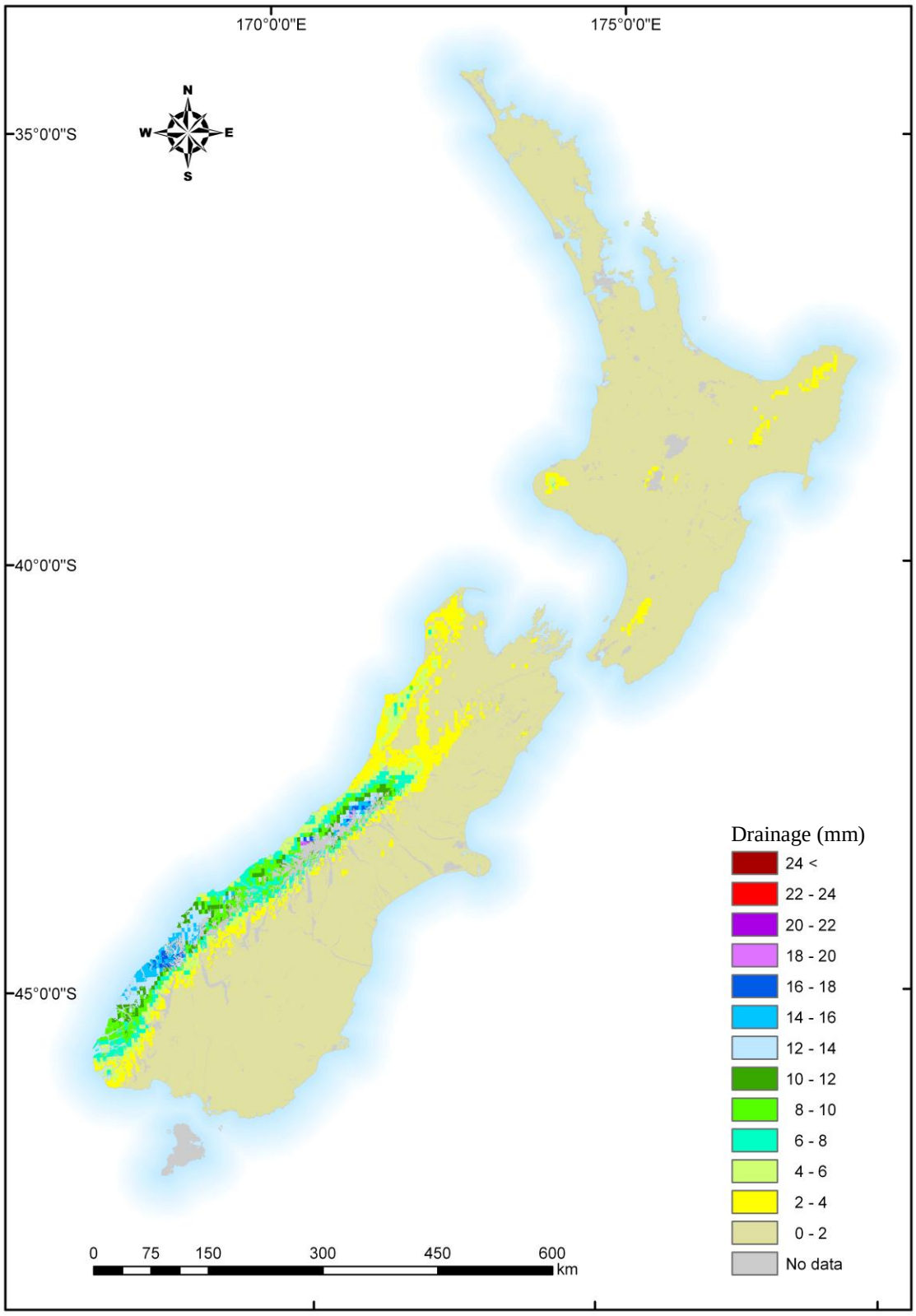
Appendix IIm: December monthly mean available root-zone water storage (W_a)



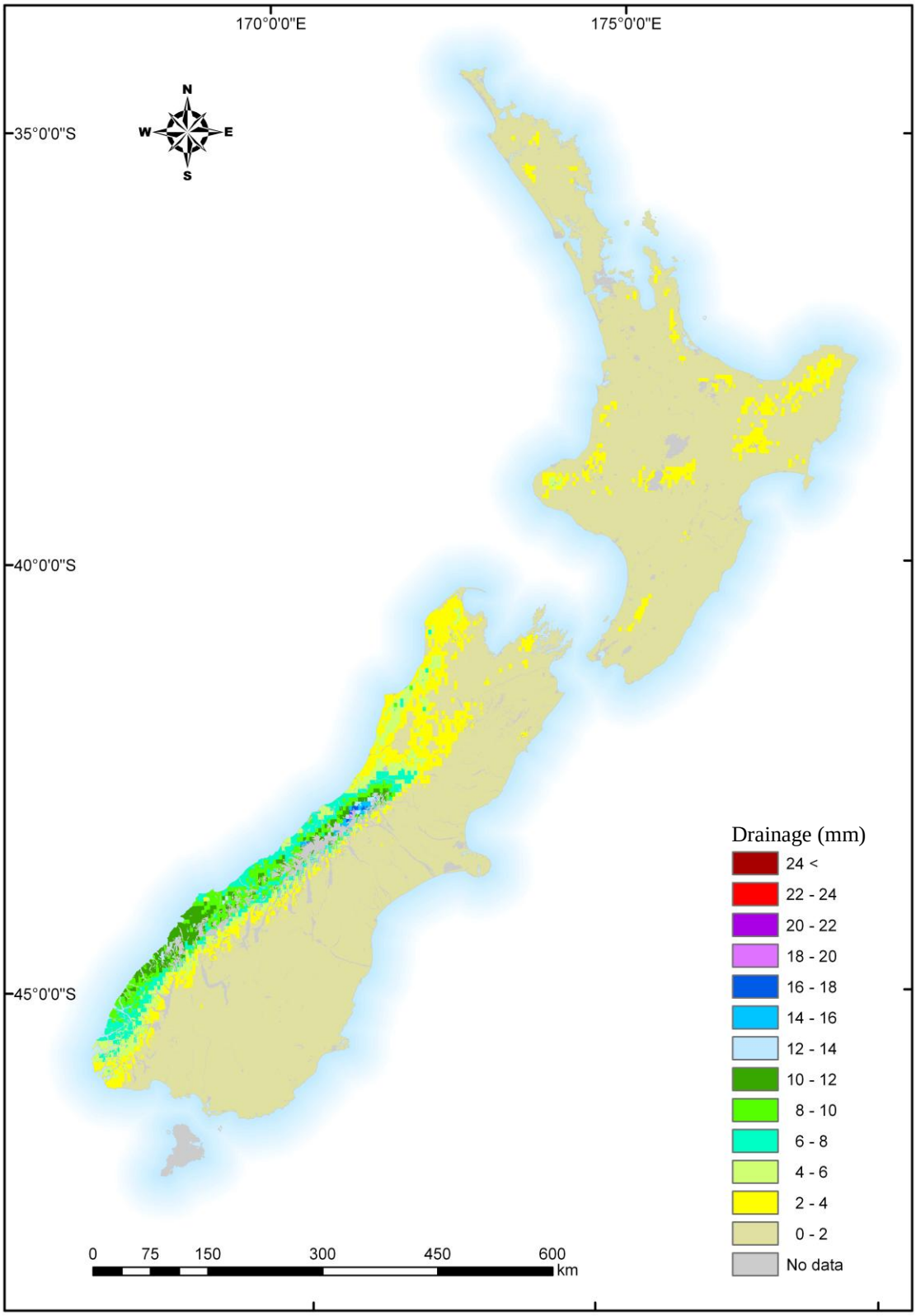
Appendix IIIa: Annual mean drainage



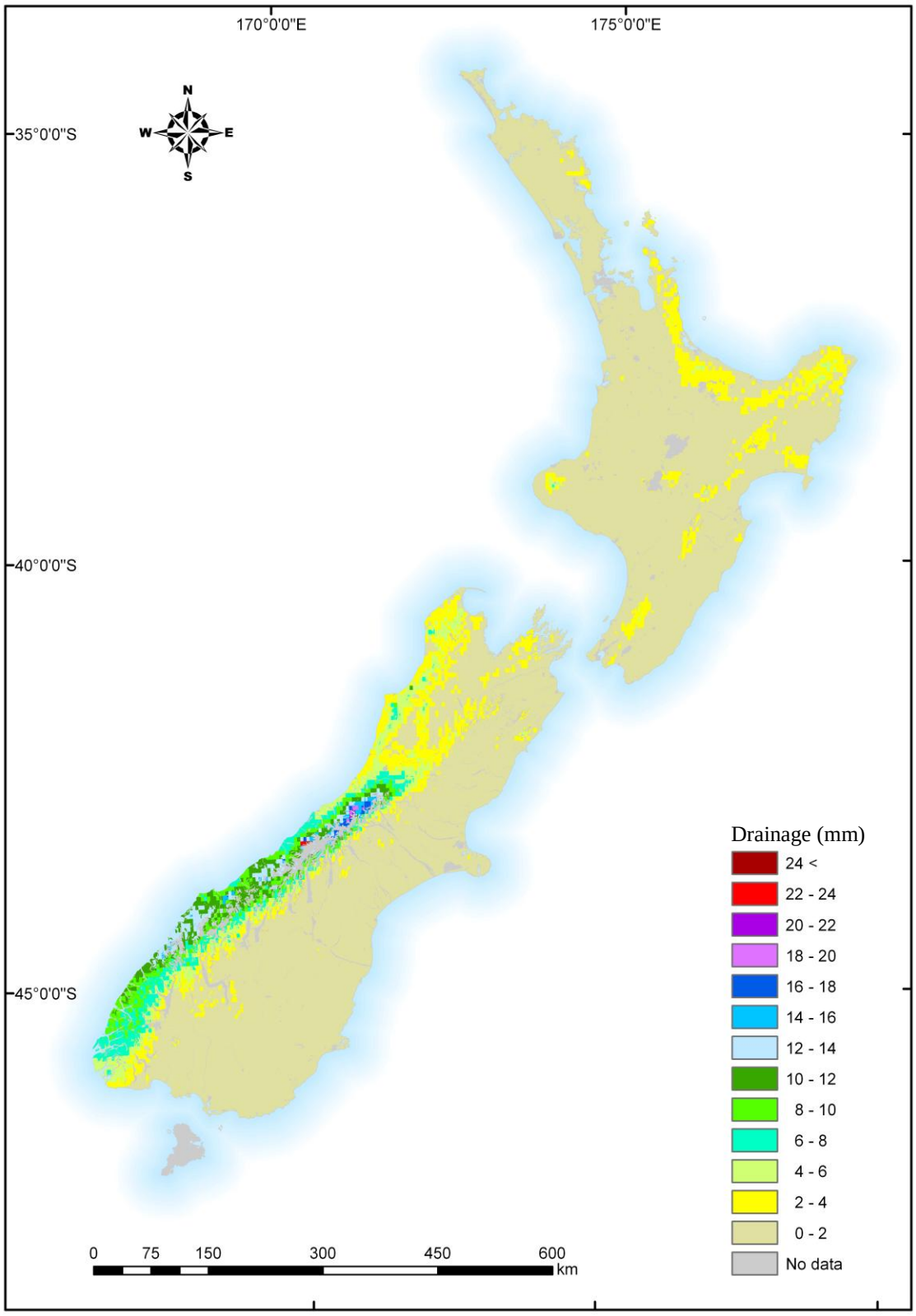
Appendix IIIb: January monthly mean drainage



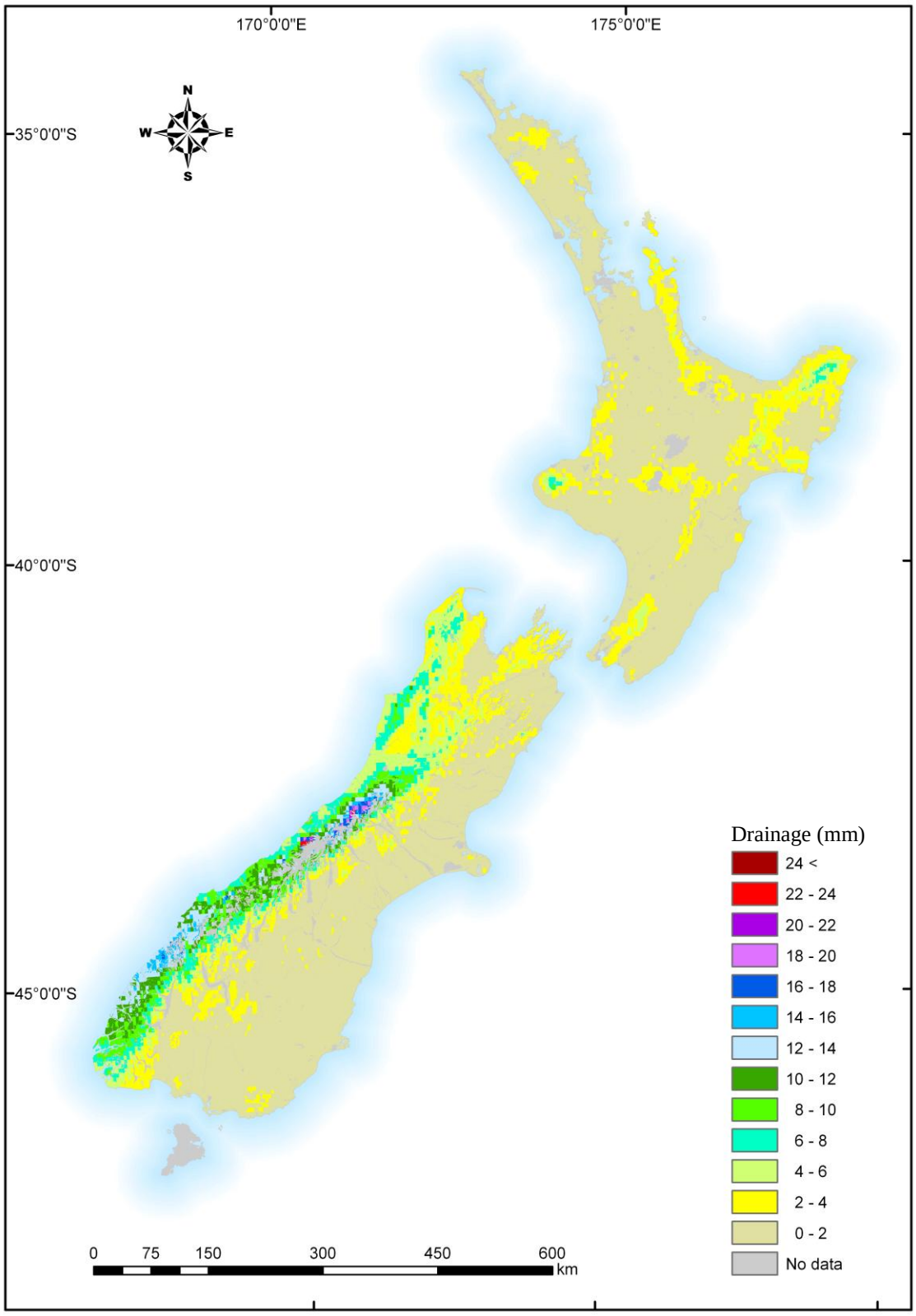
Appendix IIIc: February monthly mean drainage



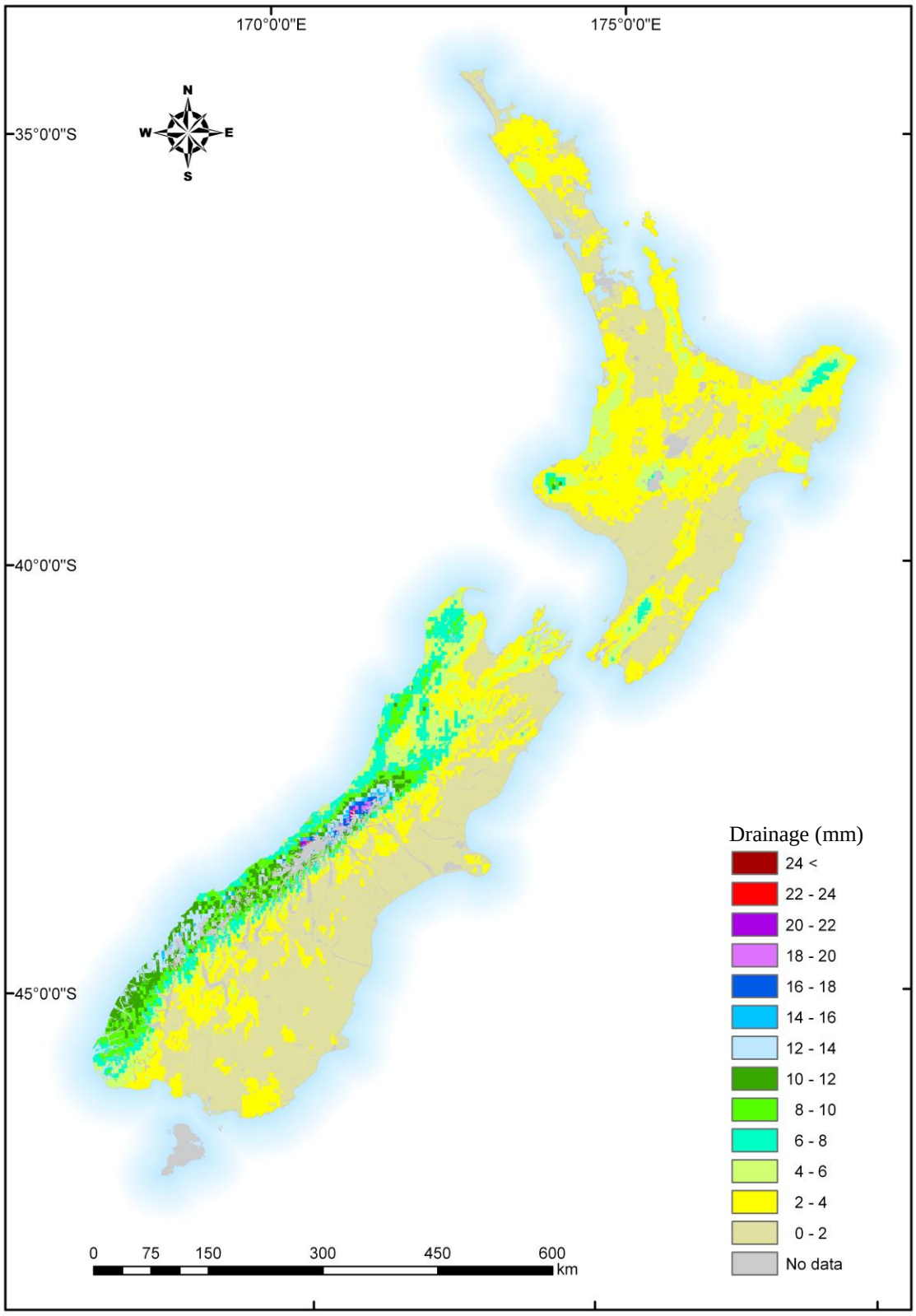
Appendix IIId: March monthly mean drainage



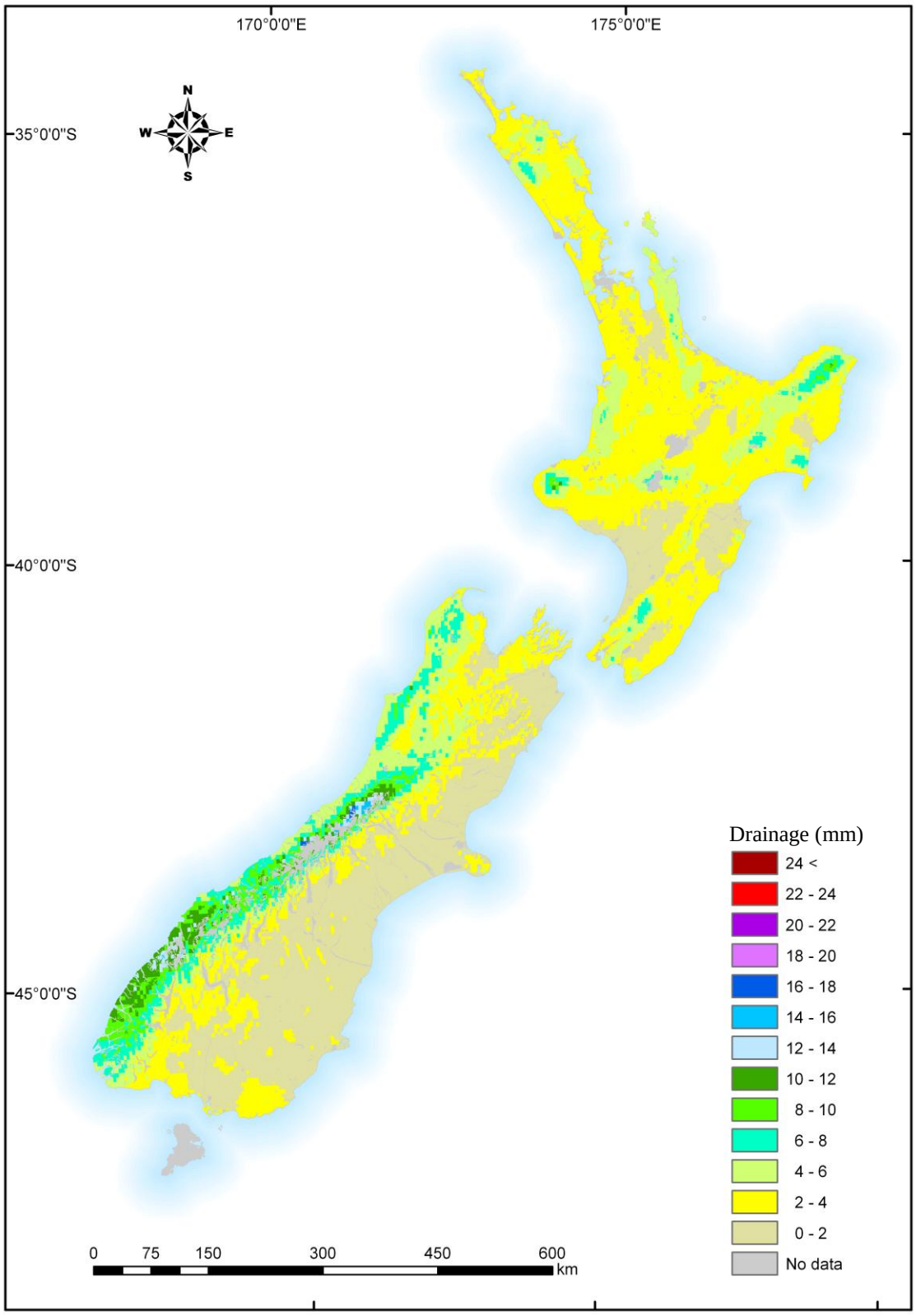
Appendix IIIe: April monthly mean drainage



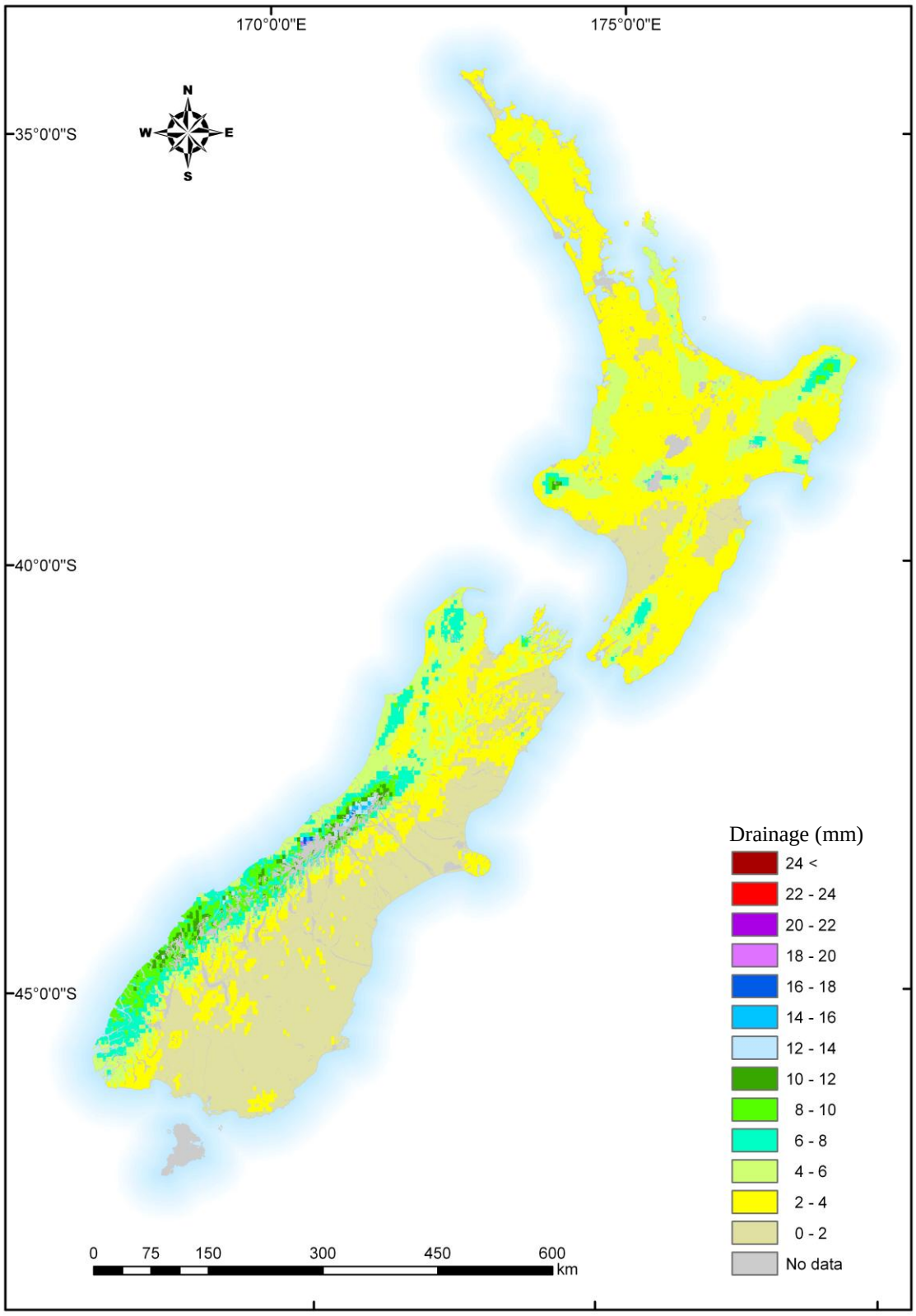
Appendix IIIf: May monthly mean drainage



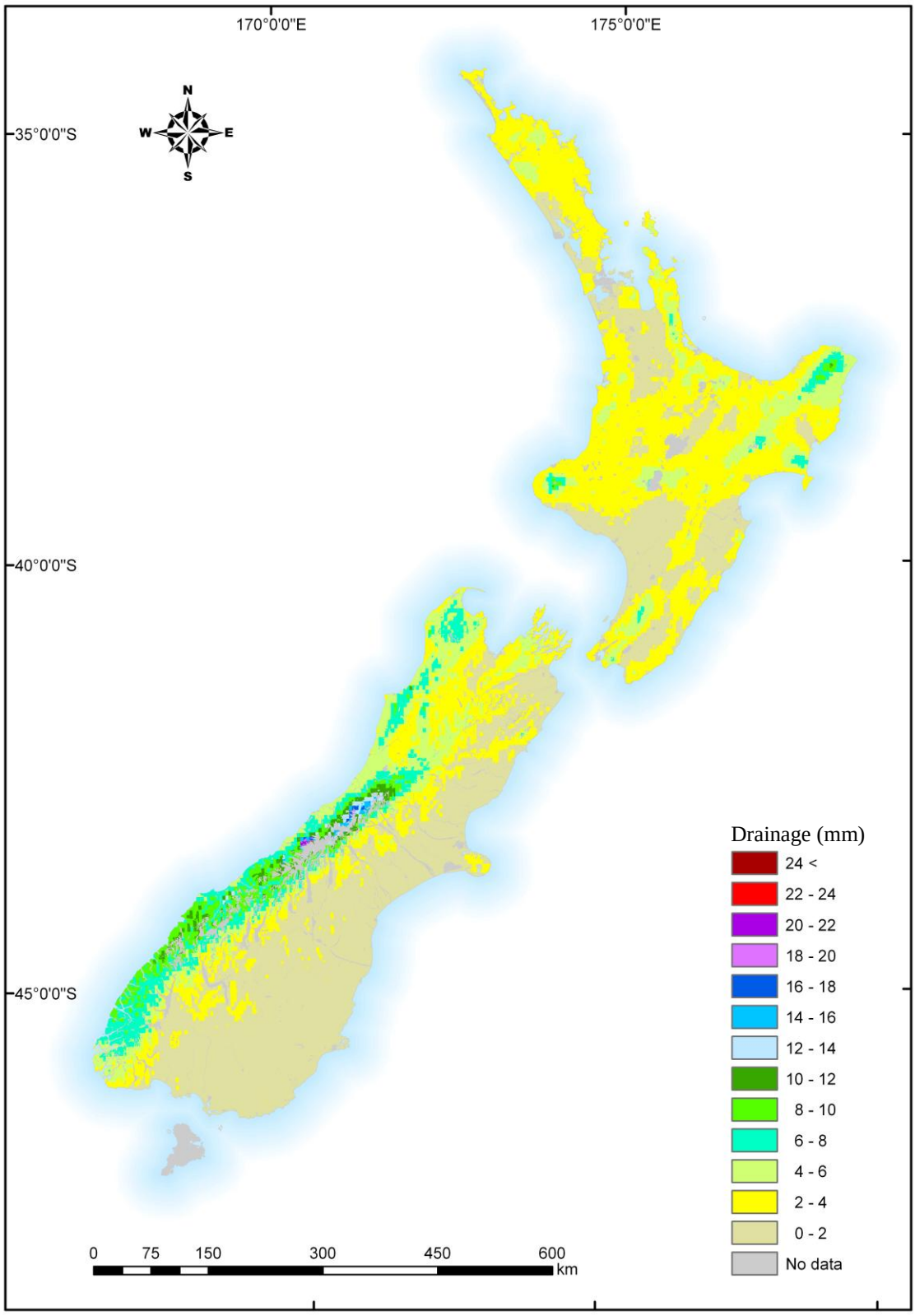
Appendix IIIg: June monthly mean drainage



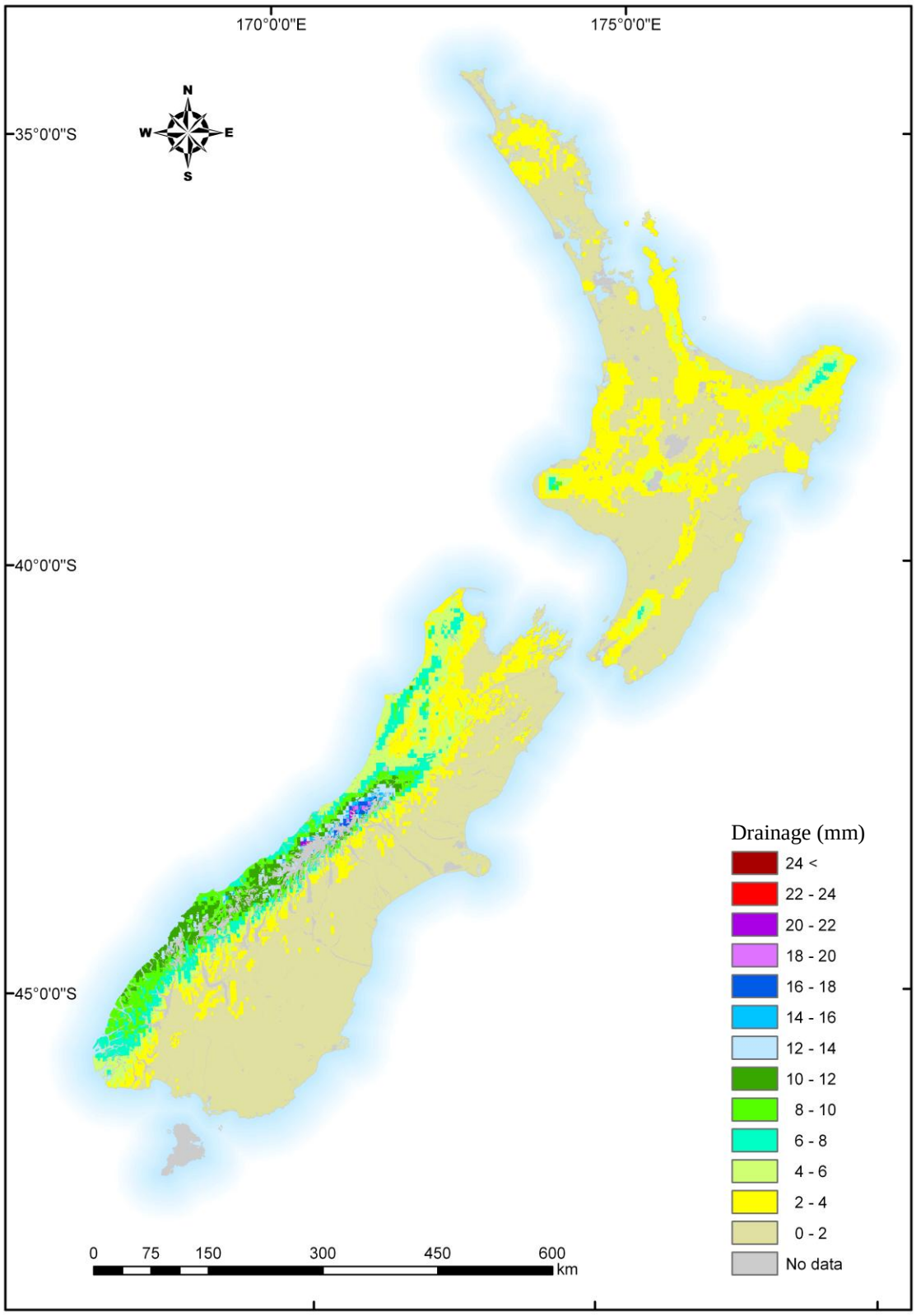
Appendix IIIh: July monthly mean drainage



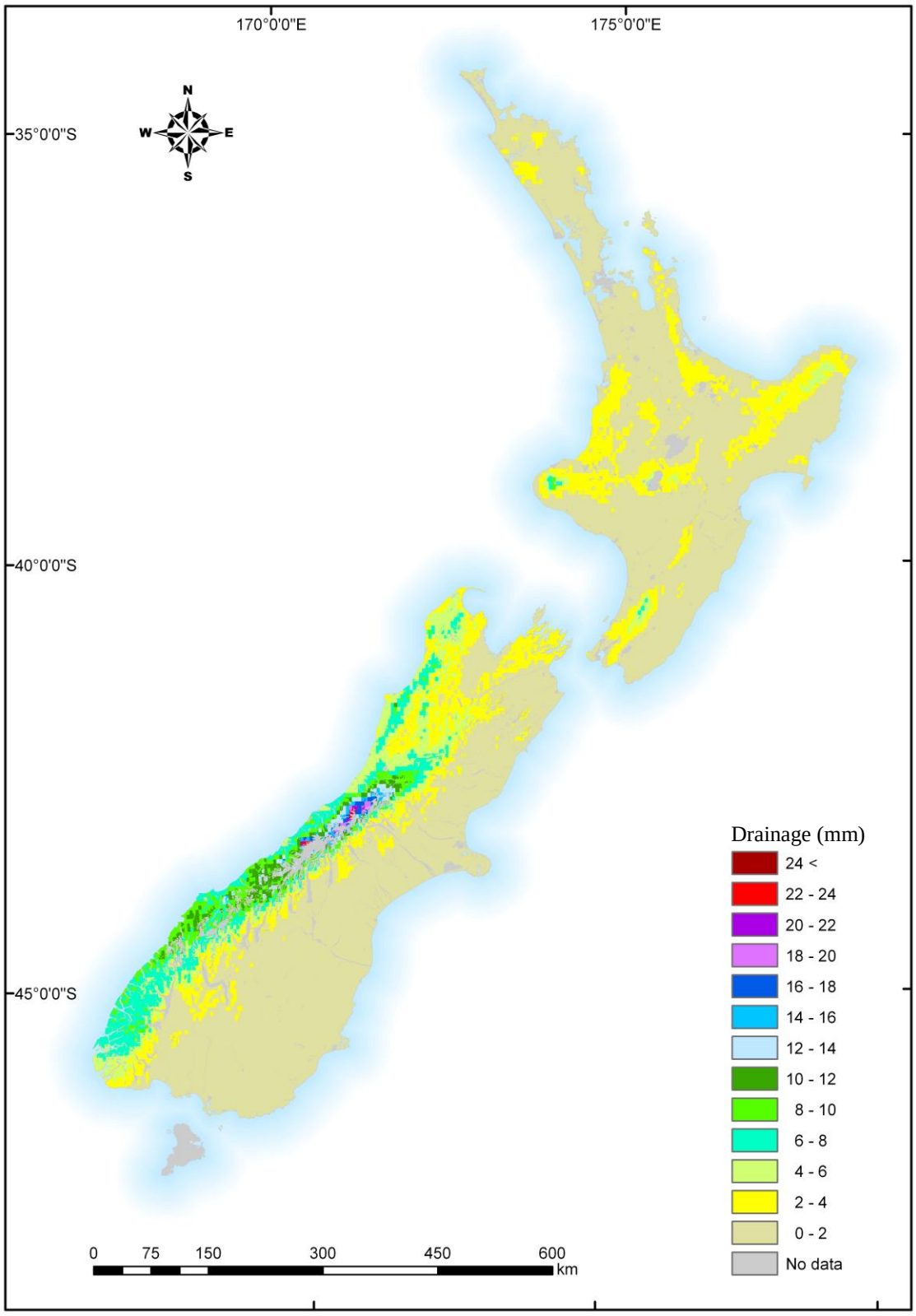
Appendix IIIi: August monthly mean drainage



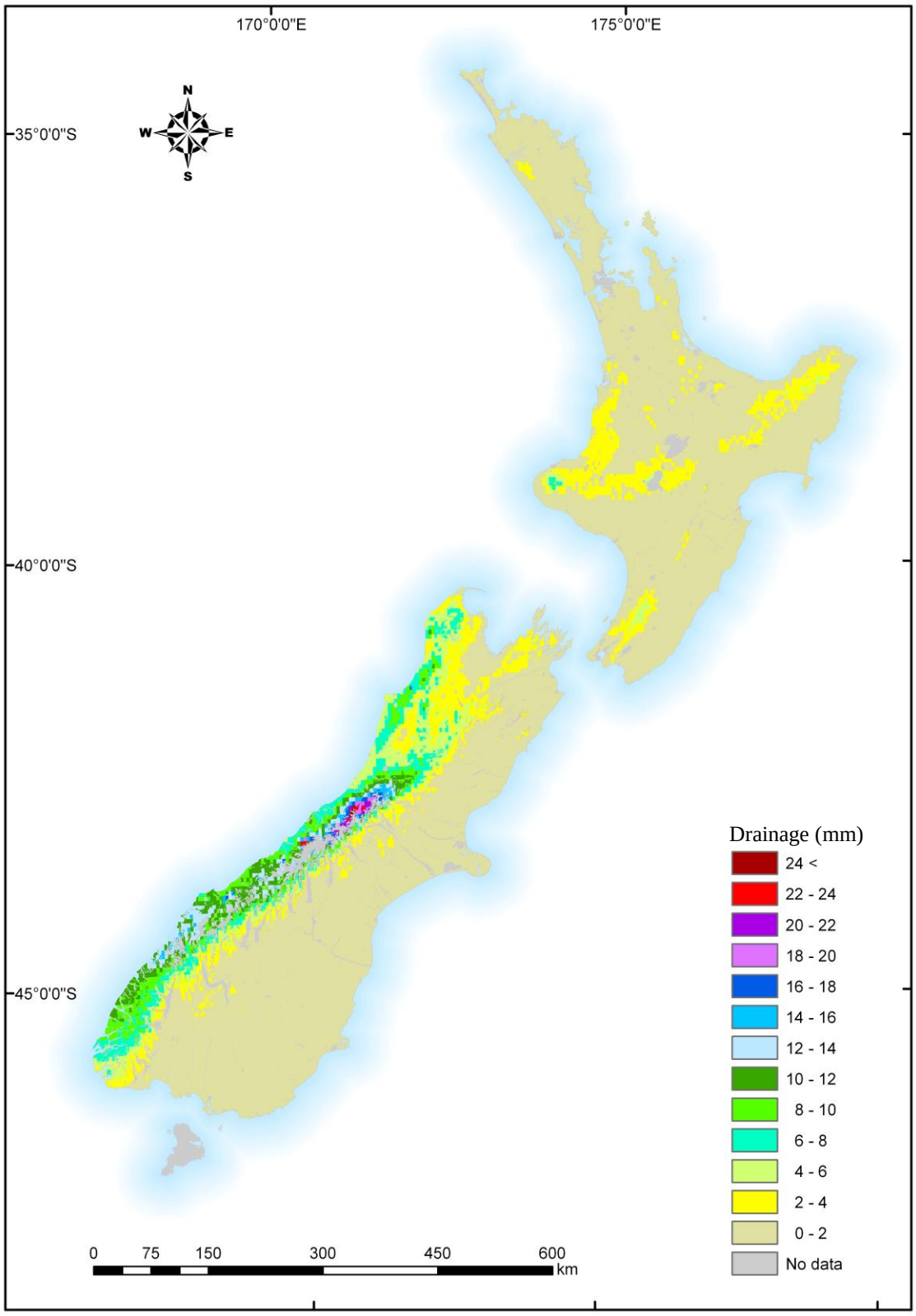
Appendix IIIj: September monthly mean drainage



Appendix IIIk: October monthly mean drainage



Appendix IIIl: November monthly mean drainage



Appendix IIIm: December monthly mean drainage