Office of the Chief Scientist & Engineer

Independent review of the climate risk method for the NSW Regional Water Strategies Program

Independent Expert Panel

2 April 2020

Note: The draft Methods Paper and background reports that the Panel reviewed for this report were not public at the time of the review.

Office of the Chief Scientist & Engineer

The Hon Melinda Pavey MP
Minister for Water, Property and Housing
GPO Box 5341
SYDNEY NSW 2001

Dear Minister

Independent Review of Climate Risk Method for the Regional Water Strategies program

In September 2019, you requested that I chair an independent expert panel to assess the proposed methodology being used by DPIE-Water to incorporate the climate risk to water resources as part of the development of Regional Water Strategies in NSW.

Please find enclosed the Panel's report. This report provides a commentary on the proposed method to improve resilience to climate risk in light of uncertainty, bringing together climate variability and change models. The Report makes a number of short- and long-term recommendations to further develop the method and improve knowledge around future climate.

I would like to acknowledge the contribution of Panel members: Professor Bryson Bates, Emeritus Professor George Kuczera, Professor Andy Pitman and Dr Scott Power, as well as the advice and support provided by staff and other stakeholders.

Yours sincerely

Dr Chris Armstrong PSM
Deputy NSW Chief Scientist & Engineer
Chair, Independent Expert Panel
2 April 2020

cc. Mr Jim Bentley, CEO and Deputy Secretary Water
Ms Rachel Connell, Executive Director, Regional Water Strategies

CONTENTS

CONTENTS

FIGURES AND TABLES

APPENDICES

EXECUTIVE SUMMARY

Highlights

- The use of stochastic models represents an important advance compared with the use of historical data and climate models alone.
- The Panel agrees with the overall approach of developing the stochastic models informed by knowledge of dominant climate drivers affecting different regions of NSW and using the characteristics of these drivers to inform understanding of the historical statistical nature of rainfall and evapotranspiration.
- The Panel notes that the paleo information captured in the current models is primarily used to better estimate the probability model of IPO periods, allowing a more representative distribution of IPO durations to be developed.
- Coupling with the stochastic models enables the practitioner to understand the joint impact of climate variability and climate change.
- Projections of rainfall and evapotranspiration in future plausible climate scenarios can inform the likelihood and consequences of events, such as droughts and storage-filling wet events, and with consideration of uncertainty, inform decision making.
- The products developed by DPIE-Water and its collaborators are data sets specific to different regions across NSW that can be used to inform decision making and preparedness. While they are based on model simulations, they enable DPIE-Water to develop robust risk-based and adaptive responses to climate variability and change.
- Any advice stemming from this methodology (as with all methods) should reflect the fact that events outside the range simulated, while unlikely, might occur.
- Overall, the DPIE-Water's methodology to use observational and paleoclimate data informed by an understanding of climate drivers to select, calibrate and test stochastic models and the factoring of NARCliM projections is consistent with best-practice approaches to climate risk management.
- The Panel is of the view that ongoing improvement of the methodology adopted by DPIE Water be given high priority. This can be undertaken in parallel with concurrent RWS work. It would include an assessment of non-stationarity in the historical record to determine whether changes in climate in recent decades affect estimates of present-day climate risk compared with climate risk based on the whole observed record.
- A further high priority of the Panel is for DPIE to convene a community of practice within DPIE and potentially broadening out, that maintains an ongoing work program to keep up-to-date with new scientific methods and findings in relation to hydrology/water availability and climate modelling. The scientific literature on these interrelated topics are becoming ever more complex. The range of staff in DPIE who use climate and hydrology data and tools is wide, and they will need to stay up-to-date as the fields advance.

Background to methodology

The Independent Expert Panel – Independent Review of Climate Risk Method for Regional Water Strategies was established to provide expert advice to the Department of Planning, Industry and Environment Water (DPIE-Water) on the proposed methodology that has been developed to account for climate variability and change in the Regional Water Strategies (RWS). The RWS strategies project over the next 20 to 40 years to determine how much water a region will need to meet future demand. Terms of Reference are at Appendix 1.

The strategies will determine future water demand of a region, as well as setting out the challenges and choices involved in meeting those regional needs. The strategies will lay out actions that can be taken to manage risks to water availability, with the goal of achieving resilient water resources for towns and communities, industry, Aboriginal communities and the environment.

The method proposed by DPIE-Water draws on continuous daily rainfall data (observed, interpolated or derived values) from the period of 1880s to 2019, evapotranspiration data from the 1970s derived back to the 1880s and paleoclimate proxy data used to infer ~400 years of climate sequences (Queensland Government, 2020c). The rainfall and evapotranspiration data were extracted from the Queensland Government SILO database.

The relevant regional data from the SILO database and the paleoproxy information are used with statistical stochastic modelling techniques to generate synthetic data sets of rainfall and evapotranspiration over a 10,000 year timeframe, providing insights into the plausible frequency, length and distribution of dry- and wet-dominant periods and insight into the natural variability of the climate beyond the observed records. These statistically generated rainfall and evapotranspiration sequences are informed by understanding of climate drivers that are important to the NSW rainfall patterns in the various regions that are the focus of the RWS, in particular East Coast Lows (ECL) in south east NSW, and the Interdecadal Pacific Oscillation (IPO) in the north of the state. The Panel notes that additional drivers are identified for the south and south-west NSW; however, DPIE-Water informed the Panel that the methodology for the southern inland NSW would be a similar approach to northern NSW.

Looking forward requires consideration of impacts under plausible scenarios of future climate and estimates of future conditions using regional climate modelling.

NSW Government agencies, in particular the DPIE Environment, Energy and Science (EES) branch and its forerunner, the Office of Environment and Heritage (OEH), have spent many years developing future climate modelling capabilities and products. The NSW and ACT Regional Climate Modelling (NARCliM) project was developed in a partnership between the NSW and ACT governments and the Climate Change Research Centre at UNSW Sydney. NARCliM is a regional climate model (RCM) that dynamically downscales global climate models (GCMs). NARCliM has produced a set of climate projections with a 10 km grid resolution for south east Australia covering NSW, ACT, Victoria and the whole of the Murray Darling Basin, and it also provides a 50 km resolution for the whole of Australasia in line with the CORDEX Framework.

Current GCMs have some shortcomings, such as in characterising short- and long-term rainfall persistence, and this shortcoming is reflected in RCMs. To overcome this, the DPIE-Water team has developed and applied a statistically defensible approach to incorporating persistence in calculations of rainfall and evapotranspiration for past and present climates, by developing a stochastic model based on historical climate to establish current baseline climate. Climate change models for future climate projections can then be used to factor the stochastic model baseline. This approach forms the basis of the methodology being considered by this review.

The methodology is summarised in the draft Methods Paper provided by DPIE-Water (Appendix 2). The Methods Paper is underpinned by three papers that develop the extended climate sequences for the three regions (Lachlan Catchment, Macquarie Catchment and Bega River Region):

- *Multisite Rainfall and Evaporation Data Generation for the Macquarie Water Infrastructure Project* by Michael Leonard, Seth Westra and Bree Bennett from the University of Adelaide (Leonard, Westra, & Bennett, 2019) (Appendix 3)
- *Development of multi-site rainfall and evaporation data for the Lachlan Regional Water Strategy* by Danielle Verdon-Kidd at the University of Newcastle (Verdon-Kidd, 2019) (Appendix 4)

- *Incorporating changes in East Coast Low (ECL) behaviour into stochastically generated hydroclimatic data for the Bega River region, New South Wales, Australia* by Anthony Kiem at the University of Newcastle (Kiem, 2019) (Appendix 5)

Note the Panel only reviewed material available to it at the time and did not review material still under development. Other than the brief description in the Methods Paper, the Panel didn't have detail of the application of NARCliM models to the stochastic data sets. This is required for the Panel to make a proper assessment of the merits and potential weakness of its implementation.

Panel assessment

The findings and recommendations of this report follow two overarching themes:

- an assessment of the method and suggestions for improvements; and
- need for a community of practice to share expertise on complexities of the topic, new developments and its communication.

The use of stochastic models represents an important advance compared with the use of historical data and climate models alone. The Panel agrees with the overall approach of using stochastic models informed by knowledge of dominant climate drivers affecting different regions of NSW. The Panel notes that the paleo information captured in the current models is primarily used to better estimate the probability model of IPO periods, allowing a more representative distribution of IPO durations. Coupling NARCliM model outputs with the stochastic models enables the practitioner to understand the joint impact of climate variability and climate change. Projections of rainfall and evapotranspiration in future plausible climate scenarios can inform the risk likelihood and consequences of events, such as droughts as well as storage-filling wet events, and with consideration of uncertainty, inform decision making.

The products developed by DPIE-Water and its collaborators are data sets specific to different regions across NSW that can be used to inform decision making and preparedness. While they are based on model simulations, they enable DPIE-Water to develop robust risk-based and adaptive responses to climate variability and change.

Overall, DPIE Water's methodology to use observational and paleoclimate data informed by an understanding of climate drivers to select, calibrate and test stochastic models and the factoring of NARCliM projections is consistent with best-practice approaches to climate risk management. The Panel is of the view that ongoing improvement of the methodology adopted by DPIE be given high priority. This can be undertaken in parallel with concurrent RWS work. It would include an assessment of non-stationarity in the historical record to determine whether changes in climate in recent decades affect estimates of present-day climate risk compared with climate risk based on the whole observed record.

A further high priority recommendation of the Panel is for DPIE to convene a community of practice within DPIE, potentially broadening out, that maintains an ongoing work program to keep up-to-date with new scientific methods and findings in relation to hydrology/water availability and climate modelling (see Recommendation 10). The scientific literature on these interrelated topics are becoming ever more complex. The range of staff in DPIE who use climate and hydrology data and tools is wide, and they will need to stay up-to-date as the fields advance.

Specific comments on the methodology raised by the Panel are described and addressed below with relevant findings and recommendations, as well as corresponding report sections that discuss the issues in further detail.

Communication of the methodology (Section 2.1.1)

Recommendation 1: DPIE-Water to prepare a single document outlining the methodologies described in each of the background papers. Currently, the description of the methodology is distributed through the draft DPIE-Water Methods Paper, but the level of detail described in the relevant sub-project documents is not reflected in the Methods Paper. The Methods Paper should explain the various subprojects, including the commonalities and differences and the reasoning behind the method and choices made by the modelling experts. A useful example, in terms of the level of detail, was prepared for the Queensland Water Modelling Network, *Critical review of climate change and water modelling in Queensland* (Alluvium, 2019). **(High priority: Medium term)**[1]

Data (Section 2.2.1)

SILO and the Bureau of Meteorology's (BOM) Australian Water Availability Project (AWAP) use the same underlying data, drawn from the BOM Australian Data Archive for Meteorology (ADAM) database, but different methods to create these products. BOM makes available the AWAP data and other quality-controlled data which goes back to 1900. The different methods to develop SILO and AWAP can lead to variations in hydrological model outputs when each of the data sets are used (e.g. level of 'wet' and 'dry' seasonally and spatially) as pointed out in literature, including Lockart et al. (2016).

Recommendation 2: DPIE-Water to investigate the difference between the SILO and AWAP data sets and justification for use of one over the other should be provided, including efforts to quality control the data. This should be clearly documented by DPIE-Water. **(High Priority: Short term)**

Evapotranspiration (Section 2.2.2)

The DPIE-Water methodology uses a range of methods to estimate evaporation and evapotranspiration (ET). Of relevance for the methodology is how the calculations of potential evapotranspiration (PET) and ET are implemented and whether that implementation introduces an error of indeterminate sign or a systematic bias that grows over time.

Recommendation 3.1: The language in the DPIE-Water Methods Paper and the corresponding background documents relating to ET and PET should be clear and anywhere that PET is used needs to be carefully evaluated for the biases it might introduce. **(High priority: Medium term)**

Recommendation 3.2: DPIE-Water, in discussion with the community of practice, consider whether 3.1 could be notionally addressed by replacing the FAO56 and Mwet approaches with a physically based model. **(Medium/High priority: Long term)**

Paleo data (Section 2.2.3)

The Panel notes that the paleo information captured in the current models is primarily used to better estimate the probability model of IPO wet and dry run lengths, allowing a more representative distribution of IPO durations to be developed. As the driver of relevance for the Lachlan and Macquarie catchments was the Interdecadal Pacific Oscillation (IPO), paleo proxy records from the Pacific Ocean region (e.g. North America, China) were chosen for inclusion in the analysis. However, this paleo-data is sourced from distant locations that may be influenced by climate drivers in a manner that differs from NSW. If high quality local records can be obtained over time, the use of proxy records from less remote countries may be preferred.

Recommendation 4: DPIE-Water engages external expertise to explore options to improve proxy records by e.g. obtaining and incorporating local proxy records. This may go some way toward improving the situation where current proxy records are derived from distant locations with possibly different climate influences. Efforts to improve both the quality and quantity of proxy records could improve our understanding of climate variability. Moreover, the (possibly

[1] Recommendations are prioritised as high, medium, low priority and by timeframe - short, medium and long term. The noted timeframes are three months (short-term); one year (medium-term) and three years (long-term).

interactive) climatic variables that affect the proxies need to be clearly identified by experts. **(Medium priority: Long term)**

Generation of stochastic data sets (Section 2.3)

The use of stochastic models represents an important advance compared with the use of historical data and climate models alone.

Using stochastic models to develop long-term data sets of climate variables aims to address the fundamental limitation of using historical observations alone, namely that relatively short records are unlikely to contain the full range of natural climate variability, including drought periods with extremes of intensity, duration and multiple occurrences that can severely stress the water resource systems.

The fundamental proposition of stochastic models is that they make better use of the information contained in the historical record. It is for this reason that stochastic models have been widely adopted in drought risk assessment for major Australian urban areas.

Principles and justification for the use of stochastic models (Section 2.3.1)

The choice of the most appropriate stochastic model for a particular region and application is best left to a skilled analyst who *inter alia* will use a range of diagnostics to validate model assumptions and to demonstrate that it is fit-for-purpose. Independent review of the models is a key step to justify the choice and applications of the models. Section 2.3.1 provides discussion points for establishing these principles for the use of diagnostics.

Recommendation 5: DPIE-Water to work with experts to develop a statement of general principles about the use of diagnostics to aid in the evaluation of competing stochastic models and to enable water resource modellers to determine which models are fit for purpose. As part of this, the stochastic models should be subject to holdout validation so that their reliability can be more accurately assessed. **(High priority: Short term)**

Climate drivers – IPO and ECL (Sections 2.3.2, 2.3.3 and 2.3.4)

The Panel agrees with the overall approach of developing the stochastic models informed by knowledge of dominant climate drivers for each of the NSW regions and using the characteristics of these drivers to inform understanding of the historical statistical nature of rainfall and ET. The primary purpose for considering drivers is to extend the length of plausible past climate variability in the stochastic model by using the available paleo data.

Although there are uncertainties in estimating IPO phases from paleo data, the method nevertheless seems to provide a more conservative estimate of the likelihood of dry and wet sequences than if this approach using the IPO were not employed. It introduces an explicit mechanism for decadal variability to produce a wider range of stochastic possibilities than would be produced using stochastic models with shorter persistence. These are useful outcomes, even if the assumption upon which the method is based requires further research to be justified.

All regions in NSW will have contributions from multiple drivers and systems, so full explanations for climate’s impacts on rainfall and ET won’t be found in any single driver and it is inevitable that multiple interacting drivers could be involved. Looking forward with a changing climate, it is unlikely that the influences of climate drivers and their interactions will be stationary, so however these major drivers may have acted historically, they may not act consistently in the future.

In relation to the RWS methodology, looking at the effects of individual and combined (joint) drivers on rainfall could be particularly important for regions where multiple drivers could play dominant roles (e.g. both IPO and ECL).

Overall, the Panel agrees that work should be undertaken in the near term on what approach to use to generate stochastic data sets where multiple drivers dominate and for future iterations of the RWS, given that climate change may alter the impact of drivers (non-stationarity), we suggest future research put effort into other ways to generate or inform the stochastic data sets.

Recommendation 6.1: The Methods Paper would benefit from further articulation of the role (or not) that the climate drivers play specifically in the development of the stochastic data sets and how uncertainty about their future behaviour may be manifested into the scenarios. **(High priority: Long term)**

Recommendation 6.2: DPIE Water, in collaboration with experts in climate science and statistics, to explore alternative approaches to generating randomised samples as part of the future research program, particularly for regions where it is not clear what the dominant driver is or where there are multiple dominant drivers. **(Medium/High priority: Short, then Long term)**

Recommendation 6.3: DPIE-Water to work with colleagues, such as the community of practice, to examine whether the future behaviour of climate drivers (e.g. IOD, IPO, ECL, ENSO, and SAM) will remain statistically consistent with the past. Climate science can provide guidance on how these may change, and a review of the latest literature every few years would be prudent. **(High priority: Short term then ongoing)**

Climate Stationarity (Section 2.3.5)

The DPIE-Water approach assumes past climate stationarity, but the Panel notes evidence of an already changing climate, in particular from Western Australia and Victoria. It is judicious to also consider this within the stochastic data set for recent years and base climate modelling from this changing position. Other entities, including Melbourne Water, are incorporating data on parameters (e.g. temperature, rainfall, etc.) into work based on an already changing climate for the past couple of decades. If climatic conditions that form the basis of the methodology’s stochastic model baseline are set at a level that is lower than it should be, then the modelling could be underestimating risk.

Addressing these issues is the highest priority for the DPIE-Water method and a two-step approach is recommended.

Recommendation 7

DPIE-Water engages external expertise to undertake a two-step approach to investigate stationarity to manage the risk that models may underestimate current and future climate risk. The objective is to capture the current climate risk and a baseline that reflects this.

<u>First Step</u> **(High priority: Short term)**

The first step is to make a careful assessment of non-stationarity in the observed record. Without pre-empting such an assessment two useful methods are described.

<u>*Step 1 Method 1*</u>

Split-sample testing or hold-out validation splits the observed record into two parts. The stochastic model can be calibrated to each part and the sampling distributions of the parameters compared to determine if there is evidence of significant differences. Alternatively, the model can be calibrated to one part and its performance in the second part assessed using diagnostics based on predictive distributions.

This approach implicitly assumes the record used for calibration is stationary, so care is required in deciding how to split the observed record. It could be split as follows:

- *1990 reference* - the observed record could be split into pre- and post-1990 intervals
- *Drought reference* - the record used for validation and prediction could be chosen with regard to the Millennium Drought and the current drought, with three different calibration periods 1) from start of record (SOR) to the beginning of Millennium Drought; 2) from SOR to end of Millennium Drought; 3) from SOR to beginning of current drought; with the three prediction periods being from the end of the calibration period to today. A variation on this scheme is to constrain the model calibration periods to be the same length, so the earliest parts of the historical record are discarded for Run numbers 2) and 3)

Using both of these approaches (*1990 reference* and *drought reference)* could be prudent to confirm the outcomes.

Step 1_Method 2

Calibrate a stochastic model conditioned on an observed covariate associated with a change signal (e.g., temperature). While this approach is still in the research domain it is expected to have more statistical power to identify non-stationarity than the first method (provided a suitable covariate can be found).

Second Step **(High priority: Medium/Long term)**

If, after application of the first step, it is concluded that the recent past may not be statistically consistent with the full observed record and, more importantly, that differences are hydrologically significant, it will be necessary to revise the stochastic modelling approach described in the Methods Paper.

A number of conceptually different approaches are available:

Step 2_Approach 1

Calibrate the stochastic model to the recent past record with the knowledge that parameter uncertainty will be elevated;

Step 2_ Approach 2

Adjust the observed record prior to the recent past so that it reproduces certain statistical features of the recent past and then calibrate the stochastic model to the full adjusted observed record;

Step 2_ Approach 3

Identify suitable change covariates and calibrate a conditional stochastic model.

Each of these approaches has strengths and weakness. As there is limited experience with this issue it is expected that methods will evolve in the future.

Parameter uncertainty (Section 2.3.6)

Uncertainty in modelling can emerge from several sources, including parameter uncertainty.

The DPIE-Water methodology starts with a set of short-term observations (from SILO) and includes additional information from paleoclimate records, then calibrates stochastic models, adjusting model parameters so that modelled statistics match the statistics of the short-term observations. Because the observed record is short, there remains uncertainty, possibly considerable, about the true value of the stochastic model parameters.

Parameter uncertainty is an ‘inconvenient’ fact - it induces an uncertainty about the probability of future events, an uncertainty that is irreducible. It has been shown that this uncertainty can have considerable influence on uncertainty in water yield, a concept that has dominated urban water resource planning. The challenge is to inform decision makers of the implications of this uncertainty.

While statistical methods for quantifying parameter uncertainty are well developed, such information is rarely, if ever, used in decision making.

Recommendation 8: DPIE-Water, through consultation with external experts and the community of practice, to look at possible sources of parameter uncertainty in various stochastic models (and other models where relevant) and continue efforts to identify and document this uncertainty, including using statistical methods to quantify the uncertainty and mechanisms to incorporate this knowledge into decision making **(High priority: Long term)**

Incorporation of climate change into stochastic data sets and future directions (Sections 2.4 and 2.7)

Climate change modelling for NSW can benefit from RCMs that downscale GCMs. For example, the NARCliM ensemble of RCMs covers south eastern Australia and uses a fine-scale 10 km grid for future climate projections over 20-year time slices.

The NARCliM product has been developed for a range of industry end-uses, including water planning, infrastructure, environmental management. When applying GCM/RCM models, such as NARCliM, to considerations of water availability and hydrology, some shortcomings in the models related to persistence mean that additional analytical processes should be used to provide a more robust understanding of risk levels. For this reason, the DPIE-Water group have developed a method to incorporate a more representative picture of persistence in climate data through using stochastic models informed by observed and paleo proxy data.

The Methods Paper recommends using the GCM/RCM combination that produces the lowest near future rainfall, scaling the stochastic data calibrated to historical and paleo data (presumably using the ratio of the daily means as the scaling factor) and stress testing the water resource system to assess its water security vulnerability. The Panel considered this a relatively cautious approach to incorporating climate change signals into the stochastic data.

The NARCliM modelling system is soon to be upgraded (NARCliM 1.5) using more recent GCMs and longer timeframes for projections. This follows feedback from practitioners and other end-users and will assist in improving utility. In addition, another update of NARCliM (2.0) is planned.

Using the stochastic data sets that have been mathematically generated and informed by localised dominant climate drivers provides a statistical picture of recent past climate. Incorporating projections from climate models into the stochastic models enables the practitioner to bring together climate variability with climate change projections. Projections of rainfall and evapotranspiration in future plausible climate scenarios can inform the risk likelihood and consequences of events, such as droughts and storage-filling wet events, and in so doing, with consideration of uncertainty, inform decision making. However, it is important to note that projections of future climate cannot capture the full range of future possibilities and therefore consideration of future scenarios outside those reflected in existing GCM/RCM simulations is worthwhile.

Using GCM/RCM simulations needs to be undertaken with care. The limited number of GCMs, differences between model simulation, model independence issues and other uncertainties need to be understood. As recommended below, an interaction between the end-users of the simulations and the climate modellers will provide insights into how best to use model products. This interaction can also explore:

- what other approaches can be used to evaluate the models’ performance, such as comparing performance with other tools and lines of evidence.

- selecting GCM/RCM scenarios in a conservative way, such as the approach used by DPIE Water, to favour lowest near-future rainfall scenarios in order to ensure simulations generate examples of low-probability, high-impact droughts.

Ongoing collaboration between developers and end-users will be valuable to ensuring the best outcomes from the models.

Recommendation 9.1: DPIE-Water work with the NARCliM developers and together begin the process of planning to incorporate NARCliM 1.5 into calculations. **(High priority: Short term)**

Recommendation 9.2: The community of practice monitor approaches used to quantify future climate risk elsewhere in Australia and internationally to ensure methods used in NSW remain at an international standard. **(Medium priority: Medium term then ongoing)**

Recommendation 9.3: DPIE-Water to work with the community of practice to explore incorporation of NARCliM 2.0. **(Medium priority: Long term)**

Forming a community of practice (Section 2.7.1)

The NSW Government has made considerable investment into tools and expertise in climate modelling and using model outputs to inform impact assessments in infrastructure, finance, agricultural, environment, emergency services and other fields. New models continue to be released with increasing resolution and performance, followed by new assessments of climate change impacts, vulnerability and adaptive management planning. These research fields are large and complex and will continue to evolve with time.

A community of practice should be convened to discuss best approaches for communicating complexities of modelling, uncertainty and related issues. The community of practice would also promote and facilitate collaboration between hydrological, hydrogeological and climate modellers and climate scientists more broadly, and would help ensure mutual understanding develops.

The community of practice could focus on a range of work, including considering new and improved ways to check the integrity of data; assessing new model developments and applications; developing collaborative projects that take full advantage of the available expertise and delivering triple-bottom-line outcomes; coordinating advice to government; etc.

Some recommended focus areas include:

a) **Identifying best practice** - Within DPIE, understanding of climate impacts into the future will be important across a range of policy areas, including in water management, ecosystem management, water planning, agriculture, infrastructure, etc. The community of practice can assist in developing a shared understanding of new scientific directions, the applications of new models into practice, experience of using different modelling approaches, as well as policy developments and their implications for NSW. It can also help the fields of hydrology and climate science to better align within DPIE's activities.

b) **Understanding uncertainty** - There is deep unquantifiable uncertainty and hence low confidence in estimates of the probability of some future climate events, particularly rare events. Where possible, multiple approaches to assessing future climate risk could be considered and adopted by the community of practice work program as relevant. Assessing the magnitude of this uncertainty is an on-going challenge. While quantifying the uncertainty is not currently a robust area of climate science, there are ways to protect against systematic biases.

c) **Communicating and managing uncertainty** - Approaches for communicating and managing risk in the face of uncertainty should be a focus area. Central to this is recognising the different types and sources of uncertainty and establishing an adaptive planning approach that builds uncertainty into decisions making, to account for an emerging knowledge base. Therefore, the primary focus of adaptation planning should be a preparedness and resilience rather than a prediction paradigm.

d) **Adaptive management** - Best practice approaches to adaptive management, where the system facilitates ongoing improvement, stress testing, comparing real world and modelled scenarios.

e) **Assessing new approaches** - Regularly monitor developments in fields of climate science and hydrological modelling, including new conditional stochastic modelling approaches that will emerge. These approaches can then be scrutinised, and a view formed about whether and how they should be considered for use in NSW.

f) **Contributing to technical improvements** - The developers of the RWS and NARCliM could engage on a defined and on-going basis to develop a common understanding of the strengths and weaknesses of alternative approaches and write a joint white paper on whether, and if so, how to merge the approaches. This should be completed by the end of 2020 so it can potentially feed into NARCliM 2.0.

g) **New approaches to ET** – Implementing physical approaches to ET and PET (e.g. replacing current approaches for FAO56 and Mwet).

h) **Observational networks and monitoring -** There should be a long-term commitment to identifying and maintaining reference/benchmark observational networks. Maintenance of regional observational networks with long-term, high-quality records is needed to facilitate detection of emerging climate-related trends; evaluation of predictive models under changing hydroclimatic conditions; and performance assessments for potential adaptation strategies.

i) **Monitoring climate trends** - Track and monitor emerging or relevant climate trends (e.g. if winter rainfall in South-Eastern Australia/NSW is decreasing) many of which are provided by the Bureau of Meteorology : http://www.bom.gov.au/climate/change/#tabs=Tracker&tracker=timeseries

j) **Research –** Identify and potentially commission research to support new knowledge.

k) **National cooperation –** promote cross-jurisdictional dialogue in relation to hydro-climatological practices with water management agencies in other jurisdictions nationally to exchange knowledge and experiences.

Recommendation 10: DPIE to convene a community of practice on a defined and ongoing basis that includes at least relevant climate science and user groups within NSW DPIE (including the broader cluster entities - Manly Hydraulics Laboratory, Water NSW, Sydney Water, Hunter Water, DPIE-Water, EES, Department of Primary Industries) and possibly other agencies that have forecasting and response roles in NSW such as Rural Fire Services, Fire and Rescue NSW, NSW Health. In addition, experts from other entities, including universities, BOM and CSIRO should be invited to participate. **(High priority: Short term, then ongoing)**

RECOMMENDATIONS

Rec #	Recommendation	Priority High Medium Low	Timeframe Short (3 months) Medium (1 year) Long (3 years) Ongoing
1	DPIE-Water to prepare a single document outlining the methodologies described in each of the background papers. Currently, the description of the methodology is distributed through the draft DPIE-Water Methods Paper, but the level of detail described in the relevant sub-project documents is not reflected in the Methods Paper. The Methods Paper should explain the various subprojects, including the commonalities and differences and the reasoning behind the method and choices made by the modelling experts. A useful example, in terms of the level of detail, was prepared for the Queensland Water Modelling Network, *Critical review of climate change and water modelling in Queensland* (Alluvium, 2019).	High	Medium
2	DPIE-Water to investigate the difference between the SILO and AWAP data sets and justification for use of one over the other should be provided, including efforts to quality control the data. This should be clearly documented by DPIE-Water.	High	Short
3.1	The language in the DPIE-Water Methods Paper and the corresponding background documents relating to ET and PET should be clear and anywhere that PET is used needs to be carefully evaluated for the biases it might introduce.	High	Medium
3.2	DPIE-Water, in discussion with the community of practice, consider whether 3.1 could be notionally addressed by replacing the FAO56 and Mwet approaches with a physically based model.	Medium/High	Long
4	DPIE-Water engages external expertise to explore options to improve proxy records by e.g. obtaining and incorporating local proxy records. This may go some way toward improving the situation where current proxy records are derived from distant locations with different climate influences. Efforts to improve both the quality and quantity of proxy records could improve our understanding of climate variability. Moreover, the (possibly interactive) climatic variables that affect the proxies need to be clearly identified by experts.	Medium	Long
5	DPIE-Water to work with experts to develop a statement of general principles about the use of diagnostics to aid in the evaluation of competing stochastic models and to enable water resource modellers to	High	Short

	determine which models are fit for purpose. As part of this, the stochastic models should be subject to holdout validation so that their reliability can be more accurately assessed.		
6.1	The Methods Paper would benefit from further articulation of the role (or not) that the climate drivers play specifically in the development of the stochastic data sets and how uncertainty about their future behaviour may be manifested into the scenarios.	High	Long
6.2	DPIE Water, in collaboration with experts in climate science and statistics, to explore alternative approaches to generating randomised samples as part of the future research program, particularly for regions where it is not clear what the dominant driver is or where there are multiple dominant drivers.	Medium/High	Short, then Long
6.3	DPIE-Water to work with colleagues, such as the community of practice, to examine whether the future behaviour of climate drivers (e.g. IOD, IPO, ECL, ENSO, and SAM) will remain statistically consistent with the past. Climate science can provide guidance on how these may change, and a review of the latest literature every few years would be prudent.	High	Short then ongoing
7	DPIE-Water engages external expertise to undertake a two-step approach to investigate stationarity over the historical record to ensure the risk that models do not underestimate current and hence future climate risk. The objective is to capture the current climate risk and a baseline that reflects this. <u>First Step</u> The first step is to make a careful assessment of non-stationarity in the observed record. Without pre-empting such an assessment two useful methods are described. *<u>Step 1_Method 1</u>* Split-sample testing or hold-out validation splits the observed record into two parts. The stochastic model can be calibrated to each part and the sampling distributions of the parameters compared to determine if there is evidence of significant differences. Alternatively, the model can be calibrated to one part and its performance in the second part assessed using diagnostics based on predictive distributions. This approach implicitly assumes the record used for calibration is stationary, so care is required in deciding how to split the observed record. It could be split as follows: • *1990 reference* - the observed record could be split into pre- and post-1990 intervals • *Drought reference* - the record used for validation and prediction could be chosen with regard to the Millennium Drought and the current drought, with three different calibration periods 1) start of record (SOR) to beginning of Millennium Drought; 2) SOR to end of Millennium Drought; 3) SOR to beginning of current drought; with the three prediction periods being from the end of the calibration	Step 1: High Step 2: High	Step 1: Short Step 2: Medium/Long

	period to today. A variation on this scheme is to constrain the model calibration periods to be the same length, so the earliest parts of the historical record are discarded for Run numbers 2) and 3) Using both of these approaches (*1990 reference* and *drought reference)* could be prudent to confirm the outcomes. *Step 1_Method 2* Calibrate a stochastic model conditioned on an observed covariate associated with a change signal (e.g., temperature). While this approach is still in the research domain it is expected to have more statistical power to identify non-stationarity than the first method (provided a suitable covariate can be found). Second Step If after application of the first step it is concluded that the recent past may not be statistically consistent with the full observed record and, more importantly, that differences are hydrologically significant, it will be necessary to revise the stochastic modelling approach described in the Methods Paper. A number of conceptually different approaches are available: *Step 2_Approach 1* Calibrate the stochastic model to the recent past record with the knowledge that parameter uncertainty will be elevated; *Step 2_ Approach 2* Adjust the observed record prior to the recent past so that it reproduces certain statistical features of the recent past and then calibrate the stochastic model to the full adjusted observed record; *Step 2_ Approach 3* Identify suitable change covariates and calibrate a conditional stochastic model. Each of these approaches has strengths and weakness. As there is limited experience with this issue it is expected that methods will evolve in the future.		
8	DPIE-Water, through consultation with external experts and the community of practice, to look at possible sources of parameter uncertainty in various stochastic models and continue efforts to identify and	High	Long

	document this uncertainty, including using statistical methods to quantify the uncertainty and mechanisms to incorporate this knowledge into decision making.		
9.1	DPIE-Water works with the NARCliM developers and together begin the process of planning to incorporate NARCliM 1.5 into calculations.	High	Short
9.2	The community of practice monitor approaches used to quantify future climate risk elsewhere in Australia and internationally to ensure methods used in NSW remain at an international standard.	Medium	Medium then ongoing
9.3	DPIE-Water to work with the community of practice to explore incorporation of NARCliM 2.0.	Medium	Long
10	DPIE to convene a community of practice on a defined and ongoing basis that includes at least relevant climate science and user groups within NSW DPIE (including the broader cluster entities - Manly Hydraulics Laboratory, Water NSW, Sydney Water, Hunter Water, DPIE-Water, EES, Department of Primary Industries) and possibly other agencies that have forecasting and response roles in NSW such as Rural Fire Services, Fire and Rescue NSW, NSW Health. In addition, experts from other entities, including universities, BOM and CSIRO should be invited to participate. Some recommended focus areas include: a) **Identifying best practice** - Within DPIE, understanding of climate impacts into the future will be important across a range of policy areas, including in water management, ecosystem management, water planning, agriculture, infrastructure, etc. The community of practice can assist in developing a shared understanding of new scientific directions, the applications of new models into practice, experience of using different modelling approaches, as well as policy developments and their implications for NSW. It can also help the fields of hydrology and climate science to better align within DPIE's activities. b) **Understanding uncertainty** - There is deep unquantifiable uncertainty and hence low confidence in estimates of the probability of some future climate events, particularly rare events. Where possible, multiple approaches to assessing future climate risk could be considered and adopted by the community of practice work program as relevant. Assessing the magnitude of this uncertainty is an on-going challenge. While quantifying the uncertainty is not currently a robust area of climate science, there are ways to protect against systematic biases. c) **Communicating and managing uncertainty** - Approaches for communicating and managing risk in the face of uncertainty should be a focus area. Central to this is recognising the different types and sources of uncertainty and establishing an adaptive planning approach that builds uncertainty	High	Short then ongoing

	into decisions making, to account for an emerging knowledge base. Therefore, the primary focus of adaptation planning should be a preparedness and resilience rather than a prediction paradigm. d) **Adaptive management** - Best practice approaches to adaptive management, where the system facilitates ongoing improvement, stress testing, comparing real world and modelled scenarios. e) **Assessing new approaches** - Regularly monitor developments in fields of climate science and hydrological modelling, including new conditional stochastic modelling approaches that will emerge. These approaches can then be scrutinised, and a view formed about whether and how they should be considered for use in NSW. f) **Contributing to technical improvements** - The developers of the RWS and NARCliM could engage on a defined and on-going basis to develop a common understanding of the strengths and weaknesses of alternative approaches and write a joint white paper on whether, and if so, how to merge the approaches. This should be completed by the end of 2020 so it can potentially feed into NARCliM 2.0. g) **New approaches to ET –** Implementing physical approaches to ET and PET (e.g. replacing current approaches for FAO56 and Mwet). h) **Observational networks and monitoring -** There should be a long-term commitment to identifying and maintaining reference/benchmark observational networks. Maintenance of regional observational networks with long-term, high-quality records is needed to facilitate detection of emerging climate-related trends; evaluation of predictive models under changing hydroclimatic conditions; and performance assessments for potential adaptation strategies. i) **Monitoring climate trends** - Track and monitor emerging or relevant climate trends (e.g. if winter rainfall in South-Eastern Australia/NSW is decreasing) and how the duration of climate trends in the future (i.e. longer periods of dry or wet) are captured in the modelling approaches. Many of the climate trends are provided by the Bureau of Meteorology: http://www.bom.gov.au/climate/change/#tabs=Tracker&tracker=timeseries j) **Research –** Identify and potentially commission research to support new knowledge. k) **National cooperation –** promote cross-jurisdictional dialogue in relation to hydro-climatological practices with water management agencies in other jurisdictions nationally to exchange knowledge and experiences.		

1 INTRODUCTION

In September 2019, the Minister for Water, Property and Housing, the Hon Melinda Pavey MP, requested that the Deputy Chief Scientist & Engineer chair an independent expert panel (the Panel) to provide advice on the incorporation of climate risk into the development of Regional Water Strategies (RWS).

The NSW Government is developing twelve RWS to determine the best policy, planning and infrastructure strategies to meet longer-term water needs for communities, the environment, aboriginal cultural values and rights, and industry (Figure 1). The first RWS for the Greater Hunter was released in November 2018 (DOI, 2018),

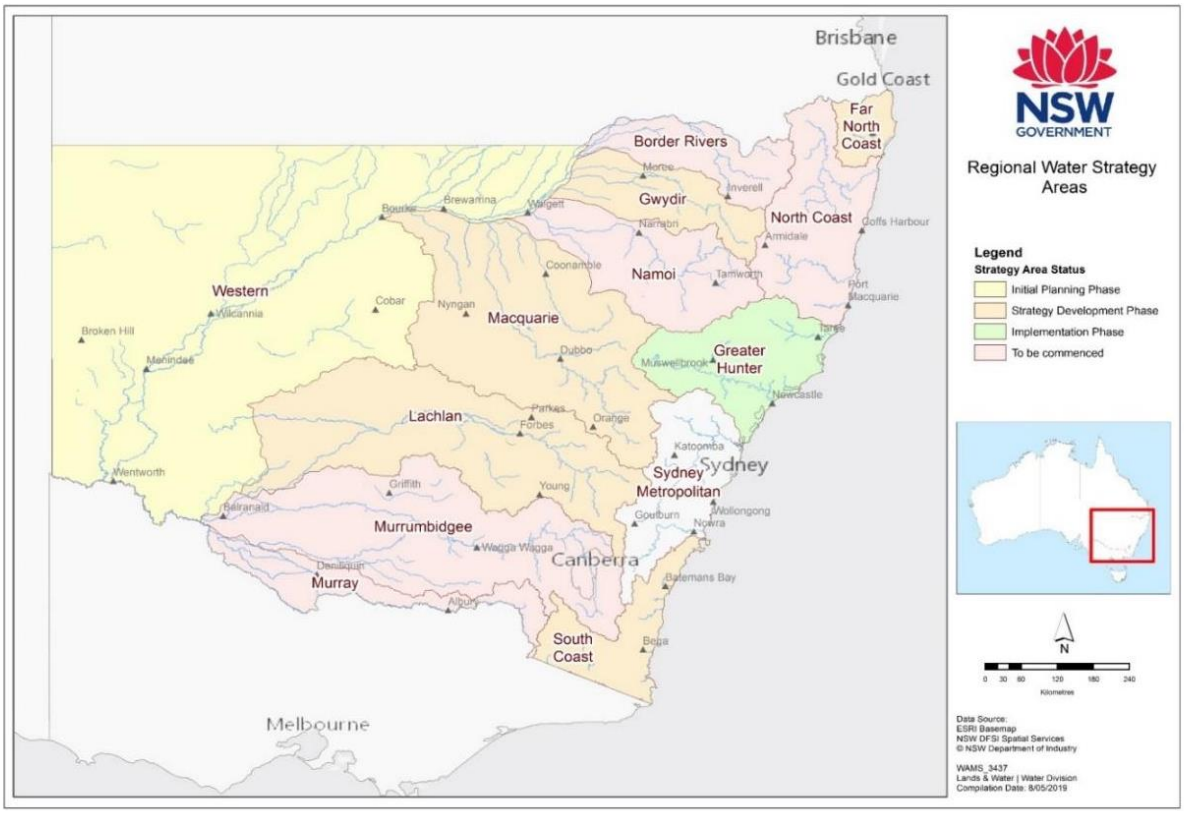

Figure 1: Regional water strategy areas

Source: DPIE-Water (presentation to the Panel)

Consistent with recommendations from the NSW State Infrastructure Strategy 2018-2038 (Infrastructure NSW, 2018), the RWS program aims to incorporate knowledge about climate variability and change in assessing risks under future scenarios and managing challenges, particularly drought.

Looking forward requires consideration of impacts under plausible scenarios of future climate and estimates of future conditions using regional climate modelling. The future scenarios and projections of risk covered in the RWS datasets are for the period over the next 20 to 40 years. The strategies will determine future water demand of a region, as well as setting out the challenges and choices involved in meeting those regional needs.

Historically, water management decisions have drawn on observed data, including rainfall and evaporation. However, the short history of recorded climate data and the uncertainty of a changing climate mean that the risk of increasingly severe droughts as well as other extreme weather events and changes in seasonality of rainfall is not well captured in models, a data limitation that could affect water security. Further, the continued emission of carbon dioxide and other radiatively active gases and associated climate changes mean that the future cannot be solely characterised by the past.

The RWS aims to address these data limitations by incorporating a statistically valid representation of past climate sequences and scenarios for future climate change into hydrological and other models used as the basis for policy decisions. The steps taken by DPIE-Water and their rationale are set out in a draft Methods Paper, *Developing climate data sets for use in climate risk assessment for Regional Water Strategies* (Appendix 2).

The Panel was established to review and provide expert advice on the suitability of the methods used to produce the extended climatic sequences and the applicability of the use of these datasets in determining long-term water security solutions. The Panel was requested to make recommendations about improvements to methods to incorporate new climate change data. The focus of the efforts of the Panel have been on the methods deployed in predicting future climate-impacted rainfall and evapotranspiration in three NSW catchments: Lachlan River catchment, Macquarie River catchment and the south coast region centred on Bega. The full scope of work is at Appendix 1.

The Panel included expertise in climate drivers, variability and change, extreme weather events, hydrology and statistics and was comprised of:

- Dr Chris Armstrong, Deputy NSW Chief Scientist & Engineer (Chair)
- Professor Bryson Bates, Adjunct Professor, School of Agriculture and Environment, University of Western Australia
- Emeritus Professor George Kuczera, School of Engineering, The University of Newcastle
- Professor Andy Pitman, Director, Australian Research Council Centre of Excellence for Climate Extremes, Climate Change Research Centre, UNSW Sydney
- Dr Scott Power, Senior Principal Research Scientist, Bureau of Meteorology.

1.1 INFORMATION PROVIDED TO THE PANEL

The Panel was provided with a number of documents to inform this Review, including:

- Draft Methods Paper provided by DPIE-Water (Appendix 2)
- Three project papers to develop the extended climate sequences for the three regions:
  - *Multisite Rainfall and Evaporation Data Generation for the Macquarie Water Infrastructure Project* by Michael Leonard, Seth Westra and Bree Bennett from the University of Adelaide (Leonard et al., 2019) (Appendix 3)
  - *Development of multi-site rainfall and evaporation data for the Lachlan Regional Water Strategy* by Danielle Verdon-Kidd at the University of Newcastle (Verdon-Kidd, 2019) (Appendix 4)
  - *Incorporating changes in East Coast Low (ECL) behaviour into stochastically generated hydroclimatic data for the Bega River region, New South Wales, Australia* by Anthony Kiem at the University of Newcastle (Kiem, 2019) (Appendix 5)

Several presentations were given to the Panel on 3 and 16 October 2019 by DPIE-Water about its methods (Appendix 6)[2], presentations to the Panel on 24 October 2019 by academic experts that had generated stochastic data to represent historical variability and tested the models under various scenarios in Bega region (Kiem, 2019) and the Macquarie Valley region (Leonard

[2] Two presentations were made as not all panel members were able to attend the first meeting at the same time.

et al., 2019). A separate presentation on work in the Lachlan region was provided to the Panel Chair (Leonard et al., 2019).

In response to questions asked by the Panel, DPIE-Water provided additional advice:

- information about climate data used by DPIE-Water Modelling Unit (Appendix 7).
- documents in relation to the implementation of Integrated Quantity and Quality Model (IQQM) software used for surface water modelling, particularly related to how flow is estimated using climate data:
  - Data Assessment and Headwater Calibration Guidelines (undated) (Appendix 8)
  - Streamflow and Climate Data Review - Murrumbidgee Valley (undated) - a report to: *"review flow and climate data (streamflow, rainfall and evaporation) in the Murrumbidgee Valley for development of headwater rainfall runoff models and flow models for residual reaches. The review covers all available data of streamflow and climate gauges that are used in the Murrumbidgee Source modelling*."
  - Headwater Flow Calibration Report – Murrumbidgee Valley – Gauging Station 410025 Jugiong Creek (last dated 19 June 2018) - a report to *"establish the rainfall-runoff model for the headwater catchment above the streamflow gauge 410025 (Jugiong Ck at Jugiong)."*
  - Headwater Flow Calibration Report – Murrumbidgee Valley – Gauging Station 410061 – Adelong Creek @ Batlow Road (last dated 24 Sept 2018) - a report to *"establish the rainfall-runoff model for the headwater catchment above the streamflow gauge 410061 (Adelong Creek @ Batlow Road)."*

In addition to NSW developments, there are a range of initiatives at national and jurisdictional level to develop climate information for decision making that the Panel was aware of. These are summarised at Appendix 9.

1.2 SUMMARY OF METHODS

The method proposed in the Methods Paper to assess climate risk incorporates the key climate drivers operating in the regions of NSW. These are dominant weather and climate systems operating in and around Australia that either independently or in combination affect the probability of or amount of rain. The dominant drivers affecting NSW include the El-Niño Southern Oscillation (ENSO)/ La Niña, Inter-decadal Pacific Oscillation (IPO), Indian Ocean Dipole (IOD), Southern Annular Mode (SAM) and East Coast Lows (ECL). A background on these climate drivers is at Appendix 10.

The methodology for the Macquarie and Lachlan regions is based on the IPO (Leonard et al., 2019; Verdon-Kidd, 2019) as the dominant driver, with the Bega region based on ECLs (Kiem, 2019). For the IPO dominated regions, the method uses observed climate data and paleo reconstruction data to develop extended past climatic sequences (stochastic data sets) of rainfall and pan evaporation/ potential evapotranspiration over a 10,000 year timeframe. Climate change risk is then factored in by DPIE-Water using NARCliM 1.0 data. For the Bega region, climate change analysis was based on climate modelling derived changes to ECL seasonal frequency. These processes are illustrated in Figure 2 and described in detail in the background papers (Appendices 3-5).

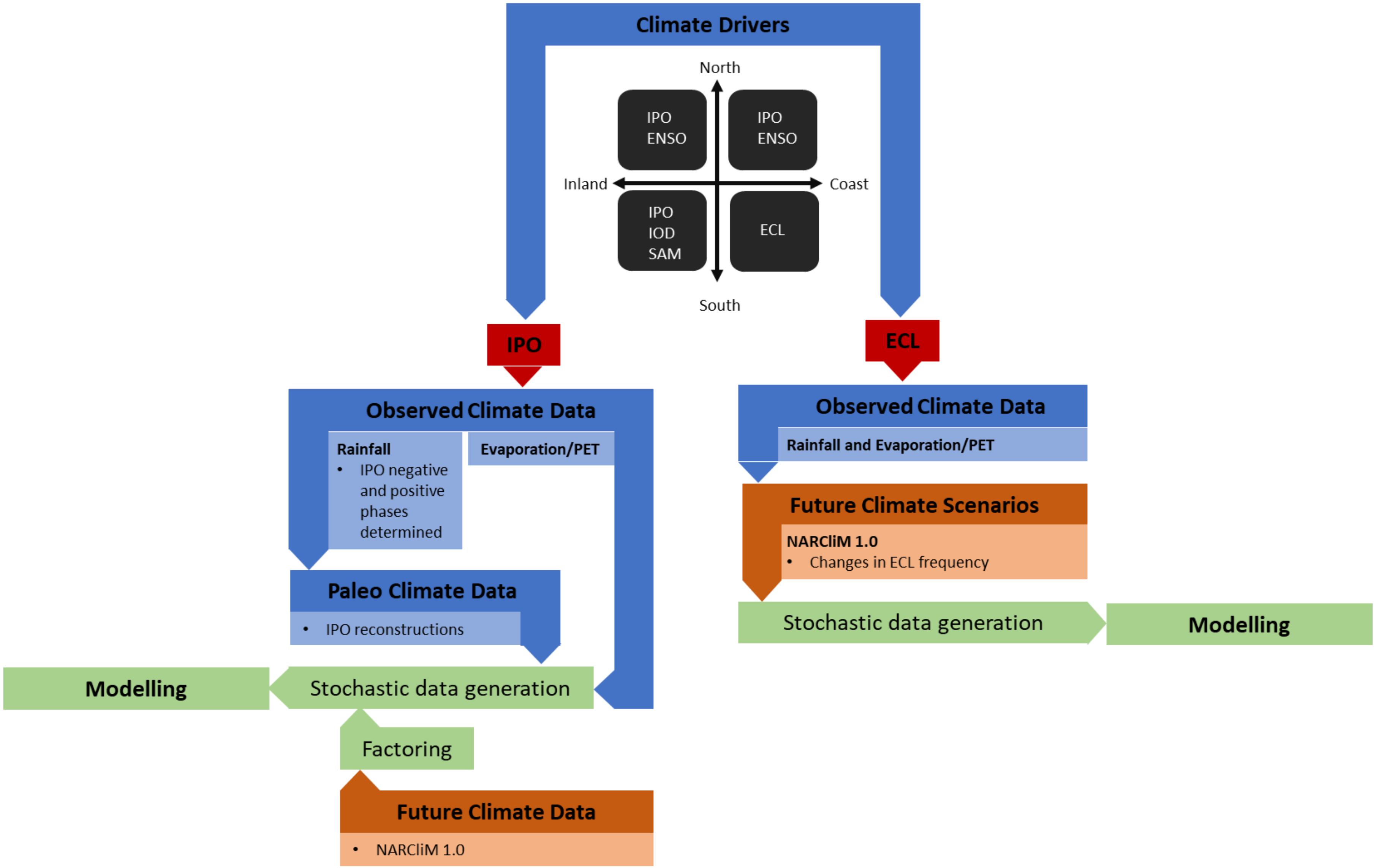

Figure 2: Overview of the development of the extended (stochastic) data sets

2 COMMENTS ON THE DPIE-WATER METHODOLOGY

This Chapter sets out the Panel's general observations about the approach used by DPIE-Water, followed by more specific comments about the methods including use of historical data, development of stochastic data sets, and incorporation of climate variability and change in water modelling, including sources of uncertainty and limitations in the method.

The observations refer to methods laid out in the covering document *Developing climate data sets for use in climate risk assessment for Regional Water Strategies (draft) version 2.0* (the Methods Paper) by DPIE-Water Branch - Water Modelling Unit (2019) as well as the three background papers describing methods to develop the stochastic data sets noted in Section 1.1 (Appendices 3-5).

The commentary on the application of the NARCliM climate predictions relies on the description in the Methods Paper by DPIE-Water.

The Panel has provided some additional observations on decision making under uncertainty (Appendix 11) that go beyond the terms of reference (TOR). Nevertheless, the Panel felt that these observations were worth noting.

2.1 GENERAL OBSERVATIONS ABOUT THE DPIE-WATER METHODOLOGY

2.1.1 Lack of technical detail in Methods Paper

Different teams within the university sector and within NSW Government have contributed to the work under review. The modelling and analysis on three valley catchments (Bega, Macquarie and Lachlan) have been reviewed by the Panel with the specific work undertaken by three university teams (led by Kiem (2019), Leonard et al. (2019) and Verdon-Kidd (2019)) and brought together in the Methods Paper by the DPIE-Water group. The DPIE-Water group undertook the work to apply the regional climate model to the rainfall and evapotranspiration data from the three catchment studies.

The Panel was provided with documents described in the preceding chapter, with the draft Methods Paper that was written by DPIE-Water being the main overarching summary of the full methodology.

The Methods Paper provides a high-level summary but refers to the reference list for the reader to seek out the detail on the background and the specific information on the approaches taken. The full reference list is quite extensive, thus is not a useful starting point for the reader to better understand specific aspects of the methodology.

The three papers by Kiem (2019), Leonard et al. (2019), and Verdon-Kidd (2019) provide some detail required to better understand the use of measured data, application of paleoclimate proxies and the development of the stochastic data sets for the three regions. However, these are complicated processes and should be reflected in the Methods Paper; too much is left to the background papers for the reader to get a clear sense of what is being done.

A useful example, in terms of the level of detail, was prepared for the Queensland Water Modelling Network, *Critical review of climate change and water modelling in Queensland* (Alluvium, 2019).

In addition, not all the methods listed in the background papers, specifically the factoring of climate models into the stochastic data sets reflected in Leonard et al. (2019), have been adopted in the final methodology set out in the Methods Paper. DPIE-Water noted that the factoring in Leonard et al. (2019) was an early expression of how to include climate change that was subsequently not used in preference for NARCliM. It would be helpful for this rationale to be explained in the Methods Paper.

Another element that should be set out in the Methods Paper, when identifying the approach taken for different regions of the state (e.g. north and south, inland and coastal), is the differences in climatological and hydrological changes across different regions and how that affects the approach taken, particularly related to non-stationary of climate discussed in Section 2.3.5.

The Methods Paper also provides little information regarding how NARCliM results are factored into the stochastic data sets. This is required for the Panel to make a proper assessment of the merits and potential weakness of its implementation. The same applies to the quality assurance of the data. These issues are discussed below.

2.1.2 General observations about methodology

Quantifying future hydroclimatic conditions for a region due to climate change is inherently difficult. Multiple models and sources of data are brought into the process, leading to uncertainty that needs to be managed going forward.

While analysis and comments on the specific aspects of the DPIE-Water methodology follow in the next sections, the Panel has made a series of general observations relevant to the approach. These include:

- Overall, the DPIE Water's methodology to use observational and paleoclimate data informed by an understanding of climate drivers to select, calibrate and test stochastic models and the factoring of NARCliM projections is consistent with best-practice approaches to climate risk management. The Panel's comments on specific issues follow in the next section.
- There is no single approach that can be considered as the 'best' under all circumstances (e.g. differences between location, dominant climate drivers, etc.) and currently available approaches have their own strengths and limitations. Therefore, having access to a suite of tools that can be selected from, with the most appropriate deployed to model a catchment based on local characteristics and information availability is important. The reasoning behind methodology choices should be clearly explained, as well as understanding the pitfalls in the method of choice and selecting additional approaches to address the pitfall where possible.
- It is good practice to run several different scenarios to understand the likelihood of events occurring, including stress testing outlier possibilities.
- It should be emphasised with the users of the data that scenarios beyond those modelled are possible.
- It is good practice to have a system that facilitates ongoing improvement, stress testing, comparing measured data with modelled scenarios.
- The skill of GCMs will generally increase with each new generation of models but rigorous assessment of these models' performance for the variables important for water security will be necessary. Tailored climate change projections are required to ensure that planning decisions are based on the latest findings of climate science.
- Water and climate science modelling should be integrated. This will require collaboration to ensure that climate modellers know what is needed and can directly examine model performance. For example, there is relatively little work on what brings drought breaking rains, etc. The issue is in part due to climate and water sciences being separated by utilisation of very different methodologies and tools.

- Uncertainty
    - Uncertainty is discussed in several places within this report, as one might expect, as managing uncertainty is a fundamental aspect of a complex decision-making environment.
    - Because instrumental records are short, the parameters in stochastic models are subject to considerable uncertainty. As this uncertainty is propagated through water planning models it produces uncertainty in key performance metrics such as water yield or risk of critically low supplies.
    - The ability to estimate future climate risk in the context of water supply is subject to considerable uncertainty.
    - For risk-averse decision makers communicating this uncertainty is important. Because droughts are widespread, if drought risks are underestimated, it will be so over a large region.
    - Knowledge about the climate system will continue to increase, but the uncertainty may not decrease as this occurs. Assessing the magnitude of this uncertainty is an on-going challenge. While quantifying the uncertainty is not currently a robust area of climate science, there are ways to protect against systematic biases. Central to this is recognising the different types and sources of uncertainty and establishing an adaptive planning approach that builds uncertainty into decision making, to account for an emerging knowledge base. Therefore, the primary focus of adaptation planning should be a preparedness and resilience rather than a prediction paradigm. Development of the stochastic models through this process enables a better ability to do this than in the past.
    - Uncertainty doesn’t need to prevent decisions being made; managing uncertainty is quite an advanced field of endeavour and research. Gathering further data and analysis, adopting conservative or flexible approaches and applying adaptive management are examples of ways to manage uncertainty. An ongoing focus for the community of practice will be to better understand uncertainty in this context and to effectively communicate and manage it. Refer to Appendix 11 for further discussion of uncertainty and adaptive management in the water planning context.

2.2 DATA

2.2.1 Sources of observed/historical data

The historic record of climate data is a mix of genuine observations and values inferred by interpolation algorithms using observations at other locations.

Over time, the range of different data sets, as well as their granularity, is increasing. In addition, different sources are infilling data missing from the observed records to create more complete data sets over a period.

Historically, BOM observed data was available for use, but there were limitations in accessing and/or using it. In response, SILO was developed by the Queensland Government drawing from BOM data and using infilling techniques to create full data sets of daily rainfall and evaporation, evapotranspiration (ET) and potential evapotranspiration (PET) going back 130 years. Since SILO was released, the BOM has also made available the Australian Water Availability Project (AWAP) data which goes back to 1900 (Data.gov.au, 2019).

More recently, BOM developed BARRA (BOM Atmospheric high-resolution Regional Reanalysis for Australia). BARRA represents 100 meteorological parameters at a 12 km grid resolution for Australia and NZ as well as 1.5 km resolution for much of NSW. BARRA uses a weather model to infill detail on atmospheric and surface variables. The parameters are represented in hourly intervals (some at 10 minutes) over a period of 29 years (1990 – 2018), which can be aggregated to daily or monthly. As a meteorological reanalysis looking historically rather than to future, its application to this setting is not scrutinised here.

The Methods Paper references both rainfall data from BOM and evaporation/ET data from SILO. These were used as input to develop stochastic data sets covering 10,000 years. It appears from the three background papers (Kiem (2019); Leonard et al. (2019); Verdon-Kidd (2019)) that rainfall and estimated evaporation/ET data covering approximately 130 years (1889 – present), used across the three catchments, is sourced from SILO. Justification to use the SILO data rather than AWAP was not explicit in the Methods Paper, which led to Panel consideration of the 'best' source of the data for the purposes at hand.

Reasons why SILO data might be used include:

- it was the first available gridded product,
- it has a slightly broader range of variables than available through AWAP
- it offers capability to drill down for data at a specific point, and
- it is relatively straightforward to obtain grids for extended times.

BOM AWAP data is quality controlled and offers a gridded product.

SILO and AWAP use the same underlying data but different methods to create their products. This can lead to variations in the data sets (e.g. level of 'wet' and 'dry' seasonally and spatially) as pointed out in literature, including Lockart et al. (2016).

Both AWAP and SILO use interpolation algorithms, with the consistency of the interpolation being affected by changes in observation networks. For example, if a rain-gauge close to a grid point drops out, the interpolation algorithm will then rely on gauges further away. This has the potential to introduce shifts in the time series that are numerical artefacts. Producing a consistent long time series is challenging. Efforts to quality control for these consistency challenges should be made.

The Panel sought clarity from DPIE-Water on the quality assurance and quality control approach deployed with the SILO data they used, given the infilling treatment that it undergoes. DPIE-Water's response is reproduced as follows:

- *"SILO Patch Point data has been adopted as the standard for rainfall data. Hydrologists can use other sources if they cannot find a satisfactory SILO rainfall site. The patch point dataset has been chosen as it consists of observed data where available and is consistently gap-filled elsewhere.*
- *Preference is given to sites that have recent (last 30 years) observed data and are still operating. This preference is driven by the need for the models to be compared with observed results (i.e. audit runs) over the recent period and to be updated each year.*
- *Rainfall sites are screened by location and length of record.*
- *Rainfall time-series are visually inspected to detect any emerging problems (e.g. spurious trend).*
- *The correlation between the rainfall and runoff is calculated as a guide to which rainfall sites to trial in the rainfall-runoff calibration.*
- *Candidate rainfall sites are compared to a regional average rainfall series to test for trends that are only present at the candidate site"* (see Appendix 8)

The difference between the AWAP and SILO data sets should be investigated. The resulting justification for use of one over the other, including efforts to quality control the data, should be clearly documented by DPIE-Water.

2.2.2 Types of evapotranspiration data used

DPIE-Water use a range of methods to estimate evaporation and ET. Each of these methods makes a series of assumptions. Where PET is used, this is commonly derived using temperature as an estimate of atmospheric demand for moisture. This is not unreasonable over wet surfaces but is non-physical over dry surfaces. The Leonard et al. (2019) and Verdon-Kidd (2019) background papers reference three sources of evaporation/ET data, two of which are Mwet and FAO56, however it is understood that these algorithms are more representative of PET so this should be clarified.

Confusion in regards to the utilisation of terms such as evaporation, pan evaporation, actual evapotranspiration and potential evapotranspiration is widespread among users of this type of information, including how these are calculated and on what assumptions. The Panel notes that it will be critical in the Methods Paper and associated documents that the language relating to ET and PET needs to be clear, and anywhere where PET is used needs to be very carefully evaluated for the biases it might introduce.

The Panel was provided with the following information about the DPIE-Water methodology for ET (Appendix 8):

- *“Morton’s wet environmental evapotranspiration has been adopted based on data availability and McMahon et al. (2013) review. Mwet is available through SILO.*
- *Station choice is based on finding the SILO patch point site that has a mean Mwet close to the catchment mean value.*
- *Mwet time-series should be visually inspected for problems.”*

The Panel is uncertain whether the use of Morton’s Mwet introduces minor errors or a systematic bias as calculations move forward in time. Morton is derived from very limited data, which is almost entirely based on temperate climate relationships.

The FAO56 approach assumes cropped surfaces, a “reference crop” of green grass, well-watered and actively growing and completely covering the ground, which is not well suited to any large NSW catchments.

The approaches also assume methods derived in the 1950s and 1960s, which are no longer suitable under the temperatures and atmospheric carbon dioxide concentrations now reached. These approaches should be modified for Australian crops and how that is done is important.

Therefore, of relevance for the methodology is how the calculations of PET and ET are implemented and whether that implementation introduces an error of indeterminate sign or a systematic bias that grows over time.

One consequence of using PET, as opposed to ET, is that as temperatures have risen and continue to rise, atmospheric demand is assumed to have increased and therefore, ET has increased. However, some consideration is needed into how plants in a drying climate respond and adapt to water stress and reduce transpiration, and whether or how the hydrological model accounts for this.

Some hydrological modelling methods using PET may overestimate the increase in actual evapotranspiration, thereby overestimating landscape drying, reductions in catchment yields, streamflow and so on. How much these methods overestimate the impact of warming is very hard to quantify. Through to 2030 it is likely to be small, but through to 2100 it is likely to be considerable. Solutions to this problem are well established and referenced in the science literature.

A physically and biologically based representation of actual evaporation exists and could be implemented. NARCliM uses such an approach, as do climate models used for future projection.

To replace PET with physically based models would be a major undertaking and is infeasible in the next 1-2 years. But as CO_2 increases and temperature increases Mwet / FAO56 becomes increasingly biased to predicting higher PET. It is unknown whether this increase would be small and insignificant or substantial and would change the resulting strategies that might be implemented.

2.2.3 Paleo data

The use of paleoclimate data across a ~400-year timeframe allows DPIE-Water to produce a wider range of rainfall and climatic variability than is available through using the historical record alone.

The Panel notes that the paleo information captured in the current models is primarily used to better estimate the probability model of IPO wet and dry run lengths, allowing a more representative distribution of IPO durations to be developed. It has not been used to inform estimation of rainfall probability models conditioned on wet and dry IPO states (see Henley et al. (2011)).

In the case of the observed data covering a period of approximately 130 years, the records hold only a small number of IPO phases since the typical length of a phase is 20-30 years. This does not permit meaningful calibration of a probability model of IPO run lengths. The main contribution of the paleo IPO reconstructions is that they increase the “sample size” of IPO runs. Therefore, for a 400-year reconstruction, it may be possible to get 25 inferred IPO runs, which is more useful than the five or so runs in the observed record of 130 years.

The Panels’ observations of multiple paleo reconstructions of IPO are that they are largely inconsistent with each other. For this reason, a composite record is constructed with the hope that some of the noise is filtered out.

However, as with measured records, while useful, paleoclimate proxies are subject to significant limitations (Sorooshian & Martinson, 1995; Franke et al., 2013), which include that:

- the climate signal recorded in the paleo record may reflect only local conditions or it may be more representative of regional or global conditions
- the accuracy with which the paleo record represents climatic variability is often unknown or untested
- the paleoclimate proxy records often reflect more than one variable (e.g. air temperature, precipitation, atmospheric pressure patterns, drought, and runoff), which makes interpretation difficult
- absolute dating may be uncertain
- spectral biases in multi-proxy climate reconstructions can lead to an over-estimation of low-frequency signals
- too few samples to provide a reliable record.

This suggests that efforts to improve both the quality and quantity of proxy records could improve our understanding of climate variability.

Most of the available paleo records are sourced from North America. The US west coast shares similar climate challenges, but it is also impacted by climate drivers that do not apply to Australia. Obtaining local proxy records, as far as possible, may go toward improving this situation.

Paleo records could provide important evidence of characteristics of ‘recent’ Australian drought prior to the observed record. These droughts may be worse in terms of intensity and duration than historical droughts.

The lack of high-resolution climate proxy records for river basins of interest to the DPIE-Water could be improved. If high quality local records can be obtained over time, it is not clear that the current use of proxy records from other locations can be justified in the long term. Moreover, the (possibly interactive) climatic variables that affect the proxies need to be clearly identified.

Note the incorporation of paleo timeseries was not included in the Bega region ECL method.

2.3 STOCHASTIC MODELS

The adoption of stochastic models represents a major advance over the use of historical records, offering, in principle, a more useful and reliable means of evaluating climate risk.

Using stochastic models to develop long term data sets of historical climate variables aims to address the fundamental limitation of using historical observations, namely that relatively short records are unlikely to contain the full range of natural climate variability, including drought periods with extremes of intensity, duration and multiple occurrences that can severely stress the system.

The fundamental proposition of stochastic models is that they make better use of the information contained in the historical record. It is for this reason that stochastic models have been widely adopted in drought risk assessment for major Australian urban areas, including Sydney and Lower Hunter, for which high levels of drought security are required (Berghout et al., 2017).

Numerous stochastic models have been developed, each having their own strengths and weaknesses. Notably, because of short historical records, the parameters in stochastic models are subject to considerable uncertainty. When this uncertainty is propagated through water planning models it produces uncertainty in key performance metrics such as yield or risk of critically low supplies. Communicating this uncertainty is important.

2.3.1 Principles and justification for the use of different stochastic models

The stochastic data sets are a synthetic representation of statistically plausible occurrences, which in the case of the RWS climate models, are based on observed and paleo data.

Stochastic models are data-driven and can be formulated at different space and time scales. This is an inherent strength in that relatively simple probability models can reproduce key statistical characteristics of the observed data. On the other hand, stochastic models need to be calibrated and tested using observed data; the statistical profile of the stochastic data reflects the statistics of the observed data. The utility of the model is critically dependent on the quality and representativeness of the data, and the care with which diagnostic testing of assumptions is conducted.

Stochastic models are heavily dependent upon the modelling assumptions, and the rainfall and stochastic data produced carry uncertainties that need to be recognised.

The three models referred to in the Methods Paper are applied to different regions (Macquarie catchment, Lachlan catchment and Bega River catchment) of NSW. The stochastic models of the three catchments represent natural climate variability without any climate change adjustments factored in, initially.

The choice of the most appropriate stochastic model for a particular region and application is best left to a skilled analyst who *inter alia* will use a range of diagnostics to validate model assumptions and to demonstrate that it is fit-for-purpose. Independent review of the models by experts is a key step to justify the choice and applications of the models.

The three reports on the Macquarie, Lachlan and Bega River catchments reveal that different diagnostic approaches were used (Kiem (2019); Leonard et al. (2019); Verdon-Kidd (2019)). While a prescriptive approach is not recommended, a statement of general principles about the use of diagnostics will aid the evaluation of competing models and inform the water resource modeller whether the stochastic model is fit for purpose. Box 1 includes points for establishing these principles.

Box 1: General principles about use of diagnostics to aid model choice

- The current model is fitted and tested over the complete historical (instrumental) record, but there is a need to have independent testing by reserving some of the data for this purpose. Relevant model-evaluation procedures are then repeated using holdout and rolling horizon/origin methods. The Panel notes that work at the University of Melbourne (Saft et al., 2016) has shown statistically significant changes in annual rainfall-runoff relationships in some catchments during the Millennium drought. Block validation or split-sample testing is an important quality assurance step, however because samples are reduced, the role of sampling distributions becomes even more important.
- Diagnostics should compare an observed statistic against its sampling distribution. This allows the assessment of whether the model behaves in a manner statistically consistent with observations. Put another way, it helps to inform whether differences between observed and simulated statistics can be accounted for by sampling variability.
- The sampling distribution can be adequately approximated using the parametric bootstrap. This involves using the calibrated model to generate many replicates of length equal to the historical record, extracting the statistic of interest from each replicate and empirically estimating the sampling distribution. The use of sampling distributions is considered statistically more rigorous than the tolerances used in the Srikanthan and McMahon (2003) assessment guidelines.
- Ideally, diagnostics should focus on statistics of observed data that are not used in calibration. It is recognised this can be computationally demanding to do. For example, if a model is calibrated to the mean and standard deviation of observed data, then one would expect the calibrated model to closely match these statistics. While a diagnostic check is advisable, it is considered a weak test of the predictive performance of the models when they are applied to data outside of the sample used for their development and fitting. Assessing model consistency with statistics not used in calibration represents a much stronger check of model adequacy. In such cases the use of sampling distributions is critical to assess the significance of difference between observed and simulated statistics.
- It is desirable that diagnostics be constructed to check key model assumptions. For example, in the Matalas multi-site model and its variants, the core of the model is the generation of independent Gaussian disturbances (or innovations). Diagnostics can be constructed to directly test such assumptions.
- It is unlikely that a stochastic model will be able to adequately simulate all key statistical characteristics from daily to decadal time scales. Given this perspective, fit-for-purpose diagnostics need to be used to guide model selection. For example, in an urban water resource system with large carryover capacity, the critical period is likely to be several years. In such a case, diagnostics based on multi-year sums seem particularly relevant. A model that reproduces daily and monthly variability but fails to reproduce multi-year variability is likely to be unsuitable.
- The generated rainfall and ET/evaporation is propagated through a rainfall-runoff model to produce streamflow time series. It is assumed that propagating stochastically generated rainfall and PET through a calibrated rainfall runoff model produces streamflow time series that are statistically consistent with observed. Where possible the veracity of this assumption should be checked using the full set of diagnostics - gauged unregulated catchments would be the obvious candidates for such a check.
- The best way to construct fit-for-purpose diagnostics is to propagate multiple replicates through the water resource system model and compare the sampling distribution of key system behaviours against observed behaviours. For example, while the use of diagnostics based on multi-year sums seems a good choice for systems with large annual carryover storage, a more relevant diagnostic would focus on the water resource system performance (such as the probability of storage falling below selected thresholds or the probability of rationing).

Beyond the issue of diagnostic approaches, several potentially significant issues, common to all the models, have been identified and can be described as follows.

2.3.2 Use of IPO

DPIE-Water has notionally divided the state into quadrants based on the key climate drivers that are relatively dominant in the locations. The RWS process will consider the twelve major catchments in NSW located within these quadrants to inform the dominant climate drivers within each of the catchments, as well as the modelling method to be used. For catchments in the north and west of NSW, the IPO driver is used to develop an understanding of the climate variability and change into the future. This is illustrated in (Figure 2).

The use of the IPO driver artificially incorporates decadal memory by creating long sequences consisting of random length sub-periods of IPO positive and negative phases. The methodology uses the paleo-record as a basis for developing stochastic data sets, but key to this approach is identifying periods of positive and negative IPO and then stochastically generating rainfall and evaporation values from the current record.

Rainfall and ET in each phase are selected using data from the corresponding phase only. This has the effect of increasing the likelihood of having a run of dry or wet multi-year sequences relative to what would occur from randomly selecting monthly data from the historical record alone.

DPIE-Water is applying the IPO method to three of the quartile regions of NSW (Figure 2). Although the Panel has not seen work related to regions other than those specified, DPIE-Water has indicated that the approach for south-west NSW, labelled in Figure 2 as being influenced by IPO, IOD and SAM, will be similar to the approach for the two currently IPO-dominated catchments of Lachlan and Macquarie that are also shown to be influenced by ENSO.

It is noted that previous studies have identified the role of IPO in modulating eastern Australian droughts (Palmer et al., 2015). Verdon (2007) demonstrated that persistent regime shifts in the hydroclimate of the Lachlan catchment are being driven by multi-decadal oscillations of the Pacific Ocean, as well as comparative analysis in Leonard et al. (2019) comparing the observed data to the modelled data confirmed the relevance of IPO, with IPO-generated model outputs fitting well with the observational record.

The use of IPO with the stochastic model has been driven by its relatively long duration signal and because there is paleo data enabling the method to extend the length of plausible past climate variability. ENSO is also listed in the Methods Paper as having an influence on the regions in focus. Recent press articles have highlighted an important role for IOD on future climate in NSW (Deacon, 2020). The IOD has not yet featured in work seen by the Panel. However, IOD phases appear to be of a shorter duration than decadal IPO phases.

Although there are uncertainties in estimating IPO phases from paleo data, the method nevertheless seems to provide a more conservative estimate of the likelihood of dry and wet sequences than if this approach using stochastic models with shorter memory. These are useful outcomes, even if the assumption upon which the method is based requires further research to be justified.

Responding to Panel questions, DPIE-Water noted that: “*reference to drivers is for contextual information. It is saying that these are the atmospheric circulation patterns that influence climate. Their influence is manifested in the observational data. The stochastic methods then use that observational data to generate the extended data. The methods are agnostic to information about anything than (sic) the data itself, and the characteristics such as variability, persistence and spatial dependency.”* (DPIE-Water, pers comm, 23 January 2020).

The Panel agrees that in principle, one does not need to include climate drivers if the interest is in hydroclimate variables. The Methods Paper would benefit from further articulation of the role that the drivers do and don’t play in the methodology, specifically in the development of the stochastic data sets and how uncertainty about their future behaviour may be manifested into the scenarios.

2.3.2.1 Issues with IPO reconstructions

The historical climate record in eastern Australia suggests there have been multi-decadal periods (or epochs) that alternate between wet and dry periods. These wet and dry epochs have been associated with positive and negative phases of the IPO climate index in the observed record. There are several paleo reconstructions of the IPO index to extend the record beyond the observed record. A characteristic of these reconstructions is that they do not agree particularly well with each other. Attempting to filter the noise in these reconstructions was one of the motivations for the composite index reported by Henley et al. (2011). Nonetheless, there remains uncertainty about the parameters of the probability distribution of the length of IPO wet and dry sequences or runs.

Two of the three models cited in the Methods Paper, Verdon-Kidd (2019) and Leonard et al. (2019), use a two-state stochastic model where the states alternate between wet and dry IPO epochs whose length is sampled from a probability distribution. Verdon-Kidd (2019) and (Leonard et al., 2019) have used different IPO reconstructions and as a result have different positive and negative periods. This may result in inconsistency in outputs that should be considered in interpretation because the IPO positive and negative periods provide the rainfall data to define wet and dry distributions. That said, small differences would not be detectable in any diagnostic analysis. The Methods Paper should provide some clarity on the choice of paleo IPO reconstructions to aid interpretation and use.

It is important not to interpret dry IPO epochs as droughts; they are periods that experience on average less rainfall than wet epochs.

In a somewhat surprising result, Zhang (2018) applied the Berghout et al. (2017) methodology to a case study based on Warragamba reservoir and found that uncertainty in the parameters defining the probability distribution of IPO run lengths contributed minimally to the total uncertainty about drought risk. This was attributed to two factors. The first was the small sample sizes for observed streamflow in the wet and dry IPO sequences. In effect, the estimation of the wet and dry epoch streamflow distributions each used about half of the observed record, further exacerbating the parameter uncertainty problem described below. The second was the fact that the reservoir drawdown time in response to severe drought was much less than the expected length of the dry IPO epoch. The severity of the drought is controlled mainly by sampling from the dry IPO streamflow distribution rather than length of the IPO dry run. This result may alleviate some concern about the uncertainty in IPO run reconstructions.

2.3.3 The ECL method for NSW south coast

The IPO methodology is not used for the south coast of NSW (Bega River region) because East Coast Lows (ECLs) are believed to be the dominant influence on rainfall. As a result, rainfall and ET modelling of future climate for this region is based on climate model informed changes to the frequency of ECLs.

The ECL method used for the south coast area of NSW relies on randomised changes in the frequency of ECLs in the observed record and recalibration of the stochastic model to represent future climate. This is in addition to the climate model informed scaling of rainfall and ET generated by the stochastic model calibrated to historical data.

The Panel is of the view that this method is likely to be as good as any other for representing rainfall on this coastal area. While ECLs are important to rainfall in the region, rainfall variability can also arise from other factors that may influence the climate, including other drivers.

Any advice stemming from this (as with all methods) should reflect the fact that events outside the range simulated, while unlikely, might occur.

Alternative approaches to generating randomised samples for the south coast should be considered as part of the future research program.

2.3.4 Future directions with climate drivers

The Panel agrees with the overall approach of developing the stochastic models informed by knowledge of dominant climate drivers for each of the NSW regions and using the characteristics of these drivers to inform understanding of the historical statistical nature of rainfall and ET, with the stochastic data also informed by observational data and the paleo record (in the case of the methods using IPO). The primary purpose for considering drivers, particularly in adding IPO to the stochastic model, is to extend the length of plausible past climate variability by using the available paleo data.

IPO drivers have a long period (multi-year to decades) with a correspondingly low frequency, while the behaviour of ECLs is such that they can occur several times per year and have a short duration of hours-days.

However all regions in NSW will have contributions from multiple drivers and systems, so full explanations for climate's impacts on rainfall and ET won't be found in any single driver and it is inevitable that multiple interacting drivers could be involved (e,g, ENSO-IPO) in rainfall events and the replenishment of surface water bodies.

In addition, there is uncertainty about whether and how climate change will influence climate drivers and their interactions. No matter how these major drivers may have acted historically, they may not act consistently in the future with a changing climate. Indeed, we do not know if they are stationary in the past record, as there is only a relatively small number (approximately 8) of IPO epochs in the measured records and more in the paleo-record.

Therefore, considering both the historical record and future changes, there is a need to look at the effects of individual and combined (joint) drivers on rainfall. In relation to the RWS methodology, this could be particularly important for regions where multiple drivers could play important role (e.g. in cases where both IPO and ECL may play a major role in rainfall).

The Panel believes work to improve stochastic series will be needed, including, but not restricted to, methods that use individual drivers, particularly for regions that may have a combination of dominant drivers. Indeed, for regions with multiple dominant drivers it may be better to explore methods that don't rely on specification of individual drivers, rather than add additional drivers into an existing method, with or without their possible nonlinear interactions. The exploration might find that using drivers is best way to go in future, but it might not.

Overall, the Panel agrees that work should be undertaken in the near term on what approach to use to generate stochastic data sets where multiple drivers dominate, and for future iterations of the RWS, given that climate change may alter the impact of drivers (non-stationarity), we suggest future research put effort into other ways to generate the stochastic data sets.

In the longer term, work with colleagues, such as the community of practice group to examine whether behaviours of climate drivers (e.g. IOD, IPO, ECL, ENSO, and SAM) into the future will remain statistically consistent with the past. Active reviews of the latest literature every few years would be prudent.

2.3.5 Assumption of climate stationarity in historical records

All the stochastic models described in the Methods Paper have been built and calibrated on the assumption of climate stationarity. They are premised on the assumptions that the historical record is representative of the current climate and that global heating will change this climate in the future. Climate stationarity in the stochastic models means the statistical characteristics of the climate over a range of time scales remain unchanged over the historical and paleo record. The inclusion of IPO state dependence in the stochastic model, while possibly providing a better description of decadal variability, does not make the model non-stationary. The distribution of IPO runs is not dependent on time, nor are the distributions of hydroclimate variables during wet and dry epochs dependent on time.

This raises the question - is the historical record already affected by global heating? If so, will using stationary stochastic models lead to a significant underestimation of climate risk in the future? The following discussion attempts to put this in perspective.

NSW climate varies markedly from year to year and decade to decade in response to ENSO, the IPO and other natural processes (Figure 3 and Figure 4). Figure 5 shows that trends in rainfall between 1950 and 2017 across NSW vary spatially and seasonally with eastern NSW showing widespread declines in winter rainfall and south-eastern NSW showing similar declines in spring and autumn.

NSW climate is also changing in response to human-forced increases in greenhouse gases, with more than 1 degree Celsius of warming over the past century (Figure 6) (CSIRO-Bureau 2015). Anthropogenic forcing is also implicated in observed rainfall declines in southwest Australia and southern Victoria over recent decades (Jakob et al., 2012; Hope et al., 2017). While rainfall declines are projected for NSW during winter and spring (CSIRO & BOM, 2015), and NSW received record low rainfall in 2019 (Figure 3 and Figure 4), the extent to which anthropogenic forcing has already influenced NSW rainfall is less clear given uncertainties. While further research is required, the possibility exists that anthropogenic forcing might already be affecting rainfall and streamflow in NSW to some extent.

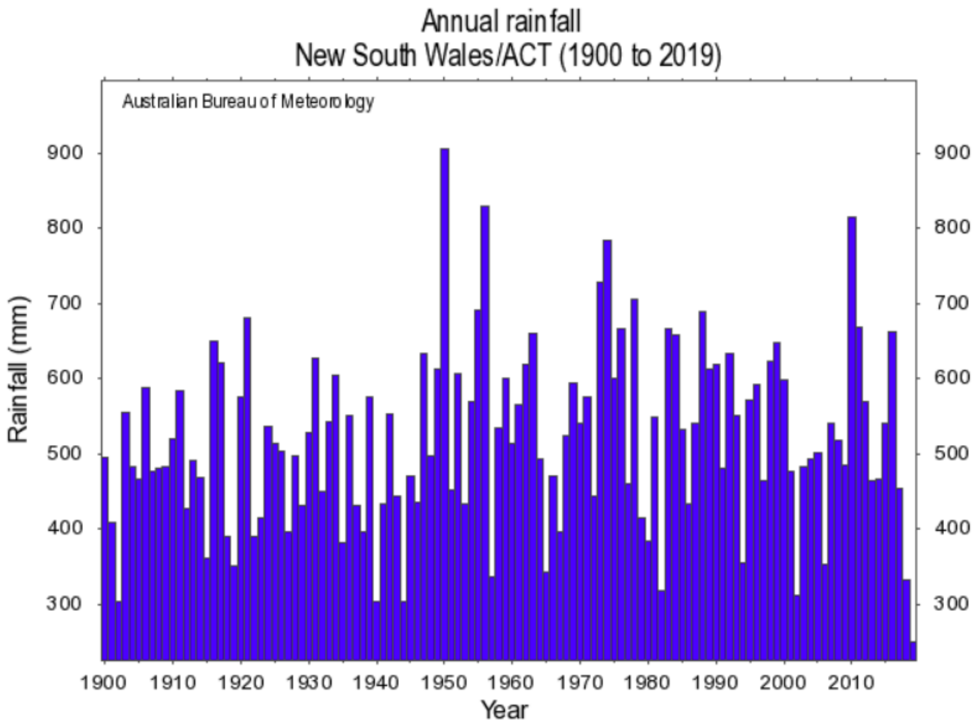

Figure 3: NSWACT annual rainfall from 1900 to 2019
(Source: BOM)

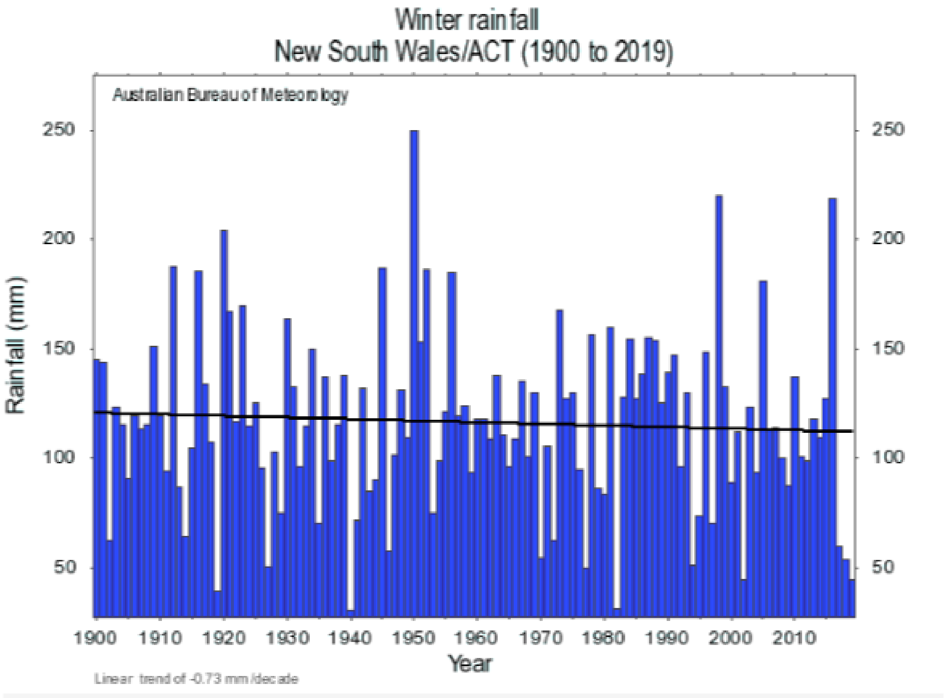

Figure 4: NSW/ACT winter rainfall from 1900 to 2019
(Source: BOM)

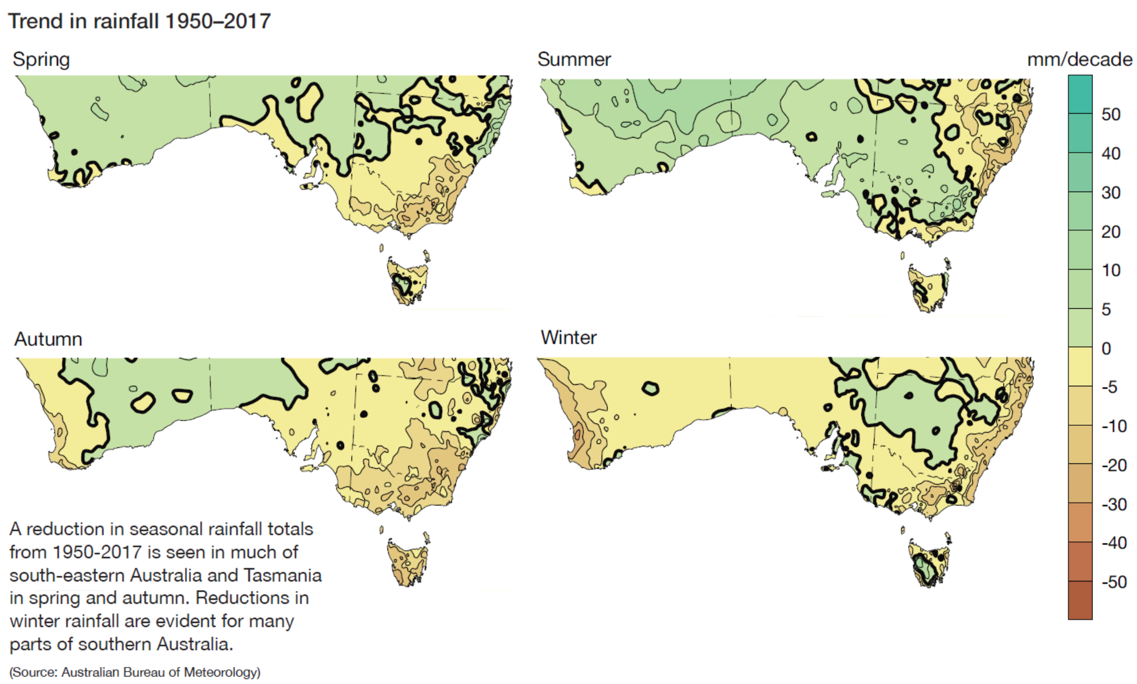

Figure 5: Spatial and seasonal trends in rainfall since 1950 over southern Australia.
(Source: BOM)

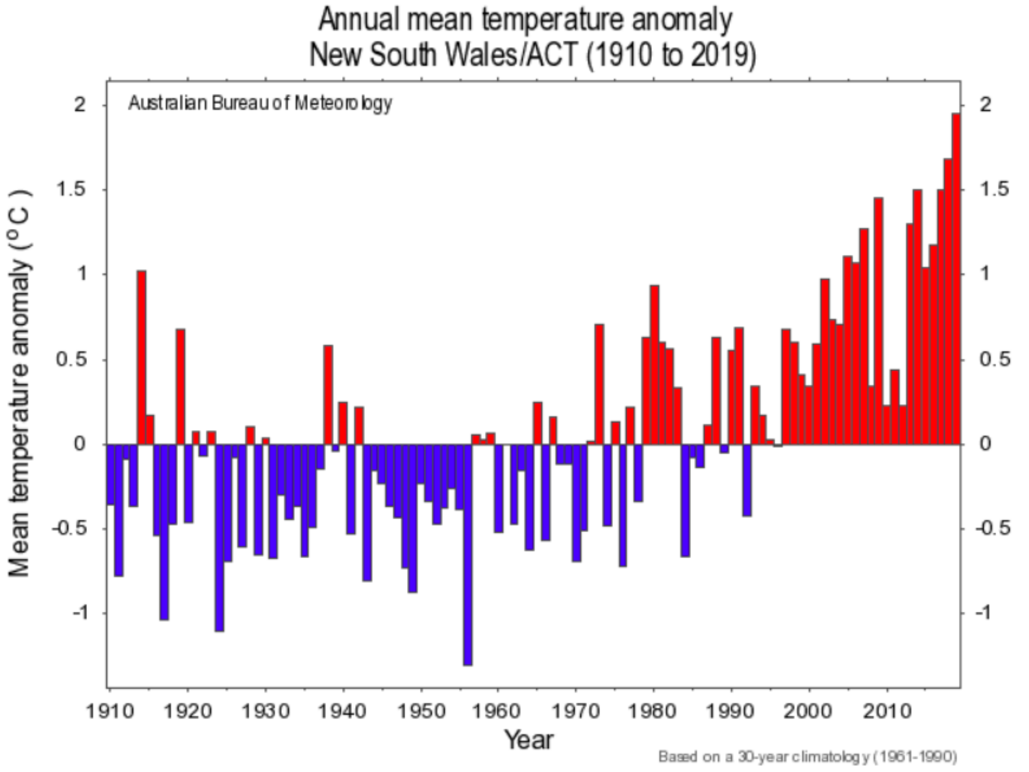

Figure 6: Annual mean temperature anomaly NSW/ACT (1910 to 2019)

(Source: BOM)

Non-stationarity in the historical record was considered in relation to the baseline conditions of the stochastic data that are factored against modelled climate change – in assessing risk if the climate has already started changing, how should that be dealt with in the modelling? Do we consider the 100+ year historical baseline to be an accurate representation of the conditions as they are today?

2.3.5.1 Associations between rainfall, temperatures and streamflow

One way to incorporate a climate change signal is to condition a stochastic model on one or more covariates that are associated with climate change and that are predicted with good skill by climate models. It is important to note that causality between the covariate and hydroclimate variables is not a prerequisite with this approach; all that is required is that the association be stable. An example of this approach was described by Wasko and Sharma (2017) who correlated the parameters of a Neyman-Scott rectangular pulse rainfall with average monthly temperature. They used this as a basis for simulating sub-daily rainfall in a warmer climate and found a reduction in the chance of the medium-sized floods that typically provide the great bulk of reservoir inflows in any one year.

Nicholls, Drosdowsky, and Lavery (1997) analysed a high-quality data set of temperature and rainfall across Australia and concluded that both annual maximum and annual minimum temperature are strongly associated with annual rainfall. Power et al. (2017) subsequently examined this work and found maximum temperature associated with rainfall, but this was not the case between minimum temperature and rainfall.[3]

Using the insight that there are possible stable associations between maximum temperature and rainfall, Kiem et al. (Under Review) proposed a stochastic model of hydroclimate variables

[3] See following paper for earlier work (1998) by Power et al https://www.researchgate.net/profile/Scott_Power4/publication/279889526_Australian_temperature_Australian_rainfall_and_the_Southern_Oscillation_1910-1992_Coherent_variability_and_recent_changes/links/566def2d08ae430ab5001e75.pdf

at seasonal time scales conditioned on average maximum daily temperature. In a case study, they applied this approach to the annual inflow into three reservoirs, including Warragamba reservoir, using as the covariate the maximum temperature at Bathurst airport, a long-term high-quality temperature record. When plotting a time series of the average daily maximum temperatures from 1910 to 2016 the authors observed that the long-term average daily maximum temperature was approximately 20ºC, while the average daily temperature between 2001 and 2016, a period that covers the Millennium drought, was approximately 21ºC. Their main findings, yet unpublished, suggest that present day average annual streamflow has dropped when compared to the historic record starting in 1910 for the Warragamba reservoir case demonstrating non-stationarity. It needs to be stressed that this result is case-specific, and the methodology is under review. In considering this finding, it should be noted that the negative correlation between annual rainfall and maximum temperature is widespread across Australia, although this correlation is by association rather than causal – more rainfall is associated with more cloud and less radiation. For the purposes of this Panel report, the relevant observation in the unpublished paper was the increase in the average temperature and decrease in the average annual streamflow in the more recent observed record versus the longer term for this case, pointing toward non-stationary conditions.

2.3.5.2 Calibrating models to incorporate non-stationarity

Some jurisdictions assume non-stationarity and calibrate their models accordingly. The Panel understands that Melbourne Water calibrate their stochastic models to the recent past due to concerns about long-term drying climate. This warrants further investigation. The consequence of ignoring the non-stationarity when calibrating the stochastic model to historical data, if climate conditions have already started to change, would be that the model produces over-optimistic (biased) estimates of the probability of drought events for the current and near future climate. The use of IPO paleo data does not fix this potential problem. The use of a stationary stochastic model calibrated to historical data may result in an underestimation of present day and future climate risk.

Overall, the DPIE Water’s methodology to use observational and paleoclimate data informed by an understanding of climate drivers to select, calibrate and test stochastic models and the factoring of NARCliM projections is consistent with best-practice approaches to climate risk management. An additional important step, however, is needed for ongoing improvement of the methodology adopted by DPIE-Water, which can be undertaken in parallel with concurrent RWS work. It includes an assessment of non-stationarity in the historical record to determine whether changes in climate in recent decades affect estimates of present-day climate risk compared with climate risk based on the whole observed record. A recent paper by Stephens et al. (2020) lists a range of different approaches to testing non-stationarity in climate and hydrology models and also approaches to addressing this including using calibration parameters. Resolving this should be the highest priority for the DPIE-Water method.

The Panel recommends:

In view of growing evidence of rainfall declines over recent decades, the observed record may be non-stationary. Because the stochastic models in the Methods Paper assume climate stationarity over the historical record, there is a distinct risk that such models may underestimate current climate risk and by implication future climate risk. To manage this a two-step approach is recommended. The objective is to capture the current climate risk and a baseline that reflects this.

First Step

The first step is to make a careful assessment of non-stationarity in the observed record. This is a specialised task requiring care in the choice of methods and an understanding of their limitations. Without pre-empting such an assessment two useful methods are described.

The Panel recommends Step 1 be undertaken by consultants with access to the software and who can interpret the calibration and validation results.

The Panel also recommends that Step 1 is undertaken immediately to assess if Step 2 if required.

Step 1_Method 1

Split-sample testing or hold-out validation is a well-established technique. It splits the observed record into two parts. The stochastic model can be calibrated to each part and the sampling distributions of the parameters compared to determine if there is evidence of significant differences[4]. Alternatively, the model can be calibrated to one part and its performance in the second part assessed using diagnostics based on predictive distributions (Figure 7).

Because this approach implicitly assumes the record used for calibration is stationary, care is required in deciding how to split the observed record. For example, in view of evidence suggesting declines in rainfall in recent decades, the observed record could be split into pre- and post-1990 intervals (*1990 reference approach*) (as undertaken by Melbourne Water).

Alternative to this approach could be that illustrated in Figure 7, where the record used for validation and prediction are chosen with regard to the two most recent droughts (*drought reference approach*). The predictive skill of the models are being tested on data they have not 'seen' during the calibration process.

A variation on this scheme is to constrain the model calibration periods to be the same length. Thus, the earliest parts of the historical record are discarded for Runs 2 and 3 (Figure 7). That way you compare 'like-with-like' in terms of the amount of data available for calibration in each run.

Using both of these approaches (*1990 reference* and *drought reference)* could be prudent to confirm the outcomes. The Panel recommends that in undertaken this work, DPIE-Water seeks independent advice from experts and uses appropriate diagnostics and quality assurance processes as appropriate.

The Panel anticipates calibration would be largely automated so re-running it with different records should be a straightforward process. The Panel anticipates that the time required to undertake the split-sample testing would be based on how much automation has been developed already to perform the diagnostic analysis.

[4]In terms of judging the difference between probability distributions, it is advised against the use of the Kolmogorov-Smirnov test which looks at the maximum vertical distance between two cumulative distribution functions. It has the disadvantage that it is more sensitive to deviations near the centre of the distributions than at the tails. It is the differences in the lower and upper tails that would be of particular interest here (i.e. droughts and floods). There are other measures of statistical distance and similarity around that are more fit-for-purpose. Also, a simple quantile-quantile plot can be quite revealing.

To differentiate the contribution of the two severe back-to-back droughts this century from any contribution from anthropogenic forcing in the recent historical record, one of the diagnostics could compare the distribution of 10 to 20 year cumulative sums predicted by a model calibrated to 20th century data against the sums observed this century.

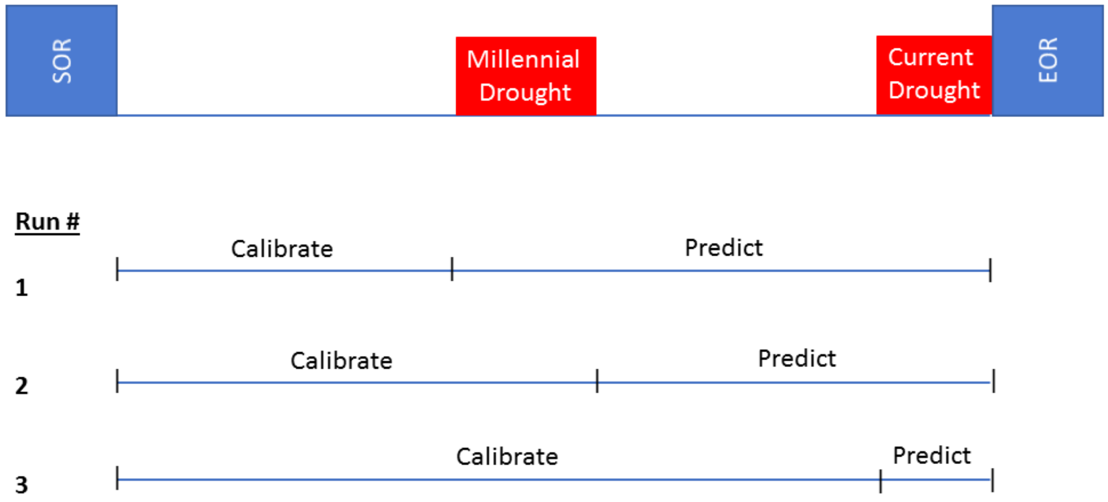

Figure 7: Schematic describing approach for Step 1 Method 1 to check for climate non-stationarity

SOR = start of historical record; EOR = end of historical record; "Calibrate" denotes the periods of record to be used jointly for the calibration of the stochastic model and IQQM; and "Predict" refers to the remainder of the record for which IQQM predictions are made.

Step 1_Method 2

The second approach is more difficult to apply than the first method and is still in the research domain. However, it is expected to have more statistical power to identify non-stationarity because it makes use of the full record and thus benefits from lower parameter uncertainty. It involves calibrating a stochastic model conditioned on an observed covariate associated with a change signal (e.g. temperature). The main challenge is to find a suitable covariate that has a stable and statistically and hydrologically significant association with rainfall.

Second Step

If after application of the first step it is concluded that the recent past may not be statistically consistent with the full observed record and, more importantly, that differences are hydrologically significant, it will be necessary to revise the stochastic modelling approach described in Methods Paper.

A number of conceptually different approaches are available if the results of Step 1 warrant Step 2:

Step 2_Approach 1

Calibrate the stochastic model to the recent past record with the knowledge that parameter uncertainty will be elevated.

Step 2_Approach 2

Adjust the observed record prior to the recent past so that it reproduces certain statistical features of the recent past and then calibrate the stochastic model to the full adjusted observed record.

Step 2_Approach 3

Identify suitable change covariates and calibrate a conditional stochastic model.

Each of these approaches has strengths and weakness. As there is limited experience with this issue it is expected that methods will evolve in the future.

The Panel suggests that Approaches 1 and 2 are in principle similar to the earlier calibration while Approach 3 would be a longer-term measure.

Once this work is complete, it should be clearly written up in the Methods Paper, including any descriptions of regional variations.

Consideration of non-stationarity does not only relate to the current model baseline, but also how future climate impacts may change and how this is captured in modelling. Stephens et al. (2020) describes non-stationarity as “a loss of predictive performance when a method that relies on past data is used to project into an unseen, significantly altered future”. Therefore practitioners will need to anticipate future climate impacts that may be non-stationary, such as interaction between climate drivers (section 2.3.4) or the response of catchments to prolonged dry periods.

2.3.6 Short historical records and parameter uncertainty

Calibration is a process whereby the parameters of the stochastic model are selected to maximise consistency with observed data. There is approximately 130 years of historical rainfall and/or evaporation/ET data available to undertake calibration. A question is whether such a short record of historical data is enough to assure the parameters of the stochastic models have been estimated with sufficient accuracy. There are limitations when calibration and bias correction are undertaken with finite data sets. A demonstration of this uncertainty is with the use of the IPO in the stochastic models. In the observed record there are only a handful of wet and dry IPO epochs. It is implicitly assumed that the rainfall/streamflow probability distribution conditioned on a dry IPO epoch is the same for every dry IPO epoch. This assumption cannot be meaningfully tested using the observed record.

Parameter uncertainty is an ‘inconvenient’ fact - it induces an uncertainty about the probability of future events, an uncertainty that is irreducible. It has been shown that this uncertainty can have considerable influence on uncertainty in water yield, a concept that has dominated urban water resource planning.

Even in the absence of climate change our confidence about the probability of future events is not high. We already know how to estimate the uncertainty about the probability of future events. The challenge is to inform decision makers of the implications of this uncertainty.

The theory of statistical inference, which seeks to quantify the uncertainty in calibrated parameters, can be used to help address the question of whether the historical records provides sufficient accuracy. By way of example, consider the problem of estimating the mean annual streamflow flowing into Warragamba reservoir, the key water storage for Sydney. The best estimate of the mean is the average of the observed annual flows. If a 100-year record is used to estimate the mean, it turns out that the true mean has a 95% chance of lying between approximately 80% and 120% of the sample average. This is rather substantial uncertainty.

While statistical methods for quantifying parameter uncertainty are well developed, such information is rarely, if ever, used in decision making. This introduces the possibility of serious under-design of water infrastructure. Box 2 illustrates this.

Box 2: Impact of parameter uncertainty

Berghout, Henley, and Kuczera (2017) investigated the impact of parameter uncertainty in a stochastic model on the yield of an urban water supply system similar to the Lower Hunter system. The yield is the annual demand that can be supplied with a specified supply security - in this case, the chance that the storage drops below 20% of capacity is no more than 1 in 1000 in any year.

Figure B2.1 shows a probability density plot of the yield inferred using 113 years of data. While the most likely value of the yield is about 80 GL/yr (green plot), the true yield may be less than 65 GL/yr. This uncertainty, if ignored, can lead to the system being exposed to elevated drought risk. Figure B2.2 quantifies the uncertainty in the probability of storages dropping below 20% in any year. This uncertainty is solely due to sampling variability arising from using a 113-year record. If the record were say 100 times longer, this uncertainty would be reduced by a factor 10.

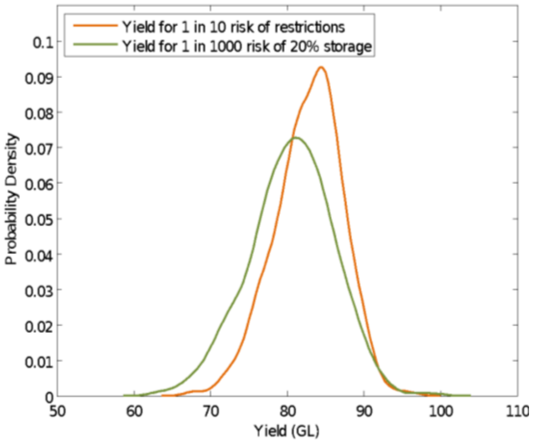

Figure B2.1: Distribution of estimated yields resulting from uncertainty in the climate parameters
Source: Berghout et al. (2017)

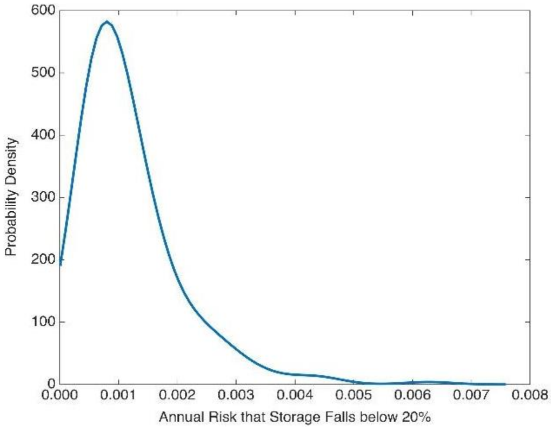

Figure B2.2: Distribution of the calculated risk of storage falling below 20% when demand is 81 GL/year
Source: Berghout et al. (2017)

Despite the practical significance of parameter uncertainty there is little research about this topic in the context of water resource management. The approach used by Berghout et al. (2017) is computationally expensive and needs to be made more efficient. It is expected that the impact of parameter uncertainty is case specific – for instance, it is expected that its significance grows the more variable the climate and the higher the required level of drought security. Arguably the biggest challenge is learning how to use this information in decision making.

2.4 INCORPORATING CLIMATE MODELS – QUANTIFYING FUTURE RISK FROM CLIMATE CHANGE

The following discussion examines the strengths and weaknesses of regionally downscaled climate models, including NARCliM, and describes their application in the context of the DPIE methodology, particularly the use of GCM/RCM climate models to factor climate change projections into the stochastic climate data sets.

While the Methods Paper relies heavily on descriptions in the three background papers, it is important to note that the description in Leonard et al. (2019) for the incorporation of climate change into the stochastic data differs from the final methodology. Instead of using CSIRO/CCIA scaling factors as per Leonard et al. (2019), NARCliM was used for the DPIE-Water methodology due to the improved spatial resolution and topography representation with NARCliM. Note that the Panel hasn't seen any detailed information on how NARCliM was specifically used in the methodology other than the brief overview provided in the Methods Paper.

2.4.1 NARCliM background

Climate projections for NSW can be obtained by using direct output from GCMs (e.g. Coupled Model Intercomparison Project Phases 5 and 6 (CMIP5/6)), although this is not advised, or by downscaling GCM output using RCMs, which have been configured for NSW. This downscaling approach is the one used by NARCliM, which used three RCMs to downscale a reanalysis dataset (1950-2009) and four GCM data sets (1990-2009, 2020-2039, 2060-2079), chosen to span a range of future projections and with consideration of model dependencies. These climate models were from the CMIP3 generation.

Within NARCliM, each GCM model was downscaled several times using different configurations of the regional model to create the best available estimate of future climate change over the state. NARCliM 1.0 used CMIP3 models and future time slices of 20 years; NARCliM1.5 uses CMIP5 models and continuous simulations over 150 years (1950 – 2100). Once NARCliM 1.5 is applied and tested, this will demonstrate whether its performance improves on NARCliM 1.0.

NARCliM has been developed as a product to inform numerous industries and end-users on the impacts of climate change on utilities, asset risk, insurance, agriculture, hydrology, finance as well as to meet the broad needs and interests of NSW Government. The model covers the whole of the Murray Darling Basin and thus extends beyond NSW. When the model was first in development, decisions were made on which GCMs to use in NARCliM at the time, focussing on GCMs with better skill at resolving the east coast climate. It is understood that this led to selecting GCMs that had particular capabilities in relation to East Coast Lows, and thus wetter projections (reinforcing the need for bias correction in relevant applications). NARCliM 1.5 includes some drier GCMs and NARCliM 2.0 will include a wider range of models, increasing the dry-dominant climate representations, that may better reflect conditions inland.

The development team behind NARCliM has maintained an approach of seeking feedback from end user groups to understand the model's performance and to inform updated versions. As such, NARCliM 1.5 version will be released in early 2020 and NARCliM 2.0 at some stage after 2022, also informed by feedback.

Table 1 sets out the characteristics of the three NARCliM versions.

Table 1: Characteristics of current NARCliM1.0 and succeeding iterations.

	NARCliM 1.0	NARCliM 1.5	NARCliM 2.0
Models	12	9	TBC
Model timeframes	20 years 1990-2009, 2020-2039, 2060-2079 (& 1950-2009 NCEP-forced simulations)	150 years 1950-2100 (& 1981-2010 ERA-Interim forced simulations)	At least 150 years 1950-2100 (+ reanalysis simulations) Bespoke regional simulations
Grid	50 km (CORDEX-Australasia) & 10 km (NARCliM)	50 km (CORDEX-Australasia) & 10 km (NARCliM)	finer resolution & multi-domain (TBC). Minimum CORDEX-Australasia (25 km), NARCliM (4 km?)

Timestep	3-hourly (all vars), hourly (precip, 2m temp, RH, wind)	3-hourly (all vars)	TBC
Variables	100+ (20 post-processed)	120+ (25 post-processed: 1-hr, 3-hr, 6-hr, daily, monthly, season)	TBC
GCMs	4 CMIP3	3 CMIP5	CMIP6
Regional models	3 regional models per GCM (WRF3.3)	3 per GCM – same RCMs as for NARCliM1.0 (WRF3.6.0.5)	Currently testing new physics (WRF4.02+)
Future emissions scenario(s)	SRES A2	RCP4.5 & RCP8.5	TBC
Example uses	Regional climate snapshots, near versus far future climate analyses for temperature, heat, snow, fire, rainfall, etc. Strategic planning (e.g. State Infrastructure Strategy, Sydney Region Plan). Transport for NSW Asset Management Authority – guidelines for rolling stock. Asset risk management (XDi). NSW Common Planning Assumptions. NSW NPW adaptation strategy, etc.	**In addition to previous iteration:** climate extremes, thresholds for impacts, compare with non-climate datasets.	**In addition to previous iterations:** hazards over cities, coastal changes, impacts of ocean warming on NSW climate.

2.4.2 Limitations of downscaled climate models

There are different approaches to downscaling climate information, including statistical and dynamical. Directly using CMIP5/6 GCM models for key variables such as rainfall is problematic as GCMs were built with coarse spatial resolution and vary widely in how well they simulate rainfall over NSW in the historical period. Work by Ukkola et al. (2018) indicates that there are marked between-model discrepancies between GCMs and they don’t simulate the severity and scale of drought events as accurately as needed. As a result, regionally downscaled GCM/RCMs are preferred over GCMs for projecting local climate impacts.

NARLCliM 1.0 is a well-designed and executed project that provides considerable information of value to NSW. However, DPIE-Water has noted limitations of using NARCliM 1.0 datasets directly for this application. These limitations include difficulty in reproducing important rainfall characteristics from interannual to decadal time scales and the shortness of modelled outputs (20-year timeframe for NARCliM 1.0) as mentioned in the previous section. The time slices, the limited length of simulations, and issues with persistence in the climate models mean the results need to be used with care.

The impact of the first limitation is illustrated in the paper from Lockart et al. (2016), which was reported on work funded by the NARCliM owners to understand its limitations. It compares the observed total storage behaviour of the Lower Hunter bulk water system from 1990 to 2010 against 12 NARCliM 1.0 reanalysis data sets using four GCMs and three 3 RCMs. The paper reports that there are differences between RCM and GCM combinations and that even with bias correction of rainfall, the GCM-RCM combinations did not adequately reproduce the observed behaviour, illustrated in Figure 8.

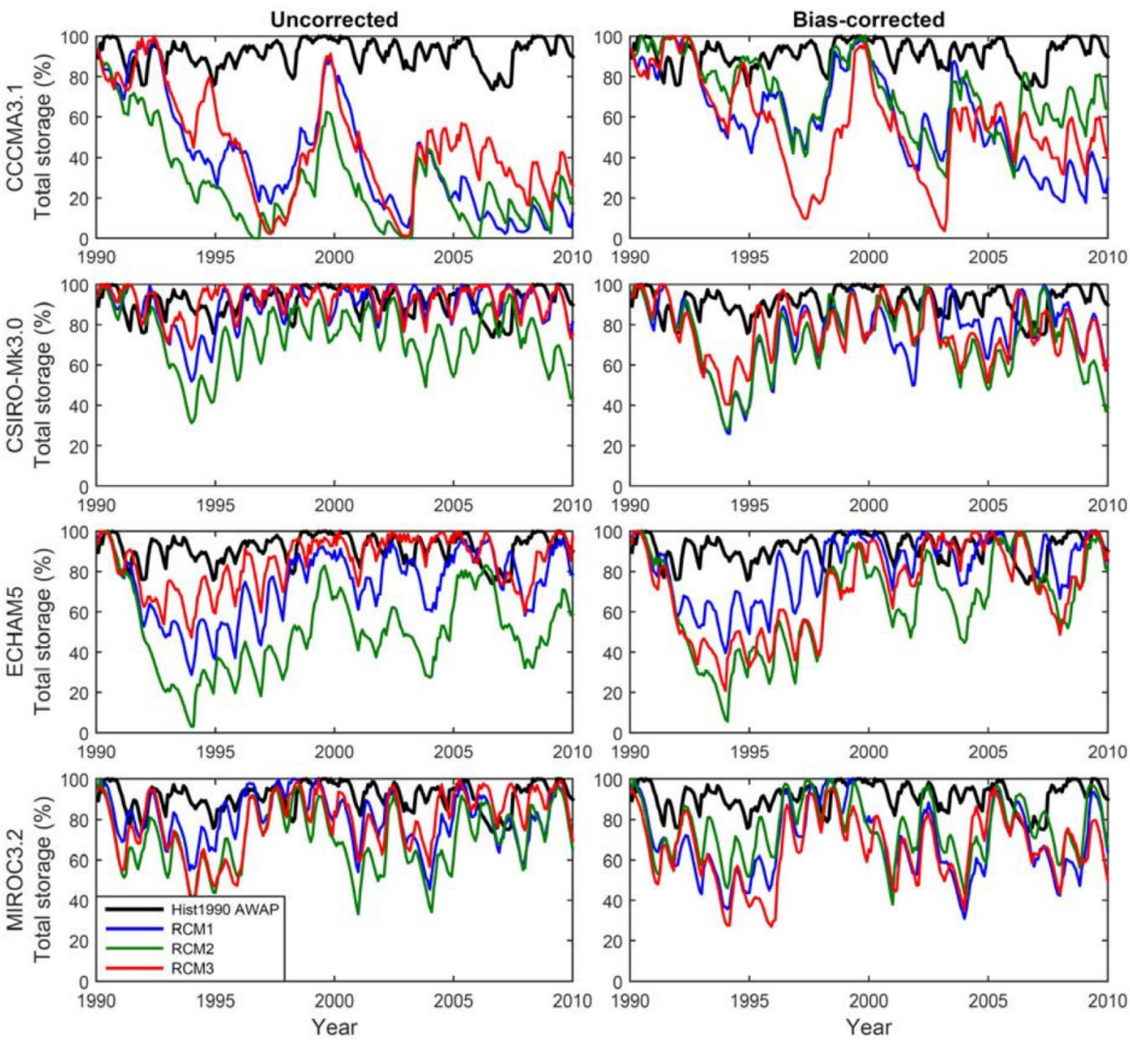

Figure 8: Total storage time for Lower Hunter bulk water supply for the period 1990 to 2010:

Note: black lines represent observed, while coloured lines represent different reanalysis combinations of GCM and RCM

Source: (Lockart et al., 2016)

The Lockart work was commissioned by the NARCliM owners to explore the limitations of NARCliM 1.0 at the time and to test how NARCliM 1.0 datasets (bias corrected and non-bias corrected) reflected reservoir and stream levels in the Hunter region when compared with observational data sets input to the hydrology models SimHyd to simulate streamflows and WATHNET5 to simulate reservoir levels.

SimHyd is a lumped conceptual model that simulated streamflow using daily rainfall and PET as input, while WATHNET5 is a bulk water simulation model. The Lower Hunter WATHNET5 model has inputs of monthly streamflow and PET for each of the five catchments in the Hunter region of NSW. The paper aimed to assess how the outputs from SimHyd and WATHNET5 using AWAP rainfall and SILO PET inputs compare with SimHyd and WATHNET5 outputs obtained using NARCliM 1.0 inputs (rainfall and PET).

The report noted the following issues that impacted the performance of the NARCliM 1.0/hydrology model outputs:

- bias correction for rainfall is available in NARCliM 1.0 however PET cannot be bias corrected, and since there is feedback between the two (as PET decreases when rainfall increases), then the model produces drier results when rainfall is bias corrected but PET is not;

- for reservoir levels, where the storage WATHNET5 hydrology model uses monthly rather than daily PET data, the impact of the daily feedback between rainfall and PET is reduced, and in combination with the lack of bias correction with the daily NARCliM 1.0 means there is, as noted in Lockart et al. (2016), a large influence on reservoir simulations – noting storages with a large surface area have relatively more impact from evaporation compared with streamflows.

The paper by Lockart reports these limitations are due to issues with the ability to perform bias correction on rainfall but not PET data. In addition, PET was averaged over a month with the model used by Lockart et al. (2016) while the rainfall data from NARCliM was averaged on a daily scale.

In commenting on the input databases and the hydrology models, Lockart et al. (2016) acknowledge that examining the performance of the hydrological models in simulating the streamflow and reservoir levels themselves was beyond the scope of their work, and also that different observed data inputs (from different sources) can also lead to different outcomes. For example, for the Chichester catchment, SILO rainfall was about 25% greater than AWAP rainfall in summer and autumn, leading to greater streamflow and reservoir volumes using SILO input data. This is of note given discussions in Section 2.2.1 on choices between AWAP and SILO, as it demonstrates that the two can differ even though they draw data from the same gauges.

DPIE-Water also noted their data QA process picked up issues with rainfall data in the Chichester catchment, concluding that rainfall data was biased due to poor sampling in high altitude/higher rainfall areas.

NARCliM 1.0 has been criticised for projecting future climate over relatively short timeframes of 20 years, which means that sampling variability may mask any climate change signal; this will be addressed in NARCliM 1.5 and 2.0 versions with 150 year projections (Table 1). Discussions with the developers of NARCliM indicates that the 2.0 version will address shortcomings relating to bias correction of ET. Another issue raised by some practitioners is NARCliM's inability to simulate persistence over multi-year/decadal scale and the short-term scale. It is this issue of persistence that was a major reason behind DPIE-Water undertaking the paleoclimate and stochastic work that is described in the methodology, as it is a method to develop statistically representative inputs of a climate, rainfall and evapotranspiration variability.

2.4.3 Application of downscaled model outputs to factor stochastic data sets

Using GCMs/RCMs as stand-alone input to hydrological modelling is not advisable at this time, but these tools can be used as part of a system to inform the hydrological modelling. The DPIE-Water methodology is doing this – drawing together multiple lines of evidence (e.g. instrumental and paleo data, stochastic data sets, climate modelling) to develop the input to hydrological modelling and using NARCliM to provide factors to scale the stochastic data sets.

2.4.3.1 IPO Stochastic NARCliM method (Lachlan & Macquarie Catchments)

The DPIE-Water methodology developed for the regions of NSW affected by IPO states (north and west NSW) is to develop the stochastic data sets and then incorporate future climate risk through re-factoring the stochastic data sets using NARCliM 1.0 results. The proposed methodology also contemplates impacts of the IOD and SAM drivers in south/west NSW (Figure 2), however the Panel has not reviewed any work for catchments in that zone.

DPIE-Water has advised the Panel that it uses the GCM/RCM combination in NARCliM rainfall data that provides the lowest future rainfall scenario to stress test the water systems. The ratios of these data (presumably daily mean rainfall data) at all locations were calculated from bias-corrected NARCliM data sets, which were used to re-factor the stochastic data sets. Similarly, the PET results from the NARCliM scenarios were used to factor the PET results from the stochastic synthesis.

The Panel considered this a relatively cautious approach to incorporating climate change signals into the stochastic data. However, more information about this is required in order to provide more detailed advice on strengths and limitations.

Given the limitations of the NARCLiM 1.0 approach regarding rainfall simulation at small scales, the proposed scaling approach represents a low risk way of extracting information from the NARCliM outputs without compromising the variability and persistence attributes replicated by the stochastic models. Simple scaling preserves the coefficient of variation and autocorrelations at all scales. Notwithstanding that judgment, the degree of confidence in this approach is understandably low at the present time due the shortcomings reflected in Section 2.4.2. It does not seem wise to rely solely on the NARCliM 1.0 rainfall-scaling approach for climate change guidance relating to hydroclimate variables; applying the NARCliM 1.0 should be undertaken in consideration of the bias in PET and in the selection of the GCMs used in NARCliM 1.0. Looking forward, a number of the shortcomings highlighted will be addressed and additional lines of evidence could be explored. Monitoring climate trends for NSW and considering these with the climate models will be an important piece of work for the community practice.

Several shortcomings with NARCliM 1.0, such as PET bias correction and 20-year projections, identified in the Lockart paper are planned to be addressed in the soon-to-be released NARCliM 1.5 or the NARCliM 2.0, which is still under development and yet to be tested (EES, pers comm, February 2020). The Panel notes that NARCliM 1.5 is due for release in early 2020, with data currently available to NSW Government agencies. The Methods Paper states that further analysis will be undertaken to test the outcomes using NARCliM data, including testing the NARCliM 1.5 data, but the Methods Paper does not specify details about how this will be done. The Methods Paper notes the result will be longer periods of simulation and potentially provide a means to better understand projected changes to inter-annual and multi-decadal variability. Further information is required in order to assess the merits and limitations to the methodology once NARCliM 1.5 and 2.0 are available.

2.4.3.2 ECL NARCliM (Bega)

The Bega River region stochastic model described in the Methods paper can be classified as a conditional stochastic model. In this application, NARCliM 1.0 is used to simulate the changes in the frequency and timing of ECLs expected in a 2050 climate. Because ECLs are synoptic scale phenomena one would expect more confidence in NARCliM 1.0 outputs. The DPIE-Water approach uses NARCliM information on ECL frequency and timing to modify ECLs in the historical record and then calibrates the stochastic model to the climate change adjusted historical record. There is scope to improve the modification algorithm. The most obvious improvement is to replicate the modifications to the historical record to reduce the impact of sampling variability.

While maximum temperature has been identified as a suitable covariate, there is scope to investigate other covariates that are reproduced with good skill by climate models.

2.5 QUALITY ASSURANCE

The Methods Paper states that additional quality control analysis of the data is being implemented in addition to quality control processes undertaken by data suppliers (i.e. the development of the stochastic data sets by Kiem (2019), Verdon-Kidd (2019) and Leonard et al. (2019). The background papers (e.g. Verdon-Kidd (2019)) make note of iterative testing of the stochastic data by DPIE-Water during their development.

In addition, DPIE-Water states in the Methods Paper that further analysis will be undertaken to test the outcomes from using the NARCliM data, including testing of NARCliM 1.5. NARCliM 1.5 will have longer periods of simulations and is expected to better reflect projected inter-annual and multi-decadal variability.

The methodology behind this work needs to be laid out before a detailed assessment can be made on the merits and limitations. All Panel comments regarding application of NARCliM are in the section above.

2.6 OBSERVATIONS ON RAINFALL & EVAPOTRANSPIRATION AND INCORPORATION INTO HYDROLOGICAL MODELS

Hydrological models developed using historical data implicitly assume stationarity – by calibrating using historical data – notwithstanding that global warming makes climate non-stationary (Milly et al., 2008). The Panel notes that work of Saft et al. (2016) has shown statistically significant changes in annual rainfall-runoff relationships in some catchments during the Millennium drought showing that drying catchments may see a change of hydrological function. Their studies in eastern Australia for Millennium drought suggest models calibrated to predrought data overpredict if there are declining water resources, which may be related to groundwater, vegetation and soil (Saft et al., 2016). DPIE-Water has also informed the Panel that this has also been their experience in previous droughts (DPIE-Water, pers comm, March 2020).

Chiew (2006) developed the concept of rainfall elasticity of streamflow, presented for 219 catchments across Australia. Elasticity is defined as the 'proportional change in mean annual streamflow divided by the proportional change in mean annual rainfall' (Chiew, 2006). The work showed that for 1% change in mean annual rainfall results in 2.0-3.5% change in mean annual streamflow, which was observed in about 70% of the catchments. This work showed that relatively small changes to annual rainfall were amplified in streamflow.

It has been known for 25 years that increasing CO_2 is important to how rainfall is translated between evaporation and runoff (Field, Jackson, & Mooney, 1995). Evapotranspiration is also not just about soil moisture and evaporation from waterbodies but also about vegetation response.

There is a huge literature highlighting how increased CO_2 changes the water use efficiency of a tree, and how this impacts runoff, including in Australia (Ukkola et al., 2015), however the role of CO_2 fertilisation and biomass accumulation appear not to be captured in the DPIE-Water Methods Paper. Current rainfall-runoff models are crude when it comes to representing vegetation dynamics, and as the dynamics seem complicated there may be competing forces at work, e.g. reductions in stomatal conductance due to higher CO_2 may be offset by increases in leaf area as vegetation seeks to maximize its biomass.

With longer periods of warmer temperature, vegetation would have longer growing seasons, and the role of CO_2 fertilisation and biomass accumulation does not appear to be captured in the work. By not using hydrological models that include the impacts of climate change – both the climate impacts and the impacts of CO_2 on stomatal control of evaporation - the model tends to include the downside risk (drying for example) but not how the vegetation adapts to water stress to reduce evaporation. As a consequence, it is very likely that in the far-future (impossible to define but beyond 2050) the models used fail to include the key process which manages the landscape response to climate change.

2.7 FUTURE DIRECTIONS

Understanding uncertainties

Any assessment of climate change for 2050 or 2100 will contain uncertainties and assessing the magnitude of uncertainty is an on-going research challenge. There is low confidence in the ability to estimate the probability of future events, at a detailed regional scale, for future climate. There are ways to protect against systematic biases however. This might include avoidance of small future samples of climate change (e.g. avoid using a small number of climate models), careful selection of RCM simulations, accounting for model independence and so on.

Communicating and managing uncertainties

Communication and management of uncertainty is a critical subject in its own right and thus should be a focus area. If communicating uncertainty is done poorly, science or engineering-based studies can be compromised and falter. A poor effort in communicating uncertainty, either to practitioners or key stakeholders will cause confusion for people trying to make sense of and respond to challenges of risk, including in this setting of climate risk and water security.

Approaches for communicating and managing risk in the face of uncertainty could be a focus area for the community of practice proposed in Recommendation 10. Central to this is recognising the different types and sources of uncertainty and establishing an adaptive planning approach that builds uncertainty into decisions making, to account for an emerging knowledge base. The primary focus of adaptive planning should be on preparedness and resilience rather than a prediction paradigm.

Test new climate models as they are released

With each new generation of RCMs, the following should be monitored and evaluated:

- Between-model differences in projections for NSW (Is the level of 'across-model' uncertainty decreasing, stabilising or increasing?).
- Between-model consensus in projected characteristics of modes of climate variability (e.g. El Niño) and intense low-pressure systems that have a major impact on water availability.
- In the context of water security, assessment of each new generation should be conducted in a fit-for-purpose context, such as a study for urban systems and a rural basin with irrigation-ecological trade-offs.

Where possible multiple approaches to assessing future climate risk could be explored and compared. An option is the use of stochastic models conditioned on climate covariates. This would be an extension of the methodology used in current stochastic models and exploits the strength of using covariates at synoptic scales. These could accept information from climate models on how associations may change in the future.

New data sets

The new BARRA product, developed by BOM, NSW Government (Rural Fire Service), Tasmania, South Australia, Western Australia and other entities, is an example of a newly developed high-resolution data set. The BARRA website describes it thus: *"A meteorological 'reanalysis' takes all available observations and uses a weather model to consistently fill in the fine detail, both at the surface where people live and in the atmosphere as well. Reanalysis datasets are extremely valuable because they provide a consistent method of analysing the atmosphere over a number of years or decades giving greater understanding of the weather over Australia, including extreme events"* (BOM, 2020). Opportunities to explore the possible information from such new datasets should be taken.

Persistence

Persistence is likely to continue to be an issue for climate models; there is evidence from the weather forecasting science that persistence requires spatial resolution of ~25 km which is infeasible in the GCMs for at least a decade. GCMs have some shortcomings that cannot be rectified with the RCM, mainly due to land surface processes and some large-scale processes. However, it may be that as GCM models improve, NARCliM will better model short-term persistence, as there is some evidence that CMIP6 models are providing better simulations of short-term persistence.

The Panel has been informed that NARCliM developers will look to include an assessment of model skill in representing persistence into the NARCliM 2.0 model selection process, while they are also looking to construct probability distributions for a range of continuous dry/wet periods across the whole NARCliM ensemble (more than 5000 modelled years when NARCliM 2.0 is complete).

Future NARCliM versions

For the RWS, as new versions of NARCliM become available, assure that the methodologies are sensitive to the new capabilities offered by later versions of the NARCliM data. DPIE-Water should begin the process of planning to incorporate NARCliM 1.5 into calculations in the short term and over the long-term, explore the capabilities of NARCliM 2.0 and investigate the use of this for RWS modelling and planning in combination with stochastic data sets as required.

2.7.1 Departmental community of practice on climate science and its application

The development and use of climate change modelling and predictions is a complex field, that will increasingly be applied in diverse industries. Climate experts will need to increase connections with the end-users of models and datasets to ensure that climate science is applied, interpreted and communicated in an appropriate accurate and balanced way.

DPIE has a wide multi-sectoral remit, including environment, energy, agriculture, water, and is a major manager of infrastructure itself. This means that new developments in climate science will be relevant across the DPIE cluster and for its stakeholders beyond.

New data sets and models are emerging, as are new innovation opportunities in climate and hydrology fields, that will require assessment and application.

Therefore, it is proposed that a cross-DPIE community of practice be convened to consider new climate and weather modelling and data products of relevance to NSW – such as the DPIE-Water methodology, NARCliM version updates, etc.

In regard to the RWS, bringing together staff within DPIE who are developing the strategies with those undertaking other forecasting activities will help clarify practitioner needs and provide insights into future model developments.

The community of practice could also be tasked with maintaining ongoing vigilance of new and different developments in fields of climate science and hydrological modelling, including approaches that are yet to emerge. These approaches can then be scrutinised, and a view formed about whether and how they should be considered in the NSW perspective.

The group would explore other methodological improvements such as incorporation of CO_2 fertilisation into modelling.

Possible focus areas could include the following:

a) **Identifying best practice** - Within DPIE, understanding of climate impacts into the future will be important across a range of policy areas, including in water management, ecosystem management, water planning, agriculture, infrastructure, etc. The community of practice can assist in developing a shared understanding of new scientific directions, the applications of new models into practice, experience of using different modelling approaches, as well as policy developments and their implications for NSW. It can also help the fields of hydrology and climate science to better align within DPIE's activities.

b) **Understanding uncertainty** - There is deep unquantifiable uncertainty and hence low confidence in estimates of the probability of some future climate events, particularly rare events. Where possible, multiple approaches to assessing future climate risk could be considered and adopted by the community of practice work program as relevant. Assessing the magnitude of this uncertainty is an on-going challenge. While quantifying the uncertainty is not currently a robust area of climate science, there are ways to protect against systematic biases.

c) **Communicating and managing uncertainty** - Approaches for communicating and managing risk in the face of uncertainty should be a focus area. Central to this is recognising the different types and sources of uncertainty and establishing an adaptive planning approach that builds uncertainty into decisions making, to account for an emerging knowledge base. Therefore, the primary focus of adaptation planning should be a preparedness and resilience rather than a prediction paradigm.

d) **Adaptive management** - Best practice approaches to adaptive management, where the system facilitates ongoing improvement, stress testing, comparing real world and modelled scenarios.

e) **Assessing new approaches** - Regularly monitor developments in fields of climate science and hydrological modelling, including new conditional stochastic modelling approaches that will emerge. These approaches can then be scrutinised, and a view formed about whether and how they should be considered for use in NSW.

f) **Contributing to technical improvements** - The developers of the RWS and NARCliM could engage on a defined and on-going basis to develop a common understanding of the strengths and weaknesses of alternative approaches and write a joint white paper on whether, and if so, how to merge the approaches. This should be completed by the end of 2020 so it can potentially feed into NARCliM 2.0.

g) **New approaches to ET** – Implementing physical approaches to ET and PET (e.g. replacing current approaches for FAO56 and Mwet).

h) **Observational networks and monitoring -** There should be a long-term commitment to identifying and maintaining reference/benchmark observational networks. Maintenance of regional observational networks with long-term, high-quality records is needed to facilitate detection of emerging climate-related trends; evaluation of predictive models under changing hydroclimatic conditions; and performance assessments for potential adaptation strategies.

i) **Monitoring climate trends** - Track and monitor emerging or relevant climate trends (e.g. if winter rainfall in South-Eastern Australia/NSW is decreasing) many of which are provided by the Bureau of Meteorology: http://www.bom.gov.au/climate/change/#tabs=Tracker&tracker=timeseries

j) **Research –** Identify and potentially commission research to support new knowledge.

k) **National cooperation –** promote cross-jurisdictional dialogue in relation to hydro-climatological practices with water management agencies in other jurisdictions nationally to exchange knowledge and experiences.

ACRONYMS

Acronym	Complete term
ACT	Australian Capital Territory
ADAM	Australian Data Archive for Meteorology
AWAP	Australian Water Availability Project
BARRA	BOM Atmospheric high-resolution Regional Reanalysis for Australia
BOM	Australian Bureau of Meteorology
CMIP3	Coupled Model Inter-comparison Project phase 3
CMIP5	Coupled Model Inter-comparison Project phase 5
CMIP6	Coupled Model Inter-comparison Project phase 6
CO_2	Carbon Dioxide
CORDEX	Coordinated Regional Climate Downscaling Experiment
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DAPP	Dynamic Adaptive Policy Pathways
DPIE	Department of Planning, Industry and Environment
EES	DPIE Environment, Energy and Science branch
ECL	East Coast Lows
ENSO	El-Niño Southern Oscillation
ET	Evapotranspiration
FAO56	Food and Agricultural Organization of the United States Irrigation and Drainage Paper No.
GCM	Global Climate Models
GL	Gigalitres
IOD	Indian Ocean Dipole
IPCC	Intergovernmental Panel on Climate Change
IPO	Interdecadal Pacific Oscillation
IQQM	Integrated Quantity and Quality Model software
Mwet	Morton's wet-environment areal evapotranspiration overland
NARCliM	NSW and ACT Regional Climate Modelling
NSW	New South Wales
OEH	NSW Office of Environment and Heritage (former)
PET	Potential evapotranspiration
PDO	Pacific decadal oscillation
RCM	Regional Climate Model
RWS	Regional Water Strategies
SAM	Southern Annular Mode
SILO	Queensland Government database of continuous daily climate data for Aust from 1889 to present
WATHNET	Water supply headworks simulation software package

REFERENCES

Alluvium. (2019). *Critical review of climate change and water modelling in Queensland - Final Report*. Prepared collaboratively with CSIRO and University of Newcastle on behalf of the Queensland Water Modelling Network, Queensland Government, Brisbane

Berghout, B., Henley, B.J., & Kuczera, G. (2017). Impact of hydroclimate parameter uncertainty on system yield. *Australasian Journal of Water Resources, 21*(2), 53-62. doi: 10.1080/13241583.2017.1404550

BOM, Bureau of Meteorology. (2020). *Atmospheric high-resolution regional reanalysis for Australia*. Retrieved 28 February 2020, from http://www.bom.gov.au/research/projects/reanalysis/

Chiew, F.H.S. (2006). Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Sciences Journal, 51*(4), 613-625. doi: https://doi.org/10.1623/hysj.51.4.613

CSIRO, & BOM, *Bureau of Meteorology*. (2015). *Climate Change in Australia Information for Australia's Natural Resource Management Regions: Technical Report*. CSIRO and Burea of Meteorology, Australia

Data.gov.au. (2019). *Australian Gridded Climate Data (AGCD)/AWAP; v1.0.0 Snapshot (1900-01-01 to 2018-12-31)*. Retrieved 25 February 2020, from https://data.gov.au/dataset/ds-bom-ANZCW0503900567/details?q=

Deacon, B. (2020, 17 February 2020). Indian Ocean Dipole linked to global warming in new research by Australian scientists, *ABC News*. Retrieved from https://www.abc.net.au/news/2020-02-13/indian-ocean-dipole-linked-to-global-warming-in-new-research/11943178

DOI, *NSW Department of Industry*. (2018). *Greater Hunter regional water strategy - Securing the future water needs of the Hunter, Central Coast and Mid-Coast areas*.

Field, C.B., Jackson, R.B., & Mooney, H.A. (1995). Stomatal Responses to Increased Co2 - Implications from the Plant to the Global-Scale. *Plant Cell and Environment, 18*(10), 1214-1225. doi: DOI 10.1111/j.1365-3040.1995.tb00630.x

Franke, J., Frank, D., Raible, C.C., Esper, J., & Brönnimann, S. (2013). Spectral biases in tree-ring climate proxies. *Nature Climate Change, 3*(4), 360-364. doi: 10.1038/nclimate1816

Henley, B.J., Thyer, M.A., Kuczera, G., & Franks, S.W. (2011). Climate-informed stochastic hydrological modeling: Incorporating decadal-scale variability using paleo data. *Water Resources Research, 47*(11). doi: 10.1029/2010wr010034

Hope, P., Timbal, B., Hendon, H., Ekstrom, M., & Potter, N. (2017). *A synthesis of findings from the Victorian Climate Initiative (VicCI)*. Bureau of Meteorology, Australia

Infrastructure NSW. (2018). *Building Momentum: State Infrastructure Strategy 2018-2038*. Retrieved from https://insw-sis.visualise.today/the-strategy.html

Jakob, D., Imielska, A., Charles, S., Fu, G., Frederiksen, C., Frederiksen, J., Zidikheri, M., Hope, P., Keay, K., Ganter, C.J., Li, Y., et al. (2012). *Indian Ocean Climate Initiative Stage 3: Summary for Policymakers*. CSIRO and Bureau of Meteorology, Australia

Kiem, A.S. (2019). *Incorporating changes in East Coast Low (ECL) behaviour into stochastically generated hydroclimatic data for the Bega River region, New South Wates, Australia*. Report prepared for NSW Department of Industry by the University of Newcastle

Kiem, A.S., Kuczera, G., Kozarovski, P., Zhang, L., & Willgoose, G. (Under Review). Stochastic generation of future hydroclimate using temperature as a climate change covariate.

Leonard, M., Westra, S., & Bennett, B. (2019). *Multisite rainfall and evaporation data generation for the Macquarie Water infrastructure project*. Report prepared for the NSW Department of Industry by the University of Adelaide

Lockart, N., Willgoose, G., Kuczera, G., Kiem, A., Chowdhury, A.F.M.K., Manage, N., Zhang, L., & Twomey, C. (2016). Case study on the use of dynamically downscaled climate model data for assessing water security in the Lower Hunter region of the eastern seaboard of Australia. *Journal of Southern Hemisphere Earth System Science, 66*, 177-202. doi: 10.22499/3.6002.007

McMahon, T.A., Peel, M.C., Lowe, L., Srikanthan, R., & McVicar, T.R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrology and Earth System Sciences, 17*, 1131-1363.

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*(5863), 573-574. doi: 10.1126/science.1151915

Nicholls, N., Drosdowsky, W., & Lavery, B. (1997). Australian rainfall variability and change. *Weather, 52*(3), 66-72. doi: 10.1002/j.1477-8696.1997.tb06274.x

Palmer, Jonathan G., Cook, Edward R., Turney, Chris S. M., Allen, Kathy, Fenwick, Pavla, Cook, Benjamin I., O'Donnell, Alison, Lough, Janice, Grierson, Pauline, & Baker, Patrick. (2015). Drought variability in the eastern Australia and New Zealand summer drought atlas (ANZDA, CE 1500–2012) modulated by the Interdecadal Pacific Oscillation. *Environmental Research Letters, 10*(12), 124002. doi: 10.1088/1748-9326/10/12/124002

Power, S.B., Murphy, B., Chung, C., Delage, F., & Ye, H. (2017, 9 February). Droughts and flooding rains already more likely as climate plays havoc with Pacific weather, *The Conversation*. Retrieved from https://theconversation.com/droughts-and-flooding-rains-already-more-likely-as-climate-change-plays-havoc-with-pacific-weather-71614

Queensland Government. (2020). *SILO - Australian Climate Data from 1889 to Yesterday*. Retrieved 19 February 2020, from https://www.longpaddock.qld.gov.au/silo/

Saft, Margarita, Peel, Murray C., Western, Andrew W., Perraud, Jean-Michel, & Zhang, Lu. (2016). Bias in streamflow projections due to climate-induced shifts in catchment response. *Geophysical Research Letters, 43*(4), 1574-1581. doi: 10.1002/2015gl067326

Sorooshian, S., & Martinson, D.G. (1995). *Proxy indicators of climate: An essay Natural Climate Variability on Decade-to-Century Time Scales*. Washington DC: The National Academies Press.

Stephens, C.M., Marshall, L.A., Johnson, F.M., Lin, L., Band, L.E., & Ajami, H. (2020). Is Past Variability a Suitable Proxy for Future Change? A Virtual Catchment Experiment. *Water Resources Research, 56*(2), e2019WR026275. doi: 10.1029/2019wr026275

Ukkola, A.M. , Pitman, A.J. , De Kauwe, M.G. , Abramowitz, G., Herger, N., Evans, J.P. , & Decker, M. (2018). Evaluating CMIP5 Model Agreement for Multiple Drought Metrics. *Journal of Hydrometeorology, 19*(6), 969-988. doi: 10.1175/jhm-d-17-0099.1

Ukkola, A.M., Prentice, I.C., Keenan, T.F., van Dijk, A.I.J.M., Viney, N.R., Myneni, R.B., & Bi, J. (2015). Reduced streamflow in water-stressed climates consistent with CO2 effects on vegetation. *Nature Climate Change, 6*, 75-78. doi: 10.1038/nclimate2831

Verdon-Kidd, D.C. (2019). *Development of multi-site rainfall and evaporation data for the Lachlan Regional Water Strategy*. Report prepared for the NSW Department of Industry by the University of Newcastle

Wasko, C., & Sharma, A. (2017). Continuous rainfall generation for a warmer climate using observed temperature sensitivities. *Journal of Hydrology, 544*, 575-590. doi: https://doi.org/10.1016/j.jhydrol.2016.12.002