Australian Rainfall & Runoff

Revision Projects

PROJECT 1

Development of Intensity-Frequency-Duration Information across Australia

Climate Change Research Plan Project

STAGE 3 REPORT

P1/S3/001

DECEMBER 2015

Australian Government

ENGINEERS
AUSTRALIA
Water Engineering

AUSTRALIAN RAINFALL AND RUNOFF
REVISION PROJECT 1: DEVELOPMENT OF INTENSITY-FREQUENCY-DURATION
INFORMATION ACROSS AUSTRALIA

CLIMATE CHANGE RESEARCH PLAN PROJECT

STAGE 3 REPORT

DECEMBER, 2015

ACKNOWLEDGEMENTS

Contractor Details

Bureau of Meteorology
700 Collins Street
Docklands, VIC, 3001

Tel: (03) 9669 4000
Fax: (03) 9669 4699
Web: www.bom.gov.au

CSIRO
Limestone Avenue
Campbell, ACT, 2612

Tel: (02) 6276 6000
Web: www.csiro.au

The University of Adelaide Water Research Centre
The University of Adelaide
Adelaide, SA, 5005

Tel: (08) 8313 3747
Fax: (08) 8313 6222
Web:
www.adelaide.edu.au/environment/wrc

UNSW Water Research Centre
The University of New South Wales
Sydney, NSW, 2052

Tel: (02) 9385 5017
Fax: (02) 9313 8624
Web: http://water.unsw.edu.au

FOREWORD

ARR Revision Process

Since its first publication in 1958, Australian Rainfall and Runoff (ARR) has remained one of the most influential and widely used guidelines published by Engineers Australia (EA). The current edition, published in 1987, retained the same level of national and international acclaim as its predecessors.

With nationwide applicability, balancing the varied climates of Australia, the information and the approaches presented in Australian Rainfall and Runoff are essential for policy decisions and projects involving:
- infrastructure such as roads, rail, airports, bridges, dams, stormwater and sewer systems;
- town planning;
- mining;
- developing flood management plans for urban and rural communities;
- flood warnings and flood emergency management;
- operation of regulated river systems; and
- prediction of extreme flood levels.

However, many of the practices recommended in the 1987 edition of ARR now are becoming outdated, and no longer represent the accepted views of professionals, both in terms of technique and approach to water management. This fact, coupled with greater understanding of climate and climatic influences makes the securing of current and complete rainfall and streamflow data and expansion of focus from flood events to the full spectrum of flows and rainfall events, crucial to maintaining an adequate knowledge of the processes that govern Australian rainfall and streamflow in the broadest sense, allowing better management, policy and planning decisions to be made. One of the major responsibilities of the National Committee on Water Engineering of Engineers Australia is the periodic revision of ARR. A recent and significant development has been that the revision of ARR has been identified as a priority in the Council of Australian Governments endorsed National Adaptation Framework for Climate Change.

The update will be completed in three stages. Twenty one revision projects have been identified and will be undertaken with the aim of filling knowledge gaps. Of these 21 projects, ten projects commenced in Stage 1 and an additional 9 projects commenced in Stage 2. The remaining two projects will commence in Stage 3. The outcomes of the projects will assist the ARR Editorial Team with the compiling and writing of chapters in the revised ARR.

Steering and Technical Committees have been established to assist the ARR Editorial Team in guiding the projects to achieve desired outcomes. Funding for Stages 1 and 2 of the ARR revision projects has been provided by the Federal Department of Climate Change and Energy Efficiency. Funding for Stages 2 and 3 of Project 1 (Development of Intensity-Frequency-Duration information across Australia) has been provided by the Bureau of Meteorology.

Mark Babister
Chair Technical Committee for
ARR Research Projects

Assoc Prof James Ball
ARR Editor

ARR REVISION PROJECTS

The 21 ARR revision projects are listed below:

ARR Project No.	Project Title	Starting Stage
1	**Development of intensity-frequency-duration information across Australia**	**1**
2	Spatial patterns of rainfall	2
3	Temporal pattern of rainfall	2
4	Continuous rainfall sequences at a point	1
5	Regional flood methods	1
6	Loss models for catchment simulation	2
7	Baseflow for catchment simulation	1
8	Use of continuous simulation for design flow determination	2
9	Urban drainage system hydraulics	1
10	Appropriate safety criteria for people	1
11	Blockage of hydraulic structures	1
12	Selection of an approach	2
13	Rational Method developments	1
14	Large to extreme floods in urban areas	3
15	Two-dimensional (2D) modelling in urban areas.	1
16	Storm patterns for use in design events	2
17	Channel loss models	2
18	Interaction of coastal processes and severe weather events	1
19	Selection of climate change boundary conditions	3
20	Risk assessment and design life	2
21	IT Delivery and Communication Strategies	2

ARR Technical Committee:

Chair: Mark Babister, WMAwater

Members: Associate Professor James Ball, Editor ARR, UTS

Professor George Kuczera, University of Newcastle

Professor Martin Lambert, University of Adelaide

Dr Rory Nathan, Jacobs

Dr Bill Weeks

Associate Professor Ashish Sharma, UNSW

Dr Bryson Bates, CSIRO

Steve Finlay, Engineers Australia

Related Appointments:

ARR Project Engineer: Monique Retallick, WMAwater

ARR Admin Support: Isabelle Testoni, WMAwater

PROJECT TEAM

Project Team Members

- Bryson Bates, CSIRO Oceans and Atmosphere Flagship (ARR Technical Committee manager)

- Deborah Abbs, CSIRO Oceans and Atmosphere Flagship

- Jason Evans, University of New South Wales

- Janice Green, Water Information Services, Bureau of Meteorology

- Aurel Griesser, Bureau of Meteorology

- Dörte Jakob, Hydrometeorological Advisory Service, Bureau of Meteorology

- Rex Lau, CSIRO Digital Productivity Flagship

- Sally Lavender, CSIRO Oceans and Atmosphere Flagship

- Eric Lehmann, CSIRO Digital Productivity Flagship

- Michael Leonard, School of Civil, Environmental and Mining Engineering, University of Adelaide

- Kim Nguyen, CSIRO Oceans and Atmosphere Flagship

- Aloke Phatak, CSIRO Digital Productivity Flagship

- Tony Rafter, CSIRO Oceans and Atmosphere Flagship

- Alan Seed, Bureau of Meteorology

- Marcus Thatcher, CSIRO Oceans and Atmosphere Flagship

- Seth Westra, Department of Civil, Environmental and Mining Engineering, University of Adelaide

- Feifei Zheng, School of Civil, Environmental and Mining Engineering, University of Adelaide

This report was independently reviewed by:

- Dr Louise Wilson, CSIRO Oceans and Atmosphere Flagship
- Dr Michael Grose, CSIRO Oceans and Atmosphere Flagship

EXECUTIVE SUMMARY

Motivation and Aims of the Study

The 2015 edition of *Australian Rainfall and Runoff* (ARR 2015) will contain new design rainfall Intensity-Frequency-Duration (IFD) curves prepared by the Bureau of Meteorology. These curves are estimated from current climate observations. The Fifth Assessment Report of the Intergovernmental Panel on Climate Change states that "extreme precipitation events over most of the mid-latitude land masses and over wet tropical regions will very likely become more intense and more frequent by the end of this century, as global mean surface temperature increases." This suggests that directly applying current-climate IFD curves may become unsuitable for infrastructure design in the decades that lie ahead. While ARR 2015 will include interim guidance on incorporating climate change into design flood estimation, the guidance is based on a 'broad brush' approach due to the paucity of published regionally-specific results that are based on the best available science.

As a first step towards bridging this gap, a pilot project titled *Rainfall Intensity-Frequency-Duration (IFD) Relationships under Climate Change* was commissioned in June 2013 to provide insight into how the new design rainfall IFD curves might be affected by anthropogenic climate change. Emphasis was placed on leveraging existing knowledge, data, methods and model output whenever possible.

Process

Estimating future changes to extreme rainfall for durations from hourly through to multi-day events represents a major scientific challenge, and numerous innovations were necessary to achieve the project aims. A multiple-lines-of-evidence approach was adopted, combining multiple historical data sources (including rainfall stations operated by the Bureau of Meteorology and water utilities, as well as the Bureau's radar data) with very-high resolution climate model output to obtain future projections. State-of-the-art statistical methods were also applied to assess historical change, to adjust climate model estimates of rainfall IFD curves to match observed data, and to develop uncertainty estimates associated with climate change projections.

To achieve the project aims within the constraints imposed in terms of budget and timelines, it was necessary to limit the number of study regions and the number of climate model runs used. Thus the scope of the study was limited to:
- Two study regions (Greater Sydney and south-east Queensland).
- Trend estimation in annual maxima series for sub-daily rainfall in the Greater Sydney region only.
- Two climate change projections for design rainfall IFD curves for the Greater Sydney region based on two high-end scenarios of greenhouse gas and aerosol emissions and two climate model combinations. The climate models consisted of a single global climate model (GCM) and a single dynamical downscaling model each: the CSIRO Mk3.5 GCM and the Weather Research and Forecasting (WRF) model run at a 2 km spatial resolution; and the Australian Community Climate and Earth System Simulator (ACCESS) coupled model version 1.0 and the CSIRO conformal cubic atmospheric model (CCAM) run at 10

and 2 km resolution.

- A single climate change projection for the south-east Queensland based on one high-end emission scenario and the ACCESS-CCAM model combination.

The realism of simulated rainfall extremes from the two model combinations described above was evaluated using blended radar and rain gauge data (Rainfields), and gridded station data from the Australian Water Availability Project (AWAP). The use of radar data in this context is a novel approach, motivated by their superior spatial and temporal resolution.

A spatial Bayesian Hierarchical Model (BHM) was developed that incorporates climate change information from CCAM and WRF by relating the parameters of the probability distributions of rainfall station data to those for the extreme rainfall generated by the climate models. The resulting information was used to 'bias-correct' climate model estimates of what might happen to extreme rainfall under climate change and produce estimates of uncertainty in the projected rainfall IFD curves. These curves were compared to those obtained through the application of the ARR Interim Guideline on Climate Change [*Engineers Australia, 2014*].

The project team consisted of research staff from the Bureau of Meteorology, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), and the Universities of NSW and Adelaide. Together, the team's expertise covered many disciplines in engineering, and the mathematical, physical and statistical sciences. Given the geographic spread of the research team (Adelaide, Canberra, Melbourne, Perth and Sydney), four workshops were held during the project life cycle to ensure scientific consistency and integration, dissemination of early research findings, and identification and remedy of any emerging impediments to progress. The project was administered by Engineers Australia and completed in June 2015.

Research Findings

1. For durations less than two hours, the intensities of observed annual maximum rainfalls have increased in the Sydney study region over the period from 1966 to 2012, while for durations greater than three hours a decrease was found to be more likely. This is explained by the changing seasonal composition of annual maxima with duration. Increasing summer maxima are more prominent at shorter timescales while decreasing winter and autumn are more prominent at longer timescales.

2. Despite differences in methodology between the BHM used herein and the new design rainfall IFD curves in ARR 2015, the results were found to be remarkably similar. In most cases, the ARR 2015 design rainfall IFD curves lie within the uncertainty bounds produced by the BHM.

3. Biases exist in both observational datasets (Rainfields and AWAP). Therefore, more than one observational dataset is recommended for use in similar evaluation studies, analogous to the use of ensembles in climate model simulation.

4. To produce credible projections of changes in rainfall extremes, GCMs and downscaling models need to be able to realistically simulate key mechanisms and characteristics of these extremes. Simulations driven by Reanalyses (global gridded data sets that represent past weather conditions) are typically a better match with observations than GCM-driven

simulations, and WRF simulations typically provide a more realistic picture than the CCAM simulations analysed. Based on the simulations considered, there is little indication that lowering the resolution from 10 km to 2 km leads to an improvement in the simulation of rainfall extremes.

5. The frequency distributions of observed and simulated rainfall exhibit marked differences. Climate models produce unrealistically high rainfall extremes, and while a simple bias correction reduced errors for modelled annual accumulations, it did not remove the tendency for overestimation of rainfall extremes and introduced artefacts. Simulated rainfall is less random in space and time than observed rainfall, which implies that not all rainfall-producing mechanisms have been correctly captured. This is important because changes in rainfall extremes are likely to be driven by changes in convective rainfall and low-pressure systems.

6. For the Greater Sydney region, neither CCAM nor WRF output is able to adequately reproduce the magnitude of extreme rainfall in the selected baseline period (1990 – 2009) although there is some relationship between patterns of extreme rainfall produced by WRF and those exhibited by rainfall station data. The BHM estimate for future extremes, which only uses the part of the WRF projection that is more reliable, suggests that the 1% annual exceedance probability (AEP) for the 24 hour duration event will increase by up to 20% by 2050. To the west, in some parts of the Blue Mountains and beyond, decreases are possible. Also, projected rainfall depths obtained from the ARR Interim Guideline on Climate Change under RCP8.5 were found to be in agreement with the BHM estimates reported herein.

7. For south-east Queensland, projected changes in the 1% AEP 24-hour rainfall can also be very large. The ACCESS-CCAM model combination projects increases in the Brisbane region and in an arc to the west, whereas it projects a decrease in extreme rainfall near the Sunshine Coast and to the northwest. For an AEP of 10%, there is little difference between current and future extremes, whereas for rarer events (1% AEP), future extreme rainfall is much larger than current extreme rainfall. Spatial patterns at different durations are similar.

Recommendations for Further Action

Building on the above knowledge gaps and research findings, the report suggests the following areas for future research:

- **Improvement of physical understanding.** The analysis of historical rainfall trends showed that short (less than 2 hours) duration rainfall extremes have experienced increases in intensity while extreme rainfall with longer durations tended to decline. Further research is required to comprehend the physical processes that drive the variations of the extreme rainfall with different durations.

- **Design rainfall temporal patterns for short and long duration storms.** Climate change may not only alter rainfall intensity, but also vary the temporal patterns of the extreme rainfall events. Further research is required to quantify the impact of climate change on these patterns.

- **Structure of spatial dependence.** It is necessary to investigate the possible changes in the structure of the spatial dependence of rainfall extremes, thereby obtaining deeper insight into the impact of climate change.

- **Implications for floods.** Possible changes to the seasonality of rainfall may change the catchment wetness prior to the flood-producing rainfall event, leading to an altered flood risk. The interaction between the flood-producing rainfall event and the antecedent catchment conditions is not well understood and should be investigated.

- **Augmentation of the Greater Sydney and south-east Queensland studies.** The climate change projections described in this report are based on a limited number of high-end emissions scenarios, global climate models and regional climate models. Therefore, it is highly likely that the uncertainty in the projections is underestimated and projected regional patterns of change are not robust. Accepted practice would involve the use of a much larger number of host global climate models and a more systematic sampling of regional model uncertainty. The use of a low emissions scenario such as RCP4.5 should also be incorporated into the experimental design.

- **Extension to other geographic domains.** The results of this study focus on the Greater Sydney and to a lesser extent the south-east Queensland regions. It is necessary to repeat this study over other geographic regions where different rainfall-generating mechanisms are likely to apply.

- **Climate model improvement.** Better understanding and modelling of the drivers of extreme rainfall from global to regional scales and their responses to greenhouse gases are required. Given the need to simulate changes in rainfall extremes at short time scales (down to sub-hourly), high resolution models (horizontal grid spacing less than 4 km) that can represent the evolution, structure and secondary development of convective storms are the tool that embodies our best knowledge of the important physical processes. Current models require further development and testing to improve their performance at these scales and hence their ability to assist in addressing the previous research needs.

- **Bayesian Hierarchical Model improvement.** The current implementation of the spatial Bayesian Hierarchical Model incorporates present and projected dynamically downscaled data in two separate steps. A more mathematically rigorous approach would be to construct a model that completely integrates data as well as the present and future climate model outputs within a unified framework.

Australian Research Capacity

The success of the project demonstrates the substantial potential embodied in the study of the behaviour of rainfall extremes under current and projected climate conditions. However, the current research infrastructure, institutional set up and funding environment in Australia create significant barriers to maximising the provision of clear, practical and science-based information and advice to support management, policy and planning decision making. With the completion of the pilot project, there no longer exists an ongoing critical mass of research activity to ensure the viability of this new field of endeavour in Australia.

TABLE OF CONTENTS

1. Introduction

1.1. Literature review

The frequency and intensity of rainfall extremes may increase as a consequence of climate change [*IPCC*, 2012]. Recently, increases in the intensity of extreme daily rainfall have been widely recognised at the global scale [*Alexander et al.*, 2006; *Min et al.*, 2011; *Westra et al.*, 2013] but changes in sub-daily extreme rainfall patterns under a future climate remain poorly understood [*Westra et al.*, 2014]. Recent reviews of the literature on future changes to extreme rainfall are provided by *Trenberth* [2011], *Westra et al.* [2014] and *O'Gorman* [2015].

Potential changes in rainfall over coming decades can be assessed on the basis of projections from Global Climate Models (GCMs). These models provide rainfall information on fairly coarse spatial scales (about 100 km) but for decision-making purposes, information on regional and local scales is required. There are, however, difficulties in understanding potential changes to sub-daily rainfall due to the combination of the highly localised nature of sub-daily rainfall mechanisms that are not well represented in large scale GCMs and the limited availability of the sub-daily rainfall data [*Boucher et al.*, 2013; *Collins et al.*, 2013].

'Downscaling' is a means for generating regionally or locally relevant data from GCMs, and involves the use of statistical and/or physics-based techniques. This approach can provide invaluable information to decision-makers about likely changes to rainfall extremes. However, model simulations for the current climate – especially those driven by GCM output – exhibit biases and require careful processing in combination with detailed evaluation based on observations. This then informs any inference concerning how shortcomings in the simulations of current conditions may affect the validity of projected changes in rainfall extremes. Projected changes in downscaled rainfall extremes exhibit a large degree of variability spatially (i.e. projected increases and decreases manifest in close proximity) and between studies [*Alexander and Arblaster*, 2009].

As a result of the changing nature of extreme rainfall, Intensity-Frequency-Duration (IFD) curves estimated from current climate observations are likely to be unsuitable for infrastructure design in a changing climate [*Cheng and AghaKouchak*, 2014], and regional climate models (RCMs) are used increasingly to develop projections of future IFDs. For example, RCMs were used to estimate changes in rainfall IFDs as part of the North American Regional Climate Change Assessment Program [*Mirhosseini et al.*, 2013, 2014; *Wang et al.*, 2013; *Zhu et al.*, 2013], with model simulations run at relatively coarse spatial resolution (50km). Studies that used station-based observations applied a bias correction to account for both model errors and the spatial scale mismatch between model grid cells and station locations [*Mirhosseini et al.*, 2013; *Wang et al.*, 2013]. This correction was applied to the rainfall depths within the IFDs so that the correction applied directly to the most extreme rainfall intensities. Biases of 50% or more were often found [*Liew et al.*, 2014; *Mirhosseini et al.*, 2013; *Wang et al.*, 2013; *Zhu et al.*, 2013] in the model results for the recent past. Several studies also applied a regional frequency analysis to increase the sample size and hence the robustness of the results [*Anthes et al.*, 1987; *Zhu et al.*, 2013].

Increasingly, numerous studies have applied RCMs at very high resolution to better resolve convective rainfall processes. The study of [*Kuo et al.*, 2015] applied an RCM at convection-permitting resolution (3km), reducing the scale mismatch and the dependence on a convection parametrization. This resolution had previously been found to be required to capture small scale convective precipitation events [see, e.g., *Westra et al.*, 2014]. While attempts to explicitly quantify uncertainty led to wide ranges [*Kuo et al.*, 2015], studies generally found increases in rainfall depths in the future, although this varied across annual exceedance probabilities (AEPs), durations and locations, emphasising the difficulty in generalising these results more broadly [*Kendon et al.*, 2012, 2014].

A number of studies have explored and quantified possible changes in sub-daily rainfall extremes as a response to a warming climate. Research into trends in extreme sub-daily rainfall has focused on individual sites or small geographic regions [*Jakob et al.*, 2011; *Lenderink et al.*, 2011; *Westra and Sisson,* 2011]. These studies found that the strength of sub-daily rainfall trends depends strongly on the geographic area being analysed, so that the results from one area are difficult to translate directly to other areas. Furthermore, the trends varied substantially based on the season and duration of the annual (or seasonal) maximum rainfall event. Finally, given the relatively short duration of historical records of sub-daily rainfall, it was found that spatial techniques (involving analysing trends in a large number of gauges simultaneously) were needed to test for trends in sub-daily rainfall.

In the current hydrological literature, the impact of climate change on IFD curves is usually assessed by comparing the change in return levels between two or more intervals of time, for example, a baseline period and a future period. While climate statistics are likely to change between such intervals, we can assume that the extremes are locally stationary if the intervals are chosen to be short, especially if they are short compared to the length of the interval between the two periods. In this work, we have chosen 20-year baseline and future periods.

Srivastav et al. [2014] categorize approaches to using information from GCMs and RCMs to assess the impact of climate change on IFD curves into one of three classes. They also provide citations in each category. The approaches include:

- Change factor or 'delta' methods that use only information from GCM or RCM output in a baseline period and a future period to estimate the impact of climate change on return levels across a range of durations. These differences (change factors) are then applied to IFD curves calculated from rainfall data from the baseline period. In this class of methods, GCM or RCM output in baseline and future periods may be downscaled to point resolution before IFD curves are calculated.
- Bias-correction methods, in which differences between observed data and GCM or RCM output for a baseline period are used to modify GCM or RCM output for future periods. IFD curves are then calculated from this modified output. Quantile matching is often used to map model output onto observed rainfall data, and this mapping is then applied to the future output.
- Bias-correction-disaggregation methods, which are based on bias-correction methods, but include a disaggregation step when only GCM or RCM output at a coarse temporal resolution is available. For example, if only daily GCM or RCM output is available,

disaggregation is required to construct sub-daily and sub-hourly sequences of rainfall data.

Change factor methods are widely used in many other contexts besides rainfall extremes, for example, in estimating changes in temperature. However, their use depends on the tacit assumption that the difference between current and future output will be an accurate indicator of what is likely to occur in the future, regardless of how well model output reproduces current climate. On the other hand, bias-correction methods account for the difference between observed data and model output, and then correct model output for future periods. Built into these methods is the assumption that whatever correction is required to map model output to observed data is also applicable in a future period.

Applications of the methods described above have been limited to single stations [*Yilmaz et al.*, 2014; *Peck et al.*, 2012]. Furthermore, bias-correction methods operate directly on sequences of daily or sub-daily rainfall. By contrast, the method we describe in this report is a spatial extremes model [*Lehmann et al.*, 2013; *Soltyk et al.*, 2014] in which (a) we use information from neighbouring gauged locations in order to obtain more precise parameter estimates, and (b) incorporate information about climate change and the difference between data and model output at the level of the parameters of the generalized extreme value distribution that is used to model extreme rainfall at different durations.

1.2.   **Objectives**

In conjunction with the revision of the Australian Rainfall and Runoff (ARR), and in particular the revision of design rainfall estimates, Engineers Australia have released an Interim Guideline [*Engineers Australia*, 2014] to "outline an approach to address the risks from climate change in projects and decisions that involve estimation of design flood characteristics". Simultaneously, the *Rainfall Intensity-Frequency-Duration (IFD) Relationships under Climate Change* project was undertaken to improve the current state of knowledge with the aim of updating and improving the interim guideline. The overall objectives of the project were twofold:

- Quantify possible changes and uncertainties in design rainfall IFD curves due to anthropogenic climate change.
- Provide interim advice to practitioners on how these changes can be included in design and planning.

1.3.   **Description of project**

Ideally, climate change projections should be based on multiple emissions scenarios and climate models to ensure that uncertainties about future greenhouse gas emissions and differences between models are captured to a reasonable extent. Such an approach was infeasible for this study due to time, computational and financial constraints. This situation led to the use of a compromise approach that:

- Built on three previous studies: 1) the NARCliM (**N**SW / **A**CT **R**egional **Cli**mate **M**odelling) project led by the University of New South Wales Climate Change Research Centre (see http://www.ccrc.unsw.edu.au/sites/default/files/NARCliM/index.html; *Evans et al.* [2014]); 2) extreme rainfall in the western Sydney region [*Abbs and Rafter*, 2009]; and 3) extreme

rainfall and coastal sea levels in south-east Queensland [*Abbs et al.*, 2007].

- Used two model combinations consisting of a single GCM and a single dynamical downscaling model each: 1) the CSIRO Mk3.5 GCM and the Weather Research and Forecasting (WRF) model [*Evans and McCabe,* 2013] run at a 2 km spatial resolution; and 2) the Australian Community Climate and Earth System Simulator (ACCESS) coupled model version 1.0 and the CSIRO conformal cubic atmospheric model (CCAM) [*Thatcher and McGregor*, 2009]) run at 10 km and 2 km spatial resolution.

- Used two Reanalysis data sets for climate model evaluation: the National Oceanic and Atmospheric Administration's Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP-NCAR) Reanalysis [*Kalnay et al.*, 1996] for the CSIRO Mk3.5 GCM – WRF model combination; and the ERA-Interim reanalysis [*Dee et al.*, 2011] for the ACCESS 1.0 – CCAM combination. (Reanalyses are global gridded data sets that represent past weather conditions at time intervals of every 6 to 12 hours. This is accomplished by blending atmospheric observations with climate model output through the use of a single version of a data assimilation system.)

- For the purpose of model evaluation, downscaled information for the current climate was produced and used for the periods 1990 – 2009 (WRF) and 1980 – 2010 (CCAM).

- Considered two high-end emissions scenarios: the IPCC SRES A2 emissions scenario was used as input to the CSIRO Mk3.5 model runs; and the representative concentration pathway RCP8.5 was used as input to the ACCESS1.0 future climate. The greenhouse gas emissions and concentrations increase markedly over time in RCP8.5, leading to a radiative forcing of 8.5 W m^{-2} at the end of the century. The warming produced using RCP8.5 is generally slightly higher than that for the A2 scenario.

- Climate projections for rainfall IFD curves were developed for a single time slice (2040 – 2059). The length of the time slice was restricted to 20 years as longer periods are not likely to be stationary due to increasing greenhouse gas concentrations.

- Used a spatial Bayesian Hierarchical Model (BHM) that incorporates climate change information from dynamical downscaling models by relating the parameters of the probability distributions of rainfall station data to those for the extreme rainfall generated by the climate models. If a relationship is found, the information can be used to 'bias-correct' climate model estimates of what might happen to extreme rainfall under climate change and produce estimates of uncertainty in the IFD curves.

- Assessed historical changes to extreme rainfall at durations from sub-hourly through to multiple days for the Greater Sydney region, using an extreme value statistical model based on max-stable process theory first adopted in the Australian context by *Westra and Sisson* [2011].

1.4.   Benchmarking

Assessments of climate model performance can be made by comparing downscaled model output to rainfall observations for the current climate. Observational data sets are typically based on

either rainfall station data or gridded information derived from station data. This approach is used herein because of its simplicity and generally wide acceptance.

The drawback in using station-based data sets, however, is their coarse temporal and/or spatial distribution. An additional resource is radar data, which provides consistent spatial and temporal information that cannot be readily gleaned from rainfall gauge information alone. Results presented below indicate that blended radar/gauge data (called Rainfields) do indeed capture the spatial distribution of high extreme precipitation better than grids derived from station data alone and are therefore better suited to evaluation of the realism of downscaled rainfall extremes. The rationale behind evaluating the performance of climate model simulations on the basis of radar data is as follows:

- In order to infer the rainfall distribution at point locations or over hydrological catchments from climate model projections, dynamical or statistical downscaling must be applied to the coarse resolution model output.

- The validity of climate model projections of rainfall extremes depends critically on how well the downscaling scheme represents the actual space-time characteristics of rainfall.

- A novel approach to assess the credibility of dynamically downscaled projections of rainfall extremes exploits the superior spatial and temporal resolution of blended radar/gauge data.

1.5.   **Report structure**

Section 2 of this report describes the study areas considered (Greater Sydney and southeast Queensland regions) and observational data sets used. The suite of methods used is described in section 3. Section 4 compares model output for current climate conditions to station and gridded rainfall observations, and the Rainfields data set. Climate change projections for design rainfall IFD curves for the study areas are presented in section 5. Finally, conclusions are given in section 6 which lists the key findings, knowledge gaps, key 'learnings' from the project and recommendations for future research.

2.   Description of Data

2.1.   Daily and recording rain gauges

The observation study concentrated on the Greater Sydney Region. For the analysis of trends in sub-daily rainfall series, a total of 69 sub-daily rainfall stations (5-minute resolution) were selected in this region, due to their relatively high quality and long record length (Figure 1). Of these, 13 stations were taken from the Bureau of Meteorology (red dots) and 56 sites were collected from water regulators under the 2007 Water Act (black dots). Figure 1 (right) shows the data length of each station, with the open circle representing the availability of the data in each year. The records for the stations from the Bureau of Meteorology are overall longer than those from the water regulators. Despite some missing years, the majority of the sub-daily rainfall stations analysed have more than 27 years of record (Figure 1, right).
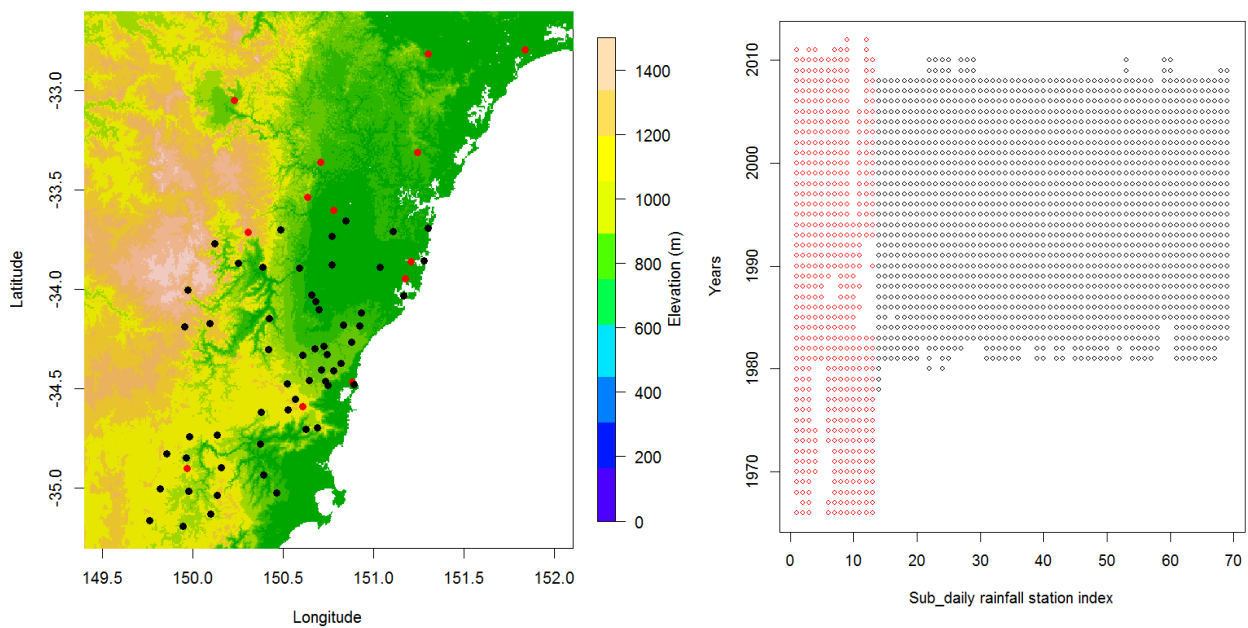
Figure 1. Locations of sub-daily rainfall stations (left) and data length of each station (right). Red and black dots respectively represent stations from the Bureau and water regulators, and the open circle dot indicates the availability of the records in a year.

For the BHM implementation, the rainfall dataset used over the Greater Sydney region is provided by a total of 872 stations (310 sub-daily and 562 daily stations, Figure 2), while the South East Queensland dataset contains a total of 1348 stations (240 sub-daily and 1108 daily stations, Figure 2). Both datasets were subjected to a thorough quality assurance process, including a number of different exploratory analyses. The sub-daily data at 5 min intervals were accumulated over 12 different durations, namely 5, 10, 15 and 30 min, and 1, 2, 3, 6, 12, 24, 48 and 72 hours. All stations within these datasets provide at least 8 years' worth of yearly rainfall maxima during the period from 1961 to 2000.
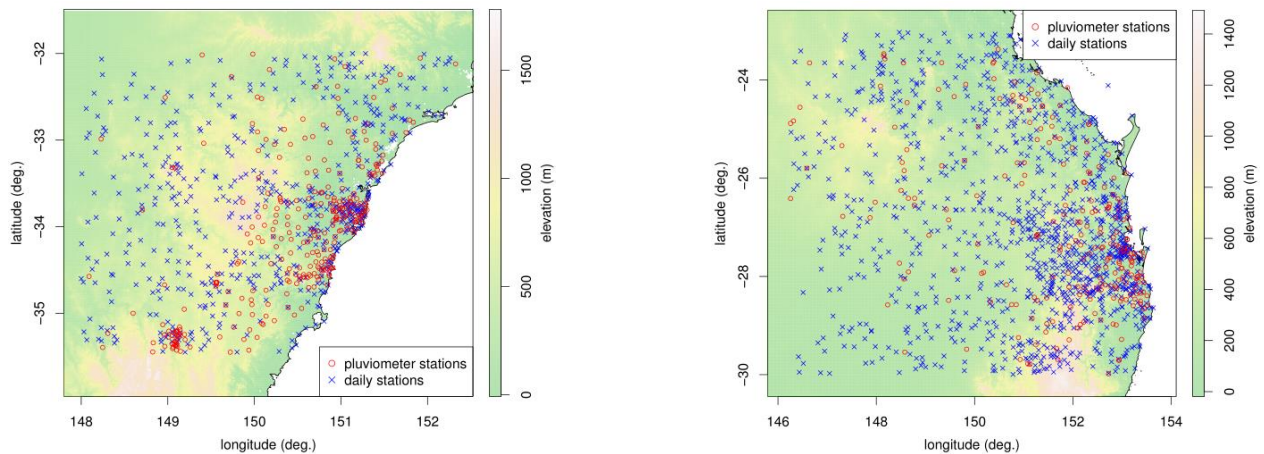
Figure 2. Locations of daily and sub-daily (pluviometer) stations used for the BHM implementation, for the Greater Sydney region (left) and South East Queensland region (right).

2.2. Australian water availability project (AWAP) rainfall

In an attempt to define a yard stick for the evaluation of gridded rainfall data and their limitations relevant to the study of rainfall extremes, analyses of daily rainfall data (24 hour accumulations from 9 am Local Standard Time) from the Australian Water Availability Project (AWAP) were used (*Jones et al.,* 2009). The resolution of the AWAP grids is 0.05° × 0.05°, which is approximately 5 km × 5 km. The AWAP grids were developed for the purpose of monitoring of climate change and climate variability for the Australian region and are produced operationally at the Bureau. The spatial analyses are based on *in-situ* data only. An anomaly-based approach was used to produce the AWAP grids, i.e. data were decomposed into the long-term average and anomaly components. A Barnes technique was used to grid the anomalies [*Koch et al.,* 1983] while smoothing splines were used for the average grids [*Wahba and Wendelberger*. 1980]. Finally, daily rainfall grids were derived by multiplying climatology and anomaly analyses. The AWAP grids are an improvement over existing analyses. However, the accuracy of these gridded data is limited by the density of the station network. Errors tend to increase where rainfall gradients are strong, such as for regions with significant orography. The mean average error is still about 50% of the average daily rainfall.

King et al. [2013] assessed the performance of AWAP grids for daily rainfall extremes across Australia. Comparisons of the frequency and intensity of rainfall events above the 95th percentile were undertaken against 119 high-quality stations. While the AWAP grid typically underestimates intensity and frequency of extreme rainfall events, the intensity and frequency of low rainfall events is typically overestimated. The magnitude of the underestimation of the average 95th percentile for the study region was of the order of 10%. Problems were identified for remote areas but considering the density of the monitoring network for the Greater Sydney region this is not a concern here. The rank correlation between AWAP and station rainfall for this region is high (0.8 to 0.9). AWAP estimates can therefore be expected to be a good indicator for where and when extremes occurred but they will be less reliable in terms of the magnitude of rainfall accumulations.

2.3. Radar data

The observations considered herein are sourced from Bureau weather radars. The Australian radar network consists of a mixed collection of about 66 radars with an average age of 9 years. Multi-radar mosaics are available for three 500 km domains (Melbourne, Sydney and Brisbane). Weather radars are operated for real-time applications rather than for climatological studies. In an operational setting, it is vital to process these radar measurements in a very short time, effectively limiting the complexity of processing algorithms that can be applied in real-time. The evaluations presented in this report are based on spatio-temporal characteristics derived from blended radar and gauge rainfall accumulations and are referred to as 'Rainfields' [*Seed et al.,* 2007] in this report. Since December 2013 operational Rainfields version 2 data have been stored in a dedicated *Radar and Rainfields Archive* at the Bureau. The accuracy of quantitative precipitation estimates is constantly being monitored. For every radar in the Rainfields product suite a monthly report is generated that provides summary statistics for 30 minute and daily rainfall accumulations based on comparison with the last three months' rain-gauge measurements. The statistics include a) the bias over the 3-month period, b) the fraction of data where the error is less than a factor of 1.6 (2 decibels [dB]), and c) the magnitude and date of the largest error for each gauge.

Algorithms have been developed to detect and remove artefacts in radar data [*Peter et al.,* 2014]. Such artefacts include partial beam blocking by local topographic effects, clutter (e.g. dust, insect, reflection on the sea surface) and – in an operational setting – bias correction that requires real-time adjustment. Station data also need quality control, and information which is obviously wrong - in particular false zeros, where the continuous rain gauge incorrectly records zero during rainfall, – should be filtered out. The conversion of radar reflectivity (Z) to rain rate (R) is one of the most critical steps in the estimation of realistic radar rainfall accumulations. There are significant differences in Z-R relationships for convective and stratiform rain, which are accounted for using complex algorithms. Instantaneous measurements of rain rate have to be integrated over time to derive rainfall accumulations.

For this project, a new version of Rainfields (version 3) was produced using a more sophisticated processing algorithm with the intention of providing improved quantitative precipitation estimates through improved quality control. As part of the new processing algorithm, a spatially varying calibration of radar against gauge data was implemented as well as measures to assess the quality of each 30-minute rainfield. These techniques have been applied to radar data for the Greater Sydney region. The resulting product is a mosaic of radar data from five radar sites (Wollongong, Newcastle, Sydney, Canberra and Kurnell), covering a 500 km × 500 km domain, centred on the Wollongong radar (150.87˚E and 34.26˚S) at a resolution of 0.01˚ (approximately 1 km) and 30 minute rainfall accumulations.

There is a fundamental difference between radar and ground based measurements: radars observe precipitation at height while rain gauges measure close to the surface. A number of meteorological processes (e.g. evaporation, drifting, fall speed) affect raindrops on their way to the surface. Due to these factors, even if the conversion from radar reflectivity to rain rate were reasonably accurate, there are difficulties in comparing rain rate at height with rain rate at the ground because the rainfall is being modified during its descent. While gauge data provide accurate information at the gauge location, the radar estimate is generally more accurate than a gauge-based estimate at distances from a gauge of more than 10 km or so.

2.4. **Reanalyses**

An atmospheric reanalysis is a combination of a numerical weather prediction model and a comprehensive set of observations of the state of the atmosphere such that they represent a best estimate of actual atmospheric conditions at any time. The formulations of the data assimilation scheme and the weather prediction model are consistent throughout the length of the simulation. The reanalysis data sets are used frequently to drive GCMs and RCMs enabling determination of the models' ability to simulate the current climate and their biases.

Two reanalyses are used in this study:

- The NCEP-NCAR Reanalysis Project [*Kalnay et al.*, 1996]. This is the longest ongoing reanalysis using past data from 1948 to the present. The data used in this study cover the period 1990 to 2009, are 6-hourly and interpolated to a 2.5° latitude-longitude grid.

- The ERA-Interim reanalysis [*Dee et al.*, 2011] covers the period from January 1979 onwards, and continues to be extended forward in near real-time. The data used in this study cover the period 1980 to 2012, are 6-hourly and interpolated to a 1.5° latitude-longitude grid.

3.   Description of Methods

3.1.   Extreme value theory

Univariate extreme value theory describes the statistical behaviour of the maximum values of a sequence of observations in a long period, such as a year [*Jenkinson*, 1955]. The statistical distribution that describes these data is known as a generalized extreme value (GEV) distribution. This distribution is usually fitted to data whose statistical properties do not change over time; however it is possible to modify the distributional parameters to model increases or decreases to the intensity of extremes over time [*Coles*, 2001; *Westra and Sisson*, 2011]. This feature is used to explore whether there are statistically significant changes in the intensity of rainfall extremes over time.

For the sub-daily rainfall trend analysis, the parameters of the model were fitted using the maximum likelihood method [*Coles*, 2001]. A spatial version of the method was used that allowed the extremes from all of the 69 sites to be estimated jointly, as this increases the precision of the parameter estimates [*Chandler and Bate*, 2007; *Westra and Sisson*, 2011].

3.2.   CSIRO Mk3.5 Global Climate Model

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) Mk3.5 GCM consists of the CSIRO spectral atmospheric GCM coupled to the Geophysical Fluid Dynamics Laboratory MOM2.2 ocean model [*Gordon et al.*, 2010]. It has been developed as a unified package including land and polar ice models. It was run using T63 resolution (about 2° × 2°) with 18 vertical levels on hybrid sigma-pressure levels. This model contributed to the Coupled Model Intercomparison Project Phase 3 ensemble (http://cmip-pcmdi.llnl.gov/) using a number of emission scenarios including the "Climate of the 20th Century" and the SRES A2 emissions scenario used here.

3.3.   ACCESS1.0 Global Climate Model

The ACCESS coupled model has been developed at the Centre for Australian Weather and Climate Research, a partnership between CSIRO and the Bureau. In this study, the ACCESS1.0 global climate model [*Bi et al.*, 2013; *Dix et al.*, 2013] is used to provide boundary conditions for the CSIRO conformal Cubic Atmospheric Model (CCAM) for the period from 1980 to 2065. ACCESS1.0 has an atmospheric horizontal resolution of 1.875° × 1.25° with 38 vertical levels. The coupled ocean model has 50 vertical levels and 1° horizontal resolution, increasing to 1/3° near the equator [*Bi et al.*, 2013].

3.4.   CSIRO Conformal Cubic Atmospheric Model (CCAM)

CCAM is a non-hydrostatic, semi-implicit, semi-Lagrangian atmospheric climate model, developed and maintained at CSIRO [*McGregor*, 2005]. The model uses a conformal cubic grid that can be stretched using a Schmidt transform [*Schmidt*, 1977] to focus the grid points over a particular area. In this way, CCAM avoids problems usually associated with lateral boundary conditions in RCMs and allows the regional climate to interact with the global circulation. The non-hydrostatic

version of CCAM allows better representation of rainfall processes at length scales below 10 km. This project is the first use of CCAM for high resolution rainfall modelling at regional scales.

Each of the sets of simulations performed progressed from a global 50 km resolution CCAM simulation, which was nudged towards the host model (or reanalysis) conditions every 3 hours via the spectral filter for the duration of the simulation (1980 to 2012 for ERA-Interim forced, and 1980 to 2065 for the ACCESS1.0-forced run), then through a continuous 10 km resolution run over south-eastern Australia performed over the same period (using the 50 km run to provide nudging). The 10 km simulations provided the boundary forcing to two 2 km domains, over the Greater Sydney region and surrounds and over Brisbane and south-east Queensland. CCAM grids used for each resolution are shown in Figure 3, while the area of each 2 km domain extracted for analysis is shown in Figure 4.
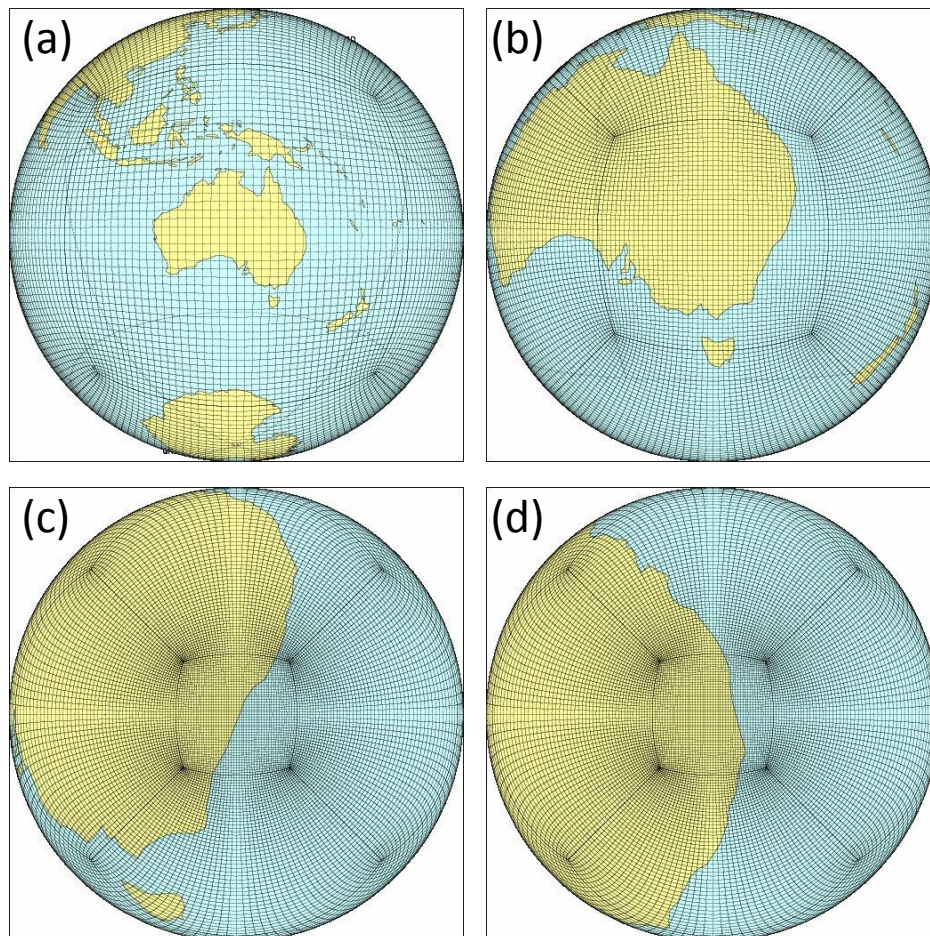

Figure 3. CCAM stretched-grid configurations for (a) global 50 km resolution simulations, (b) 10 km resolution simulations over eastern Australia, and for 2 km resolution simulations over the (c) Greater Sydney region and (d) Brisbane/southeast Queensland domains. Only every fourth grid line is shown for visual clarity.
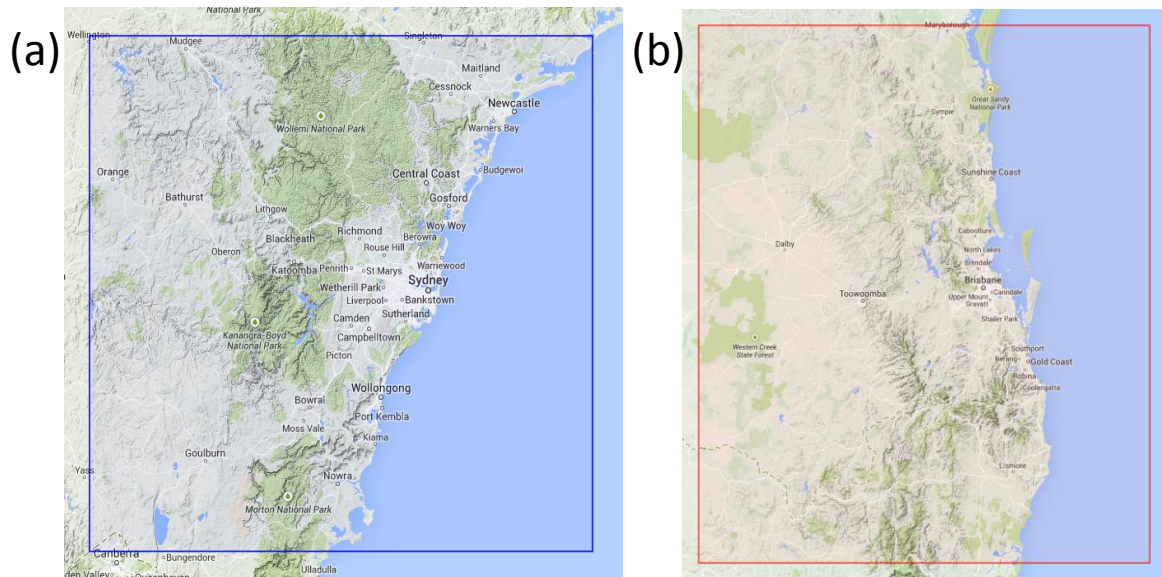
Figure 4. Geographic areas covered by the 2 km grid in high resolution CCAM simulations for the (a) Greater Sydney region and (b) south-east Queensland domains. (Not to scale.)

Due to the computational resources and time required to perform a continuous simulation at such high resolution, the ACCESS1.0-forced 2 km simulations over Sydney and Brisbane were separated into two time slices; the baseline period was run from 1980 to 2012, while the future climate period was run from 2040 to 2065. From these periods a baseline (1990-2009) and a future (2040-2059) climate period were extracted for comparison with other components of the project.

3.5. Weather Research and Forecasting (WRF) model

The Weather Research and Forecasting (WRF) modelling system is developed as a collaborative partnership between NCAR, NCEP, the Forecast Systems Laboratory, the Air Force Weather Agency, the Naval Research Laboratory, Oklahoma University, and the Federal Aviation Administration in the USA, as well as the wider research community. The version used in this study is the Advanced Research WRF version 3, maintained at NCAR [*Skamarock et al.*, 2008].

The boundary conditions for the 2 km simulation were taken from a previously performed and evaluated 10km simulation [*Evans and McCabe*, 2010]. The model used the following physics schemes: WRF Single Moment 5-class microphysics scheme; the Rapid Radiative Transfer Model longwave radiation scheme; the Dudhia shortwave radiation scheme; Monin-Obukhov surface layer similarity; Noah land-surface scheme; the Yonsei University boundary layer scheme and the Kain-Fritsch cumulus physics scheme.

For the 2 km simulation, no cumulus physics scheme is required, and the cloud microphysics scheme plays a more important role warranting the use of a more complex scheme. Several microphysics schemes were run for 1990 conditions and then evaluated against gridded and station based observations. The chosen scheme is the Thompson microphysics scheme [*Thompson et al.*, 2004]. There are 40 vertical atmospheric levels in the 2 km simulation.

There are two sources of global boundary conditions for the model. First a simulation was performed using global boundary conditions from the NCEP/NCAR Reanalysis Project [*Kalnay et al.*, 1996]. The reanalysis represents a best estimate of actual atmospheric conditions at that time. Using these boundary conditions allows this simulation to be compared directly with observations for a thorough evaluation of model performance. The second source of global boundary conditions is the CSIRO Mk3.5 GCM simulation using the SRES A2 emission scenario. Simulations for both the recent past (1990-2009) and a future time slice (2040-2059) were performed.

The modelled domain is shown in Figure 5. Two continuous simulations were performed for the 1990-2009 period. The first used reanalysis boundary conditions and provides an indication of the regional models performance given accurate boundary conditions. The second used GCM boundary conditions. Evaluation of this second simulation provides a quantification of the errors associated with the combined GCM-RCM system. The same GCM is then used to drive a future simulation (2040-2059) allowing assessment of the estimated future changes.
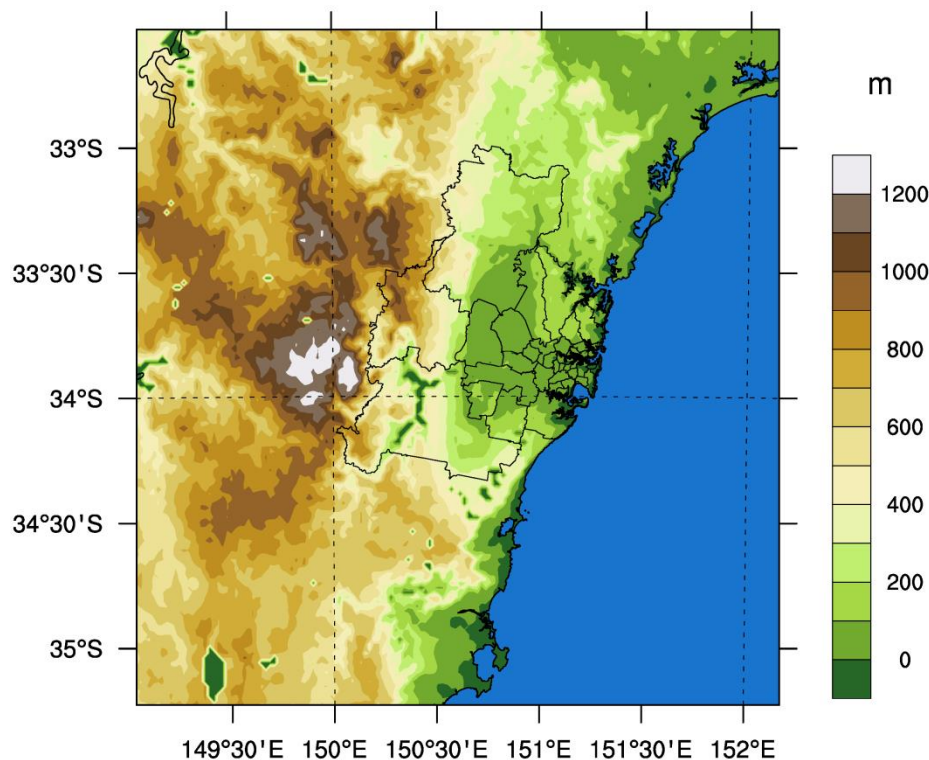
Figure 5. High resolution simulation domain for WRF model with topography. Black lines indicate the Greater Sydney region.

3.6.   Evaluation based on radar data

3.6.1. Identifying limitations in observational and model data

Preliminary investigations were undertaken to assess and understand the suitability of available observational and climate model data for the planned evaluation. These investigations indicate that biases exist in both observational data sets. It is therefore sensible - where possible – to make use of more than one observational data set for the evaluation of downscaled rainfall extremes to

gain a better understanding of the uncertainty in observed rainfall extremes. AWAP estimates are less reliable in regions with steep gradients and for sparse networks, and the intensity and frequency of extreme events is generally underestimated. Based on analyses of annual and seasonal totals, Rainfields underestimates rainfall depth. These biases are large for low and light rainfall but smaller for extreme rainfall. Statistical measures for AWAP data for the period 1990-2009 are similar to those for the period 2009-2012 for which Rainfields data are available, indicating that Rainfields data can be used to evaluate the model simulations.

The bias-corrected WRF-GCM simulations for both the current climate and the climate projections show a strong, increasing trend in annual rainfall accumulations over each of the two 20-year simulation periods, whereas no such trend is apparent in the observations for the period of investigation. This trend makes it difficult to calculate meaningful statistics over the time slice as a whole. While no such trend was found in the CCAM simulations, CCAM model output is available at 10 km and 2 km resolutions, which enabled the effect of model resolution to be investigated. It was found that the choice of resolution does not markedly affect the results of the evaluation.

3.6.2. Analysing spatio-temporal characteristics of observations and model simulations

In evaluating downscaled rainfall extremes it is desirable to answer questions like:
- How well do spatio-temporal characteristics of modelled and observed rainfall extremes match?
- Are modelled and observed events of similar intensity?
- Are modelled and observed events of similar spatial extent?
- Are peak intensities modelled in the location where they have been observed?
- Do models perform particularly well/poorly for certain seasons or synoptic situations?
- How well do models capture short duration rainfall extremes?

Maps of annual and seasonal rainfall accumulations are used to provide an initial impression of how well observations and simulations match. They are useful in identifying obvious biases in gridded data and to explore how these biases vary across the domain. If there is a consistent bias (over or underestimation) this will become very clear in the annual and seasonal accumulations. While it is informative to assess the simulations in terms of averages, it is important in the context of this study to evaluate simulations of rainfall extremes. Maps of 95th and 99th percentile of daily rainfall accumulations were used for that purpose.

Any biases identified in the average annual rainfall accumulations are not necessarily the same as those present in extreme events. Furthermore, biases for low and high intensity events may partially compensate each other. Histograms of daily and hourly accumulations, boxplots and quantile-quantile plots were used to compare the frequency distributions of observed and simulated data. Where the distribution of simulated rainfall accumulations is distinctly different from the distribution of observed rainfall accumulations, the simulated changes are less reliable and their interpretation becomes more complicated.

The set of measures used to characterise the degree of spatio-temporal organisation and lifetime of events in observations and simulations include temporal autocorrelation, wet spell duration (consecutive hours of precipitation) and feature analysis (contiguous areas with rainfall above a given threshold).

Short-duration rainfall extremes in the Greater Sydney area occur more frequently during the warmer months and the highest rainfall extremes tend to occur during summer, and changes in rainfall extremes are likely driven by changes in summer rainfall. It was therefore also assessed how well downscaled rainfall extremes replicate the seasonality of observed rainfall.

Daily rainfall totals for five rainfall events were analysed to explore how well reanalysis driven model simulations replicate the characteristics of observed events. The selection of events in this report was based on the following criteria: spatially averaged rainfall, rain cell fraction (percentage of the domain that has received rainfall), maximum rainfall intensity and information about the associated synoptic conditions and impact of an event. An important caveat in the interpretation of these comparisons is that '*the ability to model rain that resembles an observed event depends on whether the large-scale synoptic situation is sufficient to constrain the model outcome without the benefit of high-resolution initial conditions'*. (M. Thatcher, pers. comm.)

3.7. **Bayesian Hierarchical Modelling of rainfall extremes**

The statistical model developed herein is similar to the spatial BHM of *Davison et al.* [2012], but we extend it in several ways. In particular, we incorporate the scale-duration relationship of *Koutsoyiannis et al.* [1998], which allows us to model IFD curves, and we also provide a means of integrating climate model output in order to estimate changes in rainfall extremes at different durations. Appendix A describes the model in greater detail, and descriptions of earlier versions may also be found in *Lehmann et al.* [2013] and *Soltyk et al.* [2014].

The analysis of extremes is based on the GEV distribution [*Coles*, 2001]. It has three parameters that have to be estimated from station data or climate model output: location ($\tilde{\mu}$), scale (σ), and shape (ξ). For extreme rainfall of some duration d (for example, 24 hours), the scale parameter will depend on the value of this duration, but we can model this dependency using the expression in *Koutsoyiannis et al.* [1998], which introduces two additional parameters. Estimation of these parameters is carried out in a spatial Bayesian framework [*Davison et al.*, 2012]. When they have been estimated, we can then calculate quantities of interest such as return levels of extreme rainfalls across a range of durations and corresponding to different annual exceedance probabilities.

The model we have developed allows us to:

- Use information from neighbouring sites, either pluviometer or daily-read data, to increase the precision of estimates of design rainfall IFD curves.

- Provide estimates of IFD curves at locations where there is only daily-read data or at ungauged locations.

- Produce estimates of uncertainty in the IFD curves that are consistent in the sense that,

for example, the uncertainty in an IFD curve at an ungauged location should be greater than the uncertainty at a gauged location. Also, the width of uncertainty bands reflects the effects of sampling variability on estimated rainfall IFD curves.

- Integrate information from regional climate models so that it is possible to estimate rainfall IFD curves under climate change scenarios.

The results from the spatial BHM compare well with design rainfall IFD curves in ARR 2015, despite the fact that the two methods used to derive IFD curves are quite different.

4. What do the Observations Tell Us?

4.1. Sub-daily rainfall data

Figure 6 shows the sign and statistical significance of the trend in the 5-minute (left) and 12-hour (right) annual maximum rainfall totals estimated from a point-wise non-stationary GEV model for the period 1966 - 2012. The red and blue triangles respectively represent statistically significant upward and downward trends (10% significance level), and the red and blue open circles respectively represent non-significant upward and downward trends. The value of the 10% significance level was obtained using a bootstrap resampling approach applied to the observational time series at each rainfall station [*Zheng et al.*, 2013].

The majority of the rainfall stations (48 out of 69, or 70%) exhibited positive trends in five-minute rainfall extremes, of which 21 stations (30%) showed statistically significant increases at the 10% significance level (red triangles). In contrast, 48 stations out of a total of 69 have experienced declines in rainfall intensity for 12-hour annual maxima, with 15 sites showing statistically significant decreases. For annual maximums longer than 12 hours, there were similar decreasing patterns to the 12-hour results, albeit not statistically significant.
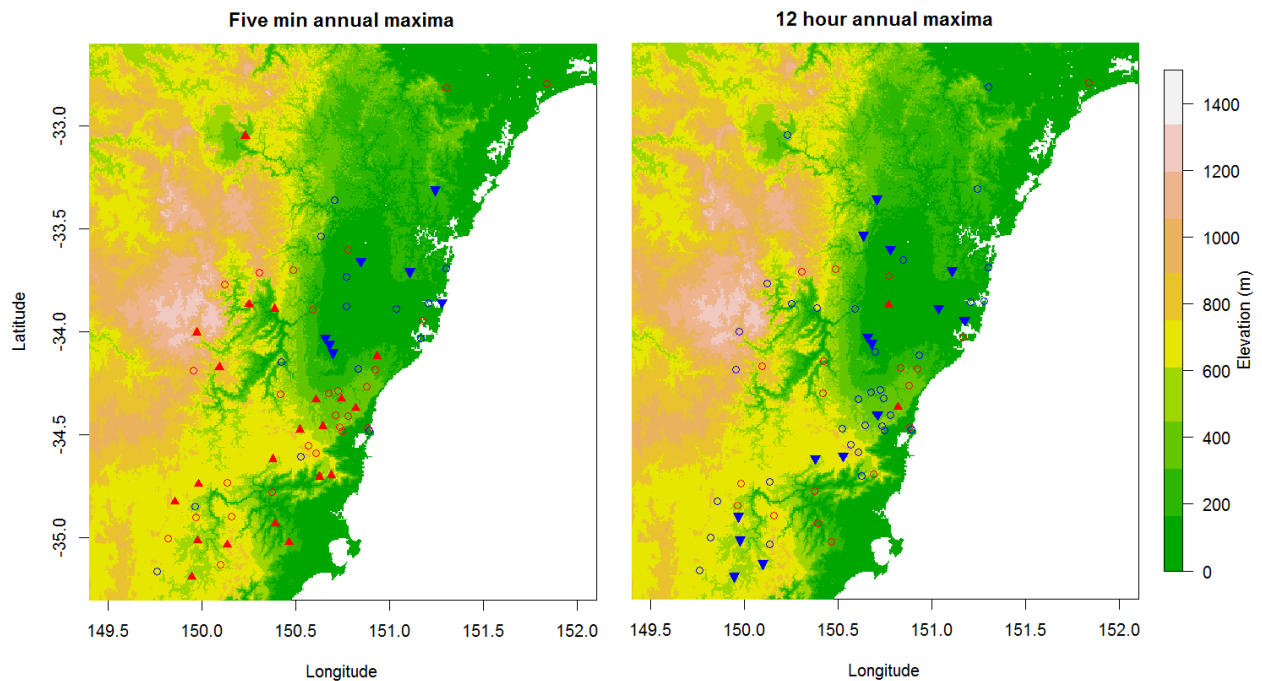
Figure 6. Trends in the annual maximum rainfall estimated from the point-wise non-stationary GEV model for the period 1966 to 2012. The red and blue triangles respectively represent upward and downward trends at a 10% significance level. The red and blue open circle dots indicate the positive and negative trends (but are not statistically significant at a 10% level).

A comparison of seasonal trends found that (a) summer extremes were significantly increasing at all durations, (b) winter and autumn extremes were significantly decreasing for durations greater than 3 hours, and (c) spring was not significantly changing for all durations. Figure 7 presents the percentage of annual maxima for each season across different durations. Extreme rainfall with short

durations was more likely to take place in summer with 50% of the sub-hourly annual maxima occurring in summer. In contrast, extreme rainfall with long durations was approximately equally likely to occur at any time of year.

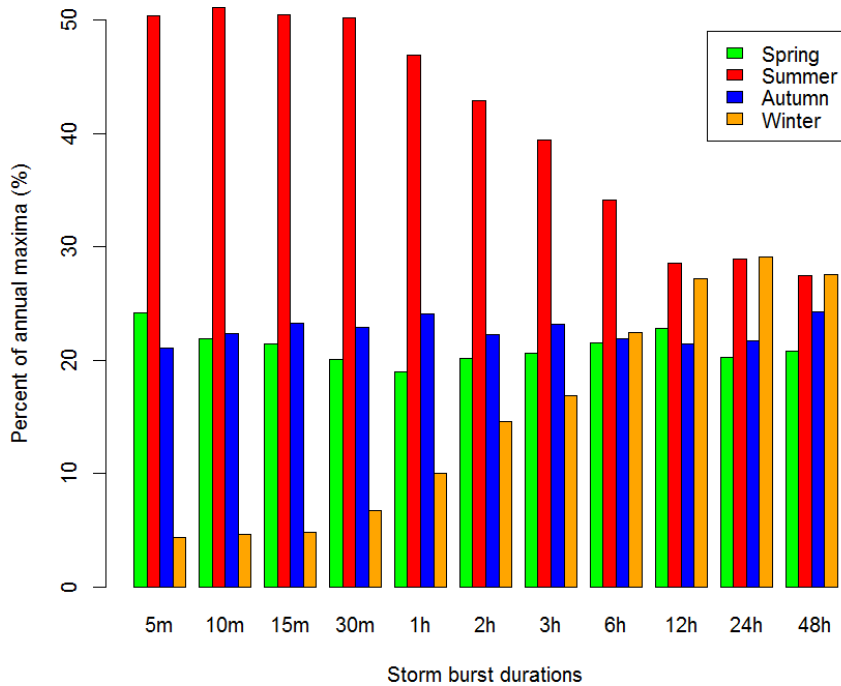**Figure 7. Percent of annual rainfall maxima for each season over the period from 1966-2012 against different durations.**

Combining the results in Figures 6 and 7, it is deduced that increases in the intensity of annual rainfall extremes with short durations are largely associated with upward trends in summer, while the declines in the annual maxima with longer durations are dominated by the decreases in autumn and winter. Further details can be found in *Zheng et al.* [2015].

4.2. Comparison of design rainfall IFD curves using regionalisation and spatial statistics

An overview of the methods used by the Bureau to estimate design rainfall IFD curves for ARR 2015 may be found in *Green et al.* [2012]. They are based on a regional frequency analysis (RFA) [*Hosking and Wallis*, 1997], which has been widely used in the hydrological sciences. By contrast, Bayesian spatial methods are relatively recent, but they have several attractive features [*Soltyk et al.*, 2014]: in particular, they are based on a flexible, unified and coherent framework in which uncertainty estimation is a natural by-product. Compared to the BHM, however, it is somewhat less straightforward to incorporate drivers of climate change into methods based on RFA, and so we have used the BHM to assess how return levels of extreme rainfall might change in the future relative to a current, baseline period. Of course, to have some confidence in the estimated changes from the BHM, we require that it produce current climate results that are similar to the design rainfall IFD curves in ARR 2015.

Our overall conclusion is that, despite the difference between the methods, the results from the BHM and the IFD revision are indeed similar in both the Greater Sydney and south-east Queensland regions, *especially in light of the uncertainty in the IFD curves*. For example, Figure 8 shows a comparison of IFD curves generated by the BHM and from the IFD curves in ARR 2015 at grid location (latitude 33.4875°S, longitude 150.5125°E) in the Greater Sydney region. Clearly the IFD revision estimates lie within the 95% uncertainty interval generated by the BHM. Naturally, the curves will not be identical, but the correspondence is good. In addition to methodological differences between the two techniques, the BHM uses only rainfall data from 5 min to 48 hours, and then extrapolates to estimate depth-duration relationships beyond that range. For durations less than 5 min this is also true for the method used in ARR 2015.
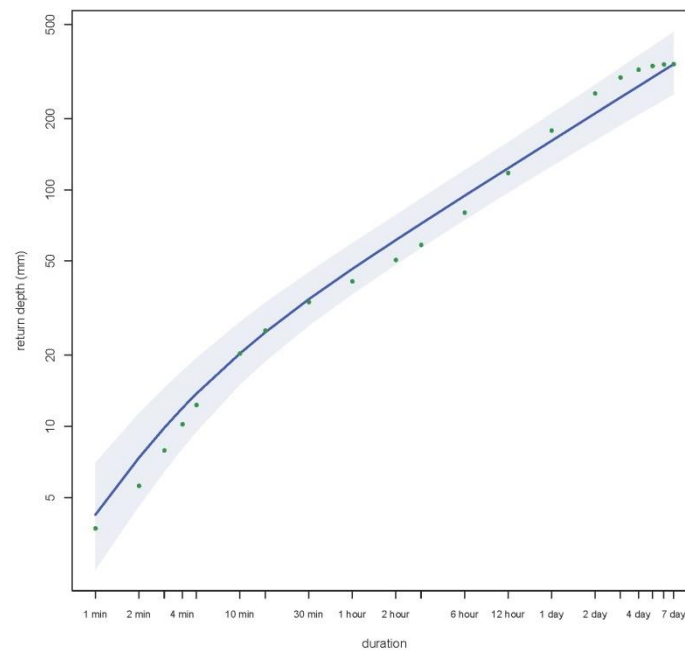

Figure 8. IFD curve for a grid location (33.4875S, 150.5125E) in the Greater Sydney region showing results from the BHM (solid line) and the IFD estimates from ARR 2015 (green points). 1% AEP and the 95% uncertainty interval is shown as a light blue band.

4.3.   **Climate model evaluation**

Maps of mean annual rainfall accumulations for the period 1990-2009, except for the Rainfields data set which covers 2009-2012, are given in Figure 9. The data sources for the maps are: AWAP, Rainfields (section 2.3); CCAM simulations at 2 km resolution forced by the ERA-Interim Reanalysis (CCAM-Re2); WRF simulations at 2 km forced by the NCEP/NCAR Reanalysis (WRF-Re); CCAM simulations at 2 km resolution with ACCESS 1.0 as the host climate model (CCAM-GCM2); and bias corrected WRF simulations at 2 km with CSIRO Mk3.5 GCM as the host climate model (WRF-GCM-bc). Perusal of Figure 9 indicates that model simulations capture some of the spatial patterns exhibited by observations. The CCAM-Re2 run appears more realistic than the CCAM-GCM2, both in terms of magnitude and in replicating spatial variability. When compared against the AWAP analyses the CCAM-GCM2 run strongly overestimates annual accumulations (by about a factor of two) while CCAM-Re2 provides a somewhat better match when compared against the AWAP analyses. Bias correction of the GCM driven WRF simulations leads to WRF-

GCM-bc performing better than the (uncorrected) reanalysis-driven WRF-Re run. The annual accumulations for WRF-GCM-bc simulations are a good match to the AWAP data, both in terms of magnitude and spatial pattern, and provide a better match than the WRF-Re simulation. The annual accumulations for WRF-Re, while overestimating the magnitude, match some of the spatial gradients shown in Rainfields.
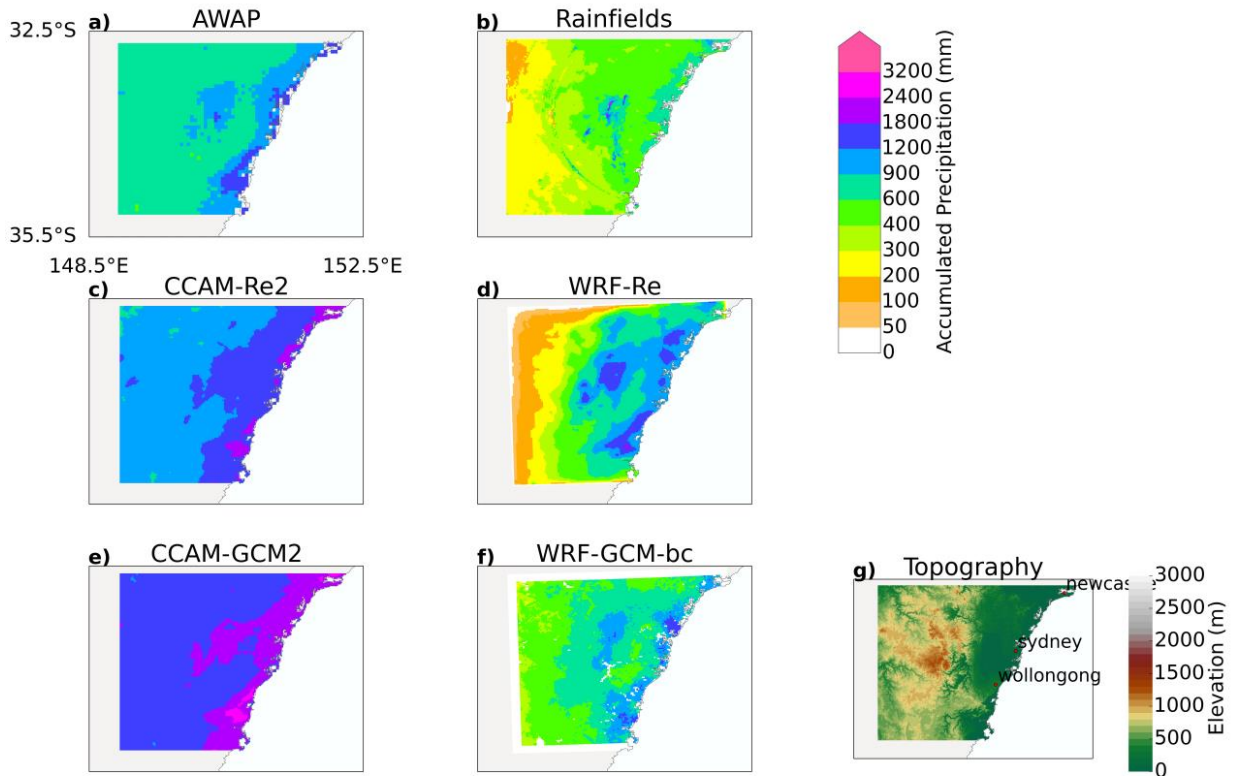

Figure 9. Mean annual rainfall for 1990-2009 over the Greater Sydney region for (a) AWAP, (b) Rainfields (2009-2012), (c) CCAM-Re2, (d) WRF-Re, (e) CCAM-GCM2 and (f) WRF-GCM-bc and (g) topography of the study area.

While it is informative to assess the simulations in terms of annual rainfall accumulations, in the context of this study it is important to evaluate simulations of heavy rainfall. Maps of the 95th percentiles of daily rainfall fields are given in Figure 10. As noted in section 2.2, the AWAP data has been found to underestimate the 95th percentile rainfall by ~10% in this area. For the AWAP data, the highest values of this percentile occur along the coastal regions. This pattern is matched by Rainfields although the magnitude of the Rainfields accumulations is notably lower. Both of the CCAM simulations exhibit consistently higher estimates than the observed but exhibit the same broad spatial pattern. CCAM-GCM2 shows little spatial variability while CCAM-Re2 looks more realistic. WRF-Re overestimates the magnitudes somewhat relative to Rainfields and does not replicate the location of the highest accumulations well. The magnitudes of WRF-GCM-bc are more realistic spatially although still too high and it appears the bias correction is compromising some of the spatial patterns seen in the raw WRF-GCM output (not shown).
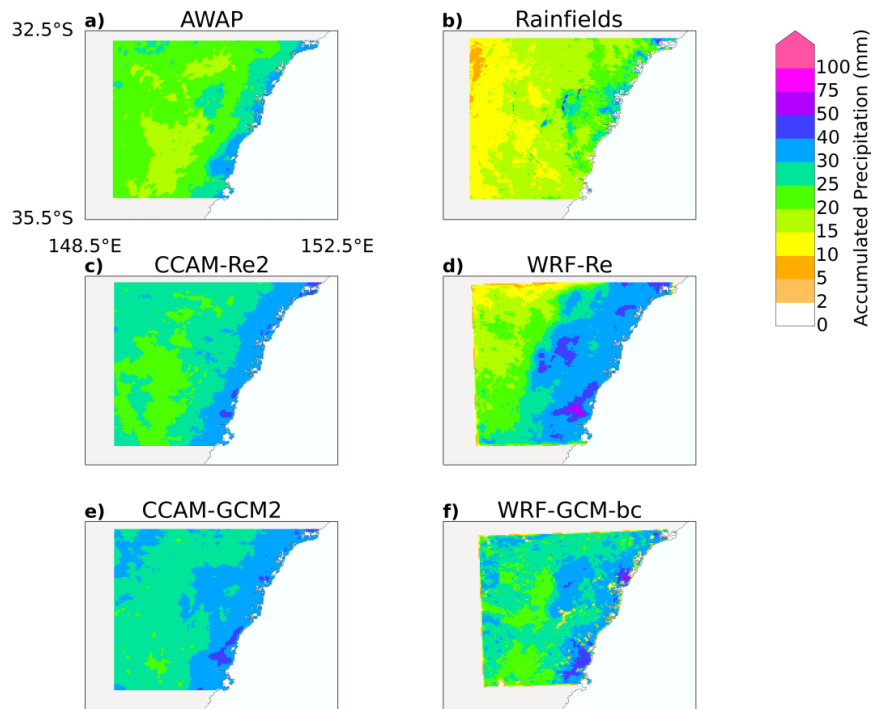
Figure 10. 95th percentile of daily rainfall in the Greater Sydney region over 1990-2009 period for (a) AWAP, (b) Rainfields (2009-2012), (c) CCAM-Re2, (d) WRF-Re, (e) CCAM-GCM2 and (f) WRF-GCM-bc.
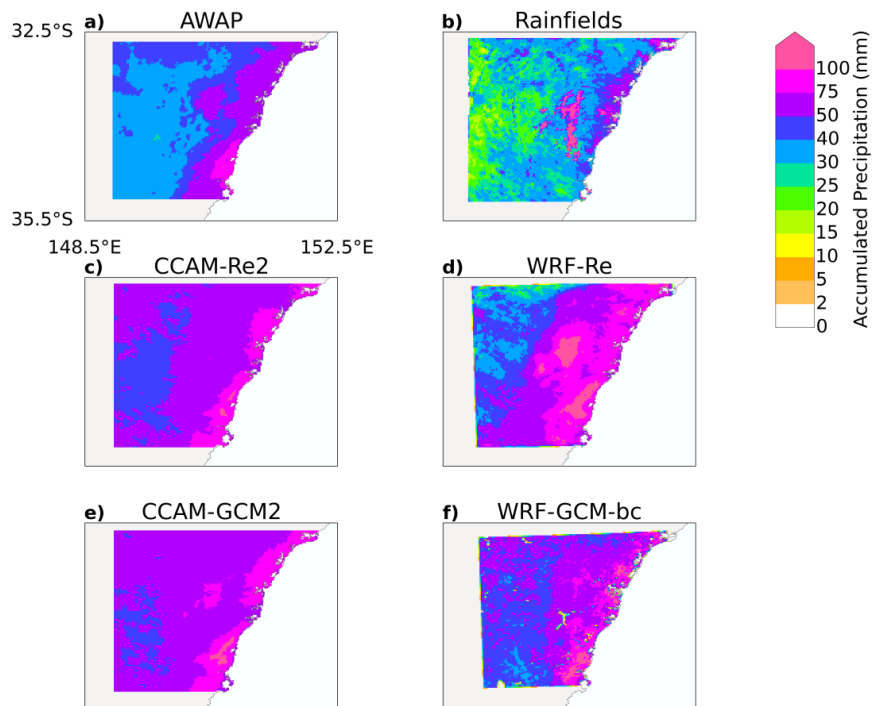
Figure 11. 99th percentile of daily rainfall in the Greater Sydney region over 1990-2009 period for (a) AWAP, (b) Rainfields (2009-2012), (c) CCAM-Re2, (d) WRF-Re, (e) CCAM-GCM2 and (f) WRF-GCM-bc.

Maps of the 99th percentiles of daily rainfall fields are given in Figure 11. While these maps exhibit the same general features as those for the 95th percentiles, they highlight the extent to which the model simulations overestimate high percentiles (≥ 95th) of daily rainfall for both the AWAP and Rainfields data.
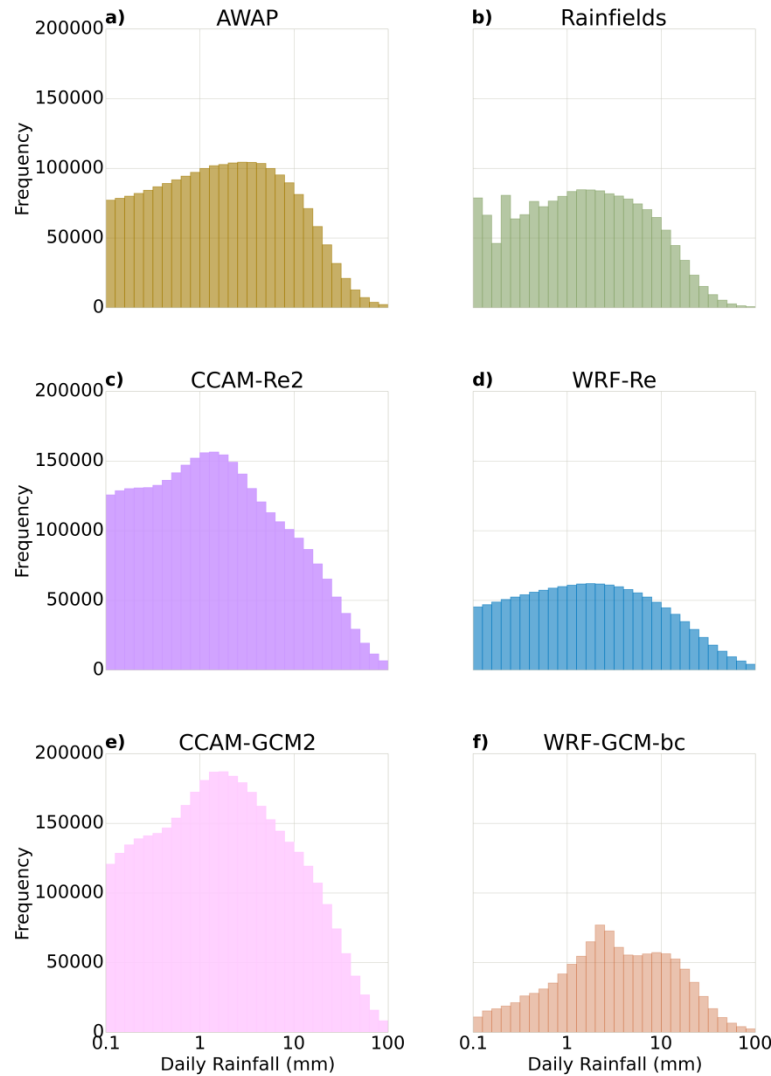
Figure 12. Histograms of daily rainfall accumulations (above 0.1 mm, regridded to 0.1°) for the Greater Sydney region (a) AWAP, (b) Rainfields (2009-2012), (c) CCAM-Re, (d) WRF-Re, (e) CCAM-GCM and (f) WRF-GCM-bc.

Histograms of daily rainfall accumulations are shown in Figure 12. The heavy lower tail of the AWAP data distribution is due to the smearing of light rainfall by the gridding algorithm. Apart from this feature, there is good agreement between the distributions of the AWAP and Rainfields data. The Rainfields data are stored at a resolution of 0.05 mm, and this granularity leads to the spikes

seen in the lowest class intervals. The distributions of CCAM daily rainfall also exhibit a heavy lower tail due to the 'drizzle problem' in which climate models typically exhibit too many days with light rainfall and too few dry days. Also, these distributions are more peaked than those for the AWAP and Rainfields data indicating that the CCAM simulations produce a higher number of moderate to heavy rainfall events. The CCAM-GCM2 simulation has an unrealistically heavy upper as well as lower tail. The distributions of the WRF simulations are much flatter than those of AWAP, Rainfields and CCAM. In the WRF-GCM-bc case the bias correction removes the heavy lower tail (perhaps too strongly) and the heavy upper tail (perhaps insufficiently). It also exhibits a bimodality that is not found in the observations.

The seasonal distributions of daily rainfall accumulations are given in Figure 13. Observations (AWAP and Rainfields) show that extreme rainfall events in the study region occur most frequently during summer (December to February). This behaviour is replicated in the simulations for both the daily and hourly accumulations. The WRF simulations achieve a better approximation of the seasonal distributions than those for CCAM. In particular, the CCAM simulations very strongly overestimate the frequency of events so that the frequency of moderate events exceeds the common scale used for the comparisons.

Figure 13. Histograms of daily rainfall accumulation. Line colour indicates season for the Greater Sydney region (a) AWAP, (b) Rainfields (2009-2012), (c) CCAM-Re2, (d) WRF-Re, (e) CCAM-GCM2 and (f) WRF-GCM-bc.
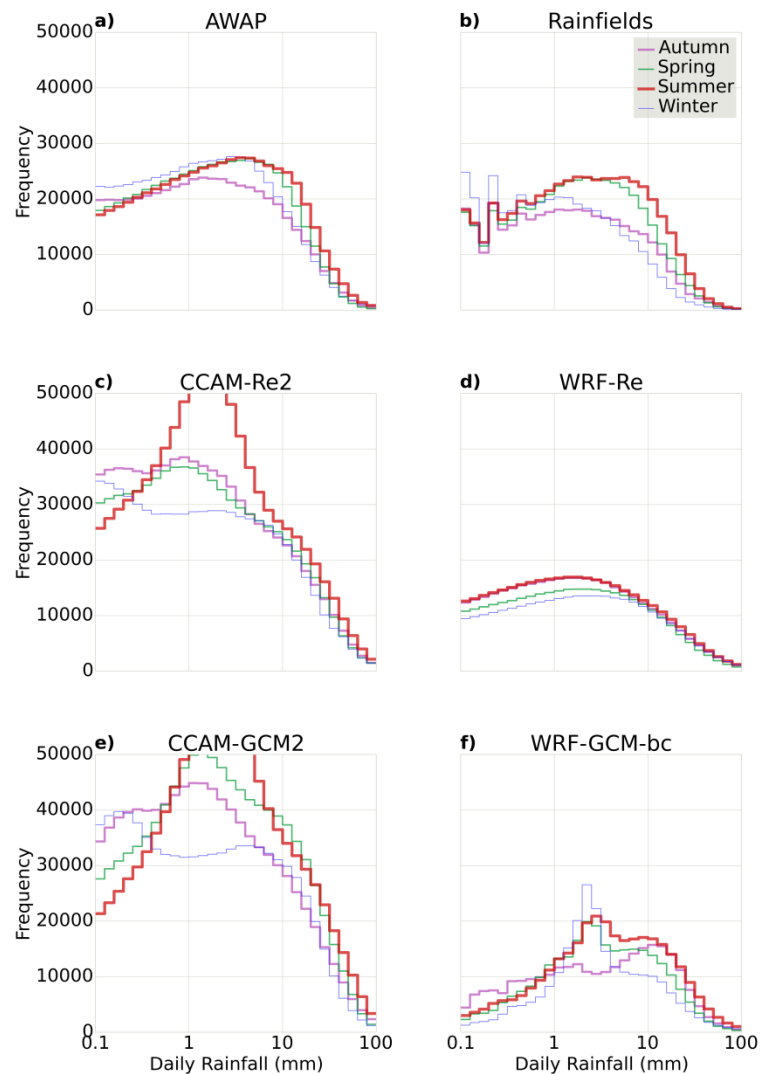
As described in Appendix A, the BHM incorporates climate change information from dynamical downscaling models by relating the parameters of the GEV distribution of the station data to the corresponding parameters obtained from the extreme rainfall generated by the climate models. If there is some relationship between these two sets of parameters, this information can be used to 'bias-correct' parameter estimates obtained from rainfall in a future period in order to obtain estimates of what may happen under climate change. Thus, in addition to some of the methods already discussed in this section, an implicit way of assessing whether a climate model can reproduce observed patterns of extreme rainfall is to compare GEV parameter estimates from station data and climate model output over the same period of time. Of the climate model runs considered in the project (CCAM: 2 and10 km in Greater Sydney and South-east Queensland; WRF: 2 km in Greater Sydney), only the scale parameter estimated from WRF output is related to the scale parameter estimated from station data.

Figure 14 shows pairwise plots of the scale parameter calculated from pluviometer data in the Greater Sydney region and the scale parameter in corresponding grid cells from WRF output for the baseline and future periods. There is a roughly linear relationship in all three plots; WRF tends to yield much larger corresponding estimates of the scale parameter than the station data. Hence, although WRF may not reproduce the magnitude of extreme rainfall in the current period, its bias is consistent across the region. The location and shape parameters do not exhibit the same strong relationships as the scale parameters. Consequently, projected changes in rainfall extremes will be driven by changes in the scale parameter.

In contrast with WRF output, GEV parameters estimated from CCAM output over the Greater Sydney and South-east Queensland regions show little or no relationship to parameters estimated from data, because the bias is neither consistent across durations nor across the region. A set of plots similar to Figure 14 for the CCAM output (not shown) display an unstructured cloud of points for all the GEV parameters. We cannot, therefore, integrate CCAM output into the BHM in order to produce modulated estimates of the impact of climate change on design rainfall IFD curves.

Spatial plots of mean annual maximum 6-hourly rainfall during the baseline period in the Greater Sydney region for data from pluvio stations and from 10 km resolution CCAM output are shown in Figure 15. Clearly CCAM overestimates extreme rainfall over most of the region; the pattern across durations is not, however, consistent – at shorter durations, CCAM underestimates extreme rainfall (not shown).
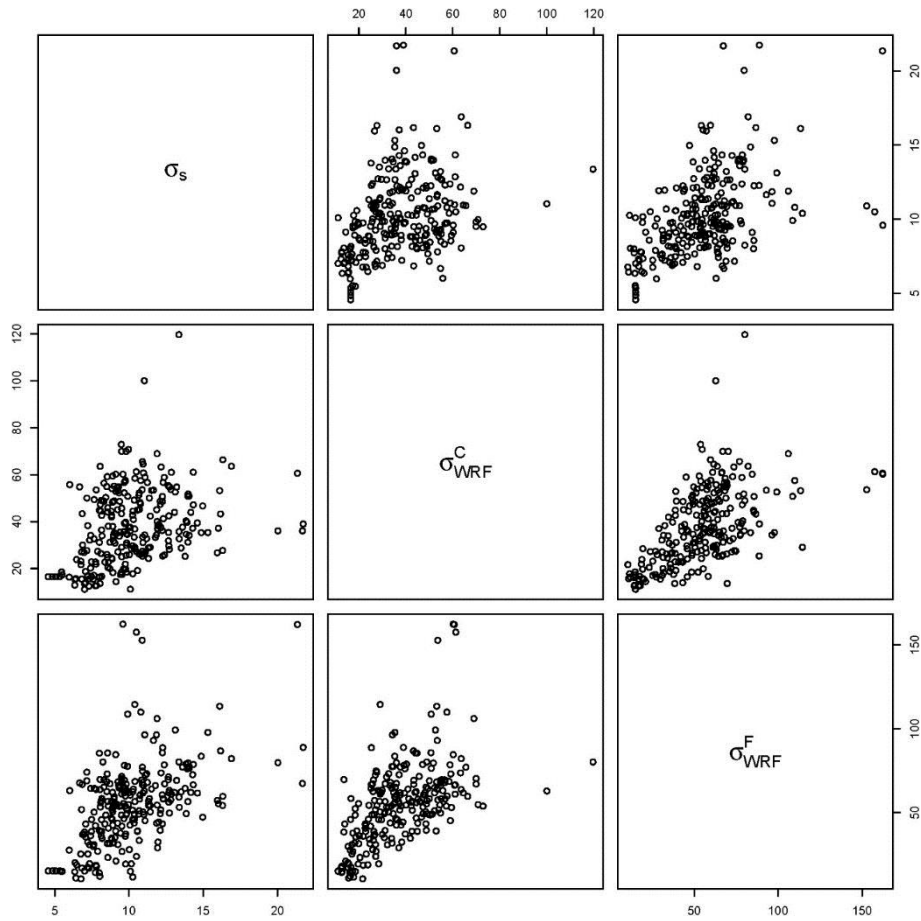
Figure 14. Greater Sydney region: Pairwise plots of GEV scale parameter obtained from pluvio station data (σ_s), WRF output in the baseline period (σ^c_{WRF}), and WRF output in the future (2040–2059) period (σ^F_{WRF}).
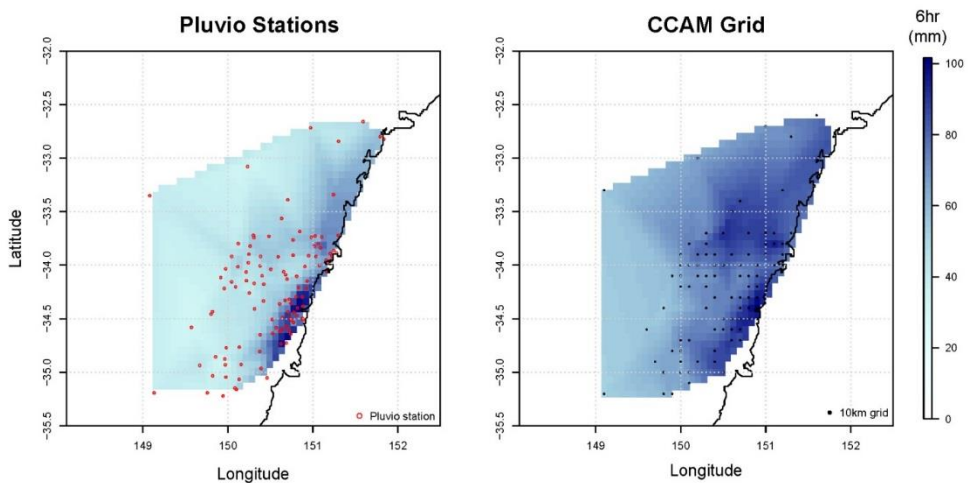
Figure 15. Spatial plots of mean annual maximum 6-hourly rainfall during the baseline period in the Greater Sydney region: pluvio stations (left) and 10 km CCAM output (right). The same colour scale has been used for both plots: CCAM overestimates extreme rainfall.

4.4. Summary

The findings for CCAM and WRF simulations for current climate conditions show a number of similarities so that it may be reasonable to assume they are more widely applicable. On the other hand, both models show similar deficiencies and may therefore not provide a good indication of the range of projected changes in rainfall extremes. To improve confidence in the likely *range* of potential changes in rainfall extremes it will be necessary to use and assess a larger number of GCMs.

The findings from this study not only provide information about projected changes but may also provide valuable guidance with regards to required bias adjustments and useful feedback for modellers to help them choose parametrisations that result in the most realistic sets of simulations, or more generally improve their models. The following statements summarise key findings and their implications.

1. The intensities of observed short duration (< 2 hours) rainfall extremes (annual maxima) increased in the Greater Sydney region over the period from 1966 to 2012. Conversely, rainfall extremes with timescales greater than three hours were more likely to exhibit a decrease. The assessment was based on the analysis of 69 rainfall gauges over the Greater Sydney region, focussing on the common period from 1966-2012 and durations from 5 minutes to 48 hours. Analysis of seasonal maxima showed increases for all durations for summer, while autumn and winter rainfall extremes exhibited decreases at all timescales. The annual maximum results could be explained by the finding that extreme rainfall events with short durations were more likely to take place in summer, whereas longer-duration extreme rainfall was approximately evenly distributed throughout the year. This highlights the importance of seasonal effects associated with possible changes in annual maximum rainfall at different durations.

2. There are deficiencies in the gridded rainfall observations that are relevant to the evaluation of downscaled rainfall extremes. AWAP grids were specifically developed for the purpose of monitoring climate change and climate variability. These grids are based on sophisticated gridding techniques and are a definite improvement over previously available gridded datasets. Errors in average and extreme AWAP rainfall are well documented [*Jones et al.*, 2009; *King et al.*, 2013]. AWAP estimates are known to be less reliable in regions with steep rainfall gradients and sparse networks. The latter is not an issue for the study region but would need to be considered if similar analyses were to be undertaken for other parts of the country. Owing to the gridding techniques (splines and Barnes) the resulting grids are overly smooth.

Radar data provide an opportunity to study the spatio-temporal characteristics of rainfall and it is difficult to ignore their usefulness in the evaluation of downscaled rainfall extremes. Limitations to the accuracy of Rainfields data are associated with the fact that radar measures rainfall at height rather than at the ground. Complex algorithms are required to convert radar reflectivity to rainfall accumulations.

Analogous to the use of ensembles in climate model simulations, using more than one observational data set helps give an indication of inherent uncertainty in the observational data set because of different data sources and analysis techniques.

3. The frequency distributions of observed and simulated rainfall exhibit marked differences. The CCAM and WRF runs evaluated here typically simulate too many light rainfall events as well as unrealistically high rainfall extremes. This leads to distributions that are flatter than observed with a median value that is typically higher than for the observations. When developing design rainfall estimates a three-parameter distribution (Generalised Extreme Value distribution) is fitted to annual maxima. These distributions are described using location, scale and shape parameters. The task of correcting the simulated extremes is complex because it is equivalent to correcting parameter estimates. This raises questions about whether models correctly replicate rainfall-producing mechanisms and affects confidence in assessments of projected changes in design rainfall estimates.

4. Compared to GCM-driven simulations, reanalysis-driven CCAM simulations are typically a better match with observations. Simulations of future climate cannot be undertaken using reanalysis data. Confidence in the (GCM driven) CCAM runs for the climate projections is therefore lower than in (reanalysis driven) runs for the current climate.

5. Bias-correction of GCM-driven WRF output (WRF-GCM-bc) leads to a marked improvement in performance but does not fully address the bias in extremes (e.g. the modelled magnitude of the 95th percentile) and introduces artefacts. A histogram-matching approach is used to post-process WRF outputs [*Argüeso et al.,* 2013]. While this leads to improvements for average rainfall accumulations, it does not remove the tendency to overestimation of extreme rainfall in WRF simulations. The bias-correction also appears to introduce a bimodality to the rainfall distribution which is not representative of observations and results in an unrealistic spatial distribution for the extremes (95th and 99th percentile).

The post-processing of model output implicitly assumes that the bias correction valid under current climate conditions is also applicable for climate projections. These findings limit the degree of certainty in the projected changes in the magnitude of rainfall extremes.

6. Model simulations may not adequately replicate extent, location and magnitude of events. Based on the limited number of events assessed, it appears that simulations perform better for large-scale events than for small-scale events. This limits confidence that all rainfall-producing mechanisms have been correctly captured and relevant changes can be correctly projected.

7. Simulated rainfall is typically more organised in space and time than observed rainfall. This finding is based on comparisons of temporal autocorrelation (dropping off more quickly in simulations than observations), wet spell duration (consistently longer for simulations than observations) and visual assessment of spatial patterns for five events. This is important because

there are indications that changes in rainfall extremes may be driven to a considerable degree by changes in convective precipitation [*Berg, Moseley and Haerter*, 2013]).

8. The seasonality of downscaled rainfall extremes from WRF simulations agrees well with the seasonality of observed extremes. Analyses of historical observations show that any long-term change in rainfall extremes is likely to be driven by changes in rainfall extremes during summer. Rainfall extremes for the study region occur most frequently during summer and this is replicated in the simulations. The seasonal distributions from WRF simulations are a far better match with the observed than those from CCAM, especially since CCAM strongly overestimates the frequency of moderate events in summer.

9. Based on the comparisons of daily rainfall from CCAM runs at two resolutions (10 km and 2 km) there is little indication that increasing the resolution leads to an improvement in the simulation of rainfall extremes (at least for this configuration of the model). Given that model runs at higher resolution are computationally much more expensive it should be assessed whether the additional effort is justified. Modelling at high resolutions is an area of active research, and is not a mature science area. The ARR results are being used to develop a better suite of parameterisations more appropriate for the study of rainfall extremes using CCAM.

10. While model simulations show some skill in matching the peak intensities of rainfall extremes on an annual scale, this skill is lower when individual simulated events are compared to observations. It may therefore be preferable to consider providing indicative estimates of projected changes for the study region as a whole without distinguishing by geographical location within the study region.

11. Comparison of GCM-driven WRF simulations for the current and the future climate (1990 – 2009 and 2040 – 2059) indicates a projected increase in annual accumulations. A comparable increase in annual accumulations is not apparent from the observations or in the CCAM simulations and is probably unrealistic. This model artefact makes it more difficult to interpret the projections derived using WRF and lowers confidence in their use.

12. **Despite differences in methodology between the Bayesian Hierarchical Model and the current methods used by BoM for estimating current climate IFD curves, the results are remarkably similar.** The BHM produces uncertainty bounds as well, and in most cases, the BoM estimates of IFD curves lie within the uncertainty bounds produced by the BHM.

5. Climate Change Projections for Design Rainfall IFD Curves

In this section, we consider changes to rainfall extremes generated by the downscaling climate models alone, and by integrating WRF with the Bayesian Hierarchical Model.

5.1. Greater Sydney region

5.1.1. CSIRO Conformal Cubic Atmospheric Model (CCAM)

The rainfall amount simulated by CCAM at 2 km resolution has been shown to be too high (Figures 9 to 13, and 15). Consequently the 2 km simulations provided no significant improvement over the results provided by the CCAM 10 km resolution simulations, from which the 2 km simulation derived its meteorological forcings. It is likely that this over prediction of rainfall is a result of the model's parameterisation settings used to configure these simulations, particularly with respect to the way in which clouds and rainfall were represented in CCAM.

As CCAM is a stretched grid atmospheric model, its physical parameterisations must handle a wide range of spatial scales, which in turn poses a challenge for their development. For the choice of parameters used in these simulations, the CCAM parameters provide reasonable results for rainfall at scales of around 50 km. However, it is evident that for sub-10 km scales, these parameter settings are sub-optimal for resolutions where the model can start to resolve rainfall events. This implies that parameterisations need to be better generalised in order to more accurately represent physical processes at these finer spatial scales. This work is already underway as a result of the analysis completed for ARR.

When compared with the observation-based AWAP rainfall data set, the shortcomings in the CCAM 2 km resolution ERA-Interim downscaled simulations is that the rainfall is generally too heavy for almost all durations. This makes it challenging to separate the effect of climate forcing by the host GCM from the effect of the poorly configured CCAM rainfall parameterisations. Nonetheless, the use of ACCESS 1.0 GCM host model does appear to further amplify the overly wet nature of the CCAM simulations over the equivalent period (Figure 13).

The projected rainfall change results from CCAM over the Sydney domain show a regionally heterogeneous pattern, with percentage change values varying greatly with location (Section 5.1). While large regions of elevated terrain to the west of Sydney (see, e.g., Figure 5 or 6) are correlated with increases in the 1% AEP, it appears that the influence of orography is not the only factor. Many land areas near to the coast also experience large increases in the 1% AEP, as do large areas offshore. As such, there does not appear to be any obvious relationship between changes in the 1% AEP and surface features. This implies that a simple statistical relationship based on the underlying orography cannot explain the changes, and instead dynamic and thermodynamic effects are the likely source of any shift in extreme rainfalls.

It is advisable that the CCAM simulation results should be interpreted over a broad region (e.g. regionally or even domain-averaged) to account for short climatological periods (20 years) being examined at such fine spatial scales. For example, although the AEP plots of CCAM output for Sydney Airport (Section 5.1.3) do not display significant changes from the current climate, we

would caution against the over-interpretation of projections based upon individual grid points.

5.1.2. Weather Research and Forecasting (WRF) model

Given the known biases present in the WRF simulations, a methodology to correct these biases, as presented in *Argüeso et al.* [2013] is applied. The method is based on fitting both the model and the observational cumulative probability distributions to theoretical functions and comparing them to calculate the model deviations. For rainfall, both the observations and model are fitted to a gamma distribution. This method can successfully correct the mean bias but tends to have relatively small impacts on the extremes of the distribution.

Here the climate change projection is found by comparing the bias corrected CSIRO-Mk3.5 driven fields from the simulation covering the period 2040-2059 to the simulation covering the period 1990-2009. In terms of rainfall, most of the domain is projected to see mean annual increases of up to 40%. These increases occur mostly in autumn, with little change in summer. Extreme rainfall (higher than 95th percentile) is projected to contribute a larger proportion of this rainfall total.

The change in IFDs estimated by the bias corrected present day (1990-2009) WRF and the future WRF (2040-2059) is shown in Figure 16. By contrast with the results discussed in the next subsection, Figure 16 shows domain-averaged IFD curves. Increases in rainfall depths can be found across all durations and AEPs. These increases are larger for smaller AEPs and reach about 50% for 12 hour durations for the 1% AEP.

Examining the seasonality of the future change in annual maximum rainfall (Figure 17) reveals that the increases in total rainfall depth from annual maxima (a combination of the number of events and the rainfall depth of each event) occur in all seasons other than winter. Winter shows little change at all durations.
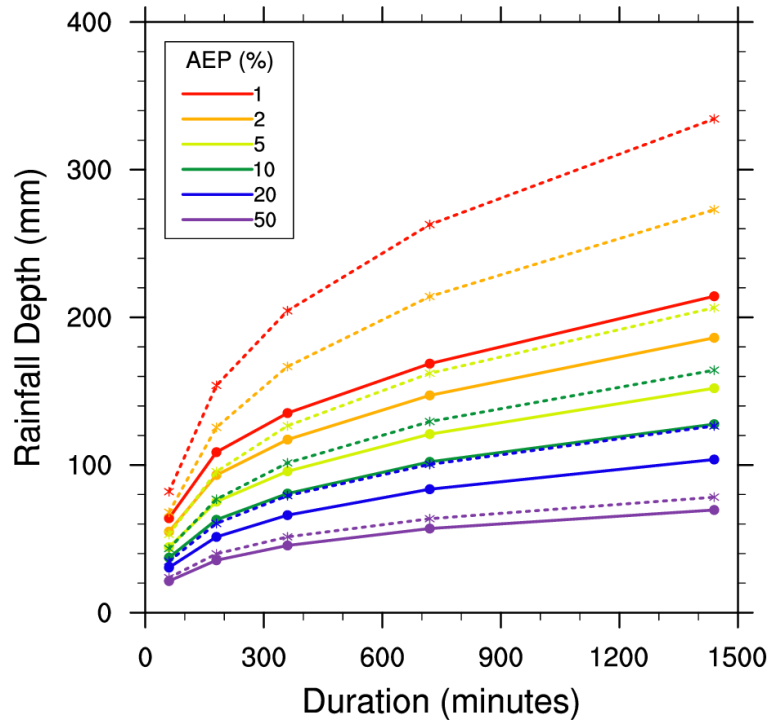
Figure 16. Domain average IFD curves for bias corrected WRF: 1990-2009 (solid lines) and 2040-2059 (dashed lines).
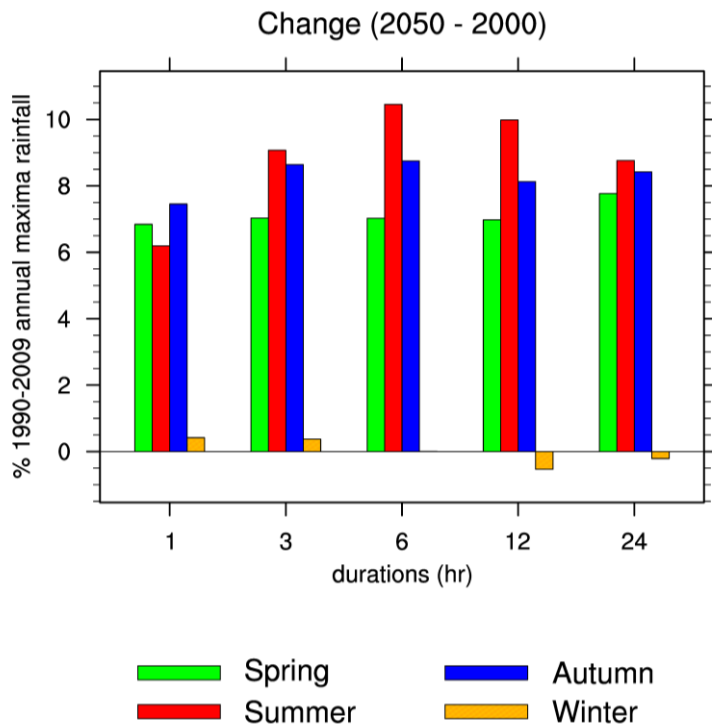

Figure 17. Seasonality of the 2040-2059 change in total rainfall depth from annual maximum rainfall expressed as a percentage of the 1990-2009 mean annual maxima rainfall depth.

5.1.3. Bayesian Hierarchical Model (BHM)

As shown in Section 4.3, neither ACCESS-CCAM nor CSIRO Mk3.5-WRF combination is able to adequately reproduce the magnitude of extreme rainfall in the baseline period, although there is indeed some relationship between patterns of extreme rainfall produced by WRF and rainfall station data, as shown in Figure 11. Therefore, it is not surprising that the model combinations alone, and the BHM integrated with CSIRO Mk3.5-WRF results, would yield different results under the two emissions scenarios considered here.

Figure 18 displays the percent difference in 1% AEP (100-year) return levels of the annual maximum 24-hour rainfall in the Greater Sydney region, calculated as the difference between future (2040 – 2059) and baseline (1990-2009) periods. Spatial plots for other durations indicate similar patterns (not shown). The top row shows results from CCAM (left) and WRF (right). The magnitude of the changes shows considerable spatial variability, and in many instances it is quite large for both CCAM and WRF. Indeed, the spatial patterns produced by the two downscaling climate models appear to be the reverse of each other in some regions. This is partly due to differences in the host model simulations and partly due to different downscaling techniques.

The bottom plot in Figure 18 shows the resulting changes when the Bayesian spatial model integrates WRF output. Clearly the WRF results are attenuated by the implicit bias correction of the BHM. It is also important to note that because changes in the scale parameter of WRF from baseline to future periods are driving the changes estimated by the spatial model, the pattern of changes we see in the WRF output is not reflected in the changes generated by the BHM. Thus, there are some regions where WRF projects decreases but the BHM projects increases. Over most of the Greater Sydney region, with the exception of the western and northern portions, the BHM projects increases of between 5 and 15%. The uncertainty is quite large, however, and we can see this more clearly in plots of rainfall IFD curves for four specific locations (Congewai, Dapto Bowling Club, Katoomba, and the Sydney Airport AMO) in Figure 19. These figures show BHM estimates of current IFD curves and future projections obtained by integration with WRF output for 10% and 1% AEP. They also include an estimate of future IFD curves obtained by using the change factor suggested in the ARR Interim Guideline for Climate Change [*Engineers Australia*, 2014].

The rainfall IFD curves in Figure 19 exhibit only slight increases (5 to 15%) in extreme rainfall, and the individual error bands obtained from the BHM overlap considerably. For RCP 8.5, the ARR Interim Guideline suggests a mid-century change factor of $1.05^{2.25} = 1.12$, an increase of 12%. Thus this result is consistent with the mid-century BHM results obtained using WRF driven by the CSIRO Mk. 3.5 GCM under the older SRES A2 scenario.
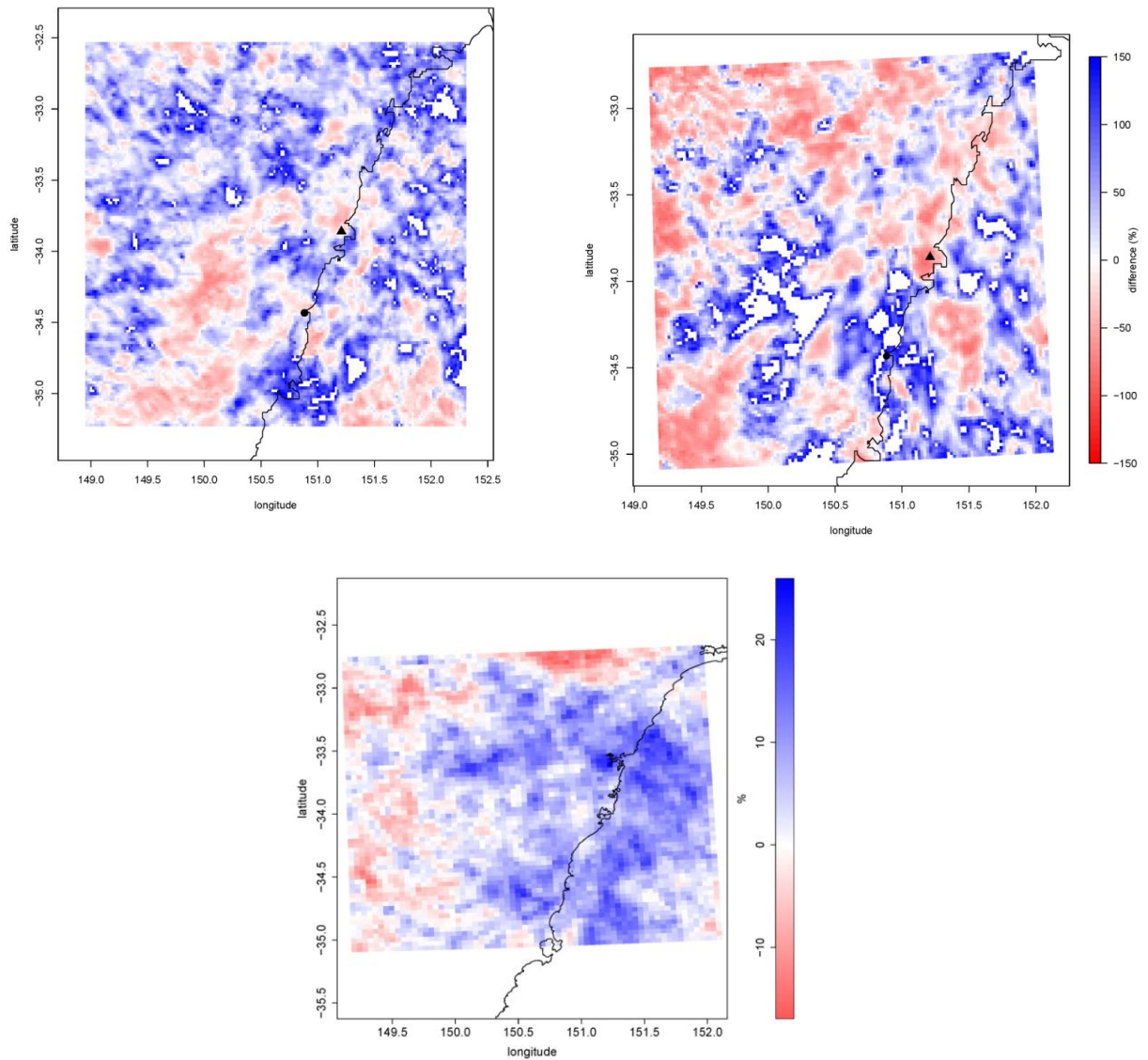
Figure 18. Greater Sydney region – Estimates of impact of climate change on 100-year (1% AEP) return levels of annual maximum 24-hour rainfall: CCAM (2 km), top left; WRF (2 km), top right; BHM integrated with WRF, bottom. Scale is the percent difference between the future period (2040 – 2059) and the baseline period (1990-2009). Contiguous white regions in the top row represent grid cells where the changes (mainly increases) exceeded the range (-150%, +150%). Filled triangle (▲) represents Sydney, filled circle (●) Wollongong.
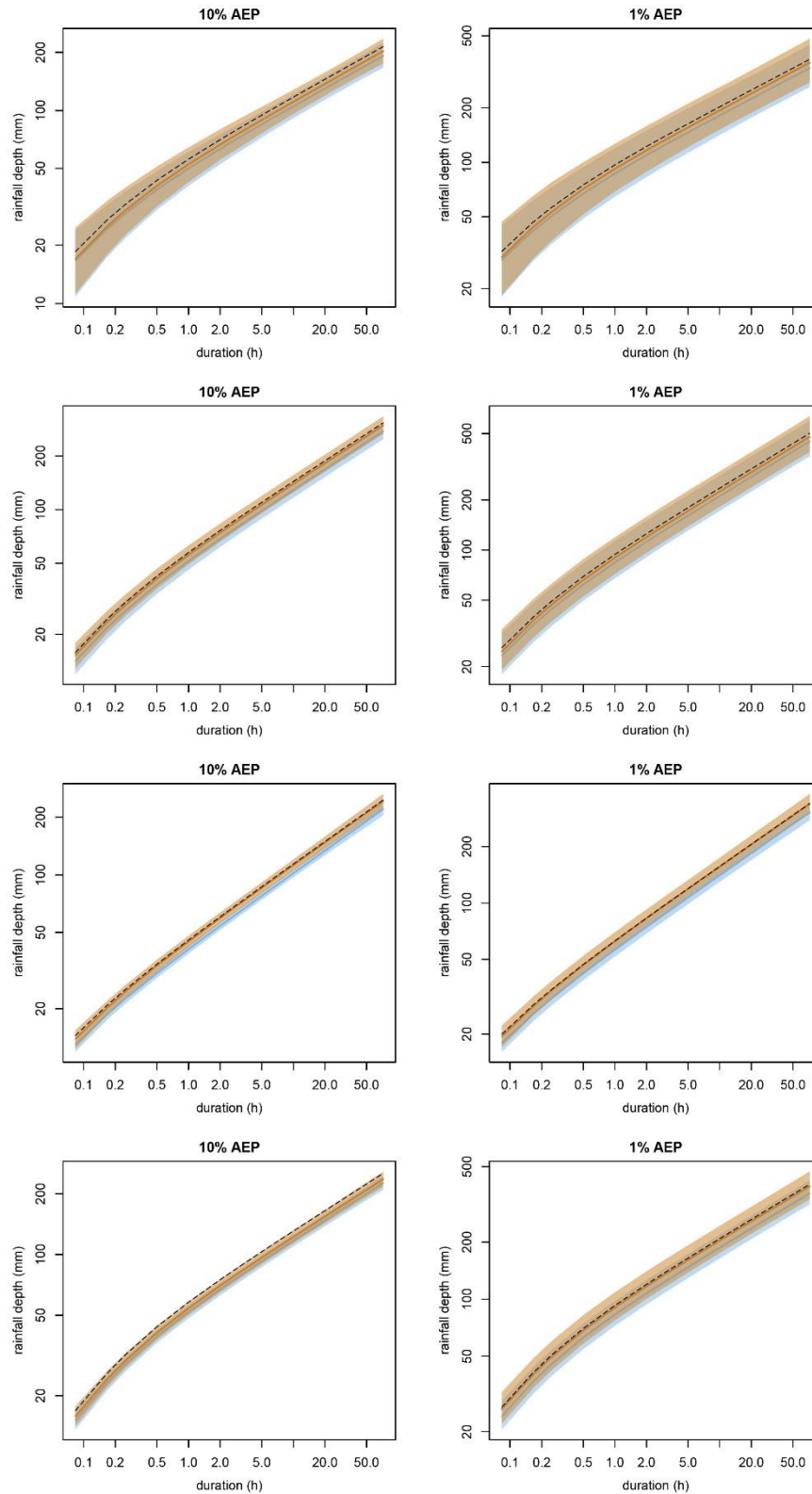
Figure 19. Rainfall IFD curves at (from top to bottom) Congewai, Dapto Bowling Club, Katoomba, and Sydney Airport AMO. Future IFD curves and their 95% uncertainty intervals are shown in orange, baseline curves and associated 95% uncertainty in blue. The dashed black line shows future IFD curves calculated from the ARR Interim Guideline for RCP8.5.
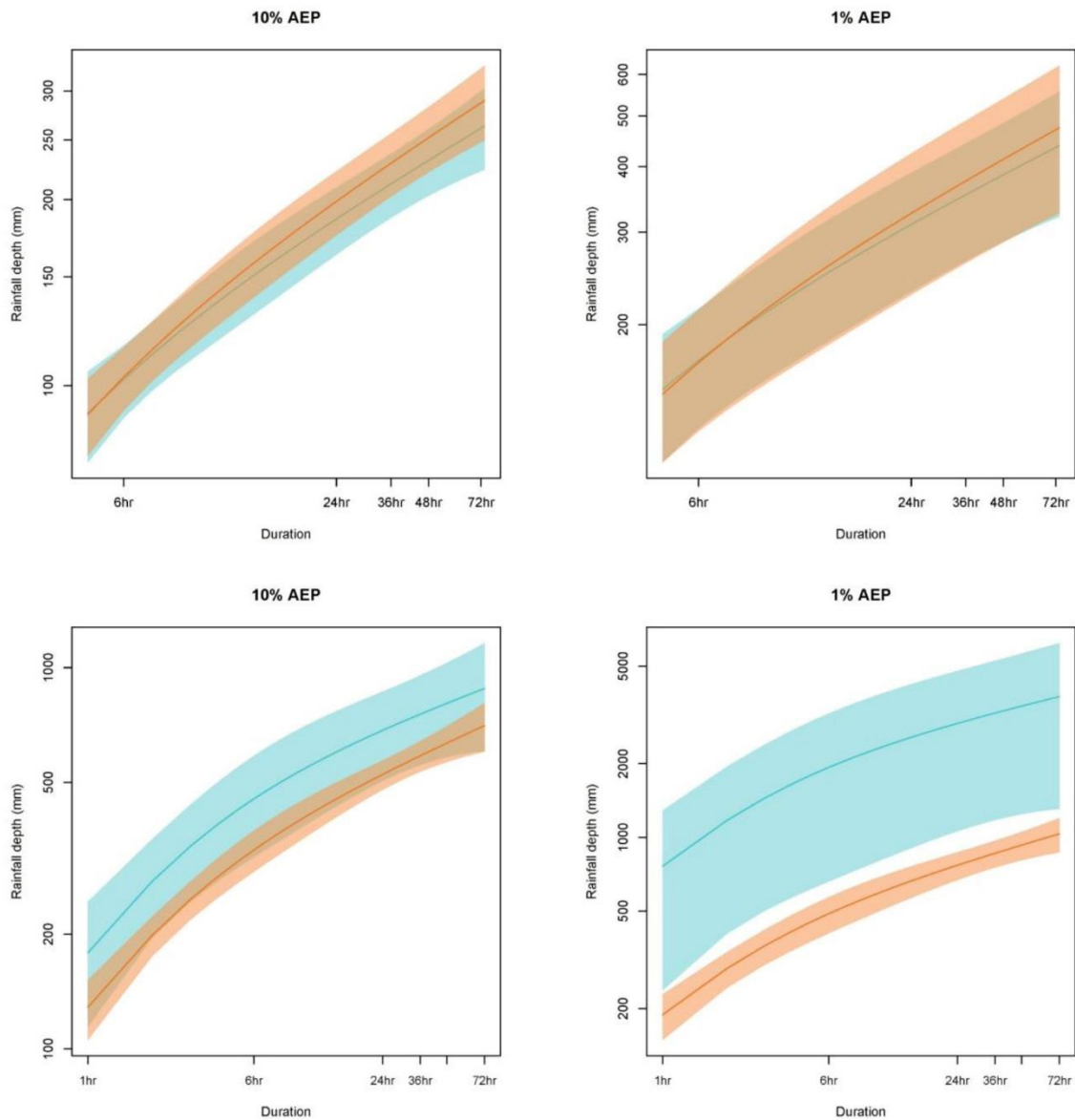
Figure 20. Rainfall IFD curves for Sydney Airport AMO produced from CCAM (top) and WRF (bottom) output. Future IFD curves and their 95% uncertainty intervals are shown in orange, baseline curves and associated 95% uncertainty in blue.

For these same four locations, plots of current and future rainfall IFD curves were produced by considering WRF and CCAM output alone at a horizontal resolution of 2 km. The output from the grid cells nearest to the rainfall stations were used to calculate the IFD curves. As shown in Figure 20, the output from CCAM and WRF are different from one another, and they both considerably overestimate the IFD curves for current climate conditions (cf bottom plot of Figure 19 for Sydney Airport AMO). In addition, WRF projects a large decrease in rainfall depth, but CCAM projects little or no change.

5.2. **South-east Queensland region**

For south-east Queensland, only CCAM results at 2 km and 10 km resolution are available. As for the Greater Sydney region, the projected changes in the 1% AEP 24-hour rainfall can be very large. As shown in Figure 21, CCAM (2 km) projects increases in the Brisbane region and in an arc to the west, whereas it projects a decrease in extreme rainfall near the Sunshine Coast and to the northwest. Spatial patterns at different durations are similar.

For a single location in the South-east Queensland region (Nerang at Gilston Road), Figure 22 shows rainfall IFD curves generated by CCAM (2 km) for baseline and future periods. For an AEP of 10%, there is little difference between current (blue) and future (orange) extremes, whereas for rarer events (1% AEP), future extreme rainfall is much larger than current extreme rainfall.
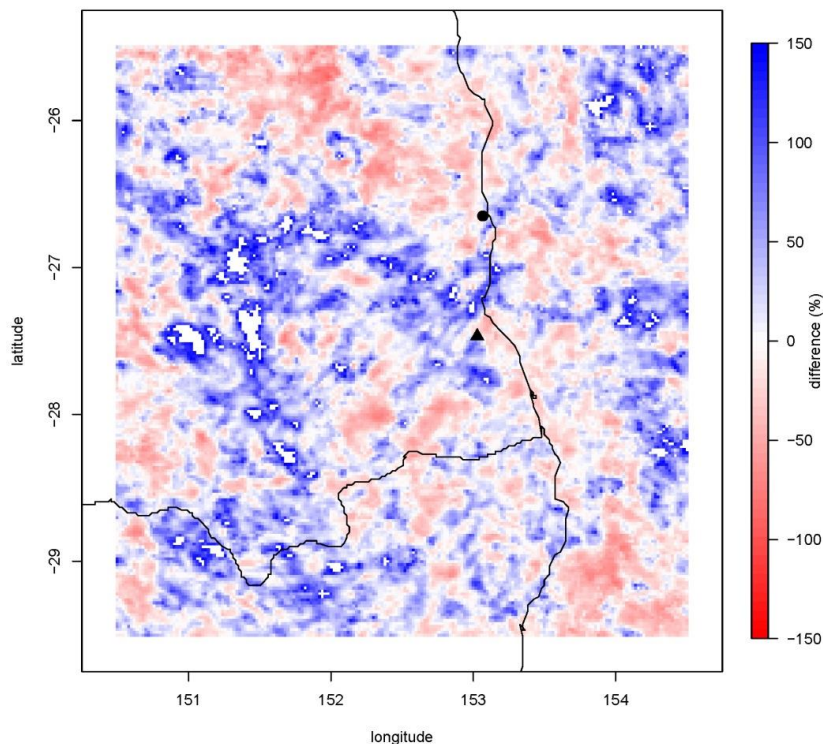
Figure 21. South-east Queensland region – Estimated change from CCAM (2 km) simulation of 100-year (1% AEP) return level of annual maximum 24-hour rainfall between future (2040-2059) and baseline periods. Contiguous white regions represent grid cells where the changes (mainly increases) exceeded the range (-150%, +150%). Filled triangle (▲) represents Brisbane, filled circle (●) Sunshine coast.

As for the Sydney region, it is advisable that the CCAM results should be interpreted over a broad region (e.g. regionally or even domain-averaged) to account for the short climatological periods (20 years) being examined at such fine spatial scales. In this case, although the AEP plots of CCAM output for Nerang Gilston Road (Figure 22) display varying changes from the current climate depending on the AEP chosen, we would again caution against the over-interpretation of projections based upon individual grid points.
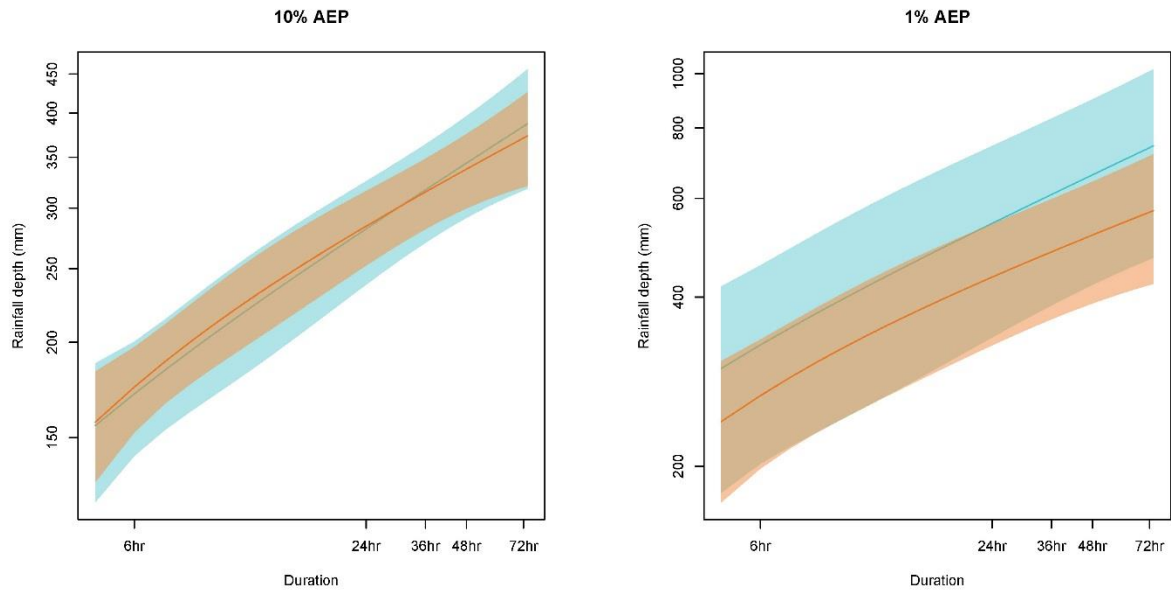
Figure 22. Baseline period (blue) and future (orange) rainfall IFD curves for Nerang Gilston Road calculated from CCAM output (2 km).

6.   Conclusions

6.1.   Key findings

6.1.1. Assessments of change from the instrumental record

- The intensities of observed short duration (< 2 hours) rainfall extremes (annual maxima) have increased in the Greater Sydney region over the period 1966 to 2012. Conversely, rainfall extremes with timescales greater than three hours more likely exhibited a decrease.

- Summer rainfall extremes experienced substantial increases across different durations, while autumn and winter maxima exhibited decreases at all timescales. No apparent changes could be observed during spring for any timescale.

- The annual maxima with short durations were more likely to occur in summer, while the annual maxima with longer durations tended to occur throughout the year. Therefore, the large increases in annual rainfall extremes at short timescales can be largely attributed to the strong upward trends in summer, while the declines in annual maxima with longer durations were attributed by the substantial decreases in autumn and winter.

- Despite differences in methodology between the spatial Bayesian Hierarchical Model (BHM) and the regionalisation approach used to obtain the new design rainfall IFD curves in ARR 2015, the results were found to be remarkably similar. In most cases, the ARR 2015 design rainfall Intensity-Frequency-Duration (IFD) curves lie within the uncertainty bounds produced by the BHM.

- Biases exist in the two observational rainfall datasets (Rainfields and AWAP) used for climate model evaluation. In particular, AWAP estimates are less reliable (i) in regions with steep gradients and for sparse networks, and (ii) in the intensity and frequency of extreme events, which are generally underestimated. Therefore, using more than one observational dataset in similar evaluation studies is recommended, analogous to the use of ensembles in climate model simulation.

6.1.2. Climate model performance and projections

- Reanalysis-driven simulations are typically a better match with observations than GCM-driven simulations, and CSIRO Mk3.5-WRF simulations typically provide a more realistic picture than the ACCESS-CCAM simulations analysed.

- The frequency distributions of observed and simulated rainfall exhibit marked differences. Climate models produce unrealistically high rainfall extremes and while a simple bias correction reduced errors for modelled annual accumulations, it did not remove the tendency for overestimation of rainfall extremes and introduced artefacts.

- For the Greater Sydney region, neither ACCESS-CCAM nor CSIRO Mk3.5-WRF output is able to adequately reproduce the magnitude of extreme rainfall in the selected baseline period (1990 – 2009) although there is indeed some relationship between patterns of extreme rainfall produced by WRF and those exhibited by rainfall station data.

- In its current form, the ACCESS-CCAM combination is less well suited to the development of projections for rainfall IFD curves than the CSIRO Mk3.5-WRF combination. GEV parameters estimated from CCAM output over the Greater Sydney and South-east Queensland regions show little or no relationship to parameters estimated from observed data. This is because the bias is neither consistent across durations nor across the region. Based on the simulations considered, there is little indication that lowering the resolution from 10 km to 2 km leads to an improvement in the simulation of rainfall extremes.

- The CSIRO Mk3.5-WRF combination results suggest significant rainfall depth increases across all durations and all seasons except winter with considerable spatial variations (noting a number of issues with the models ability to reproduce the observed extremes).

- The BHM estimates for future extremes, which only uses the part of the WRF projection that is more reliable, suggests that the 1% AEP for the 24 hour duration event will increase by up to 20% in the Greater Sydney region. To the west, in some parts of the Blue Mountains and beyond, decreases are possible. Projected rainfall depths obtained from the ARR Interim Guideline on Climate Change under RCP8.5 were found to be in agreement with the BHM estimates.

- For south-east Queensland, projected changes in the 1% AEP 24-hour rainfall exhibit large spatial variability and range. The ACCESS-CCAM model combination projects increases in the Brisbane region and in an arc to the west, whereas it projects a decrease in extreme rainfall near the Sunshine Coast and to the northwest. Spatial patterns at different durations are similar. For an AEP of 10%, there is little difference between current and future extremes, whereas for rarer events (1% AEP), future extreme rainfall is much larger than current extreme rainfall.

6.2. Unanswered questions

- Increases in the intensity of observed annual rainfall extremes for short durations are largely associated with upward trends in summer, while the declines in the annual maxima for longer durations are dominated by the decreases in autumn and winter. However, future projections indicate increases in annual maxima for all durations and all seasons, with the exception of winter. This mismatch in observed and modelled seasonal changes in annual maxima requires further investigation.

- Both high resolution models produce significant spatial variability in the rainfall extremes. Whether this is largely associated with the relatively small sample size or with real physical mechanisms (e.g. small scale topographic forcing) remains unknown.

6.3. Key 'learnings'

- Use of regional climate models at convection-permitting resolutions is relatively new. Further development and testing of these model at such resolutions is required to improve overall model performance.

- Conventional bias correction techniques are focused on correcting the mean and low order moments of the rainfall distribution. The BHM offers a more sophisticated approach by relating the parameters of the probability distributions of rainfall station data to those for the extreme rainfall generated by the climate models.

- There are a range of approaches for understanding future changes to rainfall extremes, including empirical studies that focus on analysing historical extreme rainfall patterns, and dynamical modelling studies that use high-resolution climate models to produce future climate projections. Each approach has strengths and weaknesses, highlighting the importance of a multiple line-of-evidence approach in understanding future extremes.

6.4. Recommendations for future research

There is less confidence in climate change projections for rainfall than for temperature change, even on a global scale. Rainfall is likely to exhibit considerable regional variations in its response to the enhanced greenhouse effect, because of its often-complicated interactions with orography and large-scale atmospheric circulation. Our inability to provide definitive answers to questions such as "*What is the likely effect of climate change on design rainfall IFD curves*" partly reflects the difficult nature of climate research, but also indicates that additional research involving more detailed model experiments and data analysis may assist in answering this question. Topics worthy of investigation are:

- **Physical understanding.** The analysis of historical rainfall trends showed that short duration rainfall extremes experienced increases while extreme rainfall with longer durations tended to decline. To better understand the possible changes in rainfall extremes, future research is required to comprehend the physical processes that drive the variations of the extreme rainfall with different durations. It is likely that different meteorological processes dominate the trends during different seasons, and therefore understanding the seasonality of extreme rainfall should be a priority.

- **Design rainfall temporal patterns for short and long duration events.** The analysis only considered the variations in the total rainfall amounts for a given storm burst duration. However, climate change may not only alter the rainfall intensity, but also vary the temporal patterns of the extreme rainfall events. Further research is therefore required to quantify the impact of climate change on these patterns.

- **Structure of spatial dependence.** The spatial dependence structure of the rainfall extremes in the Greater Sydney region was assumed to be constant over time, although it is not clear whether this assumption will remain valid under a changing climate. Hence, it is necessary to carry out research to investigate the possible changes in the spatial dependence structure of the rainfall extremes, thereby offering deeper insight into the impact of climate change to rainfall extremes.

- **Implications for floods.** Antecedent catchment conditions refer to the catchment wetness conditions prior to an extreme rainfall event. Possible changes to the seasonality of may mean that the catchment wetness prior to the flood-producing rainfall event may also change. For example, because of the higher evapotranspiration rates during summer, an increase in the probability that flood-producing extreme rainfall will occur in summer might mean that the extreme rainfall will occur on a drier catchment; this in turn may influence the flood risk that arises because of that event. The interaction between the flood-

producing rainfall event and the antecedent catchment conditions is currently not well understood.

- **Augmentation of the Greater Sydney and south-east Queensland studies.** The climate change projections described in this report are based on a limited number of high-end emissions scenarios, global climate models and regional climate models. Therefore, it is highly likely that the uncertainty in the projections is underestimated and projected regional patterns of change are not robust. Accepted practice would involve the use of a much larger number of host global climate models and a more systematic sampling of regional model uncertainty. The use of a low emissions scenario such as RCP4.5 should also be incorporated into the experimental design.

- **Extension to other geographic domains.** The results of this study focus on the Greater Sydney and to a lesser extent the south-east Queensland regions. Given the different rainfall-generating mechanisms that are likely to apply in other geographic regions, it is necessary to repeat this study for other locations.

- **Climate model improvement.** Better understanding and modelling of the drivers of extreme rainfall from global to regional scales and their responses to greenhouse gases are required. Given the need to simulate changes in rainfall extremes at short time scales (down to sub-hourly), high resolution, convection-permitting models are the tool that embodies our best knowledge of the important physical processes. Current models require further development and testing to improve their performance at these scales and hence their ability to assist in addressing the previous research needs.

- **Bayesian Hierarchical Model improvement.** The current implementation of the spatial BHM incorporates the present and future dynamical downscaling model outputs in two separate steps; first, it estimates the relationships between the GEV parameters from the station data and from the current-climate model output (model fitting), and in a subsequent step, the future-climate model output (GEV parameters) is used to estimate the IFD curves in the future. In both steps, the estimation of GEV parameters from the model output is carried out independently of model fitting. A more rigorous approach would be to construct a model that completely integrates data and dynamical model output, and which carries out estimation, integration of dynamical model output, and estimation of future extreme rainfall within a unified framework.

7. References

Abbs, D., K. McInnes, and T. Rafter (2007), The impact of climate change on extreme rainfall and coastal sea levels over south-east Queensland, Part 2: A high-resolution modelling study of the effect of climate change on the intensity of extreme rainfall events, *Report prepared for the Gold Coast City Council*, CSIRO Division of Marine and Atmospheric Research, Aspendale, Victoria.

Abbs, D., and T. Rafter (2009), Impact of climate variability and climate change on rainfall extremes in western Sydney and surrounding areas: Component 4 – dynamical downscaling, *Report to the Sydney Metro Catchment Management Authority and partners*, CSIRO Climate Adaptation Flagship, Melbourne, Victoria.

Alexander L V. and J.M. Arblaster (2009), Assessing trends in observed and modelled climate extremes over Australia in relation to future projections, *Intl. J. Climatol.*, *29*(3), 417–435, doi:10.1002/joc.1730.

Alexander, L. V., X. Zhang, T. C. Peterson, J. Caesar, B. Gleason, A. M. G. Klein Tank, M. Haylock, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, K. Rupa Kumar, J. Revadekar, G. Griffiths, L. Vincent, D. B. Stephenson, J. Burn, E. Aguilar, M. Brunet, M. Taylor, M. New, P. Zhai, M. Rusticucci, and J. L. Vazquez-Aguirre (2006), Global observed changes in daily climatic extremes of temperature and precipitation, *J. Geophys. Res. Atmos.*, 111(D05101), doi:10.1029/2005JD006290.

Anthes, R. A., E. -Y. Hsie, and Y. -H. Kuo (1987), Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4), NCAR Technical Note NCAR/TN-282+STR, doi:10.5065/D64B2Z90.

Argüeso, D., J. P. Evans, and L. Fita (2013), Precipitation bias correction of very high resolution regional climate models, *Hydrol. Earth Syst. Sci.*, *17*(11), 4379–4388, doi:10.5194/hess-17-4379-2013.

Berg, P., C. Moseley, and J. O. Haerter (2013), Strong increase in convective precipitation in response to higher temperatures, *Nature Geoscience, 6*(3), 181–185, doi:10.1038/ngeo1731.

Bi, D. H., et al. (2013), The ACCESS coupled model: description, control climate and evaluation, *Aust. Meteorol. Ocean. J.*, *63*(1), 41-64.

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V. -M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S. K. Satheesh, S. Sherwood, B. Stevens and X. Y. Zhang (2013), Clouds and Aerosols, Supplementary Material, in *Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Stocker, T. F., D. Qin, G. -K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley. Available from www.climatechange2013.org and www.ipcc.ch.

Chandler, R. E., and S. Bate (2007), Inference for clustered data using the independence loglikelihood, *Biometrika, 94*(1), 167-183.

Cheng L., and A. AghaKouchak (2014), Precipitation intensity-duration-frequency curves for infrastructure design in a changing climate, *Scientific Reports*, *4*, 7093, doi: 10.1038/srep07093.

Coles, S. (2001), *An Introduction to Statistical Modelling of Extreme Values*, Springer-Verlag London Limited, London.

Collins, M., R. Knutti, J. Arblaster, J. -L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A. J. Weaver, and M. Wehner (2013), Long-term climate change: Projections, commitments and irreversibility, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, Cambridge University Press, 1029-1136, doi:10.1017/CBO9781107415324.024.

Davison, A. C., S. A. Padoan, and M. Ribatet (2012), Statistical modelling of spatial extremes, *Statist. Sci.*, *27*(2), 161–186, doi:10.1214/11-STS376.

Dee, D. P., et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*, 553-597, doi:10.1002/qj.828.

Dix, M., et al. (2013), The ACCESS coupled model: documentation of core CMIP5 simulations and initial results, *Aust. Meteorol. Ocean. J.*, *63*(1), 83-99.

Engineers Australia (2014), *Australian Rainfall and Runoff Discussion Paper: Interim Guideline for considering climate change in rainfall and runoff*. Available: http://www.arr.org.au/wp-content/uploads/2013/Projects/Draft_ARR_interim_guidance_Format.pdf.

Evans, J. P., F. J, C. Lee, P. Smith, D. Argüeso, and L. Fita (2014), Design of a regional climate modelling projection ensemble experiment – NARCliM, *Geosci. Model Dev.*, *7*(2), 621–629, doi:10.5194/gmd-7-621-2014.

Evans, J., and M. McCabe (2010), Regional climate simulation over Australia's Murray-Darling basin, *J. Geophys. Res. Atmos., 115*, doi:10.1029/2010JD013816.

Evans, J., and M. McCabe (2013), Effect of model resolution on a regional climate model simulation over southeast Australia, *Climate Res., 56*(2), 131–145, doi:10.3354/cr01151.

Gordon, H., S. O'Farrell, M. Collier, M. Dix, L. Rotstayn, E. Kowalczyk, T. Hirst, and I. Watterson (2010), The CSIRO Mk3.5 Climate Model, *CAWCR Tech. Rep. No. 021*. Available http://www.cawcr.gov.au/publications/technicalreports/CTR_021.pdf.

Green, J., K. Xuereb, F. Johnson, G. Moore, and C. (2012), The revised intensity-frequency-duration (IFD) design rainfall estimates for Australia – An overview, in *Hydrology & Water Resources Symposium 2012*, pages 808 – 815. Engineers Australia, National Committee on Water Engineering, November 2012.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional Frequency Analysis. An Approach Based on L-Moments*, Cambridge University Press, Cambridge.

IPCC (2012), Summary for Policymakers, in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, edited by Field, C. B., V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D. Mastrandrea, K. J. Mach, G. -K. Plattner, S. K. Allen, M. Tignor, and P. M. Midgley. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 1-19.

Jakob, D., D. J. Karoly, and A. Seed (2011), Non-stationarity in daily and sub-daily intense rainfall - Part 1: Sydney, Australia, *Natural Hazards Earth Syst. Sci.*, 11, 2273-2284.

Jenkinson, A. F. (1955), The frequency distribution of the annual maximum (or minimum) values of meteorological elements, *Q. J. R. Meteor. Soc.*, *81*(348), 158-171.

Jones, D. A., W. Wang, and R. Fawcett, R. (2009), High-quality spatial climate data-sets for Australia, *Aust. Meteor. Ocean. J., 58*, 233–248.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts (2012), Realism of rainfall in a very high resolution regional climate model, *J. Clim.*, *25*, 5791–5806, doi:10.1175/JCLI-D-11-00562.1.

Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior (2014), Heavier summer downpours with climate change revealed by weather forecast resolution model, *Nat. Clim. Change*, *4*, 570–576, doi:10.1038/nclimate2258.

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, *Intl. J. Climatol., 33*(10), 2376–2387. doi:10.1002/joc.3588.

Koch S. E., M. desJardins, and P. J. Kocin (1983), An interactive Barnes objective map analysis scheme for use with satellite and conventional data *J. Clim. Appl. Meteor., 22*, 1487-1503.

Koutsoyiannis, D., D. Kozonis, and A. Manetas (1998), A mathematical framework for studying rainfall intensity-duration-frequency relationships, *J. Hydrol.*, *206*, 118–135.

Kuo, C.-C., T. Y. Gan, and M. Gizaw (2015), Potential impact of climate change on intensity duration frequency curves of central Alberta, *Climatic Change*, 1–15, doi:10.1007/s10584-015-1347-9.

Lehmann, E. A., A. Phatak, S. Soltyk, J. Chia, R. Lau, and M. Palmer (2013), Bayesian hierarchical modelling of extremes, in Piantadosi, J., R. S. Anderssen, and J. Boland, editors, *MODSIM2013,*

20th International Congress on Modelling and Simulation, pages 2806–2812. Modelling and Simulation Society of Australia and New Zealand, December 2013. ISBN 978-0-9872143-3-1, www.mssanz.org.au/modsim2013/L12/lehmann.pdf.

Lenderink, G., H. Y. Mok, T. C. Lee, and G. J. Van Oldenborgh (2011), Scaling and trends of hourly precipitation extremes in two different climate zones - Hong Kong and the Netherlands, *Hydrol. Earth Syst. Sci.*, *8*, 4701-4719, doi:10.5194/hess-15-3033-2011.

Liew, S. C., S. V. Raghavan, and S. -Y. Liong (2014), How to construct future IDF curves, under changing climate, for sites with scarce rainfall records?, *Hydrol. Process.*, *28*(8), 3276–3287, doi:10.1002/hyp.9839.

McGregor, J. L. (2005), C-CAM geometric aspects and dynamical formulation, *CSIRO Atmospheric Research, Technical Paper 70*, Aspendale, Victoria.

Min, S. K., X. Zhang, F. Zwiers, and G. C. Hegerl (2011), Human contribution to more-intense precipitation extremes, *Nature*, *470*, 378-381.

Mirhosseini, G., P. Srivastava, and L. Stefanova (2013), The impact of climate change on rainfall Intensity–Duration–Frequency (IDF) curves in Alabama, *Reg. Environ. Change*, *13*(1), 25–33, doi:10.1007/s10113-012-0375-5.

Mirhosseini, G., P. Srivastava, and X. Fang (2014), Developing rainfall intensity-duration-frequency curves for Alabama under future climate scenarios using artificial neural networks, *J. Hydrol. Eng.*, *19*(11), 04014022, doi:10.1061/(ASCE)HE.1943-5584.0000962.

O'Gorman, P. (2015), Precipitation extremes under climate change, *Curr. Clim. Change Rep.*, *1*, doi:10.1007/s40641-015-0009-3.

Peck, A., P. Prodanovic, and S. P. Simonovic (2012), Rainfall intensity duration frequency curves under climate change: City of London, Ontario, Canada. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, *37*, 177–189.

Peter, J., A. Seed, P. Steinle, S. Rennie, and M. Curtis, (2014), A Bayesian methodology for detecting anomalous propagation in radar reflectivity observations. *CAWCR Tech. Rep. No. 077*. Available http://www.cawcr.gov.au/publications/technicalreports/CTR_077.pdf.

Seed, A., E. Duthie, and S. Chumchean (2007), Rainfields: The Australian Bureau of Meteorology system for quantitative precipitation estimation. Proceedings 33rd AMS Conference on Radar Meteorology, Cairns, Australia.

Schmidt, F. (1977), Variable fine mesh in spectral global models, *Beitr. Phys. Atmos.*, *50*, 211-217.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X. –Y. Huang, W. Wang, and J. G. Powers (2008), A description of the advanced research WRF Version 3,

National. Center for Atmospheric Research (NCAR) Tech. Note, June 2008. Available: http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.

Soltyk, S., M. Leonard, A. Phatak, and E. Lehmann (2014), Statistical modelling of rainfall intensity-frequency-duration curves using regional frequency analysis and Bayesian hierarchical modelling, in *Hydrology & Water Resources Symposium 2014*, pages 302–309. Engineers Australia, National Committee on Water Engineering, February 2014.

Srivastav, R. K., A. Schardong, and S. P. Simonovic (2014), Equidistance quantile matching method for updating IDF curves under climate change, *Water Resour. Mgmt*, *28*, 2539–2562.

Thatcher, M., and J. L. McGregor, (2009), Using a scale-selective filter for dynamical downscaling with the conformal cubic atmospheric model, *Mon. Weath. Rev., 137*(6), 1742–1752. doi:10.1175/2008MWR2599.1.

Thompson, G., R. M. Rasmussen, and K. Manning (2004), Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis, *Mon. Wea. Rev.*, *132*, 519-542.

Trenberth, K. E. (2011), Changes in precipitation with climate change, *Climate Res.*, *47*, 123–138.

Wahba, G. and J. Wendelberger (1980), Some new mathematical methods for variational objective analysis using splines and cross validation, *Mon. Weath. Rev., 108*, 1122-1143.

Wang, D., S. C. Hagen, and K. Alizad (2013), Climate change impact and uncertainty analysis of extreme rainfall events in the Apalachicola River basin, Florida, *J. Hydrol.*, *480*, 125–135, doi:10.1016/j.jhydrol.2012.12.015.

Westra, S., and S. A. Sisson (2011), Detection of non-stationarity in precipitation extremes using a max-stable process model, *J. Hydrol.*, *406*, 119-128.

Westra, S., L. V. Alexander, and F. W. Zwiers (2013), Global increasing trends in annual maximum daily precipitation, *J. Climate*, *26*, 3904-3918.

Westra, S., H. J. Fowler, J. P. Evans, L. V. Alexander, P. Berg, F. Johnson, E. J. Kendon, G. Lenderink, and N. M. Roberts (2014), Future changes to the intensity and frequency of short-duration extreme rainfall, *Rev. Geophys.*, *52*(3), 2014RG000464.

Yilmaz, A. G., I. Hossain, and B. J. C. Perera (2014), Effect of climate change and variability on extreme rainfall intensity-frequency-duration relationships: a case study of Melbourne, *Hydrol. Earth Syst. Sci.*, *18*, 4065–4076.

Zheng, F., S. Westra, and Leonard, M. (2015), Opposing local precipitation extremes, *Nature Climate Change: correspondence, 5,* 389-390.

Zheng, F., S. Westra, and S. A. Sisson (2013), Quantifying the dependence between extreme

rainfall and storm surge in the coastal zone, *J. Hydrol.*, *505*, 172-187.

Zhu, J., W. Forsee, R. Schumer, and M. Gautam (2013), Future projections and uncertainty assessment of extreme rainfall intensity in the United States from an ensemble of climate models, *Climatic Change*, *118*(2), 469–485, doi:10.1007/s10584-012-0639-6.

Appendix A Bayesian Hierarchical Model for IFD curves

The approach that we describe here is based on the spatial Bayesian Hierarchical Model (BHM) set out in *Lehmann et al.* [2013]. If we believe that the generalized extreme value (GEV) distribution provides a good description of rainfall extremes at different durations, then changes in the behaviour of extremes under climate change will be reflected in changes in the parameters of the distribution between the baseline period and the future period. This difference will be driven by changes in the behaviour of extreme rainfall generated by dynamical downscaling model output, and we outline briefly how to incorporate this output into the spatial BHM.

We assume that annual maximum rainfall Y can be described by a generalized extreme value distribution which has location $\mu \in (-\infty, +\infty)$, scale $\sigma > 0$, and shape $\xi \in (-\infty, +\infty)$ [*Coles*, 2001], and it can be written as

$$P(Y \leq y) = F(y; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \tag{A.1}$$

where $1 + \xi(y - \mu)/\sigma > 0$. In this work, we re-parameterize the location parameter as [*Koutsoyiannis et al.*, 1998]

$$\tilde{\mu} = \mu / \sigma , \tag{A.2}$$

and also incorporate a parametric relationship between the GEV scale parameter of annual maxima at different accumulation durations and the durations themselves, as described by *Koutsoyiannis et al.* [1998]:

$$\sigma_d = \frac{\sigma d}{(d+\theta)^\eta}, \quad \theta \geq 0, \quad \eta \in [0,1] \tag{A.3}$$

The parameters $\tilde{\mu}$, σ, ξ, θ, and η represent the latent variables in the process layer of the BHM, and this process layer can be written as

$$\begin{aligned}
\tilde{\boldsymbol{\mu}} &= \mathbf{X}_{\tilde{\mu}}^T \boldsymbol{\beta}_{\tilde{\mu}} + \mathcal{P}_{\tilde{\mu}}(\ell, \alpha_{\tilde{\mu}}, \lambda_{\tilde{\mu}}) \\
\log(\boldsymbol{\sigma}) &= \mathbf{X}_{\sigma}^T \boldsymbol{\beta}_{\sigma} + \mathcal{P}_{\sigma}(\ell, \alpha_{\sigma}, \lambda_{\sigma}) \\
\boldsymbol{\xi} &= \mathbf{X}_{\xi}^T \boldsymbol{\beta}_{\xi} + \mathcal{P}_{\xi}(\ell, \alpha_{\xi}, \lambda_{\xi}) \\
\log(\boldsymbol{\theta}) &= \mathbf{X}_{\theta}^T \boldsymbol{\beta}_{\theta} + \mathcal{P}_{\theta}(\ell, \alpha_{\theta}, \lambda_{\theta}) \\
\text{logit}(\boldsymbol{\eta}) &= \mathbf{X}_{\eta}^T \boldsymbol{\beta}_{\eta} + \mathcal{P}_{\eta}(\ell, \alpha_{\eta}, \lambda_{\eta})
\end{aligned} \tag{A.4}$$

where ℓ denotes the spatial locations, $\tilde{\mu} = [\tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_s]^T$ is a vector containing the location parameters at each station $s = 1, 2, \ldots, S$ (other GEV parameter vectors are defined in the same way), and the $\mathbf{X}_{(\cdot)}$ are matrices of covariates (e.g, latitude/longitude, elevation) at each station. For each GEV parameter vector above, the term $\mathcal{P}(\cdot)$ represents a spatially correlated random process with range $\lambda_{(\cdot)}$ and sill $\alpha_{(\cdot)}$ that allows us to borrow strength from neighbouring stations in order to improve estimation. It also allows for spatial interpolation and hence for inference of parameters at ungauged locations.

In the BHM, the covariates contained in the matrices $\mathbf{X}_{(\cdot)}$ are used to model the large scale

variability in the value of a GEV parameter at a particular station, whereas the Gaussian process accounts for the spatially correlated, small scale variability. Following this line of reasoning, we use the GEV parameters estimated from a dynamical downscaling model as a "climatological" covariate. For each station, this covariate consists of the value of the GEV parameter estimated from the dynamical downscaling model output at the grid cell that contains the station. Thus, we would rewrite, for example, the first equation in eqs. (A.4) above as

$$\tilde{\boldsymbol{\mu}} = \beta_{\tilde{\mu},0} \cdot \mathbf{1} + \beta_{\tilde{\mu},1} \cdot \tilde{\boldsymbol{\mu}}_{\mathrm{RCM}} + \mathcal{P}_{\tilde{\mu}}(\ell, \alpha_{\tilde{\mu}}, \lambda_{\tilde{\mu}}) \,, \tag{A.5}$$

where $\tilde{\mu} = [\tilde{\mu}_{\mathrm{RCM},1}, \tilde{\mu}_{\mathrm{RCM},2}, \ldots, \tilde{\mu}_{\mathrm{RCM},S}]^{T}$ is the vector of $\tilde{\mu}$ parameters estimated by the dynamical downscaling model for the grid cells that contain stations $s = 1, 2, \ldots, S$. In the same way, we can rewrite the other equations in the process layer using the corresponding parameters from dynamical downscaling model output at grid cells containing the stations in the observational record.

This re-defined BHM is fitted to available precipitation data using dynamical downscaling model parameter estimates obtained over the same time period. Fitting is carried out by Markov chain Monte Carlo methods. To assess how extremes might change in the future, the GEV parameters for a future climate can then be inferred by temporally interpolating the BHM results using the corresponding dynamical downscaling model parameter estimates obtained for the future climate. This allows us, therefore, to generate IFD curves in the future climate and to compare them to IFD curves from the baseline period. Note that incorporating parameter estimates from a dynamical downscaling model into the process layer allows for an implicit evaluation of how well the model output reproduces observed extremes in the baseline period. If it does so well, we would expect the `regression coefficient' ($\beta_{\tilde{\mu},1}$ in eq. (A.5)) to be significantly different from zero, that is, there would be an approximate linear relationship between the parameter estimates from the data and from the dynamical downscaling model over the same time period.

Appendix B   List of acronyms

ACCESS1.0	Australian Community Climate and Earth System Simulator (ACCESS) coupled model version 1.0
AMO	Australian Meteorological Office
ARR	Australian Rainfall and Runoff (see http://www.arr.org.au/)
AWAP	Australian Water Availability Project
BHM	Bayesian Hierarchical Model
CCAM	CSIRO Conformal Cubic Atmospheric Model
CSIRO	Commonwealth Scientific and Industrial Research Organisation
GCM	Global Climate Model
GEV	Generalised Extreme Value distribution
IFD	Intensity-Frequency-Duration
IPCC	Intergovernmental Panel on Climate Change
NCAR	National Center for Atmospheric Research
NCEP	National Oceanic and Atmospheric Administration's Centers for Environmental Prediction
RCM	Regional Climate Model
RCP	Representative Concentration Pathway (see, e.g., https://www.wmo.int/pages/themes/climate/emission_scenarios.php)
SRES	Intergovernmental Panel on Climate Change Special Report on Emissions Scenarios (https://www.wmo.int/pages/themes/climate/emission_scenarios.php)
WRF	Weather Research and Forecasting model (see http://www.wrf-model.org/index.php)