Australian Climate Service CMIP6-Next Generation downscaled climate change projections: approach and summary

Produced by CSIRO and the Bureau of Meteorology for the Australian Climate Service Program 3

May 2025

DOI: https://doi.org/10.25919/9bde-a338

Citation

CSIRO and Bureau of Meteorology (2025). Australian Climate Service CMIP6-Next Generation downscaled climate change projections – approach and summary. Technical Report, CSIRO and Bureau of Meteorology, Australia, doi.org/10.25919/9bde-a338

Acknowledgements

This research has been undertaken with support of the Australian Climate Service. The Australian Climate Service is a partnership made of the Bureau of Meteorology, CSIRO, the Australian Bureau of Statistics and Geoscience Australia.

Key supporting repository: https://github.com/AusClimateService/

Chapters and Contributors

Leads: Michael Grose and Sugata Narsey

1. Introduction: Michael Grose and David Jones
2. Modelling strategy: Michael Grose
3. Model ensemble description: Chun-Hsu Su, Emma Howard, Marcus Thatcher, Sonny Truong, Michael Grose, Sugata Narsey
4. Evaluation and benchmarking of the ACS and CORDEX ensemble: Emma Howard, Xiaoxuan Jiang, Benjamin Ng, Christian Stassen, Michael Grose, expert advice on independence: Gab Abramowitz
5. Application-ready locally relevant datasets: Damien Irving, Alicia Takbash, Andrew Gammon, Justin Peter, Michael Grose
6. Projections methods and choices: Michael Grose and Mitchell Black
7. Mean changes and uncertainty in projections ensembles: Sugata Narsey, Acacia Pepler, Mitchell Black and Michael Grose
8. Extremes and hazards: Mitchell Black, Acacia Pepler, Richard Matear, David Hoffmann, Jessica Bhardwaj, Tess Parker, Doerte Jakob
9. Gaps, future opportunities and challenges: Sugata Narsey, Michael Grose, Benjamin Ng, Marcus Thatcher, James Risbey

Review editors: Christine Chung, David Jones, James Risbey, Tess Parker

Internal reviewers: Pandora Hope (BoM), Karl Braganza (BoM), Ian Macadam (CSIRO)

External reviewers: Ben Booth (UKMO), Peter Gibson (NIWA)

Executive Summary

Introduction and strategy

Climate projections are tools used for various purposes related to future climate change, including raising awareness, informing discussion, motivating emissions mitigation and planning adaptations. New and updated projections are motivated by advancements in modelling, including finer resolution, and a general desire to understand the latest evidence on plausible future climate change. The updated projections presented here are based on the latest international climate modelling and have a strong focus on climate extremes, supported by a multi-model Regional Climate Model (RCM) ensemble following Coordinated Regional Climate Downscaling Experiment (CORDEX) guidelines, supplemented by insights from Global Climate Models (GCMs) from the Coupled Model Intercomparison Project phase 6 (CMIP6), including large ensembles (many runs of the same model).

Methods, evaluation and choices to present projections

Climate change is examined under two Shared Socioeconomic Pathways (SSPs) that 'bracket' or 'bookend' the bulk of the likely range of future scenarios for greenhouse gas emissions, from a pathway compliant with the Paris Agreement (SSP1-2.6), to a high pathway (SSP3-7.0) used as a fallback from the highest pathway now seen as less plausible (SSP5-8.5). Strategic sampling of CMIP6 GCMs with RCMs forms a multi-model 'sparse matrix' design. GCM selection is as representative of the full CMIP6 ensemble as possible, such as sampling a spread of climate sensitivity and warming as well as increases and decreases in rainfall. But the sampling is still limited, so new simulations can be used in a 'representative climate futures' framework of key plausible futures, rather than a probabilistic range. Two RCMs have been developed and run for historical and future SSPs to 2100 by the Australian Climate Service (ACS) and are considered in conjunction with simulations from two other RCMs provided by the New South Wales and Queensland governments.

Evaluation and benchmarking analysis do not suggest rejecting any simulation outright, and independence analysis suggests equal weighting is acceptable as a basic technique (but more sophisticated ensemble-generation techniques should be pursued). Added value analyses suggest that the RCM ensemble produces consistent and credible new information for some variables and regions, but other areas show little, inconsistent or even negative added value, which needs further research. This RCM ensemble is compared to up to 35 CMIP6 models, and the outputs from two key large ensembles of the key wetter and drier representative climate futures.

Choices of presenting future contexts are for the two SSPs for time horizons consistent with the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (1995-2014 to 20-year periods centred on 2030, 2050 and 2090), or else between Global Warming Levels (GWLs) of 1.2 °C representing current warming forward to 1.5, 2 and 3 °C global warming, derived by time-sampling the SSP3-7.0 simulations (sampling the 20-year period when the model moves through that level of global warming).

Application-ready locally relevant datasets for further analysis and downstream modelling are created by calibrating model outputs to two commonly used datasets (Australian Gridded Climate Data; AGCD and the BARRA-R2 reanalysis) and presenting data on a common 5 km resolution grid. Data were produced using two techniques: 1) quantile scaling from a set of nine representative CMIP6 models, and 2) bias adjustment of all CORDEX simulations, with two methods (one univariate and one multivariate), to provide different options suited to different applications.

Projected changes to average conditions

Australia and all its sub-regions have warmed in recent decades and are *virtually certain* to continue warming under any plausible future scenario. All modelling considered consistently shows that warming is projected to approximately stabilize alongside global temperature at around mid-century under a scenario where net zero is reached and the Paris Agreement is met (SSP1-2.6), reaching around 0.7 to 1.4 °C warmer than the recent baseline (although a 'high warming' outcome approaching 2 °C can't be ruled out). In contrast, modelling suggests Australian warming continues throughout the century under SSP3-7.0, with a range of 2.6 to 4.1 °C by the end of the century relative to the recent baseline (high warming outcomes approaching 5 °C can't be ruled out), with warming continuing past 2100. Australian warming in CMIP and CORDEX models is at a rate similar or higher than global warming, but at a ratio lower than in observations (and requires further investigation). Regional and seasonal variations in warming are similar to those previously found.

Regional rainfall changes remain quite uncertain for many regions and seasons, with ongoing large climate variability. New CORDEX modelling does not systematically reduce the ranges of projected changes. An important exception is the southwest Western Australian cool season rainfall where ongoing decline remains *very likely* based on various lines of evidence and high model agreement in multiple generations of models including the new CORDEX ensemble. For all applications, a range of distinct plausible futures should always be considered when planning adaptation to the mean rainfall climate, with simulations representative of these futures available across the range of models. Key futures include an ongoing consistent decline in rainfall that is highly significant relevant to natural variability in almost all regions (provided by the ACCESS-ESM1.5 large ensemble and downscaling from this model), through to a more variable rainfall climate with the possibility of wetter periods than the historical baseline (provided by the EC-Earth3 large ensemble and downscaling from this model).

The spread of temperature and rainfall changes from the new projections span a similar range of change as previous projections for the four large super-cluster regions of Australia. In terms of warming, the low scenarios (RCP2.6 and SSP1-2.6) and the responses are similar in both projections, however we note that RCP2.6 was not strongly emphasized as it was seen as not policy-relevant prior to the Paris Agreement. The high scenario considered here (SSP3-7.0) entails lower concentration of greenhouse gases than the highest scenario considered previously (RCP8.5) but the presence of the 'low likelihood high warming' futures mean that the top end of warming is similar.

Projected changes to mean surface windspeeds are similar to previous generations of projections, including a decrease in mean wind speeds in southern mainland Australia, especially in winter.

Extremes

A sample of extremes is examined here, related to temperature and rainfall and are compared to the mean changes. We do not cover a complete list of relevant extremes for users; please see the separate report on indices, extremes and hazards.

Extreme heat and heatwaves as measured by the RXx (hottest day of the year), RXge40 (days over 40 °C) and the Excess Heat Factor (EHF) are projected to increase significantly in all regions, especially under the high scenario at further time horizons. Changes to the temperature of the hottest days could plausibly increase by more than the mean warming.

Regional models show a range of responses in projected change in heavy daily and hourly rainfall, but the overall CORDEX ensemble shows a consistent increase everywhere at a rate slightly less than the Clausius-Clapeyron relation. Extratropical cyclones (low pressure systems) that can bring extreme rainfall are projected to decrease in frequency in most seasons and regions, especially southern Australia. However, the overall impacts from the lows that do occur are projected to increase in various ways (e.g., higher rain rate, making landfall on a higher sea level, and potentially higher peak intensities).

Large climate variability at annual to decadal timescales will continue in Australia, including multi-year droughts and wet periods. Along with this, due to climate change the time in at least mild meteorological drought on average is projected to increase in the southwest, large parts of the southeast and parts of the northeast along the Queensland coast, with less certainty in the direction of change in other regions. Further analysis of drought that includes factors other than rainfall is covered in upcoming work, including changes in evaporation and land surface as well as consideration of aridity alongside changes to drought.

Forest Fire Danger Index (FFDI), taken as a general fire weather danger indicator as one part of fire risk, is projected to increase everywhere. There is a notable projected increase in the frequency of days in the severe (FFDI >50) category.

Contents

Commonly used climate variables

Short name	Long name
tasmax	Daily maximum temperature at 2m
tasmin	Daily minimum temperature at 2m
pr	Precipitation
rsds	Downwelling solar radiation
sfcWindmax	Daily peak near surface wind speed
hursmax	Daily maximum relative humidity at 2m
hursmin	Daily minimum relative humidity at 2m

Extremes Indices

Index/Metric	Definition	Variable
tasmax99	99th percentile of daily maximum temperature	tasmax
TXx	Annual maximum of daily maximum temperature	tasmax
TNx	Annual maximum of daily minimum temperature	tasmax
TXgt40	Number of days where the daily maximum exceeds 40°C each year	tasmax
tasmin01	1st percentile of minimum temperature	tasmin
TNn	Annual minimum of daily minimum temperature	tasmin
pr99	99th percentile of daily precipitation	pr
Rx1day	Maximum 1-day rainfall	pr
Rx1hour	Maximum 1-hour rainfall	pr
R20mm	Number of days when precipitation exceeds 20mm	pr
CWD	Consecutive wet days	pr
DD	Number of dry days (precipitation less than 1mm)	pr
CDD	Consecutive dry days	pr

Commonly used acronyms

ACS	Australian Climate Service
BARRA	Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis for Australia
CCAM	Conformal Cubic Atmospheric Model
BARPA	Bureau of Meteorology Atmospheric high-resolution Regional Projections for Australia
CMIP6	Coupled Model Intercomparison Project phase 6
CORDEX	Coordinated Regional Downscaling Experiment
GCM	Global Climate Model
RCM	Regional Climate Model
IPCC	Intergovernmental Panel on Climate Change
SSP	Shared Socioeconomic Pathway
GWL	Global Warming Level
ENSO	El Niño Southern Oscillation
ACCESS	Australian Climate Community Earth System Simulator

1. Introduction

- **Climate projections are used to quantify physical risks from a changing climate, used for various purposes related to future climate change – e.g., raising awareness, informing discussions, motivating emissions mitigation and planning adaptations.**
- **New and updated projections are motivated by advancements in modelling, including finer resolution, and a general desire to understand the latest evidence.**

The global climate has changed over the last century, with human influence playing an unequivocal and dominant role in many aspects (IPCC 1990-2023). The Australian climate continues to change, with increases in temperatures, shifts in rainfall, rising sea levels and other changes (BoM and CSIRO 2024).

Anthropogenic climate change presents a profound shift in how we manage climate risk. While natural climate variations have always been part of the climate system over periods of years or decades, these variations are usually modest at large scales, and the assumption of stationarity implies that conditions will tend towards a long-term average (Wilks 1996). It was on this basis that in its early days, climate science was heavily devoted to the study of average climate conditions reflected in the extensive definition of climate zones and climate classifications (Köppen 1936, Beck et al., 2018). Traditionally, observations played a major role in how we designed and planned for the future, with no use of projections of climates outside our observed experience. The accumulation of anthropogenic climate change over the past century and its recent acceleration across a range of measures, including sea level rise and temperature increase, changes how we must plan for the future. Looking at past observations alone provides an increasingly poor basis for the future.

Since we can no longer hold the assumption of stationarity, projections of future climates from dynamical Earth systems models are increasing forming the cornerstone of efforts to manage climate risk. The climate is projected to change further this century, with the rate and magnitude of change depending on the scenario of human development the world follows. Governments at all levels, businesses, non-government organisations, communities and individuals want to assess, plan for and adapt to projected climate change impacts at the regional scale. Hence, the field of regional climate projections has emerged. Regional climate projections have often focused on changes to climate classifications, changes to the mean climate for food production, water supply and natural resource management, but increasingly they focus on future extreme weather events, bringing in tools such as regional models to address this question.

Regional climate change projections are primarily derived from information from physical climate models, compared with and assessed against other lines of evidence. The success of regional climate projections as tools for decision-making can be measured in at least four dimensions: credibility, salience, legitimacy, and the decision-relevance of the information (Cash et al. 2003; Wilby and Dessai et al. 2010). Over the last 40 years, Whetton et al. (2016) noted that efforts to produce and communicate regional climate projections in Australia have faced:

- *Perennial issues* such as managing and communicating the unavoidable uncertainty in future change (sometimes termed the ‘cascade’ of uncertainty).
- *Trends* towards wider scope and sophistication in modelling and products.

- *Tensions* in aspects such as spatial resolution and use of probability in projections.

This experience has been reflected in other national projections products (USGCRP 2017, 2023). These dimensions of success of projections, and the various aspects of projections practice are relevant context to the technical description of new work here.

For regional climate projections, decision makers want and demand the latest that science and technology can offer. This includes the development in modelling itself, especially in Australia where we are exposed to physical impacts where insights from new models may be crucial. The latest research includes the Intergovernmental Panel on Climate Change Sixth Assessment Report (IPCC AR6 2023) and the Coupled Model Intercomparison Project phase 6 (CMIP6) generation of global climate models (Eyring et al. 2016) and derived modelling. Global climate models (GCMs) have been a central tool for producing regional climate projections, but there is a growing call for greater local insights, supporting the use of the Coordinated Regional Downscaling Experiment (CORDEX) and other regional modelling to gain greater spatial resolution and reduce local model biases (Gutowski et al. 2016; CORDEX 2021). Regional modelling is resource-intensive, and programs can be limited, however there is a need to cover both breadth and depth in the information produced. There is also a growing recognition that FAIR principles are useful to ensure reliability in the data channels.

Here, we lay out the strategy and technical details of new national projections for Australia. These projections update the previous national projections released in 2015 (CSIRO and BoM, 2015). They have been developed as part of the Australian Climate Service (ACS), and in partnership with others. This collaboration was supported by the National Partnership for Climate Projections (NPCP) and the creation of a National Climate Projections Roadmap.

This report outlines the motivation, background, strategy, methods, evaluation and broad results from the new generation of national climate projections for Australia. It starts with an outline of the selected modelling strategy (Chapter 2), followed by a more detailed description of the model ensemble, including the Regional Climate Models (RCMs) used (Chapter 3). Next is a summary of the model evaluation and benchmarking of the models, including an analysis of added value and ensemble independence (Chapter 4). Following this, the report gives a high-level summary of the processing to produce application-ready datasets (Chapter 5), as well as the projections methods and choices used (Chapter 6). The report presents projected changes in climate mean conditions (Chapter 7) and some selected climate extremes (Chapter 8). We finish with future work and next steps (Chapter 9). All sections draw on numerous underpinning technical documents and resources, which are linked to where appropriate.

2. Modelling strategy

- **Updated projections are based on the latest international climate modelling and have a strong focus on climate extremes.**
- **The primary source of climate model outputs is a multi-model CORDEX ensemble, supplemented by insights from Global Climate Models, including Large Ensembles.**
- **ACS Program 3 contributed modelling from two Regional Climate Models to form an important part of the CORDEX multi-model ensemble to underpin projections**

This chapter presents the strategy for projections for national applications with consistent national coverage. The focus over various past generations of Australian climate projections has been strongly on climate averages, including the effect on water, agriculture, natural resource management and similar areas, with an additional focus on extreme events. Following a series of disruptive climate events in recent years, these updated projections have a much stronger remit on extreme events and hazards. To provide useful information for managing risks from natural disasters, we must follow a strategy that addresses changes in extreme events at the regional scale, while also assessing changes to the average climate.

The current generation of regional climate projections can draw on outputs from Global Climate Models (GCMs) and Earth Systems Models (ESMs) from CMIP6, as well as Regional Climate Models (RCMs), and statistical techniques, including Artificial Intelligence/Machine Learning (AI/ML). Given the range of available options, a defensible modelling strategy with broad support from the research community is needed.

There are various methods to make climate projections regionally relevant, which Australian analysis suggests can be assessed in terms of 1) climate realism and 2) physical plausibility of change (Ekström et al. 2015). Both aspects are affected by the models' ability to simulate relevant climate processes at the full range of spatial scales. Projections need a balance between breadth (e.g., assessing a variety of structurally different models, large ensembles) and depth (e.g., spatial and temporal resolution of regional processes and extremes), with no one ideal trade-off.

To facilitate the strategy for new projections in Australia, in late 2017 the Australian projections community held the NextGen Projections Workshop, followed by a further developed NextGen Projections report. The main conclusions were:

- Projections must serve an evolving set of needs and uses, current and new users, and new developments (such as the latest sets of plausible future scenarios).
- In terms of climate modelling, there are compelling reasons to link to the new generation of IPCC Assessment and Global Climate Modelling (CMIP6), but to incorporate a coordinated program of Regional Climate Modelling through the CORDEX program as a key data source.
- A multi-model ensemble of downscaled Regional Climate Model (RCM) simulations produced through partnership can be used in a complementary way with CMIP6 itself, including the 'large ensembles' of CMIP6 models (many runs of the same model with slightly different initial conditions); further very high-resolution modelling and techniques such as machine learning.

This plan is consistent with the Climate Science for Australia's Future report from the National Climate Science Advisory Committee (NCSAC, 2019) and was further called for and endorsed by Recommendation 4.5 of the Royal Commission into National Natural Disaster Arrangements (the 'Bushfire RC'). The plan and the partnerships required were then supported through the formation of the National Partnership for Climate Projections (NPCP) and the publication of the Climate Projections Roadmap for Australia. The Australian Climate Service was set up in the wake of the Bushfire RC, with a mandate to work with states and territories to deliver suburb-scale future climate projections for Australia, covering a range of hazards, including heat waves, bushfire risk, flood risk and sea level rise.

Relative to the previous national projections (CSIRO and BoM, 2015), which focused on regional area-averages from CMIP models, higher spatial resolution through Regional Climate Modelling is given high priority in new projections to add potential insights into climate extremes and hazards. Following good practice developed by the international science community, the strategy is for a multi-model multi-scenario ensemble (MME) of regional climate model simulations following the Coordinated Regional Downscaling Experiment (CORDEX) guidelines and using simulations from four downscaling models from three NPCP partners: Australian Climate Service (ACS) Program 3 through CSIRO and the Bureau of Meteorology, which we mainly focus on in this report; the NSW and Australian Regional Climate Modelling version 2 (NARCLIM2.0); and the Queensland Future Climate Platform version 2.

Global and regional model combinations follow a semi-organised 'sparse matrix' design, where not every GCM is downscaled by every RCM. This is advantageous as it enhances the sampling of GCMs (Sobolowski et al. 2023). Downscaling by all four RCMs is present for GCM simulations representative of a few key futures – for example the most extreme plausible drying of the continent provided by ACCESS-ESM1.5, and also the minimum plausible warming future provided by NorESM2-MM – allowing a close examination of these projections, including the effect of each RCM on the projected change signal.

In the selection, there was also a special focus on climate sensitivity and avoiding oversampling the 'hot models'. Here, 'hot models' or the 'hot model problem' refers to the unbalanced sampling of climate sensitivity in CMIP6, with disproportionate sampling of models with climate sensitivity above the *likely* range assessed from other lines of evidence. These models produce warming that is plausible but must be considered a 'low likelihood high warming' future or storyline (see Hausfather et al. 2022) and shouldn't be over-sampled in an ensemble. Each group downscaled at least one 'hot model' to illustrate this possible future.

The CORDEX modelling is used along with analysis of the CMIP6 models themselves – including direct scaling of observations using the change signal from 9 GCMs, and the use of raw data from up to 35 CMIP6 models and 6 relevant large ensembles.

The selected approach of CORDEX as a core data source complemented by CMIP6 (including large ensembles) and Convection Permitting Modelling (CPM) has several advantages and disadvantages over other possible options. These other options include:

- Primary source of statistical calibration of CMIP6 (e.g., United States Fifth National Climate Assessment NCA5) - advantages of simplicity, breadth and international comparability, disadvantages of lacking potential depth and insights on extremes from dynamical downscaling.

- Use of a Perturbed Physics Ensemble (PPE) along with GCMs (e.g., United Kingdom Climate Projections 2009 and 2018, UKCP09 and UKCP18) - advantages of more fully exploring the range from physics schemes in models, disadvantages of PPE using a single model.
- Scenario-based storylines (e.g., National projections from the Royal Netherlands Meteorological Institute 2014, KNMI14) - advantage of distilling emissions and model ranges to simple, intuitive narrative-based scenarios for a given timeframe, not applicable to Australia due to our diverse climate and climate change projections.
- Presenting flexible tools to explore various model datasets for standard regions (e.g., IPCC Interactive Atlas) - advantage of providing a presentation and comparison of commonly used model sets, disadvantages of limited regionalisation and application-ready datasets.

The strategy for the new regional model ensemble outlined here is ambitious and a significant change from the modelling used to underpin Australian projections in CSIRO and BoM (2015). Using a multi-scenario multi-model ensemble (MME) of regional models following CORDEX guidelines, with multiple bias adjustment techniques applied to make the outputs application-ready and locally relevant, represents a further extension of the set of options explored by Ekström et al. (2015), reproduced below in Figure 2.1. Combining this with insights, and application-ready datasets, based on CMIP6 itself provides further support and options for different use cases.

	Ease of use	Climate realism (applicability)	Physical plausibility of change (credibility)			Potential to represent range of change	Example applications
			Spatial		Temporal		
GCM change factor method applied to fine res. obs.	Easy	Very good	May not be OK near complex topography and coastlines.	Direction of increase in approach capability	Change in mean only. OK for some applications, but not for others.	Follows GCMs considered. Strong ability to consider the full range of GCM uncertainty.	Suitable for applications not driven by rainfall extremes, and not in high topography. Appropriate for many agricultural and biodiversity applications.
As above, but enhanced methods (e.g., q-q mapping, weather generator)	Easy	Very good	May not be OK near complex topography and coastlines.		Change in variability. OK for many applications.	Follows GCMs considered. Strong ability to consider the full range of GCM uncertainty.	As above, but suitable also for hydrological applications.
Statistical downscaling (one method, multiple GCMs)	Easy	Possible bias and limited variables.	Good near high topography and coastlines.		Potentially good	Possible approach bias, but can be applied to a large number of GCMs. Strong ability to consider the full range of GCM uncertainty.	Good for many applications, but restricted by availability of relevant variables.
RCM change factor method applied to fine res. obs.	Easy	Very good	Good near complex topography.		Change in mean only. OK for some applicaitons, but not others.	Possible approach bias. Ability to consider GCM uncertainty is restricted due to number of GCMs used in the study.	Applications not driven by rainfall extremes, but can include complex topographical situations. Suitable for many agricultural and biodiversity applications.
Direct RCM (one RCM, multiple GCMs)	Medium	Biased	Good near complex topography.		Potentially good	Possible approach bias. Ability to consider GCM uncertainty is restricted due to number of GCMs used in the study.	Not recommended due to current climate biases.
Direct RCM (one RCM, multiple GCMs) with bias correction	Medium	Potentially good	Good near complex topography.		Potentially good	Possible approach bias. Ability to consider GCM uncertainty is restricted due to number of GCMs used in the study.	Suitable for many applications, including in regions with complex topography. However, climate change results may not be fully representative due to approach bias.

Figure 2.1 Summary of typical characteristics associated with downscaling methods discussed within this document. The first column denotes 'ease of use'. Note that this judgment refers to the downscaled dataset not to implementing the method. Blue coloured boxes indicate 'low complexity', 'high skill' of the listed methods, yellow indicate circumstances when methods are characterised by some limitations, and red indicates characteristics that present major challenges/restrictions in their applicability (reproduced from Figure 2 in Ekström et al. (2015)).

3. Model ensemble description

- **Climate change is examined under two Shared Socioeconomic Pathways that 'bracket' or 'bookend' the bulk of the likely range of future scenarios for greenhouse gas emissions (SSP1-2.6 and SSP3-7.0)**
- **The CORDEX-Australasia ensemble is used as a primary data source for updated projections.**
- **Strategic sampling of CMIP6 Global Climate Models into Regional Climate Models (RCMs) forms a multi-model 'sparse matrix' design.**
- **Two Regional Climate Models (RCMs) have been developed and run for the historical period and future scenarios to 2100 by the ACS and simulations from two other RCMs provided by the NSW and Queensland governments.**

The production of downscaled climate change projections must start with global greenhouse gas emissions scenarios, then follow a sequence of numerical global and regional model simulations, comparing these to other lines of evidence, through to climate services. Here we outline the choices used for updated projections, first outlining the scenarios used, the global model selection, the regional modelling, and choices to assess model performance. Regional modelling follows the Coordinated Regional Downscaling Experiment (CORDEX) guidelines.

3.1 Shared Socioeconomic Pathways (SSPs)

Beyond the near-term climate of the next 20 years, the amount of additional climate change depends predominantly on the human choices and pathways of societal development and resulting emissions. To explore future climate changes, we must take a scenario approach to addressing this uncertainty and illustrate the effects of various plausible trajectories of human influence, rather than a single pathway. It is useful to use internationally agreed future scenarios for comparability, so we adopt the current generation, the Shared Socioeconomic Pathways (SSPs, Meinshausen et al. 2019).

Regional modelling is very resource-intensive, and we lack the required resources or global model data inputs to downscale all models and all SSPs. Therefore, we don't use the full set of narrative-based scenarios in the SSPs, and instead follow CORDEX guidelines to 'bracket' or 'bookend' a range of plausible and policy-relevant SSPs and resulting climate change response using:

- **SSP1-2.6**: a net-zero pathway; this pathway is roughly compliant with meeting the Paris Agreement limiting global warming to less than 2 °C relative to preindustrial at 2100. We use this scenario as representative of the physical climate change response to this level of climate change, not only for the specific narrative behind SSP1 ('Sustainability').
- **SSP3-7.0:** a high scenario; at 2100 this pathway results in 2.8-4.6 °C global warming relative to preindustrial. We use this scenario as representative of higher levels of climate change, not specific to the narrative behind SSP3 ('Regional Rivalry'). This scenario is selected as a fallback from the extremely high SSP5-8.5 ('Fossil Fuelled Development'), which is seen as implausible in some respects by many (e.g. see Hausfather and Peters 2020).

The previous projections for Australia (CSIRO and BoM 2015) used the previous generation Representative Concentration Pathways (RCPs), with a strong focus on RCP8.5 (labelled 'very high' emissions) and RCP4.5 ('moderate' emissions), with a much smaller focus on RCP2.6 ('very low' emissions). The new projections for SSP1-2.6 and SSP3-7.0 are not directly comparable to projections presented for moderate and very high scenarios (RCP4.5 and RCP8.5) so users must adjust their frame of reference on this issue of context.

Also note that a limited number of simulations for SSP5-8.5 and SSP2-4.5 are made for some tailored applications and for the purpose of back-comparison but are not presented here. If these simulations are examined, note that there is not a direct mapping between these SSPs and RCPs 8.5 and 4.5 (e.g., SSP2-4.5 is not identical to RCP4.5 in terms of the mix of different greenhouse gases, and the trajectory of anthropogenic aerosols). The similarity of the level of extra greenhouse effect at 2100 (the last number in each code) is used to only as a rough back-comparison from CMIP6 to CMIP5 simulations.

3.2 CMIP6 modelling

The Coupled Model Intercomparison Project phase 6 (CMIP6, Eyring et al. 2016) includes the CMIP defined Historical simulations (1850-2014), and for the SSPs (2015-2100), with extensions to 2300 for some models. These simulations are core inputs into projections analysis for the IPCC Sixth Assessment Report (IPCC 2021-2023) and for various national and regional projections work in Australian and elsewhere. An assessment for Australia showed CMIP6 produced incremental improvement in the simulation of the historical climate, and only minor differences in the projected change 'signal' compared to CMIP5 but offered additional regional insights in the spatial detail of that signal (Grose et al. 2020). This assessment was led by staff in the ACS working with the Australian climate science community and is a core resource for these new projections.

The CMIP6 ensemble of simulations of future scenarios (known as ScenarioMIP) contains up to 50 distinct GCM simulations for the Historical experiment, and up to 35 simulations with daily data for the required SSPs. However, only 18 models provided the sub-daily data required for RCMs at the time of analysis, which was a main constraint on the selection for regional modelling used here. While GCMs do produce sub-daily data the very large size of these datasets and the costs of storage and data transfer means these are not always saved and shared.

Models from this subset of GCMs with data available were selected considering aspects of 1) model evaluation, 2) model independence and 3) representativeness in the projected change signal, aiming for a selection of adequate, independent models that sample the range of CMIP6 itself. Evaluation included statistics of the mean climate, climate processes and teleconnections to Australian climate, choosing to reject unsuitable models rather than select the best performers. Model independence was considered through selecting models above a threshold of similarity identified in the literature, aiming to include models that are not highly related. Representativeness considered the projection of temperature, rainfall and some key circulation features, aiming to sample the spread in the CMIP6 ensemble. See Grose et al. (2023a) for full details of the selection.

As described in Chapter 2, there was a special consideration of climate sensitivity and sampling of so-called 'hot models.

Computational limitations prevent us from downscaling all ensemble members of all models, so the ensemble is not large enough to be truly comprehensive and balanced. Therefore, we take a 'representative climate futures' approach, aiming to include at least one representative model simulation from each one of the key categories of plausible change in a two-dimensional phase space (where categories are user-defined), rather than prioritising the ensemble mean and spread in any one variable, see Whetton et al. (2012) for more on the concept. Here we aim to include at least one representative model simulation from each category of key projected change in mean warming and rainfall change (e.g., wetter or drier, warmer or hotter).

Models were selected by partners the New South Wales and Australian Regional Climate projections project version 2 (NARCLIM2.0; Di Virgilio et al. 2025) and the Queensland Future Climate Platform version 2 (QldFCP-2) using a similar process, but with some minor differences in the evaluation statistics or measures of representativeness. The selection for the whole ensemble is summarised in Grose et al. (2023a), but due to subsequent difficulties with data availability, the final model lists differ a little from the list in the paper and are shown in Table 3.1. The table highlights some of the key projections chosen as 'representative climate futures' by the selection exercises and flags the 'hot models' (climate sensitivity above the assessed likely range, but not necessarily an extreme outlier in warming for Australia, these are marked as a 'much hotter' climate future).

Table 3.1 The 'sparse matrix' of CMIP6 Global Climate Model (GCM) simulations and Regional Climate Model (RCM) simulations under CORDEX-CMIP6 that form a core dataset for new national climate projections. For RCM descriptions, see details in the following sections. GCMs highlighted in orange did not provide sub-daily data, so could not be downscaled by ACS or NARCLIM2.0. Also noted are key representative climate futures that some models represent, and models considered 'hot' (a climate sensitivity above the likely range).

	QDC scaled from GCM	BARPA-R	CCAM-ACS (v2203)	QldFCP-2 CCAM	QldFCP-2 CCAM ocean coupled	NARCLIM2.0 - WRF412 R3 and R5	Notes: Key Future, hot model, analysis of large ensemble
ACCESS-CM2_r2i1p1f1					X		Much hotter (Hot model)
ACCESS-CM2_r4i1p1f1	X	X	X				Much hotter (Hot model)
ACCESS-ESM1-5_r6i1p1f1	X	X	X	X		X	Hotter and much drier, analysis of large ensemble
ACCESS-ESM1-5_r20i1p1f1					X		As above
ACCESS-ESM1-5_r40i1p1f1					X		As above
CESM2_r11i1p1f1	X	X	X				Hot model
CMCC-ESM2_r1i1p1f1	X	X	X	X			
CNRM-CM6-1-HR_r1i1p1f2				X	X		Hot model
CNRM-ESM2-1_r1i1p1f2			X				Hot model
EC-Earth3_r1i1p1f1	X	X	X	X			Wetter and more variable, analysis of large ensemble
EC-Earth3-Veg_r1i1p1f1						X	Wetter and more variable
FGOALS-g3_r4i1p1f1				X			
GFDL-ESM4_r1i1p1f1				X			
GISS-E2-1-G_r2i1p1f2				X			
MPI-ESM1-2-HR_r1i1p1f1	X	X				X	
MPI-ESM1-2-LR_r9i1p1f1				X			
MRI-ESM2-0_r1i1p1f1				X			
NorESM2-MM_r1i1p1f1	X	X	X	X	X	X	Cooler end of warming
UKESM1-0-LL_r1i1p1f2	X					X	Much hotter (hot model)

A representative sub-set of nine GCM simulations were used in the other method of producing application-ready data directly from GCMs, that of scaling of observations (noted in Table 3.1). Here the quantile delta change (QDC) method was used, as outlined in Irving and Macadam (2024). This QDC dataset is separate from CORDEX and is a useful resource for applications where this method is better suited (see Chapter 5).

Six models provided large ensembles of 10 or more simulations for an SSP. These are models with numerous simulations using perturbed starting conditions, allowing an examination of internal variability along with the climate change response of the model. We make particular use of two large ensembles as important context for the two key 'representative climate futures' of rainfall change: ACCESS-ESM1.5 (projection of strong drying, 40 members) and EC-Earth3 (projection of greater and more variable rainfall, 58 members). Here we derive insights about the nature of the projected change in rainfall from the two large ensembles to help provide the narrative or 'storyline' behind these important representative climate futures even when only the CORDEX simulations are presented quantitatively (see Section 7.4).

3.3 CORDEX

CORDEX provides a protocol for undertaking downscaling simulations for the purpose of comparing techniques and models at an international scale. Dynamical, statistical and machine learning based methods are all included within CORDEX. CORDEX is part of the World Climate Research Program (WCRP) and is a diagnostic MIP of CMIP6. There are currently 14 recognized CORDEX domains, including Australasia (see Figure 3.1).

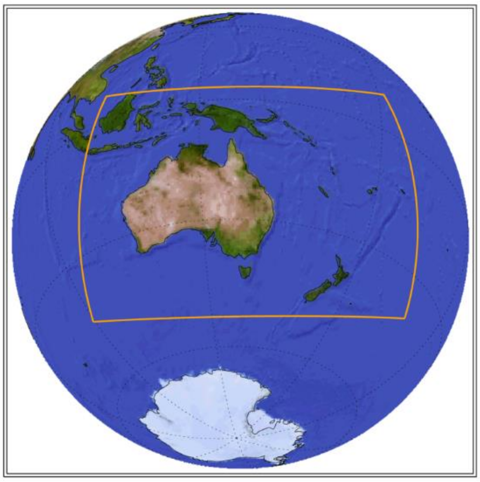

Figure 3.1 Outline of the CORDEX Australasia domain including Australia.

CORDEX provides guidance on the downscaling experiment design, suggesting spatial resolutions between 12.5 km and 25 km resolution, designated 'AUS-11' and 'AUS-22' respectively for Australasia. However, the modelling groups in Australia have used different simulation resolutions than this specification, so data are all available on different native grids, then interpolated to one of 'AUS-11i', 'AUS-20i' or 'AUS-22i'. CORDEX Australasia has also defined an AUS-20 grid at approximately 22.5 km resolution for the purposes of intercomparison of the different downscaling models.

CORDEX provides a strict standard for model output, defining variable names, units, output frequency and directory structure for the purpose of comparing different models in a standard way. Regional climate models then output their data according to this Data Reference Syntax (DRS), ensuring our output is compatible with software evaluation tools and can be readily imported by next-user models including climate hazard models (e.g., cyclones, coastal, hydrology, bushfire, etc.).

By adopting the CORDEX experiment design and output archiving specifications, the Australian Climate Service is adopting international best practice in regional climate modelling and making reuse of its modelling outputs easier. The results can also be evaluated and benchmarked by the international scientific community which helps to establish the quality of the climate simulations and the credibility of the climate projections.

The two RCMs used by ACS to conduct downscaling are BARPA-R (Section 3.3) and CCAM-ACS (Section 3.4). In parallel to the BARPA-R and CCAM-ACS downscaled projections there are two other nationally relevant dynamically downscaled projections ensembles that will be available as part of CORDEX-Australasia. The ACS regional modelling experiments were in fact designed to complement these other sets of simulations, to provide users with a sample of GCM and RCM combinations of regional projections that would help users address the range of plausible projection uncertainty (Grose et al. 2023a). This collaborative approach between the states is a cornerstone of the Australian Climate Service and underpins the National Partnership for Climate Projections (NPCP).

All CORDEX participants first run simulations to assess their climate given 'perfect' driving input sourced from ERA5 (Hersbach et al. 2020), then using the selected GCM inputs. The models contributing to CORDEX Australasia are run for at least the CORDEX-Australasia domain and are nested directly from the reanalysis of GCM in one step. All models are regional models, not Convection Permitting Models (CPM), however some subsequent CPM simulations were then nested within CORDEX simulations for further work. The models differed in the type of grid, spatial resolution, connection to the GCM atmosphere and in correction of inputs (Table 3.2).

Table 3.2 Basic details of the four Regional Climate Model (RCM) simulations contributing to CORDEX-Australasia from Australian groups. For acronym definitions and further details, see following sections.

Program	Modelling system	Grid	Spatial resolution over Australasia	Atmosphere guided by GCM (e.g., nudging)	Correction of inputs	Ocean coupling
Australian Climate Service (ACS)	CCAM	Global variable resolution	~10 km	Yes	No	Some
Australian Climate Service (ACS)	BARPA-R	Limited Area Model	~15 km	Yes	No	None
NARCLIM2.0	WRF	Limited Area Model	~18 km	Yes	No	None
Qld-FCP2	CCAM	Global variable resolution	~10 km	No	Yes – sea surface temperatures	An OC version

3.4 BARPA-R

The Bureau of Meteorology employs the Australian Community Climate and Earth-System Simulator (ACCESS) model for research, weather forecasting and seasonal prediction. ACCESS combines the UK Met Office (UKMO) Unified Model for the atmosphere, coupled to the Joint UK Land Environment Simulator (JULES) land surface model. As part of the Bureau's strategy to use a consistent modelling framework across various timescales – from historical analysis to nowcasting, multi-day forecasts, seasonal forecasts and multi-decadal projections - ACCESS was developed for regional climate modelling through the framework of BARPA (Bureau of Meteorology Regional Projections for Australia). Regional ocean modelling is not considered for BARPA at this stage.

Projections are produced using the BARPA-R (regional) configuration, based on physics modelling at horizontal grid spacings greater than 10 km (note a convective-scale version, BARPA-C, has also been developed for research purposes).

The scientific configuration of BARPA-R draws from the UKMO Hadley Centre HadREM4-GA7-05 configuration used in the UK Climate Projections program (Tucker et al., 2021). Further enhancements were made to Australian land ancillary files, advection and convection schemes as part of the ACS (Su et al., 2022). BARPA-R covers the CORDEX-Australasian domain (Figure 3.1) and is constrained by the dynamics of driving global climate simulations through forcing with sea-surface temperature and lateral boundary conditions, and grid-scale nudging of temperatures and winds above 11 km. Greenhouse gas concentrations and aerosol forcing from major historical volcanic eruptions are prescribed, and influence of aerosol radiation and cloud effects is approximated through the EasyAerosol Scheme of Stevens et al. (2017). Changes in land use and land cover are not considered, remaining fixed at current conditions, given the large uncertainties in what future land cover changes will be (following CORDEX design recommendations). Research and trials are performed using the BARPA-C model for an Australian domain (Figure 3.2) but are not used in these projections. For more details on the model descriptions, refer to Su et al. (2022), Stassen et al. (2023) and Howard et al. (2024).

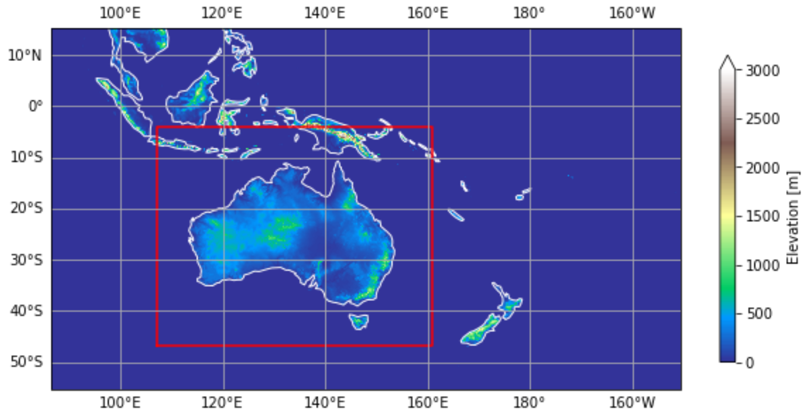

Figure 3.2 Modelling domains for BARPA-R and research trials for BARPA-C (red).

The BARPA-R downscaling approach was further tested for each selected CMIP6 (Sec. 3.1) model and a CMIP6 model was only downscaled if BARPA-R preliminary runs produced a satisfactory scorecard (Su et al. 2022). The scorecard assesses various aspects of model performance, such as the number of simulated east coast lows and tropical cyclones near Australia, temperature and precipitation biases, change signal differences between BARPA-R and global models for temperature and precipitation and temperature-precipitation correlation over Australia. Through this approach, CNRM-ESM2-1 was rejected due to pervasive cold biases in daily maximum temperature owing to upper-tropospheric temperature biases in the global model, and MPI-ESM1-2-HR was selected as a replacement. Consequently, BARPA-R provides downscaled projections for seven global model experiments (Table 3.1) for historical (1960s

014), SSP1-2.6 and SSP3-7.0 (2015-2100). ACCESS-CM2 and EC-Earth3 are also downscaled for SSP5-8.5 to aid the transition from CMIP5 to CMIP6 downscaling for impact modelling.

3.5 CCAM-ACS

The Conformal Cubic Atmosphere Model (CCAM) was developed by CSIRO and its partners for regional climate and weather research (McGregor 2005). Unlike most other regional climate models, CCAM employs a global variable resolution cubic grid that allows for the simulation to focus higher spatial resolution over a region while still maintaining a lower resolution simulation of the rest of the globe (see Figure 3.3). In this way, CCAM does not require lateral boundary conditions from the host GCM. Although CCAM is strictly a global model (a variable resolution Global Climate Model, VR-GCM), we apply the term ‘regional climate model’ for simplicity.

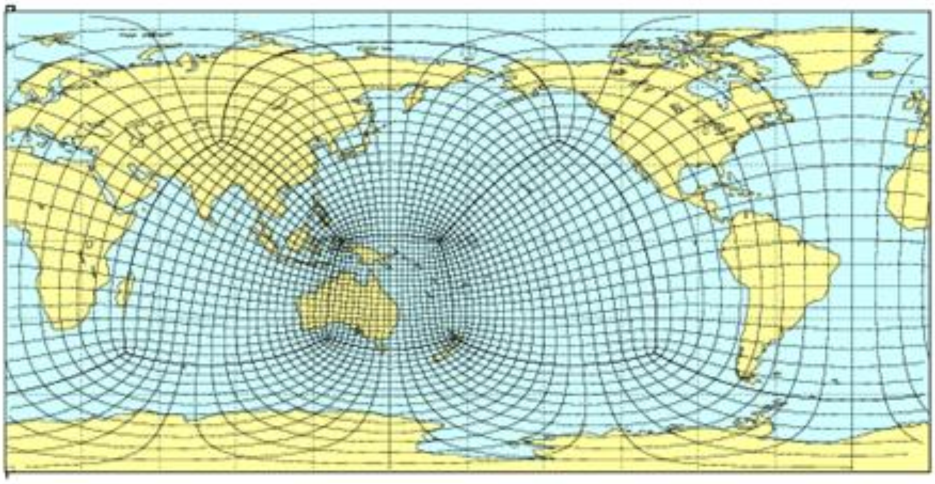

Figure 3.3 Example of the variable resolution cubic grid used by CCAM for the CORDEX Australasia grid. For clarity, not every grid point is shown.

CCAM supports a modern set of atmospheric physics parameterisations that can adapt to the changing CCAM grid resolution. CCAM also includes the CABLE land-surface model also used in the ACCESS GCMs contributing to CMIP6, as well as prognostic aerosols (due to their impact on precipitation) and an ocean model (to improve coastal regions and some extreme weather events). Land-use for crops and grazing regions was allowed to change over time as prescribed under SSPs for CMIP6, essentially representing land clearing otherwise land use was fixed following CORDEX guidelines.

Since CCAM has no lateral boundaries, it can instead optionally use a spectral nudging approach (used for CCAM-ACS) or use a sea surface temperature (SST)-only forcing approach (as used for QldFCP-2) in which the SST and sea-ice concentration drives the simulations. In the case of the spectral nudging (Thatcher and McGregor 2009), the atmosphere is constrained to agree with the host GCM at larger wavelengths (i.e., greater than 3,000 km) for air temperature, winds and surface pressure. This allows CCAM to follow the host GCM at large spatial scales but can add regional detail at smaller scales. CCAM's humidity is not directly nudged, giving more freedom for CCAM to simulate the rainfall according to its atmosphere physical parameterisations. This spectral nudging approach can ensure stronger agreement between CCAM and the host GCM (e.g., cyclones appear close to the location they would form in the host GCM) but also is more prone to including GCM errors and biases.

The CCAM downscaling of ERA5 has been evaluated in Schroeter et al. (2024), which evaluated temperature and precipitation performance, including extreme temperatures and precipitation. The study found CCAM was able to provide added value for temperature (particularly southern Australia), although tended to overestimate extreme rainfall. Nevertheless, CCAM performance was consistent with that expected from a modern regional climate model, so is not rejected for use here. CCAM has downscaled historical, SSP1-2.6 and SSP3-7.0 simulations of seven CMIP6 GCMs (see Table 3.1).

3.6 NARCliM2.0

The NARCliM2.0 (New South Wales and Australian Regional Climate Modelling Phase 2.0) program is delivered by the New South Wales State government. NARCliM2.0 uses the Weather Research and Forecasting (WRF) model to dynamically downscale CMIP6 GCM projections. NARCLiM2.0 uses a limited area model configuration for Australasia with a 20 km resolution following CORDEX guidelines. Further convection permitting modelling at 4 km resolution was performed for an inner domain over southeast Australia, and emerging simulations for southwest and northwest Australia (projections for these sub-domains are not discussed in this report, as we have a focus only on projections with national coverage).

Convection is parameterised for the 20 km resolution simulations and explicit for the smaller 4 km resolution inner domain, and both scales did not include an ocean model. The NARCliM2.0 GCM selection can be found listed in Table 3.1 above. Two configurations of the WRF model are used with each of the selected GCMs (labelled R3 and R5), with the configurations selected following a thorough assessment of many options. Further details on the NARCliM2.0 ensemble are provided in Di Virgilio et al. (2022), model evaluation and future projections are described in Di Virgilio et al. (2025).

3.7 Qld-FCP-2

The Queensland Future Climate Platform version 2 (QldFCP-2) is funded by the Queensland State government and uses the CCAM model to dynamically downscale CMIP6 GCM projections globally on a stretched grid, but with a high-resolution (approximately 10 km resolution) domain centred over Australia (Thatcher et al. 2015). While the base model system (CCAM) is the same as CCAM-ACS, these simulations were run using a very different experimental set-up and configuration compared to CCAM-ACS, hence they are treated as a separate RCM ensemble. The main difference in QldFCP-2, which also sets it apart from all the other regional models for CORDEX-Australasia, is that the GCM SSTs used as boundary conditions for the regional model are bias-adjusted against observations prior to running the downscaling simulations. The QldFCP-2 conducted a set of atmosphere-only experiments as well as a smaller set of atmosphere-ocean coupled experiments. For the atmosphere-ocean coupled experiments, CCAM SSTs were spectrally nudged towards bias adjusted GCM SSTs. For the atmosphere-only experiments CCAM was run using prescribed bias adjusted SSTs and no atmospheric nudging. Details of the QldFCP-2 model evaluation can be found in Chapman et al. (2023), and a description of QldFCP-2 projected changes can be found in Chapman et al. (2024).

3.8 Regionalisation

For evaluation, benchmarking and general reporting of projections in this report we use the eight ‘clusters’ of Natural Resource Management (NRM) regions used previously for national projections (CSIRO and BoM, 2015), shown in Figure 3.4. This regionalisation is used for back-comparison to previous work, and because the regions are quite climatologically meaningful in terms of the current climate and signals of projected change (Fiddes et al. 2021). We also use the four ‘super clusters’ (Figure 3.5), broadly similar to the four Australian IPCC reference regions of Eastern

Australia (EAU), Southern Australia (SAU), Northern Australia (NAU) and Central Australia (CAU), presented in Iturbide et al. (2020). For some analyses we also use the 15 'subclusters' (Figure 3.6) for further granularity.

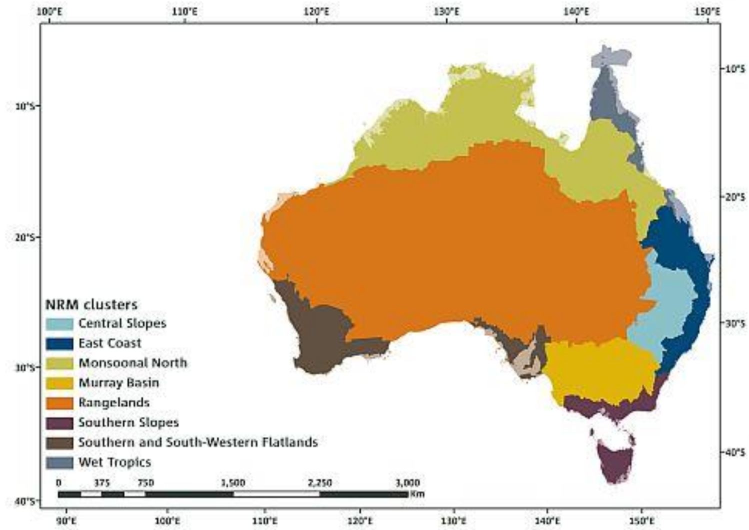

Figure 3.4. The eight NRM 'cluster' regions – abbreviated codes are (in order): CS, EC, MN, MB, R, SS, SSWF and WT. Figure 3-4 to 3.6 reproduced from www.climatechangeinaustralia.gov.au

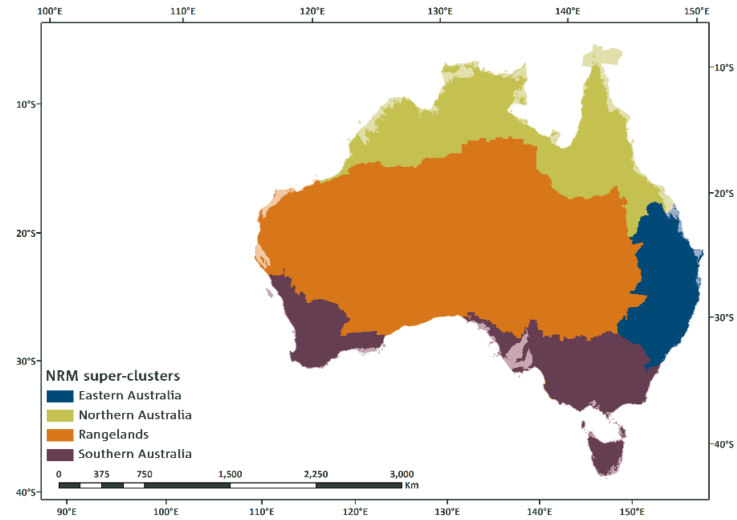

Figure 3.5. The NRM 'Supercluster' regions used for area averages in some analyses – abbreviated codes are Eastern Australia (EA), Northern Australia (NA), Southern Australia (SA) and Range lands (R).

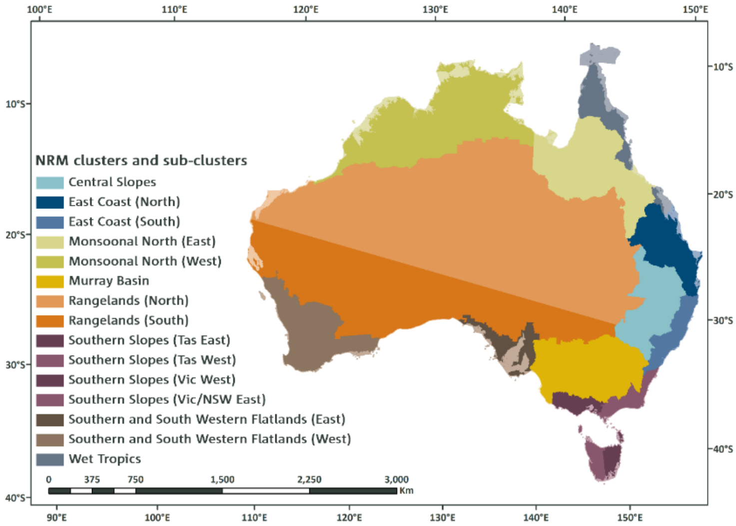

Figure 3.6. Natural Resource Management (NRM) sub-cluster regions of Australia. Abbreviated codes are (in order): CS, ECS, ECN, MNE, MNW, MB, RN, RS, SSTE, SSTW, SSVW, SSVE, SSWFE, SSWFW and WT.

4. Evaluation and benchmarking of the ACS and CORDEX ensemble

- **Regional model simulations from the CORDEX-Australasia ensemble are evaluated against observation-based datasets and benchmarked against their global driving GCM.**
- **No CORDEX Australasia ensemble members have been identified for exclusion through the benchmarking process.**
- **Added value is assessed for the CORDEX Australasia ensemble, and both positive and negative added value is found at the regional scale for the variables analysed.**
- **Through a simple model independence assessment, we do not find sufficient grounds to justify model weighting across the board but suggest further research could reveal useful weighting for targeted applications.**
- **Any decision on whether data within the CORDEX Australasia ensemble is fit for purpose will ultimately be application specific, and therefore the responsibility of the users.**

In this Chapter, we summarise the evaluation and benchmarking of the simulations used in the projections (building on the evaluation during model development described in Chapter 3). We cover evaluation of BARPA-R and CCAM-ACS, benchmarking of the entire CORDEX ensemble (including ACS but also NARCLIM2.0 and QldFCP-2 simulations), assessment of added value from BARPA-R and CCAM-ACS, and ensemble independence across CORDEX. We follow a two-pronged approach of first evaluating the downscaling models and then benchmarking the ACS ensemble. We define evaluation here as the comparison of models against observations, and benchmarking as the comparison of models against some objective standard; further details can be found below and in Jiang et al. (in review).

First, the ability of CCAM-ACS and BARPA to reproduce Australian climate given 'perfect' driving input sourced from ERA5 (Hersbach et al. 2020) was assessed separately for each downscaling model by Schroeter et al. (2024) and Howard et al. (2024), respectively. This evaluation is done to understand the RCM's ability to simulate features of the present-day climate, including temperature, rainfall, wind and pressure fields. If an RCM were to perform poorly at this first stage, confidence in its capability would be reduced and the RCM may be excluded when selecting RCMs for downscaling GCMs or examining future projections.

The second phase was to make progress on the goal of truly 'benchmarking' the performance of the CORDEX simulations of rainfall and temperature in the historical climate. Benchmarking differs from evaluation in that it aims to apply an objective framework to assess model simulation against objective and pre-determined standards and thresholds. The benchmarking framework was applied to the full CORDEX Australasia ensemble in a collaboration led by the ACS (Jiang et al, in review), but falling back on some subjective threshold benchmarks in the absence of fully developed objective thresholds. The benchmarking framework can also assess a model simulation compared to another, an RCM to the host model or a new version to a previous version of a model. For example, benchmark threshold values are calculated from the downscaled GCMs which can then be used to determine if an RCM has improved in its simulation of Australian climate relative to the GCM.

The added value and realised added value from downscaling the CMIP6 GCMs is examined for both the historical and future climate. This method involves comparing the differences between GCMs and observations against the differences between RCMs and observations to determine where downscaling has ‘added value’ or provided more information, relative to its driving GCM. Generally, RCMs add value along coastlines, areas with varying orography (e.g., mountain ranges) and around cities, as well as generally improved representation of relevant smaller scale processes - e.g. sea breezes and mountain winds, where the higher resolution provides increased information. Realised added value combines added value with the climate change signal difference between the GCM and RCM, identifying regions and variables where RCMs add value in the simulation of the historical climate and have a different climate change signal projection. This can reveal where RCMs may have plausible improvements in future projections through downscaling.

These three methods, evaluation, benchmarking, and added value, aim to assess RCM performance against observations and the driving GCMs but it is important to recognise that models are not always independent due to similar components (e.g., some GCMs share the same atmosphere model, ocean model, or coupler). Therefore, the independence between the driving CMIP6 GCMs, and additionally, the independence between the downscaled CORDEX RCMs must be assessed. This may help to determine if weighting should be applied when analysing projections.

4.1 Evaluating the ACS RCMs

For evaluation of NARCliM2.0, see DiVirgilio et al. (2025) and Ji et al. (2024). For evaluation of Qld-FCP2, see Chapman et al. (2023).

4.1.1 BARPA-R

The evaluation of BARPA-R’s downscaling of ERA5 was performed against the Australian Gridded Climate Dataset (AGCD; Jones et al. 2009) for daily maximum temperature, minimum temperature, and precipitation along with associated indices, such as maximum daily precipitation and heavy rain days. The 10-metre winds were assessed through aggregated station-based comparisons. Model performance was found to be generally on par with, or improved, when compared to the driving ERA5 model, despite the lack of data assimilation in BARPA-R compared to ERA5. Noticeable biases in BARPA-R are maximum temperature in austral winter (June to August or JJA), which is colder than ERA5 over central to southern Australia, and a wet bias over most of Australia in austral summer (December to February or DJF). These results are reported in Howard et al. (2024).

BARPA-R exhibits maximum temperature trends which are consistent with observations, however the minimum temperature warming trend is overestimated over the whole of Australia and precipitation wetting trends are underestimated over northern Australia. It was also found that downscaling through BARPA-R improved 10-metre winds in six out of the eight regions analysed; the poorer results are found over Monsoonal North and central Australian rangelands. BARPA-R generally shows improved high percentile tail values compared to ERA5, while both models underestimate ‘calm’ weather conditions with wind speeds of 0 m s^{-1}. Temperature and precipitation teleconnections of the El Niño–Southern Oscillation (ENSO), the Indian Ocean Dipole

(IOD), and the Southern Annular Mode (SAM) were well represented in the RCM when present in the driving boundary conditions. Also, large-scale atmospheric circulation features showed varied performance depending on latitude. Features such as the South Pacific Convergence Zone, northwest cloud bands and monsoon westerlies exhibited internal interannual variability and larger divergence from observations than mid-latitude features such as westerly jets and extratropical cyclones.

4.1.2 CCAM-ACS

Schroeter et al. (2024) assessed CCAM-ACS's performance in downscaling ERA5 with a focus on daily minimum temperature (tasmin), daily maximum temperature (tasmax), and daily precipitation (pr). Comparing CCAM-ACS downscaled ERA5 to AGCD (Jones et al. 2009), CCAM-ACS simulates reduced warm biases in tasmin but exhibits a strong cold bias over northern Australia in summer (DJF) which may be due to overestimation of high-level cloud. Daily precipitation in CCAM-ACS exhibits a consistent wet bias over most of Australia that is driven by an overestimation of extreme rainfall events. CCAM-ACS captures the seasonal cycle of tasmin, tasmax, and daily precipitation well.

In terms of added value, CCAM-ACS downscaling shows mixed results. For example, annual maximum tasmax over southern Australia is improved but value is lost over northern Australia due to the cold bias. Similarly, 1-day and 5-day annual maximum precipitation amounts (Rx1day, Rx5day) show added value along the coast while inland central and eastern Australia lose value due to the wet bias. Analysis of extremes, based on 5% Annual Exceedance Probability (AEP5%) also gives mixed results. CCAM-ACS adds value to extreme tasmin over western and inland central Australia but there is reduced value over eastern and south-eastern Australia.

4.2 Towards benchmarking

Traditional evaluation gives a general depiction of the fidelity of a climate simulation compared to observations and this can be used in a general sense to help inform tools central to climate projections such as confidence ratings. This is useful but still involves subjective expert judgments and can lack the objective basis for making quantitative decisions such as weighting models (including rejection or assigning a weighting of zero). Therefore, there is now interest in applying a benchmarking framework to model assessment to bring in more objectivity. Ideally, an objective set of benchmarks should be used, and passing them should indicate that the model is suitable for a given purpose. As such, benchmarks should relate to a meaningful threshold with a physical justification and should be set *a priori* (before results are examined). Meaningful, physically based thresholds can be hard to determine, and the benchmarking area is still developing. So, subjective expert judgments are used in setting thresholds here, and the thresholds are mainly statistically based rather than physically meaningful. However, the thresholds were set *a priori*, and we illustrate the process of applying benchmarks to demonstrate the principle, even if the methods are not yet well developed.

Benchmarking can also be used to objectively quantify model performance in relation to other aspects of interest, such as the performance of one model compared to another or to a previous version. Assessment of the rainfall and temperature climatologies across the ACS and the wider

CORDEX ensembles has followed a methodology informed by the benchmarking framework presented by Isphording et al. (2024) and the ILAMB project (Collier et al. 2018). This methodology defines 'benchmarks' as objective thresholds for model performance metrics, which provide a reference perspective as to whether the metric is well simulated by each ensemble member, or whether performance could ideally be improved.

Benchmarks have been defined and assessed for mean-state biases, spatial patterns, seasonal cycles, temporal distributions of spatially aggregated daily timeseries, long-term trends and extreme indices. Spatial pattern and seasonal cycle benchmarks are drawn from previous studies. Long-term trends are benchmarked relative to the observed trends, while the bias, temporal distribution, and extreme indices are benchmarked by comparing to the driving GCM ensemble mean. In all cases, the observational reference is drawn from AGCD, and an estimate of observational uncertainty is drawn from comparison between AGCD and two other observational datasets - GPM-IMERG (Huffman et al. 2014) and HadCRUT (Morice et al. 2021). We have no objective threshold for the level of test fails to reject models, so use an arbitrary cutoff of failing the majority of tests in all categories.

Figure 4.1 provides the climatological bias benchmark results for BARPA-R and CCAM along with NARCliM2.0 and QFCSP. This benchmark assesses NRM super-cluster (see Figure 3.5 for regions) mean precipitation, tasmax and tasmin biases in summer (DJF) and winter (JJA). Four super-clusters and 34 models makes for 136 tests for each season. Precipitation biases are presented as percentages of the AGCD-based climatological values. The benchmark value is derived from the mean of the CMIP6 GCM ensemble downscaled by CORDEX-Australasia across all modelling groups and requires each downscaled ensemble member to exhibit biases that are no larger than the mean magnitude of bias in CMIP6 GCMs selected for downscaling. The direction of bias does not contribute to the benchmark score but is indicated on the figure using diverging colours.

Figure 4.1 demonstrates that the benchmark is most frequently met by tasmin, and least frequently met by tasmax with only 47% of the RCM ensemble members passing the tasmax benchmark in summer (DJF) and this improves to 72.7% in winter (JJA). The GCMs are consistently too warm when simulating tasmin and thus a large majority of the RCM members improve on the GCM performance and meet the benchmark. For precipitation, approximately 69% of the RCM ensemble members pass the benchmark with 33 simulations (24%) exceeding the summer (DJF) precipitation benchmark and 51 members (38%) exceeding the benchmark in winter. The members failing tests are not excluded from the ensemble here, as we lack the firm objective basis to confidently rule them out. However, the effect of weighting and rejecting based on these tests will be followed up in further research work.

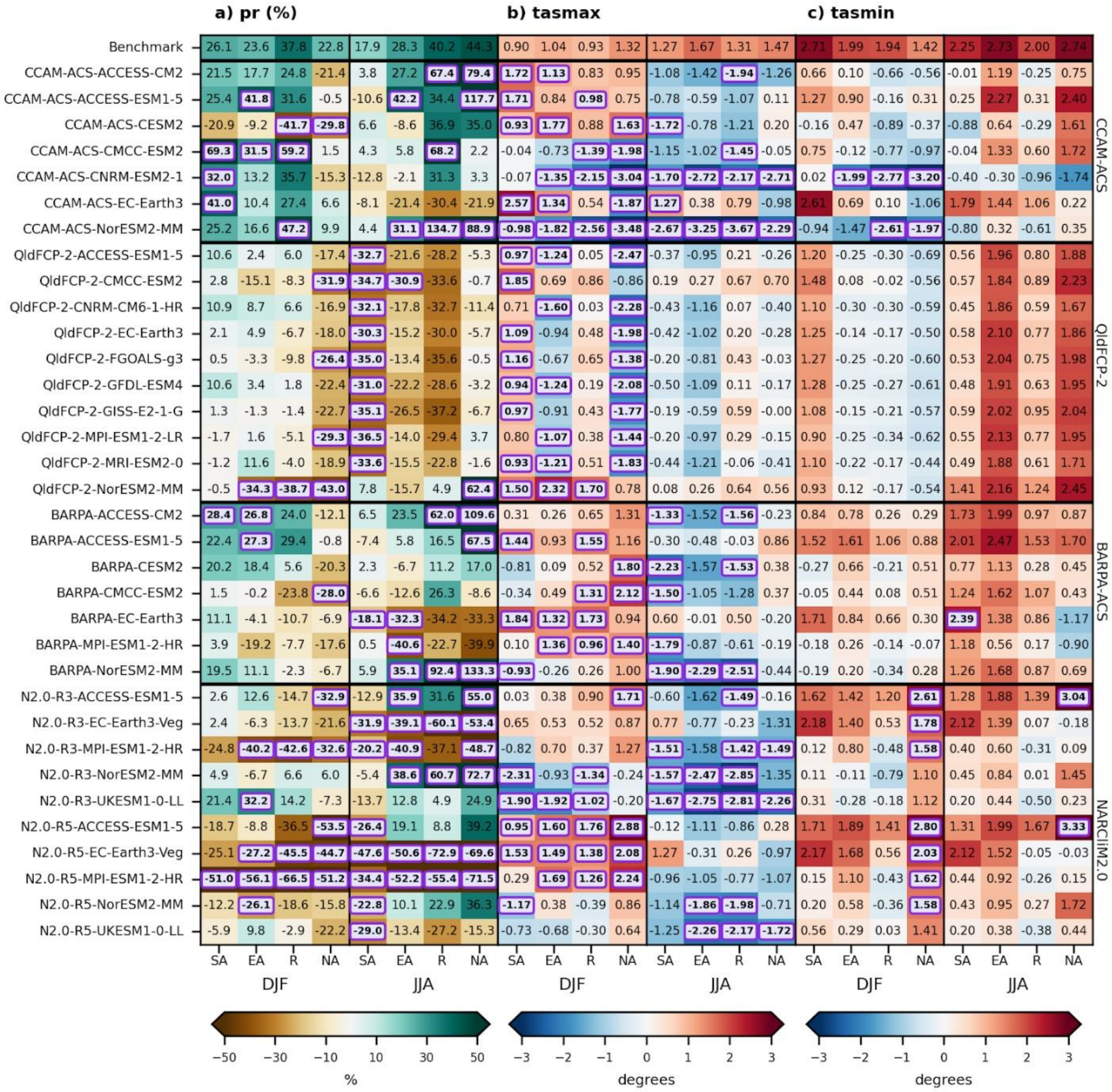

Figure 4.1 Aggregated bias benchmark for each core variable, season and GCM-RCM pair. Colours and numbers indicate the aggregated bias. The benchmark value is given in the top row. Biases larger than the benchmark are marked by purple squares. a: pr, b: tasmax, c: tasmin. Precipitation biases are presented as a percentage of the observed seasonal precipitation, while temperature biases are presented in degrees Celsius. Here, BARPA refers to BARPA-R, QldFCP-2 refers to QFCSP and N2.0 refers to NARCliM2.0. Figure reproduced from Jiang et al. (submitted).

Long-term trends from 1960 to 2014 (from the start of CORDEX simulations to the end of the historical simulation) are benchmarked against observations based on the Thiel-Sen slope estimator (Theil 1950, Sen 1968) and confidence intervals at a significance level of $p<0.05$. Note that this is again a subjective test and threshold, applied in the absence of a physically based threshold. The trend of an RCM simulation is considered to have passed the benchmark if the confidence intervals of both the RCM simulation and observations overlap. Due to the observed rainfall trends being highly seasonal, three NRM subclusters (Murray Basin, Southern and South-Western Flatlands (West), and Monsoonal North (West)) are assessed instead of the larger NRM superclusters (see Section 3.6 for region descriptions). Over these three subclusters, many RCMs are unable to capture the strength of the observed drying trend over the Southern and South-Western Flatlands (West) and this may be forced by the driving GCMs. In contrast, the RCMs appear to have more freedom to deviate from the driving GCM in the Monsoonal North (West) which may be due to internal variability of the Australian monsoon. Over the Murray Basin there is large uncertainty in the trend signal and thus all RCMs pass the benchmark.

Over the NRM superclusters, many of the RCMs and GCMs have minimum temperature trends larger than observed whereas for maximum temperature the trends are mostly positive and spread around the observations. In Eastern Australia, many of the RCMs diverge from the driving GCM which may be due to better simulation of the Great Dividing Range. Overall, 93% of the RCMs meet the maximum temperature trend benchmark and 89.8% meet the minimum temperature trend benchmark.

Performance across the remaining benchmarks is generally high across the CORDEX ensemble, with some exceptions. All RCMs pass the spatial correlation benchmarks, and most RCMs pass the seasonal cycle benchmarks, with the key exception being QldFCP-2 in southern Australia, where summers are simulated to be overly wet. Temporal distribution and extreme index benchmarks show the poorest performance, with 65% and 71% of models meeting the benchmarks, respectively. We note that these benchmarks also carry a large degree of observational uncertainty as well as the limitation of sampling short 20-year periods in very variable climate regions.

Overall, no CORDEX Australasia ensemble members have been identified for exclusion through the benchmarking process, as none fail the majority of tests in all categories. A few simulations fail all tests for a category (e.g., CCAM-ACS NorESM1-MM for tasmax, NARCLIM2.0 MPI-ESM1-2-HR for pr), or else more than ten tests across all categories (e.g., CCAM-ACS CNRM-ESM2-1 and NARCLIM2.0 EC-Earth3-Veg). Some RCMs fail the benchmark for a region and season across all GCM inputs, e.g., Qld-FCP2 for Southern Australian rainfall in winter (JJA). Further work should investigate the effect of model weighting or rejection based on benchmarking.

Special Box - ILAMB Dashboard

In addition to guiding the selection of benchmarks described in the previous section, the ILAMB methodology of Collier et al. (2018) has been adapted to provide an internal tool for science-literate data users to access evaluation results across a wide range of atmospheric variables. This dashboard provides a comparison between gridded observational datasets and the ACS ensemble members and includes the downscaled CMIP6 models as a reference. Maps of mean states, bias, root mean square errors and timing of annual maxima, and aggregated timeseries and annual cycles for 77 monthly variables may be accessed through an interactive html website. These 77 variables include 34 climate indices generated using the ICCLIM package. Observational references are sourced from satellite data, ERA5 reanalysis, and other gridded products including AGCD. Aggregation is available at NRM cluster and subcluster scale. The ILAMB dashboard is accessible at: https://auscs.duckdns.org/ilamb_latest/.

4.3 Added Value

Following the evaluation and benchmarking studies, which assess the performance and reliability of the model over a historical period, the added value (AV) from the downscaled BARPA & CCAM simulations is assessed.

Due to their higher horizontal resolution, RCMs can better resolve many aspects of the climate system than GCMs, such as the presence of mountains, urban areas, coastlines or islands and the model can simulate the atmospheric responses to these features with greater fidelity. It is therefore expected that RCMs more accurately model the representation of many key processes such as turbulence and mixing, convection, boundary layer dynamics and cloud formation, and critical phenomena such as temperature inversions. The above reasons highlight the potential of using a higher resolution regional model to supplement GCM-based climate information.

Improvements to the quality of information are not guaranteed in all aspects of downscaled model outputs. RCMs are constrained at their boundaries and can therefore inherit many synoptic scale biases from their driving model. Furthermore, a regional model has its own biases due to still yet unresolved processes, simplified parameterisation schemes and inaccurate land surface descriptions. Given both the potential and limitations of RCM downscaling, one can assess its benefits through added value (AV) analyses. AV analysis differs from model evaluation or model benchmarking and focuses on highlighting specific areas where the RCM downscaling provides more detailed and accurate climate information over the GCMs. Three concepts are used here: the improvements in the current climate (AV), the new small-scale information in the change signal, or potential added value (PAV) and where AV and PAV coincide or Realised Added Value (RAV).

This work follows the methodology of Di Luca et al. (2013) and Di Virgilio et al. (2020) and will examine two questions: (1) do the ACS RCMs provide improved climatic information at the GCM length scales (approx. 150 km), and (2) do ACS RCMs better represent regional scale (i.e. at RCM grid scales 12-18 km) variability? Here we focus on climatic characteristics of the models associated with cold, hot, wet and dry extremes, using percentiles and climate indices derived from daily maximum and minimum temperature and daily precipitation, described in Table 4.1.

Table 4.1. List of indices analysed in the added value (AV) and realised added value (RAV) research.

Index/Metric	Definition	Variable
tasmax99	99th percentile of daily maximum temperature	tasmax
TXx	Annual maximum of daily maximum temperature	tasmax
TNx	Annual maximum of daily minimum temperature	tasmax
TXgt40	Number of days where the daily maximum exceeds 40 °C each year	tasmax
tasmin01	1st percentile of daily minimum temperature	tasmin
TNn	Annual minimum of daily minimum temperature	tasmin
pr99	99th percentile of daily precipitation	pr
Rx1day	Maximum 1-day rainfall	pr
R20mm	Number of days when precipitation exceeds 20mm	pr
CWD	Consecutive wet days (more than 1 mm)	pr
DD	Number of dry days (precipitation less than 1mm)	pr
CDD	Consecutive dry days	pr

For the first research question, considering the GCM length scale, all RCM model results are first upscaled before calculating the climate indices in Table 4.1. Then a location-based approach is used to calculate AV, which closely follows the methodology developed by Di Virgilio et al. (2020).

For the second research question, considering the finer RCM length scale, a spatial distribution-based approach is used to measure the AV in extreme indices and percentiles.

The justification for this methodology is provided by Figure 4.2. This figure shows the 99th percentile of precipitation for a single RCM/GCM pair in the East Coast region of Australia. In panel Figure 4.2 a-c, it is evident that the GCM is failing to represent the high values of this index near the coast, while the RCM successfully simulates these high values, but often places them in locations not coincident with events observed, either because of stochastic variability differences in this small sample of events or possibly due to a model error. Qualitatively, the RCM performs better than the GCM, in that it simulates hazardous levels of extreme rainfall. Despite this, errors in placement cause large mean squared error values, such that the RCM has a higher RMSE than the GCM (panels Figure 4.2 lower left and lower centre). This method pools all time-aggregated statistics computed on the RCM native grid within regions of approximate spatial homogeneity and then computes the discrete spatial density function ρ, for the RCM, GCM and AGCD observations. The similarity between each of the RCM and GCM distribution and the observed distributions is then assessed using the Perkins Skill Score (PSS; Perkins et al. 2007).

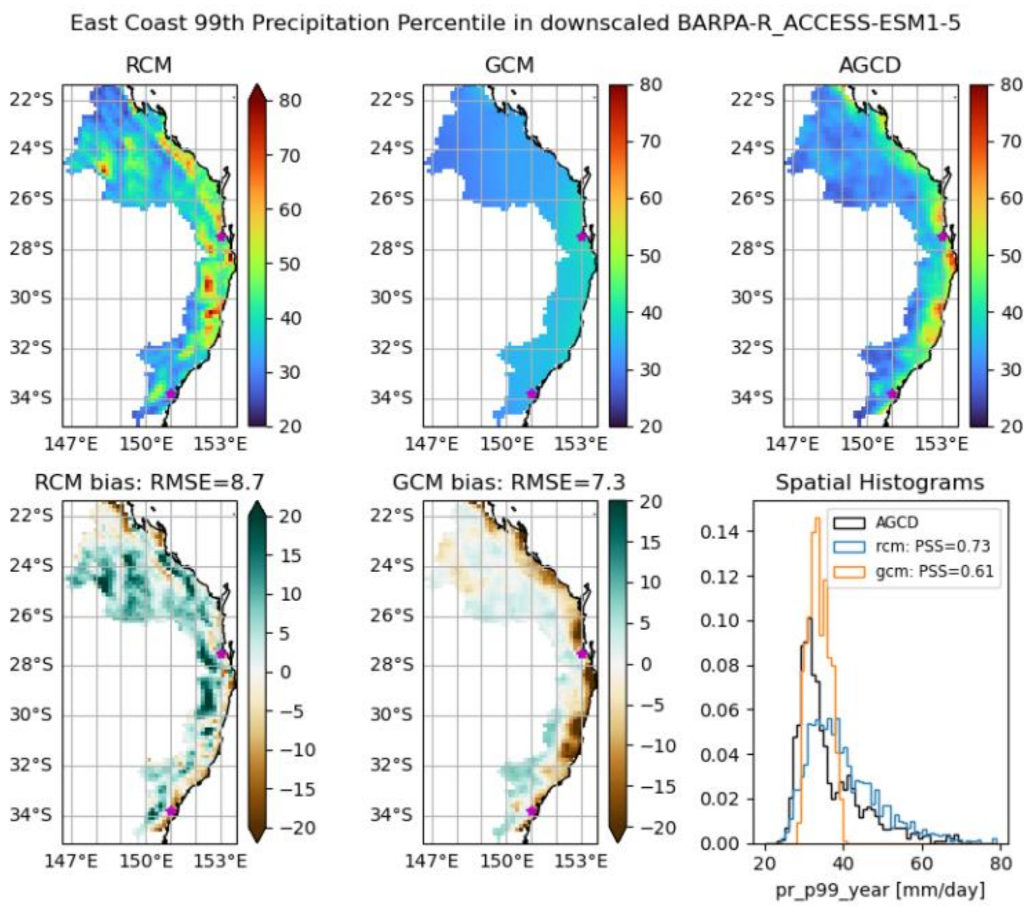

Figure 4.2. 99th percentile for the East Coast (EC) NRM region of BARPA-R downscaled ACCESS-ESM1.5 (top left), ACCESS-ESM1.5 driving data (top middle), and observations from AGCD (top right) in mm per day. The RCM's (bottom left) and GCM's (bottom middle) respective biases and spatial histogram (bottom right).

4.3.1 GCM length scale

To answer the question of whether RCMs improve the representation of the climate at GCM length scales (approx. 150 km), all model results are upscaled to 1.5 degrees resolution before calculating the climate metrics.

For each 1.5-degree grid cell, added value is calculated for annual values of the indices listed in Table 1 as the difference in the absolute difference against observations,

$$AV = |(X_{GCM-past} - X_{OBS})| - |(X_{RCM-past} - X_{OBS})| \quad (1)$$

where X_i represents the historical climate index (e.g., tasmax99). AV becomes positive (negative) when the RCM has a smaller (larger) absolute error compared to the reference dataset than the driving model.

PAV is defined as the difference between the RCM's and GCM's climate change response of the given climate index, X_i as

$$PAV = (X_{RCM-future} - X_{RCM-past}) - (X_{GCM-future} - X_{GCM-past}) \quad (2)$$

Lastly, the historical AV and future PAV are combined into the single measure, RAV using

$$RAV = \frac{AV \cdot |PAV|}{\sigma^2} \quad (3)$$

with the standard deviation of the annual time series of observations, σ. With this normalisation RAV becomes unitless and comparable across different metrics.

Additionally, we use the $AV_{fraction}$ measure to count the number of grid points where AV is positive. Naturally, this can take values between 0 (no grid-point adds value) and 1 (all grid-points add value). It is normalised by subtracting 0.5, such that $AV_{fraction} = -0.5$ when all grid points have negative AV values and 0.5 when all AV values are positive:

$$AV_{fraction} = \frac{\sum_{k=1}^{K}[sgn(AV_k);\, if\ AV_k > 0]}{K} - 0.5 \quad (4)$$

$AV_{fraction}$ is used to show that AV is not dominated by a few outliers but that the results are representative of the domain. For simplicity we call the results significant when both $AV_{fraction}$ and AV have the same sign.

For hot extremes at the GCM length scales, most RCM-GCM combinations show positive realized added value over their driving model, and most of those improvements are significant (i.e., more than half the points in the domain show positive added value). This is true for the annual maximum of daily minimum temperature (TNx), which shows positive RAV for nearly all RCM downscaling. The GCMs generally overestimate TNx while the RCMs generally reduce this warm bias. The RCMs have a general cold bias in the interior (RL) and warm biases along the coast regions (SS and SSWF-E/W), although CCAM shows colder bias than BARPA. The exception is CCAM-ACS driven by CNRM-ESM2-1 for the model, with negative temperature AV and RAV for many indices across tasmax, tasmin and pr, and in many regions (Figure 4.3). There is a strong cold bias in the driving GCM, and poor performance in tasmax in CCAM even after downscaling (Section 4.2). As such, CCAM-ACS CNRM-ESM2-1 also has negative RAV for the annual maximum of daily maximum temperature (TXx). Many of the RCM simulations do not have positive RAV for the number of days where the maximum temperature is greater than 40 °C (Figure 4.3). Many of the GCMs underestimate TXgt40, while BARPA generally overestimates TXgt40, especially for the interior and northern parts of Australia (not shown). The bias in TXgt40 for CCAM is more varied with about half the downscaled results showing positive and the other half showing negative biases.

For the cold extremes of 1st percentile minimum temperature and annual minimum of daily minimum temperature (TNn), the RCMs generally perform well with only 3 simulations having negative RAV; EC-Earth3 (both BARPA & CCAM) and ACCESS-CM2 (BARPA). All GCMs show a warm bias in the cold extremes (tasmin01 and TNn), which in the downscaled experiments are reduced in magnitude and more mixed and centred around zero (not shown). Downscaling of EC-Earth3 with BARPA and CCAM leads to only small improvements and is the only driving model where both RCMs have small negative added value over SA.

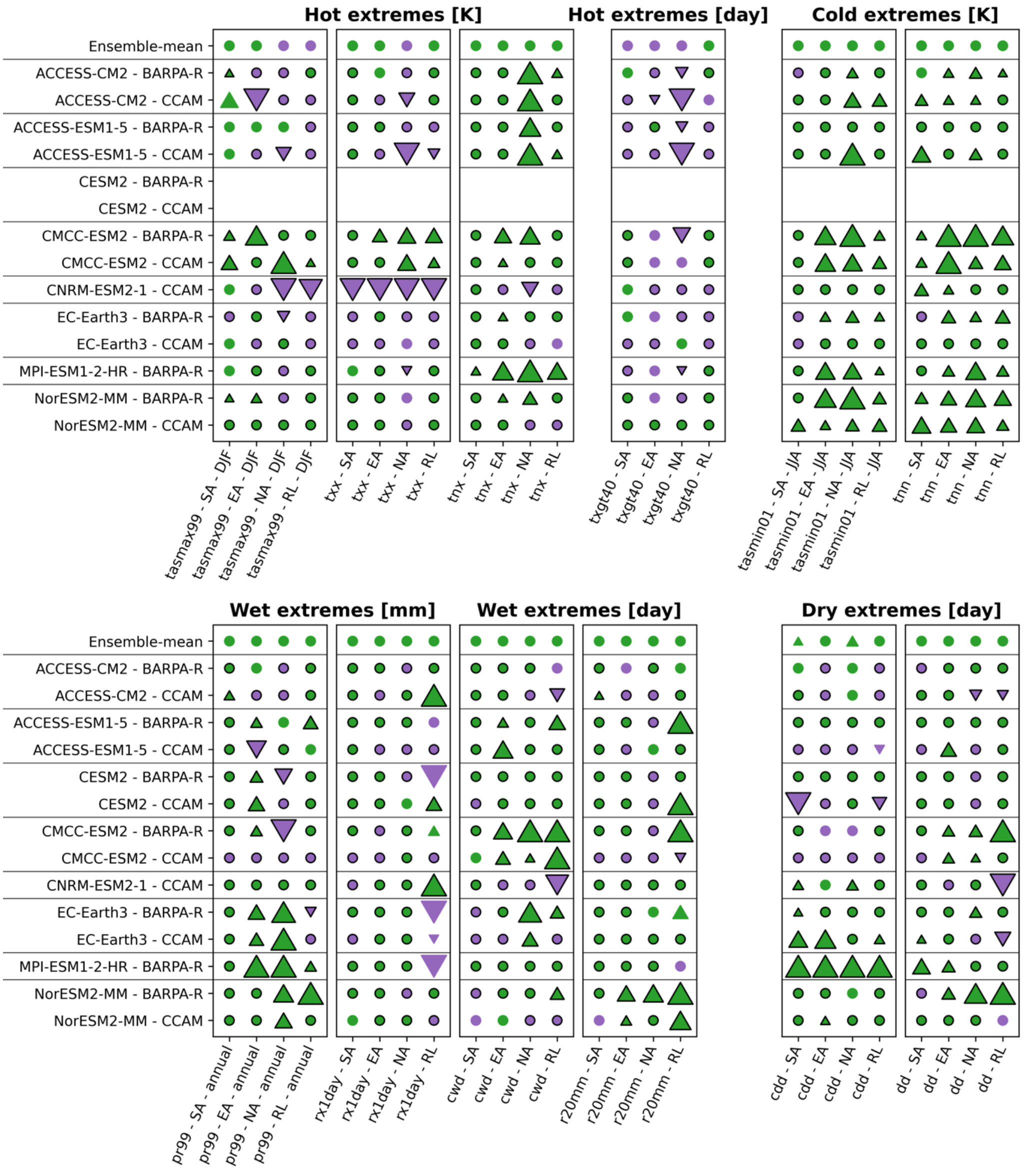

Figure 4.3. Realised added value (RAV) at GCM length scale for BARPA and CCAM downscaling for four different NRM super regions. Green circles and triangles show positive RAV, while purple circles and triangles show negative RAV. Magnitude of RAV is depicted through the size of the triangles while circles are used when triangles (i.e., magnitude) become too small. A black outline shows that more than half the grid-points in the region have the same sign as the triangle (i.e., black outline around a green triangle means positive RAV and more than 50% of grid-points are positive). Green circles in the ensemble mean row represents when value is added in more than half the grid points of all models.

For wet extremes most GCM-RCM combinations realise their added value. While maximum 1-day rainfall (Rx1day) over the Rangelands (R) and consecutive wet days (CWD) dominate the positive RAV with their large values, the agreement between the RCMs is mixed. For example, EC-Earth3, MPI-ESM1-2-HR, CESM2 (BARPA), and ACCESS-ESM1-5 show negative RAV for Rx1day while ACCESS-CM2, CESM2 (CCAM), CMCC-ESM2 and CNRM-ESM2-1 have positive RAV. For the Rangelands the observational uncertainty is large, due to sparse observations in this region, and this might partly contribute to the mixed performance in RAV. For the Eastern Australian (EA) NRM super regions nearly all GCM-RCM pairs have positive RAV. As large parts of the EA domain are coastal this could be due to better resolving the coastline as well as coastal effects, such as land-sea circulation and east coast lows, which contribute to about 25% of rainfall in this region (Pepler et al. 2021). A similar good performance can be observed for the SA region, albeit not as consistent as for the EA region.

For the dry extremes most GCM-RCM combinations have positive RAV with MPI-ESM1-2HR – BARPA performing particularly well. EC-Earth3 and NorESM2-MM are another two models where both BARPA and CCAM consistently improve on the dry extremes of the driving model. Both models stood out as ones that underestimate these dry metrics, especially over Northern Australia. The RCMs improve some of these biases but not consistently across the models or regions, indicating their fundamental limitations in simulating rainfall extremes. At their resolutions, these two models are not explicitly resolving processes that drive convective precipitation. The strongest negative RAV happens for the number of dry days (DD) in the Rangelands (R) NRM region. Like Rx1day, the mixed behaviour might stem from a large observational uncertainty in the AGCD data (see Evans et al. 2020 and King et al. 2013). The overall positive RAV in DD and CDD in the dry extremes as well as CWD in the wet extremes might also indicate a general improvement of RCMs over GCMs in the ‘drizzle bias’, general tendency of GCMs to overestimate the occurrence of light precipitation events.

4.3.2 RCM length scale

To answer the question of whether RCMs better represent regional scale variability, a spatial distribution-based approach is used to measure the AV in extreme indices and percentiles. This method pools all time-aggregated statistics computed on the RCM native grid within regions of approximate spatial homogeneity, computes the discrete spatial density function ρ, and then computes the Perkins Skill Score (PSS) of the RCM and GCM distributions as compared to the AGCD-based observed distributions. Perkins-score based AV follows

$$AV_{PSS} = PSS(RCM, OBS) - PSS(GCM, OBS)$$

with

$$PSS(a, b) = 1 - \sum_{i=1}^{N} min[\rho_a(i), \rho_b(i)]$$

which computes the differences in the density function between two datasets, a (e.g., the RCM model) and b (e.g., the observations). The PAV and RAV are computed as:

$$PAV_{PSS} = PSS(RCM, GCM)$$

$$RAV_{PSS} = AV_{PSS} \cdot |PAV_{PSS}|$$

When considering maximum temperature indices (tasmax99, TXx), hot extreme regional length-scale RAV is generally modest compared to AV (not shown). This suggests that RCMs which improve on the spatial distribution of historical temperature extremes compared to their driving models do not tend to see a large divergence in projected changes in these indices compared to their host models. This suggests that changes in hot extremes are more spatially uniform than for example wet extremes. Additionally, hot extremes are often influenced by large scale synoptics (e.g., high pressure systems), which the RCMs, due to their regional setup, hardly influence. Extreme daily minimum temperatures (TNx) show a high degree of RAV, particularly when downscaling ACCESS-CM2, ACCESS-ESM1.5 and CMCC-ESM2. Meanwhile, MPI-ESM1-2-HR and NorESM2-MM showed high AV (not shown) but more modest RAV for TNx at regional scales in northern Australia, suggesting that their projections of hot nights diverge less from the driving models. EC-Earth3 is among the driving models with the least RAV but also has the highest horizontal resolution of all downscaled CMIP6 models, suggesting that it might be able to better represent features important for hot extremes.

The largest wet extreme RAV at regional scales is present in ACCESS-CM2, ACCESS-ESM1.5 and MPI-ESM1.2-HR. Additionally, RAV is consistently high across models in the wet tropics region (MN, WT) for pr99 and Rx1day. This might be related to the RCMs better representing extreme precipitation than the driving models, which show a mixed bias in these regions. Cumulative wet day RAV is high in EC-Earth3 and in Eastern Australian clusters: EC and CS.

Daily minimum temperature extreme regional RAV is also high in the wet tropics. Elsewhere, RAV is generally positive across most cold indices, and the value is small and insignificant where it is negative.

Dry index regional-scale RAV is consistently high and similar to regional-scale AV, suggesting that divergence from GCMs is widely present. Highest values are present in the wet tropics and southern slopes regions for both CDD and DD. MPI-ESM1-2-HR shows the highest values of RAV. The overall positive RAV potentially suggests poor performance in the GCMs which has been corrected by the RCM downscaling. This may be associated with a partial resolution of the persistent GCM 'drizzle problem' achieved in BARPA-R by the inclusion of a prognostic component in the convection scheme.

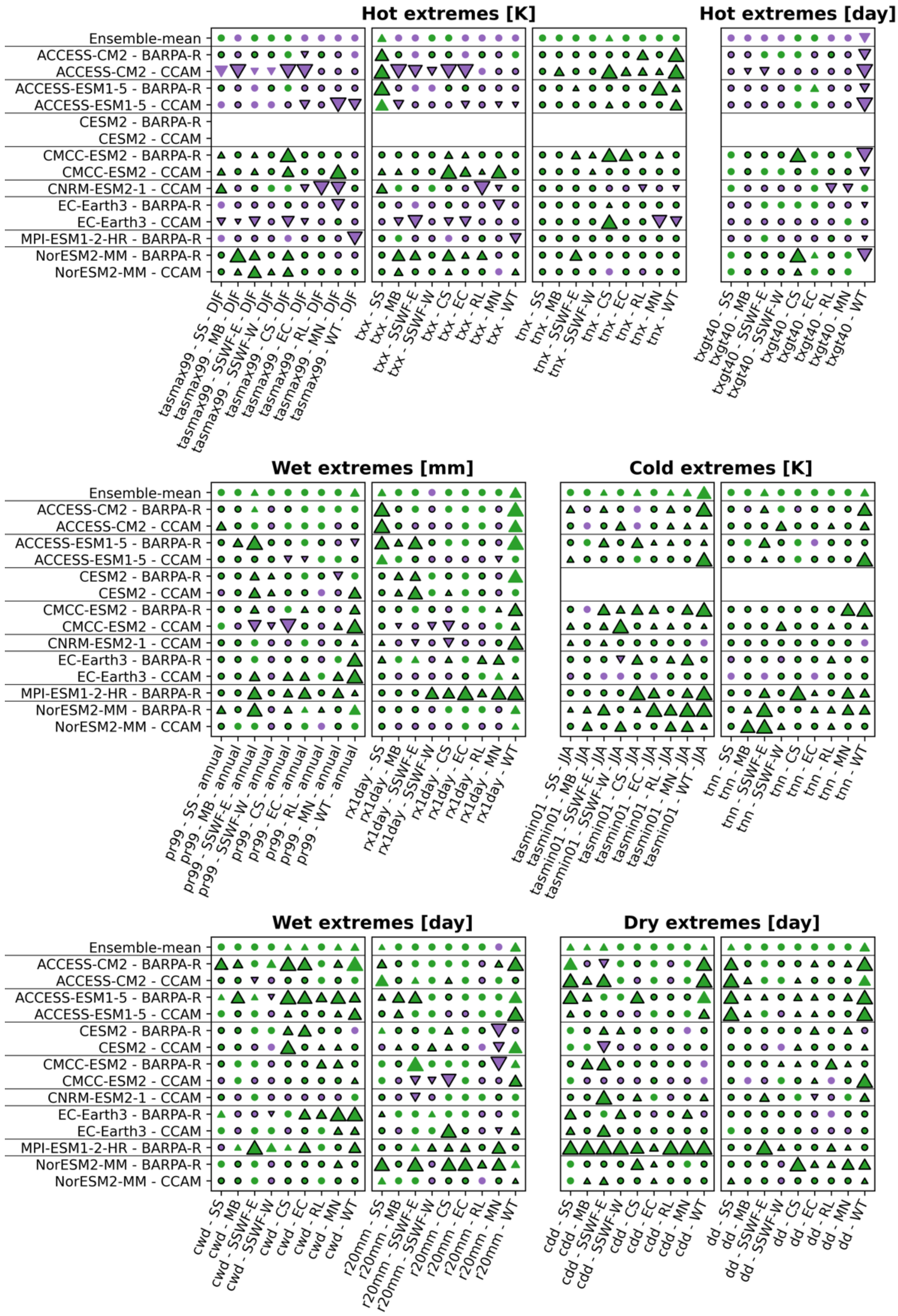

Figure 4.4 Realised added value (RAV) at RCM length scale for BARPA and CCAM downscaling for four different NRM sub regions. Green circles and triangles show positive RAV, while purple circles and triangles show negative RAV. Magnitude of RAV is depicted through the size of the triangles while circles are used when triangles (i.e., magnitude) become too small. A black outline shows that more than half the grid-points in the region have the same sign as the triangle (i.e., black outline around a green triangle means positive RAV and more than 50% of grid-points are positive). Green circles in the ensemble mean row represents when RAV is positive in more than half the grid points of all models.

In summary, while the sign and magnitude of the AV vary with various GCM-RCM experiments, the metric, season, and region considered, we found that overall BARPA and CCAM improve on their respective driving model in both length scales for the historical period as well as predicting a different climate change response. Figure 4.5 shows the ensemble mean RAV_{PSS} across all CMIP6 models and metrics. We chose to show this for the PSS based skill score as it is by construction always between 0 and 1. The absolute difference-based RAV, while also normalised, still varies more and can be biased when one metric performs very poorly. WT, with its tropical climate and SS and CS, with hot and dry summers, cool winters and complex topography perform particularly well. SSWFW, with its Mediterranean climate, has the least RAV_{PSS}. Nonetheless, all NRM regions have an overall positive RAV_{PSS}.

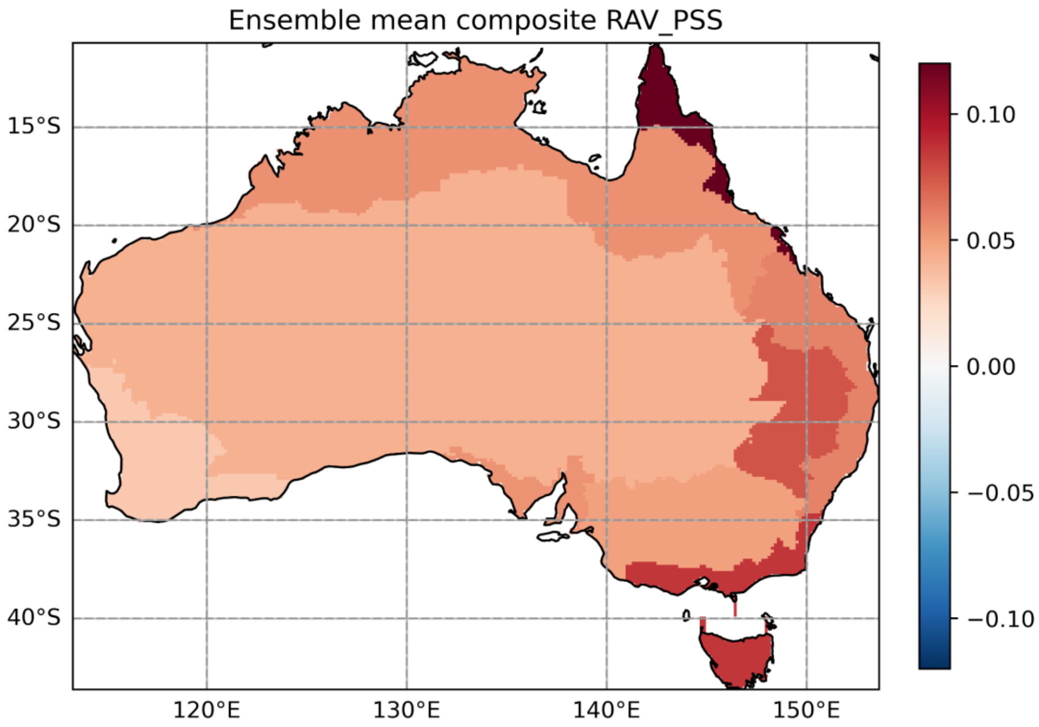

Figure 4.5: Ensemble mean of RAVPSS across all CMIP6 models and metrics.

4.4 Ensemble independence

Climate models are not fully independent - they share components and computer code, which in turn makes their outputs similar in some key respects. In some types of analysis, this can be a justified basis to weight model outputs, since their outputs are not statistically independent samples. The similarity between models can be measured and accounted for by examining the 'family tree' of model code itself, or by examining the correlations in the errors in their outputs (see Pirtle et al. (2010) and references therein for discussion of the topic). For CORDEX, the independence between simulations will be a function of the independence in the host Global Climate Models (GCMs) as well as effects from the Regional Climate Models (RCMs).

Here we use a simple but commonly used measure of model outputs to take a brief overview of the level of independence between CMIP6 and CORDEX simulations. We use the mean bias in climatological winter, summer and annual rainfall compared to the AGCD dataset (over land only), which is vectorized and the Pearson correlation between the bias is calculated for every model pair.

Correlations between annual rainfall biases between CORDEX members vary between 0.24 and 0.98 (Figure 4.5), and are comparable to CMIP6 (Figure 4.6), but with fewer pairs with very low correlations. Annual rainfall bias correlations (Figure 4.5) are more like summer biases than winter (Figure 4.7), reflecting the higher rainfall in summer in northern Australia dominating the rainfall distribution of both summer and annual rainfall. Models with the highest correlations are from QldFCP-2, and this is a function of the model setup (Figure 4.8). The process of using bias-corrected sea surface temperature (SST) and simulating the entire global atmosphere by QldFCP-2 makes the simulations highly correlated in the current climate (predominated by the RCM simulation rather than GCM inputs), but this does not link to the projection, as the SST change signal is inherited from the GCM, which is unrelated to this error correlation. Further work is needed to assess the interaction between bias adjustment, independence and the projection. The NARCLIM2.0 R3 and R5 pairs are highly correlated for three of the GCM hosts, and interestingly less so for the other two (Figure 4.8, right).

The weighting of projected change from each model based on independence could potentially lead to a more optimum ensemble result but could also make the result worse. Producing defensible ensemble weights is also complicated here due to the QldFCP-2 model setup described above. Therefore, we do not use model weights based on independence as a standard method here but leave this as a topic for further research and leave the option open for weights for specific applications.

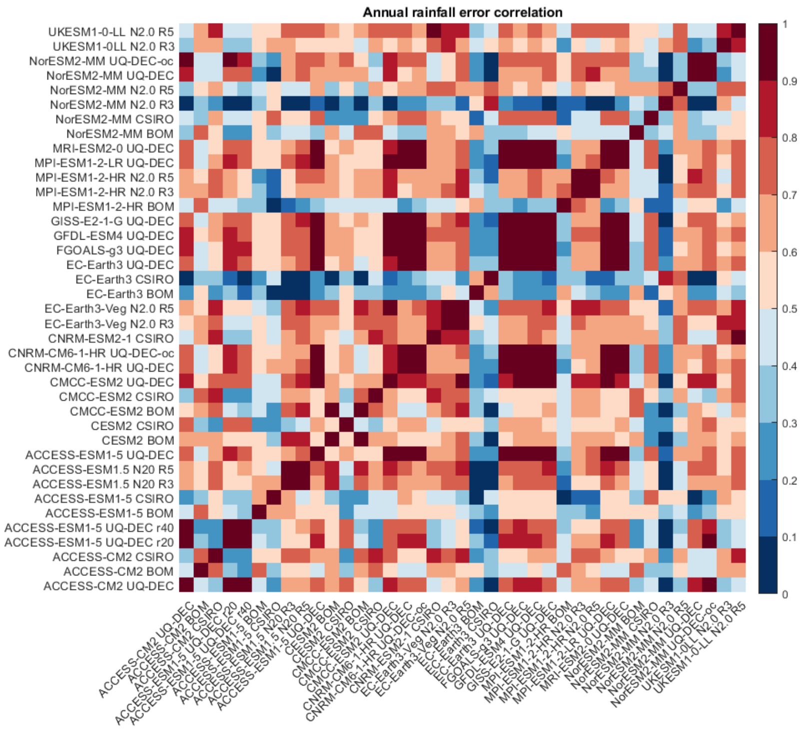

Figure 4.5. CORDEX correlations of annual rainfall biases.

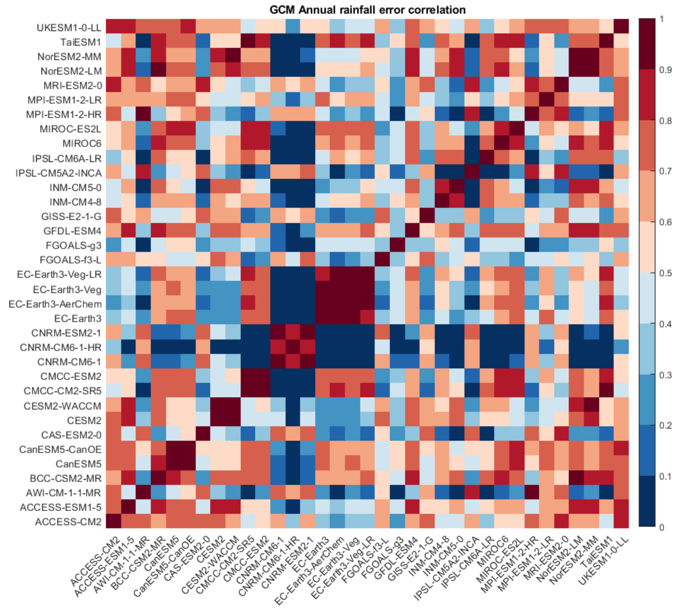

Figure 4.6. As for Figure 2, but CMIP6 GCMs.

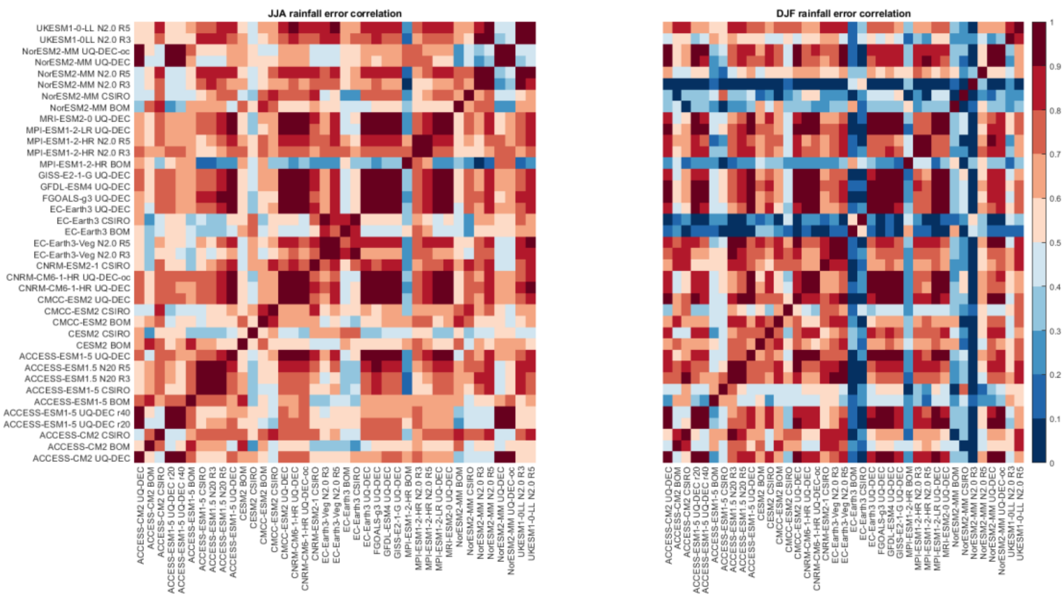

Figure 4.7. Correlations for winter (JJA) and summer (DJF).

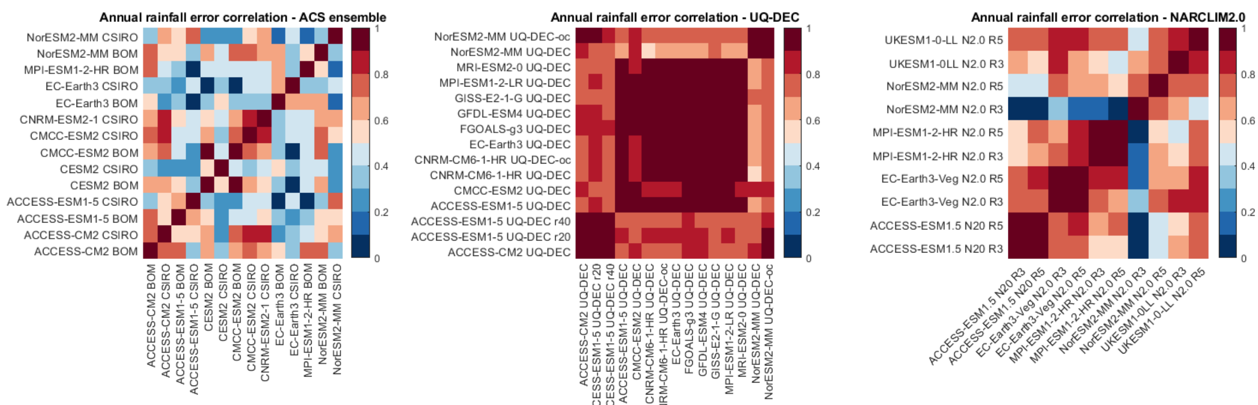

Figure 4.8. The components of Figure 1 separated by the ACS ensemble, Qld-FCP2 and NARCLIM2.0.

Weblinks Benchmarking: https://github.com/AusClimateService/Benchmarking/blob/main/README.md

5. ‘Application-ready locally relevant' datasets

- **Due to known systematic biases in climate models compared to observations of the real world, some applications may require some level of bias adjustment to climate model data before use.**
- **Here we describe the bias adjustment approach taken in the ACS to produce application-ready datasets for users.**
- **The applied bias adjustment methods range from simple univariate approaches to more complex multivariate methods. Each of these methods has its own advantages and disadvantages and is useful in different circumstances.**

Model outputs, together with other lines of evidence, are tools for understanding plausible future climate change and produce qualitative statements and quantitative ranges of change, with features like confidence levels. This can be labelled ‘knowledge of plausible regional change’, represented by the left side of Figure 5.1. Models also provide data that can be used in downstream quantitative applications, including various types of applied models (agriculture, hydrology, energy and more), and used in impact assessments, labelled ‘datasets for applications’ shown on the right of Figure 5.1. For a set of model outputs to be useful as input into quantitative models and impact assessments, they must have a few qualities:

1. Calibrated to an observed dataset that applied models use (adjusting/correcting biases)
2. At a relevant spatial scale, including downscaling and/or interpolation if needed
3. Representative of the assessment of climate change from multiple lines of evidence, without any gaps in the projected ranges of change. Including:
   a. Sampling full range of model responses
   b. Representing the projected trends from the models as they were simulated (or if the method used does alter the projected change, then a clear justification is given)

There are various methods to produce these ‘application-ready locally relevant’ datasets, including various statistical scaling and bias adjustment techniques. Here, we present high-level details of the three bias adjustment methods used in this project, selected through expert judgment and then inter-comparison. Further technical details are provided at the GitHub repository (see links at the end of the chapter) and in Irving and Macadam (2024).

Here we use the term ‘bias’ to refer to model-differences to observation in various statistical properties, not just the mean. Also, we prefer the term ‘bias adjustment’ (BA) as the techniques calibrate or ‘adjust’ the data to an observed dataset, which in itself has biases, so it is not strictly a ‘correction’ to truth. However, we also acknowledge the use of the term ‘bias correction’, including in some method names.

The data processing and applying scaling or bias adjustment can, in fact, alter the projection of climate change (or the climate change ‘signal’), in the mean, distribution, or extremes. This is a notable possibility that may be defensible and should ideally be understood and documented (some analysis of this effect is included in Chapter 7 and 8).

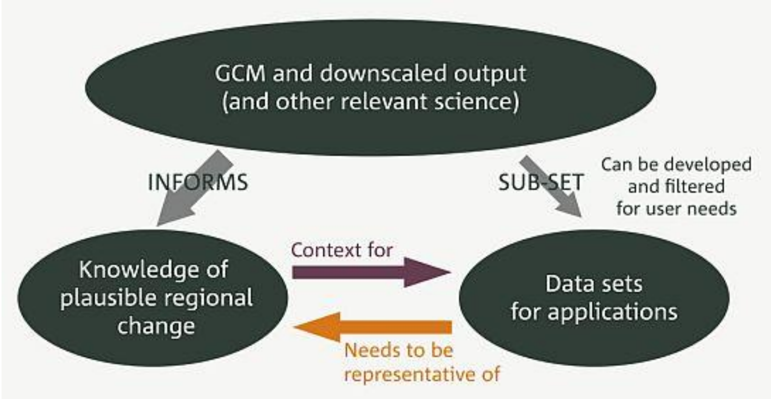

Figure 5.1. Schematic of how climate model output contributes to knowledge of plausible regional change and datasets for applications (reproduced from www.climatechangeinaustralia.gov.au)

5.1 Scaling and bias adjustment

The previous chapters have described some of the biases inherent in climate model outputs when compared to observations from the real world. This evaluation is primarily aimed at identifying models that are unacceptably biased to be credible for use in projections analyses. However, even if a model passes evaluation, the statistical properties of the data may still be unacceptably different from observations for use in some applications (e.g., impact models). These biases can be related to features such as the mean, variability, and distribution of the data, as well as temporal sequencing or spatial patterns. In most cases, these biases can be reduced and the model outputs adjusted (also termed calibrated) to an observed dataset using statistical scaling or bias adjustment methods.

The first step in a typical BA procedure involves establishing a statistical relationship or transfer function between climate model outputs and observations over a reference (i.e., historical/training) period. The established transfer function is then applied to the target model data, either the entire series or a particular time period (e.g., future model projections) to produce a 'bias adjusted' model time series. There are a wide variety of transfer functions / bias adjustment methodologies available, ranging from relatively simple methods that take a single variable as input to more sophisticated multi-variate approaches which adjust for error in multiple variables at a time.

In order to select the most appropriate bias adjustment methods for the next-generation climate projections, the National Partnership for Climate Projections (NPCP) initiated a bias adjustment intercomparison (see GitHub page, and Irving et al. Submitted). Five different BA methods were compared regarding their ability to adjust biases in temperature and rainfall, related to various aspects of the mean climatology, seasonal cycle, climate variability, daily extremes, and long-term trends. From the results of that intercomparison, three BA methods were selected (see GitHub repository for technical details and citations for each method:

- Quantile Matching for Extremes (QME; univariate)
- Quantile Delta Change (QDC; univariate)
- Multivariate Recursive Nesting Bias Correction (MRNBC; multivariate)

Both univariate methods are quantile-based, meaning the applied transfers are a function of quantiles. As opposed to the most basic form of bias adjustment, where the bias (i.e., a difference or ratio) in the mean between the model and observations is removed from all model data points of interest, quantile-based bias adjustment involves calculating the model bias for a series of quantiles and then removing that bias from the corresponding quantiles of the model data of interest. The QME method is a more sophisticated quantile method, specifically adapted to focus on the tails of distributions, as it involves scaling the data before matching the model to observations by quantile. Prior to removing the quantile biases from the model data of interest, the bias adjustment factors at the extreme ends of the distribution are also modified in the QME method to avoid potential overfitting or an excessive influence of very rare and poorly sampled events.

While it is technically a 'delta change' method as opposed to a BA method, the QDC method was also included. In contrast to bias adjustment, delta change approaches establish a transfer function between reference and future model outputs (e.g., from an historical model experiment and future model climate emission scenario experiment) and then apply that transfer function to observations to create a future time series perturbed by the model's projected change. The QDC method is conceptually very similar to basic quantile-based bias adjustment and is essentially the most basic quantile-based method available. The outputs are less subject to model biases, though they do not reflect all the changes in climate statistics simulated by models. See Irving and Macadam (2024) for more details. For this work, nine representative models were selected for QDC scaling.

Unlike the univariate approaches, the MRNBC technique is iterative, whereby a bias adjustment method is applied repeatedly until convergence is reached (i.e., until the biases are no longer getting smaller). The method is also multivariate, meaning bias correction is particularly focused on maintaining the relationship between multiple variables which can then have non-linear implications for some applications, e.g., runoff (see Vogel et al. 2023). The MRNBC method is not quantile-based, instead, it attempts to address biases in serial dependence by adjusting the data for biases in the mean, standard deviation, and lag-0 and lag-1 auto- and cross-correlations at multiple timescales (daily, monthly, seasonal and annual). Each of the methods is described in more detail in the intercomparison GitHub repository. The MRNBC method is considered the most complex one.

The effect of applying the described methods on the projected change signal in mean variables was assessed (see GitHub and Irving et al. submitted). QDC scaling produced no alteration to the mean projection, but the BA methods produce some alterations but with a magnitude always smaller than the one caused by dynamical downscaling. All methods had a negligible effect on mean temperature trends, but do affect the projections of heat indices, particularly those of absolute thresholds such as the frequency of 40 °C days (see Section 8.1). For rainfall, QME leads to generally smaller alterations to the change signal in mean rainfall, whereas MRNBC produces larger alterations (the mean absolute error of the difference between the two for mean annual rainfall is 2.4% and 4.2%, respectively). Proportional (%) alterations are generally larger for dry regions or from models with large changes to variability (e.g., from EC-Earth3 host model). The largest alterations were produced by MBCN (not discussed above; see GitHub for more information), which was one criterion for its rejection.

The QME and MRNBC methods were applied to outputs from all models contributing to the CORDEX Australasia CMIP6 dataset, producing bias adjusted sparse matrix of data (more details in the GitHub page, including a detailed report for further publications beyond this Tech Report).

The QDC method was applied to CMIP6 output as opposed to CORDEX. In the first instance, the same GCMs as those downscaled in the CORDEX project were processed, but future iterations of the dataset beyond the ACS requirements may include a wider range of CMIP6 models to allow a broader assessment of projection uncertainty and possible climate futures.

In undertaking the bias adjustment process for ACS, the Australian Gridded Climate Data (AGCD) and BARRA-R2 reanalysis were selected as reference datasets to produce two sets of bias adjusted products. AGCD is a high-resolution observational product (~5 km spatial grid) and caters to users who are satisfied with its coverage over the Australian continent (land areas only) and its limited provision for variables (tasmax, tasmin, pr). As a reanalysis product, BARRA-R2 provides a consistency across seven variables, encompassing temperature, precipitation, solar radiation, wind speed and humidity (tasmax, tasmin, pr, rsds, sfcWindmax, hursmax, hursmin were selected for this work), and covers Australia's surrounding ocean areas on a moderate 12 km spatial grid. Hence, BARRA-R2 was first interpolated to match AGCD's higher spatial resolution and then used as reference for a second set of bias adjusted products consistent with the AGCD-based outputs. Users then have the freedom to decide which of the two bias-adjusted CORDEX products is most suitable for their specific next-level analysis. A summary of the scaled and bias adjusted datasets is shown in Table 5.1.

Table 5.1 A summary of the application-ready, locally relevant datasets produced and made available, including the category of dataset, the target (what model outputs are calibrated to), methods used, ensemble and variables (see text for acronym definitions).

Category	Target dataset	Method	Variables
QDC scaled from 9 selected GCMs	AGCD	QDC	tasmax, tasmin, pr
	BARRA-R2	QDC	tasmax, tasmin, pr, rsds, sfcWindmax, hursmax, hursmin
Bias-adjusted CORDEX. 39 members	AGCD	QME	tasmax, tasmin, pr
	AGCD	MRNBC	tasmax, tasmin, pr
	BARRA-R2	QME	tasmax, tasmin, pr, rsds, sfcWindmax, hursmax, hursmin
	BARRA-R2	MRNBC	tasmax, tasmin, pr, rsds, sfcWindmax, hursmax, hursmin

FAIR Principles:

- Data release: ***https://github.com/AusClimateService/bias-correction-data-release***
- Bias adjustment method intercomparison: ***https://github.com/AusClimateService/npcp***

Other key weblinks:

- Guidance to choosing and using application-ready datasets: **link when ready**
- QDC scaled datasets: https://github.com/AusClimateService/qq-workflows/blob/main/qdc-cmip6/specs_qdc-cmip6_v1.md

6. Projections methods and choices

- **We mainly present projections using scenarios and time horizons with a standard set of choices:**
    - **SSP1-2.6 and SSP3-7.0 for standard IPCC periods 1995-2014 (baseline) and 2040-2059 (Medium term) and 2080-2099 (Long term)**
- **An alternative framework is global warming levels (GWLs), with suggested choice of:**
    - **A 1.2 °C world (a baseline centred around now) compared to 1.5 °C global warming (~2030 under all scenarios), 2 °C (mid-century under high emissions, end of century or avoided under low emissions) and 3 °C (later in the century under high emissions).**
- **For temperature projections, we separate the models within the likely range of climate sensitivity from a 'low likelihood high warming' storyline. For other variables we don't separate models based on climate sensitivity.**
- **For general presentation we use 'one model one vote' rather than weighting by evaluation, independence or other factors, but promote the use of a 'representative climate futures' where possible.**
- **We present confidence ratings using the IPCC system.**

Having described the modelling strategy, model selection and evaluation and other methods, we now discuss the choices for presenting projections for general national applications.

6.1 SSPs and Global Warming Levels

Simulations for two key Shared Socioeconomic Pathways (SSPs) of Meinshausen et al. (2019) are the main resource we use (also see Chapter 3). However, there are numerous options for analysis and presentation using the model outputs available, each with pros and cons. Because we can so clearly link cumulative carbon emissions to emissions and then to global warming, three main frameworks (labelled 'dimensions of integration' in IPCC 2021) can be used to describe future climate changes this century. The first two, SSP and time horizons, and Global Warming Levels (GWLs), are both suitable choices for most applications in assessing regional physical change and impacts, each with advantages and disadvantages (Table 6.1). The third, cumulative carbon emissions, is useful for carbon budget considerations but less useful for physical climate risk assessments.

Table 6.1 Advantages and disadvantages of climate projections for emissions scenarios and global warming levels.

Emissions scenario projections (SSP and time horizon)	**Global warming level projections**
Advantages	
• Familiar to most past users of climate projections • Gives narratives about the future world, tied to timeframes and emissions • Provides information about when in the future certain climate changes may occur (after allowing for climate variability) – can be related to planning horizons	• Helps users understand the implications of global climate change mitigation targets (e.g., the Paris Agreement 1.5 °C, 2 °C limits) • Easy to incorporate a wide range of climate model results
Disadvantages	
• More difficult to relate to global warming levels used in global climate change mitigation targets • Difficult to seamlessly incorporate a wide range of climate model results (different scenario sets, different model ensembles)	• Not appropriate for all aspects of climate change such as sea level rise where there is an important time-dependence • Cannot be easily related to time horizons for future planning without additional information

The GWL projections can be broadly related to SSP and time horizons using the assessed warming values taken from IPCC (2021), see Figure 6.1 and Table 6.2. A simple 'rule of thumb' simplification of these time windows can be used for general communication and applications (Table 6.2). Note that other SSPs show different windows, including SSP2-4.5 between these two, see IPCC (2021) for more.

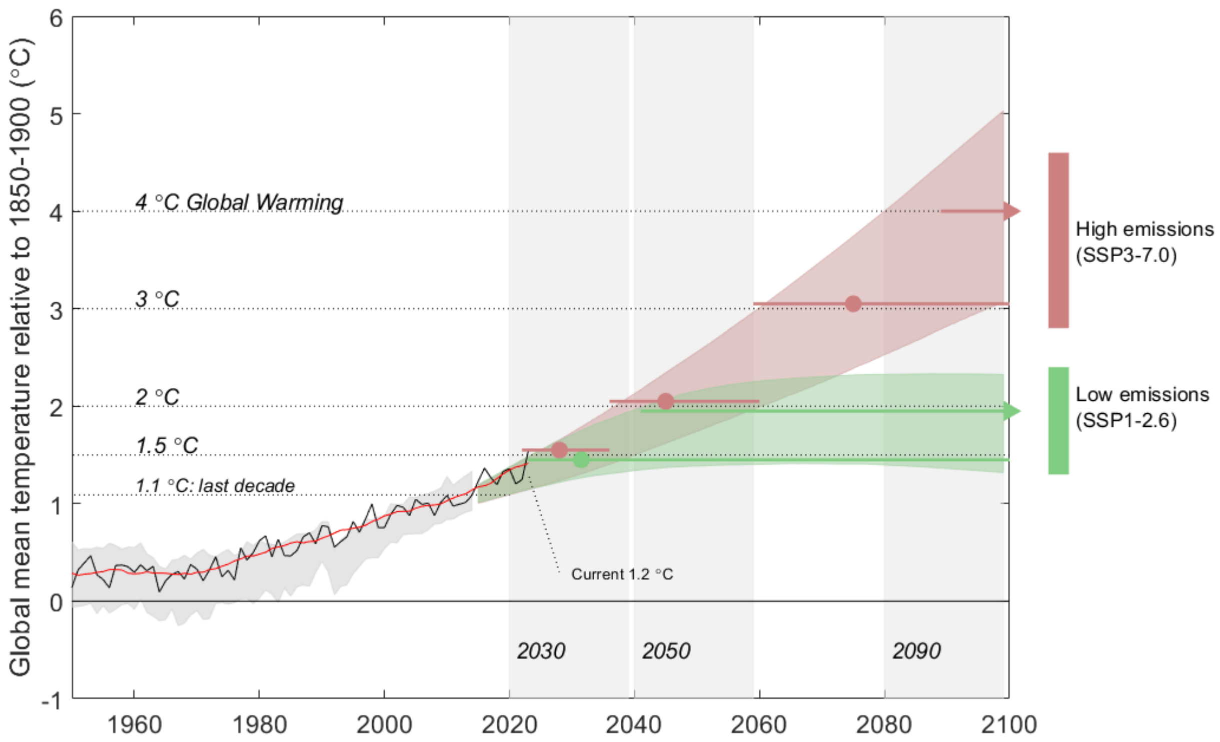

Figure 6.1. Observed and projected global surface air temperature relative to 1850-1900, observed is Berkeley Earth dataset (blackline, with Lowess smoothed series in red), coloured bands are the IPCC warming ranges assessed from multiple lines of evidence (CMIP6 models, emulators, climate sensitivity assessment, adapted from IPCC 2021).

Table 6.2. Projections of when different global warming levels, relative to the pre-industrial period, may be reached for different scenarios for future global greenhouse gas emissions. Scenarios used are Shared Socioeconomic Pathways (SSPs), adapted from IPCC (2021).

Global Warming Level	Low Emissions (SSP1-2.6)	High Emissions (SSP3-7.0)
Time windows		
1.5 °C	2024-2040	
2 °C	Possible from mid-21st Century	2037-2056
3 °C	Not possible	2055-2074
Warming Level		
Now	1.2 °C	
2030	~1.5 °C	
2050	Just over ~1.5 °C	~2 °C
2090	~2 °C	3 °C or more

To implement the GWL framework in climate models there are two options. The first is to report on the difference between a set historical baseline in time (e.g., 1995-2014) and a future warming level. This means that for each model the baseline is set in time, regardless of whether the model showed warming the same as in observations, but the future period is defined by when that model passes through a GWL. The second framework is to report between a recent warming level and a future warming level. This means that for each model, the recent historical warming level (e.g., 1.1 or 1.2 °C) is also flexible to be when the model reaches that warming level rather than assuming it matches observations precisely. The issue of baselines and different historical warming in different models is illustrated in Figure 6.2.

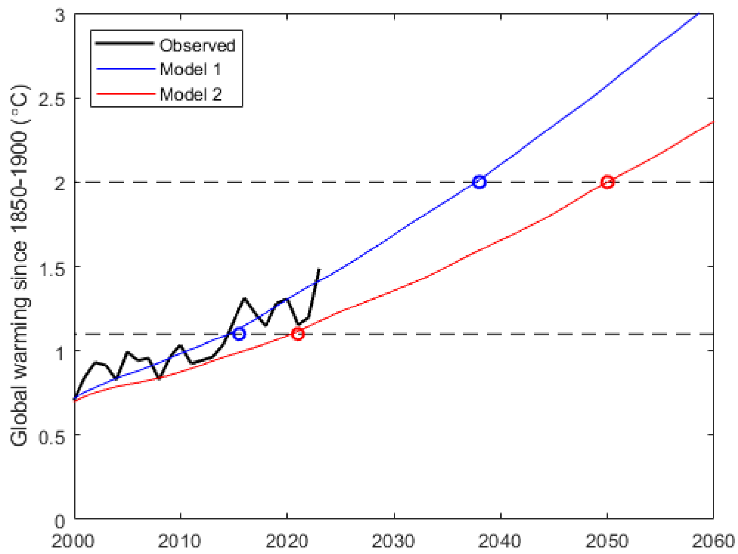

Figure 6.2 Examples of timings in reaching 1.1 and 2 °C global warming in observations and the smoothed series of two global climate models, where Model 1 crosses 1.1 °C with a similar timing to observed, but model 2 crosses later. For Model 2, the change between a baseline in time when observed reached 1.1 °C to a future GWL is in fact a period 6 years longer than for the model's change between 1.1 and 2 °C.

Each framework has advantages and disadvantages, and there is currently no single accepted standard method (Table 6.3). A database of timing for GWLs for each model is available at: https://github.com/mathause/cmip_warming_levels/tree/main

Table 6.3. The two methods of examining GWL projections, their advantages and disadvantages.

	Advantages	**Disadvantages**	**Example**
Baseline to warming level	Set historical baseline in time to orient the user, and to be consistent with existing analyses and standards	Doesn't account for model's historical warming to the baseline (models vary by >0.5 °C) – biases some results to appear lower/higher	1995-2014 baseline to 2 °C global warming - to plan adaptation from a set time of specific years in recent history to a future if we fulfill the Paris Agreement
Warming level to warming level	Accounts for different historical warming in different models, so a fairer representation of a change in climate of the stated warming magnitude – gives the user a clearer understanding of change within the GWL framework	May seem more abstract and therefore hard to interpret by users (but this can be addressed by giving each warming level a description of the time and SSP context)	Current warming of ~1.2 °C to global warming of 2 °C, similar to above but instead of specific years, the baseline is 'now' or 'current warming level'

6.2 Ensemble generation

Traditional ensemble generation follows 'one model one vote', with equal weighting of a single member from each model. To present the projections, a multi-model-mean and spread such as the 10-90th percentile range is often used (e.g., many projections plots in IPCC 2021). There are reasons for caution in using this technique with CMIP6, mainly the 'hot model problem' (see below) meaning that the range of warming from the ensemble is not balanced, but also from a lack of model independence. There are various additional considerations when using CORDEX, where there is sub-sampling of CMIP6 and more complexity in terms of independence due to multiple RCMs downscaling the same GCM.

The 'hot model' problem is most notable for temperature projections and other changes directly related to temperature change for given time horizons. However, rejecting hot models for all purposes unnecessarily throws out valuable information about plausible climate change and can produce problems such as artificially reducing projected ranges (e.g., Swaminathan et al. 2024). While the problem is very relevant to projections of SSP and time horizons it is largely bypassed by the GWL framing of projections (for which the rate of global warming is largely irrelevant). Therefore, for GWL projections we use all models with equal weighting. For time horizon projections of warming and heat indices, we present *likely* ranges of change for simulations where the host model has an Equilibrium Climate Sensitivity (ECS) between 2 and 4.5 °C taken from Sherwood et al. (2020), then models with higher ECS as 'low likelihood high warming' storylines (see Section 7.1).

We do not have a strong basis to weight models by independence (see Section 4.6), so do not employ this as standard for either the GWL of time horizon projections.

Models may be used to produce simulations with different starting conditions, known as realisations, to provide a depiction of different sequences of internal variability. CMIP6 contains 35 Realisation 1 simulations for both SSP1-2.6 and SSP3-7.0, with some models providing large ensembles. CORDEX-Australasia contains 39 simulations by 4 RCMs from 16 CMIP6 hosts (see Table 3.1). Individual RCM ensembles contain between seven and 16 simulations for each of the SSPs simulated, with the ACS ensemble containing 14 each for both SSP1-2.6 and SSP3-7.0. Given the sampling of host models and regional modelling is necessarily limited, we favour the use of the 'representative climate futures' framework (Whetton et al. 2015) wherever possible, which is one form of narrative or 'storylines' approach (Shepherd et al. 2018). This means examining all the regional model simulations based on a particular GCM as being representative of how a category of future climate could play out – such as a hot future (e.g., ACCESS-CM2, UKESM1-0-LL), a much drier future (ACCESS-ESM1.5), a future with wetter and more variable rainfall (EC-Earth3, EC-Earth3-Veg) and so on. However, in some cases we present ranges of uncertainty across different models. Where ranges are given through bar plots, we present a range appropriate to the model sample size. When using a small number of models (e.g., seven models from CCAM or BAPRA), we show 0-100% of models or else mark all ensemble members. When using a larger ensemble (e.g., all 39 CORDEX members or CMIP6) we show 10-90% range to avoid misleading impressions based on ensemble size.

6.3 Confidence assessment

For this report and further communications, we use the IPCC approach and methods for confidence and likelihood statements, documented in Mastrandrea et al. (2010) and expanded in Chapter 1 of the IPCC AR6 (Chen et al. 2021), as reproduced below (Figure 6.3). Note that a limited number of confidence statements are given in this technical report, but more are/will be given in future communication products.

Here, confidence refers to a qualitative expression of the validity of a finding (e.g., a direction or range of projected change), based on the type, amount, quality, and consistency of evidence, and the degree of agreement of that evidence. If confidence is high, then a semi-quantitative rating of probability/likelihood can then be given based on that same body of evidence.

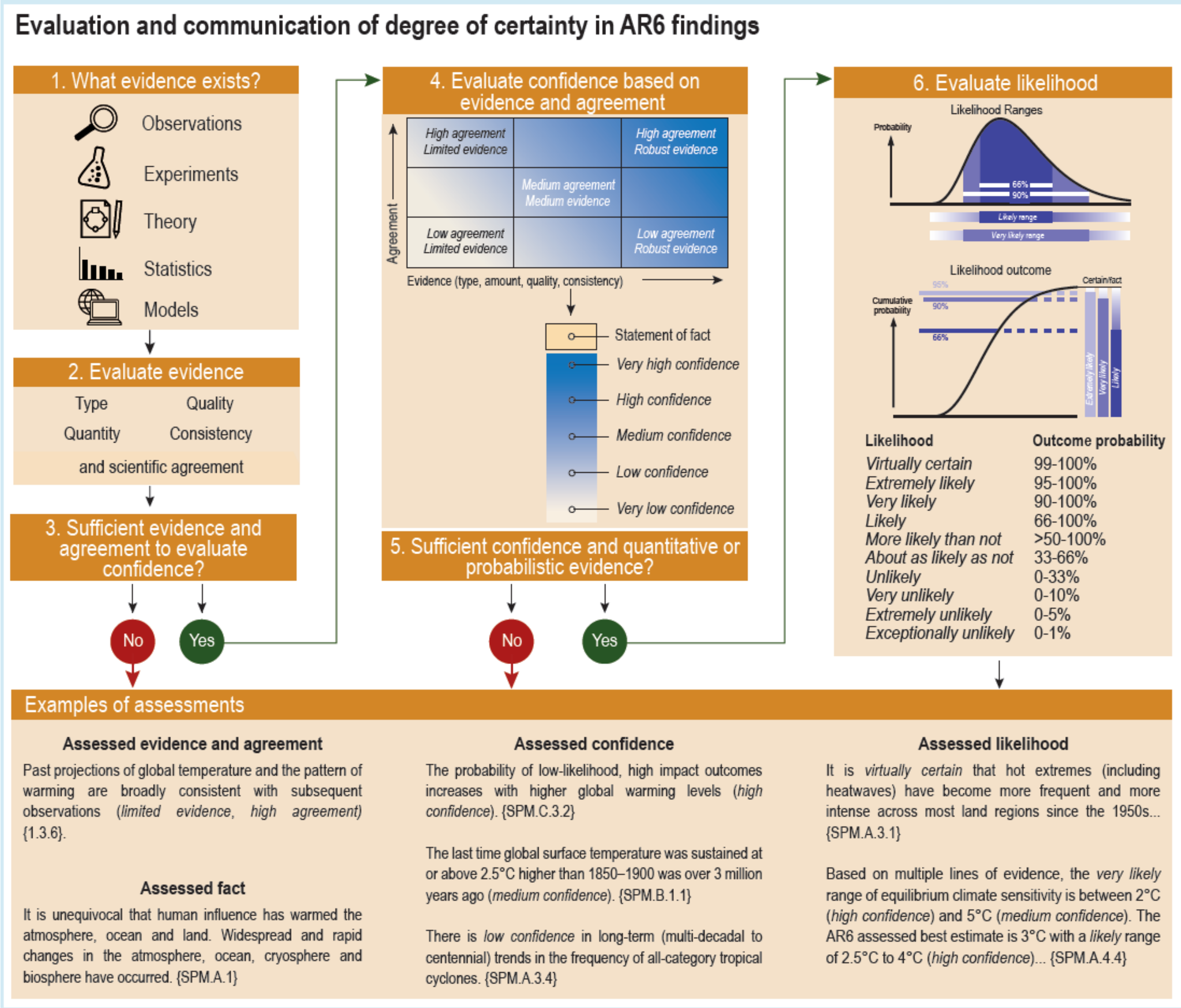

Figure 6.3. The IPCC AR6 approach for characterizing understanding and uncertainty in assessment findings. This diagram illustrates the step-by-step process authors use to evaluate and communicate the state of knowledge in their assessment (Mastrandrea et al., 2010). Authors present evidence/agreement, confidence, or likelihood terms with assessment conclusions, communicating their expert judgments accordingly. Example conclusions drawn from the IPCC Report are presented in the box at the bottom of the figure. Figure reproduced from Chen et al. (2021) and in turn adapted from Mach et al. (2017).

7. Mean changes and uncertainty in projections ensembles

- **Projected warming of Australia and sub-regions broadly agrees with other lines of evidence and previous climate modelling. The various new modelling sources broadly agree on projected warming of Australia, with likely warming by 2100 relative to the 1850-1900 baseline of 3 to 5 °C under SSP3-7.0 and 1.5 to 3 °C under SSP1-2.6.**
- **Australia is projected to warm at a rate similar or slightly greater than the global value in global climate models. Regional and seasonal variations appear as previously found (e.g., warming is greater inland than coastal).**
- **'Low likelihood high warming' storylines where climate sensitivity is very high affect the high end of projected range for the high SSPs but less so for the low SSPs – mean warming of up to 6 °C or more by 2100 for Australia relative to preindustrial under SSP3-7.0 but change above 3 °C for Australia very unlikely under SSP1-2.6.**
- **Projected change to mean rainfall remains uncertain in many regions, suggesting that a range of distinct plausible futures should always be considered when planning adaptation to the mean rainfall climate. Important exceptions include the southwest Western Australian cool season rainfall where ongoing decline remains very likely based on model projections. Other lines of evidence may help constrain regional projections but are not considered here.**
- **Projected changes to mean surface windspeeds are similar to previous generations of projections, such as generally decreases in southern mainland Australia especially in winter.**

In this Chapter we analyse and intercompare regional projections from the CMIP6 GCMs as well as the CMIP6 CORDEX-Australasia RCM ensembles. In order to concentrate on model-to-model differences we present analyses of raw model outputs (i.e. without any bias adjustment of model outputs).

7.1 Mean temperature change

Australian mean annual temperature varied by less than ±0.3°C for the period 1000–1850, and trends were all much less than in the recent period (Grose et al. 2023b). Australia warmed by ~1.6 °C between 1850-1900 and 2011-2020 (Figure 7.1), similar to the global land average and greater than the average including oceans of ~1.1 °C (Grose et al. 2023b). The total warming of Australia from 1850-1900 to the baseline used here of 1995-2014 was around 1.2 °C, and to the year 2024 is ~1.7 °C (Figure 7.1). Over the high-quality data record covered by the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset, 1910-2023, Australia warmed by ~1.5 ˚C measured as a linear trend (Grainger et al. 2023). Average annual Australian temperature has fully 'emerged' from the pre-industrial climate, and most of the warming can be confidently attributed to human influence (Gutierez et al. 2021). Extreme event attribution has shown that a notable proportion of annual and seasonal temperature records can be confidently attributed to human influence, at least from the 'Angry Summer' of 2012-13 to today (e.g., see Seneviratne et al. 2021 and references therein).

Australian average temperature will likely stabilise under SSP1-2.6 but increase throughout the century for SSP3-7.0 (Figure 7.1, Table 7.1). Note that the changes presented in Figure 7.1 and some values reported in Tabel 7.2 are relative to an early-industrial period (1850 to 1900) for this context relevant to Global Warming Levels (GWLs), but all other figures and tables are relative to a modern base period (1995 to 2014). The warming within the *likely* range of equilibrium climate sensitivity (ECS) of 2.3 - 4.5 ˚C is shown in Table 7.1, where the 10-90% of model range given. Under SSP1-2.6, we are projected to experience less than an extra 1.5 °C warming this century (or less than 3 °C from pre-industrial). Under SSP3-7.0, we see up to 4.1 °C extra warming this century (or over 5 °C from pre-industrial), with warming ongoing after 2100. For a 'low likelihood high warming' storyline where climate sensitivity is high, a single value of the median 'hot model' is given as an example (e.g., 4.7 °C by 2090 under SSP3-7.0, see Table 7.2). A high warming future could also be experienced where carbon cycle feedbacks are stronger than simulated, leading to higher greenhouse gas concentrations that those in the SSP, given the same emissions. There are also two models with climate sensitivity below the *likely* range (INM-CM4-8 and INM-CM5-0), producing global and regional warming generally at the low end of the model range (neither were downscaled by any group).

The model selection used in each RCM, and the outputs of each RCM all strongly agree on the temperature projection from CMIP6 itself (Figure 7.2), including the spread within likely range of climate sensitivity and models representative of a high warming future. There is no apparent systematic enhancement or decrease in mean warming through downscaling.

When global average temperatures reach each global warming level, Australia is projected to be at similar or slightly higher warming relative to the same baseline by CMIP5, CMIP6 and CORDEX simulations, consistent with the observed and well-understood tendency for warming over land to be higher than the warming over land and oceans combined (Table 7.2).

Values in Table 7.1 for warming since 1850-1900 show that models project a wide range in the ratio of Australian to global warming (e.g., for 2 °C global warming, Australian warming is 1.7 to 2.4 °C, meaning a ratio of 0.85 to 1.2). Model results of this ratio are almost entirely lower than the observed ratio of Australian to global warming to date, around 1.6 (Grose et al. 2023b). This difference in ratio may be due to one or more of a range of causes, such as observed uncertainties, natural variability, model deficiencies or different time dependence of processes expressed in observed change to date compared to model projections. A similar difference in the ratio was found in France (Ribes et al. 2022), which is being used as the basis for a revised approach for projections for France using 'regional warming levels' that accounts for this disparity (Corre et al. In review). Such an approach could be considered for Australia.

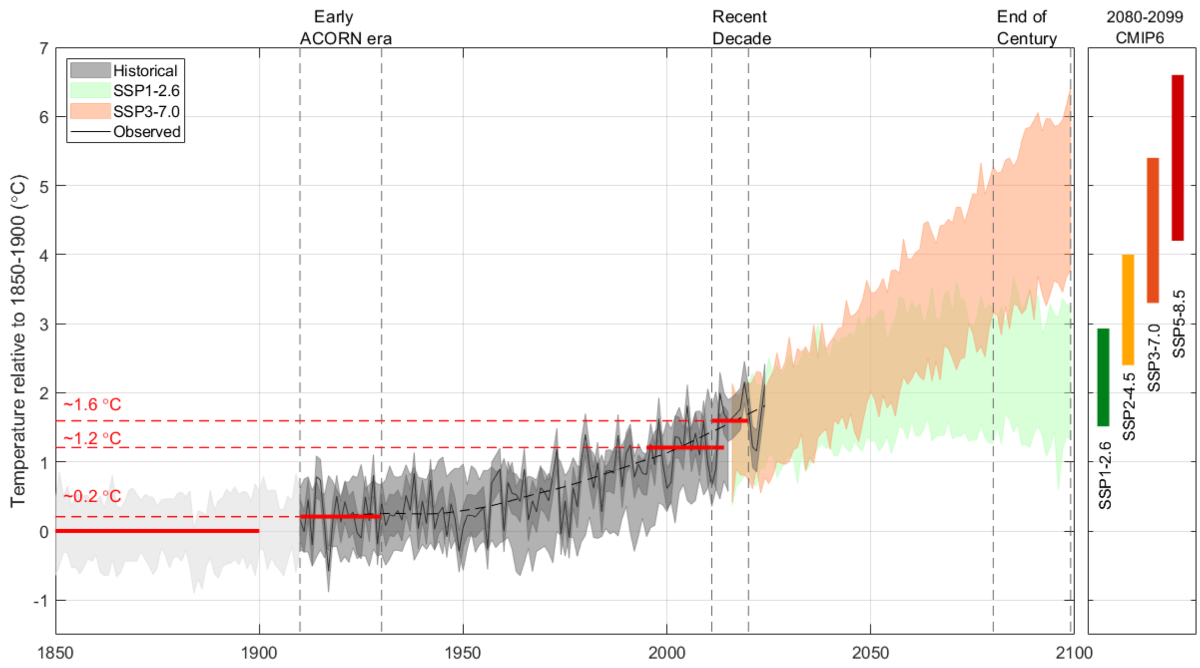

Figure 7.1. Change in Australian mean annual temperature relative to 1850-1900 in observations and CMIP6 models (without any bias adjustment to model outputs), left panel shows the time series in the ACORN-SAT observed dataset with adjustment for estimated warming prior to 1910 and 10-90% range of the first realisation of 35 models with all data available for SSP1-2.6 and SSP3-7.0, right panel shows the average in 2080-2099 for four SSPs as marked using the same data inputs. Figure adapted from Grose et al. (2023b).

Table 7.1. SSP and time horizon projection (relative to 1995-2014), *likely* range and 'low likelihood high warming' storyline in brackets

	2050	**2090**
SSP1-2.6	0.7 to 1.2 °C (1.6)	0.7 to 1.4 °C (1.9)
SSP3-7.0	1.2 to 1.7 °C (2.1)	2.6 to 4.1 °C (4.7)

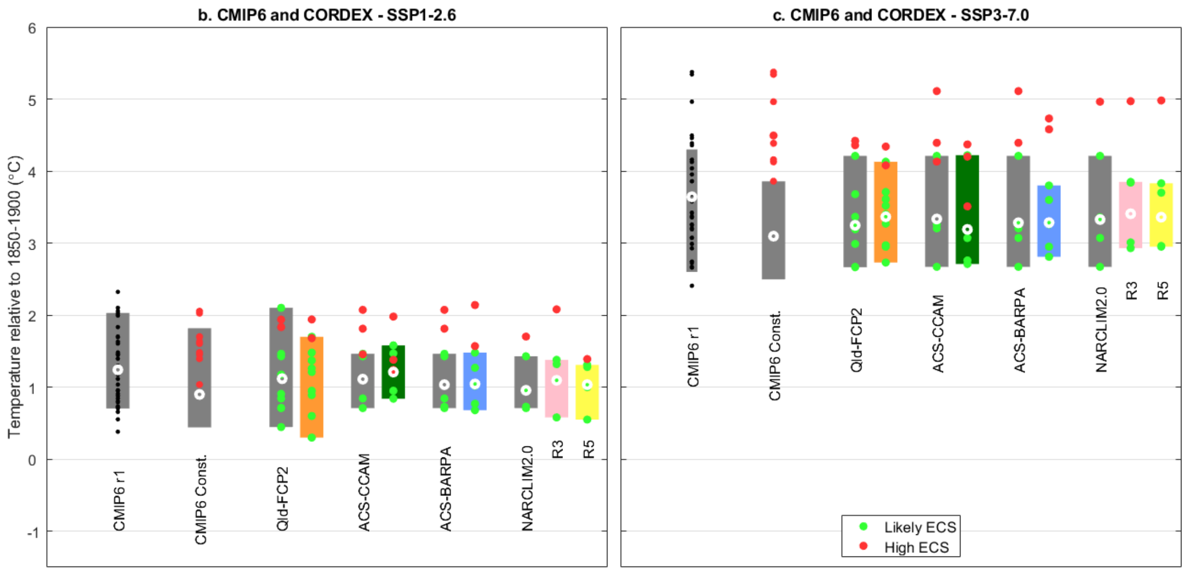

Figure 7.2 Change in Australian mean temperature between 1995-2014 to 2080-2099 for SSP1-2.6 (left panel) and SSP3-7.0 (right panel) from CMIP6 and CORDEX models, without any bias adjustment to model outputs. In each panel, the left bar shows CMIP6 using traditional methods of equal weighting of the first realization of models showing the 10-90% range (grey bar), model mean (white circle) and all 35 models (black dots). The following bar shows the 10-90% range and median when constrained by climate sensitivity and the high warming cases (likely range 2.3 to 4.5 °C, high >4.5 °C). Following this, bars show the projection from the GCM selected subset as input for each RCM including members that are not r1 (grey bars) and then the RCM result simulations themselves (coloured bars), showing individual models categorized by ECS and the 0-100% range of models in likely range of ECS.

Table 7.2. Observed change, detection and attribution (D&A) and GWL Projection of Australian air temperature over land relative to 1850-1900 (and from 1995-2014 in brackets), CMIP6 results given but results are similar in CORDEX.

Observed trends since 1850-1900	**D&A, event attribution**	**Projection**		
		1.5 °C GWL	**2 °C GWL**	**3 °C GWL**
Increase (fact) ~1.2 °C to 1995-2014 ~1.6 °C to last decade	*Robust evidence* of human contribution to trend, *very high confidence*, EEA annual and seasonal extreme events	Warmer 1.2 to 2.0 °C (0.3 to 1.0 °C from recent baseline)	Warmer 1.7 to 2.4 °C (0.7 to 1.6 °C from recent baseline)	Warmer 2.7 to 3.7 °C (1.8 to 2.7 °C from recent baseline)

Moving from national annual average to regional and seasonal changes, here again we present projected changes in temperature for the CMIP6 GCM ensemble using traditional methods (one model one vote using realization 1 from each mode), the result from the GCM sub-selections and from each RCM ensemble (Figure 7.3 and 7.4). The model projections are first coarse-grained to a common 1.5-degree lat/lon grid and are then area-averaged to the 15 NRM subclusters (see Section 3.6). Changes shown are the difference between twenty-year future and historical periods (2081 to 2100 minus 1995 to 2014) for the summer (December to February, DJF) and winter (June to August, JJA) seasons. For brevity we only present the high emissions scenario SSP3-7.0. Here we mark the 'hot models' and the downscaling of those models, and the two 'low ECS models' with climate sensitivity below the likely range (the INM models) noting that neither of these were downscaled by any RCM. The NorESM2-MM model has ECS just within the likely range, so it is downscaled by all groups as a representative climate future of low warming.

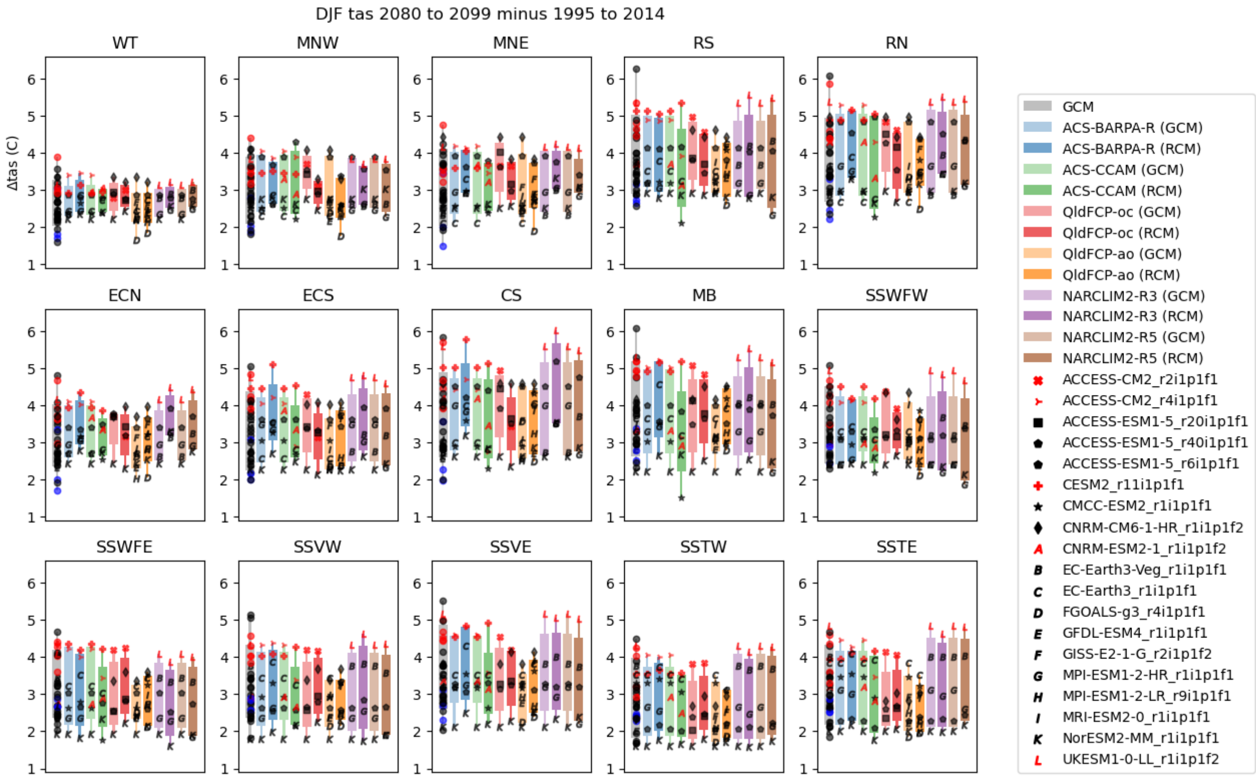

Figure 7.3. Change in summer (December to February or DJF) surface air temperature by late 21st century under a high emissions scenario (SSP3-7.0) for each model ensemble. Each panel represents an NRM subcluster region (see key in Figure 3.6). Bars show the GCM projection using traditional methods (35 models, run 1 from each model), subsequent bars show the GCM selection used by each program (including runs that are not run 1), the RCM results, red symbols indicate high ECS models while blue symbols indicate low ECS models. Units are °C.

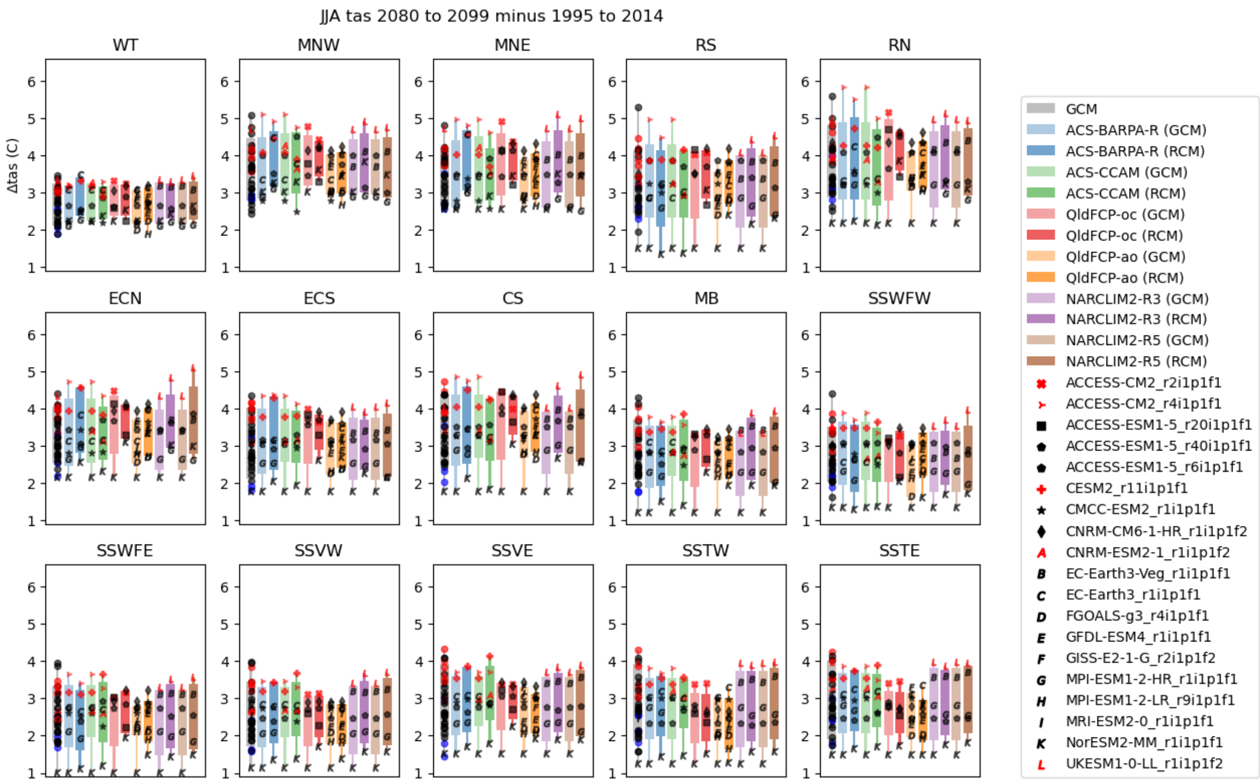

Figure 7.4. As for Figure 7.3 but showing winter (June to August or JJA).

In Figure 7.3 (December to February) and 7.4 (June to August) we explore the regional projection uncertainty for surface air temperature by late 21st century under a high emissions scenario, at the NRM subcluster scale. The range in temperature projections at the NRM region scale are broadly reflective of the national temperature projection range shown in Table 7.2, though some regional variation is apparent (e.g. higher in inland regions such as rangelands North, lower in temperate regions such as Southern Slopes). Also consistent with the national surface air temperature projections, global climate models with higher ECS and the regional models driven with those 'hot models' at their boundaries are seen to generally project the largest regional changes in temperature. At the regional scale we generally find consistency between the RCM ensembles on ranges in projected temperature change.

7.2 Mean rainfall change

Average rainfall has seen a significant decrease in the southwest in the cool season since the 1950s that has been attributed partly to human influence and is projected to continue, including a notable decrease in wet years (e.g., Rauniyar et al. 2024). Cool season rainfall in the southeast has decreased since the 1990s, with a possible emergence of a human signal explaining ~20% of the observed decline in a GCM study (e.g., Rauniyar and Power 2020). Rainfall in the warm (wet) season has increased in regions of the north in recent decades, with potential roles for greenhouse gas and aerosol forcing, along with natural climate variability (e.g., Dey et al. 2019). Detection and attribution of rainfall trends, as well as emergence, are unclear in many regions and seasons in Australia due to large natural variability and a lack of clarity on the nature of the signal.

Projections of future rainfall change remain relatively uncertain (poorly constrained) at the regional scale, and are highly dependent on location, time period, global and regional model, and scenario. In part this is because regional rainfall is highly variable, and so the signal in models is difficult to constrain with a brief observational record. In this section we focus on model projections primarily under the high emissions scenario SSP3-7.0 over the whole 21st Century, where forced response is presumably large compared to lower emissions pathways or sooner time periods.

For a broad overview of the different model results, we present the multi-model-mean projected rainfall change maps by the end of the 21st century under the SSP3-7.0 scenario in Figure 7.5, accompanied by the 10th and 90th percentile maps of rainfall change for CMIP6, CORDEX and the ACS ensemble alone. For the CMIP6 GCM ensemble as well as for the CORDEX and ACS RCM ensembles the projected rainfall change in models ranges from large decreases to large increases in most locations. The multi-model mean change in the cool season is not spatially uniform over the continent, though more than 80% of models in each ensemble project a drying over the South-West of Western Australia. The multi-model-mean rainfall change in the warm season is wetter over most of the continent, though there is little model agreement on the direction of change in most locations in any ensemble mean. Changes between 20-year periods reported here represent an estimate of the forced change signal but also decadal variability (for some analysis of this, see Section 7.4).

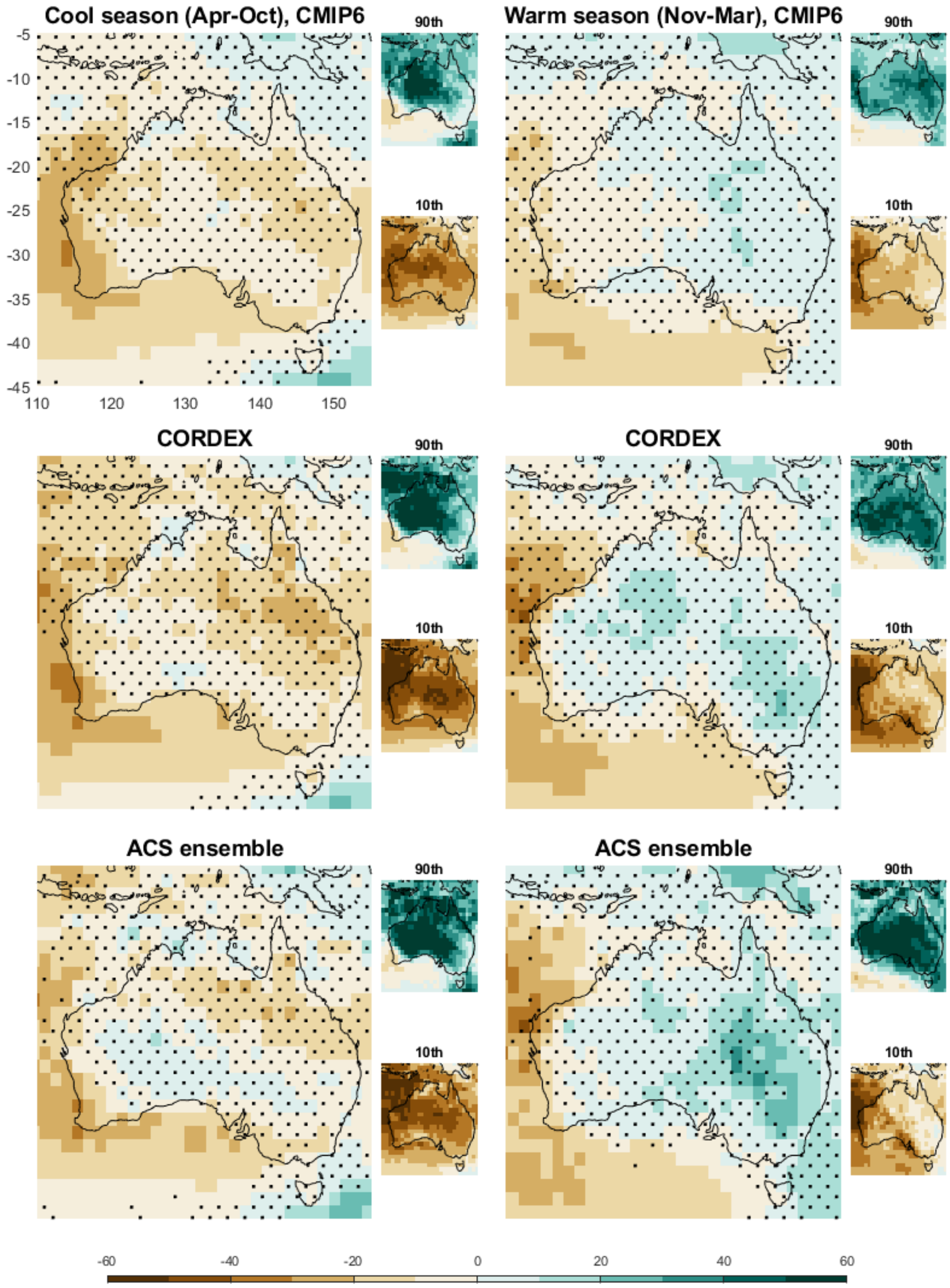

Figure 7.5. Projected rainfall change (%) between 1995-2014 and 2080-2099 under SSP3-7.0 in three model ensembles for the cool season (April-October) and warm season (November-March) using standard projections methods. Larger maps show the multi-model mean, with model agreement in the direction of change indicated by stippling (no stippling is where more than 80% of simulations agree on the direction of change, stippling indicates where less than 80% agree), smaller maps show the 10 and 90% of the ensemble at each grid cell. Note all model ensembles are plotted at a standard 1.5 ˚lat/lon grid to show a comparison of broadscale features, the CORDEX and ACS ensembles feature much finer spatial resolution than CMIP6.

The national average rainfall is the amalgam of distinct rainfall regions, so rather than presenting a national average change we instead present area averages for the 15 NRM sub-clusters. We examine area-averaged rainfall changes specifically for summer (December to February or DJF) and winter (June to August or JJA). Here again we examine change over the whole century under the higher SSP3-7.0 to provide the strongest climate change 'signal'.

The December to February changes in rainfall are highly uncertain for all NRM subclusters of Australia, with no agreement on even the direction of change found (Figure 7.6). There is also considerable disagreement at the NRM subcluster scale when intercomparing the range of projected rainfall change between the RCM ensembles. The choice of RCM ensemble at the regional scale can have important consequences for rainfall, driven by either GCM choice, RCM response, or both. For example, in Tasmania (SSTW and SSTE) the QldFCP-2 and NARCliM2.0 ensembles tend towards a decrease in summer rainfall while the BARPA-R and ACS-CCAM ensembles are centred around zero change, seemingly mainly due to GCM choice. In some other cases, the effect of the downscaling appears to be the greater source of difference between RCM ensembles. For example, there is a greater indication of rainfall increase from Qld-FCP2 compared to host models in all regions except for Tasmania. Some effects may appear quite subtle but are meaningful to downstream applications, such as the wetter end of projected change for parts of southeast Australia (e.g., SSVE) in CCAM-ACS, which has flow-on effects to summer fire danger and contributes to a different ensemble agreement on projections of drought (see Section 8.4).

In the tropical regions of Australia (WT, MNW, MNE) where much of the rainfall occurs during summer (DJF), the range of projected rainfall change varies widely between RCM ensembles, and this projected range is larger for some RCM ensembles than for the CMIP6 GCM ensemble itself. This suggests that care should be taken to consider appropriate sampling of plausible projections when considering rainfall change with global warming over the north of Australia.

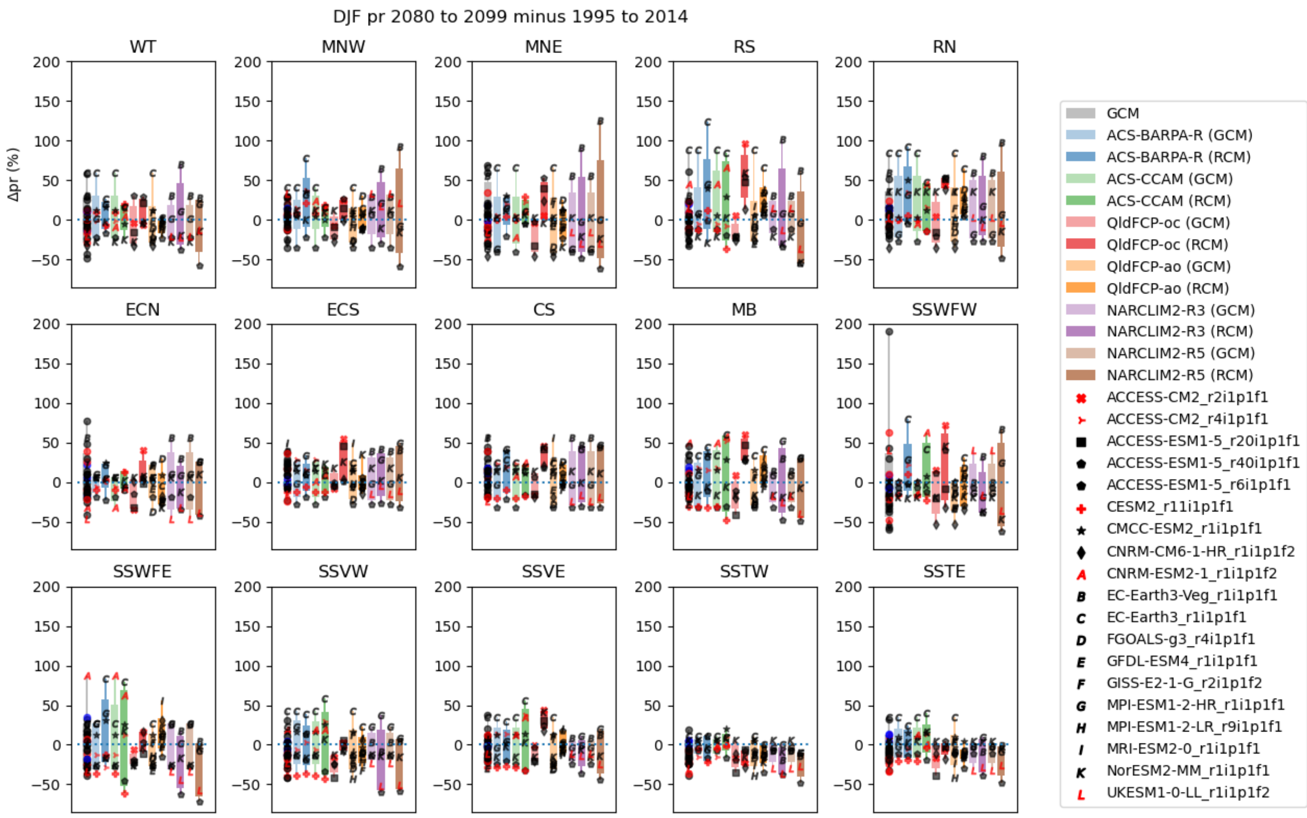

Figure 7.6. Change in December to February rainfall by late 21st century under a high emissions scenario (SSP3-7.0) for each model ensemble. Each column represents an NRM sub-cluster region. Individual models indicated by letters as marked, red symbols indicate high ECS models. Units are %, extreme outliers go off scale, version with extended y axis found in Appendix.

Similarly for the June to August season we find that for the northern regions of Australia (WT, MNW, MNE, RS, RN, ECN) there is little model agreement on changes, including direction of change (Figure 7.7). There is also a disagreement between RCM ensembles on the range and even direction of change. However, since much of the annual rain falls during the warm season these differences are less relevant than their summer counterparts (and % changes can appear large but represent small in absolute rainfall amounts). For the NRM regions located in the southern parts of the Australian continent we generally find better agreement on the direction of change, as well as better agreement between the RCM ensembles on both the range and direction of rainfall change (except for SSTW and SSTE). This may indicate a clearer signal in projected rainfall change and therefore increase our confidence in the use of any individual RCM or GCM ensemble. However, there are some differences in the projection between individual rainfall ensembles, either due to GCM choice or RCM effect on the signal. For example, the possibility of notable increase in winter rainfall is only found by Qld-FCP2 (ocean coupled version).

We generally find that projected rainfall change does not relate systematically to the ECS in models, since the high ECS models (red symbols) are not systematically the wettest or driest (Figures 7.6 and 7.7).

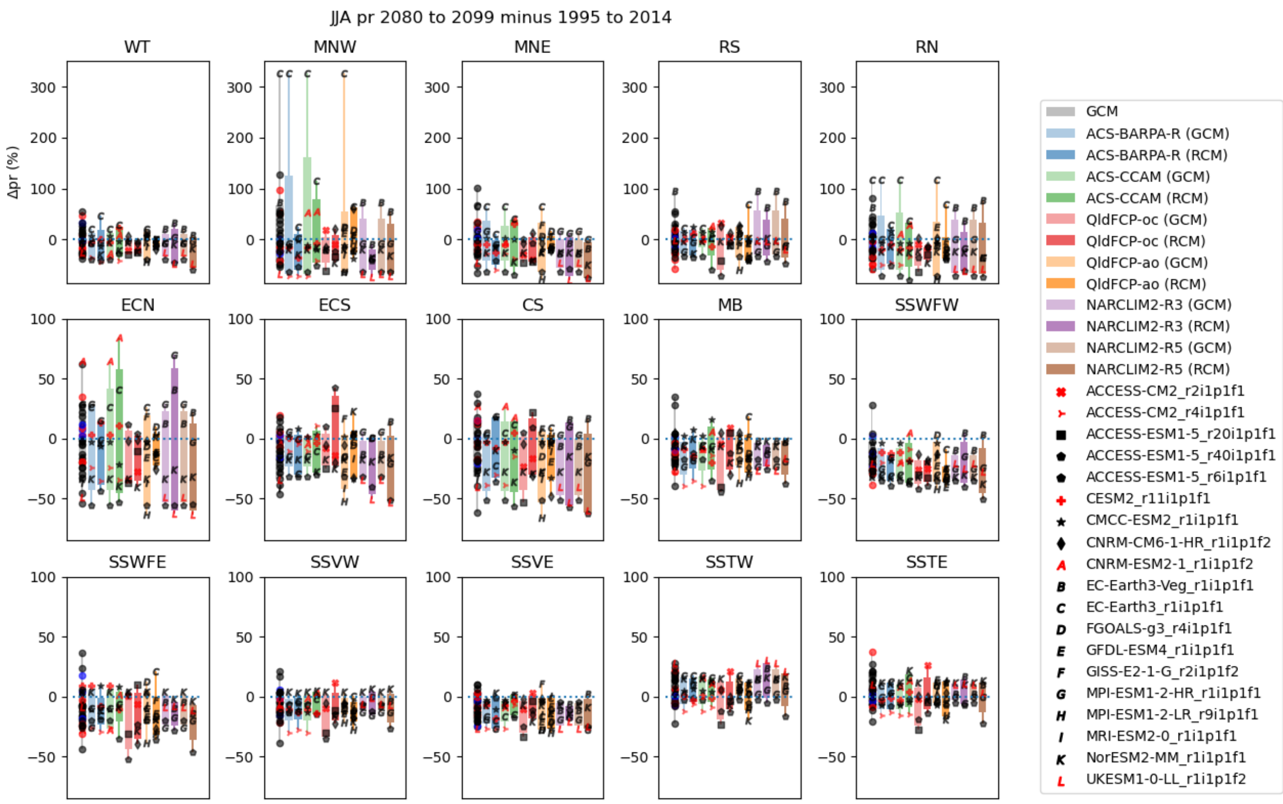

Figure 7.7. Change in June to August rainfall by late 21st century under a high emissions scenario (SSP3-7.0) for each model ensemble. Each column represents an NRM sub-cluster region. Individual models indicated by letters as marked, red symbols indicate high ECS models. Units are %, extreme outliers go off scale, version with extended y axis found in Appendix.

7.3 Projected change in rainfall interannual variability

Daily and sub-daily rainfall variability is generally projected to increase in a warmer climate, including greater extremes. This is reflected in the CORDEX-Australasia ensemble (see Section 8.3). Previous studies have shown that rainfall interannual variability is projected to increase as a global average and in many regions, but there is a spread of possibilities in most regions of Australia (e.g. Pendergrass et al. 2017). Also, there is generally a loose but notable relationship between the change in the mean and the change in the interannual variability of rainfall. Here we examine change in mean and interannual standard deviation over the whole century (1995-2014 to 2080-2099) under the higher SSP3-7.0 area averaged for the 15 NRM subclusters (see Section 3.6), in the cool season of April to October (Figure 7.8) and the warm season of Nov to March (Figure 7.9).

A relationship between the change in mean and variability is clear for many regions in both seasons, with some notable exceptions, such as the four Southern Slopes regions in the cool season. There are no regions with complete consensus in the direction of change in interannual variability, even regions where there is in the change in the mean (e.g., Southern and Southwestern Flatlands West, SSWFW in the cool season). Note that the northern regions are seasonally dry in the cool season, so some proportional (%) changes can appear very large but are not meaningful absolute amounts of rainfall, for example over northern Australia. Some extreme outliers of increase in mean and variability appear even outside the north in the cool season, including SSWFW in summer. A common wet outlier is the downscaling of EC-Earth3, inherited from the GCM and explored in the next section.

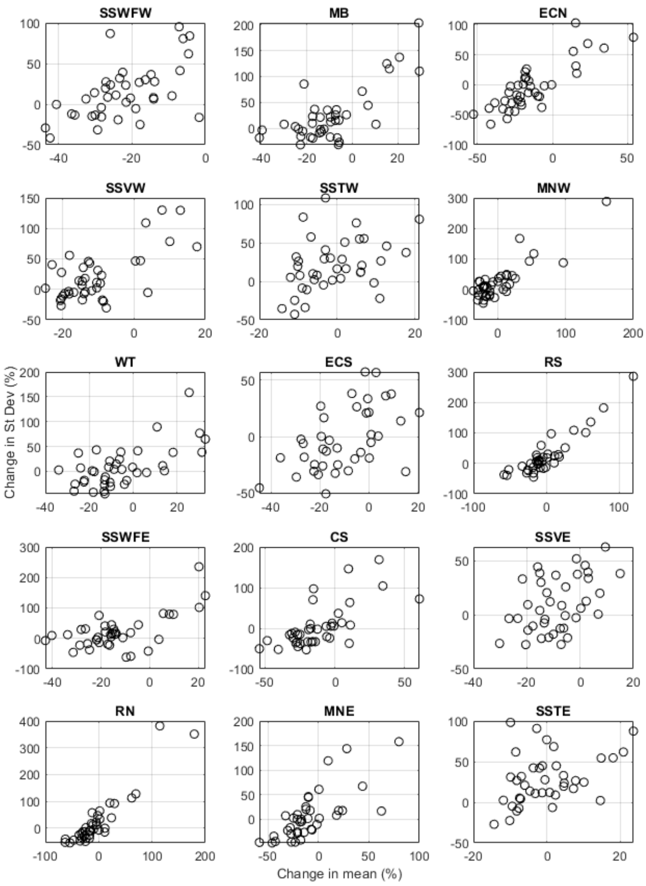

Figure 7.8. Projected change in mean seasonal rainfall and rainfall interannual variability (both %) over the whole century (1995-2014 to 2080-2099) under the higher SSP3-7.0 area averaged for the 15 NRM subclusters, in the cool season of April to October.

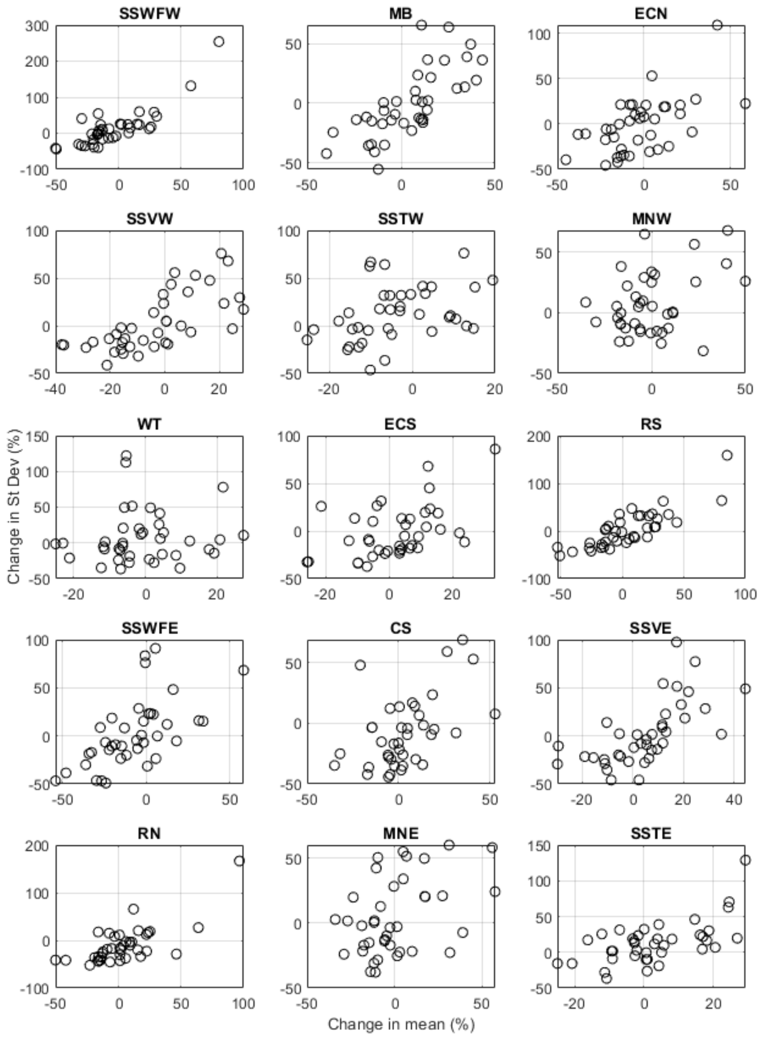

Figure 7.9. As for Figure 7.9 but for the warm season of November to March.

7.4 Insights from large ensembles – rainfall change

Model projections have often used a single member, or realization, of each model. Downscaling is mostly done for a single realization as well. Using either one realization of CMIP6 models or downscaling of one realization through CORDEX results in a small sample size for some applications regarding climate variability. However, CMIP6 includes six GCMs with large ensembles (more than 10 members for these SSPs) available to use for insights about climate variability and change (also referred to as signal and noise in this context). This analysis can provide context for the GCM realization within the large ensemble, and the downscaling of that realization.

Here we focus on the large ensemble of two CMIP6 models (not downscaled) chosen for the important 'representative climate futures' they project: ACCESS-ESM1.5 (extreme drying in most regions) and EC-Earth3 and EC-Earth3-Veg (wetter and more variable climate). The ACCESS-ESM1.5 r6i1p1f1 ensemble member was downscaled by all regional modelling groups, and the EC-Earth3 r1i1p1f1 ensemble member was downscaled by three out of four regional modelling groups, making the large ensembles of these GCMs particularly relevant. Both GCMs met the chosen thresholds for evaluation tests (Grose et al. 2023a) and the downscaling of the models perform adequately (Section 4.3), so their projection can't be rejected as implausible given this evidence (further evidence could conceivably suggest rejecting these models).

We examine the area averages for the 15 sub-clusters for the cool season (April to October) and warm season (November to March) for the higher SSP3-7.0. First looking at the ACCESS-ESM-1.5 large ensemble, we see a consistent rainfall decline across all 40 members in all subclusters in the cool season (Figure 7.11) and almost all subclusters in the warm season, but with some showing members of little change or slight increase (Figure 7.12). This model shows generally a decrease in interannual variability in both cool and warm season rainfall in most members in almost all subclusters (Figure 7.13). A few notable exceptions include Wet Tropics in the warm (wet) season and Tasmania in the cool season, where most members show an increase in interannual variability.

This suggests the projected rainfall decline from this model can be reliably described as a consistent forced response to climate change, often with a projected decrease in interannual variability. It also means that the realization downscaled (r6, and r20 and r40 downscaled by Qld-FCP2) represent a projection that is broadly consistent with all realisations in terms of the signal, but of course each with their own sequence of natural variability.

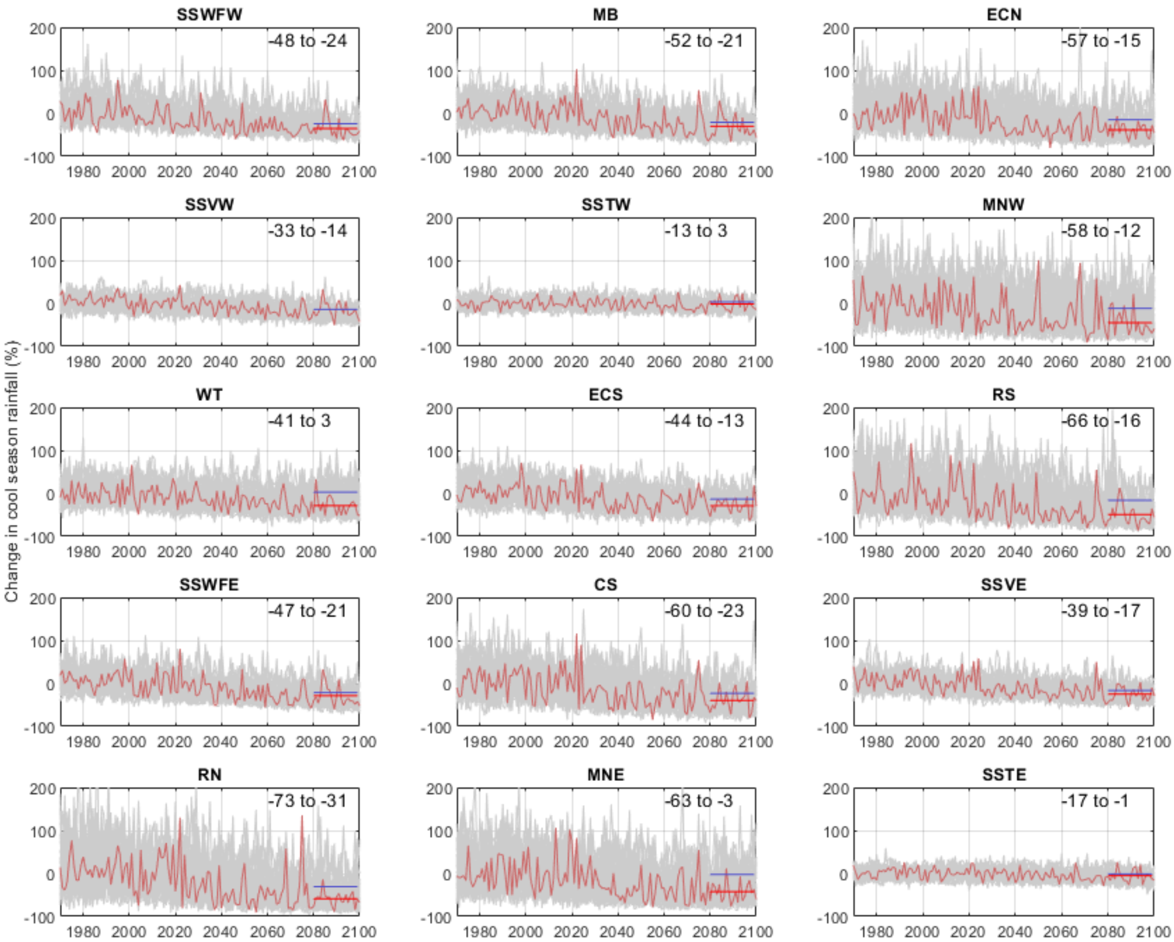

Figure 7.11 Average cool season (April – October) rainfall relative to 1995-2014 (%) in 15 subcluster regions in all 40 members of ACCESS-ESM1.5 (grey lines) and the member downscaled (r6, red line), horizontal lines show the mean in 2080-2099 for r6 (red line) and the member with the least dry projection (blue line), the full range of all members in the top right in each panel.

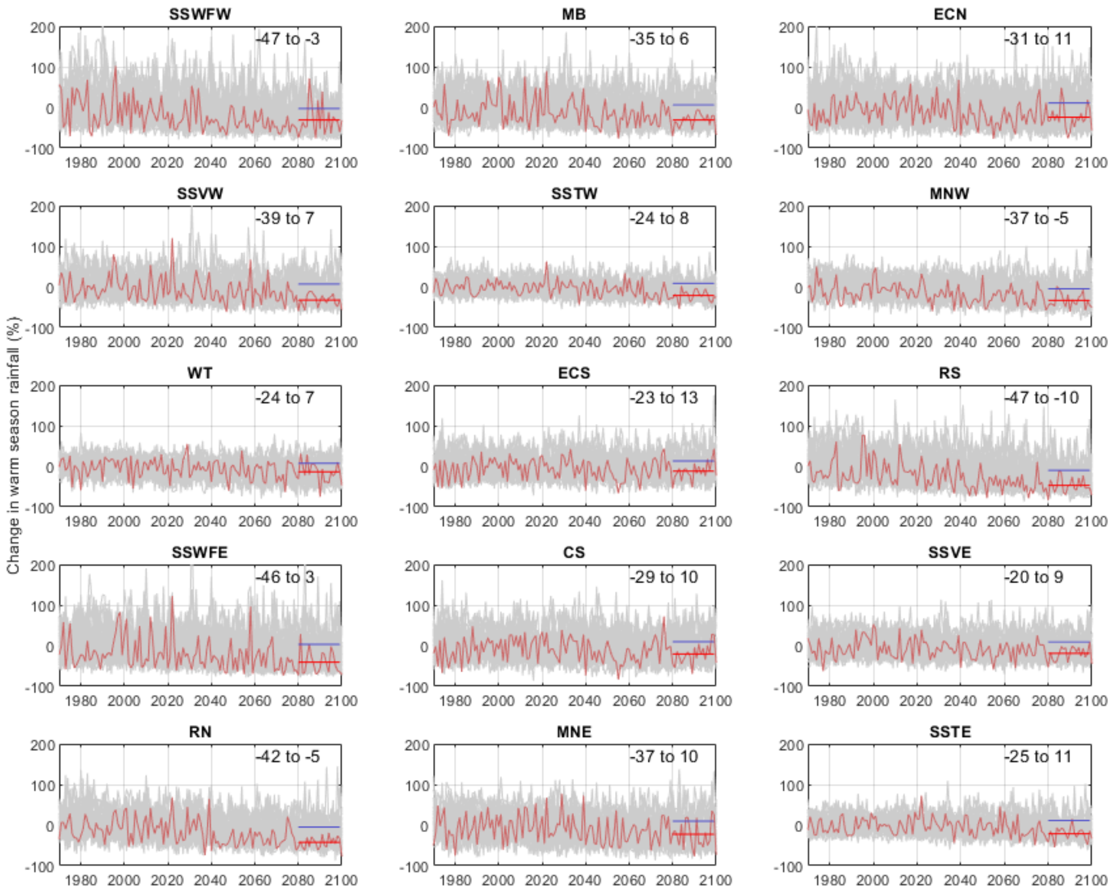

Figure 7.12 As for Figure 7.11 but showing the warm season (November to March)

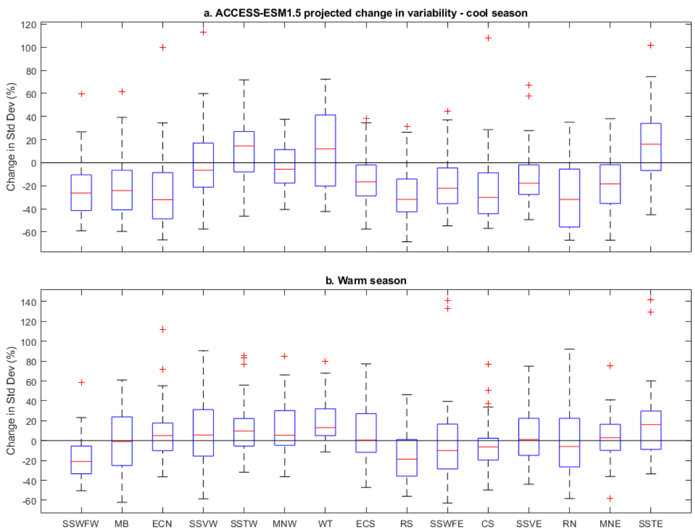

Figure 7.13 The projected change in interannual variability in cool (top) and warm (bottom) season rainfall in the 15 subclusters in the 40 member ACCESS-ESM1.5 large ensemble under SSP3-7.0 over the whole century.

Turning to the same analysis for EC-Earth3 large ensemble (n=58) we see a quite different picture. While the member downscaled (r1) generally shows a rainfall increase in both the cool and warm seasons in most subclusters, there is a notable spread across the 58 members, in most cases crossing the zero line, indicating both increase and decrease in mean rainfall (Figure 7.14, 7.15). The first realisation and many others show a notable increase in interannual variability at the end of the century in both seasons (Figure 7.16, also visible in the time series), however not all realisations do.

These results suggest that the projected change in rainfall from this model could more reliably described as a projected increase in rainfall variability allowing for wetter 20-year periods, rather than a steady increase in the mean rainfall. Therefore, the context for the downscaling of this model is of a plausible expression of a signal of wetter and more variable rainfall climate, but importantly this same signal can lead to drier periods as well.

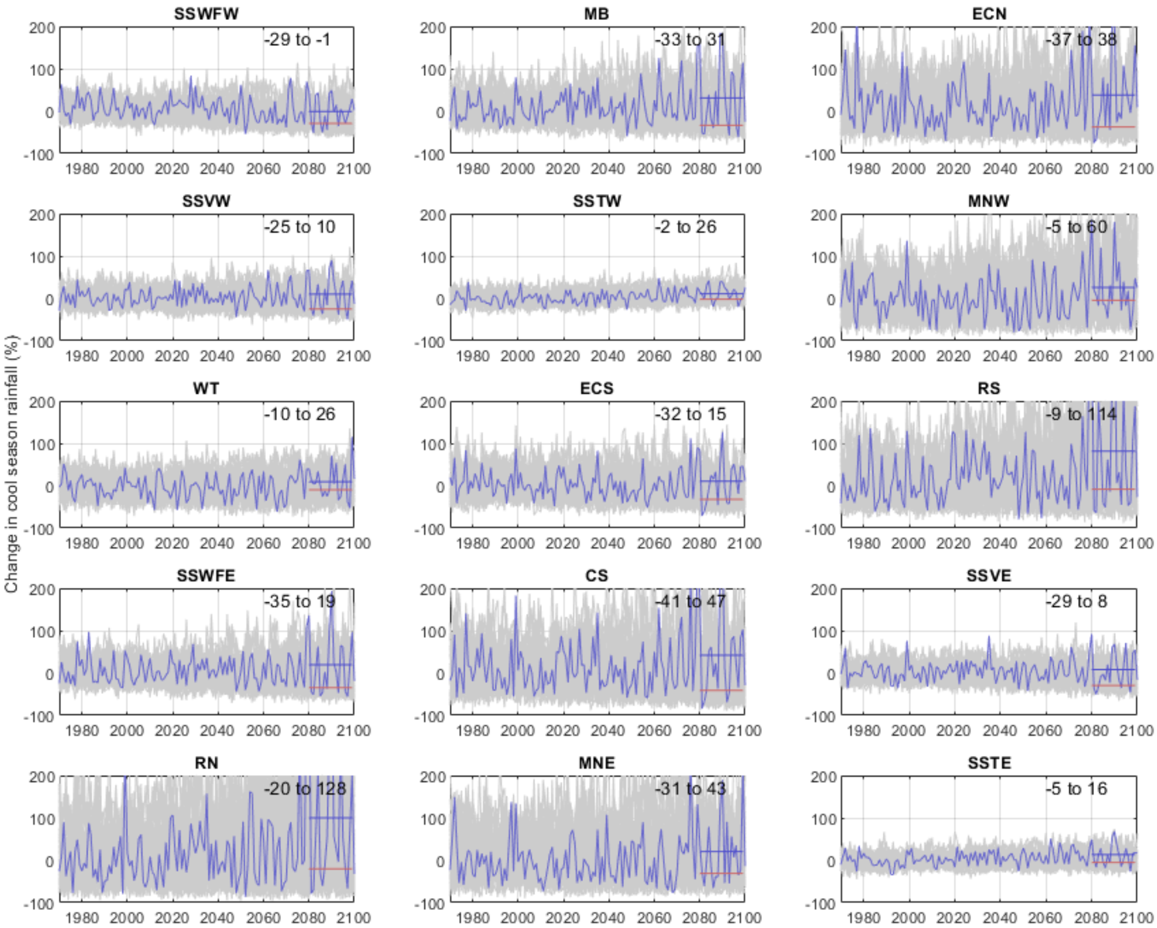

Figure 7.14 As for Figure 7.11 but using the EC-Earth3 large ensemble (58 members), and lines in 2080-2099 show the member downscaled (r1, blue line) and the member with the driest projection (red line).

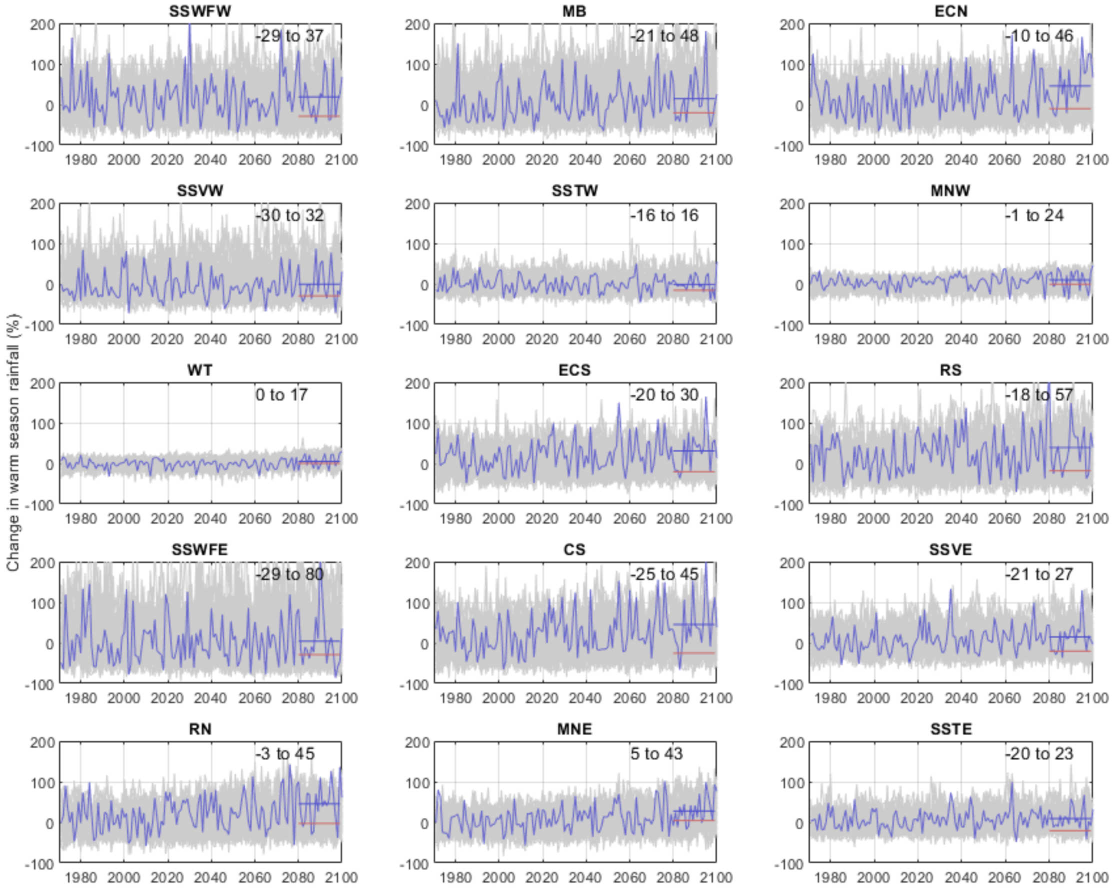

Figure 7.15 As for Figure 7.12 but using the EC-Earth3 large ensemble (58 members), and lines in 2080-2099 show the member downscaled (r1, blue line) and the member with the driest projection (red line).

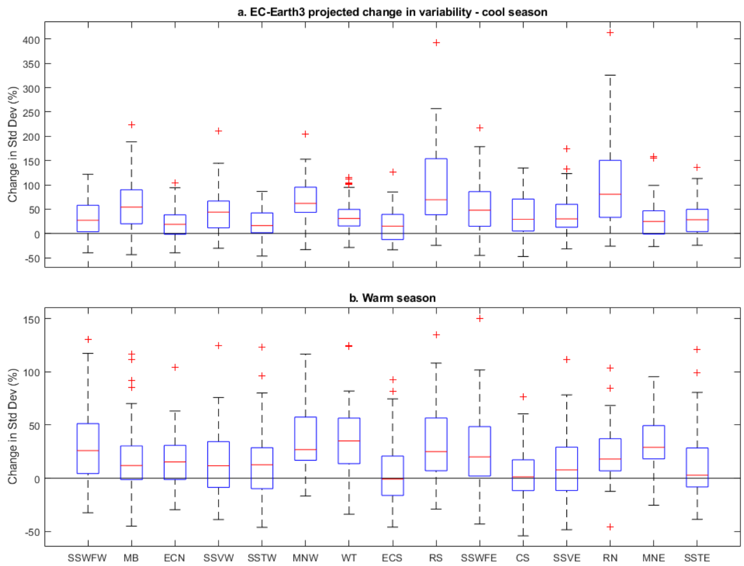

Figure 7.16 As for Figure 7.13 but using the EC-Earth3 large ensemble (58 members)

7.5 Mean surface windspeed

The climatological surface windspeeds in Australia broadly relate to location, strength and seasonal cycle of major circulation features, including the mid-latitude westerlies, trade winds and monsoon westerlies. Wind near the surface is higher over oceans than land, higher over peaks than lowlands and higher over broad open country than over forests. The CORDEX ensemble reflects these features in the 1995-2014 mean (Figure 7.18).

Projected change to seasonal mean windspeed near the surface between 1995-2014 and 2080-2099 under SSP3-7.0 (Figure 7.19) are generally similar to CMIP5 global model projections reported in CSIRO and BoM (2015) but show some added regional details. These changes include decreases in windspeed in southern mainland Australia in a spatial distribution that follows the change in circulation features, with greatest decreases in winter (JJA) at the northern edge of the current mid-latitude jet, reflecting a weakening and/or poleward movement. Tasmania being further south shows decreases in spring and autumn but in fact an increase in winter. Projected changes are generally smaller, or with lower model agreement in the direction of change, for other regions and seasons, except for a projected increase in mean surface windspeed in areas of central and eastern Australia in winter (JJA) and spring (September to November or SON).

Projected changes in mean wind speed over the century under SSP3-7.0 are generally less than 10% (compare Figure 7.18 and 7.19). Changes are likely to be smaller under the lower SSP1-2.6 and at sooner time horizons. The acute impacts from winds come from strong winds and gusts, including those from specific storms and other weather phenomena, which will not necessarily follow this mean change and need to be examined separately.

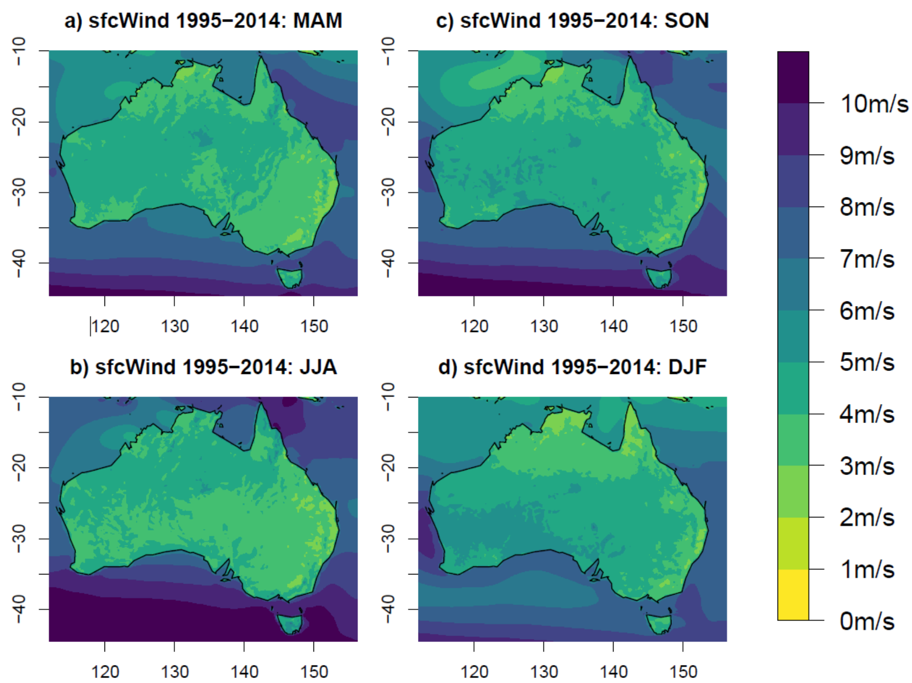

Figure 7.18 Mean seasonal surface windspeed in the mean of 39 CORDEX simulations for 1995-2014.

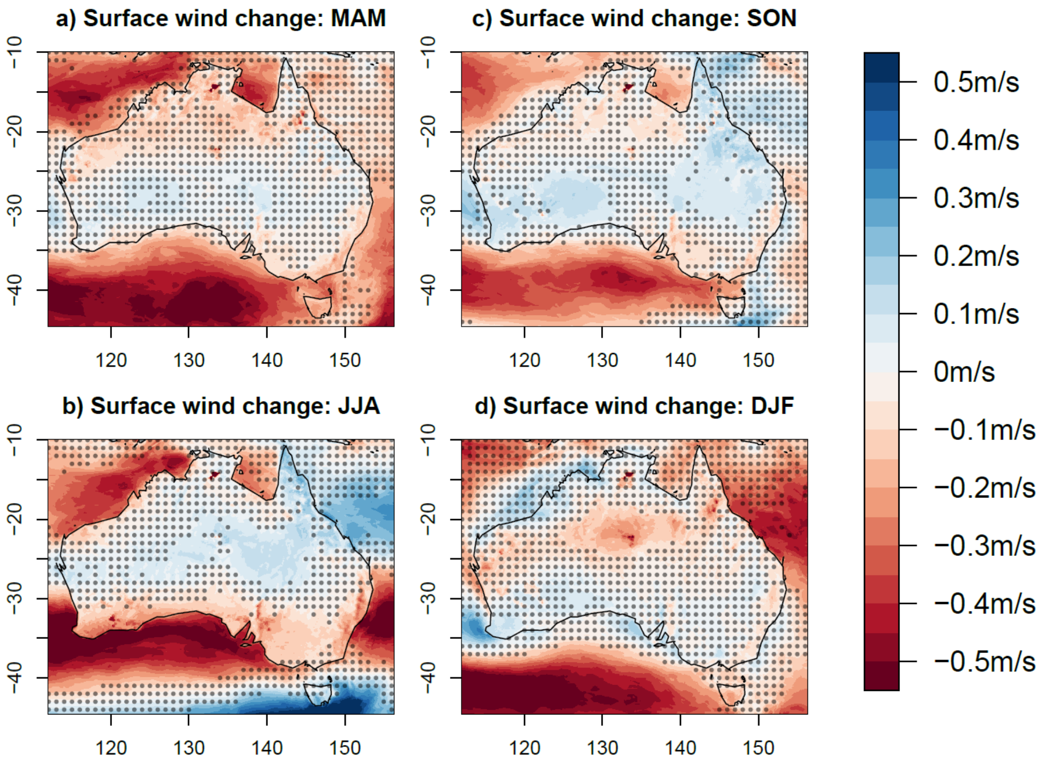

Figure 7.19 Projected change in mean seasonal surface windspeed between 1995-2014 and 2080-2099 under SSP3-7.0 in the 39 CORDEX simulations or calendar seasons. Colours show the multi-model median change, stippling indicates where there is less than 80% agreement on the direction of change among the 39 members.

7.6 Projection uncertainty

The projected changes to temperature and rainfall shown above contain large amounts of uncertainty. The range of these projected changes is understood to result from multiple sources. Hawkins and Sutton (2009) characterise those sources of projection uncertainty as stemming from 1) forcing scenario differences, 2) model-to-model differences in forced response, and 3) differences arising due to internal variability in the modelled Earth system.

For temperature projections the major source of uncertainty is the emissions pathway, with higher emissions resulting in greater warming. The model-to-model differences, even between GCMs and RCMs, is small (Figures 7.3 to 7.5). What differences do exist are partly a result of differences in climate sensitivity and can be accounted for. Internal variability does play a role, illustrated neatly by the model spread in the historical and pre-industrial temperature anomaly for Australia in Figure 7.1.

For rainfall in the near term of around the next 20 years, Australia's high climate variability is the largest source of projected spread. For a strong forcing case presented here (the high SSP3-7.0 over the whole century), the choice of CMIP6 models remains the largest source of spread in the projections on longer timescales than 20 years. In other words, we find that model-to-model differences, especially between GCMs, are the key factor for understanding rainfall projection uncertainty, (Figure 7.5 to 7.7). The RCM dimension is also a very notable source of spread, especially for some regions and seasons where different RCM ensembles produce projections that even differ in the depiction of the sign of change. An example is the projection of summer wet season (DJF) rainfall in the Wet Tropics (WT), where QldFCP2 shows high agreement on moderate drying, whereas other ensembles show uncertainty in the direction of change, with many members and even the multi-model-mean showing rainfall increase (Figure 7.6). There are also notable RCM differences in projections of heavy rainfall (see Section 8.2). By comparison, the differences in rainfall projection due to forcing scenario are small (see Narsey et al, Submitted). Internal variability does play a role at all timescales, and in some models can also change in magnitude due to anthropogenic forcing (Fig 7.13 to 7.16).

Some key caveats apply to the projection uncertainty present in the projections presented in this report. Scenario uncertainty is likely to be under-sampled in any RCM and GCM projections ensemble, since the cost of running models is prohibitive. The SSPs apply smoothly varying futures of global radiative forcing (there are abrupt changes to some regional forcings such as aerosols in SSPs, but these are remote to Australia). If events and outcomes outside the SSPs modelled futures come to pass, then warming will differ from the ranges given, for example:

- Large volcanic eruptions – temporary cooling, disrupting the warming trajectory.
- Climate engineering in which humans directly change radiative and other processes – could offset warming (or cause other changes).
- 'Overshoot' scenarios, where mitigation technologies or other process leads to not only a stabilisation but a reduction in greenhouse gas concentrations and warming, following a peak (to be addressed further in CMIP7, see Chapter 9).

Additionally, the bulk of GCMs do not contain the full range of processes that could trigger different and possibly more extreme climate responses. This means that model uncertainty may not be adequately sampled for some applications. Examples include large ice sheet dynamic, rapid biological feedbacks (such as rapid loss of carbon in the amazon) etc. Some examples of such climate responses include:

- Global climate tipping points – e.g. Atlantic Meridional Overturning Circulation (AMOC) shutdown or collapse, Amazon dieback, major permafrost thaw etc.
- High climate sensitivity or carbon feedbacks leading to high warming.
- Missing or under-represented climate feedbacks
- A 'La Niña-like' warming pattern in the Pacific (suggested in observations of the last 40 years but not produced by many model simulations) - likely to affect warming, mean rainfall and rainfall variability.

The risk of unexpected future is generally greater for higher greenhouse gas concentrations and higher global warmings than lower ones. For this reason, we urge greater caution when interpreting projections beyond at or above GWL 3.0 °C than at lower warming levels. While it is possible that some of the factor might reduce impacts, many carry risks of faster or more impactful climate change.

The results presented here suggest that response (model) uncertainty is an important concern when making choices about which set of projections to use. Depending on the variable of interest and the time horizon of interest, the choices made may subjectively bias the projected futures considered, which in turn may bias any estimates of climate related risk. In general, the best approach to sampling uncertainty is to consider the widest range of plausible futures across multiple RCM and GCM ensembles. A key aspect here is that without further careful assessment, all RCM ensembles analysed here are considered plausible, and so the use of all RCM ensembles in combination with GCM based information is perhaps best practice. We have shown that this concern is greater for rainfall than it is for surface air temperature projections.

Ultimately, the choice of GCM (or global driving model for an RCM) is likely the most important consideration since it explains the most variance in rainfall projection uncertainty. A more detailed analysis and discussion of the CORDEX-Australasia and CMIP6 GCM projection uncertainty for Australia can be found in Narsey, et al. (Submitted).

7.7 Comparison to CMIP5-based national projections in 2015

Here we provide a brief comparison of projected changes in mean temperature and rainfall from the new CORDEX ensemble with the previous generation of national climate projections (CSIRO and BoM 2015). The previous projections used CMIP5 GCM ensemble as the primary data source and were presented for the Representative Concentration Pathways (RCPs). There was a strong emphasis on RCP4.5 and RCP8.5 as moderate and high-end contexts useful for adaptation planning. The RCP8.5 pathway was seen as plausible at that time, and so a fallback to an alternative lower pathway (RCP6.0) was not taken. Projections for RCP2.6 were presented but not emphasised strongly, partly because the Paris Agreement (2015) had not been agreed on and a 'roughly 2 °C scenario' was not as policy-relevant at that time. This lack of emphasis was the case internationally, reflected in the lower number of simulations available for RCP2.6 (n = 29) compared to RCP4.5 (n = 38) and RCP8.5 (n = 42). The projections used the standard IPCC Fifth Assessment Report (AR5) baseline of 1986-2005, with the same future periods as used here (20-year periods centred on 2030, 2050, 2070 and 2090), and provided area averaging for NRM cluster regions (including sub-clusters and super-clusters).

Here we look for any notable differences in projected change to mean temperature or rainfall between the two products, due to either: 1) the different contexts presented (RCP4.5 and RCP8.5 compared to SSP1-2.6 and SSP3-7.0); or 2) the model response in CMIP5 global models compared to CMIP6-CORDEX. We expect differences in highly local scale change, and in extreme events, due to the higher resolution in CORDEX, but here we examine broad-scale change, shown in the super-cluster averages. We compare projected change in 2090 relative to the AR5 baseline of 1986-2005 for consistency and show winter (JJA) and summer (DJF) seasons.

For winter (Figure 7.19) and summer (Figure 7.20), we see a similar range of warming across the full range of scenarios in both projections noting, however, that the lower RCP2.6 was not presented as strongly for the previous generation. The similarity in the upper end of the warming range, despite the different forcing scenario (RCP8.5 compared to SSP3-7.0) is driven by the presence of the 'hot models' (high climate sensitivity and high warming). However, there are still instances when the RCP8.5 projections are higher than any in CORDEX-CMIP6.

The range of rainfall projections are generally similar between the two generations, with minor exceptions. Some differences, including the presence of outliers, is in regions and seasons where rainfall is seasonally dry, so vulnerable to misleading results as a %, e.g., Northern Australia in winter (JJA).

Conclusions about changes to mean surface wind speeds are broadly similar to CMIP5-based projections.

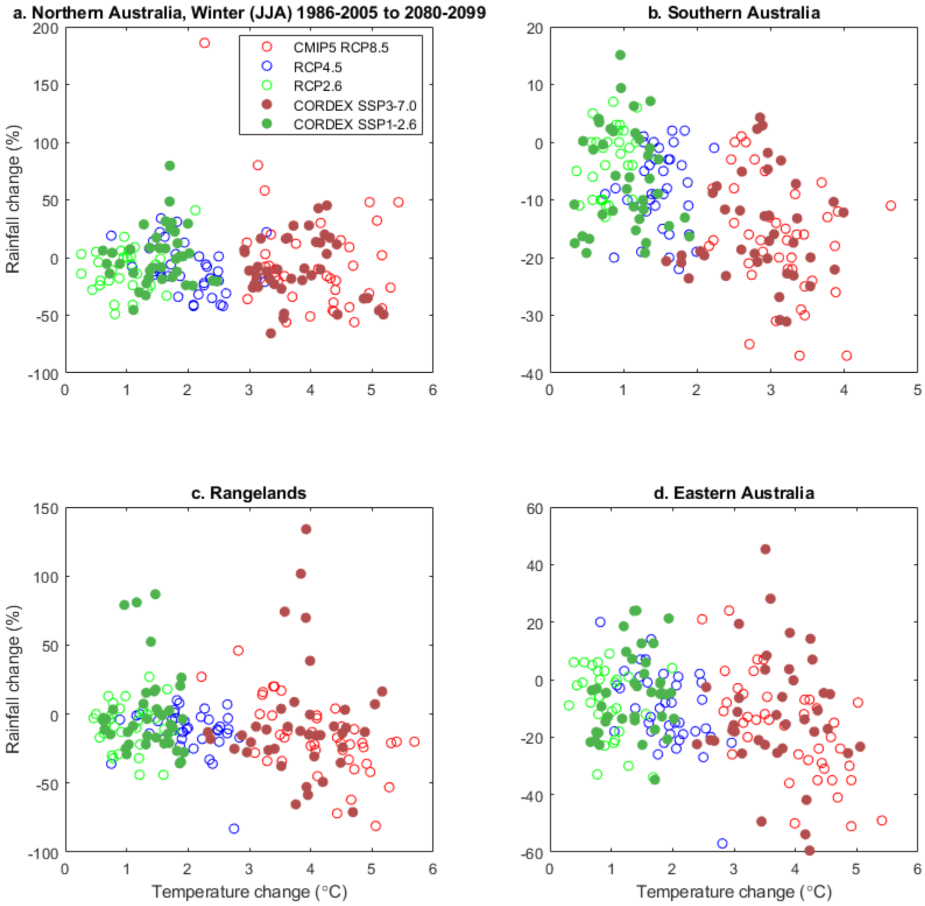

Figure 7.19 Area-averaged projected change in winter (JJA) mean temperature and rainfall in NRM Super-cluster regions between 1986-2005 and 2080-2099 from up 29 to 42 CMIP5 models and CORDEX-CMIP6 for scenarios as marked

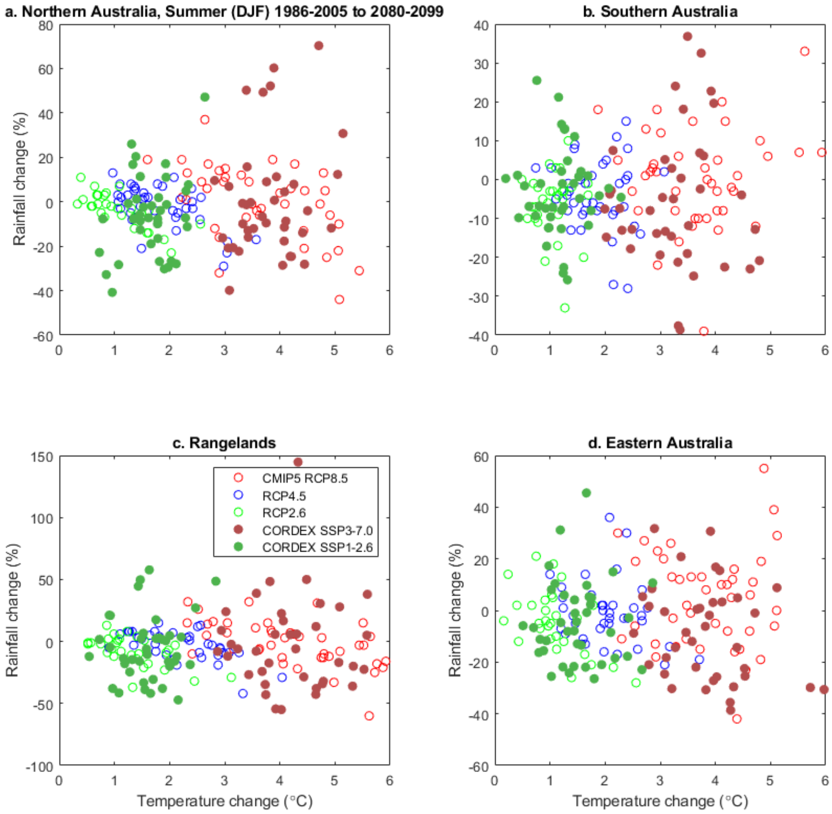

Figure 7.20 As for Figure 7.17, but for the summer (DJF) season.

8. Extremes and hazards

- **Extreme heat is projected to increase in all regions with a magnitude proportional to mean warming**
- **Extreme rainfall at hourly and daily timescale is projected to increase in most places, with some inter-model differences regarding possible regional exceptions.**
- **There is a projected decrease in extratropical cyclones (lows) in most regions of Australia, but the impact from lows might increase**
- **The proportion of time in at least mild meteorological drought at the 3-month timescale is projected to continue increasing in the southwest, as well as in large parts of the southeast and the northeast along the Queensland coast. Projected changes are less certain elsewhere, and further work will extend the work into drought accounting for evaporation and changes to aridity**
- **General fire weather danger as indicated by the Forest Fire Danger Index (FFDI), as one component of overall fire risks, is projected to increase across all of Australia, with significantly more days in the severe category (FFDI >50).**

This chapter provides examples of climate extremes derived from the new multi-model ensemble for illustrative purposes only (further work will illustrate more). We focus on extremes of temperature and rainfall (both dry and wet) following the description of changes to the mean in the previous chapter, as well as one important extreme rainfall process, that of extratropical cyclones.

8.1 Extreme temperatures and heatwaves

Extreme temperatures in Australia pose a threat to health, infrastructure, agriculture, and the natural environment. Heatwaves, characterised as extreme heat that lasts for three or more days, cause more loss of life in Australia than any other extreme weather hazard (State of the Environment 2021). In Australia, heatwaves are classified using the Excess Heat Factor (EHF, Loridan et al. 2016). The EHF is calculated based on the difference between the three-day average temperature and the 95th percentile of daily mean temperature calculated over the period 1985-2014. This is an indicator of how unusual temperatures are relative to the long-term average. The EHF is then multiplied by a factor related to the difference between the three-day mean temperature and the previous thirty days. This enhances the EHF in cases where the preceding period has been relatively cool, as an abrupt increase in temperature can have a larger impact on health than where human bodies are more acclimatised to high temperatures. Periods of consecutive days with EHF>0 are combined into heatwave events. Heatwave information is then reported using several indices, for example:

- Heatwave Frequency (**HWF**), the total number of days in a given year experiencing heatwave conditions.
- Heatwave Number (**HWN**), the total number of separate heatwave events in a given year.
- Heatwave Amplitude (**HWA**), the hottest day of the hottest heatwave in a given year.
- Heatwave Duration (**HWD**), the length of the longest heatwave, in days.

Projected changes in Heatwave Frequency (HWF) for various model choices and bias adjustment and processing choices for various NRM regions are shown in Figure 8.1. This indicates a spread of change

The ACS has also calculated several indices for monitoring changes in extreme temperatures, including:

- **TXx** – the hottest day of the year
- **TXm** – annual mean daily maximum temperature
- **TXgeN** – the number of days each year with maximum temperature ≥ N°C

Projections for TXx for over the century under SSP3-7.0 for subcluster regions (Figure 8.1) show similar results as mean summer temperature (Figure 7.3), with increases of 2-6 °C. However, there are some subtle differences, including where some models and regions show an enhanced high end of change in TXx compared to the mean. For example, in East Coast South, where the range of change in the mean is generally 2-5 °C, but TXx shows numerous changes of over 5 °C in BARPA-R and NARCliM2.0. The range of change in the host GCM subset, the raw RCM and the bias adjusted RCM are generally similar, with a few exceptions. For example, CCAM-ACS shows a suppressed increase in TXx compared to the host GCM for many southeastern regions (e.g., Southern Slopes, Murray Basin). However, all model ensembles and processing choices give broadly similar results, with GCM choice remaining the largest source of the range (from low climate sensitivity and lower warming to high sensitivity and warming). All RCM ensembles used a representative selection of GCMs, so show a similar range. However, the RCM choice does make a notable but secondary difference to the end results, with a different range and median in each RCM. There are some cases where the order of model changes is different in each ensemble, showing regional enhancement or reduction in heat relative to the driving GCM by the regional model in some cases. An example is ACCESS-CM2 downscaled through CCAM-ACS, which shows one of the highest changes in the GCM but one of the lowest in the RCM for southeastern Australia (e.g., SSVE), related to the enhanced summer rainfall increase in this model run. Importantly, the raw model output and all processing choices of the two bias adjustment methods all give similar results overall (results are also similar for the two calibration datasets, Australian Gridded Climate Data (AGCD) and BARRA-R2, not shown). Ranges of change for earlier time horizons and SSP1-2.6 are of course lower than the results in Figure 8.1.

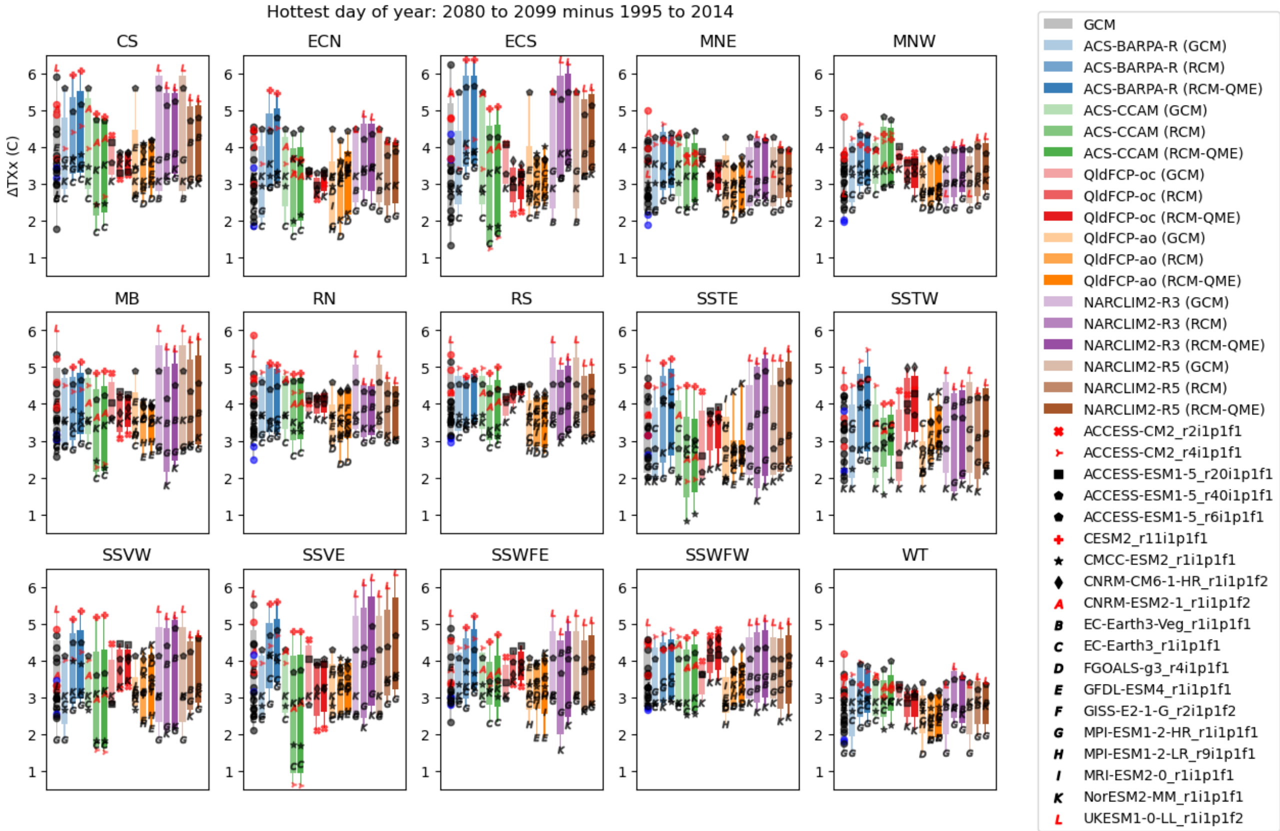

Figure 8.1 Change in the temperature (°C) of the hottest day of the year (TXx) by late 21st century under the high pathway (SSP3-7.0) for NRM subcluster regions (see key in Figure 3.6). For each panel, the left bar shows CMIP6 using traditional methods of equal weighting of the first realization of models showing the 10-90% range (grey bar), then the GCM subset used in each RCM ensemble, then the RCM result with and without bias adjustment (QME calibration to AGCD). Colours and symbols as per the legend, red symbols indicate high ECS models.

Projections of TXge40 (days over 40 °C) also show complete agreement on a significant increase (Figure 8.2), but as an absolute threshold the increase depends strongly on the current climate starting point. Projected ranges of change vary from a very small increase in Tasmania where days over 40 °C are extremely rare through to an increase of around 30 to 110 days for Monsoonal North-West, where there are currently many days in the high 30s. As for TXx, the results are broadly consistent across CMIP and RCMs, and the GCM range (including climate sensitivity) explains the greatest range. Bias adjustment generally makes a small difference in some regions (e.g., Rangelands North), but can make a very large difference in other regions, e.g., to reduce the high projections for the Wet Tropics in three of the RCMs.

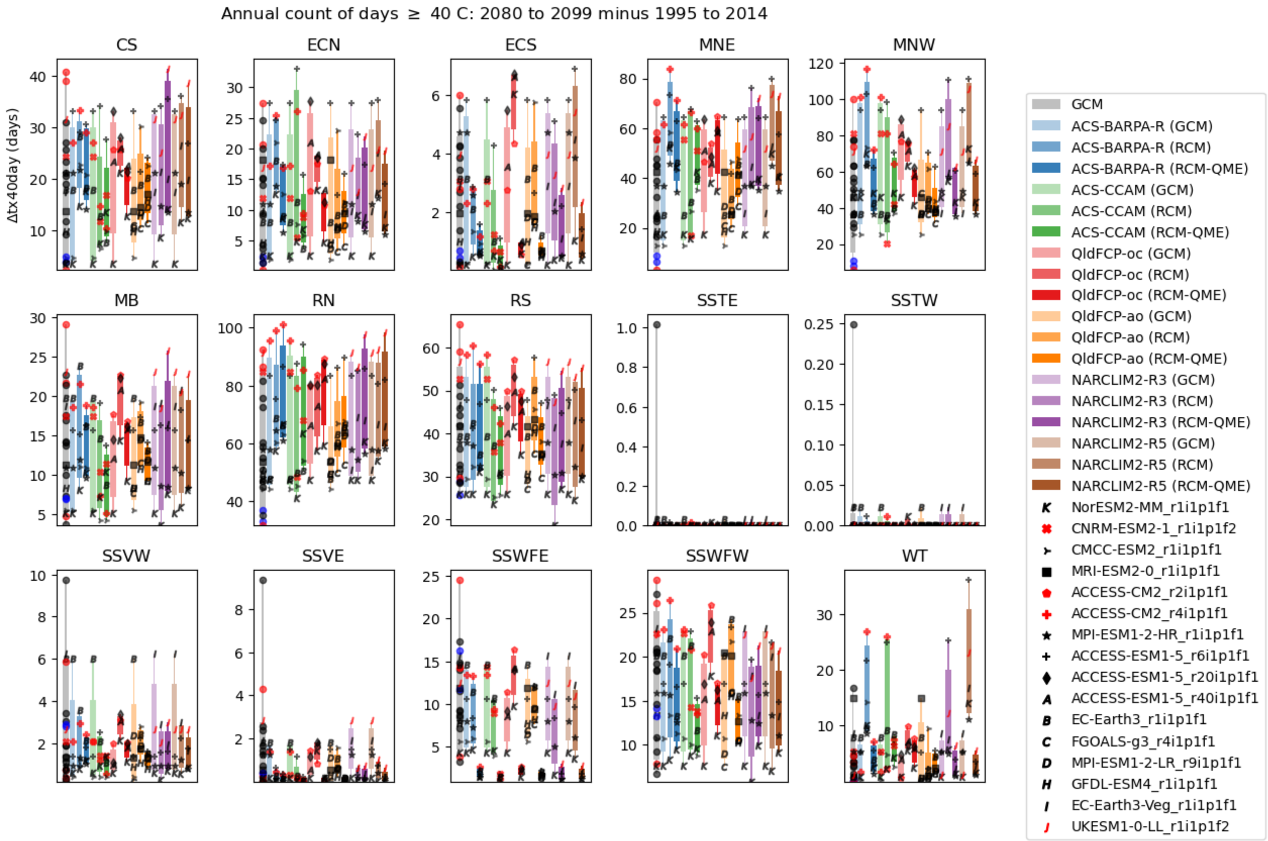

Figure 8.2 As for Figure 8.1, but showing the change in average number of days exceeding 40 °C.

8.2 Extreme rainfall

An initial set of indices for extreme rainfall were calculated following the Expert Team for Climate Change Detection and Indices (ETCCDI, https://www.wcrp-climate.org/etccdi).

- Highest annual 24-hour total (Rx1day) - this is the highest daily rainfall falling at each grid point in a year
- Highest annual 5-day total (Rx5day) - this is the highest five-day rainfall total falling at each grid point in a year
- Highest annual hourly total (Rx1hour) - this is the highest rainfall falling over one hour at each grid point in a year

Extreme rainfall is a contributing factor to flood hazards and is used for designing critical infrastructure from storm drains to dams (Wasko et al. 2024). This set of indices also enables the easy calculation of derived products such as rarer rainfall events with lower Annual Exceedance Probabilities using Generalised Extreme Value methods.

Daily extreme rain indices are calculated using the QME bias-corrected data. For Rx1hour, a bias corrected dataset of daily maximum hourly rainfall was derived by multiplying the bias corrected daily rain by the ratio between the wettest hour and daily total in the raw model output. This approach assumes that the diurnal distribution of rainfall intensity is correct in the regional model, which may not be true.

For the CORDEX ensemble, the median projection has an increase in the 20-year average Rx1day between the historical baseline and end of century across all of Australia, although model agreement is low for individual locations (Figure 8.4a). However, there is large spatial heterogeneity in individual model ensembles, with some RCMs producing largest increases in southern Australia (Figure 8.4c, d) while others have strongest increases in northern Australia (Figure 8.4e).

Averaged across Australia, there is a *very likely* (92% agreement) increase in the intensity of Rx1day, with an ensemble median increase of 11% between the historical baseline and end of century, or +4.3%/K. Both the BARPA-ACS and CCAM-ACS ensembles have an Australian average projected increase of +7%/K, which is comparable to that expected from thermodynamic principles (e.g. Wasko et al. 2024). In contrast, projected increases in Rx1day are generally below Clausius-Clapeyron for both the QldFCP-2 and NARCliM2.0 ensembles, at ~+3%/K or ~8% over the century (Figure 8.4d, e), which is comparable to the median change projected from CMIP6 models (Grose et al. 2020).

There is also a *likely* (82% agreement) increase in the Australian-average magnitude of Rx5day. This change is slightly smaller than Rx1day, with a median increase of 8% between the historical baseline and end of century, or +2.8%/K. As with Rx1day increases are broadly projected across Australia, except for a small area of declines projected in southwest WA (not shown).

The Australian-average magnitude of Rx1hour is also *very likely* to increase (Figure 8.5), with 92% agreement among models, supported by theory and process-based understanding. However, the magnitude of change is less confident than the direction of change. The ensemble-median projected change in Australian mean Rx1hour by the end of the century is +17% or +5.7%/K, with the ACS models projecting similar changes to Rx1day but an increased rate of increase for the QldFCP-2 and NARCliM2.0 models, although these remain below Clausius-Clapeyron (4-5%/K). The median change is well below the best estimate of change in this index from multiple lines of evidence at +15% per degree of warming (Wasko et al. 2024). This suggests that current regional modelling may not be able to represent key processes for the generation of extreme sub-daily rainfall such as thunderstorms and may be improved on in upcoming convection-permitting projections for Australia, and that more targeted bias adjustment of sub-daily data may be necessary.

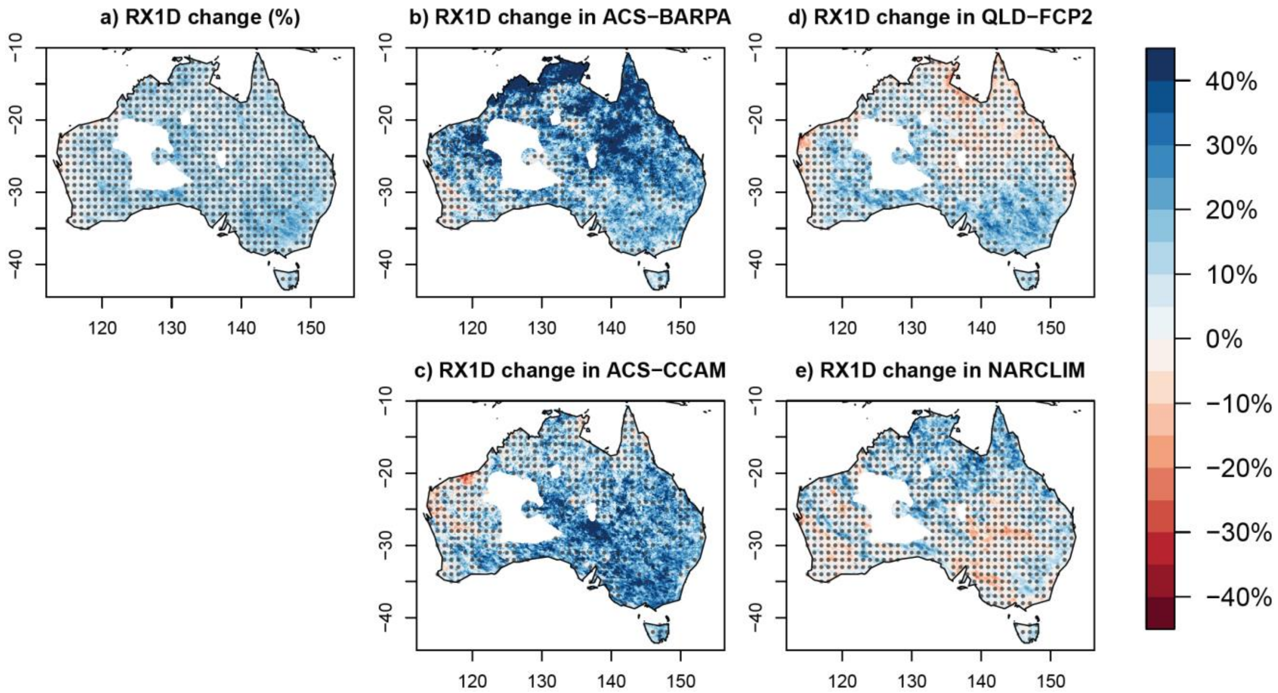

Figure 8.4. Multi-model median % change in Rx1day between 1995-2014 and 2080-2099 under SSP3-7.0 for a) the CORDEX-CMIP6 ensemble as well as the four different RCM ensembles: b) ACS-BARPA, c) ACS-CCAM d) the QldFCP-2 ensemble, and e) the NARCliM2 ensemble. *Results are presented as the ensemble median (i.e., the ensemble central estimate), with stippling indicating where fewer than 80% of ensemble members agree on the sign of the change*. The projections have been bias-adjusted using the Quantile Matching for Extremes (QME) method (Section 5). No results are shown in regions of inland Australia with low observation density where bias correction of precipitation data is challenging.

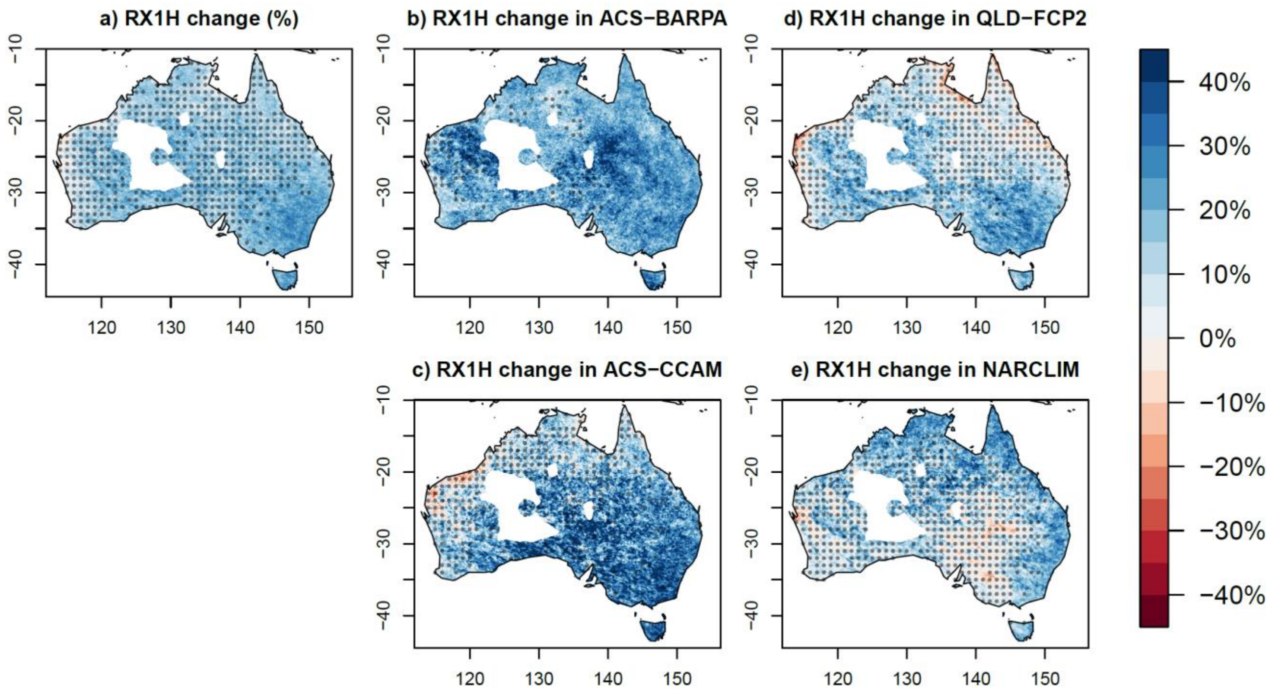

Figure 8.5 Multi model median % change in Rx1hour between 1995-2014 and 2080-2099 under SSP3-7.0 for a) the CORDEX-CMIP6 ensemble as well as the four different RCM ensembles: b) ACS-BARPA, c) ACS-CCAM d) the QldFCP-2 ensemble, and e) the NARCliM2 ensemble. *Results are presented as the ensemble median (i.e., the ensemble central estimate), with stippling indicating where fewer than 80% of ensemble members agree on the sign of the change*. The projections have been bias-adjusted using the Quantile Matching for Extremes (QME) method (Section 5). No results are shown in regions of inland Australia with low observation density where bias correction of precipitation data is challenging.

8.3 Extratropical cyclones

Projections of the meteorology behind important hazards may be derived from the CMIP6-based Next Generation of climate change projections presented earlier in this report. Often the calculation of these hazards will depend on atmospheric fields beyond simply the commonly used surface variables. An example of one such derived hazard, extratropical cyclones, is presented here for illustration purposes. Low pressure systems are a major cause of several hazards in Australia including heavy rainfall, strong winds, flooding and coastal erosion, and compound extremes (Dowdy et al. 2019).

For initial analysis in the ACS, all low-pressure systems were identified from gridded 6-hourly mean sea level pressure data using the University of Melbourne tracking scheme (Simmonds et al. 1999, Pepler and Dowdy 2022). No attempt was made to distinguish between lows of extratropical or tropical origin; however, each low was required to persist for at least 6 hours and have an associated low detected at 500hPa at least once to exclude weaker, shallower systems which are less likely to produce significant impacts, such as heat lows (Pepler and Dowdy 2020). All CORDEX simulations were considered, and for this analysis raw data was available from an additional member of the QldFCP2 ensemble, with EC-Earth3 downscaled in the coupled ocean approach (so, making a total of 40 rather than 39 simulations).

Cyclone tracking is applied to the raw Mean Sea Level Pressure (MSLP) data, and results have not been bias-adjusted. Compared to the BARRA-R2 reanalysis, the ACS model ensemble consistently underestimates the historical frequency of extratropical lows in southern Australia, particularly during the cool half of the year (Figure 8.6). This is a known bias arising from biases in global climate models and has been also identified in both global climate model output and previous generations of regional downscaling (Pepler and Dowdy 2022). Biases in southern Australia are broadly consistent between the four modelling agencies, but there are significant differences in tropical regions. While the ACS-BARPA and ACS-CCAM models tend to produce too many lows in tropical regions, the Qld-FCP2 models tend to underestimate low frequency.

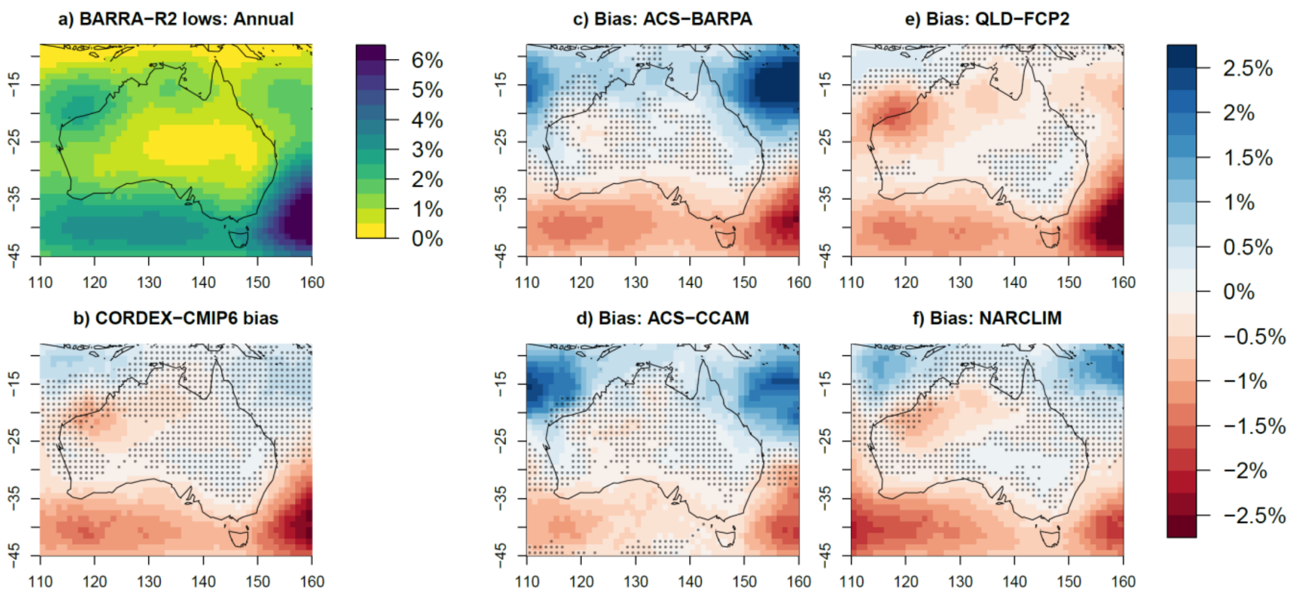

Figure 8.6 (a) BARRA-R2 proportion of annual observations with a low centre within a 5-degree radius, 1995-2014. (b) Ensemble median difference (absolute bias) between low frequency in the CORDEX-CMIP6 ensemble and BARRA-R2. Also shown the median biases for four different RCM ensembles: c) ACS-BARPA, d) ACS-CCAM e) the QldFCP-2 ensemble, and f) the NARCliM2 ensemble. Stippling indicates where fewer than 80% of ensemble members have a bias of the same sign.

In a warmer climate, there will *very likely* be a decrease in the frequency of lows in southern Australia (30-45°S, 110-155°E). This confidence is given due to very high levels of model agreement (100%), as well as consistency with projections from previous generations of global and regional climate modelling (e.g. Pepler & Dowdy 2022). The median change across all 40 RCMs between the historical baseline and end of century is -34% or -10% per degree of warming (Figure 8.7), with broadly similar declines in both the warm and cool halves of the year. In contrast, there is no model agreement on the magnitude of changes in northern Australia, where lows predominantly occur during the warm half of the year, with declines in tropical lows projected by the QLD-FCP2 models (Figure 8.7d) but little change or small increases in other RCMs.

The overall impact of lows depends not only on the frequency of lows, but also the intensity (including the proportion in the higher intensity categories), the size, translation speed, intensification near the coast, peak wind speed and the rain rate for a given sized storm. These are not reported here and will be covered in more detail in further products.

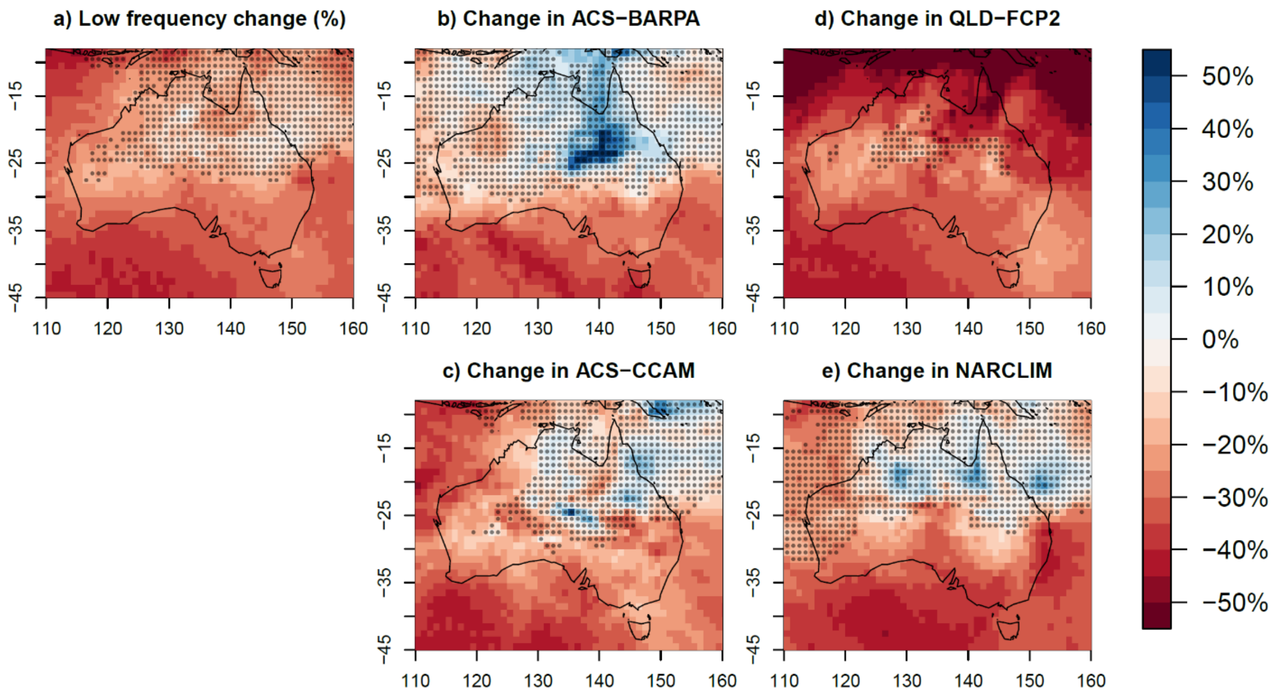

Figure 8.7 Multi model median % change in low frequency between 1995-2014 and 2080-2099 under SSP3-7.0 for a) the CORDEX-CMIP6 ensemble as well as the four different RCM ensembles: b) ACS-BARPA, c) ACS-CCAM d) the QldFCP-2 ensemble, and e) the NARCliM2 ensemble. *Results are presented as the ensemble median (i.e., the ensemble central estimate), with stippling indicating where fewer than 80% of ensemble members agree on the sign of the change*.

8.4 Meteorological drought

Drought refers to a temporary period of abnormally dry conditions and is a recurring natural phenomenon that profoundly influences Australia's agricultural sector, water resources, ecosystem health, and socio-economic stability (Van Dijk et al., 2013). Beyond its immediate effects, drought acts as a catalyst for heatwaves and severe fire seasons (Ruthrof et al., 2016).

Several indices exist to identify different drought types (Zargar et al. 2011), with the Standardized Precipitation Index (SPI) being the most widely used one. The SPI is typically used to identify and quantify the severity of meteorological droughts across various time scales (months to years) and is thus endorsed by WMO for use in water resource management, agriculture, and climate studies for its simplicity and effectiveness (WMO, 2012). The SPI measures the amount of precipitation over a specific period relative to the long-term average for that period. Though, strictly it is a meteorological drought index, the adjustable time period over which precipitation is aggregated has proven to be a good proxy for other types of droughts like agricultural (3-6 months) and hydrological (6-36 months) drought (McKee et al, 1993). Positive SPI values indicate wetter-than-average conditions, while negative values indicate drier-than-average conditions.

Here we use SPI3, meaning precipitation is aggregated over three months, and a value of -1 or lower as a drought threshold, signifying moderate drought conditions (McKee et al, 1993). While this metric does not inform about the length and frequency of drought it gives an indication whether a location will experience more or less time in moderate drought conditions.

Figure 8.8 shows the proportional (%) change in time spent in drought (proportion of months with SPI values of ≤ -1) from 1995-2014 to 2080-2099 for the full set of CORDEX simulations, using both QME (top row) and MRNBC (bottom row) bias adjustment. Results are presented as the multi-model median for the full ensemble (left column) and individual RCM ensembles, hashing shows where at least 66% of ensemble members *agree* in sign of change (note different scheme used for this purpose compared to elsewhere). Generally, RCMs generally agree on a strong increase of time in drought for south-western WA (>75%) and in large parts for other regions in southern Australia (>25%) except for ACS-CCAM. Similarly, north-eastern Australia along the Queensland coast and adjacent inland regions are projected to experience about 50% more time under moderate drought conditions. In contrast, confidence is low for changes in most of northern and inland Australia, including the Murray-Darlin Basin, due to the large uncertainty between RCMs ranging from halving (e.g. UQ-DES and ACS-BARPA) to a 50% and more increase in time spent in drought.

The bias-adjustment method applied to the data does not affect the agreement between models or sign of change but does affect the magnitude of change for regions with large negative changes, where QME results in more extreme negative changes, particularly for southwest and southeast Australia.

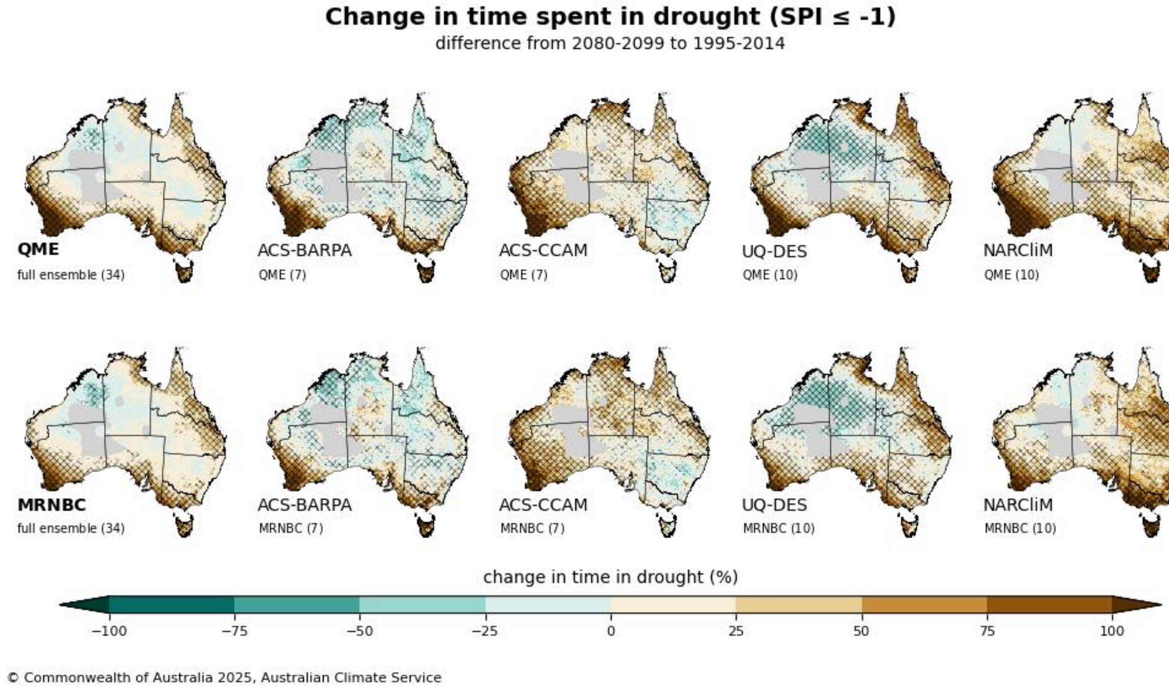

Figure 8.8 Change in 'time spent in drought' (SPI3 < -1) between 1995-2014 and 2080-2099 under SSP3-7.0. Results are presented as the ensemble median (i.e., the ensemble central estimate). Left column represents the full CORDEX ensemble but including one member of the Queensland ensemble from each CMIP6 host (so 34 rather than 39 members) with the columns to its right showing ACS-BARPA (7), ACS-CCAM (7), UQ-DES (10) and NARCliM2.0 (10) RCM ensembles. The projections in the top row have been bias-adjusted using the QME method. The projections in the bottom row have been bias-adjusted using the MRNBC method. Hashing shows where at least 66% of the models agree on the sign of change (noting different scheme used for this purpose compared to elsewhere). Inland areas with sparse observations impacting bias-adjustment have been masked out.

8.5 Bushfire weather

Bushfires and grassfires are natural and essential to most Australian ecosystems. Wildfires occur when there is plentiful dry vegetation to burn, conducive weather and climate conditions and an ignition source. Changes to susceptibility of the landscape to fire is driven by changes to vegetation from rainfall and temperature, as well as drought, heatwaves and fire weather, as well as ignitions (e.g., dry lightning), making future predictions of change in fires complex and uncertain.

The behaviour of wildfires, such as the rate of spread and intensity, are influenced by factors such as wind speed, temperature and humidity, recent rainfall, and fuel characteristics, such as the amount and how dry ('available') the fuel is. The Forest Fire Danger Index (FFDI) was originally developed to characterise fire spread in a forest vegetation, using information on the weather to predict fire spread (McArthur 1967). 'Fire weather' refers to weather that is conducive to fire spread. The FFDI calculated here provides a measure of fire weather, it is a function of wind speed, temperature, humidity, and recent rainfall. We then calculate the number of days FFDI is greater than 50, when fire weather was historically categorized as 'severe', to examine one aspect of how climate change will influence fire weather across Australia (Figure 8.9). While FFDI is most appropriate for forest environments, the index is commonly used as a general indicator of dangerous fire weather, and change in those days, across Australia (whether there is forest vegetation or not).

Since the 1950s, large parts of the country have experienced increased dangerous fire weather and a longer fire season (BoM & CSIRO, 2024).

Model analysis includes CCAM-ACS, BARPA-R and NARCliM2.0 modelling, surface wind was not available from Qld-FCP2 at the time of analysis. All the RCM ensembles (using QME bias adjustment of inputs) are similar in the historical climate, with a realistic simulation of the frequency of severe days and reproducing the large variation across the country (Figure 8.9, left column). The frequency in NRM sub-clusters ranges from zero in Southern Slopes Western Tasmania (SSTW), through 10 to 15 in Murray Basin (observed is ~15), to 65 to 75 per year in the arid Rangelands North (observed is ~75). The projected change in the frequency of severe fire weather danger days (>50 FFDI) is also similar between RCM ensembles, with a consistent increase across Australia (Figure 8.9, right columns). Changes are almost unanimously positive, but with restricted regional exceptions for individual model simulations (associated with climate variability between these 20-year periods). Multi-ensemble-mean projected change in the national average is >20 days per year (a ~50% increase), but with some regions showing more than doubling of days. Changes are lower for SSP1-2.6 (e.g., national average ~30%).

Projected increases in FFDI are broadly consistent with previous findings. An important note in interpreting the results is that the greatest increases in severe fire weather days occur in central Australia's hot, dry regions where grassland and savanna dominate rather than forest. Changes in FFDI outside forests should be interpreted only as a general indication of fire weather, rather than a meaningful index of risk. In southern Australia, where FFDI is more appropriate, severe fire weather expands south and eastward, approximately doubling the number of severe fire weather days in Victoria, NSW and southwest WA.

This section covers only projections of FFDI from the full set of new modelling, further analysis of overall fire risks incorporating all contributing factors, and of changes to bushfire classes is available in further work.

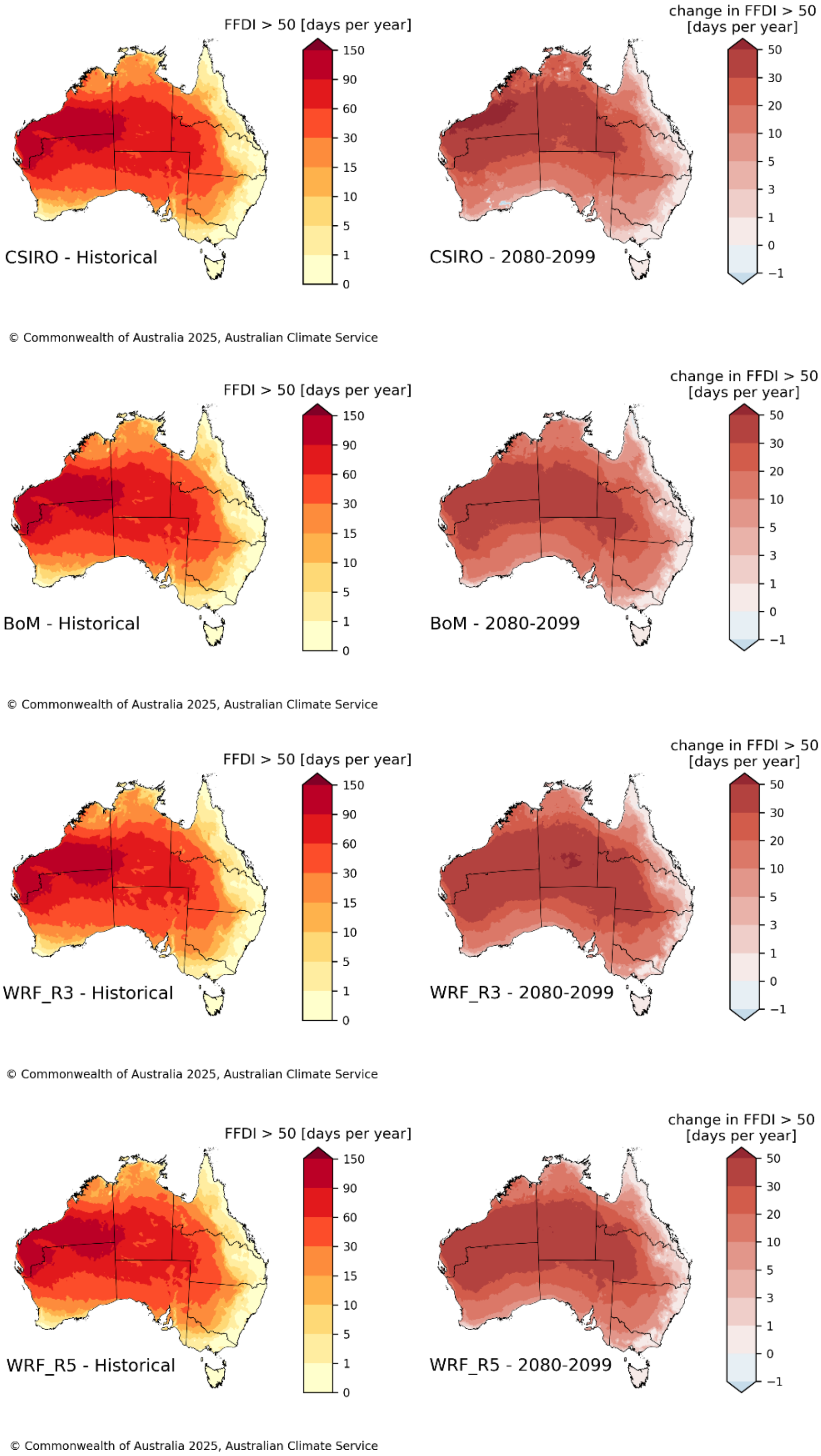

Figure 8.9. Projected annual (January to December) number of severe fire weather days (FFDI > 50) in the recent baseline period (1995-2014) and the projection change in the number of severe fire weather days by the end of the century (2080-2099) under SSP3-7.0 from members of CORDEX: CCAM-ACS (top), BARPA-R (second row) and NARCLIM2.0 R3 and R5 (bottom two rows).

9. Gaps, future opportunities and challenges

There are gaps, opportunities and challenges remaining for regional climate change projections for Australia and their use in climate risk assessment and decision making. Here we describe some key concerns that relate to the underpinning climate science of projections. We do not address many of the scientific challenges specific to the analysis of hazards and their impacts, which will be discussed in other technical reports in this series.

As noted in Chapter 1, projections should aim to be credible, salient, legitimate and decision relevant. The current projections design has a strong focus on regional modelling, aiming to enhance the credibility of projections in terms of extremes, and the salience of projections in terms of local relevance. Further improvements could be made to enhance these further. Legitimacy and decision relevance are broader questions than we tackle in this Technical Report.

9.1 Dynamical downscaling model ensemble design and configurations

Here we focus strongly on a multi-model ensemble of RCMs, partly designed by the program and partly through partnership, creating an 'ensemble of opportunity'. A more centralised and coordinated ensemble design would obviously help with more systematic sampling of the projections space. In the shorter term, exploration of model weighting schemes rather than default equal weighting (noted in Section 4.6) should be explored to see if this refines projections in any meaningful way.

As part of improvements to an RCM ensemble design, the different configurations of RCMs could be more systematically tested. At present there are multiple frameworks in use for regional dynamical downscaling which each make subjective choices on methodology. For example, whether to bias adjust the global driving model before applying the data as a boundary forcing for the regional model (as is done in the QldFCP-2 CCAM experiments), or to instead bias adjust the regional model outputs after directly applying the global climate model at its boundaries (as is done in the ACS and NARCliM2.0 experiments). Also, the effects of including a regional ocean model, and of coupling regional atmosphere model to the surface ocean that has been done needs further investigation, along with future developments in this technique. These choices, and any others involved in the production of RCM simulations (e.g., model grid configurations, use of different model schemes), could be more systematically tested and understood.

9.2 Alternative forcing scenarios

Projections must be made for plausible scenarios of future human development and resulting emissions, rather than any kind of single or probabilistic prediction. A standard set of scenarios is considered in the IPCC AR6 report, the SSPs, and here we use a simple framework of 'low' and 'high' scenario subset from this (SSP1-2.6 and SSP3-7.0). While these scenarios generally span the range of plausible future emissions trajectories and therefore allow for future planning and decision making by informing risk, they do not cover all plausible and important cases. One of the key unknowns now facing policy makers is the implications of different mitigation scenarios to reach an eventual lower greenhouse gas condition. For example, policymakers now want to understand the implications of 'overshooting' the target greenhouse gas level before applying negative emissions approaches and then stabilizing at the desired level, including what is effectively irreversible (e.g. Schleussner et al. 2024). Such scenarios are now the focus of global climate modelling efforts, and the interpretation of these scenarios for local and regional purposes may require dynamical downscaling in future.

An approach that is applied elsewhere is the 'pseudo global warming level' method (e.g. Brogli et al. 2023). Rather than downscaling global climate models directly, the pseudo global warming level approach is to apply salient climate change scaling from GCMs to a current climate control simulation (for example based on observations or reanalyses), which can then be used as a boundary forcing for regional models to produce a simulation of a hypothetical future regional climate at that global warming level. This approach has the added advantages of approximately representing realistic internal variability and can reduce the number of experiments to just a subjectively representative set. Such an approach could be conducted using the regional models available in the ACS if this framework eventually becomes an international and regional standard methodology. This approach applied to reanalysis has the disadvantage of not allowing future circulation changes, i.e., future storms are past storms with warming and thermodynamic effects added.

9.3 Unrepresented plausible regional climate responses to global warming

Our current modelling may not produce the full range of possible climate change resulting from any particular forcing scenario. One of the major strengths of the CMIP exercise is the standardization of modelling approaches, which has allowed a much clearer insight into the actual diversity of modelled responses to greenhouse gas forcing. At the regional scale this has unfortunately shown that very significant uncertainty remains in terms of constraining the regional climate change signal for many key variables including rainfall and weather extremes (if indeed such a constraint is possible). Some of the modelled climate change responses may be found to be implausible or the results of error, leading to a reduction in the projected change. However, the wide range of forced responses in models is still likely to be missing some plausible climate responses.

One well known category of climate response that may be underestimated is that of abrupt change, and so-called global climate tipping elements in the Earth system. Reviews have found that several tipping points are plausible based on evidence other than CMIP GCMs, including observations, theory, and other types of modelling (CSIRO 2024; Armstrong-McKay et al. 2023). These tipping elements include processes that are challenging to simulate or constrain in a GCM, including large-scale changes to tropical rainforests, ocean circulation, and ice sheets. A paucity of process understanding and observations for constraining them leaves open much more dramatic and rapid changes than is currently seen in our models.

Another important emerging concern is 'the El-Niño like' pattern response in climate models (where the tropical Pacific warms faster than the rest of the tropical oceans), which is in apparent contradiction to the 'La Niña-like' recent observed climate trends (where the tropical Pacific has not warmed as fast as the rest of the tropical ocean regions, e.g. Seager et al. 2019). While currently there is debate in the scientific literature on whether the observed changes are internal variability or a forced climate change response, in either circumstance there are implications for future scenarios. In the case where the models are systematically biased in their forced response pattern, we are likely underestimating regional changes to some variables and hazards, in particular wet extremes for some parts of Australia. In the case where the observed trends are due to multidecadal variability a rebound into the opposing phase of multidecadal variability will likely be associated with a very rapid warming in coming decades.

To address these unrepresented climate responses, we need better process-based understanding, more understanding of representing these processes in models, large systematic model ensembles, and more.

9.4 Other types of models or ensembles

9.4.1 Large ensembles

When investigating the climate change response in a model we subjectively choose time periods and scenarios for specific application purposes. However, the modelled climate and therefore the analysed change signal, is subject to internal variability (both forced and unforced) that may affect the calculated change between two arbitrary periods. The shorter the temporal sample the more likely it is that such an effect occurs.

One way to address this concern is to employ large ensembles of each climate model used. The large ensembles are comprised of many different runs of the same model with slightly different starting conditions. The ensemble for each model provides a representation of the role of internal variability in the model. For some variables like mean temperature change, the role of internal variability can be small relative to the forced change. However, for other variables like rainfall, the internal variability can be comparable to the forced response in the change scenarios.

In this report we provided some key insights from two important CMIP6 large ensembles (Section 7.4), but there is limited use of large ensembles in regional modelling. ACS used only a single member (run) per model and so do not provide an estimation of the role of internal variability. Qld-FCP2 provide multiple members for ACCESS-ESM1.5, but these also use two model configurations. Future work will address this issue and can draw upon the growing number of large ensemble projection simulations (Deser et al. 2020). Dynamical downscaling of all members of a large ensemble is not practical due to computational limits, but AI/ML downscaling can potentially be applied (see Section 9.4.2). Large ensemble approaches are also used now in climate change attribution to more adequately sample a representative historical period, including the `UNSEEN' approach to assessing the likelihood of extremes (Thompson et al. 2017; Irving et al. 2024). Beyond simply running standard scenarios, there are new efforts to better represent the effects of individual forcings through large ensemble approaches (Smith et al. 2022) which will aid in characterising the ways in which internal variability and forcing interact.

9.4.2 Emulators and AI/ML (including for alternative bespoke scenarios)

The practice of producing regional climate projections in Australia has traditionally used dynamical models such as GCMs and RCMs, with limited use of statistical modelling and calibration and no use of tools such as simple models and emulators. Projections haven't yet extensively made use of traditional emulators, and the rapidly emerging field of artificial intelligence and machine learning (AI/ML). Emulators and AI/ML have many advantages in the production of projections (e.g. Baño-Medina et al 2024). While there are many non-trivial challenges yet to overcome (see Rampal et al. 2024), current evidence suggests that AI/ML should be part of any strategy moving forwards, in areas such as:

- Emulation of GCMs to boost ensemble sizes and to emulate novel scenarios
- ML emulation of downscaling – where dynamical RCM data can be the training dataset for ML, then ML is used to produce projections data with greater efficiency of many orders of magnitude (see Rampal et al. 2024 and references therein).
- 'Hybrid' dynamical-ML models such as Neural GCM (Kochkov et al. 2024).
- Use of ML in bias adjustment (e.g., Yoshikane and Yoshimura 2023).
- Cutting edge and emerging uses of AI in climate modelling and across the spectrum of climate change related areas (areas not already identified).

It is clear that emulators and AI/ML will be part of any future climate projections work, given the efficiency and sample size that they offer. This means new and appropriate ensemble design strategies must be developed that are different from the current generation. As one example, we may want to run RCMs not to provide the complete production/operational dataset but rather run targeted simulations to provide a robust training dataset for ML to then produce the dataset for all the response and scenario dimensions – including for bespoke scenarios. This requires new design strategies to optimize RCM simulations for training rather than production. This is just one example of the new strategies needed.

9.4.3 Convection permitting modelling

The ACS dynamically downscaled projections presented in this technical report simulate regional climate at a much higher resolution than GCMs (~100 km grid spacing), however they are coarse enough to still require convection parameterisation (~10 km grid spacing). By advancing the resolution down to ~1 km grid spacing or finer, convection may be explicitly permitted in the models, partially addressing a key uncertainty in the processes associated with regional rainfall and circulation change. This can either be through the ambitious global k-scale modelling effort, through high resolution RCMs, or possibly using hybrid dynamical-machine learning models. However, running convection permitting models is far more expensive in terms of both computation and storage, presenting a challenge for individual programs (particularly for global k-scale modes, but also for high resolution RCMs). Additionally, new modelling biases, errors, and puzzles emerge when conducting convection permitting modelling, and further research and development would be necessary.

Nevertheless, convection permitting modelling presents unique opportunities to users of climate projections who require local information at very fine scales, or whose applications are greatly informed by certain types of hazards such as extreme sub-daily rainfall or strong wind gusts. Other climate modelling programs have therefore already provided users with convection permitting simulations e.g. the NARCliM2.0 downscaled projections at 4 km resolution. Convective permitting modelling using CCAM and BAPRA-C is well advanced, and research simulations are underway. This will yield insights with further research and development. Intercomparison and coordination of convection-permitting modelling from CCAM, BARPA, NARCLIM2.0, ACCESS simulations at

400 m from the ARC Centre of Excellence in Climate Extremes and others will be useful in progressing the area. Monitoring developments in global k-scale modelling, including through the ARC Centre of Excellence on 21st Century Weather, and preparing to use the outcomes in action is also a priority.

9.5 Historical reference for bias adjustment and evaluation of projections

The bias adjustment of climate projections, required due to inherent biases in the climate models, is an important step to produce some of the application-ready data. This process relies on accurate observational data sets serving as historical references to quantify these biases and evaluate the climate models. Gridded analyses of station-based observations and reanalysis, which assimilates many types of observations concurrently, are typically used. The Australian Gridded Climate Data (AGCD) version 2 and the Bureau's atmospheric regional reanalysis for Australia version 2 (BARRA-R2) are developed as part of the ACS, with AGCD serving as references for daily temperatures, precipitation and relative humidity, while BARRA2 can also provide references for many other less well-observed or unobserved meteorological variables (e.g., wind and gust). With the intensities and durations of weather hazards changing under climate change, these re/analysis products need to further develop to provide better information on the sub-daily timescales and extremes at these timescales. BARRA-R2 is a significant improvement to BARRA version 1 by providing a longer national view. However, BARRA-R2 remains limited in horizontal resolution, at the upper limit of convection-permitting modelling regime. It can be further improved through better assimilation of more observations, to provide more accurate analyses of the local-scale environments that led to past extreme events. The emerging km-downscaled reanalysis BARRA-C2 (Su et al. 2024) should be explored and used as appropriate.

9.6 Downstream uses of CMIP6-based regional modelling

The ACS has undertaken an extensive effort to create reliable nationally consistent dynamically downscaled projections of Australia's future climate, and alongside other significant efforts there now exists a multi-RCM CORDEX-Australasia ensemble for use in informing decision making and future planning. However, beyond an initial set of hazard indicators there is still much to be done to extend these new CMIP6 based regional projections to other areas where projections are produced.

The first set of enhancements is to further develop communication and visualization resources to understand the existing analyses and to make them accessible and usable for applications and on to decision makers. This includes further development of web-based exploration, selection and data delivery tools.

The second area for further work is to facilitate the use of the new projections into downstream modelling in areas such as agriculture, hydrology, fire management and many more. Examples include an update of the National Hydrological Projections (awo.bom.gov.au) to use the new projections, as well as updated modelling of crop yields (and presentation of agricultural indices) for the Climate Services for Agriculture platform.

9.7 Reducing uncertainty in projections

Projections cannot overcome the epistemic uncertainty of forcing scenarios, since that is the product of human actions. In fact, illustrating the effect of different scenarios to inform action is an important purpose of projections. However, there is strong interest in reducing the uncertainty in the forced response, and one of the key problems in climate science for the last few decades has been a stubborn inability to constrain the forced response of the climate at the regional scale. This is partly due to insufficiencies in our ability to model the climate reliably, but also partly due to an inability to detect a signal of climate change in observed climate records.

With time, the signal of forced climate response grows due to continued increases in anthropogenic greenhouse gases in the atmosphere, and other changes to atmosphere and ocean composition. This emergence of the forced climate response in the observed record may allow in some important cases for the climate change response to be constrained (sometimes referred to as emergent constraints). Once past projections can be verified, we will hopefully be able to better constrain the future forced response of the climate at regional scales (e.g. Shaw et al. 2024a, b). Further improvements in modelling, further emergence of climate change 'signals' with time, as well as a stronger focus on *verification* of projections (rather than just standard model evaluation and benchmarking) will be useful in coming years. Effort in this area should be pursued, although it remains to be seen how strong the resulting constraints will be – but we won't know until we try.

9.8 Computation and data challenges

Dynamical climate modelling uses a lot of resources in terms of supercomputing runtime and upgrades to the computers, as well as data storage. The needs are only growing.

The amount of data developed for the ACS project is more than 3 petabytes, with approximately 2.72 petabytes accounting for published CCAM and BARPA data as well as bias corrected data. CCAM itself generated an additional 3 petabytes of output and restart files which are important to ensure that experiments can be reproduced and need to be archived. This has required 68 million service units to generate the CCAM simulations and both the data and computational costs will only increase as RCMs move to higher resolutions. Computational costs for BARPA are approximately the same for its 17.5 km resolution simulations. For BARPA and CCAM combined, it is estimated that 4 petabytes and 216 million service units will be required for 4 km, convective permitting downscaling. This is despite the convective permitting simulations being run for a shorter period, roughly 460 simulation years versus 3200 years for the 12.5/17 km resolution data. Computation and data storage for ongoing maintenance of the existing and new model simulations is an ongoing challenge that must be managed prudently. Developments such as AI/ML may reduce the computing and storage costs compared to RCMs and should be explored fully.

The ACS project has been a major supporter and driver of the CORDEX-CMIP6 experiment. Following the CORDEX structure and specifications has helped with consistency and international credibility and comparability. The ability to intercompare projections from ACS to other members from NARCLIM2.0 and Qld-FCP2 is a valuable leap forward in consistency compared to previous *ad hoc* arrangements. There have been barriers though, including uncertainty over data specifications, publication standards, formatting, and structure. This can be addressed in future work through improved standards from CORDEX itself, and within Australia to ensure that changes are communicated clearly and timely to avoid unnecessary confusion and reprocessing of data. An example of this is data chunking required for RCM output. There is no defined or specified method for data chunking which affects the time required to compute certain metrics depending on how data is chunked. For example, CCAM chunks by time whereas BARPA chunks using a combination of time, latitude and longitude and this has adversely affected the processing time for the calculation of the Forest Fire Danger Index FFDI and TC tracking, respectively.

9.9 Next steps

The existing model ensembles and evidence base will provide a useful basis for risk assessments and climate change adaptation planning for public and private sector applications in coming years. Also, with further analyses and refinements, there will be further insights and intelligence over coming months and years.

With further refinement and developments listed in the chapter, further intelligence can be produced. Insights from further work can supplement the current work through minor ongoing updates and major updates when appropriate. As well as advances and changes from work done in Australia, international developments need to be addressed, in two areas:

1. New international standard scenarios, such as the CMIP7 ScenarioMIP (van Vuuren et al. 2025), including more focus on emissions-driven scenarios and on overshoot
2. New and emerging issues of interest, e.g., irreversible, nonlinear, self-amplifying and relatively abrupt changes driven by positive feedback dynamics etc.

The new CMIP7 ensemble (Dunne et al. 2024) will be a major new feature in the landscape over coming years, including the Assessment ‘fasttrack’ with simulations for historical and the new scenarios. Also, insights from CMIP7 CommunityMIPs will also be a valuable resource, including on particular science issues of highest interest including (but not limited to) Detection and Attribution (DAMIP), Decadal Climate Prediction Project (DCPP), FireMIP, HighResMIP, Tipping Point Modelling (TIPMIP) and WhatifMIP.

10. References

Armstrong McKay DI, Staal A, Abrams JF, Winkelmann R, Sakschewski B, Loriani S, Fetzer I, Cornell SE, Rockström J, Lenton TM, (2022). Exceeding 1.5 °C global warming could trigger multiple climate tipping points. Science 377(6611). https://doi.org/10.1126/science.abn7950

Baño-Medina, J., Iturbide, M., Fernández, J. and Gutiérrez, J.M., (2024). Transferability and explainability of deep learning emulators for regional climate model projections: Perspectives for future applications. Artificial Intelligence for the Earth Systems, 3(4), p.e230099.

Beck, H. E., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg and E. F. Wood (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data 5(1): 180214.

BoM and CSIRO, (2024). State of the Climate 2024, ISBN PRINT 978-1-4863-2125-4, pp 32.

Brogli, R., Heim, C., Mensch, J., Sørland, S.L. and Schär, C., (2023). The pseudo-global-warming (PGW) approach: methodology, software package PGW4ERA5 v1. 1, validation, and sensitivity analyses. Geoscientific Model Development, 16(3), pp.907-926.

Cash, D. W., W. C. Clark, F. Alcock, N. M. Dickson, N. Eckley and J. Jäger (2002). Salience, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making. KSG Working Papers Series RWP02-046.

Chapman, S., J. Syktus, R. Trancoso, M. Thatcher, N. Toombs, K. K.-H. Wong and A. Takbash (2023). Evaluation of Dynamically Downscaled CMIP6-CCAM Models Over Australia. Earth's Future 11(11): e2023EF003548.

Chapman, S., Syktus, J., Trancoso, R., Toombs, N., & Eccles, R. (2024). Projected Changes in Mean Climate and Extremes from Downscaled High-Resolution CMIP6 Simulations in Australia. SSRN. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4836517

Chen, D., M. Rojas, B.H. Samset, et al. (2021). Framing, Context, and Methods. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 147–286, doi:10.1017/9781009157896.003

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., et al. (2018). The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation. Journal of Advances in Modelling Earth Systems, 10, 2731–2754. https://doi.org/10.1029/2018MS001354

CORDEX (2021). CORDEX experiment design for dynamical downscaling of CMIP6. Technical report, May 2021: https://cordex.org/2021/05/24/experiment-protocol-rcms-is-published/

Corre, L., Ribes, A., Bernus, S., Drouin, A., Morin, S., Soubeyroux, J-M., (In review). Using Regional Warming Levels to Describe Future Climate Change for Services and Adaptation: Application to the French Reference Trajectory for Adaptation. Available at SSRN: https://ssrn.com/abstract=5025864 or http://dx.doi.org/10.2139/ssrn.5025864

CSIRO (2024). Understanding the risks to Australia from global climate tipping points. CSIRO Workshop Report, Australia.

CSIRO and Bureau of Meteorology (2015). Climate Change in Australia, Technical Report. Melbourne Australia. www.climatechangeinaustralia.gov.au

DCCEEW (2023). National Climate Risk Assessment, phase 1 report https://www.dcceew.gov.au/climate-change/publications/national-climate-risk-assessment

Deser, C., Lehner, F., Rodgers, K.B. et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. Nat. Clim. Chang. 10, 277–286 (2020). https://doi.org/10.1038/s41558-020-0731-2

Dey, R., S. C. Lewis and N. J. Abram (2019). Investigating observed northwest Australian rainfall trends in Coupled Model Intercomparison Project phase 5 detection and attribution experiments. International Journal of Climatology 39(1): 112-127.

Di Luca, A., R. Elía and R. Laprise (2013). Potential for small scale added value of RCM's downscaled climate change signal. Climate Dynamics 40: 601-618.

Di Virgilio, G., F. Ji, E. Tam, et al. (2022). Selecting CMIP6 GCMs for CORDEX Dynamical Downscaling: Model Performance, Independence, and Climate Change Signals. Earth's Future 10(4): e2021EF002625.

Di Virgilio, G., J. P. Evans, A. Di Luca, M. R. Grose, V. Round and M. Thatcher (2020). Realised added value in dynamical downscaling of Australian climate change. Climate Dynamics 54(11): 4675-4692.

Di Virgilio, G., J. P. Evans, F. Ji, E. et al. (2025). Design, evaluation, and future projections of the NARCliM2.0 CORDEX-CMIP6 Australasia regional climate ensemble. Geosci. Model Dev. 18(3): 671-702.

Dowdy, A. J., A. Pepler, A. Di Luca, L. Cavicchia, G. Mills, J. P. Evans, S. Louis, K. L. McInnes and K. Walsh (2019). Review of Australian east coast low pressure systems and associated extremes. Climate Dynamics 53(7): 4887-4910.

Dunne, J. P., H. T. Hewitt, J. Arblaster, F. et al. (2024). An evolving Coupled Model Intercomparison Project phase 7 (CMIP7) and Fast Track in support of future climate assessment. EGUsphere 2024: 1-51.

Ekström, M., M. R. Grose and P. H. Whetton (2015). An appraisal of downscaling methods used in climate change research. Wiley Interdisciplinary Reviews: Climate Change 6(3): 301-319.

Evans, A., D. A. Jones, R. J. Smalley and S. Lellyett (2020). An enhanced gridded rainfall analysis scheme for Australia. Bureau Research Report. BRR041.

Eyring, V., Bony, S., et al. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer and K. E. Taylor (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9(5): 1937-1958.

Fiddes, S., A. Pepler, K. Saunders and P. Hope (2021). Redefining southern Australia's climatic regions and seasons. Journal of Southern Hemisphere Earth Systems Science 71(1): 92-109.
Grainger, S., R. Fawcett, B. Trewin, D. Jones, K. Braganza, B. Jovanovic, D. Martin, R. Smalley and V. Webb (2021). Estimating the uncertainty of Australian area-average temperature anomalies. International Journal of Climatology 42(5): 2815-2834.
Grose, M. R., S. Narsey, F. P. Delage, et al. (2020). Insights From CMIP6 for Australia's Future Climate. Earth's Future 8(5): e2019EF001469.
Grose, M. R., S. Narsey, R. Trancoso, et al. (2023a). A CMIP6-based multi-model downscaling ensemble to underpin climate change services in Australia. Climate Services 30: 100368.
Grose, M. R., G. Boschat, B. Trewin, V. Round, L. Ashcroft, A. D. King, S. Narsey and E. Hawkins (2023b). Australian climate warming: observed change from 1850 and global temperature targets. Journal of Southern Hemisphere Earth Systems Science 73(1): 30-43.
Gutiérrez, J.M., R.G. Jones, et al. (2021) Atlas. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1927–2058, doi:10.1017/9781009157896.021
Gutowski Jr., W. J., Giorgi, F., et al. (2016) WCRP Coordinated Regional Downscaling Experiment (CORDEX): a diagnostic MIP for CMIP6, Geosci. Model Dev., 9, 4087–4095, https://doi.org/10.5194/gmd-9- 4087-2016
Hausfather, Z. and G. P. Peters (2020). Emissions – the 'business as usual' story is misleading. Nature 577: 618-620.
Hausfather, Z., K. D. Marvel, G. A. Schmidt, J. W. Nielson-Gammon and M. D. Zelinka (2022). Climate simulations: recognize the 'hot model' problem. Nature 605: 26-29.
Hawkins, E. and Sutton, R. (2009). 'The Potential to Narrow Uncertainty in Regional Climate Predictions', Bulletin of the American Meteorological Society, 90(8), pp. 1095–1108. Available at: https://doi.org/10.1175/2009BAMS2607.1.
Hersbach, H., B. Bell, et al. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146(730): 1999-2049.
Howard, E., Su, C.-H., Stassen, C., Naha, R., Ye, H., Pepler, A., Bell, S. S., Dowdy, A. J., Tucker, S. O., and Franklin, C.: (2024) Performance and process-based evaluation of the BARPA-R Australasian regional climate model version 1, Geosci. Model Dev., 17, 731–757, https://doi.org/10.5194/gmd-17-731-2024
Huffman, G., D. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, (2014) Integrated Multi-satellitE Retrievals for GPM (IMERG). NASA's Precipitation Processing Center.
IPCC 1990-2023 IPCC Assessment Reports 1-6. Available at: https://www.ipcc.ch/reports/
IPCC 2021-2023 IPCC Assessment Report 6. Available at: https://www.ipcc.ch/reports/
IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, In press, doi:10.1017/9781009157896.
IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, Doi: 10.59327/IPCC/AR6-9789291691647
Irving D., Takbash A., Peter J. (Submitted). An intercomparison of climate model bias correction methods across Australia.
Irving, D., J. Risbey, D. Squire, R. Matear, C. Tozer, D. Monselesan, N. Ramesh, P. Reddy, and M. Freund (2024). A multi-model likelihood analysis of unprecedented extreme rainfall along the east coast of Australia. Meteorological Applications, 31(3), 1-14. https://doi.org/10.1002/met.2217
Irving, D., Macadam, I., (2024). Application-ready climate projections from CMIP6 using the Quantile Delta Change method. CSIRO: CSIRO; 2024. csiro: EP2024-6112. https://doi.org/10.25919/03by-9y62
Isphording, R. N., L. V. Alexander, M. Bador, D. Green, J. P. Evans, and S. Wales (2024). A Standardized Benchmarking Framework to Assess Downscaled Precipitation Simulations. J. Climate, 37, 1089–1110, https://doi.org/10.1175/JCLI-D-23-0317.1
Iturbide, M., J. M. Gutiérrez, L. M. Alves, et al. (2020). An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets. Earth Syst. Sci. Data 12(4): 2959-2970.
Ji, F., G. Di Virgilio, N. Nishant, E. Tam, J. P. Evans, J. Kala, J. Andrys, C. Thomas and M. L. Riley (2024). Evaluation of precipitation extremes in ERA5 reanalysis driven regional climate simulations over the CORDEX-Australasia domain. Weather and Climate Extremes 44: 100676.
Jiang, X. et al. (submitted). Towards benchmarking the dynamically downscaled CMIP6 CORDEX-Australasia ensemble over Australia. Journal of Southern Hemisphere Earth Systems Science
Jones, D. A., W. Wang and R. Fawcett (2009). High-quality spatial climate datasets for Australia. Australian Meteorological and Oceanographic Journal 58(4): 233-248.
King, A. D., L. V. Alexander and M. G. Donat (2013). The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia. International Journal of Climatology 33(10): 2376-2387.
Kochkov, D., J. Yuval, I. Langmore, et al. (2024). Neural general circulation models for weather and climate. Nature 632(8027): 1060-1066.
Köppen, W. 1936. Das geographische System der Klimate, 1–44 (Gebrüder Borntraeger: Berlin, Germany).
Loridan T, Coates L, Argueso D, Perkins-Kirkpatrick S, McAneney J. (2016). The Excess Heat Factor as a metric for heat-related fatalities: defining heatwave risk categories. Australian Journal of Emergency Management. 31(4) pg. 31:37.
Mach, K.J., M.D. Mastrandrea, P.T. Freeman, and C.B. Field, (2017). Unleashing expert judgment in assessment. Global Environmental Change, 44, 1–14, doi: 10.1016/j.gloenvcha.2017.02.005.\
Mastrandrea, M. D., C. B. Field, T. F. Stocker, O. Edenhofer, K. L. Ebi, D. J. Frame, H. Held, E. Kriegler, K. J. Mach, P. R. Matschoss, G.-K. Plattner, G. W. Yohe and F. W. Zwiers (2010). Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties, Intergovernmental Panel on Climate Change.

McArthur, A. G. (1967). Fire behaviour in eucalypt forest. Commonwealth of Australia Timber Bureau Leaflet 107: 25 pp.
McGregor, J. L. (2005). C-CAM: Geometric aspects and dynamical formulation. CSIRO Atmospheric Research Technical Paper: 43.
Mckee, T.B., Doesken, N.J., Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th Conference on Applied Climatology. American Meteorological Society, Boston, MA, pp. 179–183.
Meinshausen, M., Z. Nicholls, J. Lewis, M. J. et al. (2019). The SSP greenhouse gas concentrations and their extensions to 2500. Geosci. Model Dev. Discuss. 2019: 1-77.
Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., et al. (2021). An updated assessment of near-surface temperature change from 1850: the HadCRUT5 data set. Journal of Geophysical Research: Atmospheres, 126, e2019JD032361. https://doi.org/10.1029/2019JD032361
Narsey S, Grose M, et al. (Submitted). Regional projection uncertainty for Australia in CMIP6 generation RCM and GCM ensembles.
National Climate Science Advisory Committee, NCSAC (2019). Climate Science for Australia’s Future. Dept. of the Environment and Energy, Commonwealth of Australia. https://www.industry.gov.au/sites/default/files/2019-12/climate-science-for-australias-future-2019.pdf
Pendergrass, A. G., R. Knutti, F. Lehner, C. Deser and B. M. Sanderson (2017). Precipitation variability increases in a warmer climate. Scientific Reports 7(1): 17966.
Pepler, A. and A. Dowdy (2020). A Three-Dimensional Perspective on Extratropical Cyclone Impacts. Journal of Climate 33(13): 5635-5649.
Pepler, A. S. and A. J. Dowdy (2022). Australia’s Future Extratropical Cyclones. Journal of Climate 35(23): 7795-7810.
Pepler, A. S., A. J. Dowdy and P. Hope (2021). The differing role of weather systems in southern Australian rainfall between 1979–1996 and 1997–2015. Climate Dynamics 56(7): 2289-2302.
Perkins, S. E., A. J. Pitman, N. J. Holbrook and J. McAneney (2007). Evaluation of the AR4 Climate Models’ Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. Journal of Climate 20(17): 4356-4376.
Pirtle, Z., R. Meyer and A. Hamilton (2010). What does it mean when climate models agree? A case for assessing independence among general circulation models. Environmental Science & Policy 13(5): 351-361.
Rampal, N., S. Hobeichi, P. B. Gibson, J. Baño-Medina, G. Abramowitz, T. Beucler, J. González-Abad, W. Chapman, P. Harder and J. M. Gutiérrez (2024). Enhancing Regional Climate Downscaling through Advances in Machine Learning. Artificial Intelligence for the Earth Systems 3(2): 230066.
Rauniyar, S. P. and S. B. Power (2020). The Impact of Anthropogenic Forcing and Natural Processes on Past, Present, and Future Rainfall over Victoria, Australia. Journal of Climate 33(18): 8087-8106.
Rauniyar, S. P., P. Hope, S. B. Power, M. Grose and D. Jones (2024). The role of internal variability and external forcing on southwestern Australian rainfall: prospects for very wet or dry years. Scientific Reports 13(1): 21578.
Ribes, A., J. Boé, S. Qasmi, B. Dubuisson, H. Douville and L. Terray (2022). An updated assessment of past and future warming over France based on a regional observational constraint. Earth Syst. Dynam. 13(4): 1397-1415.
Ruthrof, K. X., Fontaine, J. B., Matusick, G., Breshears, D. D., Law, D. J., Powell, S., & Hardy, G. (2016). How drought-induced forest die-off alters microclimate and increases fuel loadings and fire potentials. International Journal of Wildland Fire, 25(8), 819–830. https://doi.org/10.1071/WF15028
Schleussner, C.F., Ganti, G., Lejeune, Q., Zhu, B., Pfleiderer, P., Prütz, R., Ciais, P., Frölicher, T.L., Fuss, S., Gasser, T. and Gidden, M.J., (2024). Overconfidence in climate overshoot. Nature, 634(8033), pp.366-373.
Schroeter, B. J. E., B. Ng, A. Takbash, T. Rafter, and M. Thatcher (2024). A Comprehensive Evaluation of Mean and Extreme Climate for the Conformal Cubic Atmospheric Model (CCAM). J. Appl. Meteor. Climatol., 63, 997–1018, https://doi.org/10.1175/JAMC-D-24-0004.1
Seager, R., Cane, M., Henderson, N., Lee, D.E., Abernathey, R. and Zhang, H., (2019). Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. Nature Climate Change, 9(7), pp.517-522.
Sen, P.K. (1968), Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63 (324): 1379–1389, doi:10.2307/2285891
Seneviratne, S.I., X. Zhang, M. Adnan, et al., (2021). Weather and Climate Extreme Events in a Changing Climate. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1513–1766, doi:10.1017/9781009157896.013
Shaw, T. A. et al. (2024a). Emerging Climate Change Signals in Atmospheric Circulation, AGU Advances, 5(6), p. e2024AV001297. Available at: https://doi.org/10.1029/2024AV001297
Shaw, Tiffany A. et al. (2024b). Regional climate change: consensus, discrepancies, and ways forward, Frontiers in Climate, 6. Available at: https://doi.org/10.3389/fclim.2024.1391634
Shepherd, T. G., E. Boyd, R. A. Calel et al. (2018). Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. Climatic Change 151(3): 555-571.
Sherwood, S. C., M. J. Webb, et al. (2020). An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence. Reviews of Geophysics 58(4): e2019RG000678.
Simmonds, I., R. J. Murray and R. M. Leighton (1999). A refinement of cyclone tracking methods with data from FROST. Aust. Meteorol. Mag. 48: 35-49.
Smith, D.M., Gillett, N.P., Simpson, I.R., Athanasiadis, P.J., Baehr, J., Bethke, I., Bilge, T.A., Bonnet, R., Boucher, O., Findell, K.L. and Gastineau, G., (2022). Attribution of multi-annual to decadal changes in the climate system: The Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP). Frontiers in Climate, 4, p.955414. doi.org/10.3389/fclim.2022.955414
Sobolowski S, S. Somot, J. Fernandez et al. (2023). EURO-CORDEX CMIP6 GCM Selection & Ensemble Design: Best Practices and Recommendations. CORDEX White Paper, DOI 10.5281/zenodo.7673399

Srikanthan, S., Azarnivand, A., Bende-Michl, U., et al. (2022). National Hydrological Projections - Design and Methodology. In Bureau Research Report No. 061 (Issue April). http://www.bom.gov.au/research/publications/researchreports/BRR-061.pdf

Stassen, C., Su, C.-H., Dowdy, A., Franklin, C., Howard, E., Steinle, P. (2023). Development and assessment of regional atmospheric nudging in ACCESS, Bureau Research Report 086, accessed online: http://www.bom.gov.au/research/publications/researchreports/BRR-086.pdf

State of the Environment (2021). Australian State of the Environment Report 2021, Department of Climate Change, Energy, the Environment and Water (DCCEEW), https://soe.dcceew.gov.au/

Stevens, B., Fiedler, S., Kinne, S., Peters, K., Rast, S., Müsse, J., Smith, S. J., and Mauritsen, T. (2017). MACv2-SP: a parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6, Geosci. Model Dev., 10, 433–452, doi:10.5194/gmd-10-433-2017, 2017.

Su, C.-H., S. Rennie, J. Torrance, E. Howard, C. Stassen, M. Lipson, R. Warren, A. Pepler, I. Dharssi and C. Franklin (2024). BARRA-C2: Development of kilometre-scale downscaled atmospheric reanalysis over Australia. Bureau Research Report. BRR097.

Su, C.-H., Stassen, C., Howard, E., Ye, H., Bell, S. S., Pepler, A., Dowdy, A. J., Tucker, S. O., Franklin, C. (2022), BARPA: New development of ACCESS-based regional climate modelling for Australian Climate Service, Bureau Research Report 069, accessed online http://www.bom.gov.au/research/publications/researchreports/BRR-069.pdf

Swaminathan, R., J. Schewe, et al. (2024). Regional Impacts Poorly Constrained by Climate Sensitivity. Earth's Future 12(12): e2024EF004901.

Thatcher, M., J. McGregor, M. Dix, and J. Katzfey (2015). A new approach for coupled regional climate modelling using more than 10,000 cores. IFIP Advances in Information and Communication Technology 448:599–607.

Thatcher, M., J. McGregor. (2009). Using a scale-selective filter for dynamical downscaling with the Conformal Cubic Atmospheric Model. MWR. 137, 1742-1752.

Theil, H. (1950), A rank-invariant method of linear and polynomial regression analysis. I, II, III. Proceedings of the Nederlandse Akademie Wetenschappen, 53: 386–392

Thompson, V., Dunstone, N.J., Scaife, A.A. et al. (2017). High risk of unprecedented UK rainfall in the current climate. Nat Commun 8, 107. https://doi.org/10.1038/s41467-017-00275-3

Tucker, S.O., Kendon, E.J., Bellouin, N. et al. (2022). Evaluation of a new 12 km regional perturbed parameter ensemble over Europe. Clim Dyn 58, 879–903. https://doi.org/10.1007/s00382-021- 05941-3

USGCRP (2017). Climate Science Special Report: Fourth National Climate Assessment, Volume I [Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, D.J. Dokken, B.C. Stewart, and T.K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, 470 pp, doi: 10.7930/J0J964J6

USGCRP (2023). Fifth National Climate Assessment. Crimmins, A.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, B.C. Stewart, and T.K. Maycock, Eds. U.S. Global Change Research Program, Washington, DC, USA. https://doi.org/10.7930/NCA5.2023

Van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., De Jeu, R. A. M., Liu, Y. Y., Podger, G. M., Timbal, B., & Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. Water Resources Research, 49(2), 1040–1057. https://doi.org/10.1002/wrcr.20123

van Vuuren, D., B. O'Neill, C. Tebaldi, L. et al. (2025). The Scenario Model Intercomparison Project for CMIP7 (ScenarioMIP-CMIP7). EGUsphere 2025: 1-38.

Vogel, E., F. Johnson, L. Marshall, et al. (2023). An evaluation framework for downscaling and bias correction in climate change impact studies. Journal of Hydrology 622: 129693.

Wasko, C., S. Westra, R. Nathan, A. Pepler, T. H. Raupach, A. Dowdy, F. Johnson, M. Ho, K. L. McInnes, D. Jakob, J. Evans, G. Villarini and H. J. Fowler (2024). A systematic review of climate change science relevant to Australian design flood estimation. Hydrol. Earth Syst. Sci. 28(5): 1251-1285.

Whetton, P. H., M. R. Grose and K. J. Hennessy (2016). A short history of the future: Australian climate projections from 1987 to 2015. Climate Services.

Whetton, P., K. Hennessy, J. Clarke, K. McInnes and D. Kent (2012). Use of Representative Climate Futures in impact and adaptation assessment. Climatic Change 115(3-4): 433-442.

Wilby, R. L. and S. Dessai (2010). Robust adaptation to climate change. Weather 65(7): 180-185.

Wilks, D. S. (1996). Statistical significance of long-range 'optimal climate normal' temperature and precipitation forecasts. J. Climate, 9, 827–839

Wilson, L., U. Bende-Michl, W. Sharples, et al. (2022). A national hydrological projections service for Australia. Climate Services 28: 100331.

World Meteorological Organization (2012). Standardized Precipitation Index User Guide (M. Svoboda, M. Hayes and D. Wood). (WMO-No. 1090), Geneva.

Yoshikane, T. and K. Yoshimura (2023). A downscaling and bias correction method for climate model ensemble simulations of local-scale hourly precipitation. Scientific Reports 13(1): 9412.

Zargar, A., R. Sadiq, B. Naser and F. I. Khan (2011). A review of drought indices. Environmental Reviews 19(NA): 333-349.

11. Appendix

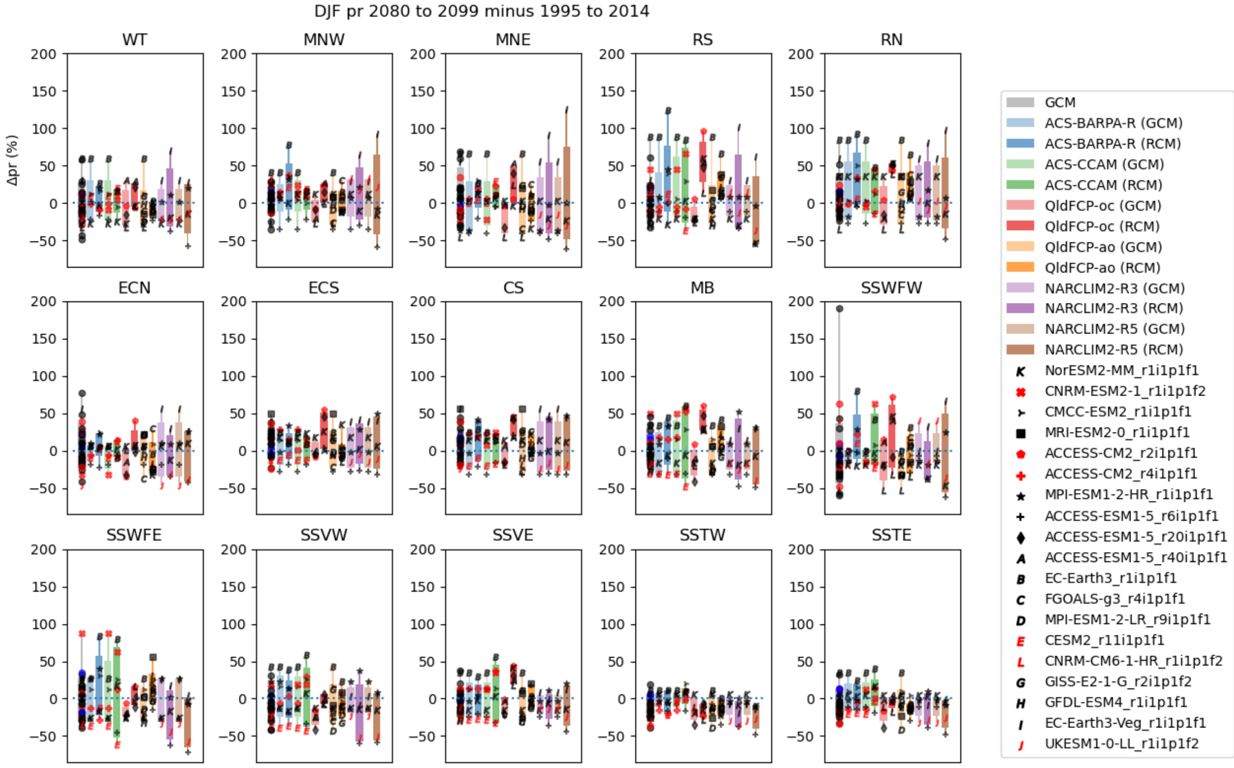

Figure 7.6 but with extended y axis

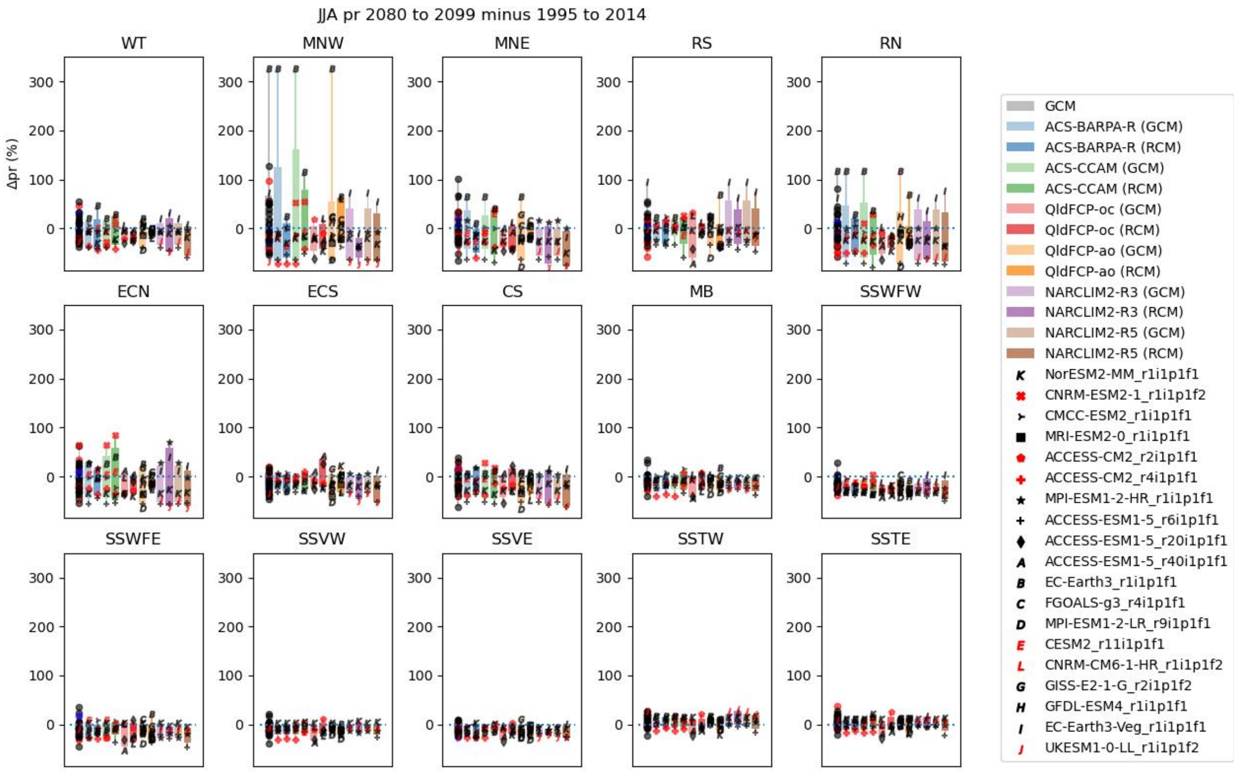

Figure 7.7 but with extended y axis

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

Contact us

1300 363 400
+61 3 9545 2176
csiro.au/contact
csiro.au

For further information

CSIRO Environment
Michael Grose
+61 3 62325345
Michael.Grose@csiro.au
csiro.au/environment

Richard Matear
+61 3 6232 5243
Richard.Matear@csiro.au
csiro.au/environment