Climate Services 4 (2016) 13–29

Contents lists available at ScienceDirect

Climate Services

journal homepage: www.elsevier.com/locate/cliser

# The method of producing climate change datasets impacts the resulting policy guidance and chance of mal-adaptation

Marie Ekström [a,*], Michael Grose [b], Craig Heady [b], Sean Turner [c], Jin Teng [a]

[a] CSIRO Land and Water, Black Mountain, Canberra, ACT, Australia
[b] CSIRO Ocean Atmosphere, Aspendale, Vic, Australia
[c] Department of Engineering Systems and Design, Singapore University of Technology and Design, Singapore

ARTICLE INFO

*Article history:*
Received 24 March 2016
Received in revised form 29 August 2016
Accepted 28 September 2016
Available online 18 October 2016

ABSTRACT

Impact, adaptation and vulnerability (IAV) research underpin strategies for adaptation to climate change and help to conceptualise what life may look like in decades to come. Research draws on information from global climate models (GCMs) though typically post-processed into a secondary product with finer resolution through methods of downscaling. Here we consider the production process as a chain of processes leading to an application-ready data set, where each step may have a significant impact on the climate change signal. Through worked examples set in an Australian context we assess the influence of GCM sub-setting, geographic area sub-setting and downscaling method on the regional change signal. Examples demonstrate that choices impact on the final results differently depending on various factors such as application needs, range of uncertainty of the projected variable, amplitude of natural variability, and size of study region. For heat extremes, the choice of emissions scenario is of prime importance, but for a given scenario the method of preparing data can affect the magnitude of the projection by a factor of two or more, strongly affecting the indicated adaptation decision. For our catchment level runoff projections, the choice of emission scenario is less dominant. Rather the method of selecting and producing application-ready datasets is crucial as demonstrated by results with opposing sign of change, raising the real possibility of mal-adaptive decisions. This work illustrates the potential pitfalls when using unwise GCM sub-sampling or the use of a single downscaled product when conducting IAV research.


**Practical Implications**

To explore possible future climates in detail, work in the impact, adaptation and vulnerability (IAV) field takes results from climate models to produce 'application-ready, locally-relevant' datasets that can be used in applied models and analysis. Datasets usually need to have fine spatial resolution and be comparable to observations. The process of producing these datasets involves numerous steps, with important choices at each step. Each choice can strongly affect the results, which can then lead to very different policy guidance. For end users with limited experience of the production of regional projections, it can be very hard to make assessment on the robustness of the information (i.e. is the regional projection physically plausible and credible). There are however a number of steps an end-user can take to critically assess the risks of ending up with a misrepresentative regional projection. These are framed around key uncertainties:

- What emission scenarios is the information based on, and are these appropriate for the context of the study?
- Are you representing the uncertainty in models ability to simulate natural and forced climate variability? This is typically done through considering either a large ensemble of global climate model (GCM) outputs or making a well-informed and representative subsampling. Is a worst case, best case, model consensus case a useful approach?
- If you are using a downscaled data set, are you familiar with the method's capabilities of capturing characteristics of the change signal as simulated by the GCM, or its abilities to add value to the GCM output? We would recommend users of downscaled information to seek information about strengths and limitations of the particular downscaling method applied. This information ought to be provided by the 'producers' of the downscaled dataset.

* Corresponding author.
*E-mail address:* marie.ekstrom@csiro.au (M. Ekström).

http://dx.doi.org/10.1016/j.cliser.2016.09.003
2405-8807/

- Is there a bias in the simulated data relative to the observed climate? If the level of bias is unacceptable to the application then it may be preferable to use a technique of scaling observations, or else employ a bias correction technique.
- Limitations of the applied model. Many models used to study bio-physical impacts (such as runoff models) are optimised based on physical relationships observed in measured variables. Does these relationships hold under climate change conditions or is there risk for method-related biases?

These are central questions to consider when engaging with regional projections. The 'application-ready' dataset must be representative of the current knowledge about future climate change and be aware of what information cannot be gained from using that particular dataset.

Here we examine case studies in the areas of heat impacts to human health and changes to water resources. For each case study we focus on information that is available to stakeholders through national and state wide projections. We find that for heat indices the choice of emission scenario naturally has large influence on projected change, so the choice of scenarios is crucial. The second largest influence for heat indices were the representation of the GCM ensemble in the regional projections, as there is still a fairly broad range of temperature responses to emissions thought plausible (termed the 'sensitivity'). Choices such as the precise geographic region, use of sophisticated downscaling or choice of complex scaling had relatively less influence that these two major choices.

In the water resources case studies, we found the use of complex methods such as statistical and dynamical downscaling compared to simpler methods had a large impact on results, sometimes affecting the sign of projected change. Differences were found particularly in small regions close to environmental features that cause a deviation of the large scale change signal by the GCMs (e.g. catchments along the Australian Alps). However, we also note that some differences are due to persistent characteristics of particular datasets, characteristics that are revealed to the user only through comparisons with other information sources (noting the risk of relying on information from a single downscaling source). As for temperature cases, GCM sub-setting also has an important effect.

Our case studies also demonstrate how the use of some simpler metrics can be sensitive to the natural range of variability of a climate variable. For example, studies using exceedances over a certain threshold can give very different results depending on where the threshold lies relative to the observed natural range of variability. An example being exceedances over a temperature threshold in the coastal tropics, where the natural range of variability is much smaller compared to the temperature range of mid-latitudes or continental climates.

Even with a limited set of case studies demonstrating the use of different climate change information sources, it is evident that under certain circumstances different datasets can provide very different guidance on 'change', and this has a large effect on the subsequent adaptation decisions. We suggest that if there is no obvious reason for why one dataset would be more credible than others, then the study must state that significant uncertainty exists around the regional signal and drawing conclusions from a particular result would not be recommended. However some understanding about the nature of uncertainty associated with a particular regional projection can be gained from considering how well that source represents uncertainties such as those associated with emission scenarios, GCM subsampling, downscaling, bias correction and robustness of the applied model under a non-stationary climate.

## 1. Introduction

Research on climate change impacts to bio-physical and socio-economic systems informs of plausible future change and can aid the development of climate change policy in industry and government. These policies subsequently influence strategic planning on time horizons relevant to climate change (which are variable and region dependant) and help to shape the future behaviour and activities of communities, businesses and governments.

Here we recognise the important albeit challenging role that impact research has in shaping policy guidance (Porter et al., 2015; Schwartz, 2012). This includes the various methods that exist of producing future climate datasets that represent plausible scenarios (Ekström et al., 2015). The production of climate scenarios and follow-on impact analysis involve many choices around method selection, where different choices may lead to advice with harmful, conflicting or even contradictory policy implications. This raises the question of methodological reliability (Smith and Petersen, 2014) of impact analysis for policy response, and the possibility of mal-adaptive responses due to inappropriate or inaccurate information on future change.

Information about climate change suitable for regional impact research typically involve several steps of modelling and analysis, each associated with particular uncertainties (Giorgi, 2008). To fully understand how a particular projection represents different sources of uncertainty require a non-trivial analysis into key elements that make up 'climate risk' or rather the influence of climate change in a particular risk analysis (Ekström et al., 2013; Brown and Wilby, 2012; Ledbetter et al., 2012; Jackson et al., 2010). To represent the range of hypothetical future climate change, it is now common practice to use simulations from global climate models (GCMs) that follow plausible future emissions scenarios (IPCC, 2013; Defra, 2009; CSIRO, 2015). In order to inform appropriate actions and avoid mal-adaptation, a study of climate change impacts should consider all projections that are considered plausible, and account for the various sources of uncertainty implicit in the generation of climate change information (Whetton et al., 2012; Stainforth et al., 2005; Foley, 2010; Conroy et al., 2011).

Assessing climate change impacts to bio-physical or socio-economical systems often require inputting GCM simulations into operational models (and other analyses). In doing so, researchers can account for complex or non-linear responses in the studied system due to expected changes in the climate. These assessments present a great challenge to the research community, as GCM output typically require translation onto finer spatial scales to be implemented in impacts, adaptation and vulnerability (IAV) studies (Fowler et al., 2007; Harris et al., 2014). We might term the climate change projection data used in applied analyses at the regional scale as *application-ready*, *locally-relevant*; reflecting the idea that different applications have different requirements in terms of what climate change information is required and what resolution it should be provided on. However, datasets for IAV analysis can be produced using a range of conceptually different methods, and whilst most of them produce output on finer resolution than the GCM, the aspects of the simulated change signal contained in the data can vary greatly (as discussed in Ekström et al. (2015) and shown in Schmidli et al. (2007)).

Thus, to develop policy guidance material that is scientifically robust from a climate projection point of view, IAV researchers should pose the following question of any dataset used in their analysis: *does it represent the full range of projected climate changes that we have confidence in,* and *does it represent them in a comprehensive and balanced way to address the question of interest for the intended application?* To demonstrate the influence that choice of regional projections has on policy guidance we examine the typical pathway for generation of application ready, locally relevant data sets. We examine the impacts of different methods for creating datasets and illustrate the effect of the choices on the end result with worked examples set in an Australian context, drawing on climate change information from the Australian national climate projections (CSIRO, 2015) and other regionally available projections (Evans et al., 2014).

## 2. Defining the credibility of application ready, locally relevant datasets

We propose that the credibility of a dataset for IAV studies must be judged in terms of the specific application. In this context, 'credibility' can be decomposed into two relevant aspects on which to assess an application-ready dataset, these being; (1) the plausible representation of a climate change signal and (2) the presence of bias compared to observations. Climate change signals include persistent changes in the mean, seasonality, daily variability, timing, periodicity, length of events and inter-annual variability of climate variables. Important aspects of bias include differences in these same statistics compared to the relevant observed dataset. Here we use the term 'bias' to describe any difference between the model outputs and observed data.

The climate change signal is typically explored using an ensemble of GCMs such as the Coupled Model Intercomparison Project phase 5 (CMIP5) of Taylor et al. (2012) run for different future emission scenarios such as the Representative Concentration Pathways (RCPs) of van Vuuren et al. (2011). Projections of variables such as temperature can thus be described in terms of: emission scenario, simulated climate response and natural variability (e.g. IPCC, 2013). Model ensembles can be used to attempt to quantify and illustrate the relative contribution from these key sources of uncertainty and thereby identify the weakest link in the simulation system, i.e. the relevant dominant uncertainty (RDU), which may vary for a specific variable, time horizon and location (Hawkins and Sutton, 2009).

Two key obstacles for implementing GCM output in IAV studies are the spatial resolution of their outputs (around 200 km) and the presence of bias relative to the observed climate. To gain greater spatial resolution for a regional projection, scaling or downscaling methods are used. We define scaling as modifying observed data by a change factor produced by models (also called the delta method, perturbation method, pattern scaling or simple scaling), and downscaling via the running of either a statistical downscaling model (SDM) or a dynamical regional climate model (RCM) to produce fine-resolution output. An appraisal of downscaling methods and a framework for assessing the potential added value brought by downscaling is provided by Ekström et al. (2015) and Maraun et al. (2015) respectively.

Scaling of observed data typically solves the second issue of bias as the bias is removed when generating the scaling factor that is subsequently applied to the baseline climate. Downscaled outputs on the other hand are not necessarily suitable for direct input into applied models and further processing may still be required (Christensen et al., 2008). This may involve using downscaled outputs to inform scaling, or employ bias correction to the downscaled output. Bias correction implies a correction of model output to make its distribution more similar to that of the observed datasets (Argüeso et al., 2013; Bennett et al., 2013; Piani et al., 2010; Teutschbein and Seibert, 2012). Whilst bias-correction is largely a standard procedure to adjust downscaled information it is not an uncontested approach as it assumes bias remains constant over time and has the potential to modify dependencies in space-time and between variables (Ehret et al., 2012). In some instances, it may even alter the climate change signals from the downscaled input (Teng et al., 2015). Nevertheless, whilst bias correction adds to the total uncertainty in climate change information, it can lead to greater distributional detail in the regional change signal and it allows the use of downscaled output in IAV models (Yang et al., 2010).

## 3. Chain of processes: producing application ready, locally relevant datasets

Different strategies are employed to illustrate impacts on communities, infrastructure and the environment due to a warming climate (UNFCCC, 2008). Research that focuses on physical impacts tend to favour a so called 'top-down' approach, which starts with existing climate model outputs and works through to the applied model. Top-down estimates of climate risk can inform policy guidance, but need to take into account the typically large uncertainty associated with the many modelling steps included in the projection. Uncertainty in top-down analysis output can be communicated through showing the spread of results, such as the 10th and 90th percentile of a model ensemble.

An alternative 'bottom up' approach (Wilby and Dessai, 2010) focuses on the decision making context, and analysis begins within the application to determine system specific vulnerabilities to climatic variables (Brown et al., 2012, 2015; Turner et al., 2014). Central to this approach is the focus on the system of interest and its ability to maintain desired performance criteria under different/-multiple climate change scenarios. A related concept is 'robust decision making' (RDM), where 'robust' implies a proposed adaptation modification that allows the 'system' to operate satisfactory under multiple future climates (as predicted by GCMs or uncertainties in the observed historical climate)(Lempert and Kalra, 2011; Weaver et al., 2013). The concept of RDM has been frequently used to address problems with 'deep' uncertainty, the term deep reflecting unresolvable lack of knowledge or fundamental disagreement amongst researchers (Lempert et al., 2006). For these reasons, the bottom-up view has become a favoured approach for adaption and vulnerability research, where typically large uncertainties associated with outputs from 'top-down' are difficult to reconcile with the urgency to develop strategies for adaption actions (UNFCCC, 2008; Mastrandrea et al., 2010).

With either approach there is often a need to produce application-ready locally-relevant datasets that contain climate change information; though its implementation in the study is fundamentally different depending on the approach taken (Ekström et al., 2013).

In producing application-ready datasets for IAV studies, the first and broadest question to assess is whether the projections from a set of GCMs using hypothetical scenarios are required for the application. If there is no need to link results to specific emission scenarios, straightforward sensitivity testing may be adequate to assess system vulnerabilities to a warming climate. For example, if there is no current knowledge about the sensitivity of a plant, animal or system to changes in mean temperature, a sensitivity study of changes of 1 °C, 2 °C and so forth is useful before consulting temperature projections. For more complex assessments, i.e. an interest in

**Table 1**
Processes for producing application-ready and locally-relevant datasets, showing stages as: discrete processes (column 2), the main issues involved in each step (column 3), key questions regarding the credibility of the process (column 4).

| | Process stages in application ready datasets | Issues for the user to consider | Key credibility issue |
|---|---|---|---|
| 1.<br>2. | Use of emission scenarios<br>Global climate response by GCMs | The three main sources of uncertainty are represented here: future emissions, model ability to simulate climate response to changing emissions, model ability to simulate natural climate variability. | Consider plausibility of change signal. Consider if the plausible range is represented |
| 3 | Sub-setting GCMs | Choices around selection of scenarios and models will influence how well these sources of uncertainty are represented in the study. Issues to consider if selecting models are:<br>• Model skill<br>• Reflecting simulated model range (does your sample reflect the full range of possibilities?) | |
| 4.a | Regionalise projections by:<br>Downscaling Scaling | Methods of different complexity can generate finer resolution projections over a specified region or location (downscaling). Are you familiar with strengths weaknesses of the method used to generate your dataset? | Consider added value of choosing downscaling (e.g., plausibility of change signal, lower bias).<br>Representativeness of range of change |
| 4.b | No Bias Correction Bias Correction | Methods other than scaling are likely to have output with a bias relative to observed data (particularly output from dynamical models). If level of bias is unacceptable, bias correction (BC) of the regional projection is needed. | |
| 5. | Applied Model | If the applied model is typically used with observed data, issues to consider are:<br>• Sensitivity to biased data<br>• Does assumptions about model parameters hold under a non-stationary climate? | Consider plausibility of change, level of bias relative to observed data and representation of key uncertainties in the application-ready data used with the Applied Model. |

change across multiple variables or change with a temporal/spatial dimension is well served by the use of application data based on physical consistent model simulations.

Table 1 outlines a pathway of decisions and issues a user need to consider when using GCM derived application ready data. The path starts with choices around GCMs and emissions and ends with the input into the applied model. The applied model can also be viewed as a vulnerability threshold (i.e. threshold relating to a level in e.g. magnitude/frequency/duration beyond which a system becomes vulnerable) of a bottom-up approach.

The initial issue to consider is the representation of core uncertainties in GCM output, *viz.* uncertainty stemming from natural variability, emission scenarios, uncertainty of climate processes (e.g. change to atmospheric circulation and physical state responses). To what degree does the inter-model range of GCMs represent a plausible range of these uncertainties? Also, model dependence and model evaluation are very practical concerns regarding the use of an ensemble of GCMs. Is it meaningful to consider rejection or weighting models depending on their performance and independence (Evans et al., 2013; Masson and Knutti, 2011; Sanderson et al., 2015). These are the generic concerns when using an ensemble of GCMs as a tool for examining future climate change.

In some instances, it may be necessary to select a sub sample of GCMs. For example, downscaling may be required and if the desired method requires significant time or computing resources, downscaling can only make use of a limited sample of GCMs (there may also be limitations imposed due to requirement of specific input variables or temporal resolution). Sub-setting introduces the questions of whether the sub-sample is representative of the full set, whether the models chosen have unacceptable biases for the application of interest and also whether the selected models are overly dependent due to similarities in model code (Knutti et al., 2010; Masson and Knutti, 2011). The choice of GCMs sets an important context for the rest of the process, and if a subset of GCMs is particularly unrepresentative then the final datasets will not be usable in terms of representing the range of plausible climate changes (or at least the plausible range that is indicated by the entire set of GCMs).

Once a set of GCM inputs are selected, many IAV studies face two major choices (Table 1, column 4a), to downscale or use simple scaling. Simple scaling by the mean can produce bias-free outputs, but can only express a change in the mean and not in variance, sequencing or duration of events. Downscaling aims to produce 'added value' in the climate change signal and offers the possibility of containing lower biases than the host GCMs. However, the downscaling may not reveal any robust regional detail and the downscaling model itself introduces new method specific effects of its own to the projection (Grose et al., 2015) e.g. due to a parameterisation choice. One approach to deal with this issue is to examine the outputs of numerous downscaling methods in a 'matrix' of GCMs and downscaling (Kendon et al., 2010; Mearns et al., 2013; Maraun et al., 2015). If only one downscaling method is used, then this may represent a skewed sample that must be acknowledged in any derived products or guidance.

If downscaling is used and a robust regional pattern of change is revealed, then outputs can either be used to produce appropriate scaling factors or bias corrected to produce application ready, locally relevant data. A key question then is whether the bias is acceptably small to allow it to be corrected. This depends on the application the data will be used for, since different applications have different data needs and different sensitivity to bias. Key questions of the output are whether the bias is satisfactorily corrected and if the climate change signal has been affected (Bennett et al., 2013).

The type of applied analysis used should also be considered. Running an applied model with downscaled outputs (even after bias correction) may not illustrate a plausible response to changes in the climate. This may be because the model may not show a plausible response in a new climate outside that it was calibrated to (Chen et al., 2011). Split-sample modelling can be used to assess how the IAV model performs when calibrated and applied to climatologically different periods (e.g. Vaze et al. (2010b)), assuming the observed data permits such analysis. Alternatively the model may have an exceedingly small tolerance for bias (e.g., due to over-fitting).

## 4. Illustrative case studies, data and methods

To demonstrate the sensitivities of selecting an appropriate data set relative to application requirements we have selected four case studies each with different regional projection dataset requirements (Table 2). All case studies are set in Australia and

**Table 2**
A summary of comparisons enabled through the example case studies in Sections 4.2 and 4.3.

| Study | Purpose | Methods compared | Key Message |
|---|---|---|---|
| *Heat stress: Melbourne (days over 42 °C) and Darwin (wet bulb temperature > 35 °C) case studies* | | | |
| 1 | Influence of emission scenario<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099) | Mean scaling factor for southern Australia super cluster imposed on Melbourne/ Darwin ACORN-SAT station record following:<br>1. RCP4.5<br>2. RCP 8.5 | Choice of RCP has large influence on projected change due to the strong relationship between increase in greenhouse gas concentration and heat increase.<br>Second largest influence is by GCM sub selection, reflecting the different degree of warming in models under similar greenhouse gas concentrations.<br>For these experiments choice of downscaling, choice of region and mean or quantile scaling appear to have lesser influence on results.<br>For regions where there is a narrow range in natural variability, studies using exceedances over a certain threshold can give very different results depending on where the threshold lies relative to the observed natural range of variability (see our Darwin case study). |
| 2 | Influence of using scaling factor from differently sized areas.<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099)<br>- RCP8.5 | Mean scaling factor imposed on Melbourne/Darwin ACORN-SAT station record based on model data from:<br>1. Super clusters: Southern Australia/Northern Australia super cluster<br>2. Sub-clusters: Southern Slopes Victoria West/Monsoonal North west | |
| 3 | Influence of sub-selecting GCMs.<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099)<br>- Common emission scenario (RCP8.5) | Mean scaling factor for Southern Slopes Victoria West/Monsoonal North West derived from:<br>1. All considered CMIP5 GCMs (∼40)<br>2. 8 selected CMIP5 GCMs<br>3. 8 CMIP5 GCMs with coolest temperature projection (low)<br>4. 8 CMIP5 GCMs with warmest temperature projection (high) | |
| 4 | Influence of change signal in complex downscaling choice.<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099)<br>- Common emission scenario (RCP8.5) | Mean scaling factor for Southern Slopes Victoria West/ Monsoonal North West derived from:<br>1. All considered CMIP5 GCMs (∼40)<br>2. SDM (22 GCMs downscaled)<br>3. CCAM (6 GCMs downscaled) | |
| 5. | Influence of scaling method in application-ready dataset.<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099)<br>- Common emission scenario (RCP8.5)<br>- Common set of 8 GCMs | Applied to the grid cell nearest to Melbourne for 8 selected CMIP5 GCMs:<br>1. Mean scaling<br>2. Quantile scaling | |
| *Water resources: Rainfall-Runoff – Catchments on Murray, Genoa and Avoca in southeast Australia* | | | |
| 1 | Influence on change signal due to different information sources.<br>- Baseline 1986–2005 for all but NARCliM (1990–2009)<br>- Future time horizon 2080–2099 for all but NARCliM (2060–2079)<br>- Emission scenario is RCP8.5 for all but NARCliM (SRES A2) | Annual and seasonal rainfall change (%) for each catchment[1] based on data from:<br>1. All considered CMIP5 GCMs (∼40)<br>2. 8 selected CMIP5 GCMs<br>3. CCAM<br>4. SDM<br>5. NARCliM | Downscaling can have a large influence on the change signal, particularly in small regions close to environmental features that cause a deviation of the large scale change signal by the GCMs (e.g. catchments along the Australian Alps).Downscaling also can give different results away from these gradients, such as consistently drier projections from statistical downscaling in the Avoca catchment (see below). The effect of mean scaling compared to quantile scaling is small compared to the choice of model inputWe note that in circumstances when different datasets provide very different guidance on 'change' and there is no obvious reason for why one dataset would be more credible than others, significant uncertainty exist around the regional signal and drawing conclusions from a particular result would not be recommended |
| 2 | Influence on change signal due to different scaling methods<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099)<br>- Common emission scenario RCP8.5 | Annual and seasonal runoff change (%) for each catchment based on application-ready data based on:<br>1. Mean scaling of rainfall and APET<br>2. Quantile scaling of rainfall, mean scaling of APET | |
| *Water resources: Water supply – urban bulk water supply system in Victoria* | | | |
| | Influence on change signal due to information source.<br>- Common baseline (1986–2005)<br>- Common future time horizon (2080–2099)<br>- Common emission scenario (RCP8.5) | Mean change in mean annual temperature (MAT) and precipitation (MAP) applied to AWAP observed data for catchment area[1] according to:<br>1. All considered CMIP5 GCMs (∼40)<br>2. 8 selected CMIP5 GCMs<br>3. SDM (21 models)<br>4. CCAM (6 models) | Same conclusions as for rainfall-runoff. The choice of rainfall projections has a large effect on the results, including the choice to subset or downscale GCM outputs. Temperature projections are more consistent, so the choice of input has less effect than rainfall. |
| 2 | Influence on change signal due to information source.<br>- Baseline 1986–2005 for all but NARCliM (1990–2009)<br>- Future time horizon 2080–2099 for all but NARCliM (2060–2079)<br>- Emission scenario is RCP8.5 for all but NARCliM (SRES A2) | Annual and seasonal rainfall change (%) for catchment area[1] based on data from:<br>1. All considered CMIP5 GCMs (∼40)<br>2. 8 selected CMIP5 GCMs<br>3. CCAM (6 models)<br>4. SDM (21 models)<br>5. NARCliM (12 models) | |

[1] Catchments and catchment area for water resource studies are mapped in Fig. 4.

represent applications with a heat or water resource focus, demonstrating the use of projections of mainly rainfall and temperature. For illustrative purposes our case studies are simple examples that can be easily replicated, but that all have particular requirements and restrictions.

Amongst the case studies, there are some discrepancies in terms of regional focus and data use. This reflects our intent to give examples where climate change impacts are more relevant (or potentially larger) than elsewhere, or because downscaling is shown to be particularly meaningful in the selected region. As a consequence there is some variation in datasets used for particular applications due to the uneven coverage of available downscaled data sets across the Australian continent.

### 4.1. Climate change projection datasets

The core GCM data set used here are outputs of about 40 CMIP5 GCMs underpinning the new climate change projections for Australia (CSIRO, 2015), available as a change signal for different agglomerations of natural resource management regions, referred to as 'clusters' (Fig. 1). Numbers of GCMs vary somewhat for different variables and RCPs (see Table 3.3.1 and 3.3.2 in CSIRO (2015)). For monthly temperature and rainfall under RCP 4.5 38 and 37 models were use respectively and for RCP8.5 38 and 39 models respectively. These projections are accompanied by a suite of datasets intended for use in applied projects (Webb et al., 2015), so called 'application-ready' data that are also utilised here (Table 2). These datasets are created through scaling of 30 years of observed data (1981–2010), where temperature and rainfall are from the Australian Water Availability Project (AWAP, Jones et al. (2009)) and areal potential evaporation (APET) data calculated from the AWAP variables using Morton's method (Chiew et al., 2009b). All application-ready datasets are at 5 km resolution across Australia.

The change signal for the application-ready dataset is derived from output data of 8 CMIP5 GCMs, representative of the larger CMIP5 ensemble: ACCESS1.0, CESM1-CAM5, CNRM-CM5, GDFL-ESM2M, HadGEM2-CC, CanESM2, MIROC5, NorESM1-M (motivation for model selection is given in Box 9.2 of CSIRO (2015)), and calculated as the difference between the 20-year baseline period 1986–2005 and relevant future period (e.g. 2080–2099). To implement the change factor on a similar resolution to that of the observed data, the coarse resolution change factors (about 200 km resolution) were bi-linearly interpolated to the 5 km resolution grid of AWAP. Further, for some variables we expect different change across different parts of the distribution of a variable (e.g. change in high values are different to mean and low values), thus scaling the observed series only by a mean change is not appropriate. For these variables, e.g. rainfall, the application-ready data is scaled using a quantile scaling approach, allowing different scaling factors to be applied to different parts of the distribution (Webb et al., 2015).

The national projections primarily consider the change signal of the full CMIP5 ensemble, but when appropriate, it also considers the change signal of two downscaled products. These are simulations of the variable resolution GCM Conformal Cubic Atmospheric Model (CCAM; McGregor and Dix (2008)) and outputs from the statistical downscaling model (SDM) developed at the Australian Bureau of Meteorology (Teng et al., 2012; Timbal et al., 2009); the latter providing only three variables, daily maximum and minimum temperature and rainfall.

The CCAM model takes the climate change signal from its host GCM model through incorporating the GCM sea surface temperature in its simulation. It has a stretchable grid across the globe, with outputs at a regular 50 km resolution globally. The SDM draws on a statistical relationship to find analogues between current observed and future modelled large-scale patterns. Having

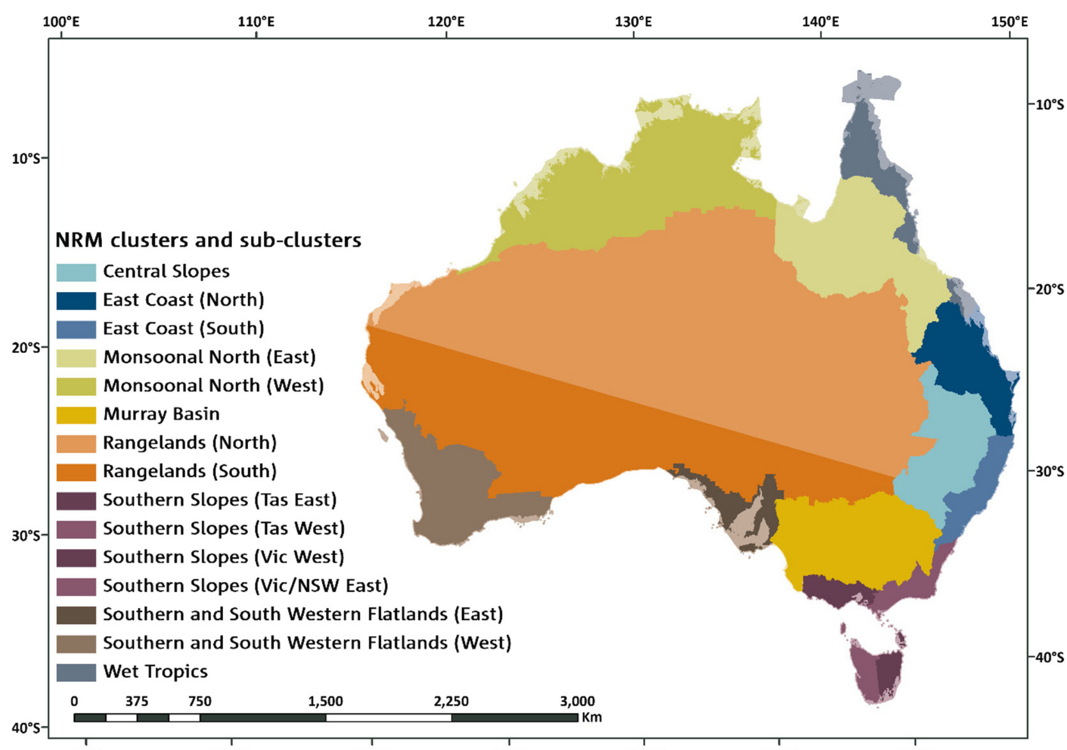


**Fig. 1.** Natural Resource Management cluster regions in Australia for which climate change projections are derived. There are 8 clusters in total, with the majority split into sub-clusters to better represent spatial variation in the climate change signal. For nation-wide projection summaries, larger 'super' clusters are used, these are: Rangelands (as depicted here), Northern Australia (agglomeration of Monsoonal North and Wet Tropics), Southern Australia (Southern and South West Flatlands and Southern Slopes), and Eastern Australia (Central Slopes and East Coast). Reprint from Fig. 7.2.13 in CSIRO (2015) (http://www.climatechangeinaustralia.gov.au/media/ccia/2.1.5/cms_page_media/178/TR_Figure2.2.png) Reproduced by permission of CSIRO Australia, © CSIRO.

found a match in analogues, the method uses the local observed surface variables (from 5 km gridded AWAP data) associated with the best matched analogue of the historical pool (Teng et al., 2012); thus the SDM is available at a 5 km resolution. We note that whilst both CCAM and SDM are available for all of Australia, the latter was implemented across regions to ensure a regionally appropriate selection of atmospheric variables as analogues. For this reason, the SDM is not a spatially homogenous field, which may be an issue to consider for users who are interested in regions that straddle downscaling boundaries. The techniques downscale 6 and 22 GCMs respectively (see downscaled GCMs in Table 3.3.1 in CSIRO (2015)). We note that neither CCAM nor SDM has been bias corrected in this study, and are used as input to scaling methods (decision 4b in Table 1 is not covered in this study).

In addition to resources available through the national projections listed above, we also consider a regional projection commissioned by the New South Wales State government, the New South Wales and Australian Capital Territory Regional Climate (NARCliM) projections (Evans et al., 2014),. This projection product is based on dynamical downscaling using the Weather Research and Forecasting (WRF) limited area model (LAM) with coverage across southeast Australia at 10 km resolution. In this setup, information about moisture, temperature and movement of the atmosphere in the GCM is feed to the LAM at regular time intervals through lateral boundary files. The NARCliM projections contain 12 models simulations, where four different GCMs (MIROC, ECHAM, CCCMA and CSIRO Mk3.0) are combined with three different setups of the WRF model (Skamarock and Klemp, 2008). These regional projections differ to those of the national product in three relevant aspects, (i) they were downscaled from the previous generation of GCMs (CMIP3, Meehl et al. (2007)) that use (ii) emission information from the A2 scenario of the Special Report on Emissions Scenarios (SRES, Nakicenovic et al. (2000)) (iii) and data is available for the baseline period 1990–2009 and future periods 2020–2039 and 2060–2079.

Despite differences, it is meaningful to include the NARCliM dataset in our comparisons when possible (depending on application requirements and geographical coverage), as it is promoted by the NSW state and the ACT. Further, it adds information about range of uncertainty in downscaling methods, which is crucial information for a region where there are very few readily available downscaled datasets. The following case studies involve comparing these datasets in general terms as projections under high emissions for late in the century, noting the difference in specifics (Table 2).

### 4.2. Heat stress examples

With regard to heat, we consider two simple examples of heat stress to illustrate the effect of different choices in preparing datasets for this purpose. First we look at the incidence of days above 42 °C in Melbourne the State Capital of Victoria (37.8136°S, 144.9631°E), a threshold at which certain birds and bats can die (Welbergen et al., 2008). For this example scaling factors from the national projections datasets are used to scale the daily station record from Melbourne from the Australian Climate Observation Network – Surface Air Temperature (ACORN-SAT) of Trewin (2013). The scaling factors are available for differently sized regions and from different model outputs. Our example tests the sensitivity in using scaling factors for the large Southern Australian super cluster and also that of the smaller south west Victorian sub-cluster (Fig. 1) (CSIRO, 2015), for the full set of CMIP5 GCMs, subsets of CMIP5 GCMs and downscaled outputs (Table 2). Mean scaling entails scaling the station record by the mean change in temperature from each model, quantile scaling uses a separate scaling of each month applied to each decile of the station record distribution (percentile bins 0–10, 11–20 etc.) with a secondary scaling to ensure that the monthly mean change is consistent with projected mean change.

The second case study illustrates similar comparisons for dataset preparation as for the Melbourne case study but for a somewhat more complex heat stress index in the tropics (Table 2). From a 'human comfort' perspective, the joint change of high temperature and humidity is of interest as increasing humidity has

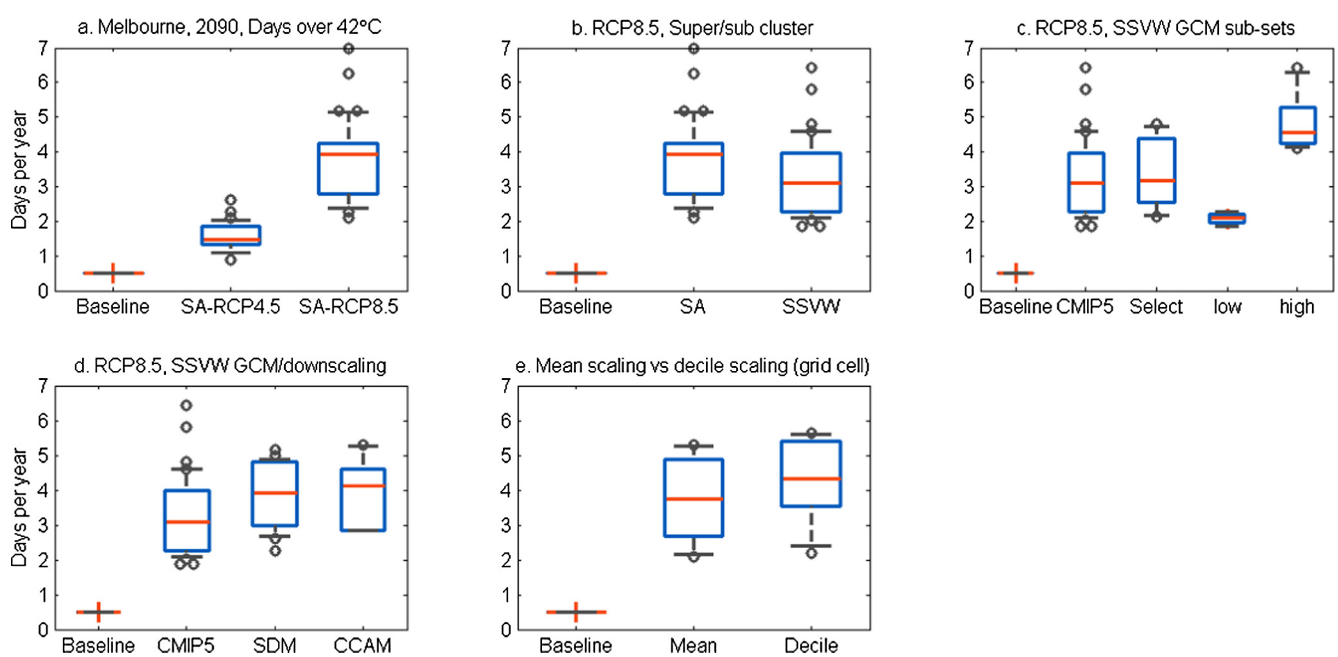


**Fig. 2.** Incidence of days over 42 °C in Melbourne in 1985–2005 (baseline) and 2090 under RCP8.5 calculated using different methods of scaling of the Melbourne ACORN-SAT station data series: (a) mean scaling factor from CMIP5 for Southern Australia (SA) super-cluster under RCP4.5 compared to RCP8.5, (b) mean scaling factor for CMIP5 for SA super-cluster and the Southern Slopes Victoria West (SSVW) sub-cluster, (c) mean scaling factor for SSVW in CMIP5, the 8 selected CMIP5 models, the 8 models showing the coolest temperature projection (low) and the eight with the hottest (high), (d) mean scaling factors for SSVW for CMIP5, BOM-SDM and CCAM downscaling, (e) mean and quantile scaling for the nearest GCM grid cell for the eight selected CMIP5 models (under RCP8.5).

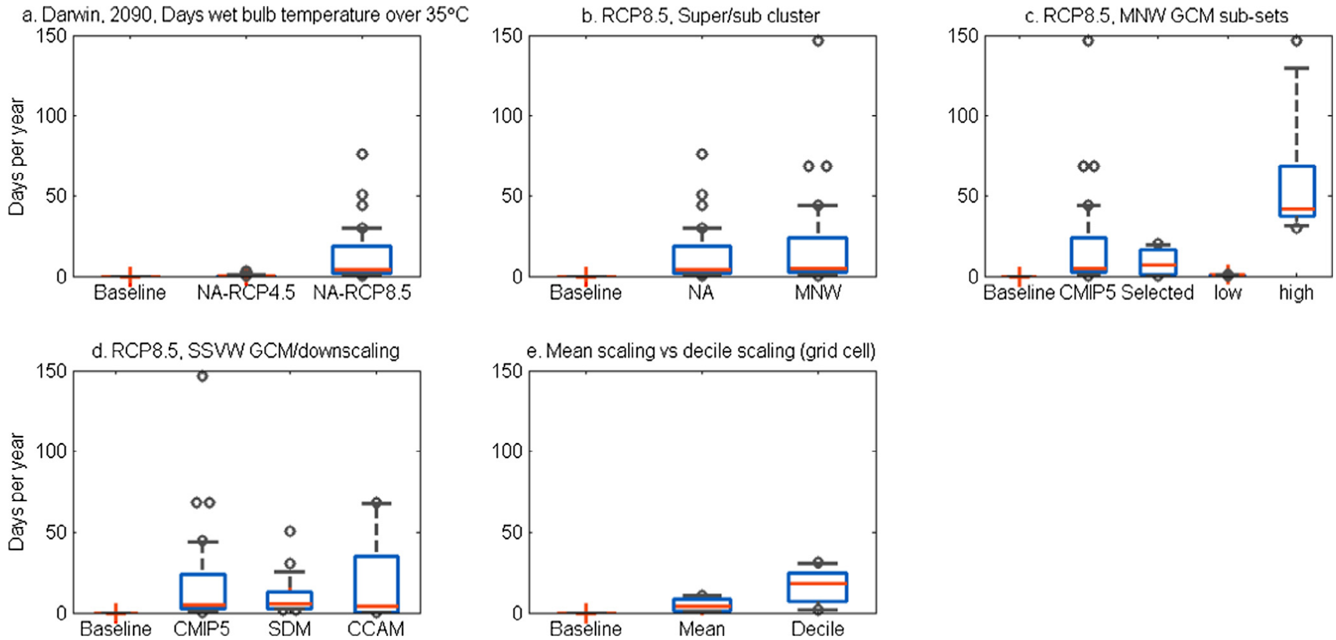


**Fig. 3.** As for Fig. 3 but in Darwin and for the number of days over 39.2 °C (approximating wet bulb temperature of 35 °C).

aggravating physiological effects during extreme temperature events. Other factors are also relevant, such as radiation, wind and behavioural factors such as metabolic rate and clothing (Epstein and Moran, 2006). However, most indices focus on changes to temperature and humidity (e.g. heat index, humidex, temperature humidity index (THI)). Here, we examine days exceeding a threshold wet bulb temperature of 35 °C, where human heat stress becomes 'extreme'. In January, Darwin in the tropical north of Australia has a daily mean relative humidity of 75.5% (mean relative humidity of 81% at 9am and 70% at 3 pm) and mean sea level pressure of 1007 hPa (Bureau of Meteorology). Projected change in relative humidity is <5% (CSIRO, 2015) and also relatively small in mean sea level pressure, so we take the climatological value of these variables and assume they remain constant for this simple example. This means that an air temperature of 39.2 °C gives a simplified wet bulb temperature of 35 °C. We note that our approach is simple and other effects may need consideration if applied in a real-world application. For example, in urban areas (not resolved by GCMs) the heat island effect will worsen the heat stress effect (Oleson et al., 2015).

### 4.3. Water sector examples

Two examples are used to demonstrate the impact of choosing different climate change information in the water sector (Table 2). Southeast Australia includes complex topography, so there is an expectation that downscaled information will add value to the regional climate change signal (Fig. 4). The first example looks at assessments of future change on runoff for three catchments in this region using a top-down approach. Our second example uses a bottom-up style approach and looks at climate change relative to decision thresholds for a complex urban water resource system in the same region (Fig. 2). These examples illustrate different approaches of including climate change information in IAV applications. The water resource example builds on the work in Turner et al. (2014) (hereafter T14), wherein the authors created a functional relationship between thresholds (related to system management) and mean annual temperature and rainfall for the relevant, dominant catchment area. Having established such a relationship, climate change information is used to examine under which emission scenarios and time horizons action thresholds are likely to be exceeded.

In comparison to the straightforward heat stress case studies, the runoff examples require some additional analysis and impact modelling. The essential details for each example are outlined below and in Table 2.

#### 4.3.1. Runoff

Our runoff example features three catchments in the state of Victoria on the rivers: Murray, Avoca and Genoa (Fig. 2). These catchments differ in size and geographical positioning, the largest being on the Avoca river at Coonooer located on the western slopes of the Great Dividing Range in western Victoria (2682 km$^2$), followed by the higher altitude catchment on the Murray river at Biggara (1257 km$^2$) and the on the Genoa river at The Gorge to the east of the Great Dividing Range (844 km$^2$). We examine change under a high emissions scenario (RCP8.5 for CMIP5 outputs, A2 for NARCliM) to the end of the century (or as late in the century that the model outputs cover) as a demonstration of a strong climate change signal (Table 2).

Runoff projections typically draw on outputs from empirically-fitted runoff models using daily series of rainfall and areal potential evapo-transpiration (APET). But there are some methods that use simpler empirical relationships (e.g. water balance equations such as that described by the Budyko curve (Budyko, 1974)). For our region of interest, a first order assessment of impact can be gained from considering the rainfall projections themselves, as rainfall is the variable with strongest predictive power in the rainfall-runoff model in this region. Therefore, by considering projected changes of rainfall, it is possible to use a 'rules-of-thumb' guidance about what the projected runoff response might be to such a change, i.e. the 'rainfall elasticity of streamflow'. Analysis of more than 200 catchments across Australia suggests that in this region, a 1% change in rainfall equate to a 2–3.5 change in runoff (as observed in about 70% of catchments) (Chiew, 2006).

For this example, this relationship is convenient as it allows us to consider downscaled data sets that would otherwise require further processing to be implemented in runoff modelling, i.e. those of the national projections and the NARCliM datasets. This comparison illustrates the effect of the first crucial decision in a study,

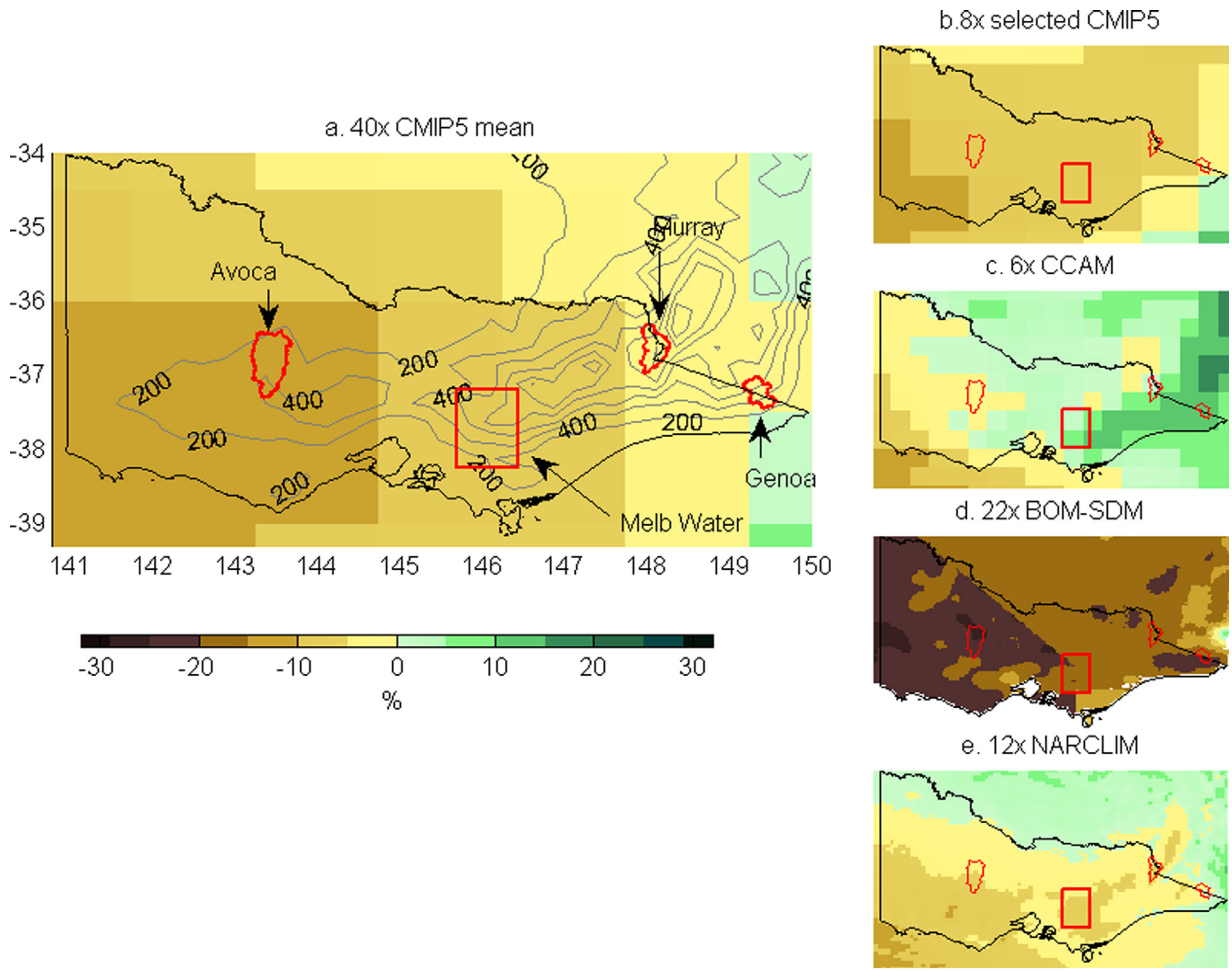


**Fig. 4.** Projected change in mean annual rainfall for CMIP5 GCMs (left panel), and then on the right in descending order: a sub selection of 8 representative CMIP5 GCMs, 6 CCAM downscaled fields, 21 BoM-SDM downscaling fields and 12 downscaled NARCliM fields. All changes are for 1986–2005 to 2080–2099 under RCP8.5, except NARCliM which is 1990–2009 to 2060–2079 under the SRES A2 scenario. The map shows the outline of selected catchments (along rivers Avoca, Murray and Genoa) used for the runoff study (4.3.1) and the rectangular region used to inform regional climate change for the Melbourne water resource case study (4.3.2), the State boundary for Victoria and the surface height of the region (height contours). The topography reveals the outline of the main topographic feature in this region, the Great Dividing Range, running parallel to the eastern coastline of Australia. The mountainous region in Victoria and NSW is commonly referred to as the Australian Alps.

the choice of input models. After this choice is explored, we illustrate the choice of method to produce datasets for the rainfall-runoff model. We also compare mean and quantile scaling of AWAP datasets of rainfall and evaporation by changes from CMIP5 GCMs available via the web portal for the national projections[1]. The hydrological model used to estimate runoff is GR4J (Perrin et al., 2003) – a lumped conceptual daily rainfall-runoff model. The GR4J model is widely used for water engineering and resource management worldwide and its performance is among the best models of its type (Vaze et al., 2010a). It has a parsimonious four-parameter structure corresponding to the maximum level of complexity that enables the optimum model performance. The modelling is carried out for each catchment using the daily rainfall and APET data averaged for the respective catchment. The model parameters are calibrated to reproduce the observed daily streamflow series over the period of 1981–2010, for the catchments described above. The GR4J model is driven with historical and future climate data to simulate corresponding runoff. The same optimised parameter values are used for both historical and future simulations.

[1] http://www.climatechangeinaustralia.gov.au/en/climate-campus/modelling-and-projections/using-projections/application-ready-data/.

*4.3.2. Water resources*

In this example we demonstrate the use of a decision scaling methodology (see e.g. Brown et al. (2012)) to assess climate risk for a large urban bulk water supply system in southeast Australia drawing on relationships developed in T14. The system comprises multiple storage reservoirs located in protected catchments, as well as a seawater desalination plant and inter-basin transfer for emergency supplies. Total system storage capacity is equivalent to about 5 years' demand. A detailed description of the modelling platform is provided in T14, here we merely note that the bulk supply system was modelled using a node-link mass balance simulation software package *eWater Source* (Kelley and O'Brien, 2012) that includes physical and operational specifications, including control curves, downstream minimum flow requirements and pump capacities.

The decision scaling approach involves identifying a vulnerability threshold, such as a critical tipping point, or a situation that would require management intervention in a system. In T14, the decision threshold relates to system performance through a yield measure. This measure defines the level of annual water supply required to meet demand without violating specified service standards and is constrained by a reliability and a vulnerability criterion (see T14 p. 3556 for further details). In T14, an experiment using the model platform with one thousand 100-year synthetic time series (based on streamflow statistics of historic inflows) was used to identify climate conditions that would sustain the required yield without violating the set criteria. This experiment enabled the construction of an empirical relationship between the system yield and the mean annual inflow from four adjacent catchments feeding the system. Combining this relationship with a second empirical relationship between streamflow and the climate variables allowed for yield to be defined as a climate response function of mean annual temperature (°C) and mean annual precipitation (mm).

In T14, assessments were made with regard to climate projections by CMIP3 models. Here we use the climate response function of T14 to relate annual yield (GL/yr) to climatic change as simulated by CMIP5 GCMs and regionally downscaled datasets available via the national projections (CSIRO, 2015) (Table 2). The change factors for CMIP5 models, CCAM and SDM are based on model output within the rectangular grid box outlined in Fig. 4 for the time periods 2020–2039 and 2080–2099 relative to the baseline period 1986–2005. This box includes the four catchments that provide the largest contribution to the yield in the bulk supply system and includes more than two thirds of the total system storage capacity.

## 5. Results

Through the use of different datasets in the four examples introduced above we can illustrate the influence on the change signal due to sub-setting GCM ensembles and use of downscaling (simple and more complex methods). An outline of comparisons are given in Table 1, and findings from each example are summarised as follows:

### 5.1. Heat stress

Our first measure of heat stress is incidence of days over 42 °C in mid-latitude and coastal Melbourne, a threshold that severely affects the physiology of birds and bats. To make an assessment on changed risk to this measure in future climate, we look at the change signal derived from different sources and make a number of comparisons (Table 2). Currently the frequency of days over 42 °C is 0.5 days per year, or around one day every second year on average. Using mean scaling, this is projected to increase differently under each RCP, with a median of 1.5 days per year for RCP4.5 and to 3.9 for RCP8.5 by 2090 (Fig. 2a). This difference is larger than for all other choices, demonstrating that choice of RCP is the most notable decision for temperature impacts by 2090. For a nearer future, the RCP is likely to be of less importance as for temperature the climate change signal is increasing in strength with time reflecting the increase in atmospheric greenhouse gas concentrations. The next most important choice is model selection (Fig. 2c), where the eight models selected to be representative show a similar projection (median 3.2 days) to the whole ensemble (median 3.1 days), but the lowest and highest 8 changes show very different results (medians 2.1 and 4.6 days). There is relatively smaller differences given by the choice to use the super-cluster or sub-cluster results (Fig. 2b: medians 3.9 and 3.1 days), or to use the downscaling (Fig. 2d: median 3.1 compared to 3.9 and 4.1 days), or to use quantile scaling over mean scaling (Fig. 2e: median 4.3 vs. 3.75 days).

It is possible that alternative methods of producing the dataset (e.g. bias correction of downscaled outputs) would give a greater difference than the choices explored here. However, for temperature we don't expect the climate change signal in bias corrected downscaling to be as different from scaled data as it could be for variables such as rainfall.

Our second heat stress index is incidence of days over 39.2 °C in tropical and coastal city of Darwin, which corresponds (under given assumptions) to 35 °C wet bulb temperature (thus also considering the air humidity). The type of comparisons for Darwin are the same as for Melbourne (Table 2).

Its proximity to the equator and coastal location gives a stable temperature regime. Thus, changes to the days over a threshold is very sensitive to the precise temperature projection. This is reflected in large differences between different model inputs (Fig. 3) and some very large outliers. The baseline is zero days, the GCM ensemble median for RCP4.5 is 0.2 days and for the higher emission scenario RCP8.5 it is 3.8 days but with a maximum of 76.5 days (Fig. 3a). Differences between CMIP5 model subsets are also large, from a median of 0.8 days for the low subset to 42 for the high subset (Fig. 3c). In contrast to the Melbourne case, the choice to use downscaling or quantile scaling also makes a large difference to the results, with medians of 3.9 and 18.8 days per year (Fig. 3d, 3e). Indeed, the inter-quartile range of mean and quantile scaling have a minimal overlap, showing that 50% of the decile scaled outputs are larger than the mean scaled outputs.

This highlights the large effect of simple decisions in preparing datasets for applications involving absolute thresholds in the tropics. Care must be taken that the change factors used are representative for the temporal and spatial scale of the application. A much more careful examination of the methods and assumptions than undertaken here is also warranted, e.g. for the consideration of change signal in variables other than temperature and the appropriateness of the chosen station time series used here. It also suggests the research question must be carefully framed, considering how meaningful a precise threshold actually is. Heat stress is a continuous variable, so it is possible to over-interpret the meaning of results either side of a precise threshold for 'extreme' heat stress. A measure of heat stress that doesn't rely on categories with precise absolute thresholds would be more appropriate in this case, e.g. using exceedances over sensibly chosen percentile thresholds.

### 5.2. Water resources

This study focuses on rainfall and runoff in three catchments in southeast Australia (as outlined in Fig. 4). Because of the input needs for rainfall-runoff modelling, actual runoff simulations are only conducted with the application-ready datasets produced by mean and quantile scaling of AWAP and APET data. This comparison draws on the rainfall elasticity of streamflow relationship, where 1% change in rainfall equate to about 2–3.5% change in runoff (Chiew, 2006).

The model mean projected change in mean annual rainfall varies considerably between the GCM and downscaled model ensembles (Fig. 4). The mean from all CMIP5 models shows a projected decrease of 0–20% in all catchments (Fig. 4a). In sub-selecting GCMs, the regional signal remains similar, but with somewhat different spatial pattern, which results in a different mean change for two of the three catchments, Avoca and Murray (Fig. 4b).

All three downscaled datasets (CCAM, SDM and NARCliM) have different spatial characteristics. Whilst all agree on projected decreases in rainfall in the west, the three datasets have a different signal over the Alps and east coast regions (Fig. 4c–e). NARCliM and to a limited extent CCAM, project an enhanced decrease in the high elevation areas over the Great Dividing Range (not seen in the SDM). Across eastern parts of the state of Victoria, CCAM suggest increases on the eastern slopes of the Great Dividing Range, a pattern somewhat supported by the NARCliM data in the southern parts of NSW. The pattern that deviates strongest from the GCM is that of SDM with large decreases, up to over 30% (implying a runoff reduction of 60 up to 100%, if implementing the elasticity relationship). However, a very linear demarcation in the SDM mapped dataset suggests existence of method dependent patterns that may require bias correction or smoothing prior to use in runoff projections.

Marked differences between these model ensembles are also found in the ranges of projected seasonal rainfall change for the three catchments (Fig. 5). On occasion, differences in projected ranges are such that they offer different guidance on direction of change in rainfall. This is the case for summer in Murray and

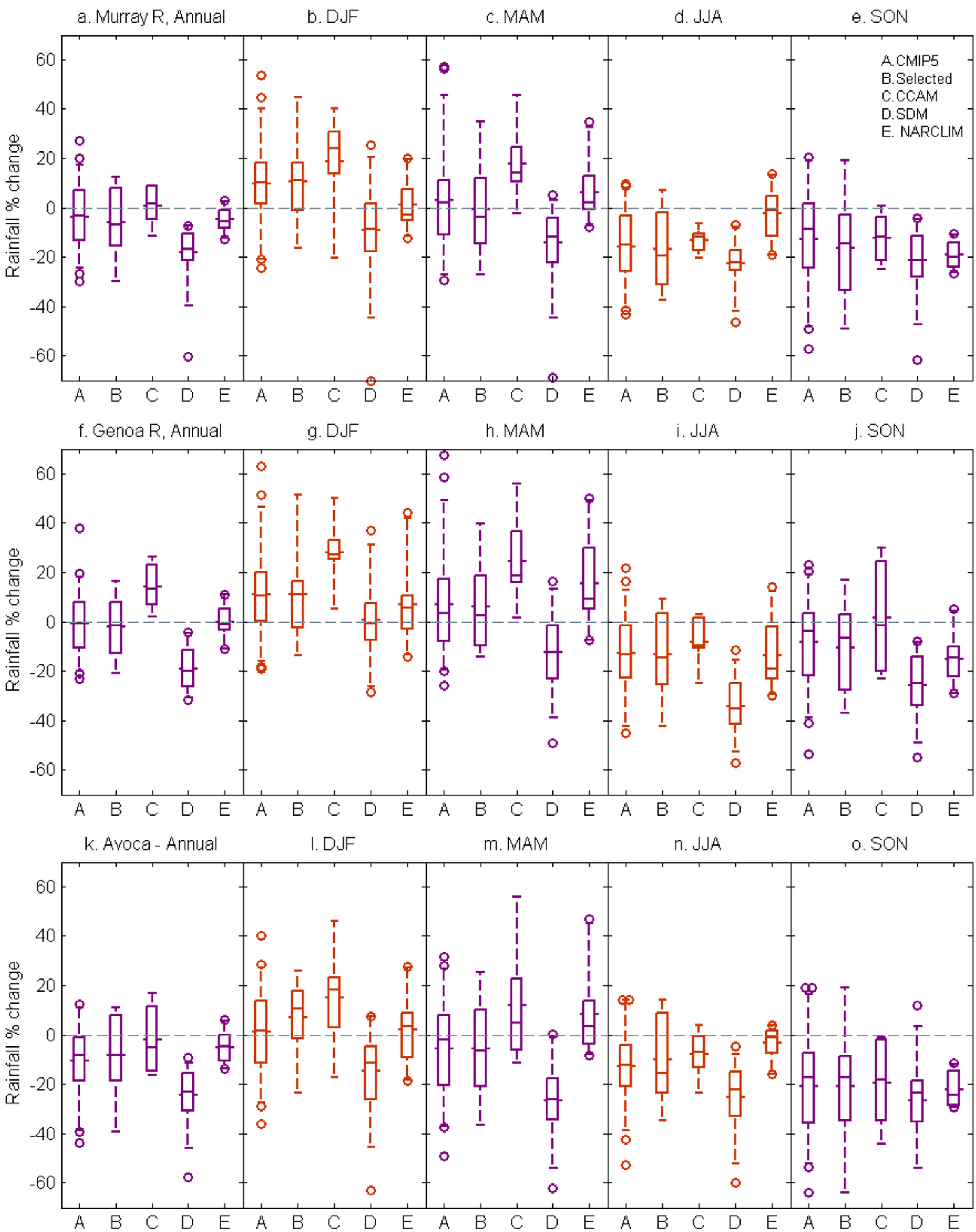


**Fig. 5.** For catchments Murray (top panel), Genoa (middle panel) and Avoca (bottom panel): annual and seasonal rainfall change (%) for 2080–2090 relative to baseline 1986–2005 following emission scenario RCP8.5. Datasets considered are listed in plot 5e.

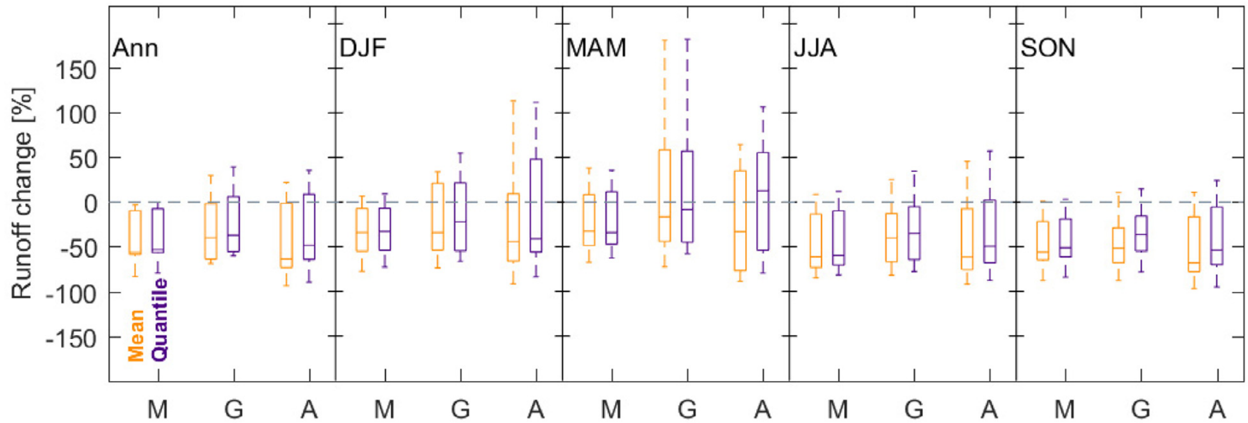


**Fig. 6.** Annual and seasonal runoff change (%) for 2080–2099 relative to 1986–2005 following RCP8.5 for three catchments (M = Murray, G = Genoa and A = Avoca, see Fig. 2). Orange bars indicate hydrological model inputs of mean scaled application-ready data for rainfall and APET, and purple bars indicate inputs of quantile scaled rainfall and mean scaled APET.

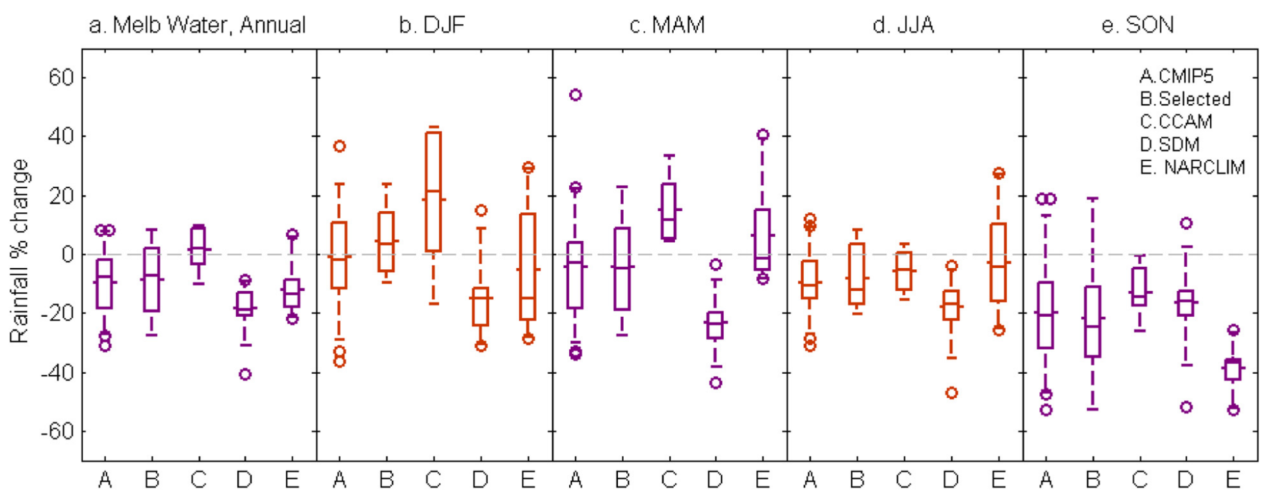


**Fig. 7.** For urban bulk supply system example (rectangular domain in Fig. 2): annual and seasonal rainfall change (%) for 2080–2090 relative to baseline 1986–2005 following emission scenario RCP8.5. Datasets considered are listed in plot 7e.

Avoca, as well as autumn in Genoa. In other seasons, particularly in winter for Genoa and Avoca, the different datasets lead to markedly different advice on the magnitude of projected change. These variations in projected change highlight the need for caution when relying on a particular downscaled dataset; noting that differences in projected runoff would be even larger than projected rainfall change given the regional rainfall elasticity of streamflow.

The use of mean or quantile scaling of rainfall for input in a hydrological model has a small effect on runoff seasonal averages (Fig. 6). Note that mean scaling of APET was used in both modelling experiments. Using the eight selected GCMs, the projected range of change is similar for each catchment in each season. Perhaps the only noteworthy differences is the slightly stronger decrease shown by the mean scaled rainfall relative to using quantile scaled rainfall; noting also that in autumn (MAM) the median of the mean scaled rainfall indicate decrease in runoff in Avoca, whilst the median of the quantile scaled rainfall indicates increase in runoff. The results illustrate the elasticity relationship, where changes are greater than those for rainfall in Fig. 5 (note different scales), and they also show a clear effect of increasing evapotranspiration where some increases in rainfall do not result in increases in runoff (e.g. evaporation losses offset increases in rainfall in Murray and Avoca in DJF). Assuming the validity of the elasticity relationship, we would expect a much larger range of results in runoff if we were also to consider change information by the downscaling methods SDM, CCAM and NARCliM.

Our last example demonstrate a bottom-up example whereby the change signal of different datasets are assessed against annual system yield First we examine projected rainfall change for context. Fig. 8 show the change in projected rainfall for the catchment region feeding our urban bulk water supply system example. The region is larger in size than the catchments considered for the runoff example (Fig. 4a), but still small relative to the typical scale of GCM output (~200 km). As was the case for the catchments above, there are marked differences in projected change in seasonal mean rainfall in the GCMs and the various downscaling sets. In particular, the sign of change is different in CCAM compared to BOM-SDM in summer and autumn.

Results from the bottom-up analysis of yield reflect the large differences in the input datasets (Fig. 8). The position of points relative to the grey lines indicate the system yield associated with a particular change in temperature and rainfall within the catchment area in Fig. 4. We note that yield lines have positive slope, so an increase in temperature (y-direction) gives a future with decreasing yield, and obviously more so with rainfall reductions Fig. 7 illustrates how the use of different datasets could lead to different conclusions about future impacts in the near (2020–2039) and far future (2080–2099) following RCP8.5. The change signal in the

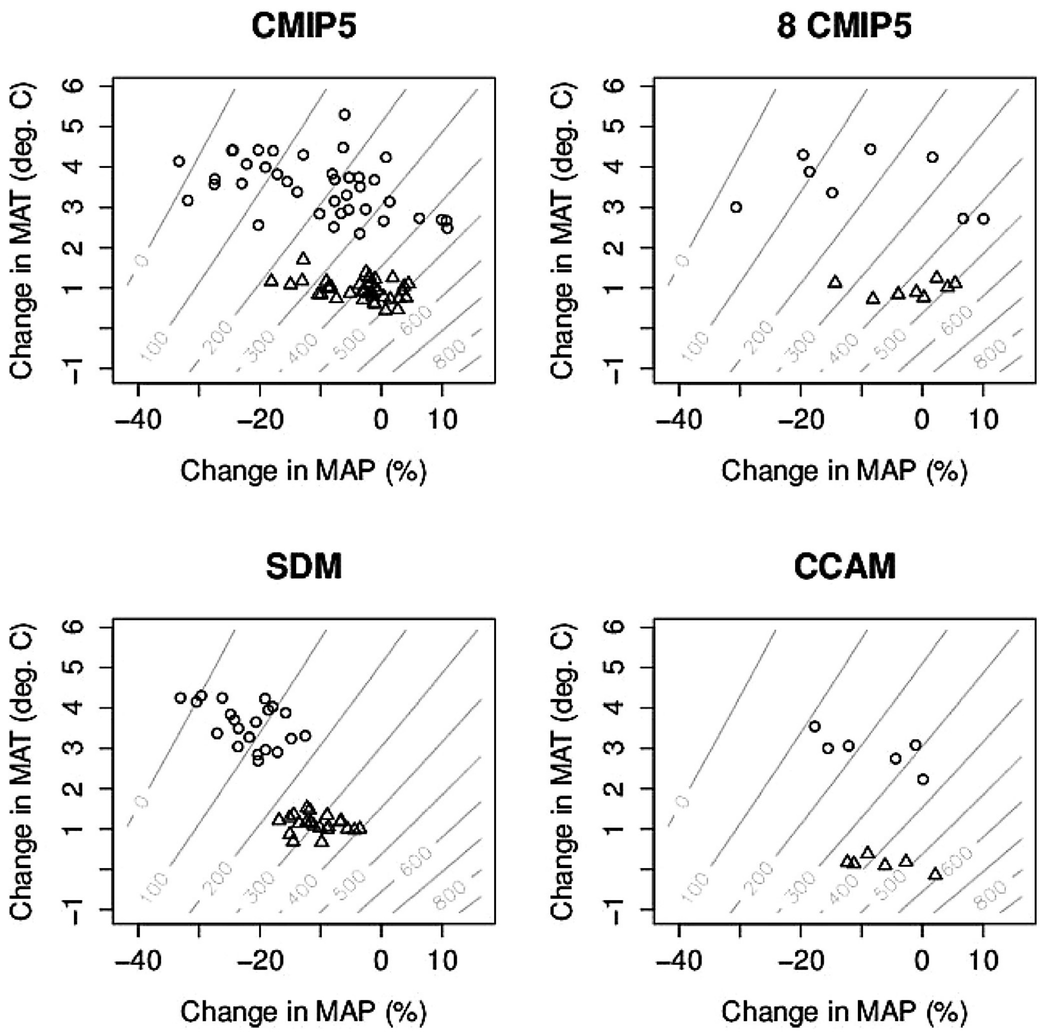


**Fig. 8.** For the catchment area (Fig. 4a): mean surface temperature (MAP, °C), mean annual precipitation (MAP,%) for 2080–2099 (circles) and 2020–2039 (triangles) relative to 1986–2005 for CMIP5 GCMs, 8 selected GCMs, SDM (21 models) and CCAM (6 models). Corresponding system yield (GL/yr) is indicated by grey lines.

GCMs indicate that the near future is likely to be drier (on an annual basis), with some models suggesting wetter, with a spread around the 400GL/yr line. The model range is larger in the later period, with some models giving larger decrease and a smaller number showing a greater increase. A similar model range (for both time horizons) is provided by the 8 models, indicating that the 8 models are broadly representative of the range provided by the full range of GCMs. This range highlights the enormous levels of climate uncertainty faced by planners. The two downscaled datasets give significantly different projections of the change signal in rainfall (less so in temperature) and consequential responses in yield. Both indicate an increase in temperature and a reduction in rainfall; in both cases the signal is stronger in the SDM data compared to CCAM, particularly so for the rainfall reduction. Thus if guided by SDM the projected range in yield change by 2090 is about 0 to less than 200GL/yr, whilst that of CCAM is about 100–400 GL/yr. Both ranges fit within that of the GCM (about 0–500 GL/yr), but results indicate a more constrained range of projected change compared to the wider GCM ensemble. This gives a more emphatic result for water managers, and could either lead to more targeted adaptation response if the constrained range is demonstrated to be suitable, or the possibility of mal-adaptation if the restricted range is not shown to be adequate. We note that in circumstances when different datasets provide very different guidance on ‘change’ and there is no obvious reason for why one dataset would be more credible than others, significant uncertainty exist around the regional signal and drawing conclusions from a particular result would not be recommended.

## 6. Discussion

Different IAV applications have different climate change information requirements, so different methods of producing regional projection dataset are more appropriate depending on the application needs. These needs may include requirements of spatial resolution, interest in capturing changes to the tails of a distribution rather than mean change, changes to seasonal timing and more. Such criteria may rule out some datasets purely because they are unlikely to contain the change information of interest. Certainly different downscaling techniques have different capabilities in adding regional detail to the climate change signal as simulated by the GCM. However, studies drawing on information from more complex techniques generally use only a limited number of emissions scenarios and GCMs, which raises the issue of representativeness and completeness of the datasets used. As shown here, using an unrepresentative set of projections raises the possibility of skewed analysis of climate change and the possibility of mal-adaptation to climate change. If a reliable constraint is found on the CMIP5 range of climate changes, or if a particular sub-set of

GCM outputs or downscaling ensemble is demonstrated to have a more plausible depiction of the climate change signal there is a case for using this in preference. There may be theoretical reasons to imply downscaling produces a more plausible change signal, such as through improved representation of topography and coastlines (Rummukainen, 2016), but it must also be demonstrated that there are no aspects of climate change specific to the downscaling method itself. However, if this has not been demonstrated then relying on theoretical and undemonstrated advantages for downscaling, such as greater resolution, is not justified.

Thus, users need to make choices around methods that have capability to capture the full plausible range of change of interest as well as making sure that key uncertainties are considered in a final assessment that could be used to guide policy development. For example, the user may want to consider appropriateness of the dataset relative to simulating change for the far or near future, is there marked topography in the region or otherwise high contrast in the land-surface environment, is temporal sequencing or extremes important to the application, is daily spatially coherent information needed, how well is multiple emission scenarios and GCMs sampled (Ekström et al., 2015).

This paper demonstrates some impacts of choosing different sources of climate change information in IAV work and the consequences it may have on subsequent policy guidance. To this purpose we use real world applications in combination with existing regional datasets. We propose a pathway for production of application-ready data sets and outline critical stages that may influence the final change signal, such as sub-selecting GCMs, choice of downscaling/scaling and requirement of bias correction. Through our examples we can demonstrate when some of these choices have fundamental impact on the change signal by comparing the use of different inputs, as outlined in Table 2.

The first stage of developing projections, or risk assessment, is to consider the ability of GCMs to represent the process that is of interest. Our case studies draw primarily on projections of temperature and rainfall. The two heat indices are primarily dependent on temperature whilst the water resource examples are strongly influenced by rainfall projections. This is a relevant distinction as GCMs agree more on temperature change compared to rainfall change, simply because the latter depend not only on changes due to a warmer atmosphere but also on the model representation on weather generating systems. Thus, temperature projections are typically given higher confidence than rainfall projections. However, the magnitude of temperature change is still quite uncertain, due in large part to our poorly constrained estimate of climate sensitivity of the earth system (IPCC, 2013). Projections of rainfall in particular must be broadly representative of plausible circulation change, and there may be great impact from model bias with consequential influences on final results. In our heat stress case studies, the direction of change is clear in all datasets, however the different choices influence the magnitude of projected change. The choice of emissions scenario is a matter for the risk management approach of the application, e.g. what future is most relevant from a risk assessment perspective.

The second stage involves the potential need of sub-setting GCMs. This requirement may arise due to a need to achieve regional resolution through scaling or downscaling, because a user is interested in a particular direction of change or because some GCMs may not have the variable of interest (or have it at the temporal resolution of interest). Our heat stress case studies illustrate the effect of sub-setting through the use of the 8 representative GCMs, and the eight models with the highest and lowest change signal for the heat stress examples. The examples show that choosing different models within the ensemble can have a strong influence on results (Chiew et al., 2009a), but if specially chosen to be representative of the larger ensemble range, a sub-sample can work well to demonstrate the range of plausible futures as simulated by the GCMs. Until there is a reliable constraint on the temperature projections, the output of the range of CMIP5 can be taken as a minimum range of possible change, so should be considered. Given a high emission scenario, the choice of the coolest models suggest around a more modest increase of 2 days per year, whereas the hottest suggest a greater increase to over 4 days per year. This difference of a factor of two in the result could potentially make a difference to the management actions taken to ameliorate the impact of this heat stress. The choice to use a representative CMIP5 subset compared to the whole ensemble, to mean scale downscaling, to use of a wider or narrower averaging region (super cluster versus sub-cluster) or to use decile scaling are much less likely to affect the adaptation decision.

Downscaling is intended to add value relative to projections by GCMs through finer resolved modelling or drawing on relationships between large and local scale variables (particularly around environmental boundaries such as coasts and complex topography). However, if using downscaled output also implies sub-setting GCMs, users need to consider the benefits of more complex downscaling relative to the need to sample a wider GCM range. Amongst our examples, the influence on the change signal from scaled and downscaled datasets was strongest when we were interested in change around a threshold for a distribution with otherwise little temporal variability (our Darwin heat stress example) and for both water resource examples due to the focus on relatively small geographical regions using projections associated with large uncertainty in projected magnitudes and change patterns. Thus, as demonstrated by our examples, downscaling may provide potential regional insights, but can also add a layer of uncertainty when different methods indicate opposing direction of change. With a small sample of methods it is difficult to attribute skill to a particular dataset without additional analysis of the physical plausibility of the regional projection. Overall, we would recommend caution when using downscaled projections that have not been demonstrated to add value relative to the regional GCM change signal. We note that demonstrating such added value is typically beyond what is expected by users of application ready datasets, rather such demonstration of added value is perhaps best made by the producers of the data sets who are familiar with the production method and its qualities. If using datasets that haven't been demonstrated to add value, users could aim to show results based on downscaling along with those based on GCM output to illustrate the full range of the physical plausibility of simulated climate change.

The application requiring the most detailed information about change is the runoff example, where all moments of the temporal distribution along with timing and periodicity of daily rainfall are key concerns as runoff has non-linear relationships to rainfall amount due to dependencies on soil wetness and connection to groundwater storage. However, we find that the choice of GCM inputs, and the associated mean rainfall projection, is in fact a far bigger source of uncertainty than the more complex downscaling sources of rainfall change. This is consistent with previous findings of Chen et al. (2011) and Frost et al. (2011). In our runoff projections, only the mean and quantile scaled application-ready data sets were used. We note that whilst these datasets are bias free (in the sense that there is no 'current climate bias', observed data is simply scaled to reflect a future change as simulated by models), they wouldn't inform a user about change in duration of runoff events as the change factor is applied to observed time series on a monthly resolution (hence change is implemented on a monthly time step). For this reason, the output from more complex downscaling is attractive as they may inform on more complex changes to the hydrological regime such as changes to drought or rain-day duration. However, as noted above, more complex downscaling typically require bias correction as hydrological models can be highly tuned to a particular observed dataset and may give a systematic bias if used in combination with a dataset different to that used for its calibration (Ekstrom and Jones, 2009). This can present a challenge to bias correction, where the correction is likely to be much greater than observed uncertainty and projected changes. Given the sensitivity to temporal sequencing, it is not surprising that methods that express a change in variability can show different results to those from simple mean scaling, as demonstrated for Tasmania for dynamically downscaled outputs Ling et al. (2014) and for the United States for statistical downscaling by Hay et al. (2000).

Finally we note that users also need to consider the validity of the IAV model under change environmental conditions. We can refer to the on-going discussions on robustness of hydrological models under a non-stationary climate (Milly et al., 2008, 2015). For example, Vaze et al. (2010b) showed that models calibrated over a wet period showed less predictive skill when applied to a dry period compared to those calibrating over a dry period and predicting in a wet period. These are implications to consider when applying operational models in a climate change context. This uncertainty can make fundamental IAV outcomes unclear, for example the choice of climate variables to use in species distribution models can make the difference between projected extinction or survival, an influence greater than the choice of GCM input (Harris et al., 2013).

Our second water resource study demonstrates an alternative to implementing projection data in IAV models by linking probabilistic risk assessment to identified system vulnerabilities (or thresholds). Different variations of this (bottom-up) approach have been demonstrated for a range of water sector applications, e.g. scenario-neutral approach to assess flood risk in UK catchments (Prudhomme et al., 2010), the decision scaling approach by (Brown et al., 2012). Others have pointed importance of considering socio-economical dimensions to managing water resource systems (considering decisions in relation to 'supply' as well as 'demand'). Indeed Korteling et al. (2013) demonstrated that taking such an approach could reveal attractive options where management decisions on the demand side could under certain circumstances (i.e. policy uptake by public) match adaptation options focusing on management of the supply side. For example Paton et al. (2013) demonstrated that demand was always the largest source of uncertainty relative to other sources (e.g. GCMs and emission scenarios used) when considering uncertainties for a southern Australian water resource system over the 2010–2050 planning horizon.

## 7. Conclusions

Central to IAV application is the representation of key uncertainties in climate projections, i.e. those relating to emission futures, models ability to simulate the climate response to changing greenhouse gas concentrations and natural climate variability. The first is perhaps simply an issue of choice as we can choose to study a low, medium or high emission future. Our ability to simulate the climate change and natural variability is however, only quantifiable through consideration of the outputs from multiple realisations of the climate systems, i.e. multiple GCMs. Thus the consideration of the spread of GCMs is fundamental to all climate change research. The complexity of many techniques to analyse climate change impacts often means a limited set of emission scenarios, GCMs or downscaling techniques are considered and this leads to a restricted and unrepresentative depiction of the climate change signal currently thought of as plausible. This increases the possibility of biased climate change projection advice and subsequent mal-adaptation to possible change. For this reason we suggest that downscaled information is used in context with the range of the GCM change signal unless assessment shows that the downscaled signal is different to that of the GCM change signal and the mechanism for that difference can be understood and deemed credible.

Through simple demonstration studies using projections of temperature and precipitation associated with the 2015 Australian national projections and the regional NARCLiM projections, the influence of different decisions when producing an application-ready data set is illustrated. From the heat stress examples we note that the strongest influence by far concerns the choice of emission scenario, directly influencing the heat retained in the atmosphere by greenhouse gases. A poor selection of GCMs could also distort the change signal as GCMs have different sensitivity to increasing greenhouse forcing. We also note that studies in environments with small range of variability are particularly sensitive to model biases or sub-setting GCM inputs, as demonstrated by our Darwin heat stress case study.

Our water resource case studies show the large differences that different climate change information can make for conclusions on future water supply. Our test regions are within and adjacent to the Australian Alps and as such, downscaling is expected to add value to the GCM projections. However, for this region (and common for many other parts of the world) very few downscaled datasets are available. As the work is conducted on relatively small spatial regions, differences between the datasets (both pattern and magnitude wise) lead to markedly different projections; even sometimes indicating a difference in projected direction of change. It may therefore be wise to first assess the regional change signal to assess if any fine scale patterns in the change signal exist and, as noted above, can be understood and deemed credible. If not, it may be more robust to consider change of a greater spatial region.

Rarely do IAV-applications have application-ready data sets produced specifically for their purpose. Rather, data sets are provided by collaborative networks or downloaded from national/global websites. However, when research feeds into policy development, considering the relevance of the change signal to the application as well as representation of key uncertainties are central to the credibility of the projection. Failing to do so may to lead to a skewed or biased projection unfit for policy development purposes.

## Acknowledgement

The Authors would like to acknowledge the internal reviewers of CSIRO for their constructive comments on this manuscript. The work presented in this paper was funded by the Regional climate projections science project of the Australian Climate Change Science Program (ACCSP).

## References

Argüeso, D., Evans, J.P., Fita, L., 2013. Precipitation bias correction of very high resolution regional climate models. Hydrol. Earth Syst. Sci. 17, 4379–4388.
Bennett, J.C., Grose, M.R., Corney, S.P., White, C.J., Holz, G.K., Katzfey, J.J., Post, D.A., Bindoff, N.L., 2013. Performance of an empirical bias-correction of a high-resolution climate dataset. Int. J. Climatol. 34, 2189–2204.
Brown, C., Wilby, R.L., 2012. An alternate approach to assessing climate risks. EOS Trans. Am. Geophys. Union 93.
Brown, C., Ghile, Y., Laverty, M., Li, K., 2012. Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour. Res. 48, 9537.
Brown, I., Berry, P., Everard, M., Firbank, L., Harrison, P., Lundy, L., Quine, C., Rowan, J., Wade, R., Watts, K., 2015. Identifying robust response options to manage environmental change using an Ecosystem Approach: a stress-testing case study for the UK. Environ. Sci. Policy 52, 74–88.
Budyko, M.I., 1974. Climate and Life. Academic Press, New York.
Chen, J., Brissette, F.P., Leconte, R., 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. J. Hydrol. 401, 190–202.

Chiew, F.H.S., 2006. Estimation of rainfall elasticity of streamflow in Australia. Hydrol. Sci. J. Des. 51, 613–625.
Chiew, F.H.S., Kirono, D.G.C., Kent, D., Vaze, J., 2009a. Assessment of rainfall simulations from global climate models and implications for climate change impact on runoff studies. 18th World Imacs Congress and Modsim09 International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences, pp. 3907–3913.
Chiew, F.H.S., Teng, J., Vaze, J., Post, D.A., Perraud, J.M., Kirono, D.G.C., Viney, N.R., 2009b. Estimating climate change impact on runoff across southeast Australia: method, results, and implications of the modeling method. Water Res. Res. 45.
Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys. Res. Lett. 35, L20709.
Conroy, M.J., Runge, M.C., Nichols, J.D., Stodola, K.W., Cooper, R.J., 2011. Conservation in the face of climate change: the roles of alternative models, monitoring, and adaptation in confronting and reducing uncertainty. Biol. Conserv. 144, 1204–1213.
CSIRO, B.O.M., 2015. Climate change in Australia Information for Australia's Natural Resource Management Regions: Technical Report. Australia.
DEFRA, 2009. Department for Environment, Food and Rural Affairs Adapting to Climate Change: UK Climate Projections. London, UK.
Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. HESS opinions: should we apply bias correction to global and regional climate model data? Hydrol. Earth Syst. Sci. 16, 3391–3404.
Ekstrom, M., Jones, P.D., 2009. Impact of rainfall estimation uncertainty on streamflow estimations for catchments Wye and Tyne in the United Kingdom. Int. J. Climatol. 29, 79–86.
Ekström, M., Kuruppu, N., Wilby, R.L., Fowler, H.J., Chiew, F.H.S., Dessai, S., Young, W.J., 2013. Examination of climate risk using a modified uncertainty matrix framework-applications in the water sector. Global Environ. Change Human Policy Dimensions 23, 115–129.
Ekström, M., Grose, M.R., Whetton, P.H., 2015. An appraisal of downscaling methods used in climate change research. Wiley Interdiscip. Rev. Clim. Change.
Epstein, Y., Moran, D.S., 2006. Thermal comfort and the heat stress indices. Ind. Health 44, 388–398.
Evans, J.P., Ji, F., G., Abramowitz, Ekstroem, M., 2013. Optimally choosing small ensemble members to produce robust climate simulations. Environ. Res. Lett. 8.
Evans, J.P., Ji, F., Lee, C., Smith, P., Argüeso, D., Fita, L., 2014. Design of a regional climate modelling projection ensemble experiment; NARCliM. Geosci. Model Dev. 7, 621–629.
Foley, A.M., 2010. Uncertainty in regional climate modelling: a review. Prog. Phys. Geogr. 34, 647–670.
Fowler, H., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27, 1547–1578.
Frost, A.J., Charles, S.P., Timbal, B., Chiew, F.H.S., Mehrotra, R., Nguyen, K.C., Chandler, R.E., Mcgregor, J.L., Fu, G., Kirono, D.G.C., Fernandez, E., Kent, D.M., 2011. A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. J. Hydrol. 408.
Giorgi, F., 2008. Regionalization of climate change information for impact assessment and adaptation. Bull. World Meteorol. Organiz. 57.
Grose, M.R., Moise, A.F., Timbal, B., Katzfey, J.J., Ekström, M., Whetton, P.H., 2015. Climate projections for southern Australian cool-season rainfall: insights from a downscaling comparison. Clim. Res. 62, 251–265.
Harris, R.M.B., Porfirio, L.L., Hugh, S., Lee, G., Bindoff, N.L., Mackey, B., Beeton, N.J., 2013. To Be Or Not to Be? Variable selection can change the projected fate of a threatened species under future climate. Ecol. Manage. Restor. 14, 230–234.
Harris, R.M.B., Grose, M.R., Lee, G., Bindoff, N.L., Porfirio, L.L., Fox-Hughes, P., 2014. Climate projections for ecologists. Wiley Interdiscip. Rev. Clim. Change 5, 621–637.
Hawkins, E., Sutton, R., 2009. The potential to narrow uncertainty in regional climate predictions. Bull. Am. Meteorol. Soc. 90, 1095–1107.
Hay, L.E., Wilby, R.L., Leavesley, G.H., 2000. A comparison of delta change and downscaled gcm scenarios for three mounfainous basins in the united states1. JAWRA J. Am. Water Res. Assoc. 36, 387–397.
IPCC, 2013. The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge, United Kingdom and New York. Cambridge University Press, NY, USA.
Jackson, J.E., Yost, M.G., Karr, C., Fitzpatrick, C., Lamb, B.K., Chung, S.H., Chen, J., Avise, J., Rosenblatt, R.A., Fenske, R.A., 2010. Public health impacts of climate change in Washington State: projected mortality risks due to heat events and air pollution. Clim. Change 102, 159–186.
Jones, D.A., Wang, W., Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. Aust. Meteorol. Oceanogr. J. 58, 233–248.
Kelley, P., O'brien, A., 2012. Source Scientific Reference Guide. Canberra, Australia: eWater Cooperative Research Centre.
Kendon, E., Jones, R., Kjellstrom, E., Murphy, J., 2010. Using and designing GCM-RCM ensemble regional climate projections. J. Clim. 23, 6485–6503.
Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G., 2010. Challenges in combining projections from multiple climate models. J. Clim. 23, 2739–2758.
Korteling, B., Dessai, S., Kapelan, Z., 2013. Using information-gap decision theory for water resources planning under severe uncertainty. Water Resour. Manage 27, 1149–1172.
Ledbetter, R., Prudhomme, C., Arnell, N., 2012. A method for incorporating climate variability in climate change impact assessments: Sensitivity of river flows in the Eden catchment to precipitation scenarios. Clim. Change 113.
Lempert, R., Kalra, N., 2011. Managing climate risks in developing countries with robust decision making. World Res. Rep. Uncertainty Ser.
Lempert, R.J., Groves, D.G., Popper, S.W., Bankes, S.C., 2006. A general, analytic method for generating robust strategies and narrative scenarios. Manage. Sci. 52, 514–528.
Ling, F.L.N., Robinson, K.A., Bennett, J.C., 2014. Projections of Storage Reliability Under Future Climate. Hydrology and Water Resources Symposium, Perth, Australia.
Maraun, D., Widmann, M., Gutiérrez, J.M., Kotlarski, S., Chandler, R.E., Hertig, E., Wibig, J., Huth, R., Wilcke, R.A.I., 2015. Value: a framework to validate downscaling approaches for climate change studies. Earth's Future. 2014EF000259.
Masson, D., Knutti, R., 2011. Climate model genealogy. Geophys. Res. Lett. 38.
Mastrandrea, M.D., Heller, N.E., Root, T.L., Schneider, S.H., 2010. Bridging the gap: linking climate-impacts research with adaptation planning and management. Clim. Change 100, 87–101.
Mcgregor, J., Dix, M., 2008. An updated description of the conformal-cubic atmospheric model. In: Hamilton, K., Ohfuchi, W. (Eds.), High Resolution Numerical Modelling of the Atmosphere and Ocean. Springer, New York.
Mearns, L.O., Sain, S., Leung, L.R., Bukovsky, M.S., McGinnis, S., Biner, S., Caya, D., Arritt, R.W., Gutowski, W., Takle, E., Snyder, M., Jones, R.G., Nunes, A.M.B., Tucker, S., Herzmann, D., McDaniel, L., Sloan, L., 2013. Climate change projections of the north American regional climate change assessment program (NARCCAP). Clim. Change 120, 965–975.
Meehl, G.A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J.F.B., Stouffer, R.J., Taylor, K.E., 2007. The WCRP CMIP3 multimodel dataset – A new era in climate change research. Bull. Am. Meteorol. Soc. 88, 1383–1394.
Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: whither water management? Science 319, 573–574.
Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., Krysanova, V., 2015. On critiques of "stationarity is dead: whither water management?". Water Resour. Res. 51, 7785–7789.
Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grübler, A., Jung, T.Y., Kram, T., la Rovere, E.L., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, P., Shukla, P., Smith, S., Swart, R., van Rooijen, S., Victor, N., Dadi, Z., 2000. Special Report on Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom.
Oleson, K.W., Monaghan, A., Wilhelmi, O., Barlage, M., Brunsell, N., Feddema, J., Hu, L., Steinhoff, D.F., 2015. Interactions between urbanization, heat stress, and climate change. Clim. Change 129, 525–541.
Paton, F.L., Maier, H.R., Dandy, G.C., 2013. Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system. Water Resour. Res. 49, 1643–1667.
Perrin, C., Michel, C., Andreassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279, 275–289.
Piani, C., Weedon, G.P., Best, M., Gomes, S.M., Viterbo, P., Hagemann, S., Haerter, J.O., 2010. Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J. Hydrol. 395, 199–215.
Porter, J.J., Demeritt, D., Dessai, S., 2015. The right stuff? Informing adaptation to climate change in British Local Government. Global Environ. Change Human Policy Dimensions 35, 411–422.
Prudhomme, C., Wilby, R., Crooks, S., Kay, A., Reynard, N., 2010. Scenario-neutral approach to climate change impact studies: application to flood risk. J. Hydrol. 390, 198–209.
Rummukainen, M., 2016. Added value in regional climate modeling. Wiley Interdiscip. Rev. Clim. Change 7, 145–159.
Sanderson, B.M., Knutti, R., Caldwell, P., 2015. Addressing interdependency in a multimodel ensemble by interpolation of model properties. J. Clim. 28, 5150–5170.
Schmidli, J., Goodess, C.M., Frei, C., Haylock, M.R., Hundecha, Y., Ribalaygua, J., Schmith, T., 2007. Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. J. Geophys. Res. Atmos. 112.
Schwartz, M.W., 2012. Using niche models with climate projections to inform conservation management decisions. Biol. Conserv. 155, 149–156.
Skamarock, W.C., Klemp, J.B., 2008. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. J. Comput. Phys. 227, 3465–3485.
Smith, L.A., Petersen, A.C., 2014. Variations on reliability: connecting climate predictions to climate policy. In: Boumans, M., Hon, G., Petersen, A.C. (Eds.), Error and Uncertainty in Scientific Practice. Pickering & Chatto, London.
Stainforth, D.A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D.J., Kettleborough, J.A., Knight, S., Martin, A., Murphy, J.M., Piani, C., Sexton, D., Smith, L.A., Spicer, R.A., Thorpe, A.J., Allen, M.R., 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. Nature 433, 403–406.
Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498.

Teng, J., Chiew, F.H.S., Timbal, B., Wang, Y., Vaze, J., Wang, B., 2012. Assessment of an analogue downscaling method for modelling climate change impacts on runoff. J. Hydrol. (Amsterdam), 472–473.

Teng, J., Potter, N.J., Chiew, F.H.S., Zhang, L., Wang, B., Vaze, J., Evans, J.P., 2015. How does bias correction of regional climate model precipitation affect modelled runoff? Hydrol. Earth Syst. Sci. 19, 711–728.

Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456, 12–29.

Timbal, B., Fernandez, E., Li, Z., 2009. Generalization of a statistical downscaling model to provide local climate change projections for Australia. Environ. Modell. Software 24, 341–358.

Trewin, B., 2013. A daily homogenized temperature data set for Australia. Int. J. Climatol. 33, 1510–1529.

Turner, S.W.D., Marlow, D., Ekstroem, M., Rhodes, B.G., Kularathna, U., Jeffrey, P.J., 2014. Linking climate projections to performance: a yield- based decision scaling assessment of a large urban water resources system. Water Resour. Res. 50, 3553–3567.

UNFCCC, 2008. Compendium on methods and tools to evaluate impacts of, and vulnerability and adaptation to, climate change.

van Vuuren, D., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S., Rose, S., 2011. The representative concentration pathways: an overview. Clim. Change 109, 5–31.

Vaze, J., Chiew, F., Perraud, J.-M., Viney, N., Post, D., Teng, J., Wang, B., Lerat, J., Goswami, M., 2010a. Rainfall-runoff modelling across southeast Australia: datasets, models and results. Aust. J. Water Res. 14, 101–116.

Vaze, J., Post, D.A., Chiew, F.H.S., Perraud, J.M., Viney, N.R., Teng, J., 2010b. Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. J. Hydrol. 394, 447–457.

Weaver, C.P., Lempert, R.J., Brown, C., Hall, J.A., Revell, D., Sarewitz, D., 2013. Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. Wiley Interdiscip. Rev. Clim. Change 4, 39–60.

Webb, L., Clarke, J., Hennessy, K., Heady, C., Erwin, T., 2015. Climate Change in Australia Data Delivery. CSIRO, Australia.

Welbergen, J.A., Klose, S.M., Markus, N., Eby, P., 2008. Climate change and the effects of temperature extremes on Australian flying-foxes. Proc. R. Soc. London B: Biolog. Sci. 275, 419–425.

Whetton, P., Hennessy, K., Clarke, J., McInnes, K., Kent, D., 2012. Use of representative climate futures in impact and adaptation assessment. Clim. Change 115, 433–442.

Wilby, R.L., Dessai, S., 2010. Robust adaptation to climate change. Weather 65, 180–185.

Yang, W., Andréasson, J., Phil Graham, L., Olsson, J., Rosberg, J., Wetterhall, F., 2010. Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. Hydrol. Res. 41, 211–229.