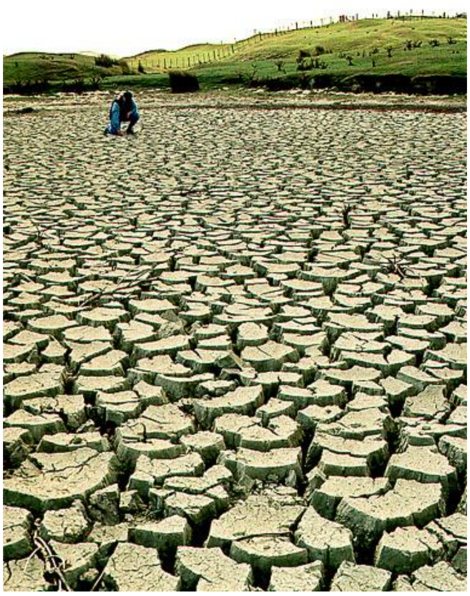
Scenarios of Regional Drought under Climate Change

Prepared for Ministry of Agriculture and Forestry

June 2011



Authors/Contributors:
Dr Anthony Clark, Dr Brett Mullan, Mr Alan Porteous

For any information regarding this report please contact:

Dr Anthony Clark
Climate Variability
+64-4-386 0509
a.clark@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
301 Evans Bay Parade, Greta Point
Wellington 6021
Private Bag 14901, Kilbirnie
Wellington 6241
New Zealand

Phone +64-4-386 0300
Fax +64-4-386 0574

NIWA Client Report No:	WLG2010-32
Report date:	June 2011
NIWA Project:	SLDR093
FRST Contract:	C01X0818

Photo credits: Hill country farmer east of Gladstone surveys the paddocks after 14 consecutive months of no significant rainfall. [Alan Blacklock, NIWA 2008]

Contents

Tables

Figures

Reviewed by

AB Tait

Approved for release by

David Wratt

Executive summary

Key finding

This report describes the methodology and results of an assessment of New Zealand dry land agronomic drought under future scenarios of climate change. Based on this work, the more likely scenario is that New Zealand can plan for around ten percent additional time spent in drought by the middle of this century for key eastern agricultural regions. Less likely but identifiable upper and lower drought scenarios for the middle of this century range from: a strong shift toward a more drought prone climate over most agricultural regions with well over a doubling of time spent in drought across most of New Zealand; to an environment of small increases from current levels isolated to eastern agricultural regions. By the end of the this century the projections establish that much of New Zealand's agricultural zone will experience some increase in drought, even under the milder scenario. While planning and action can focus on the more likely scenario which will also encompass the lower projection, contingencies also need to be considered to manage the less likely upper range.

Methodology

This study follows previous work undertaken by NIWA in 2005 as well as a series of consultations on the methodology. The updated estimates are based on a new and expanded set of climate change data developed by: accessing the climate model ensemble used in the Intergovernmental Panel on Climate Changes (IPCC) Fourth Assessment Report (AR4); using the three major global greenhouse gas emissions scenarios, B1, A1B and A2; and improved statistical downscaling methods to provide localised climate scenarios for New Zealand. The final estimates of drought probability were derived from a nationally comprehensive soil moisture indicator that is sensitive to moderate to severe intensity droughts greater than one month in duration. Verification against the NIWA monitoring network showed that for quantifying drought at these scales it is appropriate to use a simple water balance. In addition the occurrence statistics of severe drought were not sensitive to the changed plant water use efficiency expected under increasing carbon dioxide, so this process was not factored into the study. The methodology also includes a rigorous evaluation of the ability of climate models to replicate observed drought statistics. This built additional confidence in the approach as the levels of change detected were typically greater than any inaccuracies introduced by the climate models and downscaling.

Detailed findings

This study establishes distinct regional differences across New Zealand in changes to drought vulnerability projected under future climate change—with an increase in drought on the east coast of the North and South Island being the most plausible and consistent outcome. This is consistent with the result of the 2005 study which used Third Assessment Report (TAR) climate models and different downscaling methods. The overall conclusion based on these new estimates concurs with the conclusion drawn from the 2005 study that: drought risk is expected to increase during this century in all areas that are currently drought prone, under both the 'low-medium' and 'medium-high' scenarios.

Based on the new estimates developed in this study, the strongest evidence for increased future exposure to drought was found on the Canterbury Plains where the climate is projected to shift towards a more drought prone setting even under very mild future climate change. An equally strong signal was found for no significant change in drought exposure on the West Coast of the South Island. Although the projections were not as consistent across the different climate models, strong evidence for increases in time spent in drought were also found around Hawkes Bay-Gisborne and Northland extending to the Waikato. This overall geographic pattern of change was largely insensitive to climate model but with small differences between the lower B1 and the A1B-A2 emissions scenarios. For the remainder of the North Island and Nelson-Marlborough on the South Island, the prospect of increased exposure to drought ranged from minimal change through to over a doubling in time spent in drought in these regions depending upon the climate model and scenario considered.

While the general direction and geographic pattern of change found in the 2005 and this study were similar, the magnitude of change is less than reported previously when the full range of outcomes is considered. When considering the lowest (B1) emissions scenario and the lower end of the range projected by climate models, this study found evidence of a plausible milder scenario—small increases in drought confined to the Canterbury Plains. In addition two distinct groups of future scenarios were found given the method of downscaling potential evapotranspiration. The first relied on temperature to downscale evaporation and establishes a future with relatively severe change, a strong shift towards a more arid climate with well over a doubling of future exposure to drought in some regions as early as 2030-50. The second included radiation in the downscaling and establishes a future with a milder degree of change.

General implications

The results obtained in this study point to a number of areas where future adaptation analysis and responses might be targeted. The changing nature of drought through the 21st century highlights that basing response on an historically determined understanding of what is normal will increasingly put Governments and farm managers in a weakened position to manage drought risk. Given the most likely scenario found for increasing drought in eastern regions, planning efforts to ascertain and improve resilience to drought and climate variability are warranted. This is a dual measure that would manage the most likely and less likely low end scenarios found in this research, as well as improve management of current variability. Developing a contingency to manage the upper level scenario of a distinct shift toward a more drought prone climate requires some careful and detailed consideration, as potentially this may warrant the exploration of deeper transformational changes in the primary sector with higher implementation costs.

These new estimates bring into sharp focus the need to assess the vulnerability and sustainable management of New Zealand's irrigation water resources as this will be a key drought management tool in the future, particularly critical in consistently exposed regions like the Canterbury Plains. Resilience based research focussing on adaptation is a much broader undertaking than the climate focussed work undertaken here, and efforts which improve understanding of drought within this broader framework are required to support planning.

Key uncertainties and limitations

All climate change vulnerability analyses are subject to uncertainties surrounding change in the climate system and the response of land based production systems. An uncertainty in all studies is representing regional and local changes in precipitation in a warming global climate. This study quantifies considerable diversity in model projections, and these are a result of technical limitations in models, observation networks and downscaling schemes. Although the levels of accuracy in downscaling are acceptable in New Zealand's agricultural landscape, the diversity of model results between GCM's is wide, with particular importance for the North Island. Downscaling rainfall for the Southern Alps of New Zealand is difficult and the levels of precision quantified in this report suggest more robust approaches are needed for future hydrological studies.

There is debate surrounding the degree to which the model ensemble used in this study captures the full range of outcomes possible in a warming global climate. It may be that equally plausible outcomes exists outside this range and may be quantified in future projections. In addition, the assumption here is to take the average of the model ensemble as the most likely outcome, and there is debate as to weather this is a reasonable assumption. It may be that the upper or lower level scenarios deserve more weight in assessing the overall vulnerability of the country to changes in drought.

This study established two methods for downscaling potential evaporation, and does not identify which approach provides a more reliable or physically plausible projection. While the higher quality downscaling model included radiation in the scheme, future projections may be unreliable due to concerns over the quality of radiation projections from the global climate models. Future projections derived using the radiation downscaling model illustrate a far more moderate change in drought than those found using the temperature only model. This difference based on methodology is consistent with recent international research in this area, which shows that choice of method and assumptions used for determining evaporation can influence the results of scenario analysis.

An important response uncertainty in this study is the degree to which management of land based production systems can be changed to reduce or avoid the negative impacts identified. Through the choice of drought indicators and the analysis approach, this report assumes a fixed response of land based systems under climate change with no adaptation, nor does it factor in current levels of resilience. Irrigation for example is not factored into the assessment. As such, this report is best used for raising risk awareness about potential impacts, but should not be considered the final analysis when assessing how to manage risks and opportunities posed by climate change.

Many of these limitations can be addressed by future research. While there needs to be a focus on actions that improve resilience to drought in such work, progress can also be made to improve the physical climate science supporting vulnerability and adaptation analyses. Some key areas for New Zealand include better understanding and downscaling of Southern Alps rainfall mechanisms, improved monitoring and estimates of evaporation, and continued updating of this type of scenario analyses with physically based regional scale climate modelling and new global emissions scenarios.

1. Introduction

1.1 Drought in a changing climate

Rainfall deficits are a common feature of New Zealand's climatic environment and it is not unusual to experience short duration dry spells as isolated regional level events. Geographically widespread rainfall deficits with durations over one to two months are less frequent, and are usually considered to be 'agricultural droughts' as modern pastoral and crop production systems are vulnerable to rainfall deficits of this scale and duration. Recently the widespread rainfall deficit spanning late 2007 to the end of autumn 2008 was estimated to cost the New Zealand economy around $2.8 billion (MAF 2009), mainly from negative on-farm impacts but also from smaller but detectable negative impacts on regional economies.

Given the occurrence of agricultural drought and the importance of the primary sector to the New Zealand economy, the prospect of future global level climate change modifying exposure to drought warrants careful examination. Current global level assessments suggest that droughts are expected to both increase and decrease under future climate change depending upon geographic location (Wang 2005). Generally, as the global climate systems warms it is expected to adopt a more tropical structure, with increased rainfall at the equator and in higher latitudes, but drying in mid latitudes. This structure is illustrated by the geographic distribution of projected winter rainfall change in 2080-99 in Figure 1. New Zealand spans the higher latitude zone where rainfall is projected to increase, but also touches the mid latitudes where rainfall is expected to decrease.

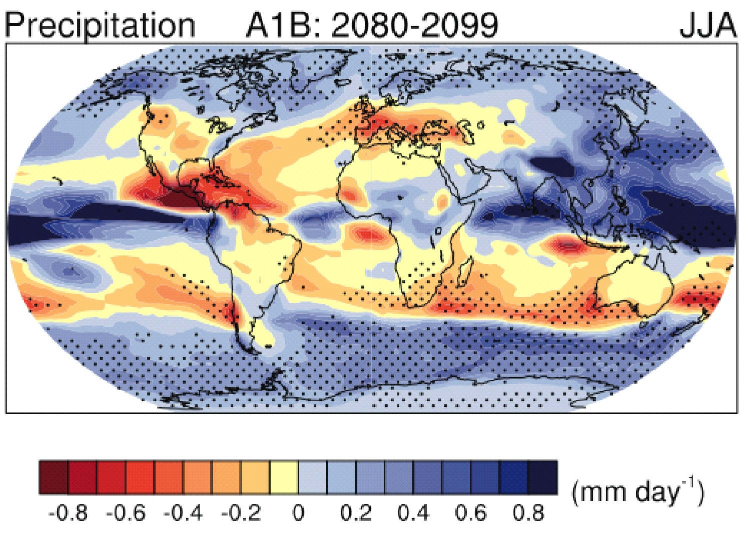


Figure 1: The global distribution of winter (JJA) rainfall change for the A1B emissions scenario. Colour is the mean of 21 AOGCMs. Dots illustrate where models are more consistent. Source: Meehl et al. (2007).

As this global pattern will be modified by New Zealand's maritime climate and complex topography, a more detailed view than in Figure 1 is required to develop plausible scenarios of future drought occurrence for this region. The key finding of previous NIWA work which developed regional scenarios of drought in New Zealand (Mullan et al. 2005; Porteous 2005) was that under 'low-medium' and 'medium-high' regional climate change scenarios: 'Drought risk is expected to increase during this century in all areas that are currently already drought-prone'. The range of change was from a doubling of drought frequency by 2080 in the most exposed areas (eastern regions of both Islands) for a low-medium scenario to almost a four fold increase in the most exposed areas under a medium-high regional climate change scenario.

1.2 Stakeholder consultation

The authors of the 2005 study warned that these projections were subject to uncertainty and listed some of the constraints and assumptions of their study. Key uncertainties were further discussed and opportunities to improve the analysis identified at a stakeholder consultation workshop in 2008, reported in Clark and Tait. (2008). Larsen (2005) also provides a critique of some aspects of the methodology, focussed on the example analysis on the Canterbury Plains reported by Porteous (2005). The main directions emerging from this consultation and critique include:

- update the Mullan et al. (2005) drought analysis, which used model results from the IPCC's Third Assessment Report (TAR), by using the IPCC's Assessment Report Four (AR4) climate modelling;
- use a more process based approach to simulating soil moisture for calculations of potential evapotranspiration deficit and other drought indicators;
- utilize newly constructed national standard data sets to develop drought indicators and implement a climate change analysis methodology;
- investigate if improvements can be made to the current empirical downscaling methods;
- utilize data emerging from NIWA's regional climate modelling program to investigate changes in drought risk; and
- continue to integrate climate change scenarios with agricultural impact models.

1.3 Project objectives and scope

This report documents research undertaken by NIWA during 2009/10 aimed at responding to this stakeholder feedback. NIWA submitted a project proposal to the research foundation (PROP-20264-SLMACC-NIW) under the Ministry for Agriculture and Forestry 1 year program - Sustainable Land Management Mitigation & Adaptation to Climate Change. The project was supported with the following aims and objectives:

- use the latest (2007) IPCC global model projections, which lead to some differences in projected changes in rainfall seasonality and spatial distribution over New Zealand compared to the IPCC (2001) projections used in previous studies;

- devise regional scenarios from a much wider range of climate models (12 or more global models, compared to just 2 models in the 2005 study), and development of a risk assessment methodology based on these results;
- improve downscaling of potential evapotranspiration (PET) to explicitly include solar radiation as a predictor;
- compare a physically-based calculation of potential evaporation (PET) with the method used with the previous work, for example derive the full Penman equation using daily meteorological variables output by NIWA's regional climate model;
- consider a more complex physically based soil moisture balance calculation and different drought indicators, and assess this against the Potential Evaporation Deficit index and simpler water balance used in the previous study;
- undertake a sensitivity analysis of plant responses to climate change, including stomatal resistance increases under higher carbon dioxide levels; and
- produce more regional information in terms of maps and time series plots of changes in drought risk through the 21st century.

There were a number of broader issues identified at the stakeholder consultation workshop in 2008 that fall outside the scope of this study. Specifically:

- assessments of vulnerability and specific adaptations to manage drought given a socio-economic perspective;
- continued improvement to the risk communication of drought, for both tactical management through to development of drought decision support systems;
- research focussed on steps needed for improved drought management;
- assessment of hydrological drought , particularly the availability and status of irrigation water from groundwater and surface water resources; and
- improvement in the monitoring of baseline climate data, particularly in catchments of national importance and alpine regions which are currently under sampled.

1.4 Limitations of this study

No one analysis of drought will provide a satisfactory account for every stakeholder with interest in this phenomenon. As described in Chapter 4 this is a well established aspect of drought research, more formerly termed 'operational indeterminacy'. Hence, this research project has a number of acknowledged limitations, specifically:

- the study examines medium to long term drought scenarios given anthropogenic forcing of global climate. In an historical context, the previous 5-7 years has seen an increased drought frequency across New Zealand's agricultural land. The degree to which this is attributable to anthropogenic forcing compared to other factors including internal forcing from the El Niño

Southern Oscillation (ENSO, see Mullan 1995; Kidson and Renwick 2002), the Interdecadal Pacific Oscillation (IPO, see Thompson 2006) and Southern Annular Mode (SAM) are not analysed in this study. This is an important question which would require a substantially different approach and set of resources to examine;

- the study takes a predominantly climatic view of drought, extending this to generalised agronomic indices that account for rainfall effectiveness. This provides a partial estimate of total drought risk by examining the likelihood of climatic events only. No attempt has been made to quantify impacts on physical production or the cumulative economic impacts of droughts or their management. Understanding these effects would also require a more global change approach to account for carbon dioxide fertilisation during times when water is non-limiting. Constructing scenarios using a full definition of drought risk has high relevance, but would require significant investment to develop an integrated assessment methodology;

- as this is a climate focussed study, the range of factors influencing the vulnerability and resilience to drought in the New Zealand agricultural sector are not examined. These include: farm economic factors like debt, equity and the degree of profit orientation for a given farm business; farm physical factors like system intensity and agronomic drought management; and social factors including education, innovation models, mental health and the capital provided by networks across communities (White 2009). Even if climatic vulnerability or exposure to drought remained stable or within expected historical variability, changes to these factors could increase the total vulnerability of farm businesses, communities and families to drought impacts;

- the research does not explicitly examine hydrological drought, as the indicators used do not measure surface and or ground water resources. There is need to develop scenarios of hydrological drought for New Zealand on a nationally comprehensive basis, as undertaken by Bright et al. (2008) for central Canterbury ;

- the methodology used is an example of top-down climate change impact analysis. This approach provides both a nationally and regional level assessment. However, local scale factors that are not fully represented in the study can influence drought vulnerability. Examples are between and within paddock soil variability, and the extent to which different farm management practices modify drought exposure under the same regional scale climate regime;

- in order to examine a wide range of climate change scenarios the analysis in this report is based primarily on empirical downscaling of climate models. This reflected delays in commissioning of the NIWA supercomputer where work to dynamically downscale climate from a physically based approach is underway. Robust and comprehensive dynamic downscaling was not yet available at the time the analysis was undertaken;

- in New Zealand rainfall deficits are formerly declared droughts by a consensus approach by Central and Regional Governments. The drought probabilities quantified by objective indicators in this study could differ from those determined by these processes; and
- there are ongoing limits in all global and regional scale climate change scenarios studies relating to the dynamics of and change to precipitation regimes. While small steps are taken to improve the representation of precipitation variability in this analysis, there remains scope to improve this aspect of scenario studies.

1.5 Guide to chapter structure

The research is presented as four chapters, and readers not interested in the work undertaken to develop the methodology may prefer to go straight to Chapters 2 and 3. Chapter 2 provides an overview summary of the work specifically for a general audience. Chapter 3 provides a more detailed presentation of the final results, including a short summary of the methodology. Chapters 4 and 5 detail the development of the methodology including: a review of drought indicators (Chapter 4); an examination of water balance models including verification (Chapter 4); and development of regional scale climate scenarios (Chapter 5).

2. Regional scenarios of drought under climate change: an overview

Key outcomes

- Based on the latest climate and impact modelling, New Zealand can expect more droughts in the future in some locations. The new evidence described in this report highlights key agricultural regions on the Eastern seaboard particularly the Canterbury Plains, as well as Northland, as being the most consistently vulnerable areas.
- Although there is more uncertainty, other regions in the North Island also experience similar vulnerabilities, including key primary industry regions like the Waikato.
- There is a wide range of potential changes in the frequency of major drought events, dependant on global emissions scenario, climate model and location. It is now possible to ascertain the likelihood of changes grouped as a most likely middle range scenario, as well as least likely upper and lower scenarios.
- Although a wide range in potential outcomes were quantified in this study, the A1B emissions scenario (Figure 2) provides a useful guide to the direction and likely magnitude of changes.

Future scenarios of drought

Moderate increases in drought—a more likely scenario

The more likely scenario is that New Zealand can plan for around ten percent additional time spent in drought by the middle of this century for key eastern agricultural regions and parts of the North Island. The main locations of change include the Canterbury Plains stretching into eastern Southland as well as most of the North Island in the first half of the century (2030-2050). The expectation created by the projections is that this overall pattern will intensify in the second half of the century (2070-2090). This is the mid range outcome when considering 19 global climate models and three emissions scenarios.

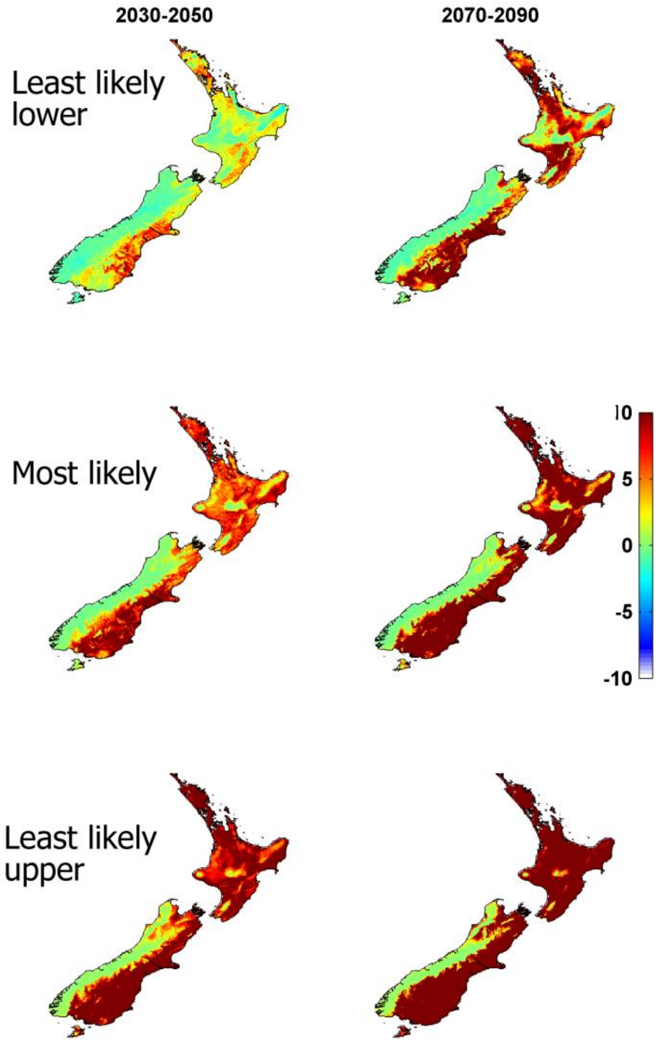


Figure 2. Projected increase in percentage of time spent in drought from 1980-99 levels for the A1B emissions scenario. Results summarise 19 global climate models.

Widespread severe change—a less likely scenario

A least likely but severe change scenario is found when the more extreme projections of drought developed in this study are examined. Here New Zealand experiences a distinct shift towards a more arid environment, with well over 10 percent additional time spent in drought. This level change is also widespread across the country affecting all regions except the West Coast of the South Island.

Mild change constrained to key regions—a less likely scenario

A less likely but identifiable lower range drought scenario is also evident in the projections. Here any increased time in drought during the first half of the century is well constrained below an additional 10 percent and confined to the Canterbury Plains, Hawkes Bay to Northern Wairarapa, parts of the Waikato, and Northland. These changes intensify in the second half of the century, becoming slightly more widespread but are still constrained to these regions.

Responding to scenarios

Appropriate use of scenarios relies on understanding that they are not predictions of the future, but plausible descriptions of the range of what could happen given current levels of knowledge. Specific responses to these scenarios need to be explored throughout the primary sectors, but principally occur at the farm, industry organisation and government levels. Translating these new drought scenarios into tangible outcomes requires detailed planning and implementation by managers, and it is not appropriate to be prescriptive. However the new scenarios suggest that managers could focus on two strategic areas when framing specific responses:

1. No regrets planning for the most likely and least likely low end scenario, which would also include the increased frequency of droughts experienced in New Zealand in the 2007-2010 period. This should focus on known adaptations, and the extent to which small modifications to current primary industry practices can manage drought. Immediate implementation would benefit industry now, in terms of managing current variability.

2. Development of a contingency to manage the least likely upper end scenario. The severity of this scenario may warrant high levels of innovation to be considered in the mix of adaptive responses, thereby developing completely new approaches to management of drought in the New Zealand primary sector. Because of likely high implementation costs, formal consideration of the costs and benefits of this type of adaptation need to be undertaken.

Uncertainties and limitations

Developing the new drought scenarios has relied on research undertaken by NIWA into the methodologies to quantify drought and downscaling climate scenarios for New Zealand. This research was undertaken in response to stakeholder consultation surrounding a previous set of drought scenarios developed in 2005. But like all studies of this nature there are some ongoing uncertainties and limitations surrounding these new scenarios. The key limitations are:

- In order to generalise across the country the study takes a predominantly climatic view of drought, only examining changes to large scale events greater than one month in duration. A production focussed view of drought would require a more specific focus on individual industries, a focus on evaluating impacts at a farm or production unit level, as well as consideration of other factors like carbon dioxide fertilisation when water is non-limiting.

- A key ongoing uncertainty is the exact behaviour of the climate system, and more specifically the ability to model regional to local changes in precipitation in a warming global climate. This is reflected in the range of outcomes quantified in this study. There are also numerous technical and theoretical uncertainties associated with convective rainfall and cloud feedbacks. This is an ongoing area of research in the broader climate research community.

- There are differences in the drought projections depending upon the method used to downscale potential evaporation. When radiation is included in the downscaling scheme a more conservative projection was found than when temperature only is considered. While the inclusion of radiation yielded an improved downscaling model, future projections based on it may be unreliable because of concerns over data for radiation projections from the climate models.
- There are also ongoing uncertainties relating to the response of primary production systems to drought. Through the choice of drought indicators and the analysis approach, this report assumes a set management response under climate change with no adaptation. It also does not quantify current levels of resilience. An approach which factors in capacity to adapt is warranted in future research.

3. Scenarios of drought under climate change

3.1 Summary

This chapter describes future projections of drought for New Zealand, based on the work to develop drought indicators and climate change scenarios described in the previous chapters. Despite a range of influences from different climate models, emissions scenarios and downscaling methods, the most common direction of change was toward increased exposure to drought in some key agricultural regions of New Zealand. The analysis confirmed the geographic pattern of change found by Mullan et al. (2005), as regions on the east coast have consistently increased drought in future projections. Similarly there is consistently no change in drought projections for the west coast of the South Island. Although not as consistent as found in these regions, strong evidence for increased likelihood of drought in the future was observed in the Hawkes Bay-Bay of Plenty region as well as Northland extending to the northern Waikato. Projections for the remainder of the North Island and parts of the South Island have a wider range, from minimal or no change through to over a doubling of drought probability by the end of this century.

This work also establishes that the magnitude of increase reported by Mullan et al. (2005) is plausible, of at least a doubling in drought risk by the end of this century in the most exposed regions. Given the exploration of a broader set of emissions scenarios and models in this study, a plausible scenario of milder change below a doubling in the more exposed regions is also identified. Two groups of drought projections were found based on methods used to downscale potential evaporation:

- using a temperature driven approach, there is a distinct shift toward aridity where the most severe climate models showed drought risk could increase by well over two fold by as early as 2030-2050 in the most exposed regions.
- introducing radiation established a plausible scenario for milder change in drought probability. The least severe climate models showed minimal or in some regions decreased drought probability as late as 2070-90.

These can be used to define the less likely absolute upper and lower range of outcomes around the more likely scenario previously identified.

The projections of future drought developed in this study are best considered and applied in the context of scenario planning that is firmly grounded in risk management. Some general guidance is provided surrounding steps that will help to achieve this outcome.

3.2 Methodology

The methodology for this analysis is built on work described in more detail in Chapters 2-3. A short summary of this work is presented here to illustrate how the final analysis is constructed.

3.2.1 Drought indicator

The analysis is based on a dry land agricultural definition of drought which uses a soil water based indicator based on the Phillips and McGregor Drought Index [Table 2 and section 4.2.6]. This is a compromise between a purely meteorological indicator based on rainfall, hydrological indicators examining water resources and use of trialling agricultural drought indicators like pasture growth or a socio-economic index. Management factors which influence drought vulnerability and resilience have a greater bearing on these later drought indices, so it was deemed that they are not a practical choice for the generalised analysis pursued in this study.

While meteorological and agronomic indicators are useful, they do not always describe the impact of droughts so they provide only a partial estimate of risk. No matter which drought indicator is used they can all be questioned in terms of their relevance, particularly by those seeking information outside of the context in which a drought analysis is undertaken. This is a well known aspect of drought research more formerly termed ‘operational indeterminacy’ [section 4.2.3]. Soil water with no other transformation is used in this study because it is relatively simple and transparent, is practical and consistent with constructing nationally comprehensive future scenarios, provides a universal metric across many industries, and approximates rainfall effectiveness.

The drought indicator is based on a simple one layer soil water balance that can be parameterised to reflect the main soil controls on drying and field capacity [section 4.3]. This model was chosen in favour of a more complex framework because of its demonstrated level of precision for quantifying observed soil water. While there is evidence that more complex models can be more precise, the level of precision obtained by this simpler model was adequate for quantifying drought at the durations and intensities specified below. More complex frameworks also come with the added cost of parameterisation, limiting their extension in a national framework. This model is used instead of NIWA’s current operational framework that was also used in the previous study. The new model provides a small improvement given the ability to parameterise soil control on drying. For national scale simulations the model was set up with a reanalysis of the Fundamental Soil Data Set, where field capacity and the soil drying coefficient were derived at the 0.05 degree resolution of NIWA’s Virtual Climate Station Network (Tait et al., 2006).

The drought indicator normalises for site specific variability by using historically determined thresholds to signify drought entry and termination [section 4.2.6]. This assumes that agricultural systems are geared to each environment through farm size, economies of scale and other adaptations. This was achieved by modifying the Phillips and McGregor drought index. Drought entry occurred when soil moisture was below the historically established 10th percentile for the given time of the year for a period greater than one month. Drought termination occurred when soil moisture was above the 10th percentile for one month. The historical thresholds are derived from a base 1980-1999, and these thresholds are also used in the calculation of drought probability in the future scenarios. The thresholds are based on a pragmatic view that events of this scale become significant in New Zealand primary sectors—they are relatively severe events in the context of the New Zealand climatic environment. While shorter duration droughts (1-2 week events) do feature in New Zealand, they may not be as significant across the country in terms of nation wide impact. Once drought events are defined using these thresholds, drought probabilities are determined using a non-parametric bi-modal probability density function of duration and intensity, with the final probability determined by integration. These are transformed to represent the percentage of time that a given site is deemed to be in drought. The technical approach is based on Kim et al. (2003), Gonzalez and Valdes (2003) and Ramsay et al. (2008), but adapting algorithms described by Martinez and Martinez (2002).

3.2.2 Simulation design

There are many different factors and sources of influencing the projections, and these are clearly defined by the overall experimental design in Table 1. It is structured as: a *two* (PET method), by *three* (emissions scenario), by *two* (time slice) by *nineteen* (climate model) design. The variable analysed across this design is the additional percentage of time spent in drought from a base period defined as twentieth century climate from 1980-99.

Table 1: Design used in the drought scenario experiment. .

	VARIABLES			
	PET Downscaling Method	**Emissions scenario**	**Time slice**	**Global Climate Model**
FACTORS	1. PLS regression of MSLP and surface temperature 2. PLS regression of MSLP and surface temperature and radiation	1. B1 marker scenario 2. A1B marker scenario 3. A2 marker scenario	1. 2030-50 2. 2070-90	1. BCM2 2. CGMR 3. CNCM3 4. CSMK 5. ECHOG 6. FGOALS 7. GFCM21 8. GIAOM 9. GIER 10. HADCM3 11. HADGEM 12. IPCM4 13. MIHR 14. MIMR 15. MPEH5 16. MRCGCM 17. NCCCSM 18. MCPCM 19. NCPCM

Note: See Table 1 in Mullan et al. (2011) for further description of these GCMs. Only the 19 models listed here (of the total of 24 available) had archived sufficient output data to apply the downscaling algorithm and drought calculation.

The emissions scenarios used in this study represent two separate 'higher end' scenarios (A2 and A1B) and a lower emissions scenario (B1) from the Special Report on Emissions Scenarios (SRES). The main difference between the two higher end scenarios is the timing of peak emissions, a result of different underlying population growth patterns and technologies. This is shown clearly by the global temperature projections in Figure 3 where at around the middle of the century the A1B marker scenario represents the more extreme level of warming. By the end of the century the A2 marker scenario has the greatest warming.

Each specific scenario is drawn from a particular emissions story line where a range or family of scenarios have been specified (SRES, see Nakicenovic and Swart 2000). Each story line represents a particular pathway where the global economy balances development with emissions reductions. The SRES story lines are described generally as:

- A1 storyline: a future world of very rapid economic growth, global population that peaks in mid-century and declines thereafter, and rapid introduction of new and more efficient technologies.
- A2 storyline: a very heterogeneous world with continuously increasing global population and regionally oriented economic growth that is more fragmented and slower than in other storylines.
- B1 storyline: a convergent world with the same global population as in the A1 storyline but with rapid changes in economic structures toward a service and information economy, with reductions in material intensity, and the introduction of clean and resource-efficient technologies.

The greenhouse gas concentrations used as anthropogenic forcing in climate models are taken from illustrative marker scenarios, a subset of 6 of the 50 emissions scenarios available across all the storylines. For the Coupled Model Intercomparison Project (CMIP3) which informed the IPCC's AR4, three of the six marker scenarios were used to run climate models (B1, A1B and A2). These are the emissions scenarios adopted in this study. Hence the scenarios used here do not represent either the most extreme or lowest rate of climate change that would be present if all the SRES emissions scenarios were used to force climate models. Potentially the different climate futures derived in the CMIP3 experiments may not necessarily represent the full range of plausible future climates under anthropogenic forcing. Similarly the SRES marker scenarios do not directly represent mitigation policy actions to manage current and future emissions.

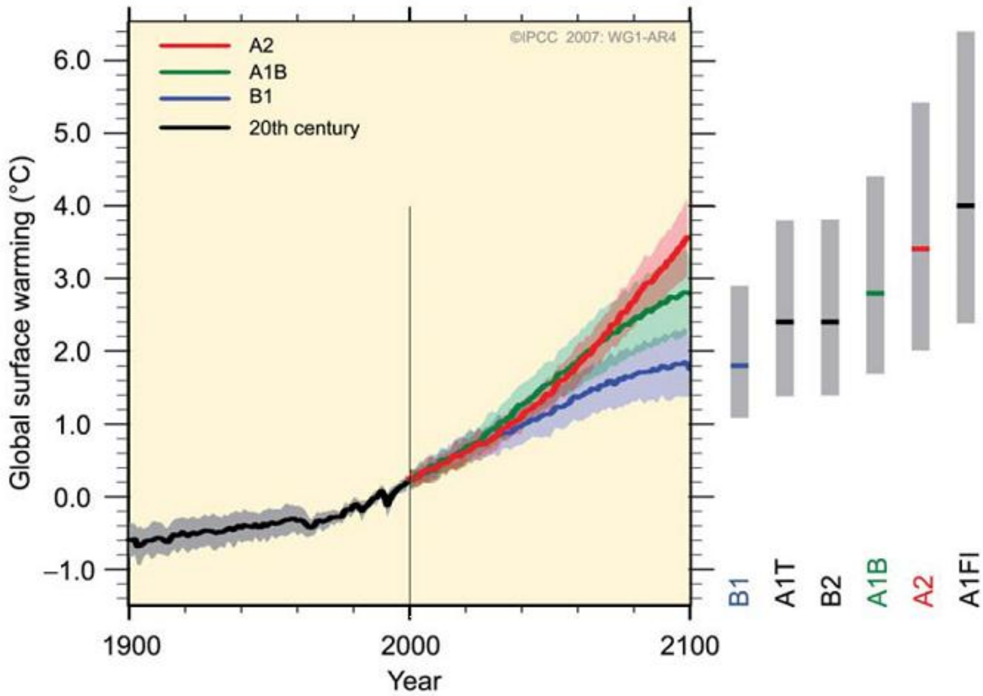


Figure 3: The rate of future global surface warming projected by climate models under the three marker scenarios B1, A2 and A1B. Source Meehl et al. (2007).

3.2.3 Downscaled climate change scenarios

For each PET downscaling method, future scenarios and twentieth century base climates were developed for 19 different climate models. The future scenarios cover the period 2000-2100 for the three emissions scenarios. The variables were monthly rainfall and potential evapotranspiration (PET) which are inputs to the water balance. The GCM data are from CMIP3 (Meehl et al., 2007) and downscaled using empirical transfer functions developed by Partial Least Squares (PLS) regression [Chapter 5]. These are based on relationships between broad scale data from the NCEP reanalysis and finer scale gridded climate data from the Virtual Climate Station Network [sections 5.4.6 and 5.4.7]. A temporal downscaling method is also used that allow the month-to month variability across the future 100 years produced by GCMs to be preserved [section 5.5]. This provided a comprehensive set of future climate scenarios across the 5 km grid for New Zealand. Evaluation of the ability of these scenarios to replicate 20th Century (1980-99) drought probabilities suggests that although there are biases introduced by the downscaling and the GCMs, they are acceptable for quantifying drought.

Two downscaling models for PET are used as no approach was clearly superior to the other [section 5.4.7]. Based on tests with observations, downscaling based on Mean Sea Level Pressure (MSLP), surface temperature and radiation appeared to provide an improved model over the MSLP and surface temperature only approach. However, when resolved with twentieth century GCM data there were minimal differences between the two, and for some GCMs the radiation model appeared to introduce greater bias. Hence it isn't clear to what extent radiation represents a redundant variable in projection studies, and there is risk that it could introduce spurious results given the limited ability of some GCMs to adequately represent this surface variable.

3.2.4 Constructing drought scenarios

Developing scenarios in an experiment with multiple climate models, emissions scenarios and downscaling methods creates a data intensive problem and careful computational design is required. It was estimated that the full experiment would involve the manipulation of around 6000 million data points. To develop a practical system the experiment is run by computationally integrating the algorithms which resolve the downscaling, the water balance and calculate the drought probabilities. A full country wide simulation including 20th Century (1980-99) and 100 years of future scenarios for B1, A1B and A2 takes about 2.5 hours of CPU time on a standard unix terminal. This type of integration also means that variables and data used in intermediate calculations, such as the climate inputs to the model, soil water deficit time series and drought indexes, are not routinely retained and stored. Only the initial inputs, including the downscaling coefficients, broad scale model fields and the final outputs are saved in a data structure. To provide a convenient basis for reporting two time slices are used, 2030-2050 represents a medium range time frame and 2070-2090 a long range time frame. Where relevant transient simulations are also analysed and provided.

In most analyses the results are reported as percentages, representing the deviations in drought probability from the 1980-99 baseline. Thus the changes reported are additional to existing 20th century exposure to drought mapped in Figure 48. For example, a site which is currently in drought 5 percent of the time may have a projected change of 10 percent, leaving a final drought probability of 15 percent. The final scenarios are also a data rich source of information and require a considered approach to ensure that analysis is both practical and succinct. Different approaches and levels of aggregation are used to assess and report the changes to drought probabilities under future scenarios, and these are described in the relevant parts of the results section.

Analysis of multiple GCMs ensembles as undertaken here requires careful consideration as a number of methodologies are available (Carter 2007; Knutti et al. 2010). Some strategies include from simple averaging of all GCMs, screening of GCMs to exclude those that are deemed less reliable through to weighting of GCMs based on a measure of quality. Discussion of techniques focussing on climatology applications can be found in Murphy et al. (2004); Katz (2002) provides an overview of approaches used in regional scale impact studies; and Ghosh and Mujumdar (2010) provide an example of non-parametric methods used for GCM scenario analysis of droughts. New approaches are emerging such as Bayesian likelihood analysis which attempt to formally quantify uncertainty as part of formal model verification (see Annan and Hargreaves 2010; Knutti et al. 2010). As presented in the next section a range of techniques are used here, from derivation of probability density functions through to quantifying the range and central tendency of the model ensemble used.

3.3 Results

3.3.1 Base drought risk

Late twentieth century (1980-99) drought probabilities are shown in Figure 48 as the percentage of time spent in drought. This is the base period of the study, and the future projections are expressed as the additional level of drought risk above or below that described in Figure 4. This reveals the expected geographic distribution of New Zealand droughts, with higher levels of exposure on the eastern sea board and key regions like the Canterbury Plains, Gisborne-Hawkes Bay and Northland being some of the most exposed regions in New Zealand. There is also a degree of geographic complexity, evidence of sharp gradients of change over small distances and the presence of climatic niches, in terms of the reliability of climate across the country. It also suggests that some regions not normally considered to be exposed to drought have higher levels of exposure than might be anticipated. For example west coastal regions like Taranaki and Buller-Greymouth experience similar drought exposure to parts of the Canterbury Plains. This re-enforces a key assumption of the drought indicator used in this study, where drought is calculated relative to historical thresholds [section 4.3.4]; namely that enterprises experience drought exposure relative to each environment.

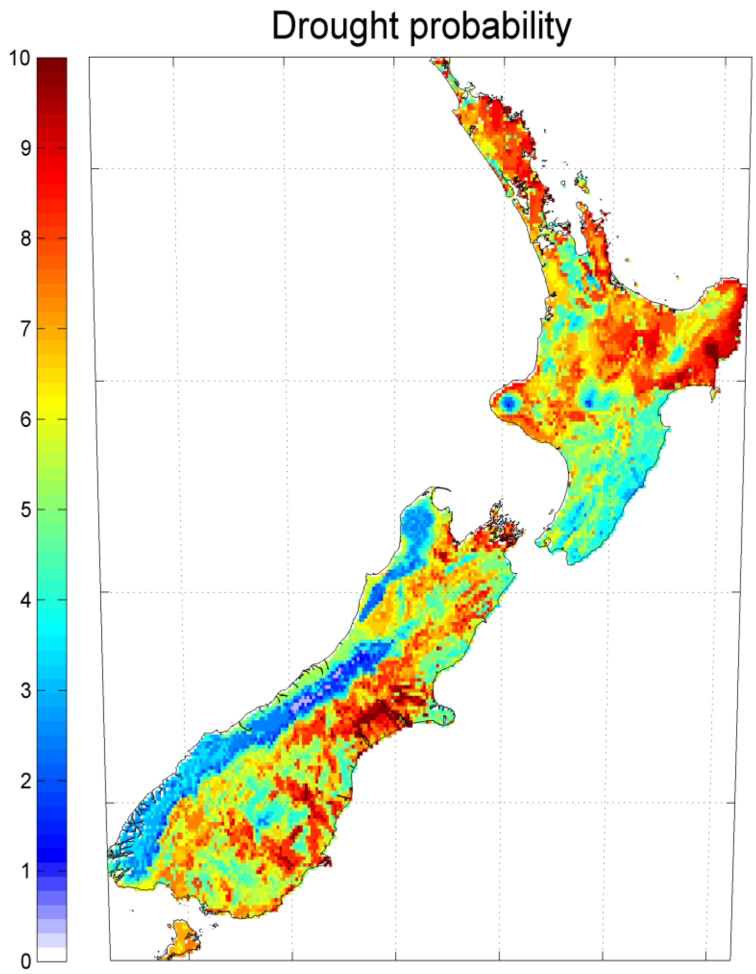


Figure 4: Base 20th century drought probabilities expressed as percentage of time in drought for the period 1980-99.

3.3.2 Probability density functions of change

The scenarios describing additive changes to twentieth century drought exposure are summarised nationally in Figure 5 and Figure 6. These are probability density functions

(PDFs) depicting the likelihood of change stratified by emissions scenario and time slice. They summarise the range of results for the 19 different models for all sites across the country. Probability density functions are used because they summarise a large range of data succinctly—each graph is based on around 650 000 individual data points.

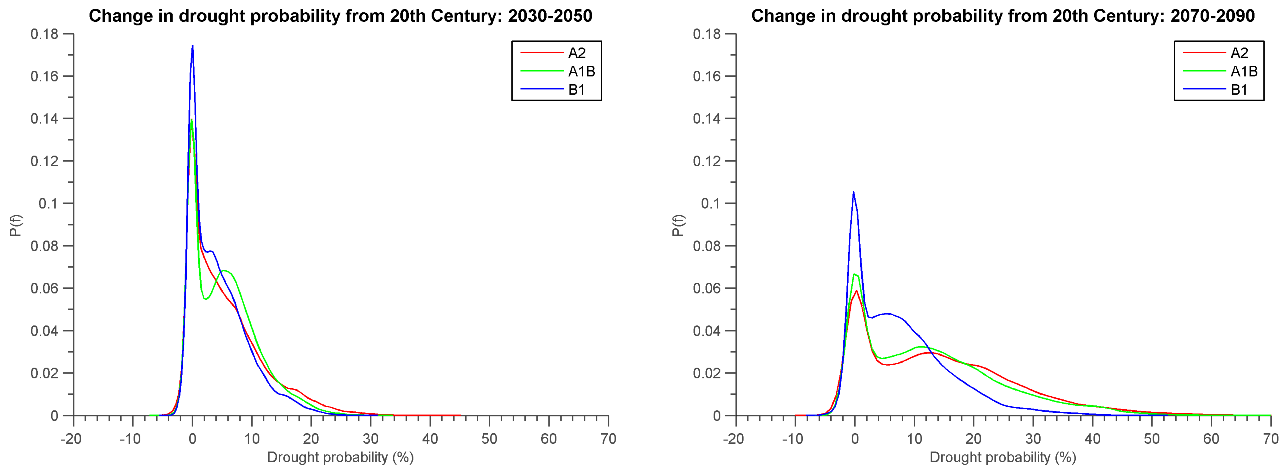


Figure 5: Probability density functions of projected changes to drought probabilities, accumulated over all of New Zealand, and based on the temperature PET downscaling method.

The scenarios developed using the temperature based PET downscaling (Figure 49) indicate a bi-modal effect, where the most probable outcome is for no significant change in drought probability, but with a large chance of increasing drought exposure. In the 2030-50 time slice the range of positive change spans a 5-20% addition to the time spent in drought, with a detectable but low chance of more severe change. There are slight differences between the emission scenarios with B1 providing the mildest degree of change. A2 and A1B are almost identical except for small differences like the higher density of change at around 10 percent for A1B, and the very small probability of large increases at around 40 percent for A2. The degree of positive change becomes stronger and the differences between scenarios more pronounced in the 2070-90 time slice. Changes in drought probability of 10 percent or more, and at worst up to 50-70 percent in the 2070-90 time frame, represent a very strong shift toward are more arid environment. Under this scenario large parts of New Zealand are exposed to more prolonged and intense droughts that are well outside drought conditions experienced over the twentieth century period (1980-99).

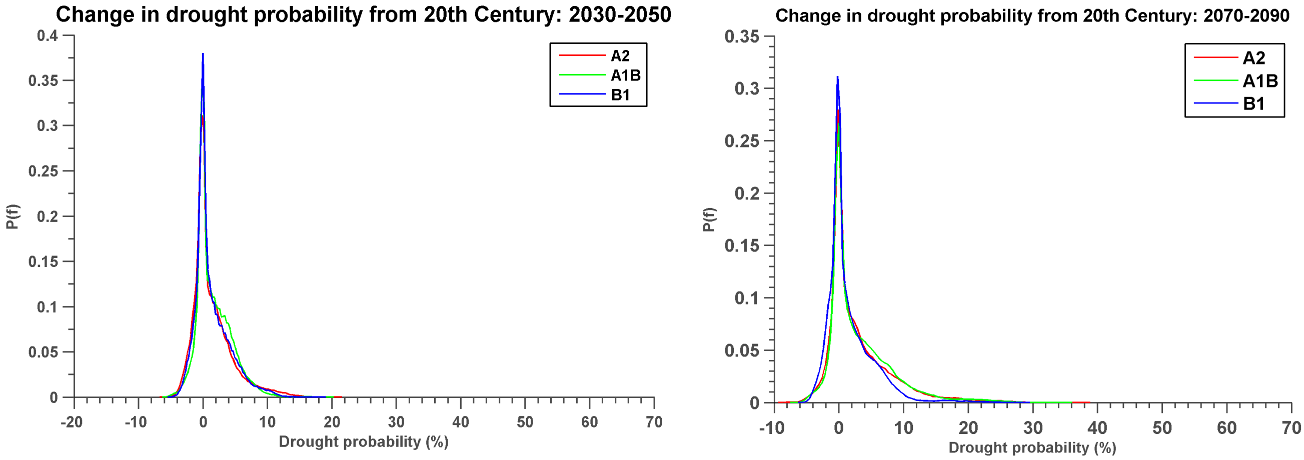


Figure 6: Probability density functions of projected changes to drought probabilities accumulated over all of New Zealand, and based on the radiation PET downscaling method.

The probability density functions summarising the results of the radiation PET downscaling method illustrate a smaller change in drought than the scenarios from the temperature based downscaling method (Figure 50). Like the previous result the highest probability density was found around the zero percent change. But there was also a detectable increase in drought probability, in the order of 3-10 percent for the 2030-50 time slice and 5 to 15 percent at 2070-90. Increases of the order of 20 percent at 2030-50 and 30 or more percent at 2070-90 are detectable but not frequent. Small differences are evident between the scenarios, with B1 having the mildest change. A2 and A1B have a similar PDF, except for higher magnitude changes were the A2 appeared to slightly increase the probability of change at this level. While these scenarios show a lower order magnitude increase in drought than those developed with the temperature PET downscaling, the levels of positive change are not inconsequential. An additional 5-10 percent of time spent in drought in some locations by 2030-50 represents almost a doubling over that experienced in the 1980-99 base period. The 5-20 percent increases found by 2070-90 are well over a doubling.

At this stage it is useful to reflect on how the projected magnitude of changes compares to the downscaling and GCM biases quantified in the previous chapter. Generally the positive shifts detected in the order of 5-10 percent are larger than the biases on agricultural land quantified of around 2-5 percent in the temperature downscaling method and 1-2 percent in the radiation downscaling method. This leads to the observation that the climate change signal detected in this analysis is greater than the bias quantified by downscaling and GCM. However, as there was a distinct spatial distribution in the bias it is appropriate to examine the scenarios geographically to further examine confidence in this observation.

3.3.3 Geographic analysis of change

A mapping analysis of the scenario experiment is presented in Figure 51-Figure 52. Here the 10th, 50th (median) and 90th percentiles of the 19 model ensemble are mapped for each emissions scenario and time slice. These levels represent:

- the high end (90th percentile) of the climate model ensemble;
- the middle (median or 50th percentile) or most likely projection; and
- the low end (10th percentile) of the climate model ensemble.

In the literature there are different methods and theoretical assumptions being developed and evaluated for analysing large multi-model ensembles (see Annan and Hargreaves 2010). This study adopts an approach which assumes that the middle of the ensemble represents the most likely projection in terms of climate model outcomes, with the high and low ends representing plausible but less likely projections. To allow comparability, the contour range in each map is restricted to a range of plus to minus 10% additional time spent in drought beyond the base climate. As shown by the probability density functions there are positive changes well above 10%, so in the maps large changes above this level take a deep red colour.

A key observation across all the results is the consistent spatial pattern of change, where for the most part an increase in drought probability is found for the Canterbury Plains. A similarly consistent result for little change in drought was found on the West Coast of the South Island. This geographic pattern of change was evident largely irrespective of climate model, scenario and time slice. Although the signal is not as consistent as found for Canterbury, droughts also increased for most models on the east coast of the North Island and in Northland. This is the same spatial pattern of change found previously by NIWA (Mullan et al. 2005). The region where the scenarios diverge the most is the remainder of the North Island, which ranged from minimal change through to dramatic increase in exposure to drought. This is not unexpected given the location of the North Island in the context of the global scenario (refer to Figure 1), a shift to a more tropically structured climate with mid latitude drying—and the known geographic and resolution based inconsistencies between the GCMs in identifying the precise region to which the mid-latitude rainfall decrease extends south.

For the temperature based downscaling experiment (Figure 7) drought increases of 5-10% (additional time in drought) across Canterbury were found, even for the least severe projection, the low end of the GCM ensemble at 2030-50 for the lower B1 emissions scenario. For this region the increase in time spent in drought and the geographic extent of this effect increased with GCM, scenario and in the longer term (2070-90) time slice. While it is the Canterbury Plains where agricultural land use is concentrated which appears to be consistently exposed to increased drought, this result does not hold for many of its main surface water catchments located in the Southern Alps. At the low end of the climate model ensemble little change is projected in the catchment areas, but for the mid and high points of the ensemble increases of 10% or greater are projected. This observation holds for all scenarios and time slices, except for the A2 and possibly the A1B low end of the climate model ensemble at 2070-90.

The diverse range across the GCM ensemble for the North Island (outside of the east coast) identified previously is pronounced in this experiment, with projected droughts ranging from little or small change at the low end of the climate models through to large often greater than 10% increase at the high end. At the mid range of the model ensemble, projected change on the North Island has a complex geographic pattern with regions of little or no shift through to large (greater than 10%) increases in drought exposure. Northland provides a good example of this diversity, with little or no change (0-2%) at the low end of the GCM ensemble, moderate (5-10%) change in the mid range and large increase (greater than 10%) in drought exposure at the high end of the GCM's.

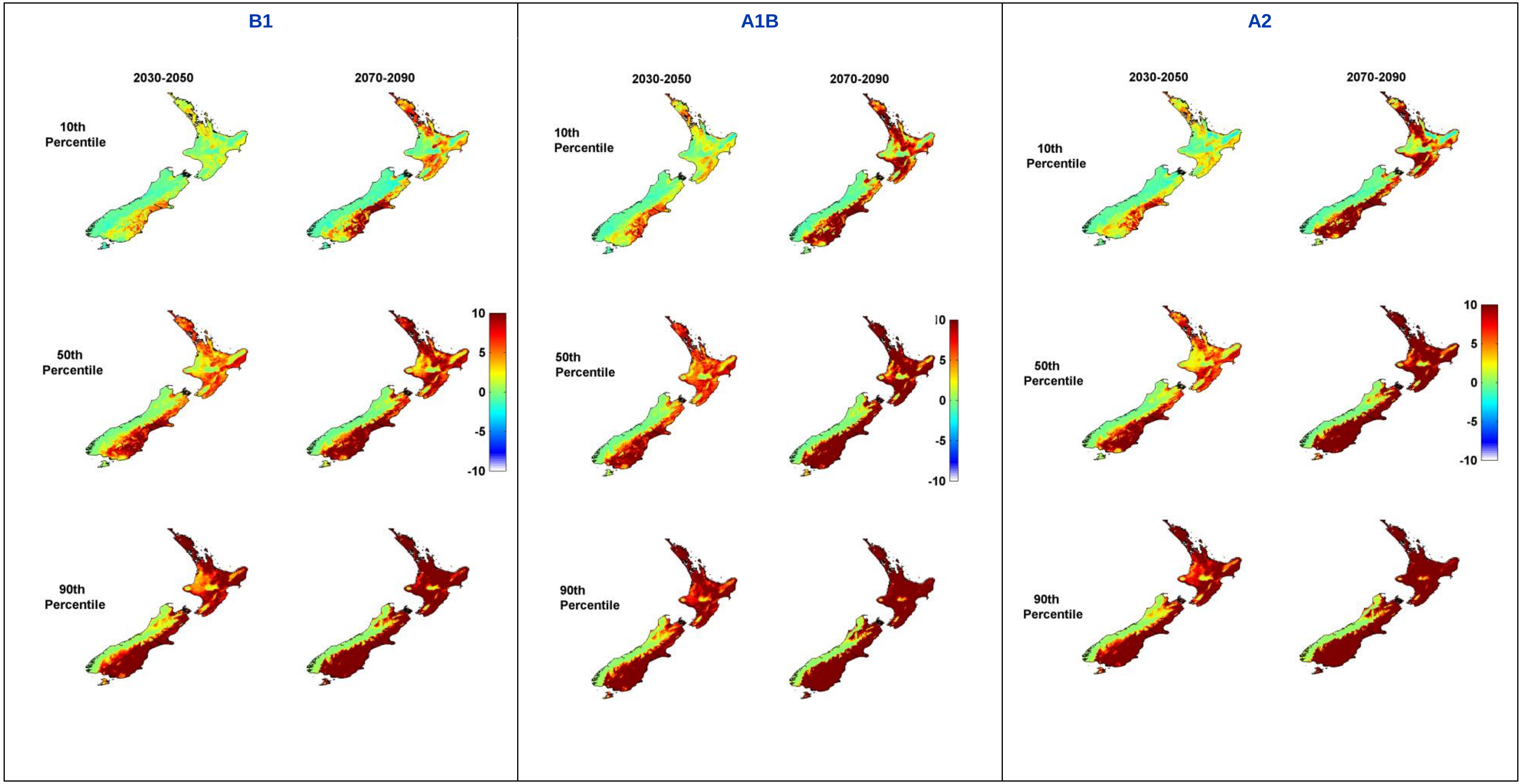


Figure 7: Additional time spent in drought from 1980-99 for the temperature based PET downscaling method. Percentiles summarise the range of outcomes across 19 climate models.

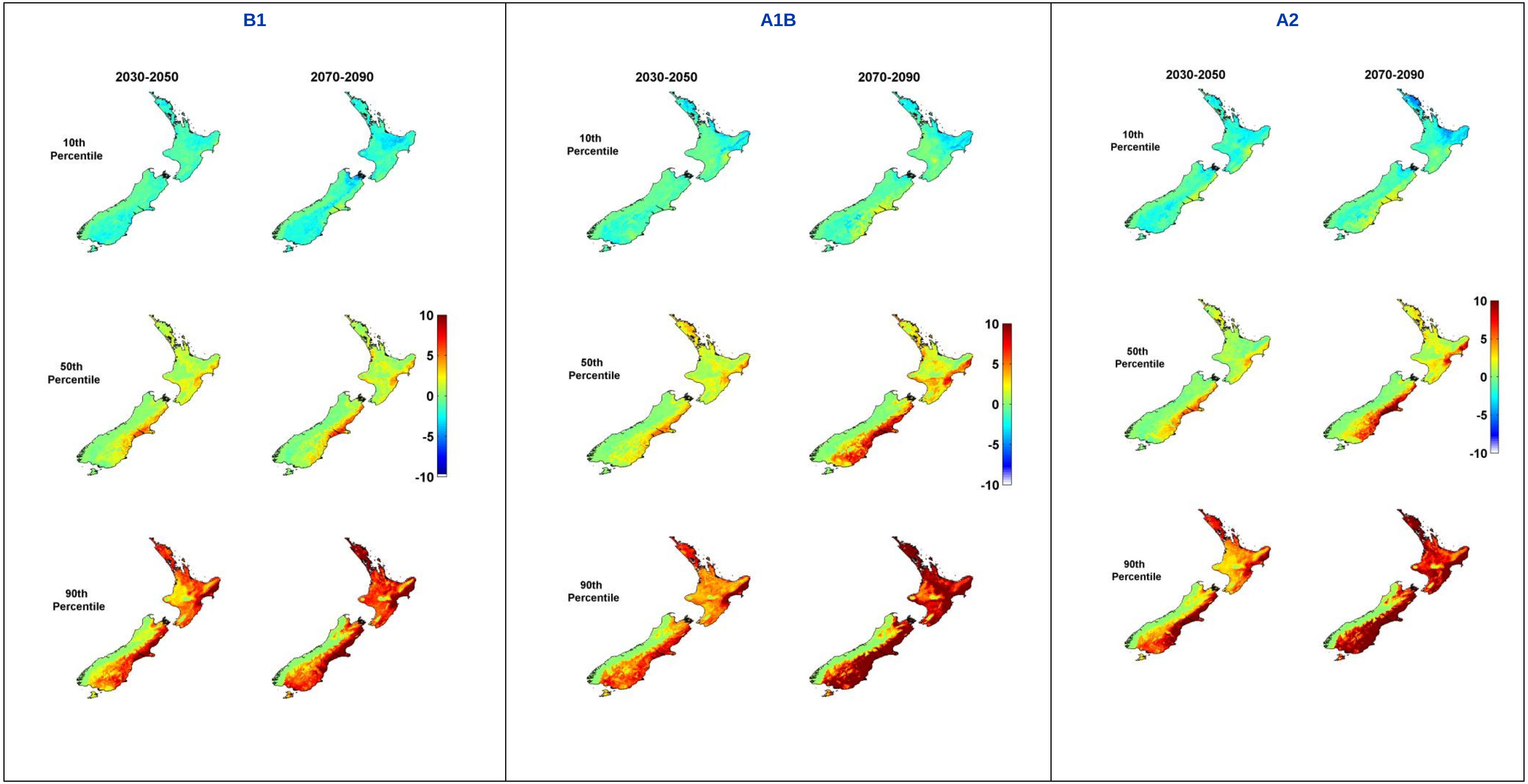


Figure 8: Additional time percentage of spent in drought from 1980-99 for the radiation based PET downscaling method. Percentiles summarise the range of outcomes across 19 climate models.

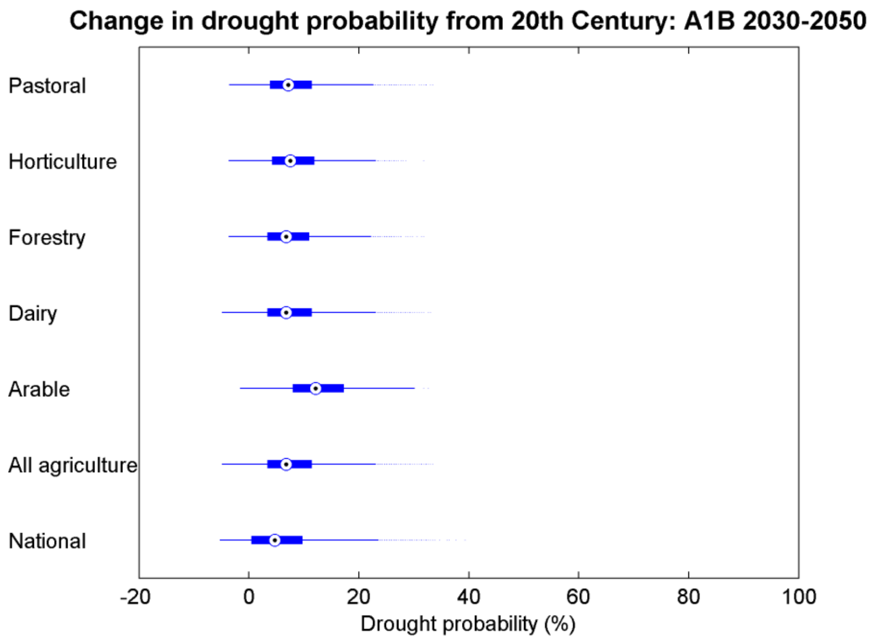
The radiation downscaling experiment introduced more diversity across the GCMs, which may be reflecting the diverse ability of GCMs to simulate the surface energy budget. At the low end in Figure 8 droughts remained stable or slightly decreased, even in the 2070-90 time frame in some regions. The general geographic structure of change in drought described previously was not evident in the 2070-90 timeframe for the B1 emissions scenario, but was evident as a mild increase (2-5 %) in the A1B and A2 scenarios. The general spatial distribution of change to New Zealand drought exposure was clearer at the middle and high end of the GCM ensemble. Similarly the diversity of change to drought exposure for the North Island was also reflected, ranging from stable and decreasing through to strong increases depending upon GCM.

The geographic analysis of the scenario experiment also shows that the regions where the strongest increases in time spent in drought coincide with the regions shown to have least downscaling and GCM bias, also the main agricultural regions of New Zealand. This reaffirms the observation that in general the shifts projected by climate change are greater than the biases introduced by downscaling. This is a positive result for the methodology and builds confidence in the projections as well as inferences that can be drawn from them.

3.3.4 Industry distribution of change

To examine the projected changes in a more specific way they are stratified by current industry land use in accordance with the 2007 agricultural census. The land use was determined for each 5 km grid cell by calculating the dominant industry given the underlying mesh block classification. Where a clear dominant land use could not be identified the result for each grid cell is double counted.

The results for the A1B emissions scenario stratified by this industry classification are presented in Figure 9 and Figure 10 as box plots. The boxplots illustrate the range of outcomes across the GCM ensemble for all locations relevant to each sector. These clearly show the differences in magnitude between the two downscaling methods, with the radiation based approach (Figure 8) having a lower and more constrained degree of positive change and a greater proportion of area experiencing a reduction in drought. For the radiation based method the median changes were in the order of 2-5 percent increase across the industry sectors while for the temperature based method the median change was in the order of 5-10% more time spent in drought.



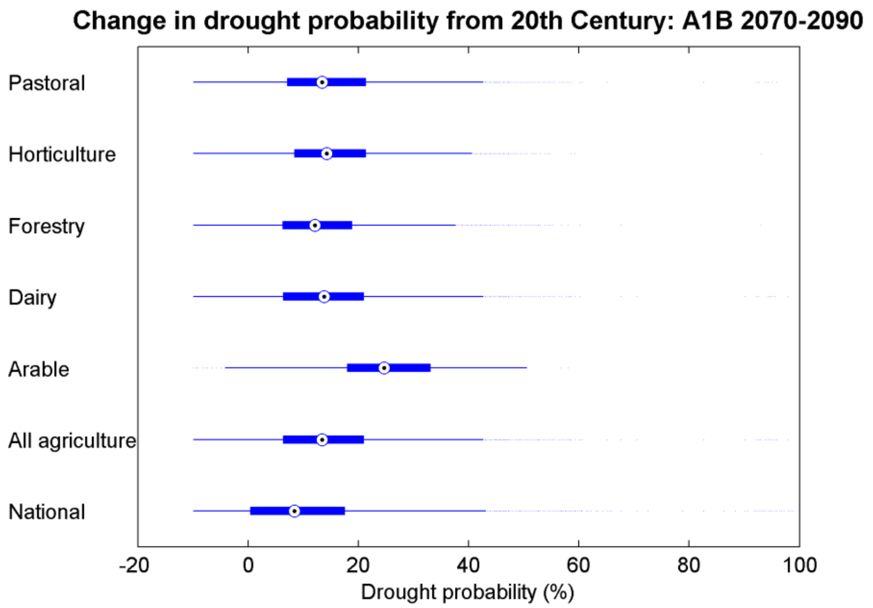


Figure 9: Range of projected changes in drought probability for each sector based on current land use patterns for the temperature PET downscaling method. Changes represent additional time spent in drought relative to the 1980-99 base period. The mid-point (circle) is the median of the GCM ensemble, with solid bars showing the inter-quartile range. Thin lines extend to the 10th and 90th percentiles, with outliers shown as dots.

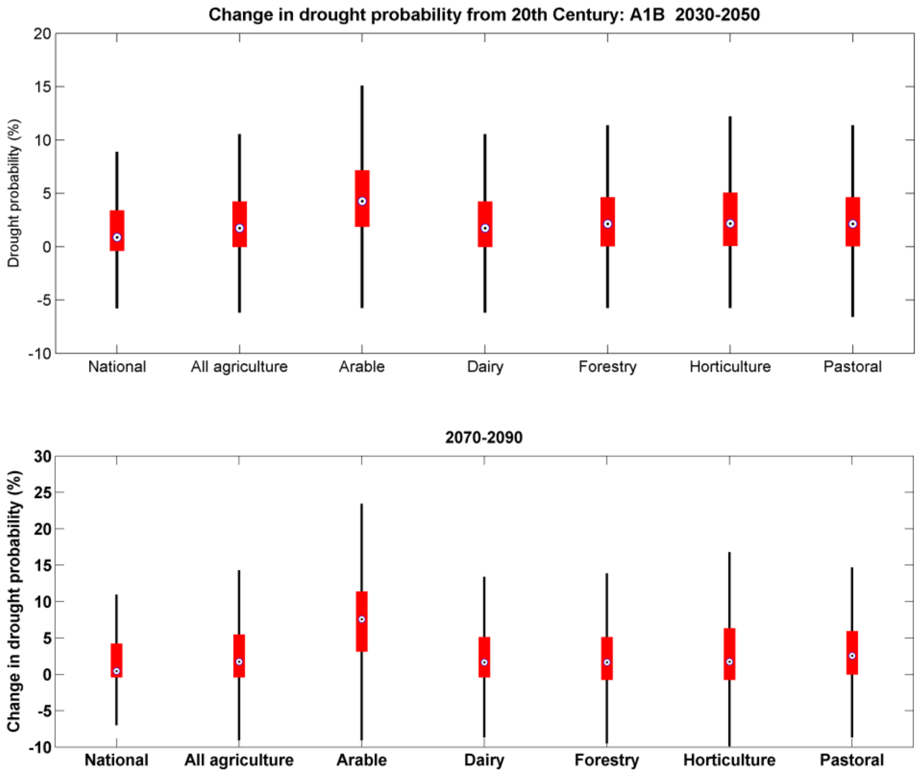


Figure 10: As Figure 9, but using the radiation PET downscaling method.

Examining all agricultural land suggests that the whole sector is slightly more exposed to increasing drought under climate change than the country as a whole—but examining the overlap in the interquartile range shows that this is a small difference and not likely to be significant if a statistical test was applied. With the exception of the arable farming sector there were no major differences between the individual sectors and all agricultural land. Closer examination suggests that there are some marginal differences between sectors, for example in the radiation experiment the horticulture sector appears to experience a wider range of outcomes than the other sectors, with the 90th percentile being at a plus 15 percent level and the 10th percentile below the negative ten percent.

3.3.5 Transient analysis of change

As an alternative to the time slice analysis, a transient analysis was undertaken to examine the behaviour of drought probabilities over the 100 years 2000-2100. It was also an opportunity to focus on two aspects of drought that have not been addressed in the previous approach; the geographic scale of the event and the likelihood of receiving two ‘back to back’ droughts within a twelve month time-frame. These are of interest as they are likely to be droughts of national economic significance. To analyse drought in this way a national indicator was constructed with the following modifications:

- for each 5 km grid the drought occurrence index was calculated according to the approach and thresholds described previously;
- given this base input a national drought occurrence index was calculated by firstly excluding all land not used in primary production;

- a widespread drought was deemed to have occurred when greater than 40 percent of the agricultural land was in drought. Drought was revoked when this figure fell below 40 percent. The drought event also needed to be continuous spatially, that is greater than 90 percent of the grid cells deemed to be in drought needed to be adjacent;
- a national scale drought was deemed to have occurred when two events of this nature were recorded within a twelve month window. This was applied as a moving window recursively.
- at this point the bivariate distribution of duration and intensity could be integrated and as before the percentage of time in drought calculated to represent the probability.

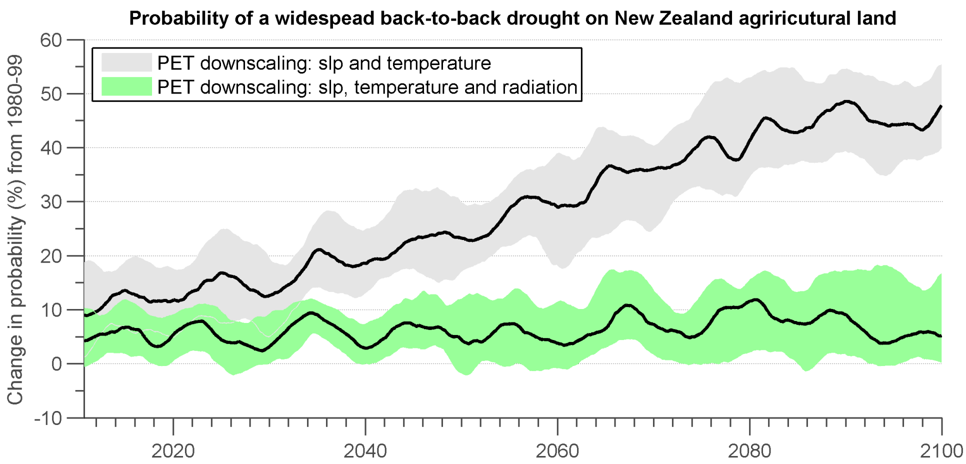


Figure 11: Transient analysis of widespread back to back droughts for the two downscaling methods, for the A1B emissions scenario over 2010-2100. Shaded areas summarise the interquartile range across 19 GCMs and the black lines are the median.

The results of this analysis for the A1B emissions scenario are shown in Figure 11. The graphs show the most likely change in the GCM ensemble as a black line (the model ensemble median) and the range or 'risk envelope' around this as shading. Both methods illustrate that the most likely projection is for increased time spent in back-to-back drought across the century out to 2100. It is a further illustration of the differences in magnitude between the two downscaling approaches, as by around 2040 the differences between the two approaches become greater than the intermodel differences between GCMs.

The temperature method suggests a dramatic rise in the order of an addition 40-50 percent in the prospect of widespread back to back droughts by the end of the century. Examining the median change the rate of increase appears to peak at around 2080 and plateau out to the end of the century—reflecting the emissions profile in the A1B marker scenario described in section 3.2.2. This represents a strong and significant shift to a much more arid climatic regime, where the type of droughts that are currently challenging to manage for the national economy become a regular feature of New Zealand's climatic environment.

While the radiation downscaling method also suggests an increase in drought probability over time, the magnitude is far lower. The median of the GCMs peaks at around an additional ten percent in 2070 then plateaus and declines by the end of the century to around a 5 percent increase. Interestingly this effect beyond 2070 is consistent with the A1B marker scenario were global population peaks in the middle of the century with a declining rate of emissions increase from 2060. While the range of models suggest an increase in drought exposure for all GCMs using the temperature downscaling, the lower range of the models in the radiation method shows minimal change or even periods of decrease out to 2100.

There is also presence of inter-decadal climate variability in the transient analysis. This is a reminder that the GCMs are sensitive to this shorter range variability, but it should be stressed that there is likely to be little predictive skill for identifying the timing of short range variability—it is not appropriate to use this to pinpoint which future decades might or might not have increased exposure to drought. Interestingly the magnitude of longer term change projected by the temperature only method is far greater than the decadal variability, but for the radiation method it may not be. These types of subtleties have a bearing on how adaptation planning might be undertaken. As mentioned in the introduction to this study, decadal scale variability has not been analysed here, but its presence, predictability and relationships with both shorter term variability (ENSO and SAM) and longer term change are worthy of attention by researchers.

3.4 Discussion

3.4.1 Relationship with previous studies

The broad level result of this study is confirmation of work by Mullan et al. (2005), who found that under future scenarios constrained by GCMs vulnerability to drought in New Zealand increases in regions that are currently more exposed, specifically the eastern seaboard of the North and South Islands. This east-west divide in projected change has now been found in two studies with different climate models, downscaling methods and approaches used to quantifying drought. In this study, the most consistent result was for minimal change on the West coast of the South Island and for increased future exposure on the Canterbury Plains. Although not as consistent, generally increases in drought vulnerability were also found in the Hawkes Bay-Gisborne and Northland regions. Like the 2005 drought study, and for rainfall projections in general, there was also a wide range in the magnitude of projections for the majority of the North Island, where little or no change through to large increases were found, sometimes well over a doubling in exposure to drought.

Where this study differs from the previous work is in the quantification of the magnitude of change and identifying its associated uncertainty. At one level this study used approaches where downscaling and model error could be tracked more specifically than the previous work. This strengthened the overall confidence in the projections particularly for use on agricultural land as the climate change signal detected was greater than the error introduced by downscaling and GCMs. This is an important result for the overall approach, and should also add weight to the previous study given the consistencies found. As described in section 3.4.3, tracking error in this way points to some areas of future work where uncertainties might be reduced.

The key difference between the studies is in the range of magnitudes found in future drought scenarios. The 2005 work reported increases to drought probability of at least twice as often in some areas, through to four times as often for a one in twenty year drought. While this study provides further evidence that this range of change is plausible over the next 100 years, results from the lower end of the GCM ensemble provide scenarios of milder change. These equate to less than a doubling of a 1 in 10 year drought in the more exposed regions described previously. More significantly, this study also explicitly identified and pursued methodological uncertainty in the downscaling of potential evapotranspiration. This resulted in two distinct future projections of drought in New Zealand:

1. a ‘temperature driven scenario’ which pushes the country well towards a climate that is more arid with large increases in drought exposure, well over a doubling in the more exposed regions;
2. a ‘radiation driven scenario’ where there is milder climate change, ranging from no movement or even slight decreases in drought exposure through to a future where the change is tightly constrained to at worst around or just below a doubling in the most exposed regions.

Although there are small geographic differences, generally both of these projections have the same spatial pattern of exposure described previously. Differences in results between methods for estimating potential evaporation are consistent with results from the recent scientific literature. Two key examples are: in a global level climate change study, Kingston et al. (2009) found that although there were consistently increasing signals of PET associated with a warming climate, there was almost a 100% difference in the magnitude of signals depending on the PET derivation method; in a hydrological study of the United Kingdom, Kay and Davies (2008) compared one temperature based PET method with a physical method (Penman-Monteith) and found large differences between the two approaches when they are resolved with climate change scenarios. Similar uncertainties surrounding the quantification and representation of surface water flux are also found independent of climate change analysis, for example in model development (Oudin et al. 2005a,b; Droogers 2000; Green et al. 2008), trend analysis of observation (Rayner 2007) and derivation of a fully transportable approach to quantifying potential evapotranspiration (Lu et al. 2005; Chavez Eguez 2008).

The discussion of both these results by Kingston et al. (2009) and Kay and Davies (2008) identifies a methodological issue also found in this study, and one that is also common in all estimation and modelling of PET—the choice between simpler temperature based PET methods that have fewer inputs but are more empirical and less faithful to known processes, versus methods where there is more process information but are more complex and more data intensive. As found here and in the other studies, when resolved with projections from GCMs the two methods can behave quite differently and un-expectedly, both in terms of their 20th century diagnostics as well as the nature of climate change scenarios that they produce. In this study the methodological differences created a wide range of outcomes, outside of the range of the GCMs when evaluated nationally, and potentially changing the sign of the response for some GCMs when evaluated regionally.

3.4.2 Probability, uncertainty and risk management

At this point it is important to re-consider the purpose of scenario analysis in this and other contexts; they are developed to create plausible futures for the process of scenario planning. The foremost job of a scenario analysis is to uncover the most likely projection, the future given most weight in a planning exercise. In this case plan for around 10% more time spent in drought in the coming century in key eastern agricultural regions of New Zealand. The second task is to examine the range of outcomes and establish less likely but plausible future scenarios. For New Zealand contingencies can be developed for a low range outcome a 'no change in drought' projection, and a high range outcome a strong shift towards a more arid climate. These can be considered but given less weight than the most likely scenario in planning.

The information from scenario analysis is only valuable if it is used in planning exercises—an active dialogue by managers about what tangible steps can be considered under the different scenarios. The value is in developing new insights, clearer strategies and importantly contingencies to manage the full range of plausible outcomes. This is sometimes known as 'what-if planning' or 'futures management'. This is best achieved within a planning framework, and because projections have a probabilistic nature, this is an exercise that should be firmly grounded in risk management. In this context, and using the Australian New Zealand Standard, this research and resulting drought scenarios represent the very first stages of this process. They are best used to raise risk awareness and guide constructive action-oriented dialogue by managers.

Moving forward from a scenario plan to tangibly improving actual risk preparedness for climate change, in other words building resilience or reducing vulnerability to drought, is done by developing and implementing risk management activities. This involves identification and costing of potential affirmative actions; and implementing and monitoring those actions. This is a very different activity to scenario planning and requires different information than the scenarios developed here, such as cost-benefit analysis of different options.

This study has clearly shown the contribution that downscaling GCMs can make to improving risk awareness under climate change in New Zealand. The consistent geographic pattern of change to drought exposure has been found now in two studies. Without this effort to downscale GCMs, New Zealand would be entering the future largely unaware that this is a physically plausible outcome. We would also argue that quantifying the range of results from multiple climate models and identifying less likely but plausible scenarios is also valuable. By carefully tracking the range of results we were able to identify two physically plausible sets of futures, one with mild and the other with severe increases in drought exposure. For scenario planning it means options need to be identified for both outcomes. Given the effort to downscale GCMs both monitoring and further research to reduce uncertainty can now be more targeted.

Updating the scenarios and GCMs from the 2005 study is an illustration that uncertainties in climate change are not irreducible, and further improvement to GCM based climate change scenarios can be anticipated over time. There is also a level of certainty in the results that can be used to identify and guide decisions that can be made now. Knowledge of the range of outcomes also helps to target decision making in the future, for instance clarity around which actions might remain purely as contingency plans for now until a greater level of precision is attained in the future. All of these responses are consistent with a risk management approach.

The use of GCMs for developing scenarios does not in any way preclude or diminish the value of sensitivity analysis independent of climate projection, or a more 'bottom up' approach that might take a whole farm and examine its vulnerability. In the context of risk management, information from this approach, like cost-benefit analysis, is likely to be more useful as the process matures and help guide the tangible actions needed for becoming more risk prepared. Similarly this should not be seen as an activity that should preclude the redevelopment of ‘top-down’ GCM based scenarios in the future. Both approaches have a contribution to make toward climate change risk management, which in itself is an iterative process that should be continually updated with new information as it improves.

It is also tempting and sometimes a requirement of analysts to conclude by identifying what tangible and specific management actions should now be taken. This has been avoided in this study in favour of the general discussion about scenario planning and risk management presented above. Options for improving agricultural drought risk management are certainly available at the farm level, for example diversifying income or improving access to irrigation water. They are also available at the broader institutional level such as improving early warning systems or re-examining the types and timeliness of support provided during a drought. These and other options may be discussed by managers in a scenario planning exercise, and prove to be a tangible means of adapting to the future drought scenarios posed in this study. Going further along this theme is outside the scope of this work, and more valuably undertaken by managers who are seeking to build resilience to climate change in their own specific contexts.

3.4.3 Further research

It is within the scope of the study to comment on potential avenues for further research within the framework of GCM based scenario analysis, and climate science more broadly. Some of these are consistent with the risk management action of seeking to reduce uncertainty through monitoring and increasing the precision of a risk analysis. These can be stratified into smaller targeted areas of work that follow immediately from this study and some broader longer term substantive lines of research that could be pursued. The more targeted options for further work are:

- there is clearly opportunity to extend the partial least squares regression downscaling framework developed in this study to operate at a daily timescale and for more variables;
- further to this there is also opportunity to continue working with other methods for empirically downscaling daily rainfall, for example application of non-linear regression or analogue models;
- there is also opportunity to utilise the results from NIWA’s physical regional climate model (RCM). This hinges on the timing of different model runs and model improvements, as well as work underway in bias correction of these simulations;
- in the short term there is also need to undertake further experiments and inter comparison of different PET downscaling methods given both empirical and dynamic modelling;
- collectively these initiatives would prepare for the IPCC’s 5th Assessment Report, where the next CMIP5 experiment will run updated Global Climate Models with new emissions scenarios based on representative concentration pathways;

- it is also possible to downscale alternative GCM experiments such as the doubling carbon dioxide simulations in order to provide a basis for impact sensitivity analysis or relating system responses to a per degree global warming basis.
- consistent with efforts internationally there is scope to improve and update New Zealand's drought early warning information systems. This might include expansion of drought indicators, inclusion of new forecasting methodologies and improvements to underlying models.

Some potential broader lines of research include:

- the results provide further illustration that there is a case to improve the monitoring of climate in some of New Zealand's regions, notably the alpine zones of hydrological catchments. Such action is a feasible way to reduce uncertainty under present and future climate but will take time to accrue its full value. This has been known for some time and was clearly identified in the stakeholder workshop preceding this study (Clark and Tait 2008). NIWA has a fledgling snow and ice network and has also been trialling innovations to improve the modelling of rainfall and temperature fields in these zones;
- this parallels the need to work on fundamental physical knowledge concerning the behaviour, modelling and downscaling of the precipitation mechanism in the Southern Alps. Given the uncertainty identified in this study, improving the geographic precision of projected climate in this region is clearly important to future management of the water resource;
- there is need to work toward an overall improvement to monitoring and resulting downscaling of potential evapotranspiration, so as to inform decision makers which of the two climate futures identified in this study is the more plausible. This also relates to uncertainties surrounding the field response of plants under elevated carbon dioxide. We believe that substantive improvement in this area requires a more precise and long term monitoring campaign for this variable. NIWA has already undertaken a small step toward this goal with a permanent eddy covariance system soon to be installed at one site on the Canterbury Plains, integrated with greenhouse gas monitoring activities.
- although not thoroughly analysed there is need for further work on decadal scale climate variability as it relates to New Zealand and interactions with both longer term climate change and also shorter run variability. While it is convenient to summarise climate across 20-30 year windows, managers will experience climate change as the integral of these sources of variability. Although targeted research in this area is at a very early stage, there are opportunities emerging such as global climate model experiments set up to track decadal signals. In the future there may be scope to build this information into a drought early warning system for New Zealand.

- there is clearly need for general improvements to the modelling of convective precipitation to improve the precision of studies like this one. This has been identified in numerous occasions and forums, and is a community wide challenge for climate research. There are a diverse range of innovations that are being pursued, too many to describe sensibly here (see Trenberth et al. 2003 for a discussion). Additionally the difficulty of this task should also not be understated.

4. Review of drought indicators

4.1 Summary

This chapter provides a review of drought definitions, indicators and also examines the selection of a water balance for producing estimates of drought. The chapter focuses on methodology, with a key consideration the parsimony of drought indicators and simulation models with the temporal and spatial scales of climate projections. The work was undertaken to examine some of the factors surrounding drought analysis raised in stakeholder consultation (Clark and Tait 2008).

4.1.1 Drought indicators

The review illustrates that drought does not have a universal definition, and compared to other natural hazards it is ‘insidious’ or lacks specificity, and hence there are uncertainties in defining it for both research and management. Drought definitions and the construction of drought indicators are context specific. International evaluations have shown that no matter which indicator is selected it is likely to be open for critique by some stakeholders. This is a common and unavoidable aspect of drought research and management, more formerly termed ‘operational indeterminacy’;

Following an assessment of drought indicators, development of a simple dry land agronomic index of soil water is pursued in this study. This is a compromise between a purely climatic indicator which may not account for rainfall effectiveness and hydrological, production and economic indicators. The later set of indicators have the advantage of being more relevant and direct estimates of drought impact for specific industries. But it would be difficult and resource intensive to build the type of universal indicator sought after for this study based on hydrological, agricultural and socio economic indicators of drought. This is because there is neither sufficient data nor a modelling methodology to generate future projections of drought under climate change on a comprehensive national to regional basis using these indicators. A dry land soil water based index provides an agriculturally relevant drought indicator that is generalisable both across regions and industries and supports projection analysis.

The overall structure of the Phillips and McGregor Drought Index (PMDI) has been used in this study as it provides a framework where the duration and intensity thresholds used to define drought are transparent. This replaces the use of potential evapotranspiration deficit (PED) used in the previous study (Mullan et al. 2005). The PMDI has some advantage over the other indicators considered, as in most cases drought defining thresholds were not transparent but inherent in complex technical procedures. The PMDI has been set up using thresholds built from historical analysis, rather than physically based thresholds, providing a description of relative drought across New Zealand. For this study the PMDI was modified to be sensitive to severe events in the New Zealand context, the types of drought durations and intensities which have potentially strong impacts on management and production. The following specific thresholds are used: drought occurs when soil water is below the 10th percentile of 20th Century variability (1980-2000) for one month, and revoked only when it is above this threshold for more than one month.

4.1.2 Water balance models

A second modification was made to the published PMDI framework, where instead of accumulated rainfall anomalies soil water was used as the physical input. As soil water is to be derived from climate variables through simulation, a verification study of a number of alternative simulation models was also undertaken. This builds an evidence base to guide model selection and addresses some of the key questions raised about methodologies in stakeholder consultation.

It was found that when considering the 'severe level droughts' approximated by the thresholds detailed above, differences between simple and more complex models were small. Detailed verification using a national network of observations also pointed to a similar conclusion—although a more complex framework could be devised to attain more precision, such gains are unlikely to effect the overall calculation of long term probabilities of severe drought.

These results provide evidence that some assumptions taken in the 2005 study and those underpinning NIWA's current operational system are valid. Specifically a simplified water balance is an appropriate framework for analysis of drought events of around one month or more in duration at national to regional scales. The final model applied is a one layer water balance with an exponential decay function to calculate loss by evapotranspiration. The decay function differs slightly from NIWA's operational model which uses a step-function with a set relationship—this small refinement was implemented so that the simulations are sensitive to soil texture across New Zealand, as described by current Fundamental Soil Layer datasets.

Sensitivity tests were also carried out using a more detailed biophysical water balance to assess the influence of stomata response under rising carbon dioxide in the context of calculating long term drought. The tests illustrate that stomata response is not an important factor when considering the severe drought events examined by this study. This is consistent with recent field evidence which illustrates that the closure of stomata and associated plant growth response under elevated carbon dioxide is not pronounced at times of water limitation. Carbon dioxide fertilisation and plant responses are not examined further or used to produce projections of drought in this study of drought, but as discussed in more detail it is an important factor when considering total production variability.

4.2 Defining drought

Drought is generally defined as a severe moisture deficit below expected levels that restricts some type of activity (Wilhite et al. 2006). It is important not to confuse drought with related concepts like aridity where low rainfall is a permanent feature of the climatic environment, seasonality where water shortage is a normal part of the annual climatic regime, or desertification where a region of aridity shifts because of poor management (White and Walcott 2009).

Moving beyond this general definition of drought is problematic as it is a complex insidious event when contrasted to other natural phenomena (Wilhite et al. 2006; Hisdal and Tallasken 2005): it does not have clear entry, duration and termination points compared to floods, fires and storms. There are random properties of drought that do not lend themselves to standard analysis (Gordon 1992). Drought is 'context specific' as both its biophysical properties and physical impacts vary considerably between socio economic groups, regions and industries—it is possible that two adjacent farms experiencing the same climatic drought conditions will be experiencing different economic impacts.

As a result defining and quantifying drought is a research field in its own right, and there are many alternative definitions. Despite this diversity, two general frameworks are identified in this report as a way of understanding the numerous drought definitions, which lead to the methods of quantifying drought. The first and more widespread framework defines drought as a natural hazard. This is based on the notion that precipitation deficit leads to a water shortage and subsequent negative impact on an activity, group or environmental process (White and Walcott 2009, Wilhite 2000). The second framework examines drought as a risk that can be managed (White and Walcott 2009). This is based on the assumption that the impacts of drought can be reduced or avoided by taking actions, or that systems have an inherent resilience that can be adjusted.

4.2.1 Drought as a natural hazard

Research viewing drought as a natural hazard supports responses that manage the event during its occurrence, thereby alleviating and or repairing damage and loss. Examples include the formation of regional drought committees; the provision of income and other forms of support to farmers during a drought; and in some countries the alleviation of food shortages in the broader community. Best practice in this approach relies on: establishing planned and funded responses well before the drought event occurs; the use of well targeted drought indicators as an early warning system that guide the timely implementation of these responses; and ongoing monitoring to promote an early recovery.

The widely cited work of Wilhite (2000) on defining drought draws heavily on the view that droughts are a natural hazard. This stratifies drought definitions and indicators into a series of general categories:

Meteorological: based on the degree of rainfall deficit relative to what is considered normal for a given climatic environment;

- Hydrological: based on the physical water supply shortfall such as runoff, river flows and dam levels.
- Agronomic: based on levels of biophysical production, principally plant growth and crop yield, but also animal condition and numbers for livestock industries
- Economic: based on estimates of farm gate income, debt, equity and other business performance estimates. At a macro level it is referring to a reduction to the supply of goods and services.
- Social: based on measures like farmer stress, the number of forced sales of properties attributable to drought and or perceptions of climate variability.

4.2.2 Drought as a manageable risk

Defining drought as a manageable risk supports the identification of vulnerability and improvement of drought resilience. In agriculture, this is achieved by adopting a more holistic systems view of farming leading to transformational changes such as: shifts to the intensity of management; new or alternative production systems more appropriate to the climatic variability of a given region; and business restructuring to adapt to the dual risk of price and climate variability. Figure 12 is an example of how drought can be viewed as a manageable risk (White and Walcott, 2009). Here meteorological droughts are considered causes or drivers of hydrological processes like the soil water balance. The functioning of these processes gives rise to drought biophysical impacts like lack of irrigation water, loss of pasture growth and animal condition. In turn these biophysical impacts give rise to set of socio-economic drought effects for example loss of production and reduction of farm income.

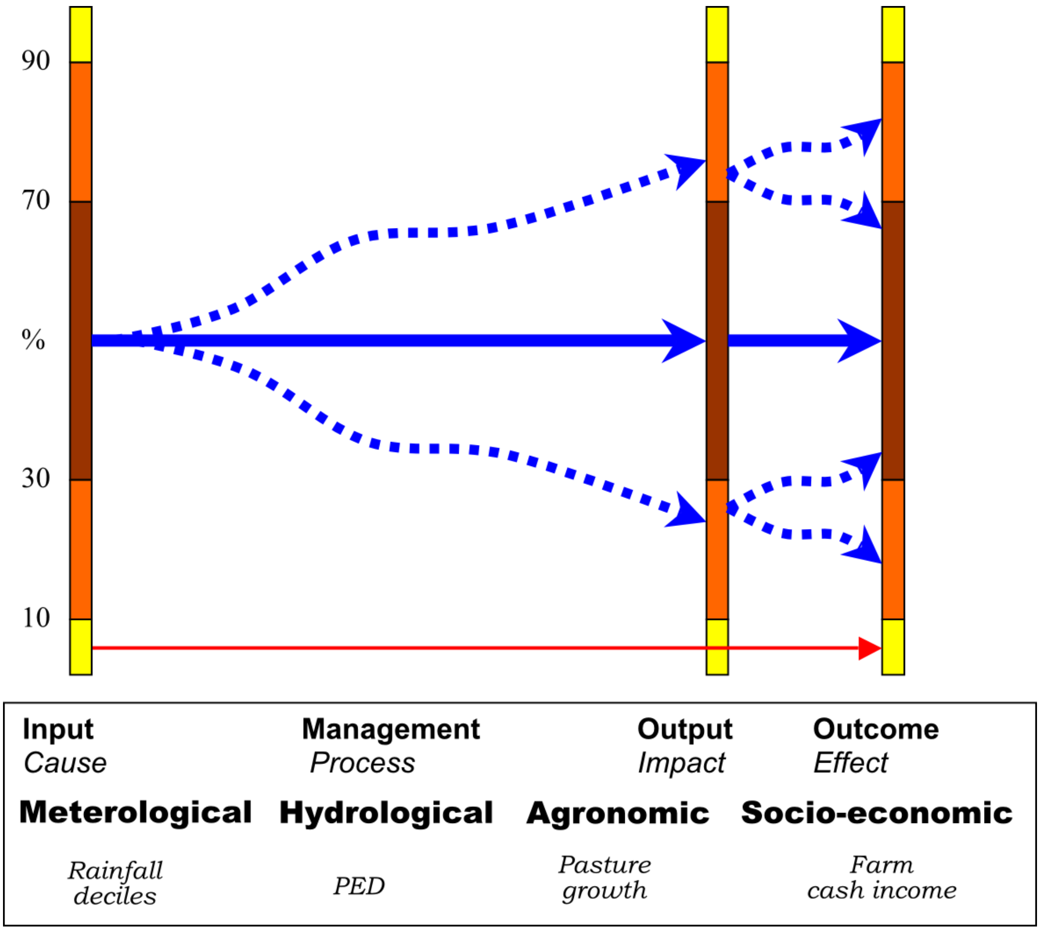


Figure 12: Drought as a manageable risk: a framework. Used and modified with permission from White and Walcott (2009).

In an agricultural landscape the meteorological causes of drought are increasingly modified by management as processes, as impacts and production outputs and outcomes are considered. This is illustrated by the blue arrows in Figure 12 which illustrates that an average climatic season could end up producing an above or below average farm income depending upon management practices. If a system is resilient to drought, climatic inputs lower than 30% of historical range might yield a farm cash income of 50% or greater than historical returns. The red arrow is indicative of a drought that cannot be managed; suggesting that a meteorological threshold exists below which management actions cannot mitigate negative impacts.

The presence and quantification of such climatic thresholds and their relationship with management is a key uncertainty when defining and assessing drought (White and Walcott 2009). Similarly, drought risk management is difficult to define because it is context specific: it is a management process that is highly situation dependant and evolves over time given the almost unique climatic and financial pressures experienced during each drought event (Clark 2001). It is possible to build illustrative examples (typologies) of the type of management options that are available, but it is difficult to be specific and prescriptive without real context. Some example typologies include;

- institutional governance like national to regional adaptive planning, research management, crisis response management (Nelson et al. 2008; Steinemann and Cavalcanti 2006); and
- specific farm level responses, such as: manipulating production system intensity; early sale of stock to avoid impacts on the animal and feed costs; opportunistic purchase of fodder in good times; use of forward selling and diversifying income so that a farm business is less sensitive to rainfall deficits.

4.2.3 Objective drought indicators

Drought indicators provide an objective basis for quantification and analysis of drought, and can be a source of information in both of the main frameworks used to define drought. Importantly the selection of specific indicators and degree of integration are different between the two frameworks. The hazard definition usually relies on one or two universal or generic drought indicators specific to a given response, for example meteorological or soil moisture metrics. In the manageable risk framework, multiple indicators are used highlighting climate inputs, biophysical function and outputs. In this case it is important that the indicators are tightly integrated, and that there is a high degree of consistency in their derivation and analysis.

Beyond the general framework used to define drought, other more specific properties of drought need to be considered to construct objective indicators:

- every individual drought event will have a unique duration or persistence, and intensity. But the duration and intensity of drought also have a degree of positive dependence;
- drought goes through distinct phases including an entry or onset phase as well as a recovery, exit or termination phase. Depending upon prevailing climatic conditions, these phases may be rapid and obvious, while in some events they develop slowly in a way that is not always observable without sophisticated analysis;
- drought recovery can be particularly difficult to define (Day et al. 2005) as it is in part dependent on the severity of preceding conditions—a few rainfalls may constitute recovery from a low intensity drought but will not constitute recovery from a severe drought;
- the impacts and effects will lag behind the meteorological causes of a drought, for instance when a region is well into a climatic recovery the financial effects may be at their most intense and persist for some time;

- drought may also be defined in terms of its geographic scale, for example individual properties, regions or at a national level;
- droughts can also be dependant upon the temporal scale of examination, for example do two distinct dry periods within the same 18 month period constitute one long drought or two separate events? At what point does a succession of individual drought events constitute a shift to a more arid climatic environment?

The literature documents a large and diverse range of drought indicators that have been developed over the past twenty years. Some examples of overviews of drought indicators include: Rao and Voeller (1997) White et al. (1998); Steinemann and Cavalcanti (2006) and Wilhite (2006); White and Walcott (2009). Some examples of applied agricultural drought analysis include: Stafford Smith and McKeon (1998); Stephens (1998); Arena et al. (2006); and Yoo et al. (2006). Examples of hydrological analysis include: Clausen and Pearson (1995); Ramsay et al. (2008); Thompson (2006); and Arena et al. (2006). A review by Steinemann (2005, cited in Steinemann and Cavalcanti 2006) identified 150 different drought indicators, noting that this may not be a comprehensive list.

A key insight common in this body of work is that there is no universally accepted drought indicator, and they need to be developed for a specific context, are suited to different purposes and are motivated for a variety of reasons. The appropriateness of a given indicator for its stated purpose is the key factor to consider when examining a drought analysis. Some of the contexts in which drought indicators are developed include:

- to provide a description of drought for use in climate research or event based risk analysis (e.g., Mullan et al. 2005; Blenkinsop and Fowler 2007; Hennessy et al. 2008);
- to provide information for planners and practitioners to trigger responses as part of regional drought plans or national response (Steinemann and Cavalcanti 2006; White and Bordas 1997);
- as part of an early warning system to build anticipation and preparedness into drought response;
- to trigger on farm management responses;
- in formal drought declaration and monitoring systems which are sometimes linked to Government interventions designed to alleviate effects.

The high degree of arbitrary choice in selecting an appropriate indicator is a source of controversy in many drought analyses and management activities. It is likely that no matter what indicator is chosen and developed for a given context some stakeholders will not be entirely happy with the choice. It is often difficult to anticipate the needs of every individual stakeholder, or the context in which they interpret or wish to use an indicator. This has been described more formerly by Steinemann and Cavalcanti (2006) as 'operational indeterminacy' and is an unavoidable aspect of drought research and management.

A degree of pragmatism is also required in selecting indicators as data quality and availability, the timeliness of information and the nature of supporting models influence choice. For this reason, and because they are the leading expressions of drought in the landscape, meteorological indicators are the most widely used. Many countries have well established near real time climate monitoring networks that provide a rich source of current and long term historical data (White and Walcott 2009). An increasing number also have the modelling infrastructure needed to produce derived agronomic indicators like simulated soil water and crop yields.

Given data availability, going beyond meteorological drought indicators to agronomic indices based on field observations or socio-economic indicators based on survey or simulations can be problematic. They are less widely used because they are not monitored as intensively or routinely and may not have lengthy historical data sets by which to establish benchmarks. Thus well founded measures of impact are usually more difficult to obtain and analyse, particularly in near real time. Economic studies of impacts (e.g. MAF 2009) are usually undertaken well after an event has passed and usually it is difficult to undertake a longitudinal analysis over many decades.

4.2.4 Candidate indicators

The purpose of this study is to undertake a generalised national to regional scale analysis of drought where scenarios under projected future climate can be developed. Hence the study naturally follows a hazard and largely meteorological definition of drought. It is also important to recognise that the selection of indicators in this context is undertaken with a considerable degree of stakeholder guidance. The 2008 consultation (Clark and Tait 2008) ratified some general factors identified in the previous industry consultation reported by Porteous (2005). 'Desirable properties' developed through these consultation processes for a New Zealand drought indicator should be :

Universal – the index should be applicable to all parts of the country, and be suitable for both nation-wide and regional analyses. The index should also be sufficiently versatile to cope with varying thresholds or scales of severity;

- easily interpreted;
- supported by readily available data to enable calculation of robust anomaly and recurrence statistics;
- should enable improved advice to farmers for land use planning;
- an indication of production loss;
- suitable for subsequent research needs – eg. land use, social implications, water policy;
- based on parameters for which projection of future change can be plausibly developed, given current knowledge;
- able to be linked to decision trigger points (eg. the kinds of thresholds that farmers might use to implement drought mitigation actions), and;
- represent both the duration and intensity of droughts, as both are important.

Within this overall context Table 2 introduces a small subset of drought indicators described in the literature that were considered as candidates for this study. Table 2 focuses on indicators that can be derived from climate data, thereby supporting the desired properties of universality, data availability for national-regional scale analysis and ability to integrate them in a projection study. They are organised by general drought definition to represent the major alternatives that were considered for this study. This suite of indicators is firstly assessed by general appraisal to highlight some properties of the candidate set. A more detailed examination of assumptions, including some supplementary analysis, is undertaken to assess the assumptions of the main candidates in the New Zealand context.

4.2.5 General appraisal of indicators

A common criticism of all the rainfall or meteorological based indicators listed in Table 2 is that they provide only a partial measure of agricultural drought as they don't account for rainfall effectiveness for plant growth and yield. For instance rainfall may occur outside the optimal temperature window for growth or at a rate where runoff is high. Simple rainfall aggregates may depart from yields or complex calculation procedures developed to try to account for these factors, such as the cumulative rainfall deficiencies (CRD). Rainfall based indicators do however provide a practical, general leading index because this is the dominant driver of drought for many industries and natural systems. Sometimes because it is a leading index and there are good climate monitoring networks, rainfall based indicators are chosen for early warning—the WMO has recently adopted the standardised precipitation index (SPI) as its universal drought indicator; and rainfall deciles (RD) have been used under Australia's National Drought Policy as a leading indicator to 'trigger' a more detailed appraisal of agronomic and socio economic impacts.

Indicators based on soil moisture are usually considered more relevant as a general drought indicator for agriculture because they account for rainfall effectiveness. Examples from Table 2 include the Palmer Index, Potential Evapotranspiration Deficit and the Soil Moisture Deficit Index. Ideally soil moisture based indicators should be derived with monitored field data, but this is not always practical or possible. For instance New Zealand is one of the few countries which has a national soil moisture monitoring network but the data set is not yet comprehensive enough to build national level drought indicators for a full climate period (20-30 years). Often drought indicators based on soil moisture are derived from a simulated water balance driven by climate observations.

Although soil moisture provides a more agronomic drought indicator it only goes part of the way to accounting for rainfall effectiveness. Temperature restrictions on crop and pasture growth during winter for example can influence the timing and duration of droughts. More specificity for a given farm or industry can be obtained by basing drought indicators on direct measures of production such as crop yield and pasture growth or even socio economic indicators like farm cash income. However with specificity comes the penalty of losing universality, a desirable property of a drought indicator for this study.

Table 2: Candidate indicators considered for use in this study.

Meteorological		
Rainfall deciles Gibbs and Maher (1967)	$TP_{(n)} = P_{(n)} + \sum_{i=t}^{n-i} P_{-i}$	Ranks accumulated rainfall (TP_n) over an integration period (*t*, e. g 3, 6 or 12 months) relative to historical data.
Hutchinson Drought Severity Index (HDSI) Smith et al. (1993)	$HDSI_{(n)} = \sum_{i=t} RP_{(n-)i}$	The sum of rainfall percentiles (*RP*) over the integration period *t*.
Cumulative Rainfall Deficiencies (CRD)	$CRD_{(n)} = \sum \left(P_{mean} - P_{(n)}\right)$ where $P_{(n)}<P_{mean}$	Sum of the mean (P_{mean}) minus the observed rainfall ($P_{(n)}$) when the observed is less than the mean.
Drought Severity Index (DSI) Phillips and McGregor (1998)	$DSI_n = \frac{\left(\sum_i X_y - X_m\right)}{\sigma} \times 100$ Entry when $X_y<X_m$ Terminate when $X_y<X_m$ and $X_y>\Sigma X_m$	Standardised accumulation (divided by standard deviation then multiplied by 100) of precipitation anomaly for a period *i*. Drought entry when negative. Drought termination when positive the *ith* month running mean (ΣX_m).
Standardised Precipitation Index (SPI) Cancelliere et al. (2007)	$SPI = \frac{X_y - X_m}{\sigma}$	The observed rainfall (X_y) minus the mean rainfall (X_m) divided by the standard deviation (σ)
New Zealand Definition (NZD) Mosley and Pearson (1997)	$NZD_{(n)} = \sum_{P(n-1)>15}^{\Pr<0.1} P_{(n)}$	Drought is a 15 day period where there is no measurable rainfall (0. 1mm).
Hydrological		
Surface Water Supply Index (SWSI)	$SWSI_{(n)} = \frac{aP_{snow} + bP_{prec} + cP_{strm} + dP_{resv} - 50}{12}$	*a, b,c,d*, are weighting factors for snow (P_{snow}), precipitation (P_{prec}), stream (P_{strm}) and reservoir (P_{resv}) anomalies respectively.
Agronomic		
Plant Growth Index (PGI) McDonald (1994)	$PGI_{(n)} = \sum Pgr_{(n)}$ where *Pgr>t*	Accumulation of plant growth rate (*Pgr*) when it exceeds a critical threshold (*t*)
Palmer Drought Severity Index (PDSI) Palmer (1968)	$PDSI_{(n)} = 0.897 PDSI_{(n-1)} + \frac{1}{3} Z_{(n)}$	Palmer Moisture Anomaly Index, a function of monthly anomaly indexes for evaporation, loss, runoff, recharge and runoff determined from a water balance
Potential Evaporation Deficit (PED) Porteous et al. (2005)	$PED_{(n)} = \sum_t PED_{(n-1)} + \left(Ep_{(n)} - Ea_{(n)}\right)$	Accumulation over period *t* (twelve months July-June) of the difference between potential (*Ep*) and actual (*Ea*) evaporation determined from a model.
Soil Moisture Deficit Index (SMDI) Narasimhan and Srinivasan (2005)	$SMDI_{(n)} = 0.55\left(SMDI_{(n-1)} + \frac{SD_{(n)}}{50}\right)$	The accumulation of soil moisture deficit scores (SMDI, value -4 to +4) SMDI is determined by the percentile soil water deficit (SD0

A key criticism of many drought indicators is that complex calculation schemes are devised in an attempt to account for rainfall effectiveness or remove volatility. At this point drought indicators can loose their transparency and are not readily interpretable. For example, the Palmer Drought Severity Index and the Soil Moisture Deficit Index have lengthy computation procedures that attribute different weightings to parts of the water balance (evaporation, runoff, drainage, transpiration). Once soil water input is established, the drought index is produced by re-calibration to a scale from -4 to 4. While this produces a standardised metric, further work needs to be done by users to track this back to physical processes (soil water deficit) to trigger management actions.

Drought indicators do not always provide a valid metric or have enough sensitivity in environments where the rainfall distribution is highly skewed, such as rangelands, arid zones or deserts. This is particularly true when assumptions about ‘normality’ are built into the indicators calculation. For example, the Phillips and McGregor index in Table 2 is based on anomalies from the monthly mean rainfall. In some arid environments the monthly mean is close to zero, so this index would be highly unstable in these environments. Similarly some indices are not sensitive enough in extremely wet environments where there are very short duration rainfall deficits.

Drought indicators are also built on assumptions about the spatial and temporal scale of events. For example the choice of climate integration period in a rainfall decile analysis builds in assumptions about what drought durations are considered severe. In this case the assumptions are not wholly transparent and the end user needs to carefully consider the integration period, making interpretation of rainfall deciles difficult.

4.2.6 Detailed appraisal of indicators

Occurrence thresholds

An important aspect of all drought indicators are the assumptions used to determine occurrence thresholds, or the entry and exit criterion. The critical distinction is whether they are based on a measure of a ‘historical’ or a ‘physical’ threshold. Physical thresholds like plant wilting point or a critical dam level are appealing because they can be linked directly to a physical impact or socio economic effect, and readily support design and engineering analysis. They also have a scientific basis because the thresholds can be measured or established from experiments and or scientific literature.

Examples from Table 2 include the Potential Evaporation Deficit (PED) and the plant growth index (PGI). A limitation is that it may be a normal event for a physical threshold to be exceeded; hence care needs to be taken not to confuse a physical threshold based indicator with an index of aridity or seasonality.

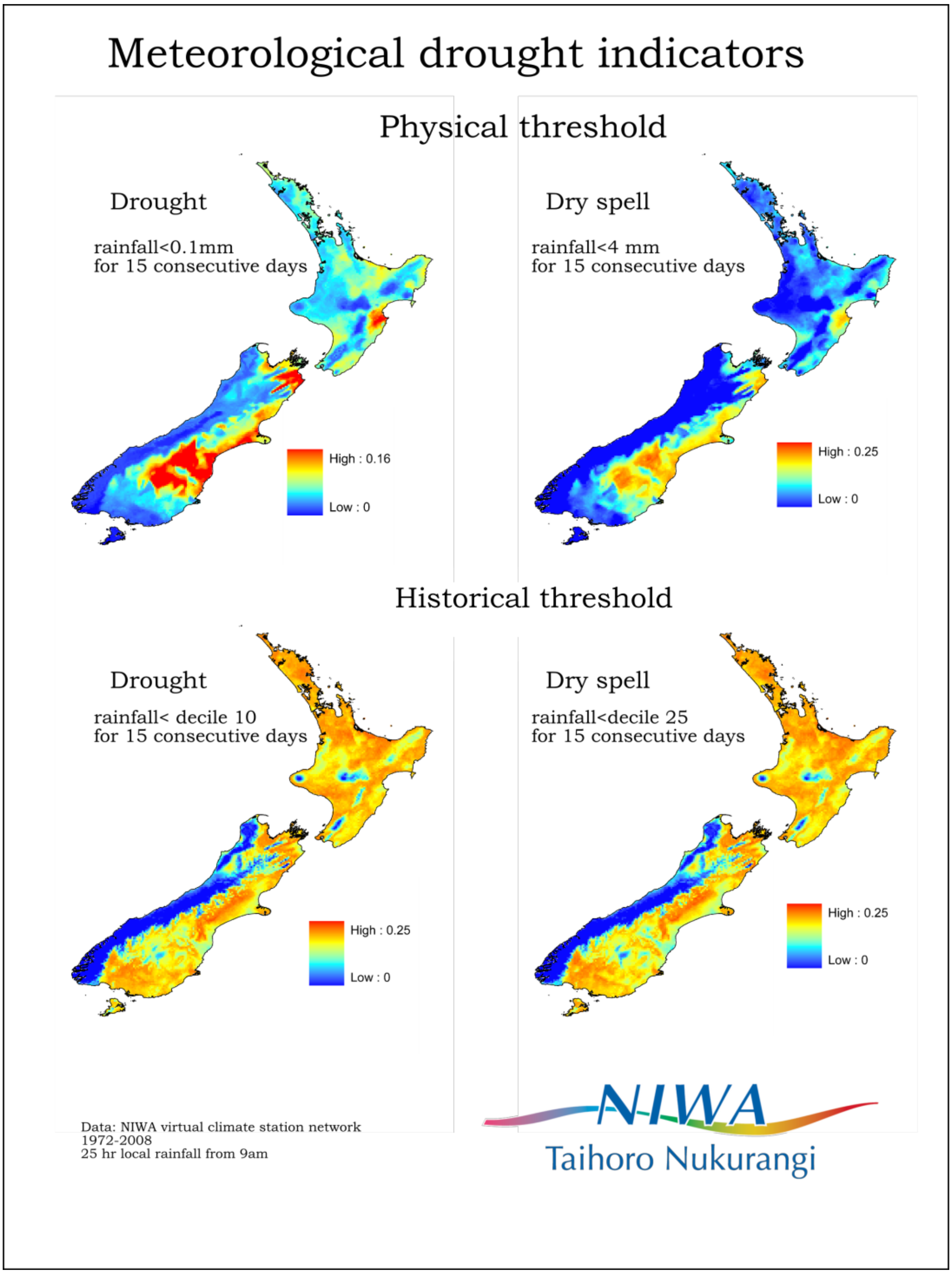


Figure 13: Examples of meteorological drought probabilities calculated using either physical or historical approaches to determine occurrence thresholds.

Basing thresholds on ‘historical’ analysis overcomes this deficiency be establishing entry and exit thresholds relative to past variability. This is the basis of percentile or anomaly based indicators of rainfall and agronomic variables, such as the Standardised Precipitation Index (SPI), the Hutchinson Drought Severity Index or the Phillips and McGregor Drought Index. As thresholds are based on the historical distribution, the approach is reliant on the availability of long term historical data sets of 30 years or more. This approach can also be criticised because the choice of historical threshold is largely arbitrary—does a drought constitute the 25th 10th or 5th percentile of this historical distribution. The accuracy of determination is also dependant on length of historical data set as well as the statistical estimator used to determine the distribution.

Figure 13 illustrates some differences between the two approaches for meteorological-based indicators in New Zealand using the period 1972-2008. The analysis assumes that drought occurs if rainfall deficits accrue for 15 days below a certain threshold. Using the physical assumption (0.1 mm and 4 mm, Mosley and Pearson (1997)), areas of higher drought probability are confined to the east coast of both islands, particularly the South Island. According to this, droughts would be a relatively rare occurrence in the Waikato and Northland region. Using historical assumptions (10th and 25th percentile for the week of the year), the area of higher drought probability expands to covering a larger proportion of the North Island. Using this approach the Waikato and Northland regions appear to have similar event frequency to the Canterbury Plains.

Under longer term climate change, the relevance of drought indicators based on 'historical normality' can be debated. On one hand the statistics of hazardous drought derived by analysing the 20th Century provide a useful analogue and benchmark for users. It is common to use concepts like 'within living memory' or the 'experiential timeframe under which a management system evolves', and use the last 30 years to establish historical drought benchmarks. However, in a changing climate very severe droughts in the past may be more common or be within a new normal range of variability (Mullan et al. 2005, Hennessy et al. 2008). This would make a drought threshold based on the previous 30 or even 100 years irrelevant, assuming that management is continually adapting to the new climatic regime and reducing drought vulnerability. This raises a number of practical issues about how to establish historical drought thresholds for this study:

- from a management perspective what has happened in the past is an increasingly poor guide for the climate variability of the future when climate is changing, and particularly if there is rapid change.
- should drought benchmarks and thresholds be established for the 20th Century and used to analyse projected droughts in the future?
- alternatively has the projected climate of the future reached a new normal state and should drought thresholds be established based on a new climatology?

Drought duration and intensity

All the drought indicators listed in Table 2 have assumptions about the intensity and duration of drought events (Shiau and Shen 2001). In some indicators prior assumptions are made about the length of an event, for example the 15 day period in the New Zealand meteorological definition or the choice of integration period in the Hutchinson Drought Severity Index. Alternatively assumptions can be made that drought occurs when a specific indicator falls below a given intensity threshold, for example below a certain rainfall percentile (Gibbs and Maher 1967). More sophisticated indices build in different thresholds to distinguish between entry and exit, for example the Drought Severity Index (DSI) proposed by Phillips and McGregor (1998) has an entry criterion of two consecutive months of negative rainfall anomaly and a termination criterion of three consecutive months positive anomaly.

Technically 'characteristic drought' is more correctly considered as a joint function between intensity and duration (Gonzalez and Valdes (2003); Kim et al. 2003; Thompson 1996; Ramsay et al. 2008), where each environment has its own particular set of drought characteristics. Although the West Coast of the South Island has very high rainfall and could be considered 'drought free', rainfall deficits events do occur and are detectable as short duration events. They are very short lived with the most intense events lasting no more than 10-15 days. In contrast on the Canterbury Plains, more prolonged (around 6 month periods) but low intensity rainfall deficiencies are detectable, with the most common drought being a high intensity event of 1-2 months in duration.

A simple matrix is proposed that characterises durations and intensities as the 'characteristic droughts' typically experienced across New Zealand's environments (Table 3). Assumptions about the historical percentiles that may be used as threshold intensities as well as time frames for durations are also presented. For example, a short duration high intensity (SH) drought is one that has duration less than a month and intensity less than the 10th percentile of historical variability. The framework in Table 3 is used throughout this project to more clearly define various drought analyses.

Table 3: A matrix of characteristic droughts experienced in New Zealand with associated assumptions about duration (drought length in days or months) and intensity thresholds (percentiles based on historical variability). The shaded area focuses on the characteristic droughts examined in this study. .

		Intensity		
Duration		**High** (<=10th)	**Moderate** (<=25th)	**Low** (<=50th)
	Short 1month (>0 days, <31days)	Short duration high intensity **(SH)**	Short duration moderate intensity **(SMo)**	Short duration low intensity **(SLw)**
	Medium 1-3 months (>31 days, <=93 days)	Medium duration high intensity **(MH)**	Medium duration moderate intensity **(MMo)**	Medium duration low intensity **(MLw)**
	Long 6 Months (>93 days)	Long duration high intensity **(LH)**	Long duration moderate intensity **(LMo)**	Long duration low intensity **(LLw)**

Spatial and temporal scale

Beyond duration and intensity drought has a definable spatial scale. This dimension is often overlooked or not discussed thoroughly in the literature describing drought indicators. For a country like New Zealand where agriculture is a significant proportion of the economy, droughts of economic importance are those more geographically widespread across the agricultural landscape. There are also regions of more significance, such as hydroelectricity and irrigation catchments.

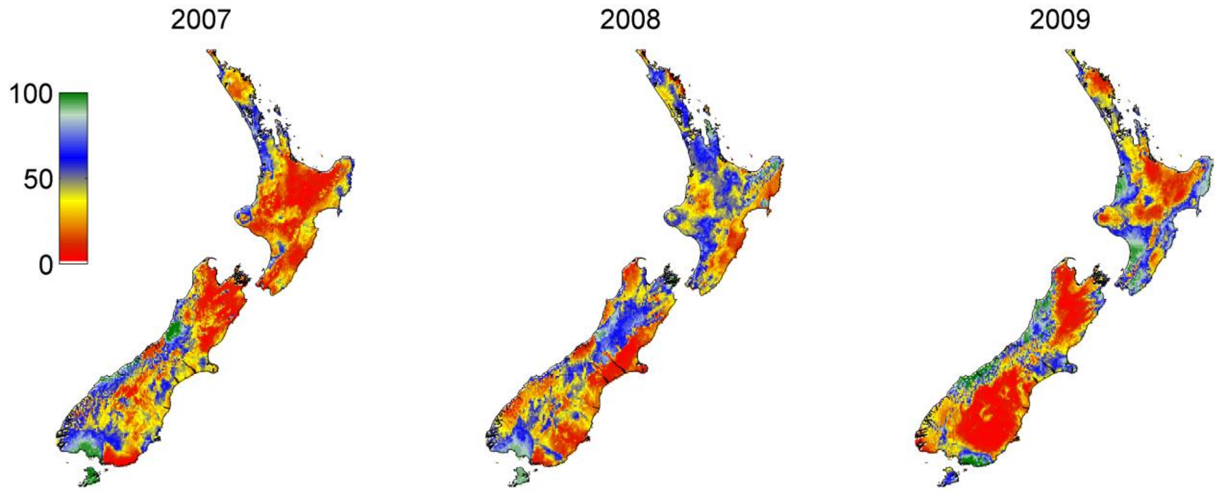


Figure 14: The regional structure of spring (SON) droughts experienced during 2007-2009. The drought indicator is pasture growth percentiles based on simulations for the period 1972-2009. Simulations undertaken for this study using an experimental pasture growth model .

Droughts have a complex geographic distribution as each drought event has a unique extent across the country. This is illustrated clearly in the example maps of pasture growth percentiles (Figure 14). The three calendar years are an example of widespread drought in New Zealand, but clearly some individual regions experienced favourable climatic conditions for production in the Springs of 2007, 2008 and 2009. Some regions experience three consecutive failures during spring, such as the Northern Waiarapa.

The temporal scale of drought events can also be assessed by developing spatially aggregated indices. The indices in Figure 15 were produced by using only agricultural land from the virtual climate station network (Tait et al. 2006), and averaging the indices calculated at each grid. These national indicators are sensitive to the duration and intensity of drought, as well as its spatial scale. For instance the period 2007-2009 is expressed as a long duration moderate intensity drought in the national indicators—as some regions experienced favourable conditions in that period, the intensity was moderate when analysed at the national level. However a deep intense and widespread drought was evident in the period up to the end of May 2010, a widespread and intense event.

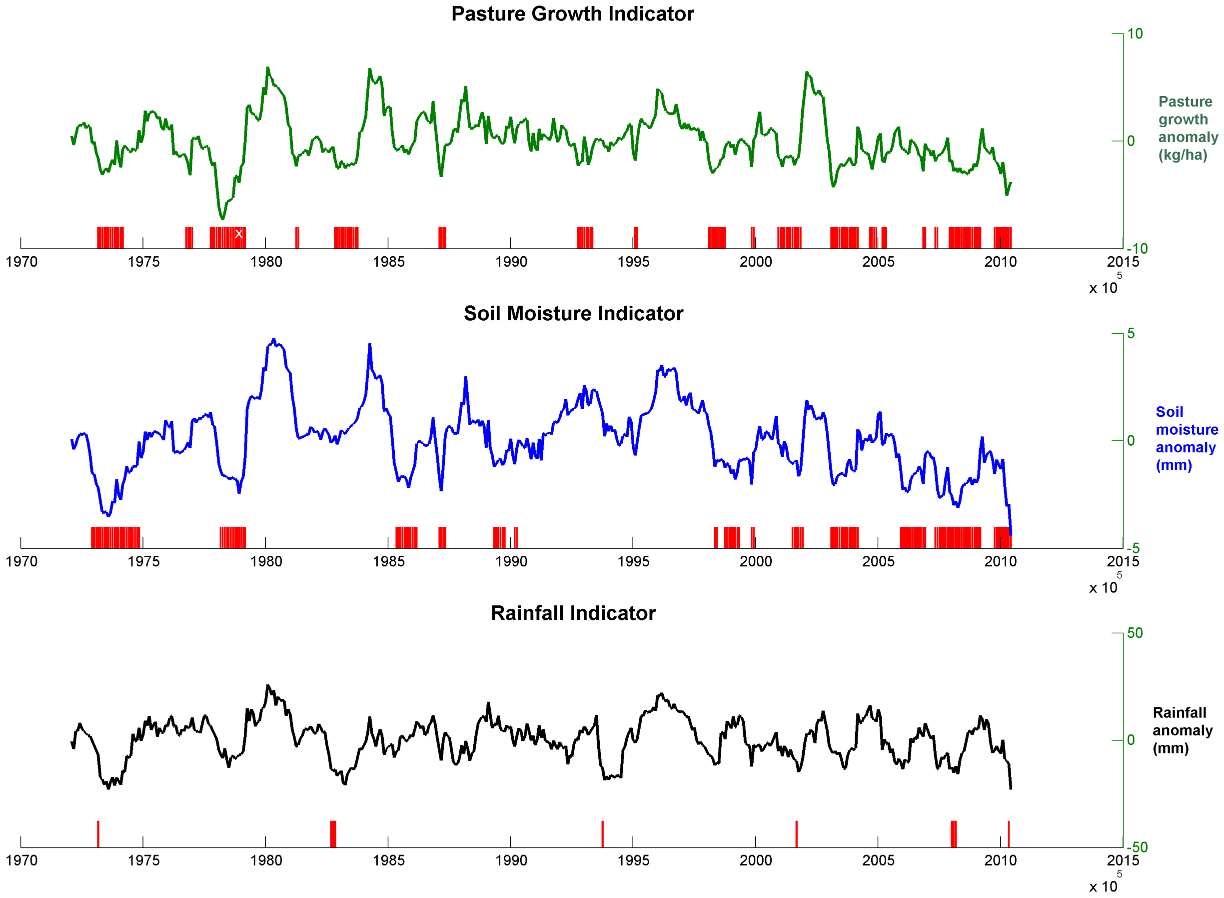


Figure 15: National scale indicators of drought for New Zealand from 1972-May 2010. The red bars under each graph display the timing of drought entry, its duration and revocation.

4.2.7 Previous New Zealand studies

In previous work undertaken by NIWA (Porteous 2005 and Mullan et al. 2005), a methodology for assessing New Zealand agricultural drought and undertaking climate change impact assessment was developed. While acknowledging the benefits and limitations of drought indices a choice needed to be made on indicator selection. The studies of Porteous (2005) and Mullan et al. (2005) selected Potential Evaporation Deficit (PED) as a key indicator for drought based arguing that it has most of the desirable properties listed above. The 1972-2009 average of PED accumulated over a year for is shown in **Figure 16**.

Stated advantages and limitations of the index are listed in Table 4, and methodological testing was undertaken from a site near Lincoln on the Canterbury Plains. An important rationale was that PED, like all indicators that are based on a water balance calculation, integrates a broader suite of climate variables than rainfall alone, and to some extent accounts for rainfall effectiveness for agricultural production.

Table 4: Stated advantages and limitation of Potential Evaporation Deficit (PED) as a drought indicator, after Porteous (2005) and Mullan et al. (2005).

Advantages	Limitations
▪ Based on testing at Lincoln, PED provided a useful comparison with a number of recently experienced droughts ▪ Data readily available for underlying model platform ▪ Correlated with days of soil moisture deficit which explains variability in stock numbers and seasonal milk production and the normalised vegetation index. ▪ Ability to be modelled spatially ▪ The underlying water balance found to reasonably closely represent soil moisture conditions in the pasture root zone in previous studies	▪ Uncertainty about the relationship between PED and surface flow, and hence does not reflect water availability for irrigation. ▪ Difficult to quantify relationships between PED and socio economic effects and other environmental impacts of drought. ▪ Uncertainty if relationships between PED and productivity will hold in the future.

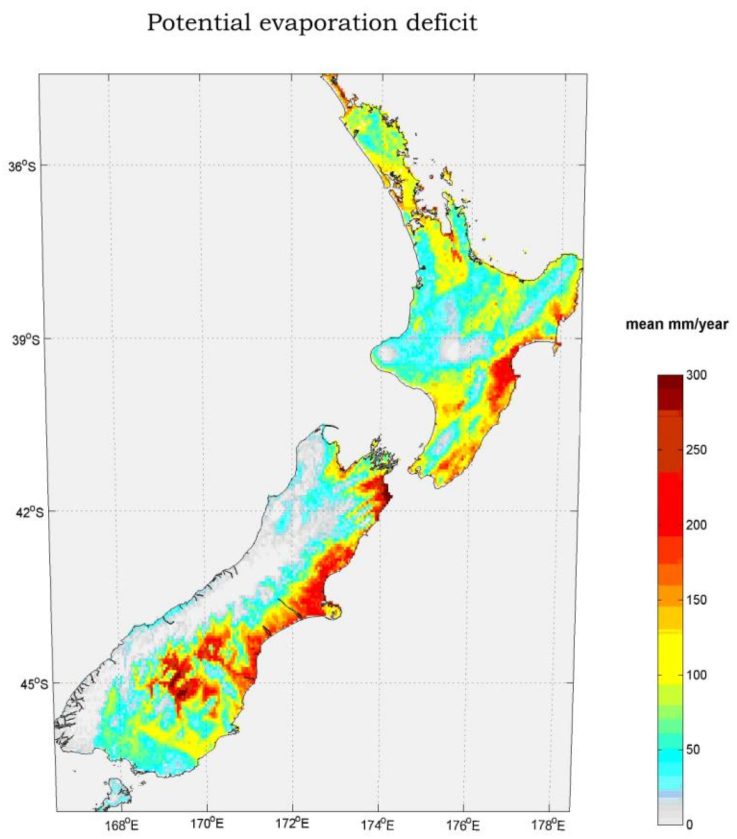


Figure 16: Average annual Potential Evaporation Deficit (PED), 1972-2008.

4.2.8 Indicator selection

Drought indicators are to be used in this research for determining changes to national-regional scale drought statistics given global scale projections of climate. There is need for a small number of indicators for this purpose, more consistent with the natural hazard framework. While there is considerable merit in use of the manageable risk framework, use of a suite of integrated indicators and projecting their response under climate change scenarios is beyond the scope of this study. Examining drought resilience and adaptation requires either a whole farm modelling or broader integrated assessment methodology and is also beyond the scope of this study.

To account for rainfall effectiveness the agronomic definition of drought is used rather than a purely meteorological one. This is consistent with the previous 2005 study as well as the feedback from the consultation workshop (Clark and Tait 2008)—so the rationale is not discussed in detail here. This means utilising either a soil water or plant production metric derived from a simulation model with climatic inputs.

The choice was made to continue with a soil moisture rather than a pasture based drought indicator, because of the availability of monitoring data to support model selection. While it is technically feasible to model pasture growth at the regional and national level, it was decided not to in this study because of a paucity of readily available verification data at the regional to national scale. In addition, a pasture growth indicator would focus the study heavily on the beef, sheep and dairy industries, at the expense of arable crops and horticulture which would require a more specific yield indicator.

The next decision is selection of the specific soil moisture based metric. It was decided to not use the PED index from the 2005 study in favour of a standard analysis of soil water. The rationale was that: this is more consistent with the relative soil water indicator provided by NIWA as part of its ongoing advice to national and regional drought committees and as public information for drought early warning; the major concern about the PED is that it may be an index of aridity—it measures the degree to which the climatic environment falls below its capacity to maintain moisture or in other words irrigation deficit—so it may be overly sensitive to droughts in some regions and insensitive to drought in others; it is common international practice to use actual soil water (mm) as part of national drought strategic alert systems.

The next key choice in developing a drought indicator in this research is whether or not to pursue one which is based on a physical or an historical threshold. An historical based threshold established by analysis of a 20th Century base period (1980-2000) has been chosen because this builds in the assumption that agricultural production systems are geared to the climate variability of a given region. Establishing the intensity threshold in this way will normalise the drought indicator to local historical variability, and provide a more balanced indicator across New Zealand's regions. The most striking case is the Waikato, which clearly does experience droughts that impact on dairy production, but as shown by the examples in Figure 13 a threshold based indicator may not be sensitive enough to detect these adequately. The judgement was made the historical threshold would provide a simpler and more generalisable index.

Related to this is the choice of analytical framework for the analysis of the soil water metric. For reasons of simplicity and transparency, the broad framework defined by Phillips and McGregor (1998) was chosen because it allows specific control of both drought entry and termination—as described in the previous sections many of the drought indicators have in built assumptions that are not always obvious and importantly controllable in an analysis. The original work by Phillips and McGregor used rainfall anomalies with the sign of anomaly over specified durations to signify drought. The approach was adapted here so as to utilise soil moisture and historical thresholds over assumed durations.

The next choice is to decide which characteristic drought types (Table 3) are more or less relevant and feasible to analyse in the context of this study. It was decided to analyse only the probabilities of receiving medium duration and moderate intensity droughts or greater in this study (shaded region in Table 3), not the full suite of drought types. The key thresholds are therefore droughts of greater than one month in duration and intensities below the 10th percentile of historical variability.

Analytically this was implemented by modifying the Phillips and McGregor drought index to produce event time series of drought given entry and termination thresholds. Drought entry occurred when the soil water metric fell below the 10th percentile for a period greater than one month and drought was terminated when it was greater than the 10th percentile for one month. Given these drought event time series long term probabilities were calculated by integrating a non-parametric bi-modal distribution of the duration and intensity of these events. This procedure is described by Kim et al. (2003), Gonzalez and Valdes (2003) and used by Ramsay et al. (2008). One important modification was made to this methodology to account for computational limitations found when using non-parametric statistics across the national scale data set—the more computationally efficient kernel density estimators described by Martinez and Martinez (2002) were adapted and used in this study. For simplicity the probabilities were converted to units so as to express the time spent in drought in a given region.

These drought threshold assumptions were made for two reasons: firstly following a pragmatic argument that while droughts of one to two weeks in duration do occur they are do not always have a large agronomic impact at regional scale, whereas droughts greater than a month become a concern to the sector; secondly there is a methodological rationale as given the research undertaken on climate scenario downscaling (Chapter 3), it was not possible to build robust scenarios at the daily times scale for rainfall—hence the statistics of change in the shorter run droughts under the climate scenarios produced would not be valid.

4.3 Soil water balance

Given the continued use of a soil water based drought indicator, it is appropriate to examine the simulation model used to derive it. Feedback from the 2008 consultation workshop (Clark and Tait 2008) identified some key research questions about the representation of processes in soil water balance models in this context. The principal questions focused on the relatively simple approach used to determine the water balance by Mullan et al. (2005), assumptions surrounding the way models represent the process of evapotranspiration, and related to this the role of carbon dioxide fertilisation or the behaviour of leaf stomata under climate change. These issues are explored here through intermodel comparison, model verification with observations and sensitivity testing. Before this, a brief description of data sources used in the methodological development, and steps taken in data quality control are presented. Data homogeneity was a key concern raised by Larsen (2005) in a critique of previous methodology development work undertaken at Lincoln (Porteous 2005).

4.3.1 Data and quality control

A soil moisture model evaluation data set was developed using NIWA's climate monitoring network. In the late 1990's NIWA began establishing a network of soil moisture monitoring sites using electro-magnetic flux tape probes, which provide sub-daily estimates of soil water flux. The majority of probes are installed on flat pastoral land both for practical reasons and also for the purpose of monitoring agricultural drought. The probes are usually installed in conjunction with a "tier two" climate station so a full suite of climate variables are also available. The network has now grown to over 60 sites, although for the purpose of this study 11 of these were discarded because of short records or quality problems. The locations of sites used in this study are in Figure 17.

Figure 17: Site locations of the NIWA soil moisture monitoring network.

The primary climate and flux probe data are stored in base form on NIWA's national climate data base (CLIDB) . The following processing and quality control procedures were applied to develop the model evaluation data:

- extraction of electromagnetic flux probe data from CLIDB for the first metre of soil;

- as part of the extraction, aggregation of the flux probe observations to a daily time step;
- screening of the data to eliminate sites with records less than 12 months in duration, and to eliminate obvious outliers, deemed as observations more than half a standard deviation above the 99th percentile of the time series and below zero.
- the flux probes take 1-2 years to stabilise and provide a reliable *in situ* estimate. The first twelve months of all series were discarded, and a linear trend fitted. All series with a positive significant trend were manually inspected and a decision was made to: discard the data set entirely (3 sites); discard a further twelve months of the series (2 sites); remove the linear trend from the data (2 sites).
- following these procedures further aggregation of the daily soil water observation data to 7 day and 10 day observation periods.
- where available, extraction of 18 years of observed daily climate variables (rainfall, maximum and minimum temperature, radiation and potential evaporation, 1990-2008) from the nearest climate station to the soil water observation site. For the majority of sites, the climate station was in the immediate vicinity of the soil moisture probe (within 500m), but for around 10 of the sites the climate station was 1-7 km away. Sites without a climate station within 7km were discarded;
- extraction of the Virtual Climate Station Network data (VCSN) for the closest grid point to the monitoring probe. For the variables listed above the VCSN data was bias corrected using a cumulative distribution mapping method. The bias corrected VCSN series were used to infill any missing data, creating a complete patched set at each site.

The main uncertainty using the data is the validity of the standard in-built calibration of the aquaflex probe. There is concern that actual soil moisture determined by the probe could be different to that measured by neutron count, weight by sample extraction or other methods for determining *in situ* soil moisture.

4.3.2 Inter model comparison

Four models are compared in this study (Table 5), all simple water balances that were chosen because they require minimal information about the landscape and are predominantly climate driven. In the context of work in this area, all of these models are low in complexity and more physically based water balances are available and used in New Zealand (e.g. Fowler et al. 2008; Palmer et al. 2009; Wilson et al. 2003; Woodward et al. 2001; and Andrew and Andrew 2007) . Each model relaxes some key assumptions and expands the system boundary of the water balance. The base model is NIWA's current operational system for drought reporting and the water balance used in the Porteous (2005) study. It is abbreviated here as 1L_STEP, as it is a 1 layer tipping bucket model which uses a set step function of soil water deficit to regulate the relationship between potential and actual evaporation (PET:AET). This step function is set so that above a soil water deficit of 50 percent of water holding capacity (75mm) AET is half of PET. Below this threshold AET is set equal to PET. In its operational form the model is set up assuming a 150mm water holding capacity (WHC) across New Zealand, arguably representative of the rooting zone of a white clover-ryegrass sward. This is a simple 1 parameter model that is largely controlled by climate variability, assuming that the main landscape factor is potential depth of water in the profile.

The 1L_EXP model introduces an exponential decay function to this framework to regulate the PET:AET relationship. This is a simple way of incorporating soil physical properties into a model framework, that is the tension at which water is held in the soil matrix and hence the rate of drying as water availability moves from wet to dry. This is a two parameter model that is largely climate driven.

The single layer assumption is relaxed in the 2L_STEP model where preferential evapotranspiration occurs from the surface soil layer and drainage occurs as a slugging process. The step function is also set up to provide a linear PET:AET relationship to a water deficit threshed where AET is equal to PET. Unlike 1L_STEP, these thresholds and the linear rate of drying can be set. This also introduces a more complex parameterisation with additional functions and terms introduced to govern the flux of water between the layers, the AET:PET mechanism and an assumed plant root distribution for preferential extraction.

The final model examined (2L_VEG) tightly couples the two layer scheme to an ecophysiological vegetation model with a canopy (big leaf) rather than single leaf structure. In this case it was set up as best as possible to represent a ryegrass-white clover sward. The evapotranspiration process is represented explicitly by partitioning it into transpiration and bare soil components, thus moving away from the functional structure to regulate PET:AET. Stomatal conductance is represented by the parameterisation of the canopy along with the simulated flux of seasonal growth and maturation. Along with the tension at which soil water is held in the soil matrix and the degree of moisture deficit above the canopy governs the rate of flow, potential transpiration is regulated by the canopy flux. Potential bare soil evaporation is an energy balance. Actual evapotranspiration rate is assumed to be the maximum of these components. This is the most complex of the models examined with 15 parameters required, not including specific plant parameters which are fixed to represent white clover/ryegrass.

Table 5: Description of water balance models examined in this study.

	1L_STEP	1L_EXP	2L_STEP	2L_VEG
General description	One layer step function•	One layer step function	Two layer step function	Two layer with dynamic vegetation
Model family	Tipping bucket	Tipping bucket	Force-restore	Force-restore
Number of parameters	1, WHC assumed 150mm	2	7	15
Vegetation subroutine	None	None	None	Dynamic (growth and plant function simulated)
Rainfall/runoff mechanism	Lumped[+]	Lumped[+]	Saturation threshold	Wetness, vegetation
Evapotranspiration mechanism	Lumped Step function of soil water deficit	Lumped Negative exponential function of soil water deficit	Lumped Step function of soil water deficit	Partitioned Transpiration and soil evaporation simulated separately using resistance terms
Drainage mechanism	One layer using saturation threshold	One layer using saturation threshold	Flow rate between layers	Flow rate between layers

The first level of examination is to compare the national level drought characterisation of each model by running each framework with comparable assumptions (parameterisation) on the same set of daily input data (Figure 18). In this case the final drought probabilities were separately for each of the into the different characteristic drought types listed in Table 2. This analysis highlights that there is minimal difference between models when used to characterise long duration droughts at all three intensity levels. This suggests that droughts of long duration are largely climate controlled, and that variability in rainfall the dominant factor determining the drought probabilities at this scale.

Differences between the 1L-STEP model and the other three frameworks were evident in the quantification of medium duration droughts, notably this model results in an underestimation of probabilities. These are most pronounced across the North Island for low and moderate intensity droughts and underestimation of high intensity medium duration droughts on the east coast of the South Island. This illustrates that there is a degree of soil control on drought probabilities at shorter durations—that is the set rate of drying in the structure of the 1L_STEP model appears to introduce a systematic bias when compared to the other models where PET:AET is either a more continuous function or simulated dynamically.

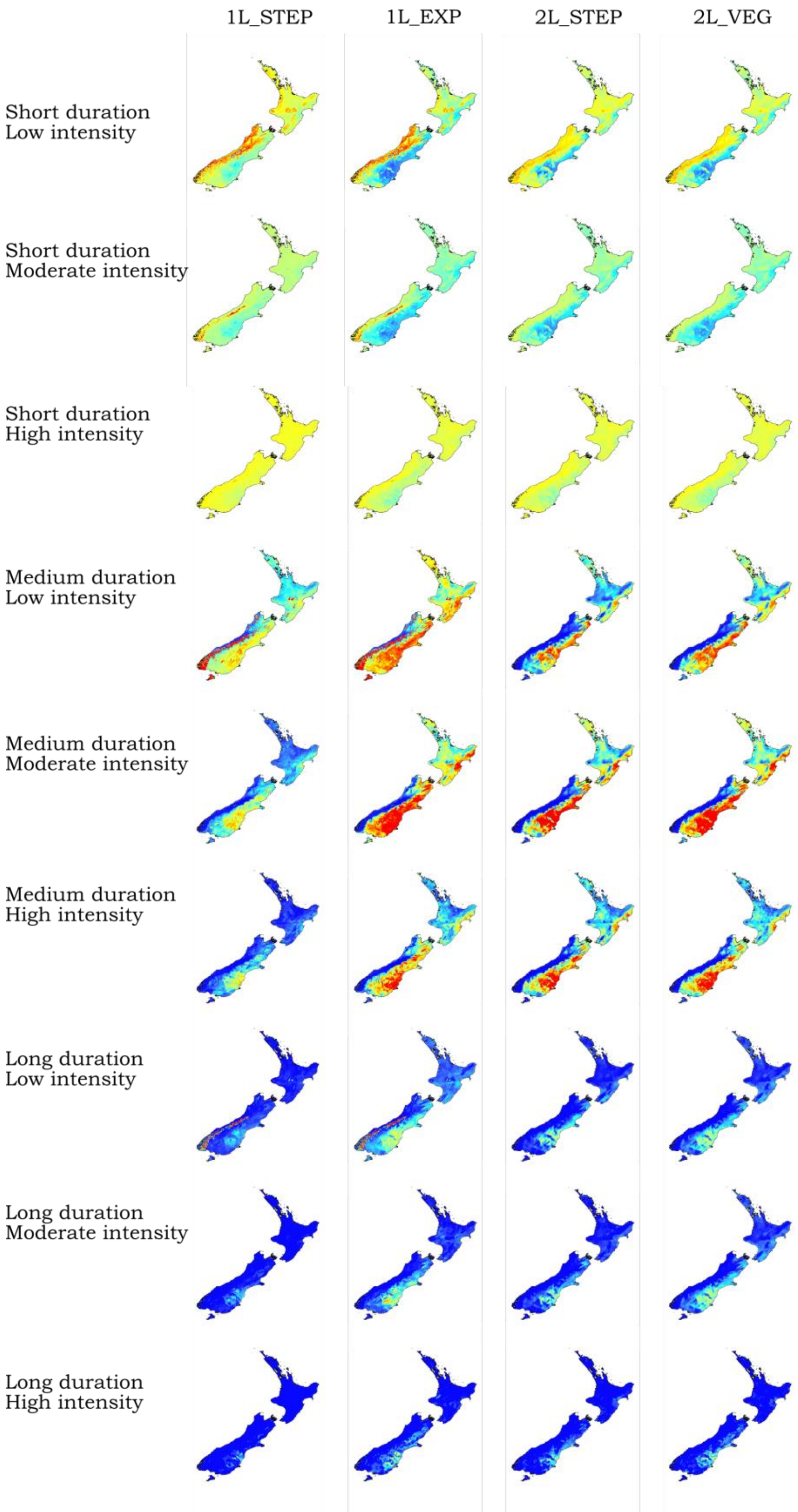


Figure 18: Probabilities of characteristic drought types determined by each soil water balance model. Red is 20 percent of time spent in drought grading to dark blue less than 1 percent of time.

There are regional differences between all the models in the quantification of short duration droughts, although a systematic difference is not obvious. High intensity short duration droughts appear to be slightly over estimated on the east coast of the South Island by 1L_STEP compared to the other three models. 1L_SPEP and 1L_EXP appear to overestimate short duration low intensity droughts compared to the other two models. This highlights that there is a complex interaction between the model structures, their parameterisation and climate variability at this time scale.

4.3.3 Model optimisation

Given this result it was decided to perform a model optimisation experiment in order to investigate the behaviour of the different model structures further. In this context model optimisation is not used as a means to obtain a set of preferred parameters, but rather a diagnostic tool to examine the models behaviour. Technically this was implemented by minimising the norm of the residuals using a non-linear technique known as the Levenberg algorithm (described in Aster et al. 2005). Standard errors for the parameters were obtained from the co-variance matrix of the objective function. The optimisation was not constrained in any way and the initial starting parameters defined as for the intermodal comparison of drought probability calculations above. The procedure was allowed to iterate either until a minimum was found (generally about three seconds CPU time) or for 6 minutes CPU time.

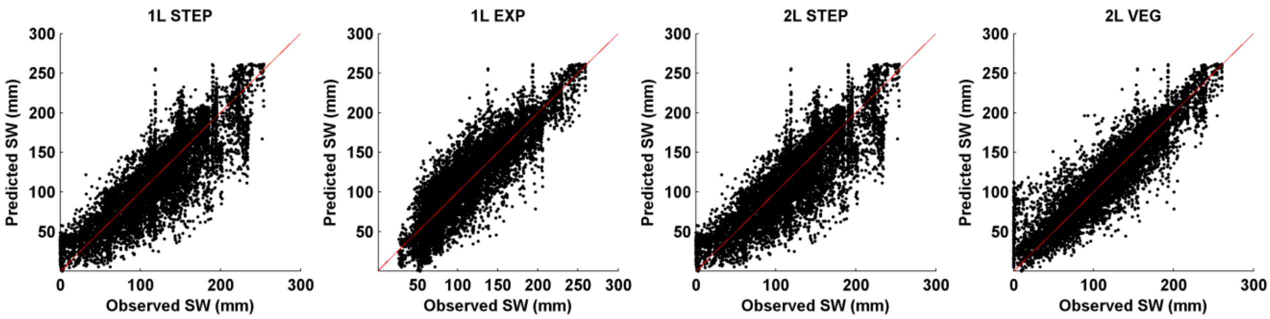


Figure 19: Scatter plots of optimised soil water balance models at all sites.

The results are summarised by the scatter plots in Figure 19. Generally a stable estimation of model parameters was possible at a large number of sites for the four models. There were also a smaller set of sites where the models could not be optimised, reflecting questionable observations. Systematic bias was found in the in the 1L_STEP model where the standard error of the WHC capacity was large (cross site mean of 75) and often greater than the value found, that is the shape of the objective function was flat and broad and it was not possible to find a robust global minima. This is indicative of an under specified model where more process information is needed in the model structure to accurately account for the observed variability.

This problem appeared to be addressed by the 1L_EXP model where the AET:PET relationship was a continuous function. Lower standard errors were obtained for its two parameters at a high number of sites that were generally within a physically plausible range This was a relatively stable optimisation, suggesting that the model described the main processes driving the observed variability. There is however further potential to explain more variability, as the upper limit of its precision appeared to be in the vicinity of eighty five percent. Hence this model is able to adequately resolve the seasonal and month to month variability in the observations, but more a more detailed model would be needed to explain the day-to-day differences in soil water flux.

The 2L_STEP model appeared to be the most unstable, having a large number of sites where the model could not be optimised or a local rather than a global minima found. Where it could be optimised a number of the parameters were well outside a physically plausible range, for example a negative value for the point at which AET is set to PET. This is indicative of a model that is not parsimonious, where a part of the model's structure is described in significant detail (the two layers) but another is ill-specified. In this case it is likely to be the lack of partitioning between bare soil evaporation in the two layer scheme where there is preferential evaporation from two distinct layers, as well as the lack of plant dynamics.

The dynamic vegetation scheme in the 2L_VEG model appeared to address this problem at a select number of sites. It provided a stable and accurate solution at the one third of the sites where it could be optimised. In these cases, the optimisation demonstrates the improved precision capable with a more detailed model, where it peaked at around 90-95 percent of observed variance. It captures some of the day-to-day variability in soil water flux that the 1L_EXP structure did not. This may not be the true upper level of precision possible for the model itself, as the quality of the climate input data is likely to limit further gains beyond this level. The results and difficulty in optimising this model at the other sites illustrate that the added complexity introduces another common modelling problem, high dimensionality. Given the degrees of freedom introduced by the detailed structure and model dynamics, the objective function was considerably noisier and in many cases a global minima could not be found. While the next logical step is to change the optimisation methodology or the models structure, taking this step is not the goal of this research.

4.3.4 Model verification

The first level of verification examined observed and simulated soil water at individual sites. Figure 20 provides three examples of the type of estimate emerging from this exercise, time series obtained for one model at three sites chosen randomly. In all three cases the modelled soil water captures the month to month and inter-seasonal variation in soil water well. This type of graphical display of data is a relatively standard way of illustrating the verification or validation of soil water balance models.

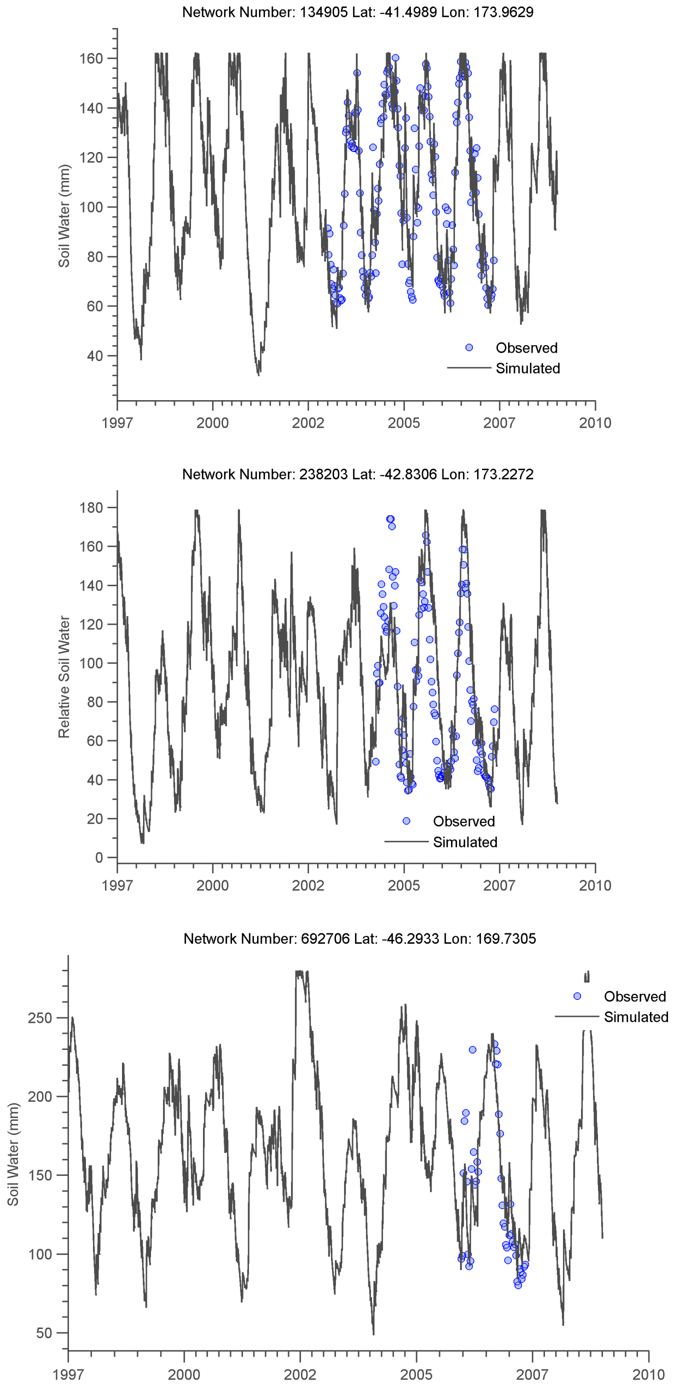


Figure 20: Examples of site level verification of soil water balance for 1L_EXP model that has been optimised.

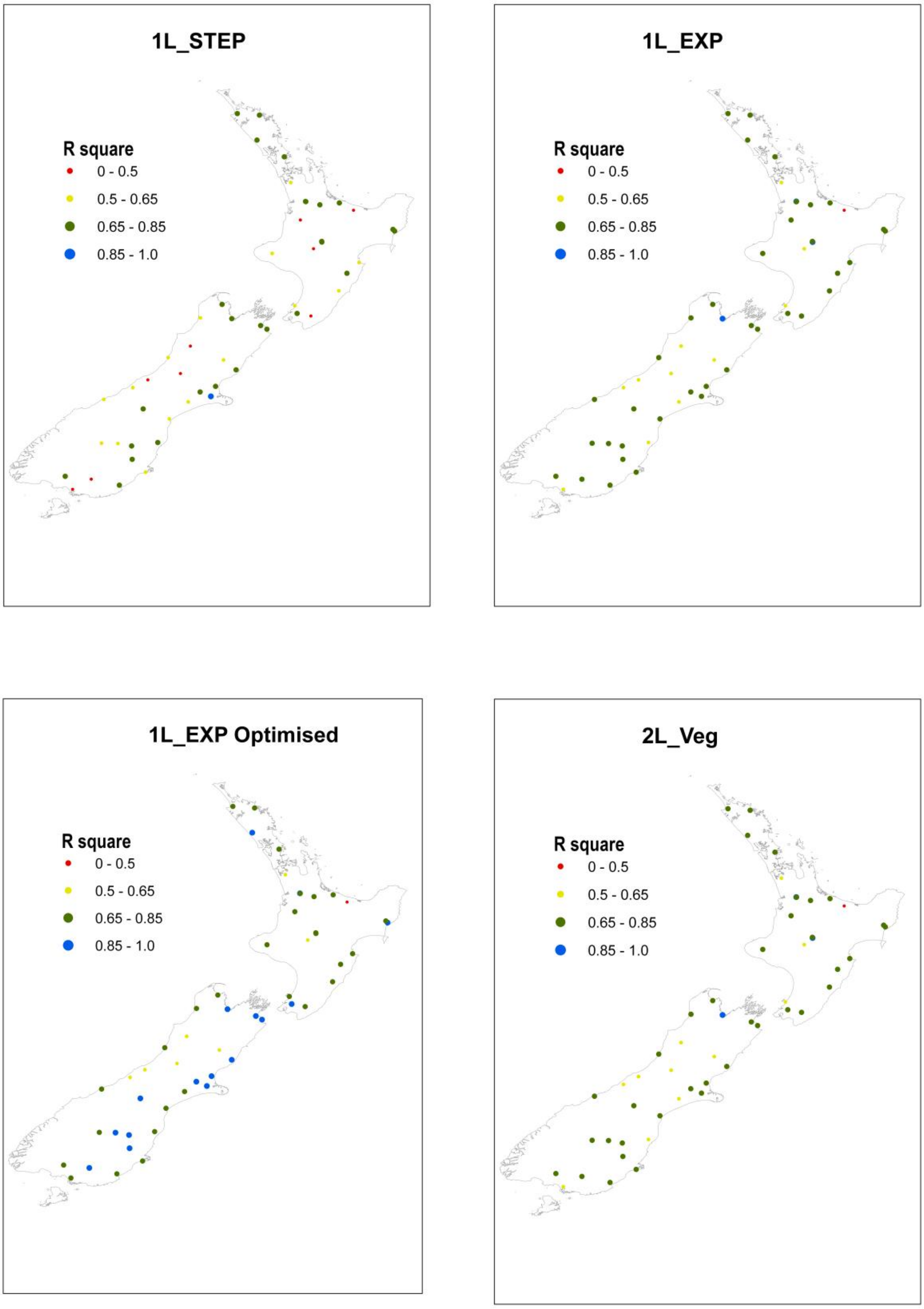
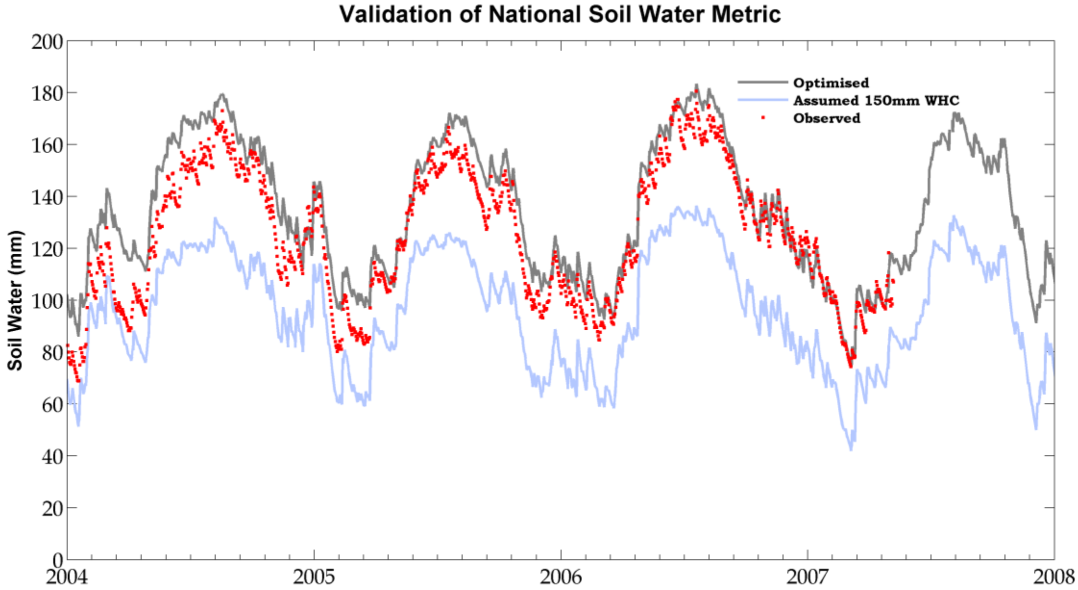


Figure 21: Correlation coefficients between models and all sites in the monitoring network .

To summarise, the site level results for the entire network the correlation coefficients between observed and predicted obtained at each site are mapped in Figure 21. For the 1L_STEP model the majority of sites had correlation coefficients in the range of 0.5 to 0.85. The scores were slightly improved for the 1L-EXP model, the 2L_STEP (not shown) and the 2L_Veg model, all with the majority of sites yielding a correlation coefficient in the range 0.65-0.85. The 1L_EXP model optimised (described below) illustrates the type of improvement obtained when local parameterisation is implemented at each site. Generally this shows that the correlation coefficients are improved by about 0.1-0.15, that is by 10 to 15 percent of variability by taking local estimates into account.

The second level of verification examines the implication of aggregation of soil water metrics on model selection, a strategy that would be typical in the construction of regional and national drought indicators. The first graph (Figure 22) shows the influence of optimisation (in this case for 1L_EXP) on a national metric, whereby the optimised version clearly tracks the observed soil water well, whereas the model with assumed 150mm WHC and set AET:PET modification rate does not. The next graph displays the data as a relative soil water metric, that is the observed variability expressed as a proportion of its observed maximum. Importantly for 1L_STEP, this is the form that the data is used in to provide ongoing advice about drought. This clearly shows that in relative form both the 1L_SPEP and 1L_EXP model provide a satisfactory account of seasonal variability.



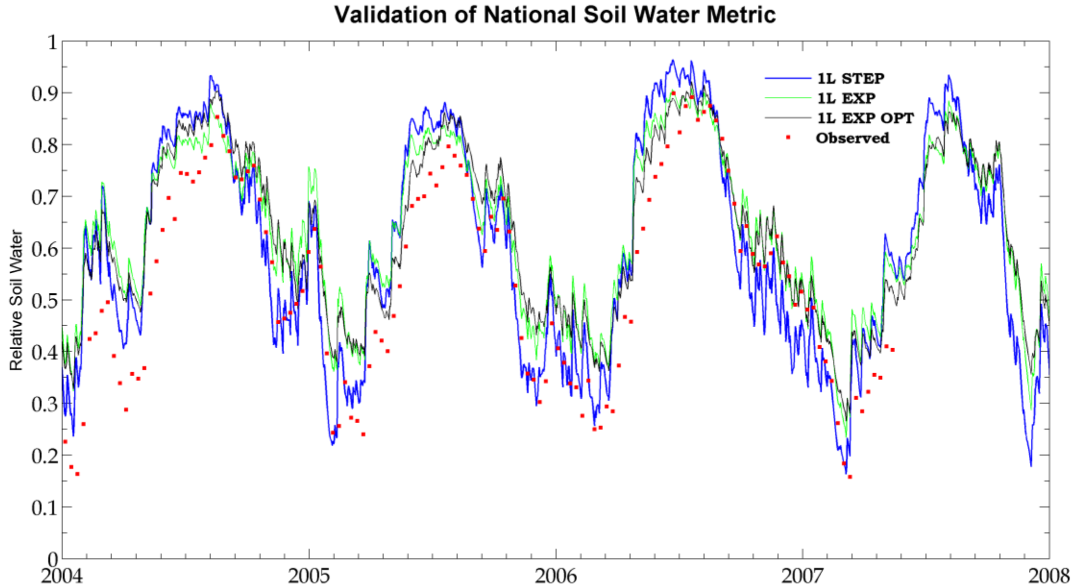


Figure 22: Validation of the national soil water metric for absolute and relative estimates of soil water.

4.3.5 Sensitivity of drought to changes in pasture water use efficiency

The 2L_VEG model was used to perform a sensitivity experiment evaluating the effect of changes in stomatal conductance on drought probabilities. The experiment incrementally changed the canopy resistance and some of the photosynthetic response parameters of the model to reflect the response of closing stomata under elevated carbon dioxide (or increasing the water use efficiency). The results are summarised in Figure 23, as the changes in drought probability (years per decade) per 15 percent increase in water use efficiency. Similarly to the intermodel comparison the results are partitioned into characteristic drought types.

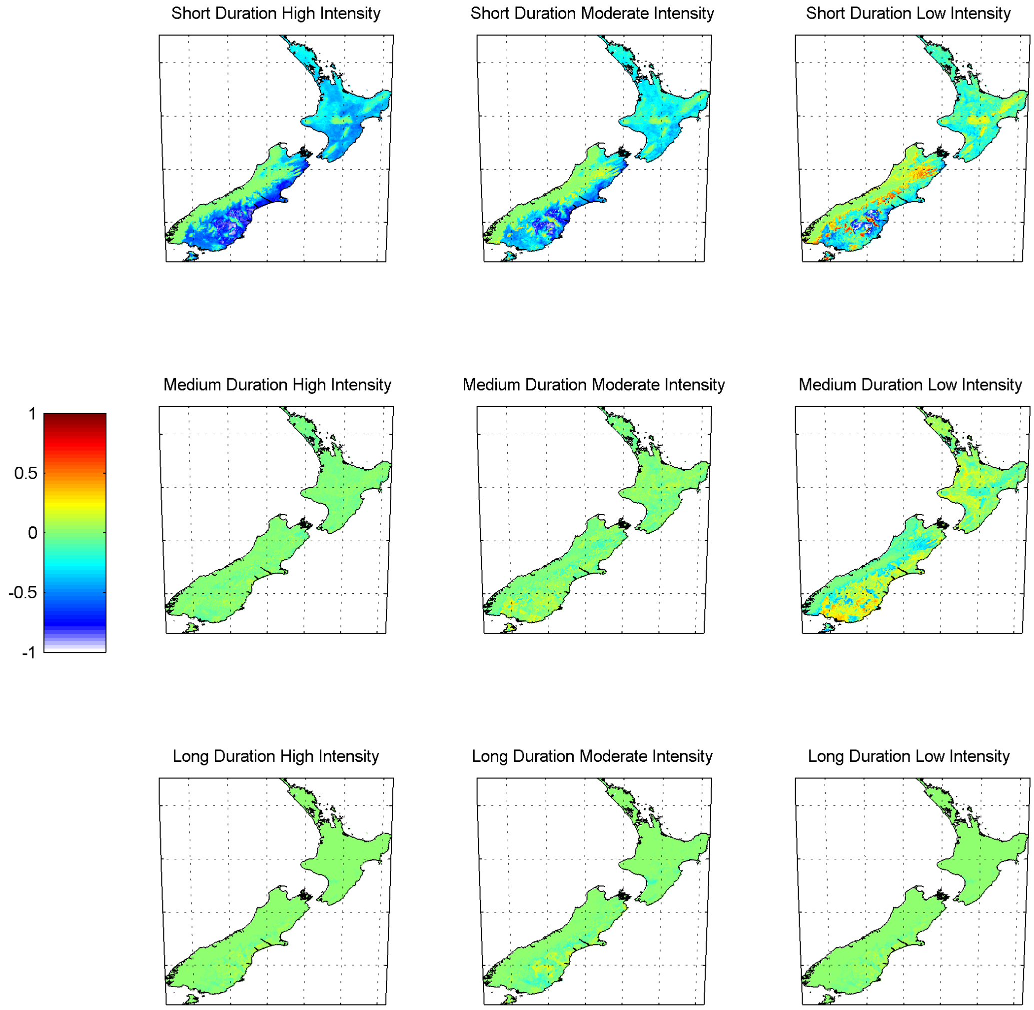


Figure 23: Sensitivity of drought indicators to increasing water use efficiency (reduced stomatal conductance). Figures are changed drought probability (years per decade) for a 15 percent increase in water use efficiency.

The results for short duration droughts are the expected response of improving water use efficiency of plants under carbon dioxide fertilisation. Pastures are transpiring less water given closure of the stomata and enhanced photosynthetic response, so for short duration droughts at all intensities the long term probability of drought decreases in many regions as more water tends to be left in the soil matrix. However, the results for medium and longer duration droughts could seem at first instance outside of the expected response, as drought probabilities at this scale appear to be insensitive to changes in stomata closure. The explanations for this result is that droughts of this scale are predominantly driven by the climate signal rather than processes at the soil and plant level. Simply put, a climate driven water deficiency of one month or greater appears to create a circumstance where most of the water has been lost from the system over that time frame.

This sensitivity analysis is consistent with recent field based research using Free Air Chamber Experiments (FACE). In a recent meta-analysis of published FACE work Kimball (2010) describes results which show that water limitations essentially override the stomata response under carbon dioxide enrichment. The summary of this evidence describes the physical mechanism clearly: if plants are growing under limited water supplies at seasonal time scales, their stomata will be closed a large proportion of the growing season, so elevated CO2 will have comparatively little effect on stomatal conductance (Kimball 2010). When active the carbon dioxide fertilisation process as expressed by physiological and cellular level responses is complex and there are feedbacks given soil fertility (Edwards et al. 2005), detectable species level variability in response (Ghannoum 2001) and **???**

It is extremely important not to confuse this mechanism with an evaluation of the validity of the carbon dioxide fertilisation, or to diminish the importance of the process when considering future productivity. Potential increases in plant growth and improved water use efficiency under carbon dioxide enrichment have been well demonstrated in chamber, FACE and other approaches, and they can be expressed strongly when water is non-limiting (Kimball 2010). Other limiting factors like nitrogen, different capacities to respond across species and plant functional groups, and feedbacks on canopy temperature and leaf area, have also been demonstrated (Kimball 2010). It is important to be aware of the full range of potential responses, particularly when considering production variability in a whole farm management for adaptation under climate change. The sensitivity analysis simply illustrates that there are times of climatic limitation where this process is not expressed as strongly—assessing the likelihood of these climatic events is the primary focus of this study.

4.3.6 Model selection

Given the results outlined in this chapter the 1L_EXP model was selected as the water balance for further operational use in this study. The model evaluation clearly illustrates that for seasonal scale events simpler water balance models are appropriate, given that they are constrained to examining droughts at the ‘climatic scale’ of around one month in duration. There isn’t a strong case to move beyond this level of model complexity for the purposes of constructing regional to national level drought indicators based on soil moisture. Using a more complex model with more parameters would lead to another problem, the difficulty of accurately specifying model parameters across the country—it would require a larger number of assumptions to be made or analysis of the spatial distribution of land use, grazing management, soil fertility and pasture species distribution. However, if drought indicators need to be precise enough to represent either localised (farm level) or shorter run (day-to-day) variability, there is a case for using an alternative model.

Marginal improvements to NIWA's current operational model (1L_STEP) can be made by utilising a model that can better reflect the influence of soil properties on the rate of drying (1L_EXP). This is because the current model has a structural bias in the estimation of the rate of drying of the profile due to the set AET:PET function. This does not invalidate the 1L_STEP models use in the previous study nor as a basis for ongoing advice about seasonal drought at regional to national scales. However given the set parameterisation it is particularly important to use a metric of relative rather than absolute soil moisture when examining output from this model.

The sensitivity tests found that probabilities of medium to long duration droughts based on soil moisture are not highly sensitive to changes in plant water use efficiency. This reflects a mechanism where water limitations essentially override closed stomata under enriched carbon dioxide. Given the focus on climatic scale drought events in this study, it was not appropriate to pursue the development and implementation of an operational modelling framework that is sensitive to carbon dioxide enrichment.

This chapter demonstrates the importance of considering the parsimony of drought indicators and simulation models with the temporal and spatial scales of a planned analysis. There is clearly shorter run variability like week long droughts that are beyond the capacity of some drought indicators and simulation models to track accurately. For this reason the choice of indicator, statistical analysis and model selected to simulate water deficits targets medium to longer duration droughts at regional to national level. The issue of scale is also evident in the next chapter where downscaling of climate from General Circulation Models (GCMs) is undertaken.

5. Downscaling global climate change scenarios

5.1 Summary

This chapter describes methodological work undertaken to develop a generalised empirical-statistical downscaling scheme (ESD), suitable for providing regional scale climate change scenarios for New Zealand. The scheme resolves monthly climate variables across a five kilometre grid over the whole country. Rainfall and PET are the key focus variables of this study as they are the input variables to the water balance used to determine drought probabilities in Chapter 4. Monthly temperature and some initial work with daily variables are also presented.

The final ESD scheme is developed using partial least squares regression. This is an alternative to more traditional regression approaches, whereby the dimension reduction of climate predictors is optimised. The key result is an improvement in the stability of the downscaling model and a detectable improvement in precision particularly for monthly rainfall compared to multiple regression. The approach also supports a key change from previous downscaling work for New Zealand in that the month-to-month variability is preserved in the projections—this provides more plausible scenarios for the purpose of drought analysis.

Thorough testing shows that the empirically downscaled scenarios provide a satisfactory estimate of 20th century drought probabilities when compared to those derived from actual observations. While this signal degrades slightly when 20th century fields from GCMs are used, scenarios derived from the climate models are within acceptable levels of precision for application in this study, particularly for New Zealand's agricultural land.

Two approaches for downscaling potential evapotranspiration (PET) are developed and evaluated, each with its own strengths and weaknesses. A clearly superior candidate to go forward in the scenario analysis of droughts was not found. While including radiation as a predictor of PET in the transfer functions appeared to improve results based on observed climate, there was some instability when the functions are resolved with 20th century radiation from the GCMs. Both approaches are used in Chapter 4 to produce families of drought scenarios.

While the work undertaken here provides adequate scenarios for use in this study, a number of constraints and opportunities remain. Downscaling climate for the alpine zone remains problematic, likely due to inadequacies in the observation model in these regions. Similarly downscaling in regions where a high proportion of rainfall is generated from localised systems (convection) does not perform as well as regions with a greater degree of synoptic rainfall (frontal). Encouraging results for empirically downscaling daily climate variability were obtained, and although not used further in this study because of poor results in estimating the full rainfall distribution, these initial trials suggest there is potential to develop an operational scheme.

5.2 Background and rationale

Downscaling is necessary in climate change impact analyses that seek to constrain regional climate by information from global climate models (General Circulation Models, GCMs). It is particularly important in the New Zealand context, as given maritime, topographic and convective climate processes, local to regional scale variability is not always well represented by the broader global scale features simulated by GCMs. GCMs have a broad resolution of approximately 50-200 km and are known to represent broad synoptic features well. This is known as the ‘skilful scale’ of these models and it can be exploited to provide finer mesoscale scenarios, usually between the 1-30 km resolution. This finer scale is more relevant for assessment of hydrological and agricultural system responses in New Zealand.

Two approaches are commonly used for downscaling, empirical-statistical downscaling (ESD) and regional climate modelling (RCM). In RCM the GCM provides boundary conditions for a limited area physical model. In ESD broad scale climate predictors are related to predictands (both observations) by empirical transfer functions which are then applied to projections from GCMs. When examined in detail there are considerable technical differences between RCM and ESD based schemes, and is not relevant to describe these here. Benestad (2008) summarises the main tradeoffs between each methodology:

- RCM provides a more physically plausible basis for developing future climate change scenarios at the meso-scale;
- RCM requires more computational resources than ESD. As a result it is usually only applied to a small number of GCMs. Current practice uses bias corrections similar to ESD schemes to produce reliable projections for a downstream model. This is because of limitations in the cloud parameterisation in many physical climate models;
- by comparison ESD is computationally efficient, can be readily applied to a number of GCMs and has been shown to exhibit similar local level precision compared to RCM;
- ESD usually under estimates variance (extremes), and care needs to be taken to ensure that the physical relationship is both strong and physically plausible as it is based on the assumption of its future stationarity (Schmith 2008).

The emphasis in this study is on developing an ESD scheme, although this work runs parallel to efforts currently underway to develop RCM based projections for New Zealand by NIWA. Increased demand for climate change impact analysis has seen ESD become well described in the international literature over the last five years. There are subtle differences between almost every ESD scheme relating to localised environmental features, and also advances in methodologies including improved data collection and availability. Generally ESD schemes fall into the following categories:

- interpolation methods;
- pattern scaling where current climate is adjusted on a per degree warming basis given sensitivities from GCMs ;
- regression type methods where linear transfer functions are developed (e.g., Wilby 1998; Wilby and Wigley 2000; Timbal et al. 2009; Li and Smith 2008; Charles et al. 2004);
- analogue based methods where past climate are used which match potential future climate regimes (e.g., Imbert and Benestad 2005; Wetterhall 2005);

- stochastic weather generation where climate is simulated given estimated of current and future probability distributions (e.g., Semenov and Barrow, 1997); and
- emerging non-linear approaches where alternative model structures and optimisation methodologies are employed (Cannon 2009; Li and Sailor 2000).

This breadth of approaches in ESD is well described in general level texts (Benestad et al. 2008), and also in technical guidance for use of scenarios in impact studies undertaken as part of the IPCC process (Wilby et al. 2004). ESD methodologies are also an active area of climate research, with new techniques emerging recently (e.g. Cannon 2010; Timbal et al. 2009; Benestad 2010). Recent work has focussed on the downscaling of the daily rainfall distribution with attempts to examine extremes (Benestad 2010; D'onofrio et al. 2010; Cannon 2009).

For end users of climate change projections ongoing changes to methodologies underpinning climate projection can be frustrating, as new versions appear to change the basis for impact and adaptation planning. The counter argument is that ongoing innovation and improvement of climate modelling and downscaling is critical to the plausibility of climate change scenarios at the regional level where impacts occur. It has also been common to compare different approaches within the ESD framework and dynamical downscaling in an effort to choose the optimal approach. As described above there are tradeoffs between methods and none are completely superior. The reality is that the range of methodologies are best seen as complementary, and focussed on the difficult task of improving the physical plausibility and accuracy of climate change scenarios.

In the New Zealand context ESD has been used for some time. The CLIMPACTS system (Kenny et al. 1995) has in the past used a pattern scaling where current climate is adjusted on a 'per degree of global warming' basis given a climate sensitivity ratio derived from models. Mullan et al. (2001) used multiple regressions to develop an ESD scheme given Trenberth Indices (Z1 and M1) as predictors and monthly temperature and rainfall as predictands. Projections for two time slices form the basis of MfE guidance material on climate change for New Zealand: 6 IPCC Third Assessment Report (TAR) models centred on the 2030s and 2080s were used in the first Guidance Manual (MfE 2004); and 12 IPCC Fourth Assessment Report (TAR) models centred on 2040 and 2090 in the second Guidance Manual (MfE 2008). The previous work on drought analysis used two of the IPCC Third Assessment Report models (Mullan et al. 2005). In an application outside of climate change analysis, Renwick et al. (2010) developed an ESD scheme relating the broad scale National Centre for Environmental Prediction (NCEP) ensemble with local scale observation to produce 15-day probabilities of rainfall and temperature. The methodology is based on stepwise regression with a range of synoptic variables screened for skill.

ESD is based on the assumption of stationary as it uses transfer functions that are developed on a past relationship to project into the future. Stationarity refers to the assumptions that the relationships between broad scale predictors and finer scale predictands do not change in the future. Given this assumption there is ongoing debate surrounding the appropriateness of using ESD to develop future scenarios (Schmith 2008). The core concern is that given feedbacks and nonlinearity in climate processes the relationships in the transfer functions could change and therefore they may not fully capture the nature of future change. More specifically, the change may move climate to a level that is outside the range upon which the ESD scheme was developed and shown to be valid, thereby providing a weak and possibly biased basis for inference. Despite this there are some general benchmarks for an ESD scheme to meet in order for it to be more confidently applied in scenario studies. These include:

- there needs to be a strong and physically plausible relationship between the broad scale predictors and finer scale predictands. In efforts to model complex processes or seek improved accuracy, some ESD schemes have moved well away from a transparent physical basis and place too much emphasis on the underlying data sets.   This has been avoided in this study;

- related to this the transfer functions built with data from a training period should be independently verified given data outside the training period. This implies the use of either block partitioned observation sets or the design of an appropriate cross validation;

- the implication of this is that a downscaling model should be subject to some standard fundamental tests including: goodness of fit to quantify precision; have normally distributed and randomly structured residuals; and have minimal residual trends to examine the presence and strength of non-stationarity; and

- an ESD scheme must also yield sensible physically plausible projections that reflect the change projected by the GCMs. Care needs to be taken to ensure that the conclusions drawn in a projection study are not due to bias or error in the ESD scheme. A study should reflect the climate change signal downscaled with a stable physically plausible scheme. This is an application specific issue and can be achieved by quantification of error in the variables of interest, in this case the probabilities derived from the drought indicator.

5.3 Technical considerations

A defining feature of any ESD scheme is the selection of predictands and how they are treated prior to model fitting. This is the first stage of developing a physically based relationship as deductive hypotheses need to be formed about which subset of broad scale predictands might explain a high proportion of the local predictor variability. Single through to multiple predictand relationships might be hypothesised and tested, and it is tempting to select many variables in order to explain a complex process like rainfall. However, a model that has a large number of predictors and seemingly accurate fit can be statistically invalid. An overly complex scheme has higher risk of over fitting, illustrated by poor predictive skill when tested independently. Even if a complex ESD scheme has good predictive skill it may also have redundant variables and the transfer functions may yield unusual or physically unrealistic projections when resolved with future GCM fields. A common strategy is to undertake screening experiments or apply a step-wise procedure to evaluate which individual predictors or which combinations perform best.

Climate predictor fields also contain many individual time series that are spatially and temporally correlated and have lot of noisy redundant information. This creates a high dimensional problem even if only a few individual predictors are selected. Physical and or statistical transformations of predictors are usually undertaken to reduce the dimensionality of the problem. For example, the use of the Trenberth Indices Z1 and M1 by Mullan et al. (2005) is a form of transformation which aims to reduce dimensionality given physically based inference. This is similar to the use of indices like the North Atlantic Oscillation and other teleconnection indices in European ESD schemes. Statistical transformations are also used widely in ESD, with Principal Component Analysis (PCA) now a routine approach used to reduce the dimensions of the predictor space (Benestad et al. 2008). The empirical orthogonal functions (EOFs) used by Renwick et al. (2010) are an example of this approach. Approaches which use PCA are collectively described in this study as 'the common EOF framework'.

5.4 Model experiments

To build an appropriate ESD scheme for this study of drought under climate change a number of incremental experiments were undertaken and critically examined. These investigated some methodological choices underpinning a practical working scheme for the purpose of projecting drought. From the beginning a regression based scheme was pursued as it was seen to be the most promising framework in which to develop a generalised national to regional level approach. This does not preclude future development of advanced interpolation methods or non-linear approaches. Stochastic weather generation was not pursued because it was deemed to be more suitable for individual site or small to medium catchment scale analysis. Given the time series of observations available, there is insufficient record to develop a nationally comprehensive analogue based scheme. Before the downscaling experiments are detailed, the common data sets underpinning the work are described.

5.4.1 Observation and projection data sets

The observed predictors used to build the ESD scheme in this study are broad scale data from the NCEP reanalysis (Kalnay et al. 1996). The NCEP reanalysis is a state of the art data assimilation program providing global atmospheric and surface variables on a 2.5 degree grid from 1948-present. A higher quality re-analysis is available from 1979 to present that utilises some advanced remote sensing products in the data assimilation, and this set was used in this study. The variables obtained for this study are pressure, temperature, precipitation, geopotential height, humidity and wind (u and v) at multiple levels as well as surface radiation. Selection and further treatment of predicators specific to each predictand are described in following sections.

The regional scale observations used as predictands are a nationally comprehensive set of station observations which have been interpolated across New Zealand to a regular 0.05 degree grid fpr the period 1972-2009. The application of thin plate smoothing splines to interpolate observations in New Zealand is described by Tait et al. (2006) for precipitation; Tait (2008) for temperature; Tait and Liley (2009) for radiation, and Tait and Woods (2009) for PET. This body of work and resulting data set is known as the Virtual Climate Station Network (VCSN). All daily values are aggregated to the monthly time step for the purpose of developing transfer functions at this time scale. For some of the initial experiments a subset of 10 sites selected from the full grid were used to evaluate and test the methodology (Figure 24).

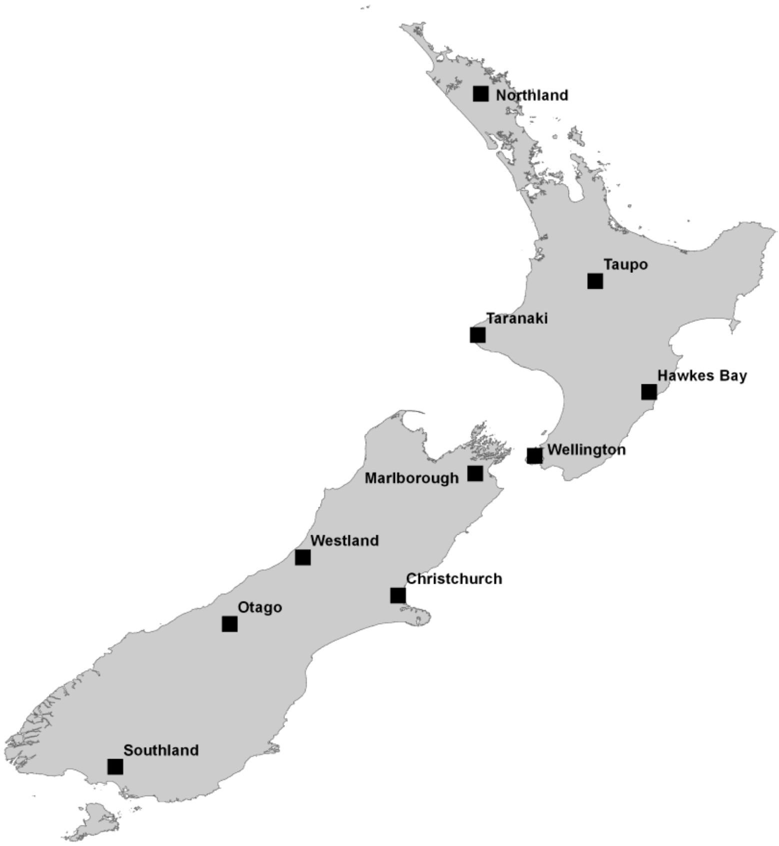


Figure 24: Location of test sites used in the initial development of the ESD scheme.

The predictors used to develop the future scenarios were obtained for up to 24 GCMs based on simulations undertaken for IPCC's Fourth Assessment Report (AR4). Data were from two data distribution centres: for monthly fields the IPCCs data distribution centre via the World Climate Data Centre; and for daily fields from the Coupled Model Intercomparison Project (CMIP3).

The data acquisition and quality screening aspects of obtaining the GCM fields are not described in any detail, but did represent a substantial commitment of resources. Monthly projection data from 1900-2100 are available for most of the GCMs for only a subset of surface variables: precipitation, temperature, radiation, humidity and wind. The availability of daily data is more restricted both in terms of the models, level, time span and variables. Data at upper atmospheric levels are not as widely available as surface data across this model ensemble at both monthly and daily time steps, to the point where it was not practical to build a downscaled Multi Model Ensemble (MME) based on these as predictors. This is a data availability constraint on the ESD scheme, as it has been shown in other applications that inclusion of upper level humidity and stability data can improve rainfall downscaling. In order to build a consistent basis for constructing future scenarios, all GCM fields were interpolated to the 2.5 degree NCEP grid.

5.4.2 Common EOF Framework

An initial trial of the common EOF framework was undertaken to evaluate the approach for downscaling monthly temperature and rainfall. The training window was from 1980-99 and the seasonal component of both the predictand and predictor sets removed. For the monthly mean temperature the initial predictor evaluated was temperature at the 850hPa level. Common EOFs were derived and initially the leading 12 principal components retained. A screening approach known as stepwise regression was then used to evaluate the predictors by determining which improved the models accuracy. In these experiments the stepwise regression assumed no predictor term in the initial model. Predictor terms were included and retained if there was an increase in model precision at the $p<=0.05$ level (the “p to enter” rate). If the number of predictors was greater than 10 the stepwise regression was re-run at a $p \leq 0.01$ level.

Results from this approach are shown as scatter plots between observed and predicted data at six of the ten evaluation sites in Figure 25, illustrating a strong relationship with relatively low Residual Mean Square Error (RMSE) and high correlation coefficients (R^2). This distribution of residuals appears to be random, with no seasonal biases or truncation points. There is under-prediction of high summer and spring temperature anomalies. The number of predictors retained by the screening regression varied between sites, ranging from ten down to five. Importantly no independent validation was set up to verify the predictive skill of the modelled relationship at this stage.

Given the results of this initial trial the scheme was set up to run across the VCSN network, fitting to each grid point on an individual basis with no spatial dependency. Computationally this procedure was completed in around 45 minutes on a standard personal computer.

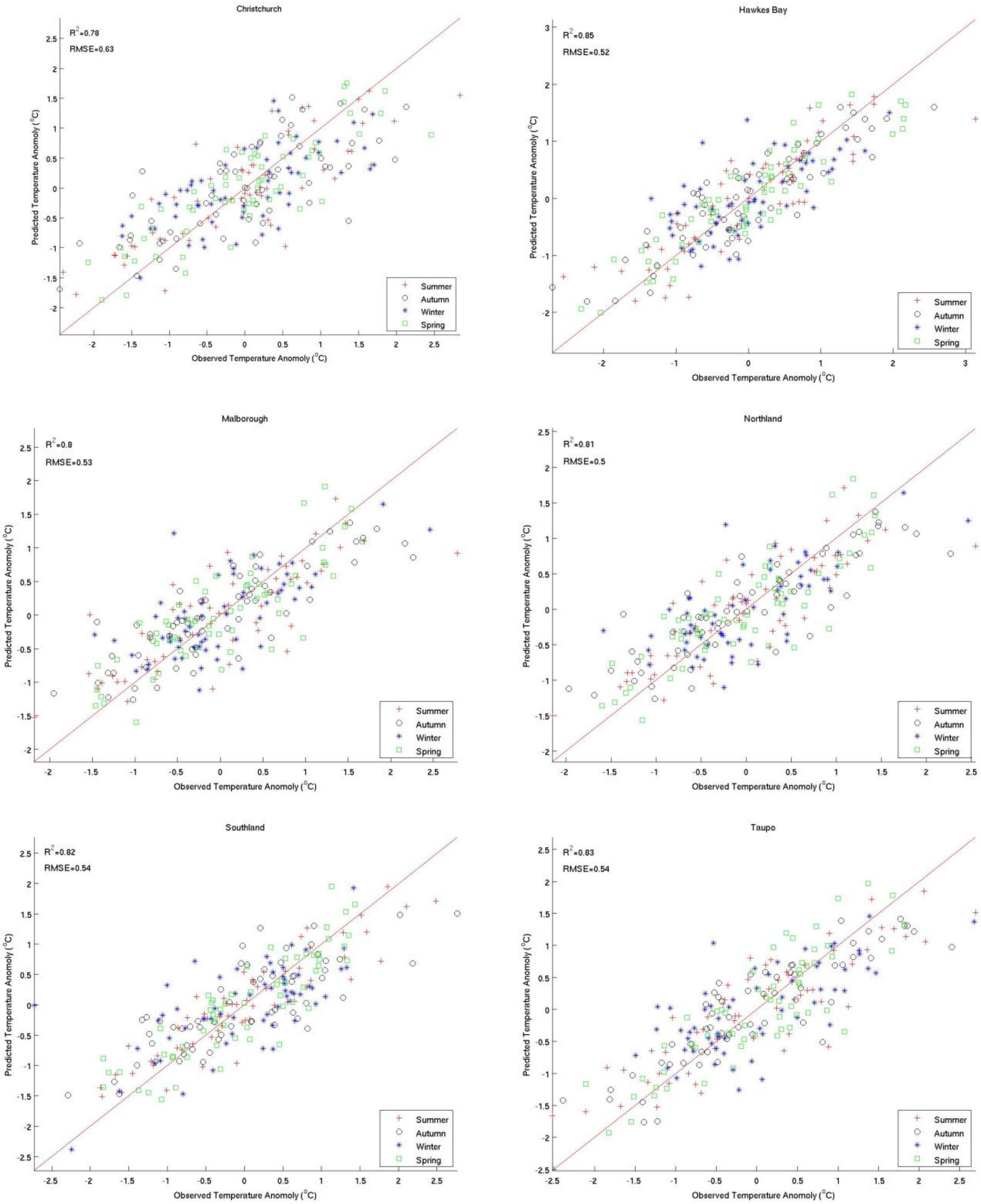


Figure 25: Scatter plots of observed and predicted temperature anomalies based on the initial trial of the common EOF downscaling methodology. .

The summary results are mapped in Figure 26 highlighting moderate to high levels of model precision on low lying (agricultural) regions in both Islands, with higher error on the mountains of the South Island. The regional distribution of error may be related to the observation model, as there is known increased error at higher altitudes due to a low site density. The stationarity test is summarised by the significance level of the linear regression fitted to the model residuals. This suggests that for the most part there are significant trends in the residuals (in this case mostly positive) and hence strong evidence of non-stationarity over the twenty year time frame.

An important result of this experiment is the map of the number of variables retained by the stepwise regression procedure. There are large sometimes abrupt changes across New Zealand to the number of variables retained at each site, and although there appears to be regional structure the changes are abrupt and do not follow a smooth transition nor have a clear physically basis in the spatial pattern. Without going to the step of resolving the scheme with actual future projections this is a clear indication of over-fitting in some regions. Instability of this type would lead to highly irregular and inconsistent projections at the regional level.

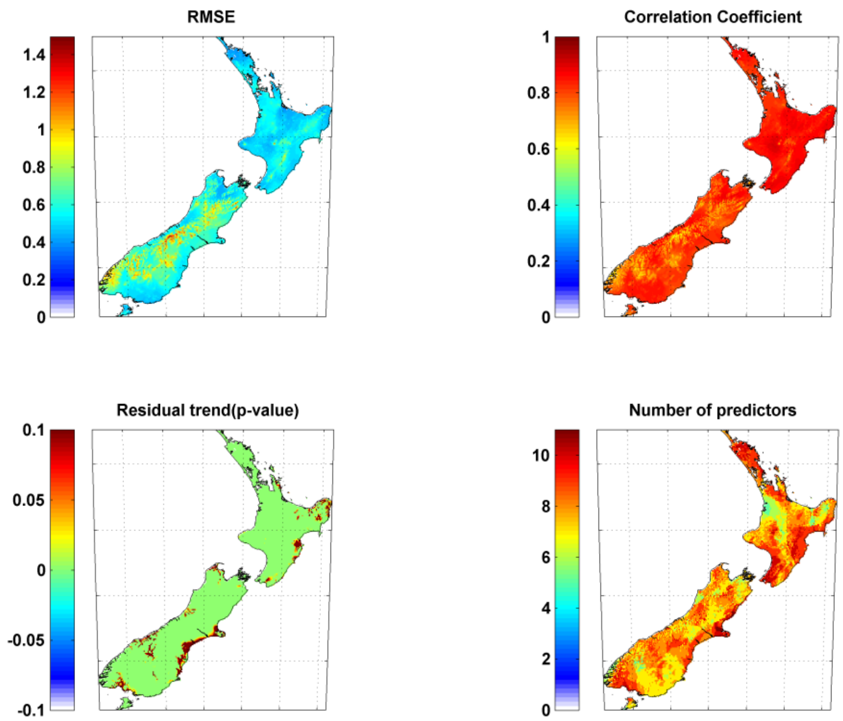


Figure 26:   National level results for downscaled surface temperature under the initial trial of the common EOF methodology.

The common EOF methodology was used to develop a downscaling scheme for monthly rainfall. The predictand and predictor data were transformed by taking the square root, consistent with many treatments of rainfall using statistical modelling procedures that rely on the assumption of normality. To reflect the main synoptic and local scale drivers of precipitation, candidate predictors included surface pressure (MSLP), geopotential height at

500 hPa, relative humidity at 850hPa and relative vorticity at 850 hPa. Further trials where also undertaken which restricted this predictor set to reduce the number of variables, but for the purpose of the initial experiment the full predictor set was used. Common EOFs of the predictors were taken and the leading four components retained, based on an assessment of the variance explained. The PCA is illustrated by the spatial patterns of the leading 6 EOFs, the first two principal components and the variance explained by each component for MSLP (Figure 27) and relative humidity at 850 hPa (Figure 28). The PCAs illustrate the expected patterns for New Zealand in that the leading three to four components illustrate the dominance of north-south and east-west flows in explaining the New Zealand regions variability. No terms were included in the initial screening model, and p-to-enter rate set to the p≤0.05 level. If the number of predictors was greater than 15 the stepwise regression was re-run at a p≤0.01 level.

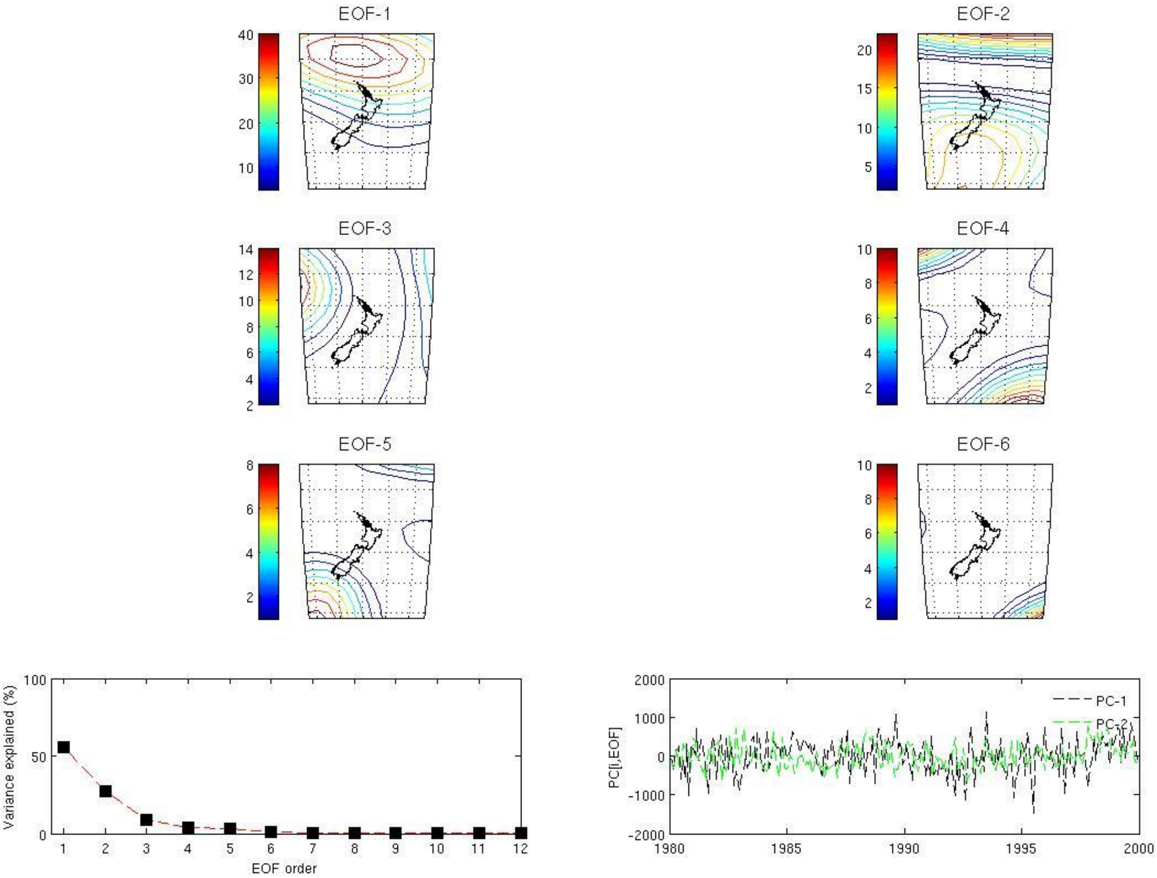


Figure 27: Principal component analysis of NCEP reanalysis sea level pressure (MSLP).

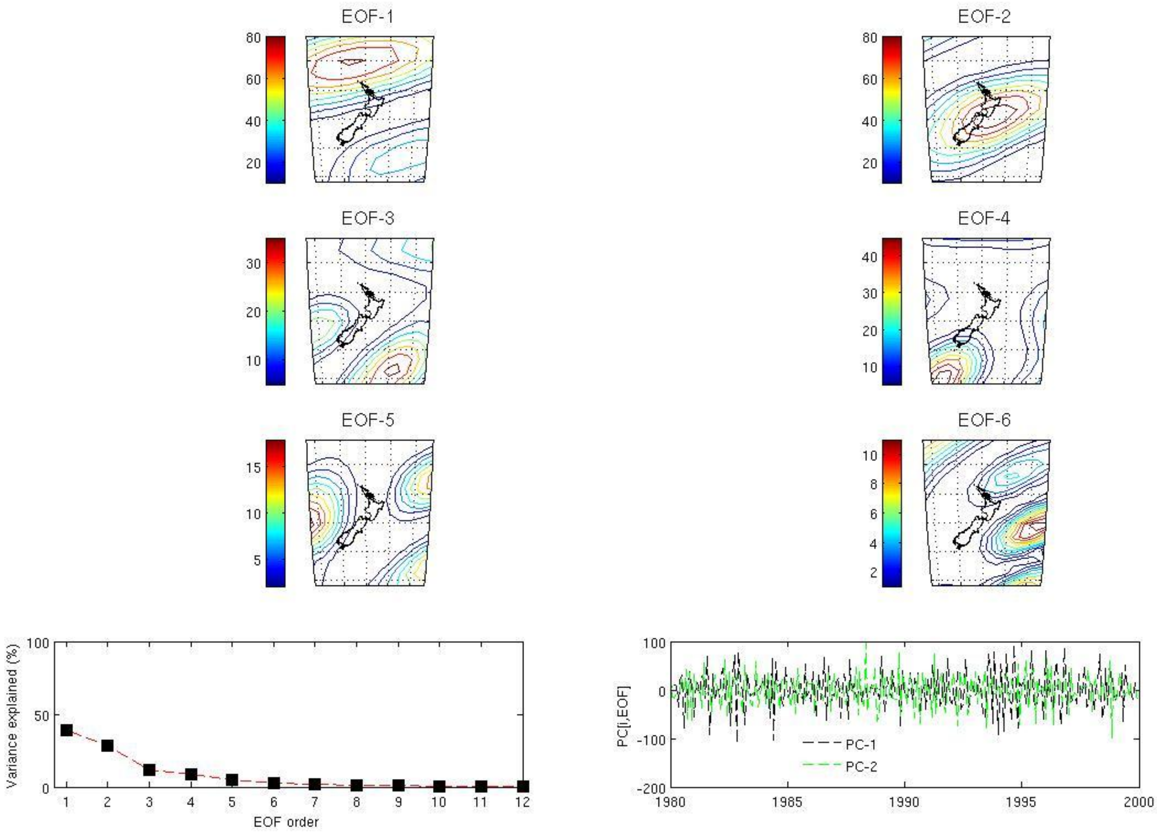


Figure 28: Principal component analysis of NCEP reanalysis humidity at 850 hPa.

The results for monthly rainfall at six of the 10 trial sites are shown in Figure 29. Similar to temperature the procedure yielded moderately accurate fits to observed rainfall at these sites, with relatively low RMSE and high correlation coefficients. The distribution of residuals appears to be random with no obvious seasonal biases or truncations. Like temperature, the number of variables retained across the sites varied considerably. The most consistent result was the inclusion of the leading components of MSLP in the model, while the inclusion or exclusion of terms from the other predictor variables seemed to follow no discernable pattern.

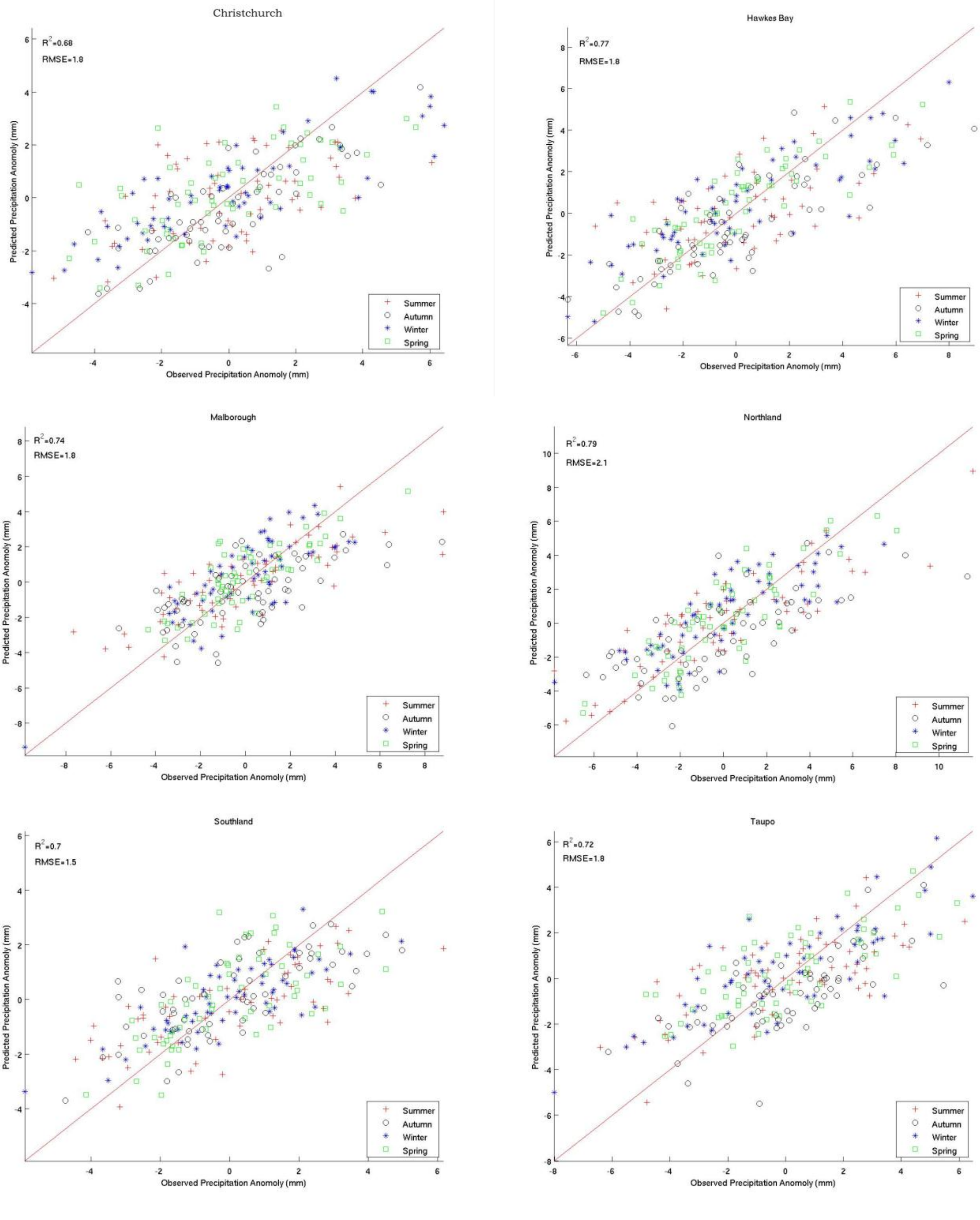


Figure 29: Scatter plots of observed and predicted rainfall anomalies based on the initial trial of the common EOF downscaling methodology.

When the procedure was run on an individual grid cell basis for the VCSN network promising levels of accuracy were obtained, with relatively low RMSE and correlation coefficients (Figure 30). Interestingly there was a different spatial structure to these estimates to that found for temperature, with low correlation coefficients in the central and coastal Otago region and higher errors across the alpine zone of the South Island. Unlike temperature, there was for the most part no significant trend in the residuals. However, there was significant instability in the underlying statistical model. Firstly, the number of predictors (>10) is very high. Secondly, the number of predictors retained changed almost randomly across the country with little evidence of a physical basis. Like the temperature model this instability would invalidate its general application in scenario development because there would be abrupt changes between individual grid cells in the future projections. Similar to the individual site studies the most common predictor retained was MSLP, with no discernible pattern in the selection of other predictors.

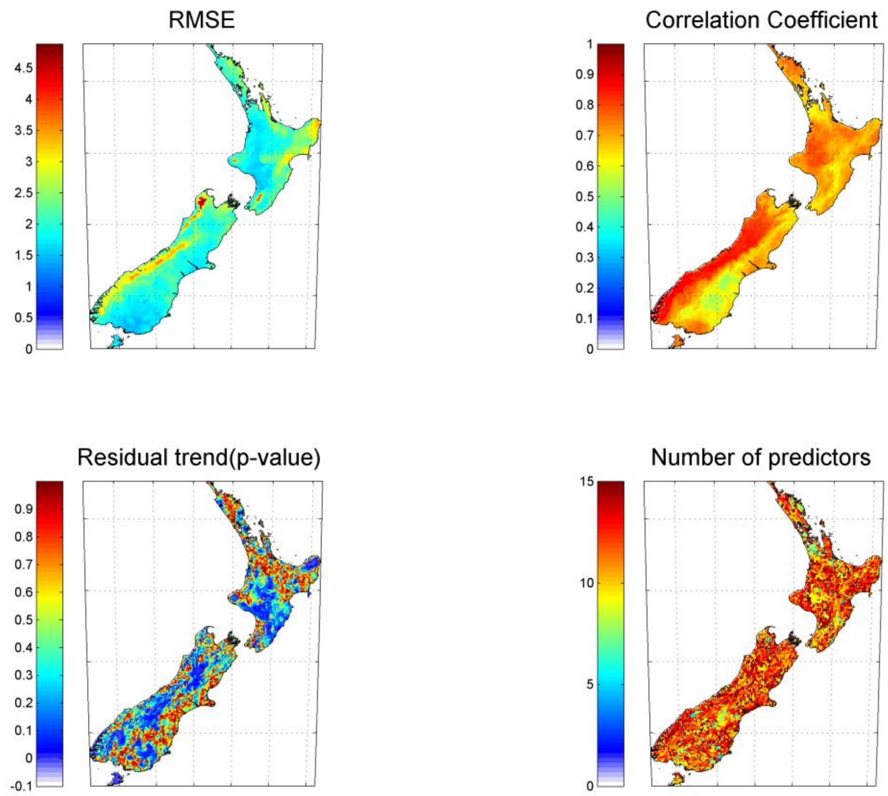


Figure 30: Initial results for monthly rainfall when applying the common EOF methodology. Note rainfall not back transformed.

While the accuracy measures of these initial trials of the common EOF framework appear encouraging, they did not yield a stable predictive model that could be satisfactorily used to project drought across the country. Although the results are not presented in this report further refinement of the common EOF approach was carried out beyond this initial trial. This included:

- entering fewer variables as candidates in the screening regression and differing combinations of variables;
- retaining fewer leading principal components for the initial model;
- using rotated rather than common EOFs;
- adjusting the p-to-enter rate to provide either a stricter or looser rejection criterion;
- using the leading three components of MSLP as the initial terms in the model;
- using some alternative predictors such as locally interpolated zonal and meridional (u and v) wind to calculate vorticity in a site specific way;
- different data transformations including the cube root for rainfall, detrending the series and not removing the seasonal component;
- trials of the procedure for other predictands including monthly radiation, PET and wind;
- reframing the VCSN-wide procedure as a multi-site problem, thereby building spatial dependency into the co-variance matrix of the regression.

In summary these alternatives, which included some combinations of the changes, either maintained or deteriorated the model precision estimates from those described previously. In all but one case they did not address the problem of model instability found in the original trial without significantly degrading the strength of the relationship. A stable and sufficiently accurate model was found for temperature based on MSLP and temperature at 850hPa, but the alterations for rainfall either deteriorated accuracy or introduced instability. Some researchers choose to ignore this problem, or address it by building a much more regionally constrained ESD scheme by creating distinct zones where different predictor sets are used (e.g Timbal et al. 2009). This would be an unsatisfactory avenue to pursue in this study which seeks to establish a consistent regional to national basis for the analysis of changes to drought probabilities, including the identification of spatial gradients in the regional patterns of change. Therefore an alternative means of addressing this problem was explored.

5.4.3 Statistical models

Partial least squares (PLS) regression was trialled as an alternative statistical procedure to address the problem of model instability. In contrast to the common EOF methodology, PLS regression is a two block modelling scheme which simultaneously optimises the dimension reduction of the predictor space and the transfer functions. This is achieved by developing new functions for the predictors which are linear combinations of the original high dimension predictor space. This makes PLS regression useful for high dimensional problems like downscaling climate where there is a high degree of spatial and temporal correlation in the predictor fields.

Technically this is achieved by establishing geographic (X-scores) and temporal (L-loadings) weights that maximise the co-variance of the predictors and minimise the error of the final transfer functions. While this is analogous to taking an EOF, the dimension reduction phase is optimised so there are fewer a-priori choices, such as the EOF components for each predictor variable to enter into the model and the p-to-enter ratios. In PLS regression the only information required is the number of individual predictor components to retain in the dimension reduction. This can be investigated by undertaking an exploratory analysis to evaluate the behaviour of the scheme over a large number of components, then re-running the procedure with a refined number to balance complexity with the strength of the relationship. The application of PLS regression in fields outside climate downscaling has been shown to make modest to large improvements to the predictive skill of linear models when predictors are highly correlated (Stone and Brooks, 1990).

Given the successful use of the common EOF framework in building a temperature model but identification of some problems with rainfall, monthly rainfall was the focus of the initial trial of PLS regression. In these experiments evaluation of the predictive strength and model stability is achieved by both independent testing and cross validation. The cross validation partitions the data into a training and independent validation set of equal length, by taking continuous but randomly located 4 year periods from the 1980-2009 time frame without replacement. The partitioning and model fitting was repeated 50 times and the statistics of model accuracy averaged at each of the 10 sites. In these trials the PLS regressions retained 5 model components. This was compared to a range of common EOF based models, where different predictor variables were paired with sea level pressure and either 4 or six components retained. These represent the best rainfall models found in the initial experiments described above.

Table 6: Comparison of PLS regression and the common EOF framework (step wise regression) for monthly rainfall.

	Training set (1980-1999)		Cross validation		Independent test (2000-2009)	
Experiment	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**
Stepwise regression						
SLP: PC(4)	2. 4	0. 53	2. 5	0. 49	2. 7	0. 42
SLP: PC(6)	2. 2	0. 62	2. 3	0. 55	3. 2	0. 22
SLP: PC(4), Rh_{850}:PC(4)	2. 3	0. 58	2. 5	0. 48	2. 8	0. 33
SLP: PC(4), Hgt_{500}:PC(4)	2. 2	0. 59	2. 4	0. 51	3. 7	0. 31
SLP: PC(4), RV_{700}:PC(4)	2. 3	0. 58	2. 5	0. 49	2. 9	0. 27
SLP: PC(6), RV_{700}:PC(4)	2. 1	0. 63	2. 3	0. 54	2. 7	0. 48
SLP: PC(6), Ta_{850}:PC(4)	*2. 3*	*0. 57*	2. 4	0. 49	3. 3	0. 37
Partial least squares regression						
SLP	2. 0	0. 69	2. 4	0. 55	2. 4	0. 57
SLP,RV700	2. 0	0. 68	2. 5	0. 59	2. 4	0. 57
SLP,Rh850	1. 9	0. 74	2. 2	0. 64	2. 0	0. 62

Summary results for the monthly rainfall experiment. Data are mean diagnostics for the 10 test sites. Variables SLP is sea level pressure ; Rh850 is relative humidity at 850 hPa; Hgt500 is geopotential height at 500 hPa; RV700 is relative vorticity at 700 hPa. PC (n) is the n number of principal components retained in the prediction scheme.

The results of these experiments are summarised as the mean of the model accuracy scores across the 10 trial sites (Table 6). This illustrates that the PLS regression method yielded a generally stronger prediction relationship that was relatively stable when compared to those developed with the common EOF framework. Importantly higher model diagnostics were found for the PLS downscaling models in both the independent testing and cross validation, where there was large degradation in all the models developed using the EOF approach. The number of variables selected by the stepwise regression varied between sites from 5 through to 10 again with no apparent systematic pattern in the selection. This suggests an overall improvement to the predictive skill of the PLS developed models, and seems to address some of the predictor instability found in the common EOF approach.

5.4.4 National implementation of PLS regression

Given this result, the PLS procedure was set up for the VCSN grid with temperature and monthly rainfall as predictands. The regression was solved as a multi-site problem thereby including spatial dependence in the model fitting. Generally the PLS regressions took around one to two minutes CPU time on a standard unix mainframe to optimise for the entire country. To evaluate independent prediction skill, the training window was set at 20 years (1980-1999) with the remaining nine years (2000-2009) used as the independent set.

For both rainfall and temperature the procedure was run initially retaining 20 components and the behaviour of the model assessed. The final models were developed by re-running the procedure where inclusion of further components did not improve model performance. An example of model behaviour analysis is shown for the monthly temperature procedure in Figure 31. In this case the mean square error (MSE) of the response was reduced substantially after inclusion of only 4 components, with further but smaller MSE reductions out to 10 components. Adding more components did not further reduce response MSE. For the predictor function MSE reduced considerably at 10 components, and fell further at a smaller rate out to 20 components. The overall variance explained by the model rose sharply to over 50 percent with the inclusion of 5 components through to 60 percent at 10 components. The variance explained continued to rise to well over 70 percent at 20 components.

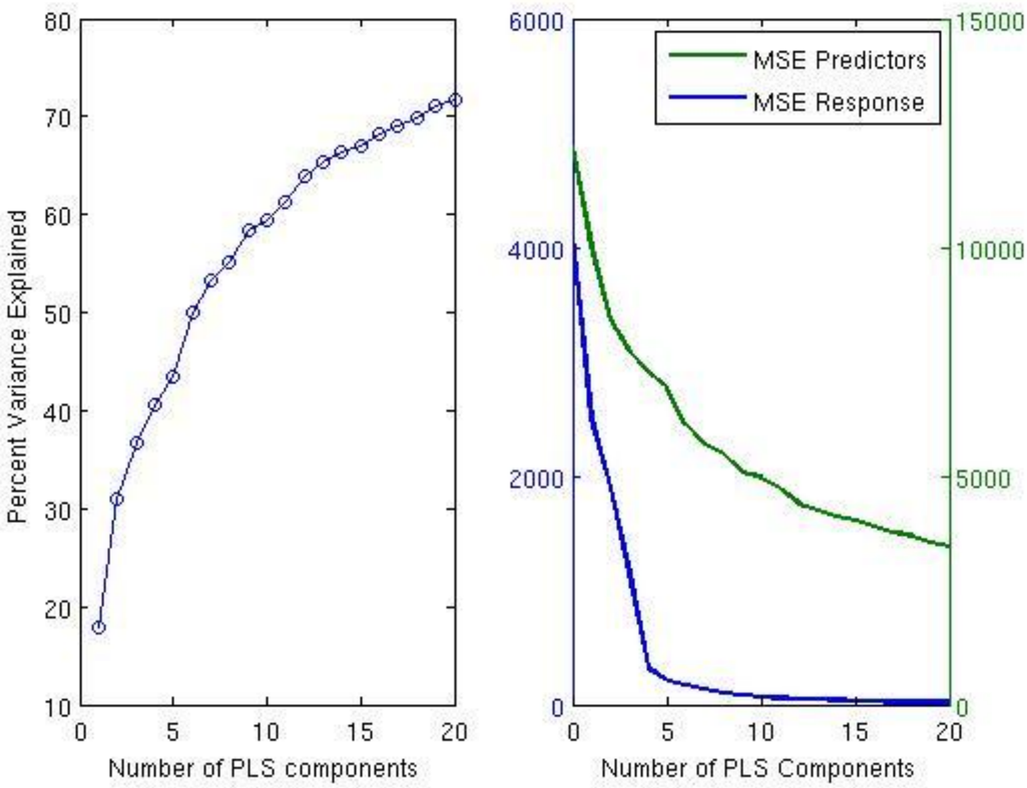


Figure 31: Model behaviour analysis for PLS regression of monthly temperature.

At this point the analyst is faced with the usual trade-off between model complexity and precision, where the model diagnostics help to isolate where this trade off occurs. An overtly complex model while explaining a higher proportion of the variability runs the risk of over fitting and basing the downscaling on an empirically convenient but physically unsound relationship. A too simple model may not be precise enough for a given application leading to reduced usefulness. A conservative approach was adopted at this stage, with preference toward a simpler model sometimes at the expense of precision. The diagnostics in Figure 31 also aggregate spatial variability in model precision to the national level, so the model may be performing better or worse across the agricultural land of interest in this study. A further level of diagnostic analysis is required to choose a reasonable final model, where the residuals and the spatial distribution of errors are examined thoroughly.

In the temperature example described, the PLS procedure was re-run at the four, five and six component levels and the diagnostics of the model assessed with residual and spatial diagnostics. Figure 32 is presented as an example of the type of residual diagnostic analysis undertaken in all subsequent model development. In this case it is for monthly temperature with 4 PLS components retained, using MSLP and temperature at 850 hPa as predictors. In this case only the independent evaluation set is shown. The scatter plot is a 500 grid point sample of the potential 11492 VCSN grid points constrained to agricultural land only. There is a positive and reasonably accurate relationship between observed and downscaled temperature anomalies for this sample of sites, with summary correlation coefficient of 0.69 and RMSE of 0.7 °C The second graph is a qq-plot where the blue line is a rank ordering of residuals according to their position on the observed distribution and the red line is the theoretical standard normal distribution of the models residuals. This shows that the models residuals provide a good estimate of normal distribution except for over prediction at high extremes and under prediction at low extremes. This is indicative of a well founded model with few systematic biases, good evidence that it is based on a plausible physical relationship.

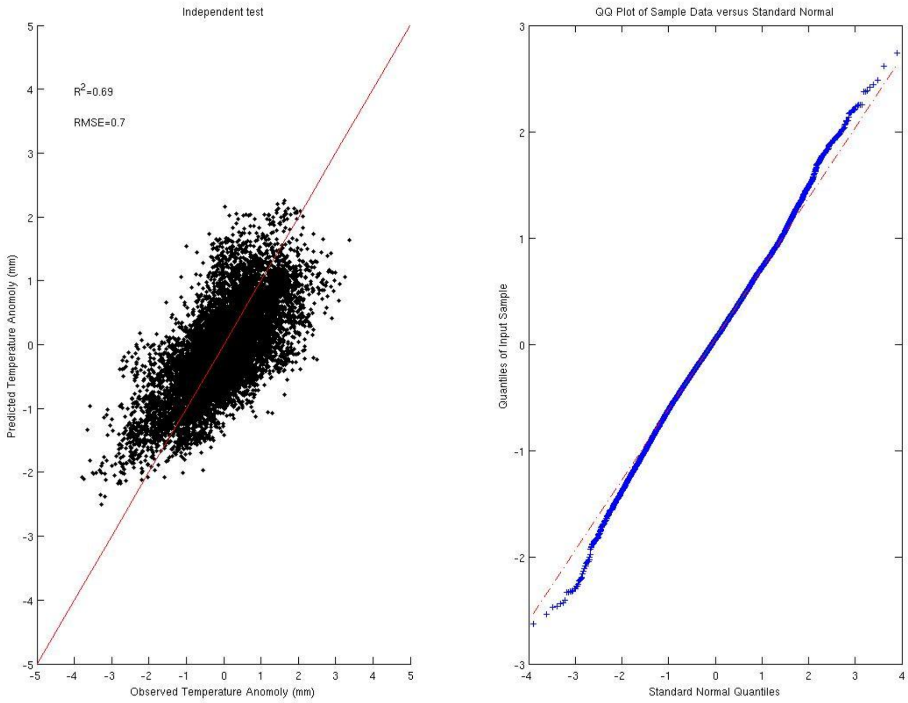


Figure 32: Example of residual diagnostics: initial PLS results for surface temperature anomaly using MSLP and temperature at 850 hPa as predictors. 4 components retained.

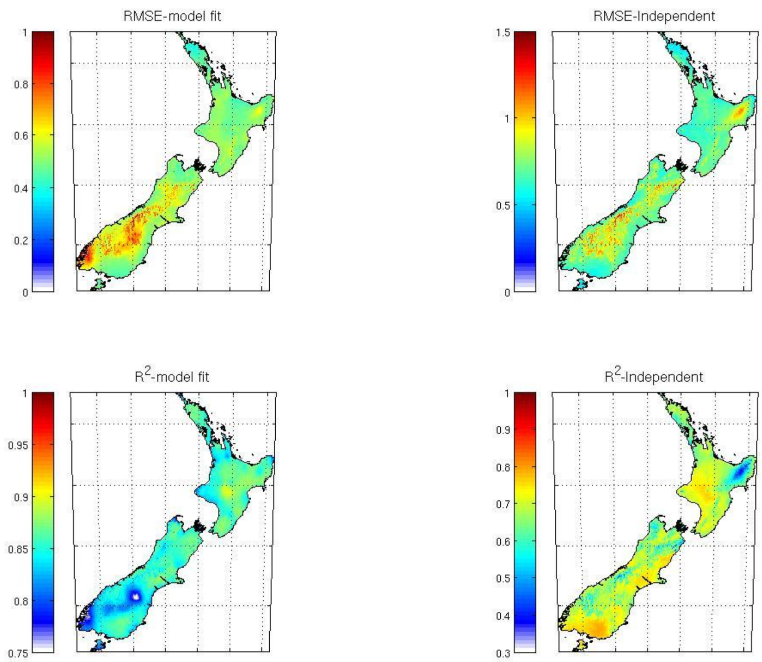


Figure 33: Spatial diagnostics: initial PLS results for surface temperature using SLP and temperature at 850 hPa as predictors. Number of components retained was 4.

An example of the spatial diagnostic analysis undertaken in all subsequent model development is shown in Figure 33. Here the temperature downscaling models precision scores are mapped across the country for both the independent and training period ('model fit' in this figure). These show that there are discernable spatial patterns in the model diagnostics, where generally reasonable results are found across most of the agricultural land: RMSE generally below 1 °C; and correlation coefficients between 0.6-0.7. The increased error in the Southern Alps found with the EOF approach persisted with this model, re-enforcing that the modelling process is sensitive to the quality and comprehensiveness of the observation system. There are also some other features, with a distinct area on the north east cape of the North Island where the model consistently performed poorly. This is an another location where there is a paucity of climate monitoring sites in New Zealand's national climate data base.

To maintain brevity the diagnostic results of these initial trials of PLS regression for monthly rainfall are not presented in any detail. In summary reasonable results were found using MSLP and humidity at 850 hPa as predictors and retaining six PLS components. Like the site level results, encouraging outcomes for monthly rainfall were obtained with independent verification scores around 20mm for RMSE with a 0.61 correlation coefficient on a sample of agricultural land. The residuals were normally distributed, but there were more severe biases at the extremes of the distribution than reported in the initial model for temperature. The final PLS based downscaling model developed for rainfall is described in more detail below, but prior to this a another experiment is briefly described where the geographic bounds of the predictor space are examined.

5.4.5 Geographic bounds of predictors

One of the choices need in both RCM and ESD downscaling schemes are the geographic properties of the predictor space. For the initial experiments described above this was assumed to be a rectangular box around New Zealand spanning 25 to 65 degrees south and 150 to 200 degrees east. Given the relative computational efficiency of the multi-site PLS regressions in developing national scale models it was decided to investigate the influence of the geographic bounds of the predictor fields on the final temperature model (described below).

A centroid for New Zealand was chosen at approximately the coast just north of Christchurch. The smallest bounding box was a field running plus and minus 7.5 degrees in longitude and latitude from this centroid, denoted the '0km boundary'. Successively larger fields were created by uniformly expanding the coordinates of this grid box by 5000 km increments. The PLS regression model procedure was run for each successively larger boundary based on variables from the NCEP re-analysis.

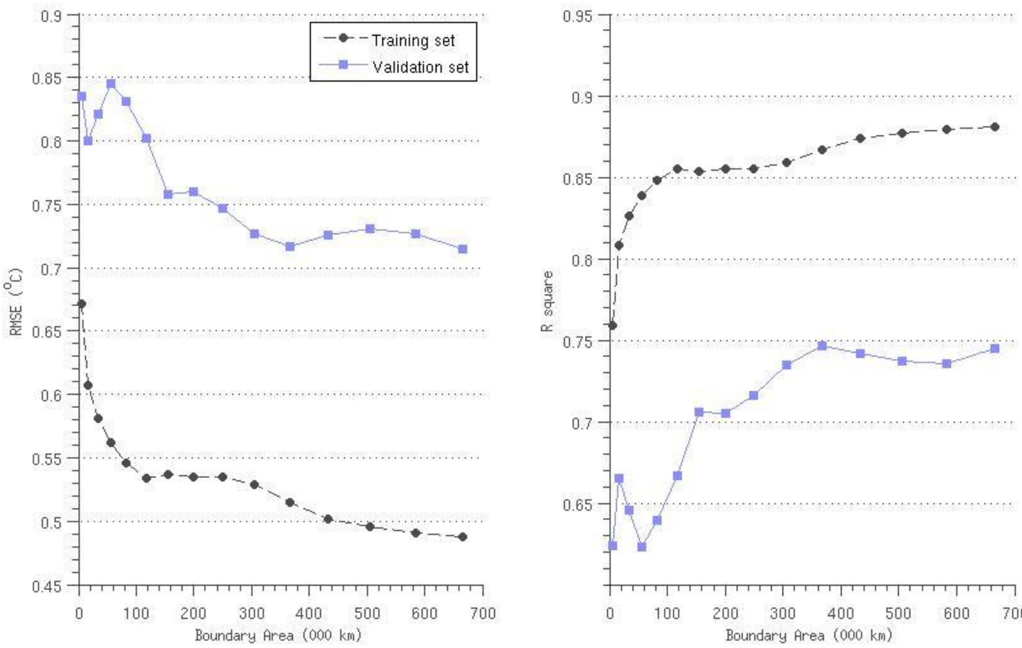


Figure 34: Evaluation of the geographic extent of the PLS regression predictor space. Experiment for temperature using sea level pressure and surface temperature.

The results of this experiment are summarised in Figure 34 as nationally average model diagnostics plotted against the enlarged boundary area above the minimum box. This shows that there is a continual improvement in model diagnostics out to the 100000 km^2 boundary area, where the rate of improvement seems to decline. Close inspection of the plots reveals a slight bi-modal pattern in the results—that is a peak in precision at about 300000 km^2 area then a small decline in precision with an seeming small increase again at km^2 kilometre area. This is most obvious in the correlation coefficient score for the validation period.

Although there was a gradual pattern of change in the experiment, it did not reveal distinct optima above the initial improvement at the 100000 kilometre area whereby a clear best candidate zone for the ESD scheme could be selected. This could be due to the sensitivity of the model diagnostics used, and or the behaviour of the PLS regression in reducing the dimensionality of the predictor space. Given this result the issue of computational efficiency became the predominant factor in guiding the selection of the boundary, so a relatively fine box was used for all remaining model development 25 to 55 degrees south and 160 to 190 degrees east.

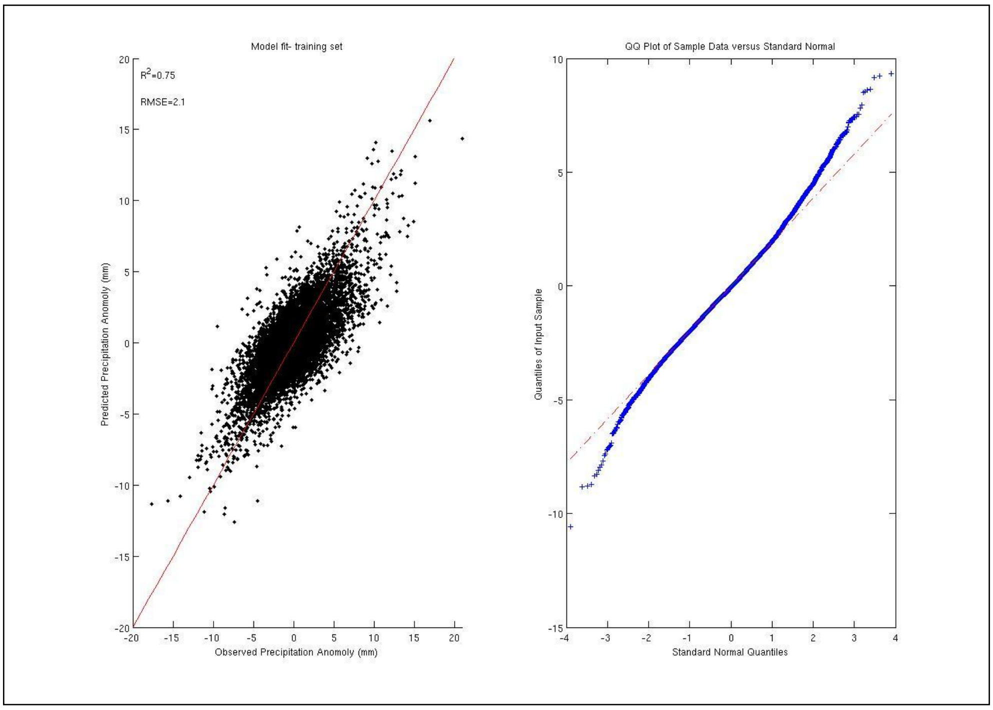
While this experiment was successful it was decided to not pursue this line of reasoning further at this stage. This was a pragmatic decision related to resources and timeframe available for this particular study. This does not preclude further refinements to this aspect of ESD schemes in the future, such as: investigating alternatives for different predictands; changing the model's complexity; and changing the proportions of the bounding box or even the rectangular shape.

The next sections summarise the development of the monthly rainfall and potential evaporation downscaling models used in the applied analysis. A brief summary of the PLS regression models applied to other variables is also given, although these models have no further application in this study. Finally a short experiment investigating the utility of the approach for downscaling daily fields is detailed.

5.4.6 Monthly Precipitation

The mid to lower atmospheric predictors used in the initial rainfall and temperature models are not comprehensively available for all the AR4 climate models. Hence the final ESD scheme used some alternative predictors than used in previous trials, so as to support development of a larger downscaled model ensemble in the final drought analysis. Although some alternatives were examined, the final rainfall model used MSLP and the surface precipitation field as predictors. A key observation made during this entire model development process was that MSLP appeared to be the most useful predictor for all variables and on its own explained a large proportion of the variance, which is not surprising given its general usefulness in synoptic diagnosis and known accuracy. The inclusion of surface precipitation did make a small but detectable improvement to the ESD scheme. Based on the evaluation process three components are retained in the final derivation of the PLS regressions.

The residual based model diagnostics for the final monthly rainfall model are shown in Figure 35. This highlights a strong relationship between the observed and downscaled rainfall for the training set, and although the strength of the relationship degraded the relationship appears robust when tested on the independent period. The qq-plots illustrated a good approximation of the theoretical normal distribution in both the training and independent period, but with a degree of bias in the downscaling of the tails. Similarly this bias is stronger in the independent testing period, although the model still has normally distributed residuals.



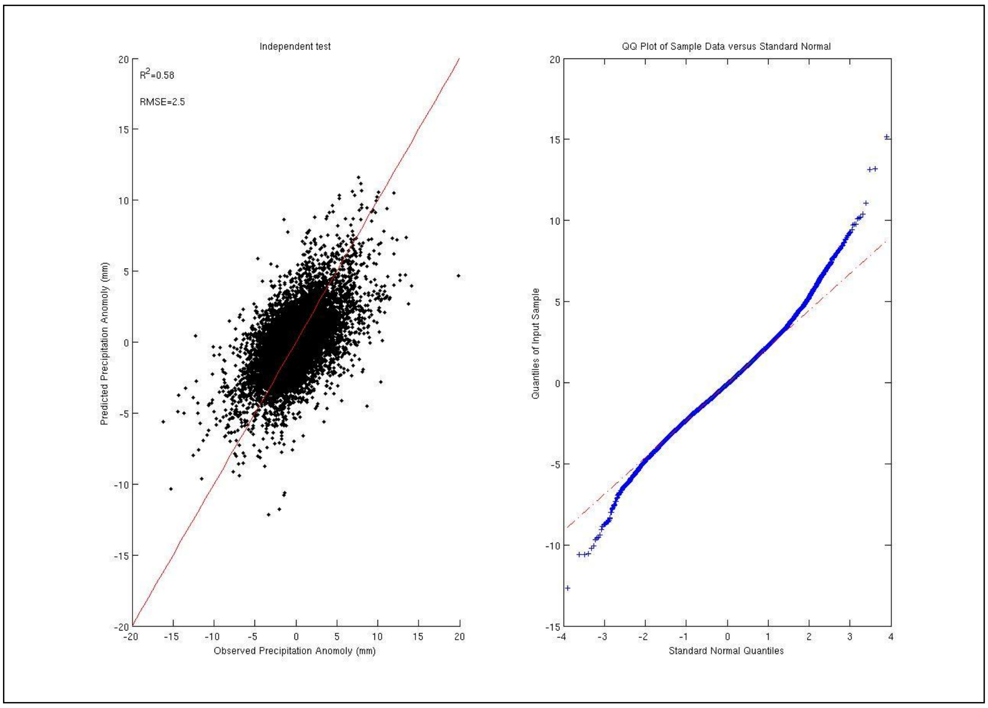


Figure 35: Residual based diagnostics for the final monthly rainfall model.

.

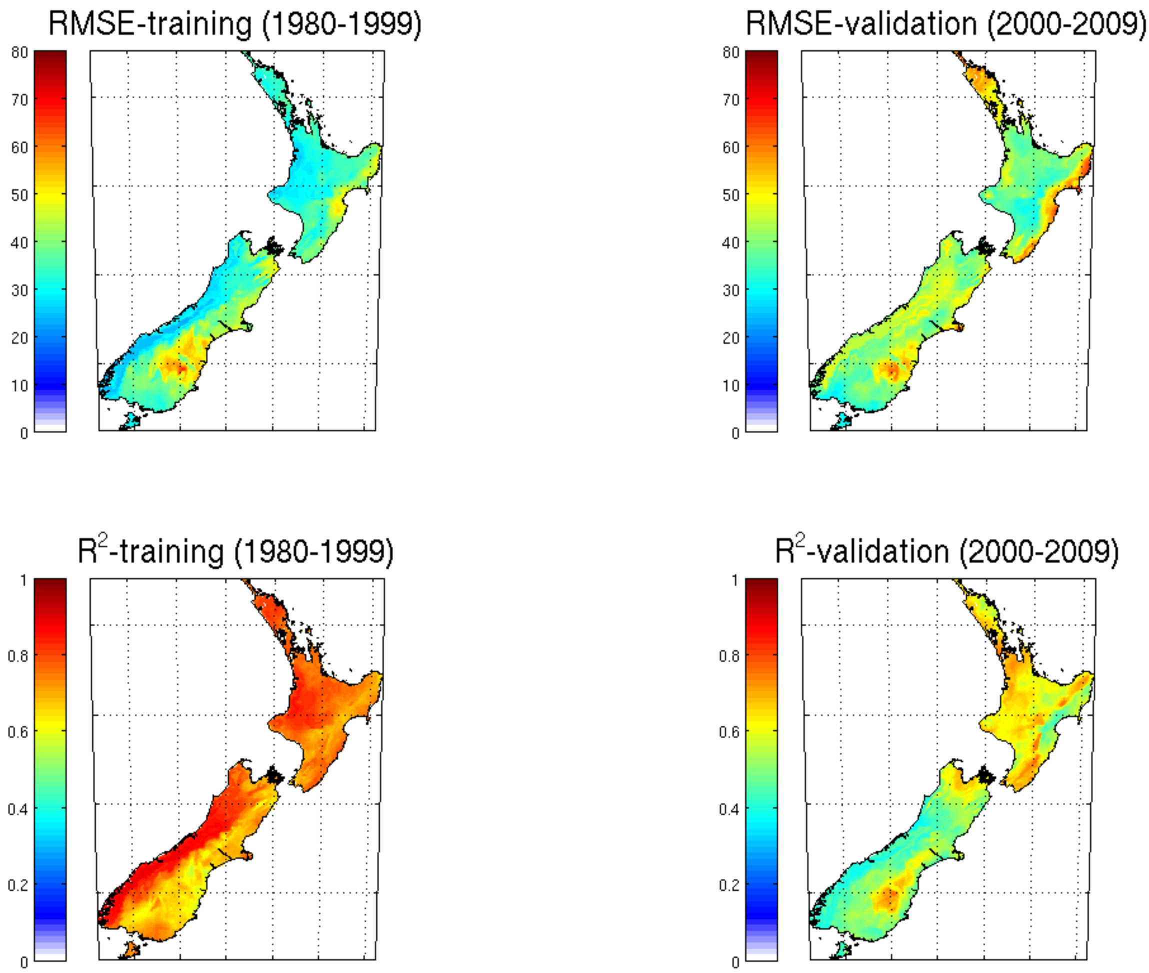


Figure 36: Spatial diagnostics for the final monthly rainfall model. .

The spatial diagnostics in Figure 36 identify a distribution of error that reflects some of the known properties of the New Zealand precipitation regime. Principally the Southern Alps and particularly high altitude zone on the West Coast of the South Island have a distinct error structure when compared to the rest of the country. Generally the RMSE across the agricultural landscape was in the range of 20-30mm (note this has been back transformed), although there are two zones where the errors are higher: the east and north coast of the North Island particularly in the validation set and the lower Otago (Timaru) region of the South Island. The results are not highly precise—with for example the correlation coefficient ranging between 0.6-0.7 at best for the validation set—but they do suggest that the model is not over-fitted and based on a signal that can be used for producing scenarios within sensible consideration of these errors. This level of model precision is consistent with results obtained in many other modelling of precipitation.

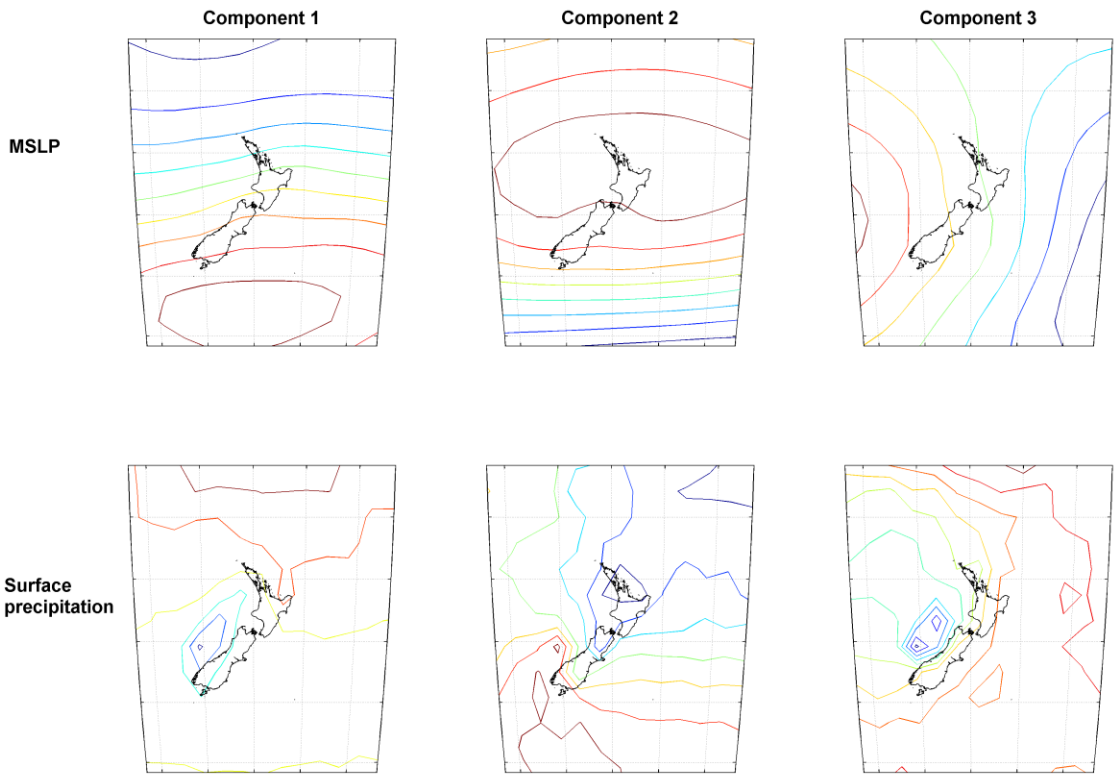


Figure 37: PLS predictor weighting for the final monthly rainfall model. Negative weightings grade to a deeper shade of red. Positive weightings grade to a deeper shade of blue.

To further evaluate the physical basis of the rainfall model the PLS X-weights were mapped across the predictor space (Figure 37). This is analogous to interpreting the spatial EOFs in a principal component analysis. The geographic distribution of weights highlights that the dominant aspect of the model are the north-south gradients of the pressure fields (components 1-2 of MSLP) followed by the east-west gradient (component 3 of MSLP). Although they are not identical, this type of spatial pattern is consistent with the main three Kidson Synoptic types for New Zealand, blocking, trough and zonal systems (Kidson 2000, Kidson and Renwick 2002). For the precipitation field a strong positive weighting occurs for rainfall events centred off the West Coast of the South Island (component 1 and component three), with the second component suggesting there is strong positive influence on the final model with rainfall events that have a northern sub-tropical origin. In aggregate these spatial patterns are broadly consistent with the known major influences on New Zealand rainfall regime, providing increased confidence that the model has a sound physical basis.

5.4.7 Monthly potential evapotranspiration

The flux of water from the land surface to the lower atmosphere is an important factor in the hydrological cycle and particularly in governing the behaviour of drought beyond the actual precipitation regime. Potential evapotranspiration (PET) is the climatically determined upper limit of water flux from the land surface to the atmosphere, a function of air temperature and radiation modified by relative humidity and wind. Using PET as the climatic upper limit is one way to examine land-atmosphere water flux. In some applications PET is not used, but the processes of bare soil evaporation and plant transpiration estimated given a physical energy budget and vapour pressure deficit driven models.

In climate impact analysis, downscaling of PET from broad scale fields is a special case as it is not routinely undertaken. There has been considerably more research focus on downscaling temperature and precipitation and more recently work on downscaling of precipitation extremes. Aside from direct downscaling of PET from broad scale atmospheric variables it can also be estimated in climate change studies indirectly, by downscaling other variables like temperature, radiation and wind. These variables are then used to drive a numeric calculation to PET, for example Penman, Penman-Monteith, Priestley-Taylor and Hargreaves PET formulations (Weiss and Menzel 2008).

In this study the direct downscaling of monthly PET is explored. The observation model supporting the downscaling is a version of Penman PET, where site based solutions to the equation are interpolated across New Zealand. To provide additional sites adjusted Priestley-Taylor PET estimates are also used. This is because there is low number of sites where all the variables are available for Penman PET calculation in New Zealand. This procedure is described by Tait and Woods (2009).

Following some initial experiments two PET downscaling models were developed:

- PET method A: where broad scale MSLP and surface temperature are used with three PLS components retained. This is described here as the “temperature driven model”;
- PET method B: where broad scale MSLP, surface temperature and incoming radiation at the surface are used, with three PLS components retained. This can be described the “radiation model”.

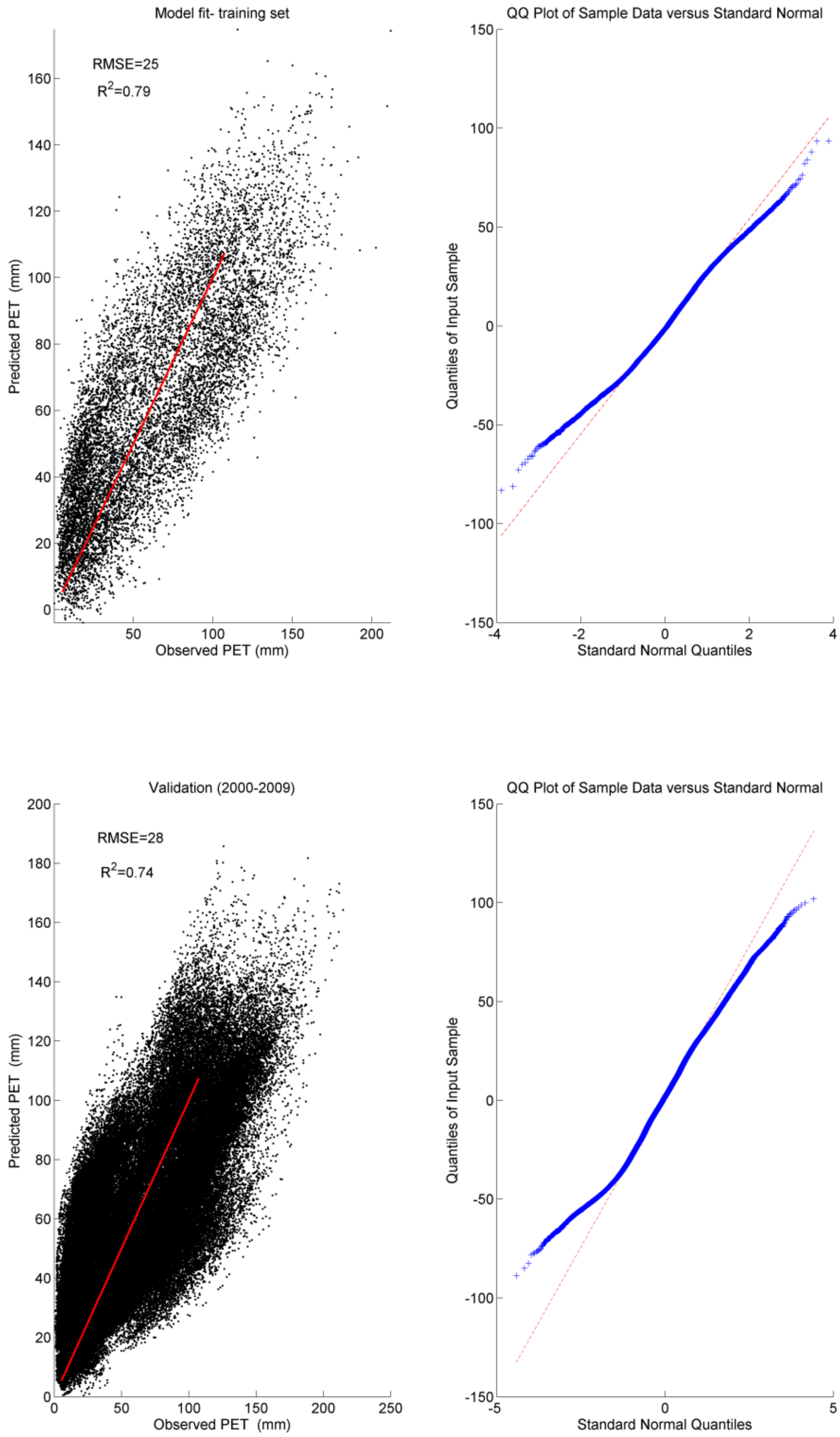


Figure 38: Residual analysis of PET downscaling method A which has MSLP and surface temperature as predictors.

The results for PET method A are presented in Figure 38 and Figure 39. The residual analysis reveals a strong downscaling relationship, with correlation coefficients of 0.79 and

0.74 for the training and independent testing periods. There was no strong geographic distribution in the correlation coefficient, apart from a slight degradation in the alpine regions. RMSE of around 20-30 mm were found, again with no major geographic differences, although a slight increase in error in the lower half of the South Island was evident. The residuals provided a good approximation of the normal distribution with some deviation around the extremes. Given these results the temperature driven model appears to be an acceptable form for use.

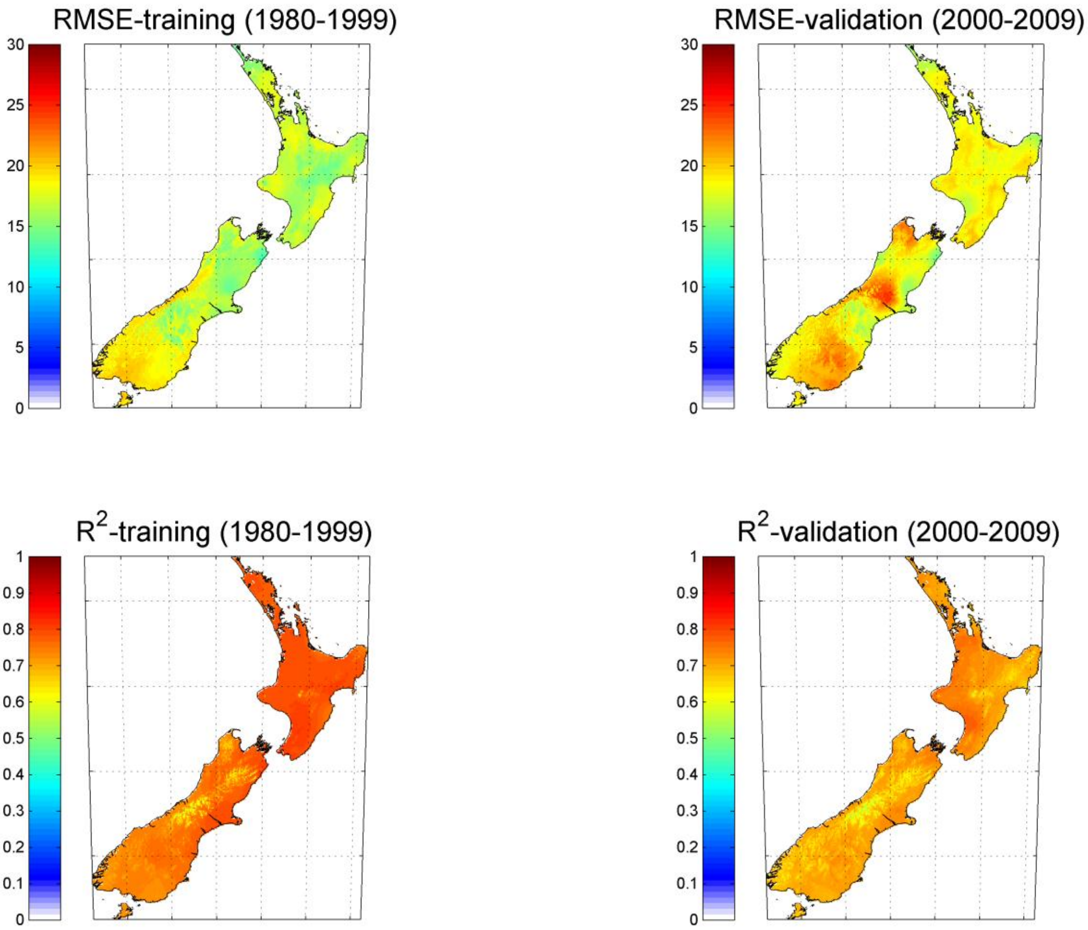


Figure 39: Model diagnostics for PET downscaling method A.

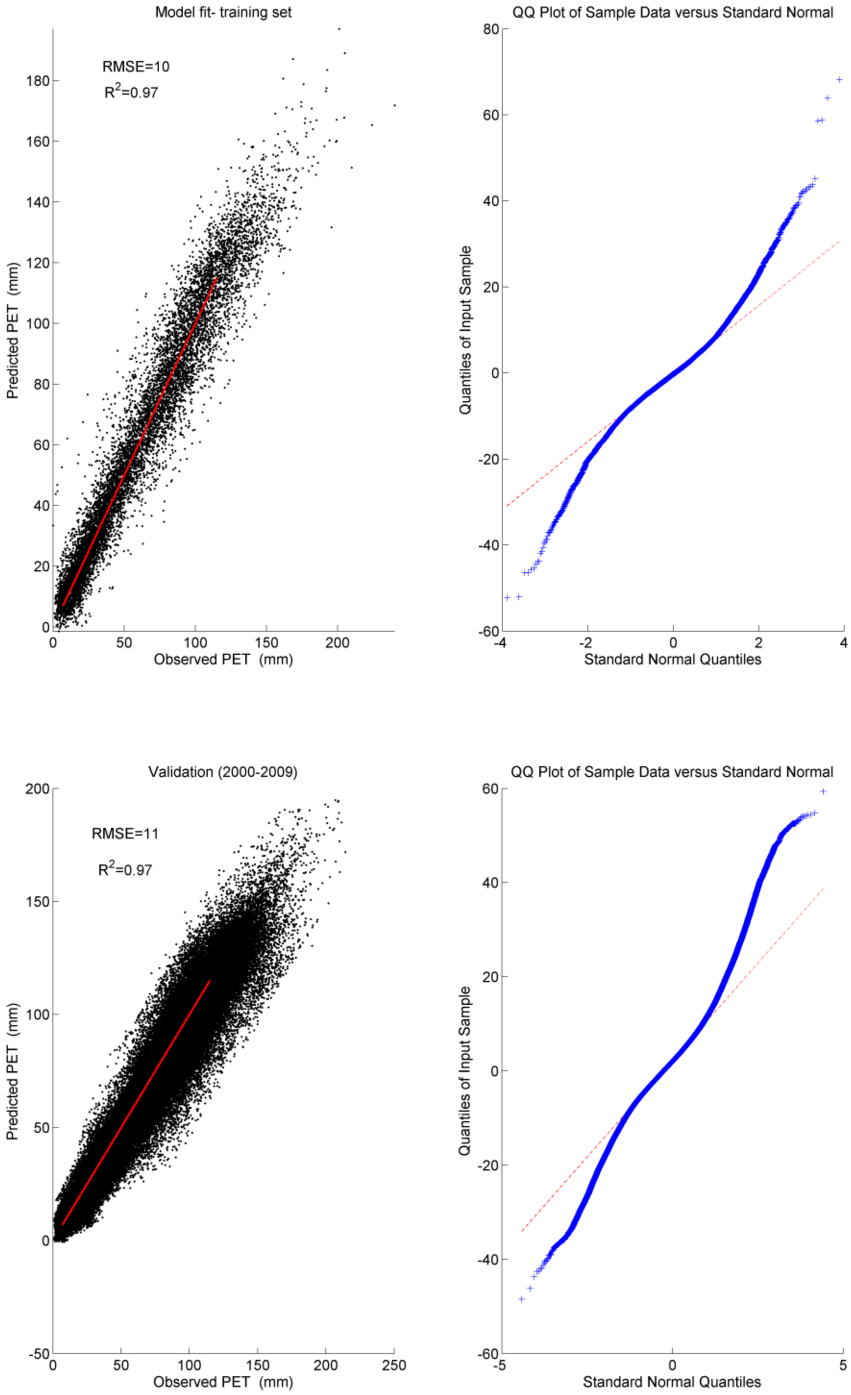


Figure 40: Residual analysis for PET method B which has MSLP surface temperature and surface radiation as predictors.

The results of PET method B are in Figure 40 and Figure 41. Inclusion of radiation resulted in an improvement to the downscaling model performance. Generally the correlation coefficients were above 0.95 for both the training and independent window uniformly across the country, providing a very strong relationship. The RMSE values were low in the order of 10mm on agricultural land. A slight increase in RMSE was evident in the alpine regions, but the increased error in Southland found in method A was not present for Method B. The theoretical normal distribution of residuals was not as well approximated for method B compared to method A. Despite the better performing distribution of residuals, the radiation method (B) appeared to have a stronger under-estimation of high evaporative events compared to the temperature method (A).

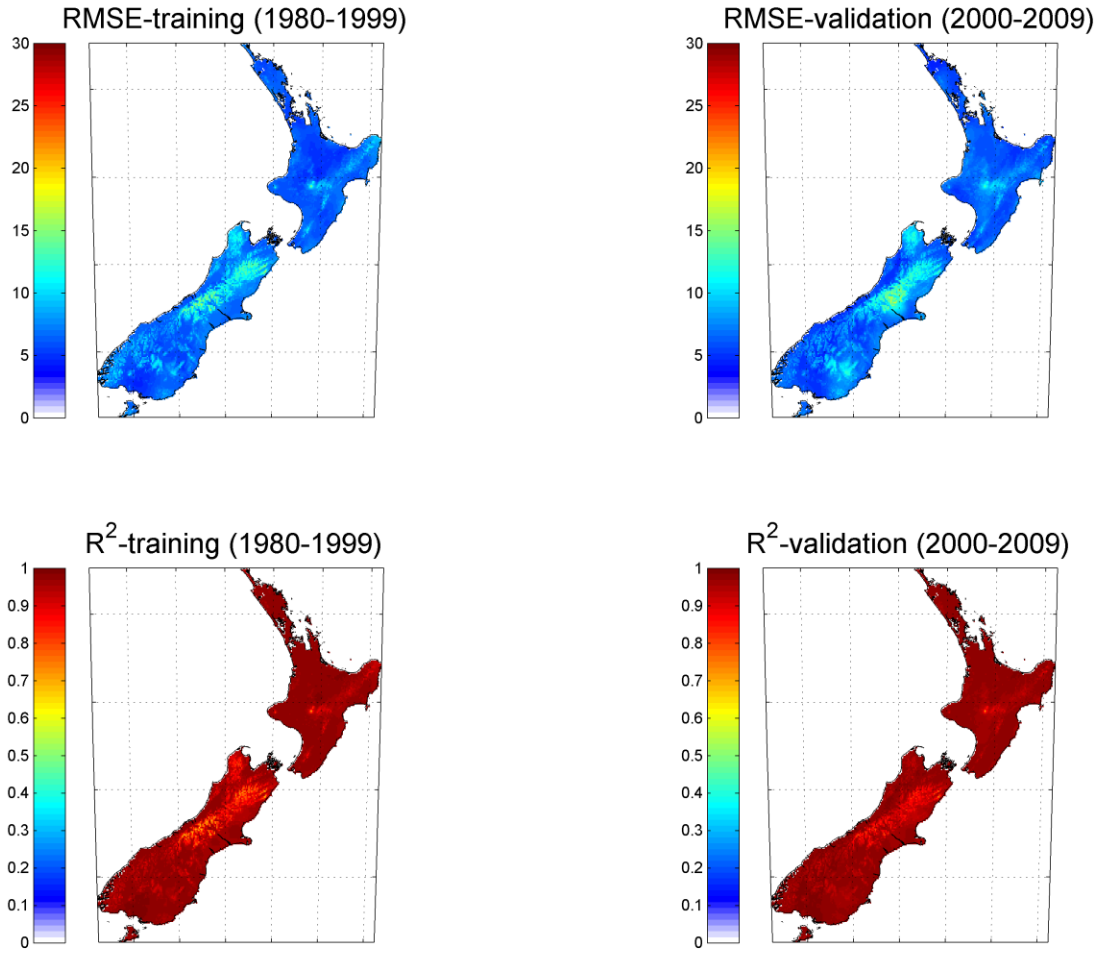


Figure 41: Model diagnostics for PET downscaling method B.

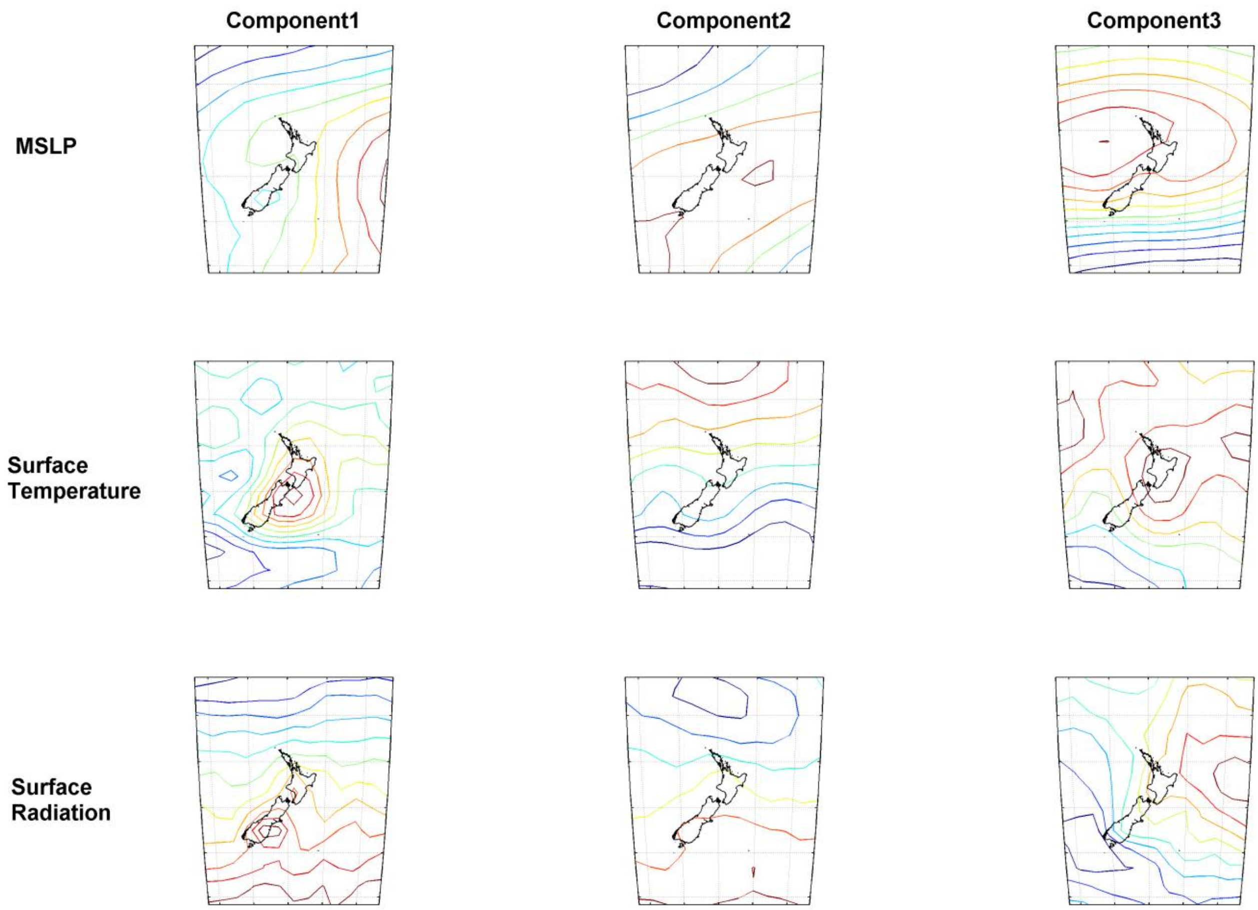


Figure 42: PLS predictor weighting for PET downscaling method B. Negative weightings grade to a deeper shade of red. Positive weightings grade to a deeper shade of blue.

The components of the radiation based PET downscaling model (Method B) are mapped in Figure 42. Clearly the strong negative weightings for surface radiation centred on Southland in component 1 become an important aspect of the model and explain the improved result for Method B in this part of New Zealand. Similarly a strong negative weighting for temperature in component 1 off the coast from Christchurch are an important factor in explain PET variability. The pattern of weightings for MSLP appears to be physically plausible with strong south east to north west gradient in components one and two and a pronounced north south gradient in component three.

These results indicate that there are two plausible direct downscaling models for PET. On one hand the temperature only method yields poorer overall diagnostics, but appears to provide a better approximation of above and below average PET events than the radiation method (B). This makes method A seem a more suitable option for use in the analysis of drought, but this would be made at the expense of a seemingly more accurate model. Further consideration is given to both these methods below, comparing twentieth century drought probabilities calculated from each approach. Before this two additional downscaling experiments are briefly described.

5.4.8 Monthly Temperature

Surface temperature is not used as part of the drought analyses undertaken in Chapter 4. However the results of PLS regression applied to monthly temperature downscaling are detailed briefly, assuming there will be general interest in this variable and to further examine the downscaling approach. It was also an opportunity to investigate the influence of the underlying observation model on the downscaling scheme. Here two alternatives to interpolating observed temperature are examined: the first observation set is derived from the VCSN data product, where daily temperature records are interpolated, and the monthly data obtained by aggregating the grids; the second set used the same underlying observations but the data were aggregated to monthly at each site and the resulting monthly observations interpolated using the same three dimensional spline♣. The downscaling procedure used monthly MSLP and surface temperature as predictors with the first three PLS components retained.

The overall downscaling model was very good for mean monthly temperature, and the diagnostics suggest that it is based on a strong and accurate relationship. The observation model in this case had little effect on the overall quality of the downscaling, although using the monthly interpolated data resulted in a slight spatial smoothing, evident when comparing the RMSE of the two approaches (Figure 43 and Figure 44). RMSE scores were in the order of 1-1.5 °C for the agricultural land, and around 2 °C in the alpine region of the South Island, for both observation models in the training and validation sets. High correlation coefficients were evident, above 0.85 across the country, again for both models and time partitions.

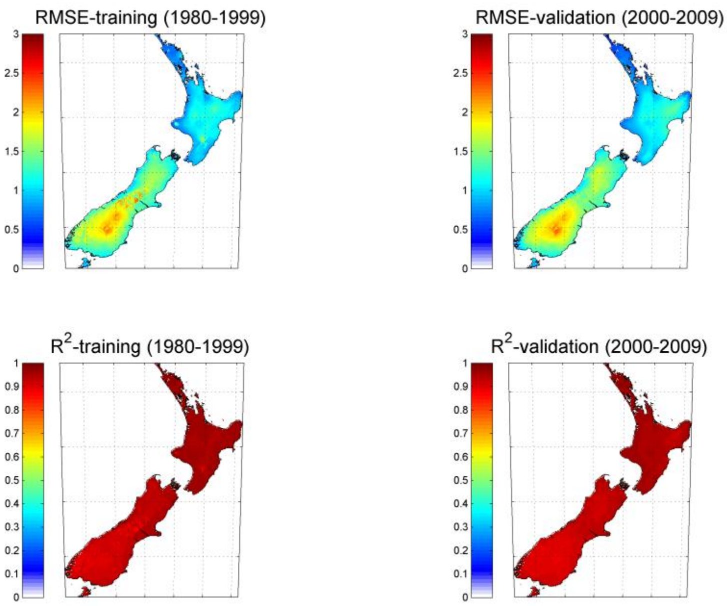


Figure 43: Diagnostics of PLS regressions downscaling for mean monthly temperature. Predictand is daily interpolated observation data aggregated to monthly.

♣ the author would like to acknowledge Dr Andrew Tait for developing the second set of data specifically for this experiment

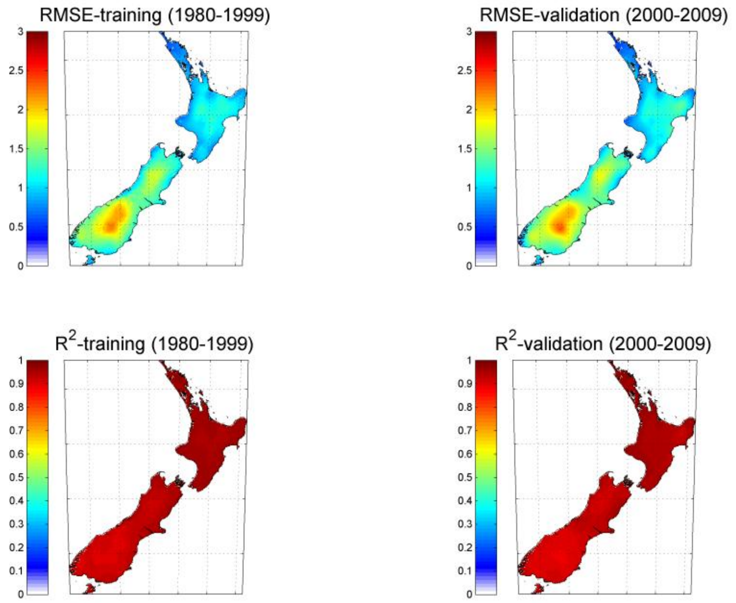
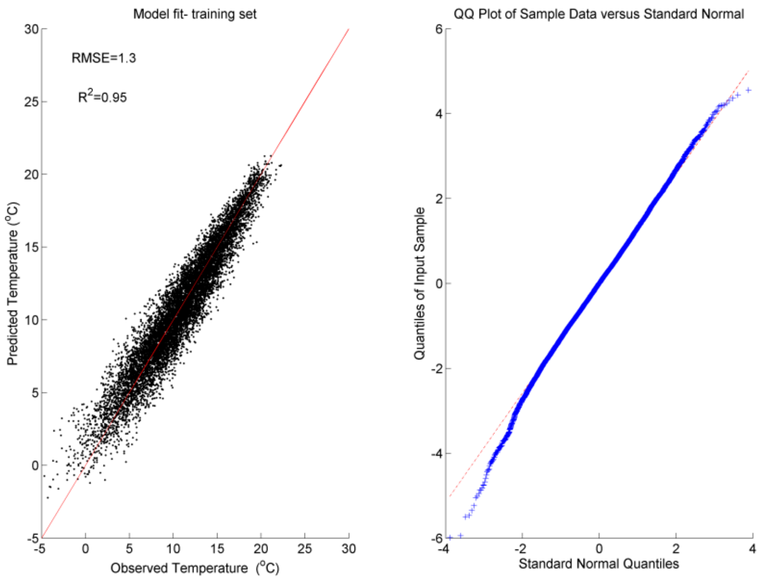
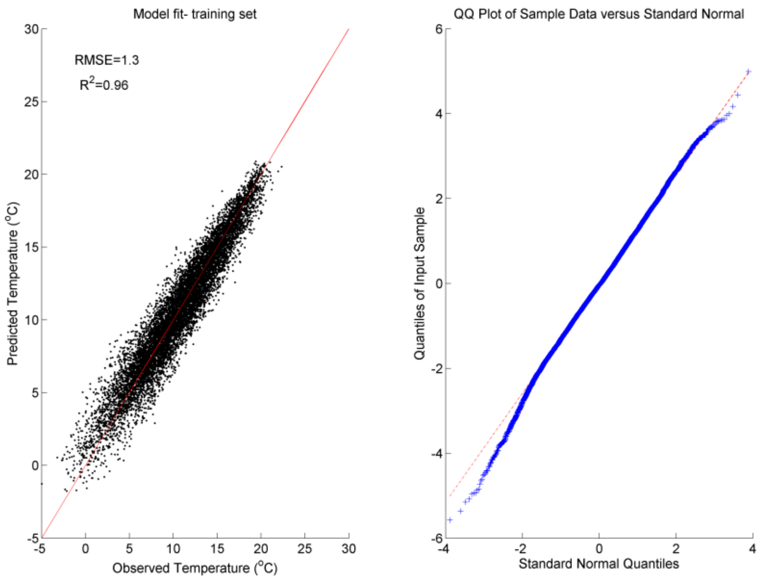


Figure 44: Diagnostics of PLS regressions downscaling for mean monthly temperature. Predictand is daily site observation data aggregated to monthly then interpolated.

Similarly the residual analyses cannot be differentiated based on observation set (Figure 45, Figure 46). Both provided a good approximation of the theoretical normal distribution, with good relationships between observed and predicted on agricultural land.

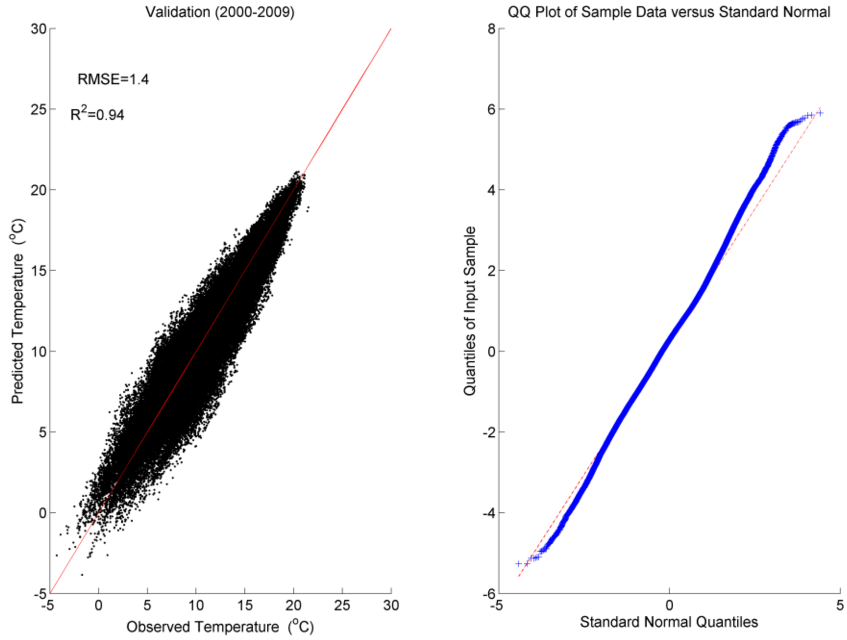


(a)

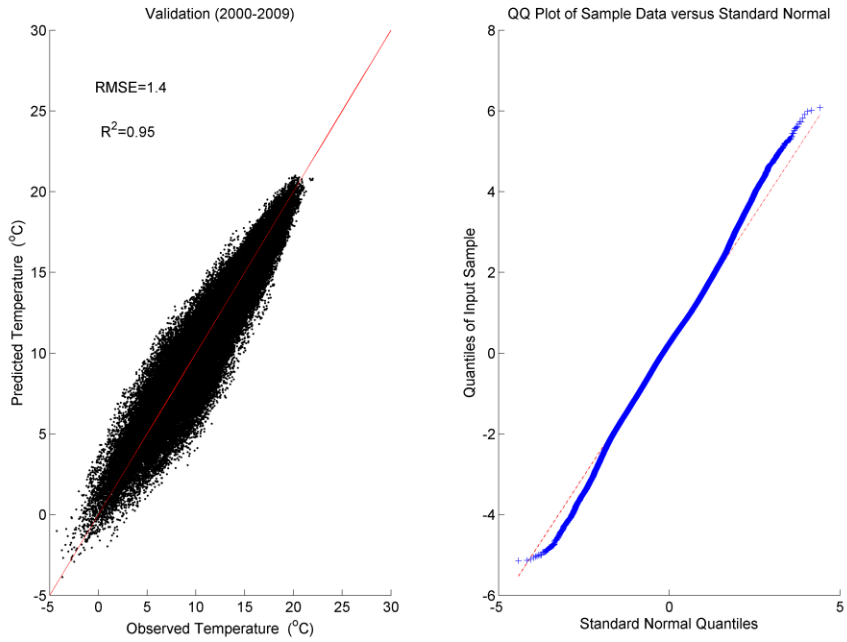


(b)

Figure 45: Residual analysis for the training window for the interpolate then aggregate observation set (a), and the aggregate then interpolate observation set (b).



(a)



(b)

Figure 46: Residual analysis for the validation window for the interpolate then aggregate observation set (a) and the aggregate then interpolate observation set (b).

5.4.9 Daily resolution

Climate resolved at a daily time step is usually of more interest in climate impact analysis, particularly those that use a downstream hydrological or agricultural model to simulate the behaviour of the system of interest. This is because the daily time step is often regarded as the base temporal resolution required to adequately simulate the dynamics of many biophysical systems—for instance the majority of agricultural systems models and associated water balance subroutines require daily resolution data to represent rainfall-runoff mechanism and key plant functions. In NIWA's current operational ESD scheme, assumptions are made to apply monthly change fields to adjust daily resolution data, thereby assuming the same temporal structure in the future as observed in the base period. There is clearly interest in relaxing this assumption so as to provide a more physically plausible set of future projections.

However the requirement for daily resolution data can create challenges for regression based ESD because the mathematical foundations of schemes (the optimisation methodology) rely on the assumption that the underlying distributions are Gaussian (normal). It is well known that the finer the time and spatial resolution of climate variables the more their distribution becomes non-Gaussian. This feature of climate scaling is particularly strong for rainfall, as it is governed by the interplay between complex physical mechanisms (frontal systems, stability of airflow and local convection). This leads to a well known property of many linear ESD schemes when applied to rainfall, that they do not resolve the tails of the distribution well— this result was evident even at the monthly time scale experiments with PLS regression described above as they required a moderate transformation so that the problem became Gaussian.

We expect that daily temperature, radiation and PET will to some extent be normally distributed, but we do not expect this for daily rainfall. Thus the use of linear PLS regression will require experimentation with some standard approaches to addressing this problem, such as transformations. Application of an alternative methodology, non-linear techniques and/or reforming of the optimisation problem is too resource intensive to meaningfully pursue as part of this research program. It needs to be recognised that this level of innovation is the current leading edge of ESD (see Cannon 2009), and is also one of the rationales to pursue a more physically based approach through RCM.

Despite these challenges it was decided to undertake some trials of PLS regression to downscale daily climate variability. These need to be viewed as limited initial trials, undertaken to examine if there is potential to further develop PLS regression for downscaling at the daily timescale. With the exception of daily temperature the experiments simply repeat the modelling strategies developed for monthly variables. Hence there remains considerable scope for further refinement. Although the daily downscaling models are not used in the final analysis, the results are detailed here as overall they were encouraging.

Daily rainfall

Daily rainfall was downscaled using daily MSLP and surface precipitation fields from the NCEP reanalysis as predictors and the daily VCSN fields as the predictand. The PLS regression retained 4 components. The only difference with the monthly strategy was the application of a stronger transformation, use of a cube root rather than a square root, to normalise the data. The RMSE scores mapped across the country are shown in Figure 47, suggesting that overall there was a minimal level of error on agricultural land with RMSE of 5mm or less in the training window and 10mm or less in the validation period. There was a strong spatial gradient in the distribution of error with evidence of topographic control, as very high RMSE were found in the Alpine regions of the North and South Islands, where there is also higher daily rainfall on average.

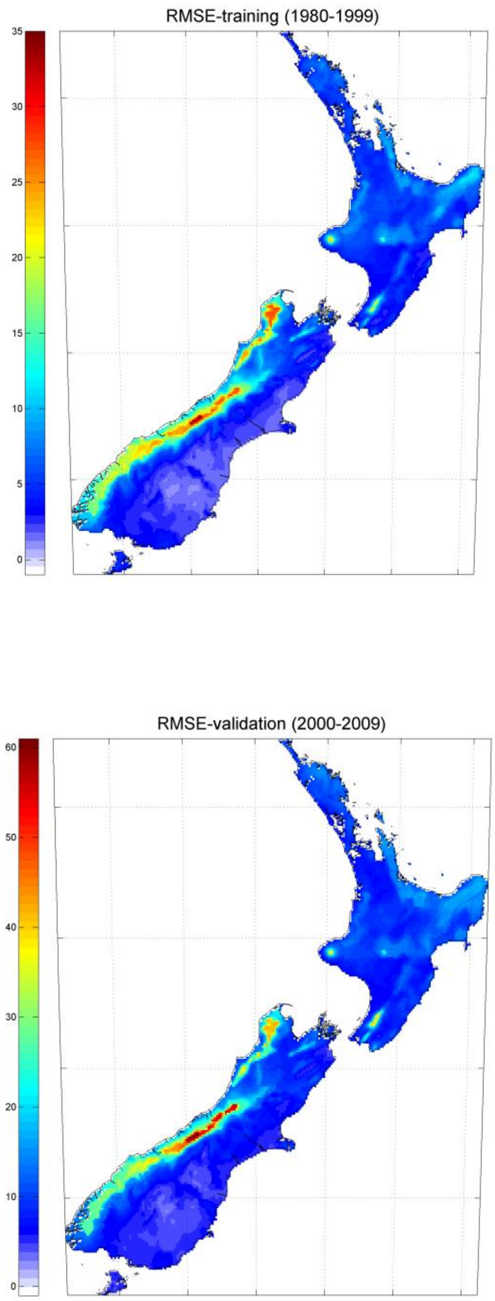
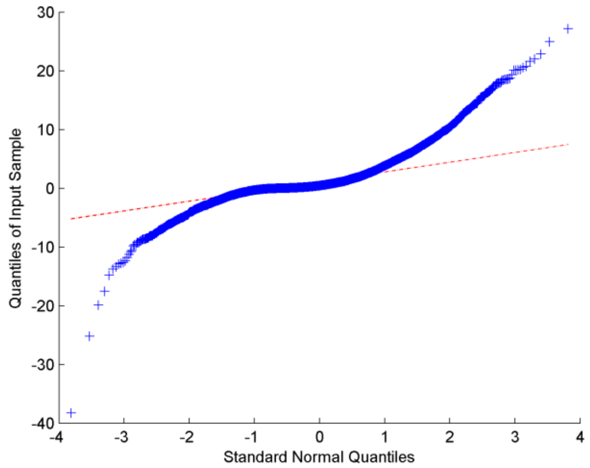
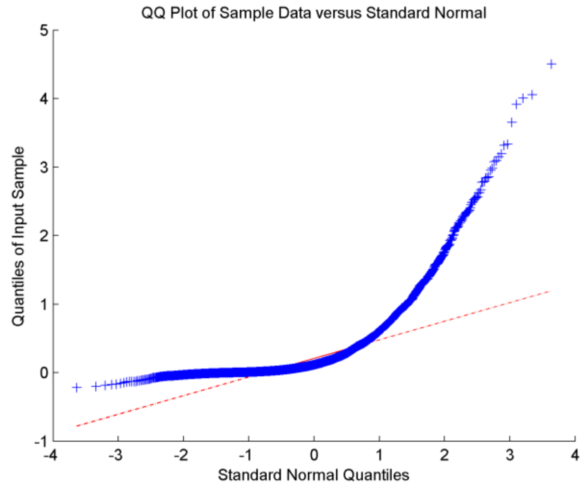


Figure 47: RMSE for daily rainfall for the training and validation period.

The residual analysis of the daily rainfall model is shown as qq plots in Figure 48. For the training period (a) the results were as anticipated, a strong tendency for the model to simulate the central moments of the distribution but not capture the tails or the extremes well. In the independent validation period (b) the model did not perform satisfactorily outside of the range of the first quantile, and in particular over-estimated variability in the lower quantile range. This is indicative of the model not capturing the number, persistence and magnitude of dry spells (days with zero rainfall). Given these results it would not be appropriate to use this downscaling model in an examination of drought.



a



b

Figure 48: Distribution of residuals for rainfall in the training (a) and validation (b) periods.

Despite these results they provide encouraging evidence to extend the PLS approach for this level of downscaling. Even this initial trial provided a downscaling model which had a physically plausible spatial distribution of error. Some relatively simple options that have been used successfully in downscaling daily rainfall elsewhere were not trialled here, for instance:

- developing separate models for downscaling dry and wet days;
- using stronger transformations;

- segregating the model into monthly or physically (synoptically) defined periods like Kidson synoptic types to target models towards the rainfall mechanisms;
- using different predictands and different transformations such as filtering; and
- lagging predictands in time.

There is also potential to pursue more difficult innovations at a fundamental mathematical level. Briefly this would involve integrating non-linear models and an appropriate optimisation methodology into the overall PLS strategy.

Daily potential evapotranspiration

Daily PET was downscaled using daily MSLP, surface temperature and incoming surface radiation as predictors and the interpolated daily Penman product from the VCSN set as the predictand. The RMSE results in Figure 49 indicate that the downscaling relationship is strong, with errors generally below 1.2 mm in both the training and validation period. There is evidence of some topographic influence with the highest errors occurring in higher altitude zones. There are also coastal areas on the east coast of both islands where the errors are higher, regions where high wind events associated with warm clear days during summer and autumn are known to increase the evaporative flux of water. Outside of these zones on a large portion of agricultural land RMSE was low, less than 1mm per day.

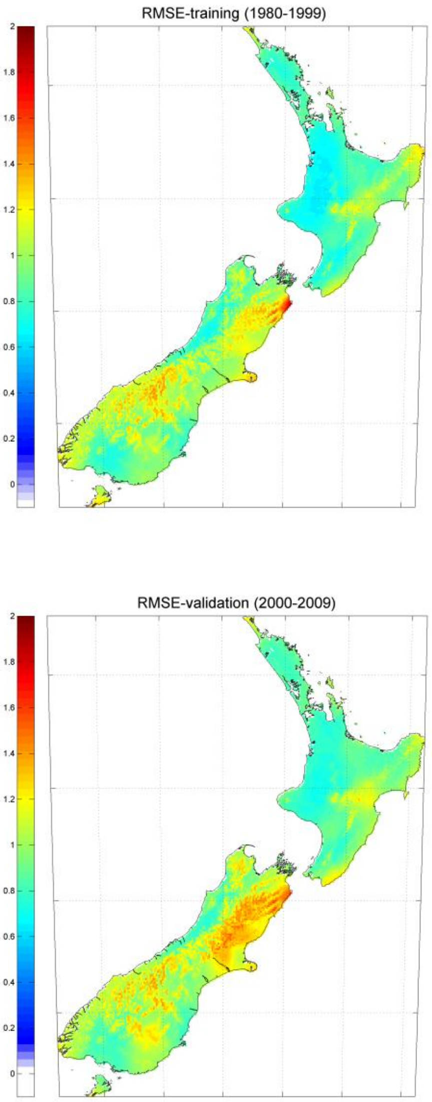
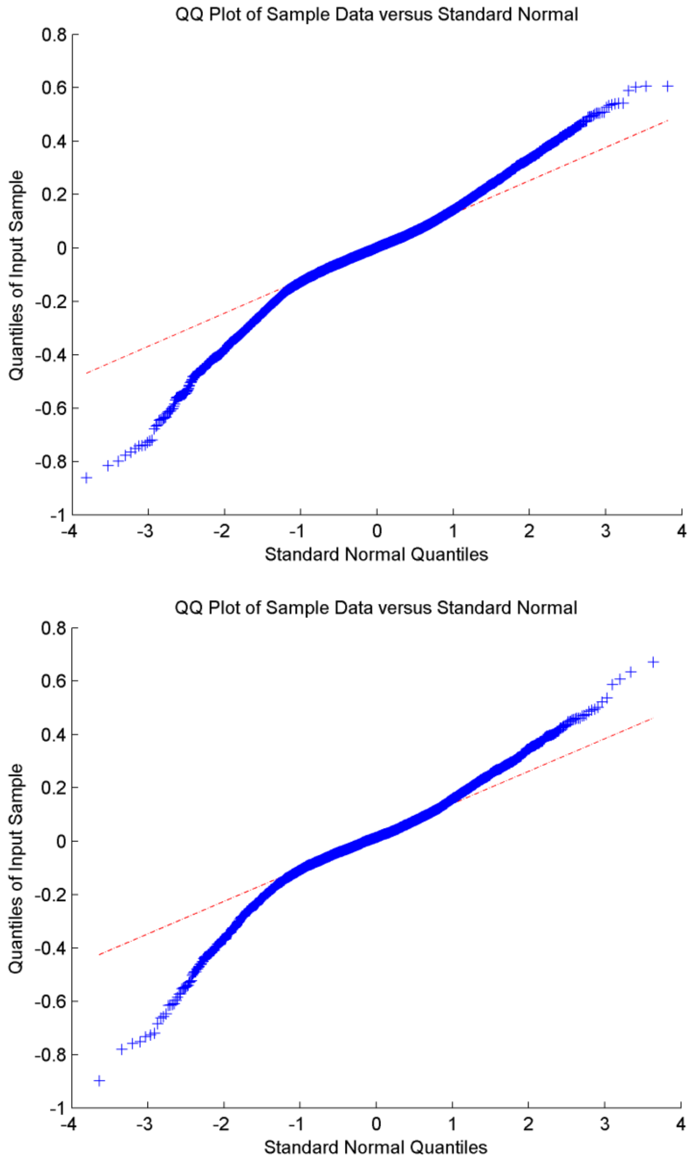


Figure 49: RMSE for daily PET in the training and validation period.

The qq plots in Figure 50 suggest that the model provided a reasonable approximation of the theoretical normal distribution of residuals, but as expected there were departures at the higher quantiles. The model appeared to perform slightly worse at the lower tail of the distribution, indicating that low evaporative events may not be well approximated in some regions. Although not examined there may be potential to address this problem by application of a mild data transformation. In aggregate these results are encouraging and suggest that this model could be used in a scenario analysis, particularly for drought which is likely to be more sensitive to high evaporative events which appear to be approximated well at least in some regions.



(a)

(b)

Figure 50: Distribution of residuals for daily PET in the training (a) and validation (b) periods.

Daily temperature

In some applications daily maximum and minimum temperatures are needed to simulate aspects of the biophysical system like plant functions, for example response to frost. If independent downscaling of these variables is implemented a common problem is that due to random model error on some days the daily maximum can be below the daily minimum. As a result a different strategy was developed for downscaling daily maximum and minimum temperatures than used for mean monthly temperature. Here three temperature variables are downscaled, the mean daily temperature and the anomaly above the mean and anomaly below the mean. Constraints are built into the optimisation such that the negative anomaly always needs to be below the mean and the positive anomaly always needs to be above the daily mean. Daily maximum and minimum temperatures are then recalculated after the downscaling procedure. Like the monthly scheme NCEP MSLP and mean surface temperature were used as predictors (but at a daily timestep) and the number of PLS components retained was four.

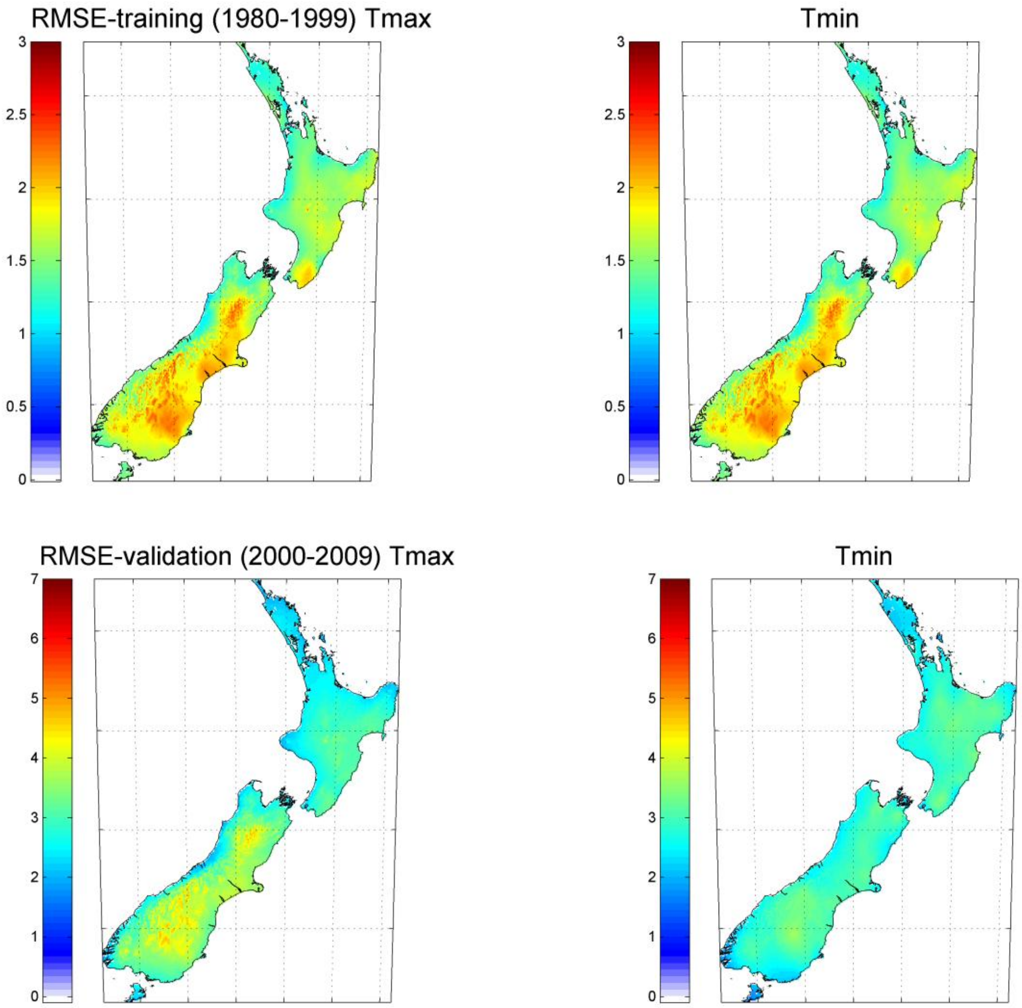


Figure 51: RMSE for daily temperature in the training and validation period.

The RMSE results for this approach to downscaling daily temperature are in Figure 51, and illustrate on the whole that the model is based on a strong and accurate relationship. Typically errors of around 2-3 °C were found for both the training and validation period, and like the downscaling of monthly mean temperature there appears to be some topographic control on the error structure. Again errors increased in the higher elevation and true alpine zone consistent with the hypothesis that in general the observation model is of reduced quality in these regions of New Zealand. Interestingly the lowest error occurred for minimum temperature during the independent test period, which runs counter to previous results where there is an expected deterioration of model accuracy scores in this test.

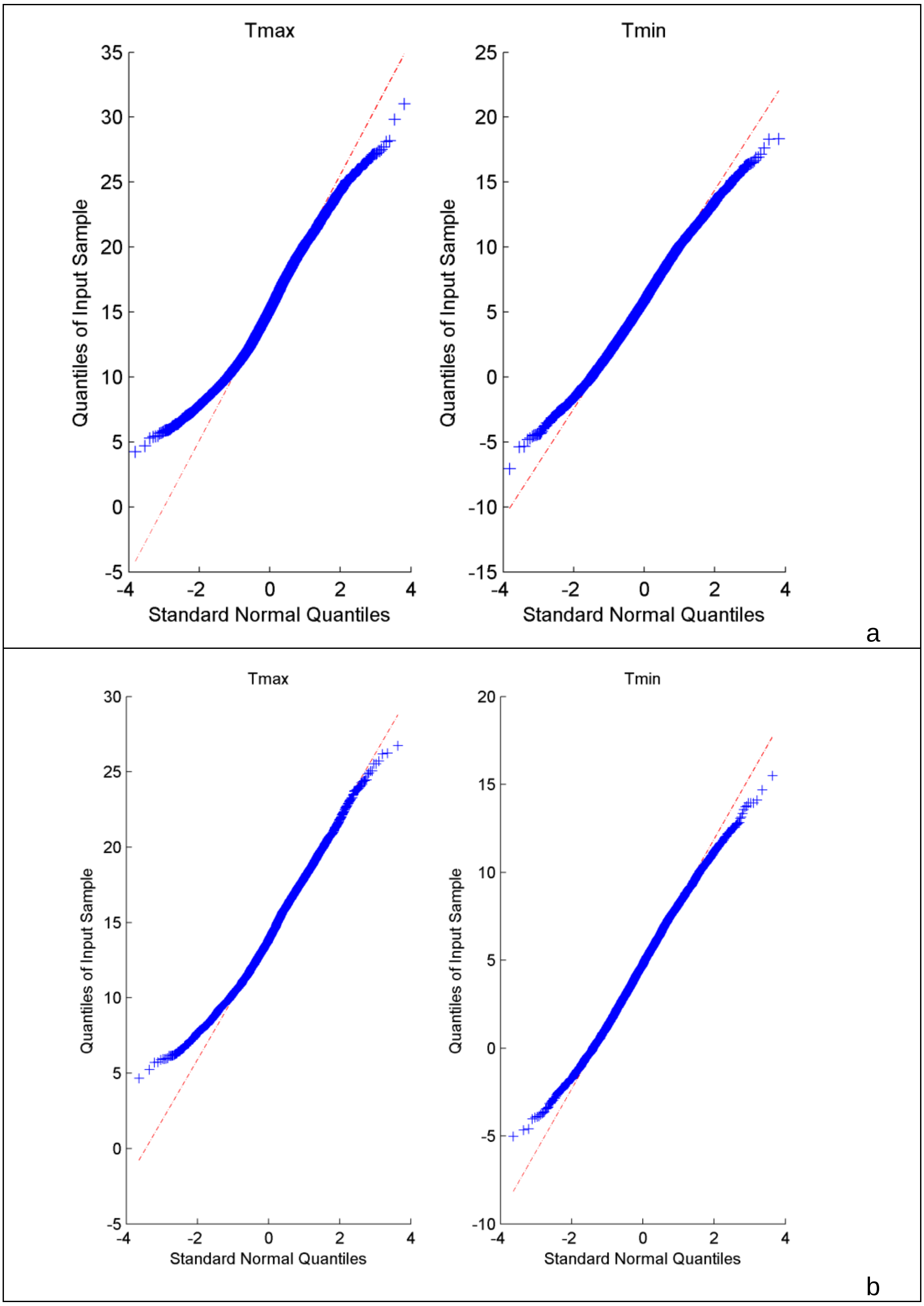


Figure 52: Distribution of residuals for daily temperature for the training (a) and validation (b) periods.

A small variation to the analysis of residuals is reported in Figure 52 as we wanted to evaluate the actual temperatures at which departures were strongest. The residuals were not rank ordered by observation quantile but by observation threshold so to assess the approach for some specific applications like projection of frost or high temperature events. Overall the tests illustrate that the model provided a good approximation of the theoretical normal distribution, but with some departure at the tails and for genuine extremes. Encouragingly the departures were not strong for minimum temperature at the 1 and 0 °C thresholds, suggesting that the downscaling model might be used with some confidence for evaluating projections of future frost. Similarly the model appeared to be satisfactorily resolving the estimation of the upper tail of the maximum daily temperature distribution. On the whole this downscaling scheme appears to be satisfactory for operational application.

5.5 Temporal downscaling

The use of projections at the monthly timescale in this study necessitates the use of a temporal downscaling method to provide climate variables at the resolution of the impact model. The utility and trade-offs between monthly versus daily climate data in the context of climate impact studies and statistical downscaling is discussed by Maurer and Hidalgo (2008). In previous New Zealand studies, including past work on drought (Mullan et al. 2005) mean monthly adjustments have been used to develop daily time series by adding the relevant monthly temperature adjustment to an observed daily record. For rainfall, adjustments are made only on days where rain occurred, so that the final change accumulates to match the monthly adjustment. This procedure for scaling daily time series is described by MfE (2008) and used in numerous other studies of climate change impact where daily resolved data are required, for example Bright et al. (2008). In a study of drought this is a limitation, as the future temporal variability (daily and monthly) is assumed to be largely unchanged or not responsive to GCM based changes in variability, around an imposed shift in the mean made with guidance from GCMs.

An alternative temporal downscaling procedure is applied in this study whereby the monthly variability of changes simulated by GCMs is preserved. This is based on the use of interpolation methods developed by Hutchinson (per comm. Hutchinson 2010) and described by Clark (2008). The overall strategy is to use the output of the PLS regression temporally downscaled to a weekly time step and run the water balance at this temporal resolution. For both potential evaporation and rainfall:

- the cumulative sum of the monthly data are taken to create a continuous positively incrementing series;
- a non-linear spline is fitted to the monthly cumulative series which has periodic end conditions, so the first and second derivatives are matched at the start and end of the function;
- the spline function is then evaluated at weekly intervals. The date conventions for calculating weekly series with consistent adjustments for leap years are adapted from Keig and McAlpine (1974);
- the evaluation at weekly intervals also approximates the 1st derivative of the accumulating monthly function using finite differences. This provides the weekly climate time series.

This approach to temporal downscaling is best applied when the results of the model runs are re-aggregated to monthly or seasonal time periods. Testing on a range of pastoral production indices for Southern Australia illustrated that the method provides an acceptable approximation of changes to monthly and seasonal means (Clark 2008). For this study, where the interest is in changes to droughts around the 10th percentile level, the temporal downscaling methodology is verified by comparing drought probabilities calculated using a common observation data set (VCSN, 1972-2009) where:

3. the soil water balance model is run at daily time steps, accumulated to months and then drought probabilities determined;
4. monthly climate inputs are temporally downscaled to weekly and the water balance run with this input data. The soil water data are aggregated to monthly for calculation of drought probabilities.

A comparison between the two approaches is shown in Figure 43, where the differences in drought probabilities are normalised, expressed as a percentage of the long term drought probability (time spent in drought) from the daily observations. The results are expressed in this way because the actual differences between the two methods are very small, usually below 0.01 percent of time spent in drought (a very small drought probability around 10^{-4}).

For most of New Zealand there are very small normalised percentage differences between the two approaches, between -2 and +2 percent of the daily derived drought probability (Figure 53). A key feature is the over estimation of drought probabilities by the temporal downscaling on the peaks of the Southern Alps (red in Figure 43) and underestimation in some microclimates such as Central Otago, the Kaikoura coast and coastal Wairarapa (white in Figure 43). These features are at worst around one fifth (20 percent) of the signal of that estimated from the daily observations, while in some areas this equates to only around 5 percent (light blue regions in Figure 43). It is important to recognize the differences between the methods equate to very small raw changes, around 0.06 of time spent in drought. Given the relative stability of the methodology and the small magnitudes of biases found, the weekly downscaling methodology can be applied in this study with a degree of confidence.

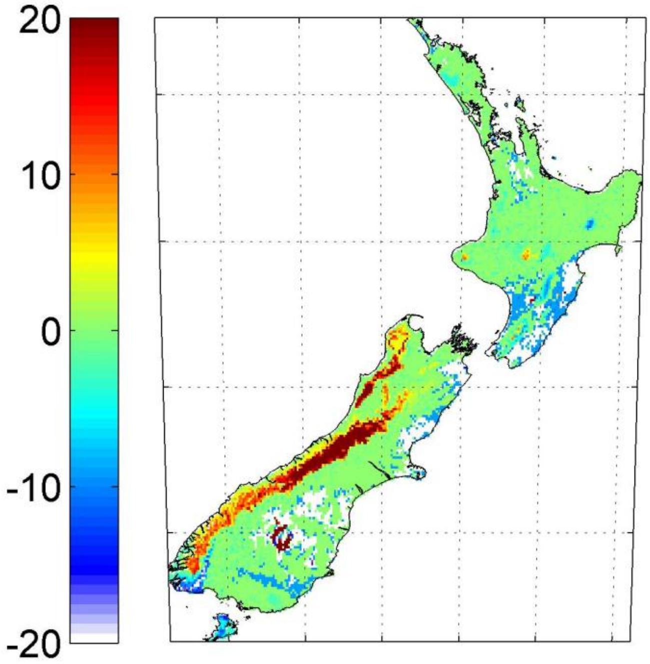


Figure 53: Percentage difference between long term drought probabilities (1972-2009) derived from daily observations and those derived from monthly observations temporally downscaled to weekly.

This approach to temporal downscaling along with the monthly PLS regression downscaling opens up a number of opportunities in the study. In particular the capacity to undertake transient 100 year simulations or locate the study in any time slice. Importantly the month-to-month variability simulated by the GCMs is preserved, which will ensure the analysis is sensitive to changes in a proportion of the temporal variability as well as shifts in the mean state under future projected climates.

5.6 Verification of drought probabilities

As described previously an important feature to investigate is the amount of error the downscaling scheme introduces into an analysis, so as to ensure that final inferences are made with respect to projected climate change, rather than the properties of the downscaling scheme itself. This is achieved by assessing the biases between drought probabilities calculated by downscaling the NCEP reanalysis data and then the GCM data, relative to those derived from observations. All the analyses here use the period 1980-1999 as the base period and drought probability expressed as the percentage of time a given location was deemed to be in drought.

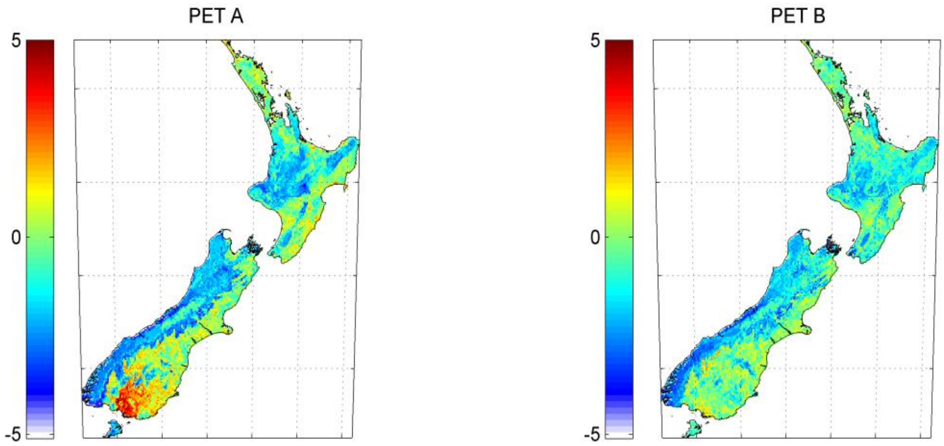


Figure 54: Percentage difference between drought probabilities calculated with downscaled NCEP data and those calculated with observations. PET A is the 'temperature' downscaling model while PET B is the' radiation' downscaling model. The colour scale shows differences between drought probabilities (percentage of time in drought) with those derived from observations.

The bias in the NCEP based downscaling is shown in Figure 54, and reflects the error structure found in both the rainfall and PET downscaling. The overall magnitude of the bias is relatively low, typically between 1-2 percent with the largest biases being in the order of 5 percent (e.g., a change in drought probability from 0.100 to 0.105 is a 5% bias). For both approaches there is a detectable bias associated with topography, where generally droughts in the alpine and higher altitude zones were underestimated, at worst around 5 percent. In the lowland agricultural regions biases were lower, typically an over estimation of around one to two percent, but there were large areas where this was close to zero. The radiation downscaling model performed better overall, particularly in low land agricultural areas with a tighter overall constraint on the bias. Reflecting the results for PET downscaling, including radiation appeared to address a large over-estimation of drought in Southland with the temperature only model.

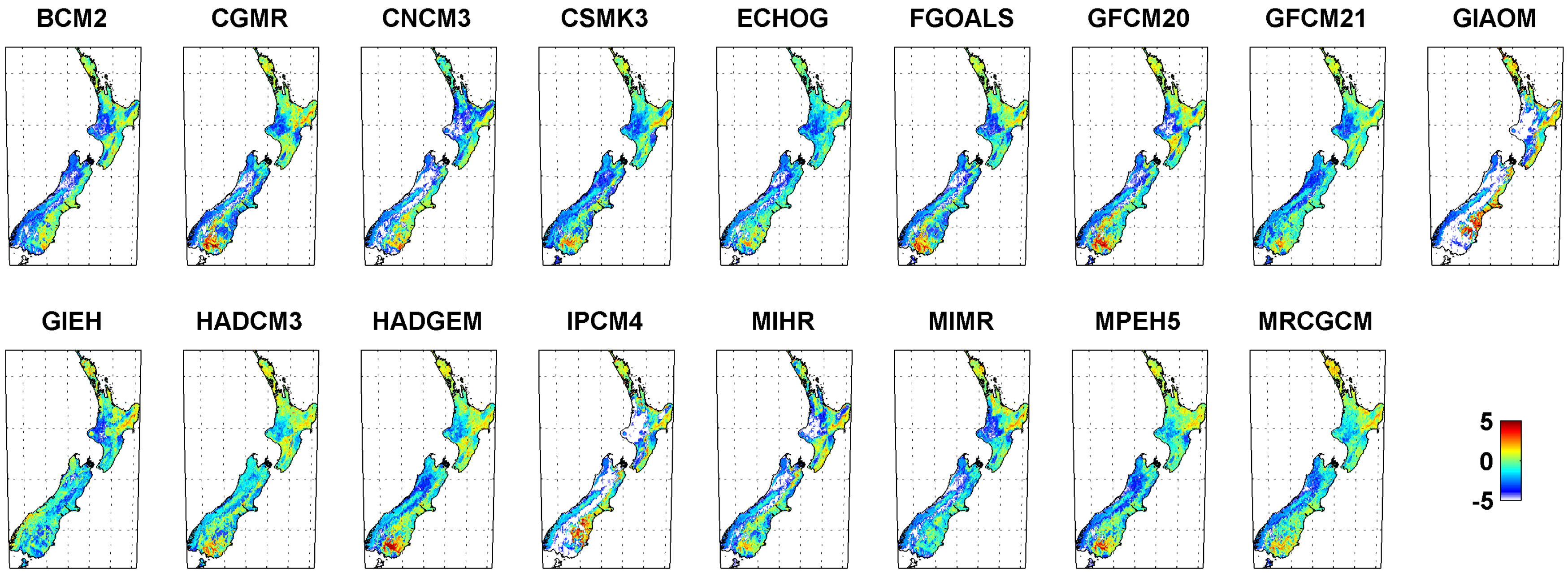


Figure 55: Difference between drought probabilities calculated with downscaled GCM data and those calculated with observations using PET method A.

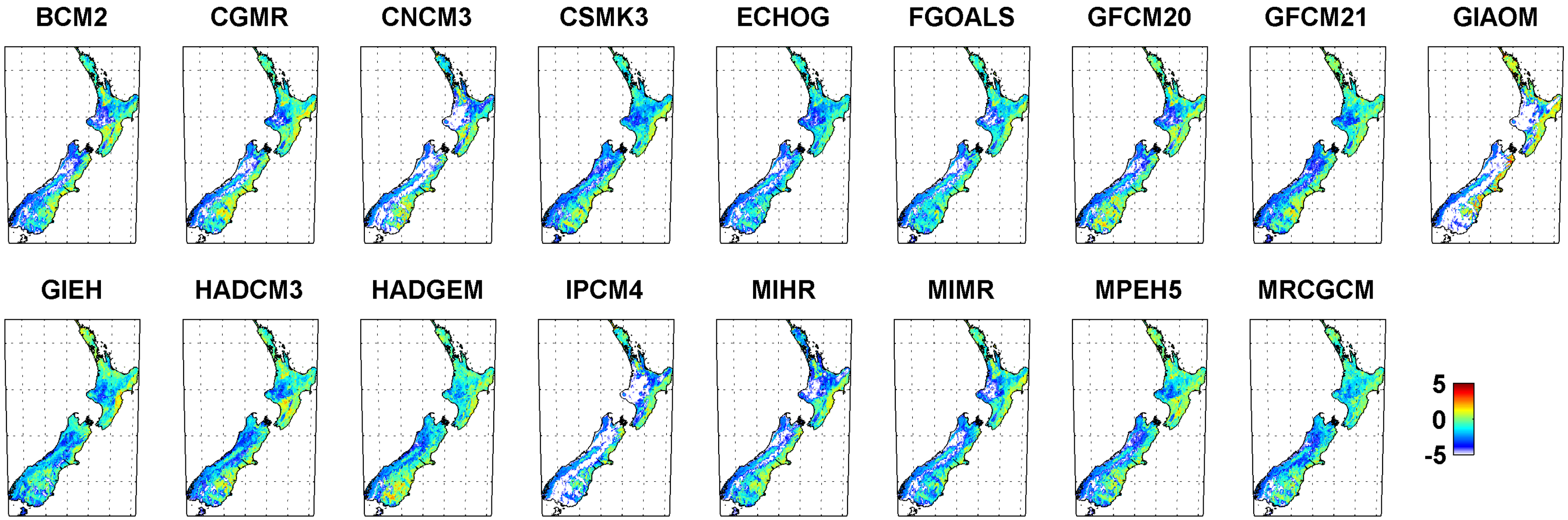


Figure 56: Difference between drought probabilities calculated with downscaled GCM data and those calculated with observations using PET method B.

The use of GCM fields to resolve 20th century drought probabilities introduces some additional biases, although overall the results were supportive in terms of establishing the validity of the downscaled AR4 model ensemble in this study (Figure 45 and Figure 46). Generally the same spatial gradient as the NCEP based downscaling was evident, and biases typically constrained by plus or minus five percent. The regions of least bias were typically the agricultural zones notably the east coasts of both islands. Within this overall spatial structure there are also subtle differences between GCMs with region to region variations in the biases—this is an expected feature when using GCMs, particularly precipitation fields.

The most obvious additional feature introduced by some of the GCMs and downscaling methodology is the under-estimation of drought probabilities on the alpine zone of the South Island, notably for GIAOM, CNCM3 and IPCM4[1]. With the temperature only PET model the large positive bias reported previously for Southland is present in all the GCM resolved drought probabilities to some extent. Interestingly the location of this bias seems to vary, with some of the models placing it neatly in Southland, while others such as GIAOM locating it on the lower east coast. This is a good illustration of the spatial biases found when using GCM data, related to slightly different spatial positioning of the large synoptic features.

When applied using GCM data, the radiation downscaling model (B) removed the Southland bias, but it did not result in as strong an overall improvement compared to the NCEP downscaling. This is likely due to the poorer estimation of surface radiation by the GCMs. In some models resolved with GCM output, the introduction of the radiation term degraded the signal compared to the temperature based approach. This is an important consideration going forward, as although there appears to be a superior approach to downscaling PET using radiation from the observation sets, it may not provide the same confidence when building future projections using GCM data.

[1] The GIAOM model (from Goddard Institute for Space Studies, USA) and the IPCM4 model (from Institut Pierre Simon Laplace, France) were deleted from consideration in the MfE (2008) climate change Guidance Manual, because of their poor simulations of New Zealand 20th century climate.

6. References

Allen, R. G.; Pereira, L. S. (1998). Crop evapotranspiration - Guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56. Rome, Food and Agriculture Organization of the United Nations.

Alley, W.M. (1984). The Palmer Drought Severity Index: limitations and assumptions. *Journal of Climate and Applied Meteorology, 23*:1100-1109.

Andrew, R.M.; Andrew, J.R. (2007). A distributed model of water balance in the Motueka catchment, New Zealand. *Environmental Modelling and Software, 22*: 1510-15.

Annan, J.D.; Hargreaves J.C. (2010). Reliability of the CMIP3 ensemble. *Geophysical Research Letters 37,* L02703, doi:10.1029/2009GL041994.

Arena, C.; Cannarozzo, M.; Mozzola, M.R. (2006). Multi-year drought frequency analysis at multiple sites by operational hydrology- a comparison of methods. *Physics and Chemistry of the Earth, 31*: 1146-1163.

Aster, R.C.; Borchers, B.; Thuber, C. (2005). Parameter estimation and inverse problems. Elsevier, 320p.

Barringer, J.; Lilburne, L. (1999). Scale issues in developing regional scale water balance surfaces. The 11th Annual Colloquium of the Spatial Information Research Centre.

Benestad, R.E. (2002). Empirically downscaled temperature scenarios for Northern Europe based on a multi-model ensemble. *Climate Research, 21:* 105-125.

Benestad, R.E. (2010). Downscaling precipitation extremes. *Theoretical and Applied Climatology, 100:* 1-21.

Benestad, R.E; Hanssen-Bauer, I.; Chen, D. (2008). Empirical-Statistical Downscaling. World Scientific Publishing Co., Singapore, 215 p.

Blenkinsop, S.; Fowler, H.J. (2007). Changes in drought frequency, severity and duration for the British Isles projected by the PRUDENCE regional climate model. *Journal of Hydrology, 342:* 50-71.

Blenkinsop, S.; Fowler, H.J. (2007). Changes in European drought characteristics projected by the PRUDENCE models. *International Journal of Climatology, 27*: 1595-1610.

Bright, J.; Rutter, H.; Dommisse, J.; Woods, R.; Tait, A.; Hendrikx, J.; Diettrich, J. (2008). Projected effects of climate change on water supply reliability in Mid-Canterbury. Aqualink Report No, Co8120/1.

Cancelliere, A.; Mauro, G.D.; Bonaccorso, B.; Rossi, G. (2007). Drought forecasting using the standardised precipitation index. *Water Resources Management, 21:* 801-819.

Cannon, A.J. (2009). Negative ridge regression parameters for improving the covariance structure of multivariate linear downscaling models. *International Journal of Climatology, 29:* 761-769.

Carter, T.R., (2007). General guidelines on the use of scenario data for climate impact and adaptation assessment. Task Group and Scenario Support for Impact and Climate Assessment (TGICA), Intergovernmental Panel on Climate Change.

Charles, S.P.; Bates, B.C.; Smith, I.N.; Hughes, J.P. (2004). Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. *Hydrological Processes, 18:* 1373-1394.

Chavez Eguez, J.L; Neale, C.M.; Prueger, J.H.; Kustas, W.P. (2008). Daily evapotranspiration estimates from extrapolating instantaneous airborne remote sensing ET values. *Irrigation Science, 27:* 67-81.

Clark, A.J. (2001). Risk management for climate agriculture and policy, Bureau of Rural Sciences, Commonwealth of Australia.

Clark, A.J. (2008). Simplified models for analysing the impact of climate change on Australia's sheep pastoral zone, PhD Thesis, Australian National University.

Clark, A.J.; Tait, A.B. (2008). Drought, agricultural production and climate change – a new way forward to a better understanding. NIWA Client Report WESIP1011.

Clausen, B.; Pearson, C.P. (1995). Regional frequency analysis of annual maximum streamflow drought. *Journal of Hydrology, 173:* 111-130.

Day, K.A.; Rickert, K.G.; McKeon, G.M. (2005). Monitoring agricultural drought in Australia. In: 'Monitoring and Predicting Agricultural Drought: a Global Study'. (Eds V.K. Boken, A.P. Cracknell and R.L. Heathcote), pp.369–385, Oxford University Press.

D'onofrio, A.; Boulanger, J.P.; Segura, E.C. (2010). CHAC: a weather pattern classification system for regional climate downscaling of daily precipitation. *Climatic Change, 98:* 405-427.

Droogers, P. (2000). Estimating actual evapotranspiration using a detailed agro-hydrological model. *Journal of Hydrology, 229:* 50-58.

Edwards, E.J.; McCaffery, S.; Evans, J.R. (2005). Phosphorus status determines biomass response to elevated CO2 in a legume: C4 grass community. *Global Change Biology, 11:* 1968-1981.

Fontaine, M.M.; Steinemann, A.C. (2009). Assessing vulnerability to natural hazards: impact-based method and application to drought in Washington State. *Natural Hazards Review, 10:* 11-18.

Fowler, A.; Aiken, S.; Maree, K. (2008). Vulnerability of New Zealand pastoral farming to the impacts of future climate change on the soil water regime. Client report prepared for the Ministry of Agriculture and Forestry, School of Geography and Earth Sciences, The University of Auckland.

Ghannoum, O.; von Caemmerer, S.; Conroy, J.P. (2001). Plant water use efficiency at 17 Australian NAD-ME and NADP-ME C4 grasses at ambient and elevated CO2 partial pressure. *Australian Journal of Plant Physiology, 28:* 1207-1217.

Ghosh, S.; Mujumdar, P.P. (2007). Nonparametric methods for modelling GCM and scenario uncertainty in drought assessment. *Water Resources Research, 43:* 1034-56.

Gibbs, W.J.; Maher, J.V. (1967). Rainfall deciles as drought indicators. Bureau of Meteorology Bulletin No. 48, Commonwealth of Australia, Melbourne.

Gifford, R. (2004). The CO2 fertilisation effect – does it occur in the real world? *New Phytologist, 163:* 221-223.

Gonzalez, J.; Valdes, J.B. (2003). Bivariate drought recurrence analysis using tree ring reconstructions. *Journal of Hydrologic Engineering, 8:* 247-258.

Gordon, A.H. (1992). The random nature of drought: mathematical and physical causes. *International Journal of Climatology, 13:* 497-507.

Green, S.; Clothier, B.; Dijsell, C.V.; Deurer, M. (2008). Measuring and modelling the stress response of grapevines to soil-water deficits. Chapter 12. Responses of Crops to limited water: understanding and modelling water stress effects on plant growth processes. Advances in Agricultural Systems Modelling Series 1.

Hayhoe, K.; Wake, C.P.; Huntington, T.G.; Luo, L.; Schwartz, M.; Sheffield, J.; Wood, E.; Anderson, B.; Bradbury, J.; DeGaetano, A.; Troy, T.; Wolfe, D. (2007). Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dynamics, 28:* 381-407.

Hennessy, K.; Fawcett, R.; Kirono, D.; Mpelasoka, F.; Jones, D.; Bathols, J.; Whetton, P.; Stafford Smith, M.; Howden, M.; Mitchell, C.; Plummer, N. (2008). An assessment of the impact of climate change on the nature and frequency of exceptional climatic events. CSIRO and Bureau of Meteorology, 33 p. http://www.bom.gov.au/droughtec.

Hisdal, H.; Tallasken, L.M. (2000). Drought Event Definition, Technical Report to the ARIDE project No. 6. Department of Geophysics, University of Oslo.

Imbert, A.; Benestad, R.E. (2005). An improvement of the analog model strategy for more reliable local climate change scenarios. *Theoretical and Applied Climatology, 82:* 245-255.

IPCC (2001). *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881 pp.

IPCC (2007). Summary for Policymakers. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S.; D. Qin, M.

Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Kalnay et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin American Meteorological Society, 77:* 437-470.

Katz, R. W. (2002). Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Research, 20:* 167-185.

Kay, A.L.; Davies, H.N. (2008). Calculating potential evaporation from climate model data: a source of uncertainty for hydrological climate change impacts. *Journal of Hydrology, 358:* 221-239.

Keig, B.; McAlpine, J.R. (1974). WATBAL: A computer program for the estimation and analysis of soil moisture from simple climatic data (second edition). Technical Memo 74/4. Queensland, Australia: Division of Land Research, Commonwealth Scientific and Industrial Research Organization.

Kenny, G.J.; Warrick, R.A.; Mitchell, N.D.; Mullan, A.B.; Salinger, M.J. (1995). CLIMPACTS: An integrated model for assessment of the effects of climate change on the New Zealand environment. *Journal of Biogeography, 22:* 883-895.

Kidson, J.W. (2000). An analysis of New Zealand synoptic types and their use in defining weather regimes. *International Journal of Climatology, 20:* 299-316.

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

Kim, T.W.; Valdes, J.B.; Yoo, C. (2003). Non-parametric approach for estimating return periods of droughts in arid regions. *Journal of Hydrologic Engineering, 8:* 237-246.

Kimball, B.A. (2010). Lessons from FACE: CO2 Effects and Interactions with Water, Nitrogen and Temperature, in Handbook of Climate Change and Agroecosystems, Hillel, D and Rosenzweig, C, Imperial College Press.

Kingston, D.G.; Todd, M.C.; Taylor, R.G.; Thompson, J.R.; Arnell, N.W. (2009). Uncertainty in the estimation of potential evapotranspiration under climate change. *Geophysical Research Letters, 36:* L20403.

Knutti, R.; Furrer, R.; Tebaldi, C.; Cermak, J.; Meehl, G.A. (2010). Challenges in combining projections from multiple climate models. *Journal of Climate, 23:* 2739-2758.

Larsen, S.H. (2005). Evaluation of Drought at Lincoln, Canterbury, Environment Canterbury, Report No. R05/29.

Lawrence, J.E.; Hornberger, G.M., (2007). Soil moisture variability across climate zones. *Geophysical Research Letters, 34*: L20402, doi:10.1029/2007GL031382.

Li, X.; Sailor, D. (2000). Application of tree-structured regression for regional precipitation prediction using general circulation model output. *Climate Research, 16:* 17-39.

Li, Y.; Smith, I. (2008). A statistical downscaling methodology for Southern Australia winter rainfall. *Journal of Climate, 22:* 1142:1158.

Lu, J.; Sun G.; McNulty, S.G.; Amatya, D.M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. *Journal of American Water Resources Association, 41*: 621-633.

MAF (2009). Regional and National Impacts of the 2007-2008 Drought. Prepared for MAF Policy by Butcher Partners Ltd, July 2009.

Martinez, W.L.; Martinez A.G. (2002). Computational statistics handbook with Matlab. Chapman and Hall CRC.

Maurer, E.P.; Hidalgo, H.G. (2008). Utility of daily vs monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrological and Earth System Sciences, 12*: 551-563.

McDonald, C.K. (1994). Calculating climatic indices affecting plant growth. Tropical Agronomy Technical Memorandum No. 83, CSIRO Division of Tropical Crops and Pastures, Brisbane, Queensland, Australia.

Meehl, G.A;, Covey, C.; Delworth, T.; Latif, M.; McAvaney, B.; Mitchell, J.F.B.; Stouffer, R.J.; Taylor, K.E. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin American Meteorological Society, 88:* 1383–1394.

Meehl, G.A.; Stocker, T.F.; Collins, W.D.; Friedelinstien, P.; Gaye, A.T.; Gregory, J.M.; Kitoh, A.; Knutti, R.; Murphy, J.M.; Noda, A.; Raper, S.C.B.; Watterson, I.G.; Weaver, A.J.; Zhao, Z.C. (2007). Global climate change projections. In: Climate Change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Ministry for the Environment. 2004. *Climate Change Effects and Impacts Assessment: A Guidance Manual for Local Government in New Zealand.* Prepared by: David Wratt, Brett Mullan and Jim Salinger (NIWA), Sylvia Allan and Tania Morgan (MWH New Zealand Ltd), and Gavin Kenny (Earthwise Consulting). Ministry for the Environment Report ME 513, Wellington, 153p.

Ministry for the Environment. (2008). *Climate Change Effects and Impacts Assessment: A Guidance Manual for Local Government in New Zealand 2nd Edition.* Prepared by: Brett Mullan, David Wratt, Sam Dean and Michele Hollis (NIWA), Sylvia Allan and Tania Williams (MWH New Zealand Ltd), and Gavin Kenny (Earthwise Consulting), with MfE. Ministry for the Environment, ME 870, Wellington, 156 p. http://www.mfe.govt.nz/publications/climate/climate-change-effect-impacts-assessments-may08/index.html (2 April 2008).

Mosley, M. P.; Pearson, C.P. (1997). Introduction: hydrological extremes and climate in New Zealand. *In:* Mosley, M.P.; Pearson, C.P., editors. *Floods and Droughts: The New Zealand Experience*. Wellington: New Zealand Hydrological Society, pp 1–14.

Mullan, A.B. (1995). On the linearity and stability of Southern Oscillation-climate relationships for New Zealand. *International Journal of Climatology, 15*: 1365-1386.

Mullan, A.B.; Wratt, D.S.; Renwick, J.A. (2001). Transient model scenarios of climate changes for New Zealand. *Weather and Climate 21*: 3–34.

Mullan, B.; Carey-Smith, T.; Griffiths, G.; Sood, A. (2011). Scenarios of storminess and regional wind extremes under climate change. NIWA Client Report WLG2010-31 prepared for Ministry of Agriculture and Forestry, March 2011, 80p.

Mullan, B.; Porteous, A.; Wratt, D.; Hollis, M. (2005). Changes in drought risk with climate change. NIWA Project MFE05305, Report Prepared for the Ministry for the Environment and Ministry of Agriculture and Forestry, May 2005, 58p.

Murphy, J.M.; Sexton, D.M.H.; Barnett, D.N.; Jones, G.S.; Webb, M.J.; Collins, M.; Stainforth, D.A. (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature, 430:* 768-772.

Nakicenovic, N.; Swart, R (Eds). (2000). Special Report on Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 599 p.

Narasimhan, B.; Srinivasan, R. (2005). Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agricultural and Forest Meteorology, 133:* 69-88.

Neitsch, S.L.; Arnold, J.G.; Kiniry, J.T.; Williams, J.R. (2005). Soil and Water Assessment Tool Theoretical Documentation, Grassland Soil and Water Research Laboratory, Texas.

Nelson, R.; Howden, M.H.; Stafford-Smith, M. (2008). Using adaptive governance to rethink the way science supports Australian drought policy. *Environmental Policy and Planning, 11:* 588-601.

Oudin, L.; Michel, C.; Anctil, F. (2005a). Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1 – Can rainfall-runoff models effectively handle detailed potential evaporation inputs. *Journal of Hydrology, 303:* 275-289.

Oudin, L.; Hervieu, F.; Michel, C.; Perrin, C.; Andreassian, V.; Anctil, F.; Loumagne, C. (2005b). Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall runoff modelling. *Journal of Hydrology, 303:* 290-306.

Palmer, D.J.; Watt, M.S.; Höck, B.K.; Lowe, D.J.; Payn, T.W. (2009). A dynamic framework for spatial modelling of *pinus radiata* soil water balance (SWaTBal) across New Zealand. Forest Research Bulletin No 234.

Palmer, D.J.; Höck, B.K.; Lowe, D.J.; Dunningham, A.G.; Payn, T.W. (2009). Developing National Scale Terrain Attributes for New Zealand (TANZ), Forest Research Bulletin No 232.

Palmer, W.C. (1968). Keeping track of crop moisture conditions, nationwide: The new Crop Moisture Index. *Weatherwise 21:* 156-161.

Phillips, I.D.; McGregor, G.R. (1998). The utility of a drought index for assessing the drought hazard in Devon and Cornwall, South West England. *Meteorological Applications, 5:* 359-372.

Porteous, A. (2005). Changes in drought risk with climate change. Part 1: Defining a climate change sensitive drought index. NIWA Client Report WLG2004-49 prepared for New Zealand Ministry for the Environment, June 2004, 39p.

Ramsay, D.; Thompson, C.S.; Mullan, B.; Burgess, S. (2008). Kiribati Adaptation Programme. Phase II: Information for Climate Risk Management, High Intensity rainfall and drought. NIWA Client Report WLG2008-12, prepared for Government of Kiribati, July 2008, 103p.

Rao, A.R.; Voeller, T.L. (1997). Development and testing of drought indicators, *Water Resources Management, 11:* 119-136.

Raupach, M.R.; Briggs, P.R.; King, E.A.; Paget, M.; Trudinger, C.M. (2007). Australian Water Availability Project, CSIRO Marine and Atmospheric Research Component, Final Report Phase 2. CSIRO, Australia.

Rayner, D.P. (2007). Wind run changes: the dominant factor affecting pan evaporation trends in Australia. *Journal of Climate, 20:* 3379-3394.

Renwick, J.A.; Mullan, A.B.; Porteous, A. (2009). Statistical downscaling of New Zealand climate. *Weather and Climate, 29:* 24-44.

Schmith, T. (2008). Stationarity of regression relationships: application to empirical downscaling. *Journal of Climate: 21:* 4529-4537.

Semenov, M.A.; Barrow, E.M. (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change, 35*: 397–414.

Sheffield, J.; Wood, E.F. (2006). Global trends and Variability in Soil Moisture and Drought Characteristics, 1950-2000, from observations driven simulations of the terrestrial hydrological cycle. *Journal of Climate, 21:* 432-458.

Shiau, J.; Shen, H.W. (2001). Recurrence analysis of hydrological droughts of different severity. *Journal of Water Resources Planning and Management, 127:* 30-40.

Smith, D.I.; Hutchinson, M.F.; McArthur, R.J. (1993). Australian climatic and agricultural drought: Payments and policy. *Drought Network News 5:* 11-12.

Stafford Smith, D.M.; McKeon, G.M. (1998). Assessing the historical frequency of drought events on grazing properties in Australian rangelands. *Agricultural Systems, 57:* 271-299.

Steinemann, A.C.; Cavalcanti, L.F.N. (2006). Developing multiple indicators and triggers for drought plans. *Journal of Water Resources Planning and Management, 132*: 164-174.

Stephens, D.J. (1998). Objective criteria for estimating the severity of drought in the wheat cropping areas of Australia. *Agricultural Systems, 57:* 333-350.

Stone, M.; Brooks, R. (1990). Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. *Journal of the Royal Statistical Society, Series B, 52(2):* 237-269.

Sturman, A.; Tapper, N. (1996). The weather and climate of Australia and New Zealand. Oxford University Press.

Tait A.B. (2008). Future projections of growing degree days and frost in New Zealand and some implications for grape growing. *Weather and Climate, 28:* 17-36.

Tait A.B.; Liley, B. (2009). Interpolation of daily solar radiation for New Zealand using a satellite data-derived cloud cover surface. *Weather and Climate, 29:* 70-88.

Tait A.B.; Woods, R. (2009). Spatial interpolation of daily potential evapotranspiration for New Zealand using a spline model. *Journal of Hydrometeorology, 8: 430*-438

Tait, A.; Henderson, R.; Turner, R.; Zheng, X. (2006). Thin plate smoothing splines interpolation of daily rainfall for New Zealand using a climatological rainfall surface. *International Journal of Climatology, 26:* 2097-2116.

Teuling, A.J.; Troch, P.A. (2005). Improved understanding of soil moisture variability dynamics. *Geophysical Research Letters, 32:* L05404, doi:10.1029/2004GL021935.

Thompson, C.S. (2006). Relative influence of the Interdecadal Pacific Oscillation on drought occurrence and severity. *Weather and Climate: 26:* 35-66.

Timbal, B. (2004). Southwest Australia past and future rainfall trends. *Climate Research, 26:* 233-249.

Timbal, B.; Fernandez, E.; Li, Z. (2009). Generalization of a statistical downscaling model to provide local climate change projections for Australia. *Environmental Modeling and Software, 24*: 341-358.

Trenberth, K.E.; Dai, A.; Rasmusson, R.M.; Parsons, D.B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society, 95:* 1205-2017.

Wang, G. (2005). Agricultural drought in a future climate: results from 15 global climate models participating in the IPCC 4th assessment. *Climate Dynamics, 25:* 739-753.

Weiss, M.; Menzel, L. (2008). A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments. *Advances in Geoscience, 18:* 15-23, doi:10.5194/adgeo-18-15-2008.

Wetterhall, F.; Halldin, S.; Xu, C. (2005). Statistical precipitation downscaling in central Sweden with the analogue method. *Journal of Hydrology, 306:* 174-190.

White D.H.; Walcott J.J. (2009). The role of seasonal indices in monitoring and assessing agricultural and other droughts: a review. *Crop and Pasture Science, 60:* 599-616.

White, D.H.; Bordas, V.M. (1997). Indicators of Drought Exceptional Circumstances, in Proceedings of a Workshop Held in Canberra on 1 October 1996, (Eds. D.H. White and V.M. Bordas). Bureau of Resource Sciences, Commonwealth of Australia.

White, D.H.; Howden, S.M.; Walcott, J.J.; Cannon, R.M. (1998). A framework for estimating the extent and severity of drought based on a grazing system in south eastern Australia. *Agricultural Systems, 57:* 359-270.

White, T. (2009). SFF Northland farmer resilience: a literature review. Client report prepared for MAF Sustainable Farming Fund, AgResearch, Hamilton.

Wilby, R.L. (1998). Statistical downscaling of daily precipitation using daily airflow and teleconnection indices. *Climate Research, 10:* 163-178.

Wilby, R.L.; Wigley, T.M.L. (2000). Precipitation predictors for downscaling: Observed and General Circulation Model relationships. *International Journal of Climatology, 20:* 641-661.

Wilby, R.L.; Charles S.P.; Zorita, E.; Timbal, B.; Whetton, P.; Mearns, L.O. (2004). Guidelines for use of climate scenarios developed from statistical downscaling methods. Task Group and Scenario Support for Impact and Climate Assessment (TGICA) Intergovernmental Panel on Climate Change.

Wilhite, D.A. (2000). Drought planning and risk assessment: status and future directions. *Annals of Arid Zone 39(3):* 211-230

Wilhite, D.A.; Svoboda, M.D.; Hayes, M.J. (2006). Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness. *Water Resources Management, 21:* 763-774.

Wilson, D.J.; Western, A.W.; Grayson, R.B.; Berg, A.A.; Lear, M.S.; Redell, M.; Famiglietti, J.S.; Woods, R.A.; McMahon, T.A. (2003). Spatial distribution of soil moisture over 6 and 30cm depth, Mahurangi river catchment, New Zealand. *Journal of Hydrology, 276:* 254-274.

Woodward, S.J.R.; Barker, D.J.; Zyskowski, R.F. (2001). A practical model for predicting soil water deficit in New Zealand pastures. *New Zealand Journal of Agricultural Research, 44*: 91-109.

Yoo, C.; Kim, S.; Kim, T. (2006). Assessment of drought vulnerability based on the soil moisture PDF. *Stochastic Environmental Resource Risk Assessment, 21:* 131-141.

Yurekli, K.; Kurune, A. (2006). Simulating agricultural drought periods based on daily rainfall and crop water consumption. *Journal of Arid Environments, 67:* 629-640.

Acknowledgements

This work was funded by the New Zealand Ministry for Agriculture and Forestry as a 1 year program - Sustainable Land Management Mitigation & Adaptation to Climate Change, under Contract C01X0818.