

Assessment of non-stationarity for stochastic time series generation in the southern basin

Anjana Devanand, Michael Leonard and Seth Westra



Disclaimer

The Intelligent Water Decisions research group has taken all reasonable steps to ensure that the information contained within this report is accurate at the time of production. In some cases, we may have relied upon the information supplied by the client.

This report has been prepared in accordance with good professional practice. No other warranty expressed or implied is made as to the professional advice given in this report.

The Intelligent Water Decisions research group maintains no responsibility for the misrepresentation of results due to incorrect usage of the information contained within this report.

This report should remain together and be read as a whole.

This report has been prepared solely for the benefit of the NSW Department of Climate Change Energy, the Environment and Water.

Department reference number: PUB24/330

No liability is accepted by the Intelligent Water Decisions research group with respect to the use of this report by third parties without prior written approval.

Contact

Intelligent Water Decisions research group
School of Civil, Environmental and Mining Engineering
University of Adelaide

Physical: Engineering North Building North Terrace Campus University of Adelaide, Adelaide

Postal: University of Adelaide, Adelaide, 5005

Telephone: (08) 8313 5451

Contents

Executive summary	7
Stationarity in temperature, rainfall and evapotranspiration.....	7
Rainfall	7
Evapotranspiration.....	8
Temperature.....	8
Influence of climate drivers on meteorological variables in the southern basin	8
The importance of stochastic generation calibration period on key statistics	9
Recommendations for stochastic data generation.....	9
1 Introduction	10
2 Large-scale patterns of variability and change relevant to south-east Australia	12
2.1 Influence of large-scale drivers: ENSO, IOD, SAM and IPO	12
2.2 Influence of east coast lows on the south-east coast.....	13
3 Climatic trends in the historical record of south-east Australia	14
3.1 Rainfall.....	14
3.2 Pan evaporation.....	18
3.3 Temperature.....	19
4 Future climate projections for south-east Australia	21
5 Pilot study: assessment of the historical record	23
5.1 Data	23
5.2 Methodology	26
5.2.1 Trend detection	26
5.2.2 Split sample stochastic simulations	28
5.3 Results of non-parametric trend testing.....	30
5.3.1 Trends in rainfall attributes	30
5.3.2 Trends in evapotranspiration	41
5.3.3 Trends in temperature	42
5.4 Results of split sample testing	44
5.4.1 Rainfall.....	44
5.4.2 Evapotranspiration	48
6 Summary of key findings from assessment of non-stationarity	50
7 Options for stochastic time series generation in the southern basin	52

7.1	Use the entire historical record to calibrate the stochastic model.....	53
7.2	Use a climatological baseline to calibrate the stochastic model.....	53
7.3	Use a hybrid baseline of the historical climate to calibrate the stochastic model	54
7.4	Use the inverse method to generate a current or future climate	54
7.5	Use weather typing to capture specific mechanisms of generation of precipitation.....	55
7.6	Recommendation for stochastic generation.....	57
8	References.....	58
	Appendix A – Supplementary material.....	62

List of figures

Figure 1 Southern basin of New South Wales.....	10
Figure 2 Anomalies of April–October rainfall for south-east Australia (south of 33°S, east of 135°E inclusive).....	14
Figure 3 Projected changes (compared to 1986–2005) in annual mean rainfall in the Ovens and Murray catchments for medium emissions (top) and high emissions (bottom).....	22
Figure 4 Location of rainfall, evapotranspiration and temperature pilot sites.....	24
Figure 5 Location of pilot temperature sites and 3 ACORN–SAT v2 stations in the vicinity	25
Figure 6 Annual mean Tmax and Tmin at site 82039 (Rutherglen Research) from raw station data and homogenised ACORN–SAT v2 data with linear trendlines.....	26
Figure 7 Split sample experiments.....	29
Figure 8 Short-term (1950 to 2018) seasonal trends in number of wet days (in days/decade).....	40
Figure 9 Short-term (1950 to 2018) seasonal trends in total rainfall (in mm/decade).....	41
Figure 10 Histograms of the mean differences in total rainfall in MAM during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites	45
Figure 11 Histograms of the mean differences in total rainfall in DJF during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites	46
Figure 12 Histograms of the mean differences in the number of wet days in MAM during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites	47
Figure 13 Histograms of the mean differences in total annual evapotranspiration during the calibration and validation time periods from stochastic simulations (in mm) at 30 pilot sites.....	49
Figure A.1 Annual mean short-term (post-1960) Tmax and Tmin at site 82039 (Rutherglen Research) from raw station data and homogenised ACORN–SAT v2 data with linear trendlines ...	62
Figure A.2 Histograms of the mean differences in total rainfall in JJA during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites	63
Figure A.3 Histograms of the mean differences in total rainfall in SON during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites	64
Figure A.4 Histograms of the mean differences in number of wet days in JJA during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites.....	65
Figure A.5 Histograms of the mean differences in number of wet days in SON during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites	66
Figure A.6 Histograms of the mean differences in number of wet days in DJF during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites	67
Figure A.7 Histograms of the mean differences in annual extreme rainfall intensity during the calibration and validation time periods from stochastic simulations (in mm/day) at 49 pilot sites	68

List of tables

Table 1 Summary of reported precipitation trends in the historical record of south-east Australia	16
Table 2 Summary of reported evapotranspiration trends in the historical record of south-east Australia.....	19
Table 3 Summary of reported temperature trends in the historical record of south-east Australia	20
Table 4 Summary of number of different observation time series by variable type	24
Table 5 ACORN–SAT v2 stations used for analysis.....	26
Table 6 Attributes of hydroclimatic variables and the periods used for trend analyses.....	28
Table 7 Number of sites that exhibit significant trends in annual and seasonal attributes of rainfall	31
Table 8 Trends in rainfall attributes in the pilot sites and the changes reported in literature	36
Table 9 Median magnitude of trends at 49 pilot sites	39
Table 10 Number of sites that exhibit significant trends in annual and seasonal total evapotranspiration (Total M _{wet} in mm)	42
Table 11 Trends in evapotranspiration in the pilot sites and the changes reported in literature ...	42
Table 12 The magnitude of trends in T _{max} at the ACORN–SAT v2 sites and the trends reported in literature (°C/decade).....	43
Table 13 The magnitude of trends in T _{min} at the ACORN–SAT v2 sites and the trends reported in literature (°C/decade).....	44
Table 14 Summary of options for stochastic time series generation in the southern basin, with advantages/disadvantages described in the context of a non-stationarity climate signal.....	56

Executive summary

The New South Wales Department of Climate Change Environment Energy and Water (the department) has adopted a risk-based methodology to account for climate variability and change in developing its regional water strategies, in which the risk assessment is informed by the use of stochastically generated long-term sequences that reflect climate variability beyond that contained within the instrumental record. An independent expert panel review of this climate risk method commended many aspects of the methodology, but recommended:

- **further review of the potential non-stationarity of historical climate data used to inform the stochastic models**, 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, by
 - assessing non-stationarity of the historical record
 - split sample testing of the stochastic model
- **further articulation of the effects of multiple climate drivers**, to develop the methodology for stochastic generation in regions where multiple climate drivers influence the regional hydroclimate.

These recommendations are addressed in this report in the context of rainfall, evapotranspiration, and temperature for the southern basin, using a 'multiple lines of evidence' approach to determine the presence and potential causes of any non-stationarity. Specifically, this report documents a review of literature on the physical mechanisms influencing the regional climate of south-east Australia, reported trends in hydroclimatic variables, and future projections in this region. This review is complemented by a pilot study using data from the Ovens, Upper Murray, and Snowy catchments to assess trends in historical record and implications for stochastic simulations generated using the data. The pilot analysis used 49 rainfall time series and 30 evapotranspiration time series from these catchments, and 3 temperature time series at nearby locations from a homogenised dataset. The trends in seasonal and annual evapotranspiration, temperature and attributes of rainfall are studied using the Mann Kendall test. Split sample stochastic simulations are performed on the rainfall and evapotranspiration time series.

The main findings are summarised below.

Stationarity in temperature, rainfall and evapotranspiration

Rainfall

Literature documents decreasing trends in cool season (April to October) rainfall by 10–20% in south-east Australia since the mid-1990s, predominantly in autumn and early winter. There are accompanying decreasing trends in the number of wet days during the cool season. Literature documented that 'the decline in rainfall across south-eastern Australia was at least partly attributable to climate change' (CSIRO 2012, p. 4) and 'drying across southern Australia cannot be explained by natural variability alone' (CSIRO and BOM 2015, p.45). Victoria Climate Projections 2019 (VCP19) project median annual rainfall decreases in the Ovens catchment during the 2020–2039 period amounting to about 6% under the medium emission scenario and about 11% under the high emission scenario (uncertainty range from –3% to –18%), with respect to a 1986–2005 baseline.

The results of the trend assessment using data from the pilot sites are in general agreement with literature, showing decreasing trends in cool season rainfall and the number of wet days. The median total autumn rainfall trends amount to -5.5% per decade at the pilot sites over the period from 1950 to 2018. Thus, multiple lines of evidence indicate that the historical rainfall record in this region is non-stationary due to the changes in the cool season. The pilot sites also show a short-term decline in spring (SON) rainfall and increase in intensity of extreme rainfall intensity. These trends are not in agreement with all literature, but consistent with regional studies in nearby catchments.

Evapotranspiration

Literature documents decreasing trends in pan evaporation over the period 1975–2002, whereas studies using more recent data (up to 2018) report insignificant increasing trends. There are long-term increasing trends in annual Morton wet evapotranspiration at the pilot sites, whereas short-term trends are insignificant. The pilot study results are not directly comparable with available literature based on pan evaporation data given differences in processes that drive pan evaporation and Morton wet evapotranspiration. Thus, non-stationarity in the historical record of evapotranspiration remains highly uncertain. VCP19 projects 8 to 10.8% increases in pan evaporation over the period 2020–2039.

Temperature

Literature documents temperature increases in the southern basin region, especially post-1960. The Climate Change in Australia initiative reports an increase of $0.8\text{ }^{\circ}\text{C}$ in mean annual temperature in the Murray basin cluster (which contains the pilot catchments) over the period 1910–2013 assuming a linear trend, with higher trends for temperature minimums than for maximums. This is broadly consistent with a mean annual temperature increase across Australia of just over $1\text{ }^{\circ}\text{C}$ during the slightly longer period 1910–2018. Moreover, climate projections indicate that increases of $0.6\text{--}1.3\text{ }^{\circ}\text{C}$ are expected in the Murray basin cluster in the near term (2020–2039) with respect to a 1986–2005 baseline. Trends in homogenised temperature sites near the pilot catchments are broadly consistent with these findings, with increases of $0.8\text{--}1.5\text{ }^{\circ}\text{C}$ in maximum temperatures and $1.9\text{--}2.6\text{ }^{\circ}\text{C}$ in minimum temperatures during the period 1913–2018. These multiple lines of evidence are in agreement that there is non-stationarity in the temperature record of the southern system.

Influence of climate drivers on meteorological variables in the southern basin

The documented influence of climate drivers on key meteorological variables in the southern basin is summarised in this report. The declining cool season rainfall is associated with an expansion of the tropics, increasing intensity of the subtropical ridge over the continent and positive trends in the Southern Annular Mode (SAM). Literature indicates that these changes in large-scale patterns during the cool season are at least partly attributable to climate change. Other climate drivers, notably El Niño Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD), influence the interannual variability of regional rainfall, primarily affecting rainfall in winter and during the warm season (CSIRO 2012; CSIRO and BOM 2015; Hope et al. 2017).

The importance of stochastic generation calibration period on key statistics

Consistent with the trend assessment, the split sample tests demonstrate that the period used for calibration can significantly influence the statistics of the simulated data. In particular, the split sample tests show that the Millennium Drought is a 'high-leverage' event, in the sense that statistics of the simulated rainfall can vary significantly depending on whether the drought is included in the calibration period or validation period.

When the drought is included in the calibration period, the extent of similarity between the simulated rainfall and observed data following the drought is dependent upon the statistic and season under consideration. For example, inclusion of the drought in the calibration period brings the autumn rainfall in the simulated data close to recent (2010–2018) observations (reduction of biases in simulations from +24% to –4%) but can cause larger deviations in simulated summer rainfall from recent observations (increase in biases in simulations from +16% to +29%).

The implication of non-stationarity is that the calibration period needs to be carefully considered in the context of any southern basin stochastic analysis, with the appropriate approach likely to depend on whether the objectives are to represent risk over historical, current or future periods.

Recommendations for stochastic data generation

The calibration period should be as long as possible while also providing consistency with the method of assessment of climate projections. NARcliM 1.5 (1950–2005) has a suitable minimum baseline length for hydrological studies, which is significantly longer than the NARcliM 1.0 baseline.

The complexity of change in climate attributes gives rise to the possibility that scaling by the application of change factors to seasonal or annual totals may not lead to appropriate adjustments of the other statistics. For this reason, quantile scaling methods should be applied in preference to simple scaling.

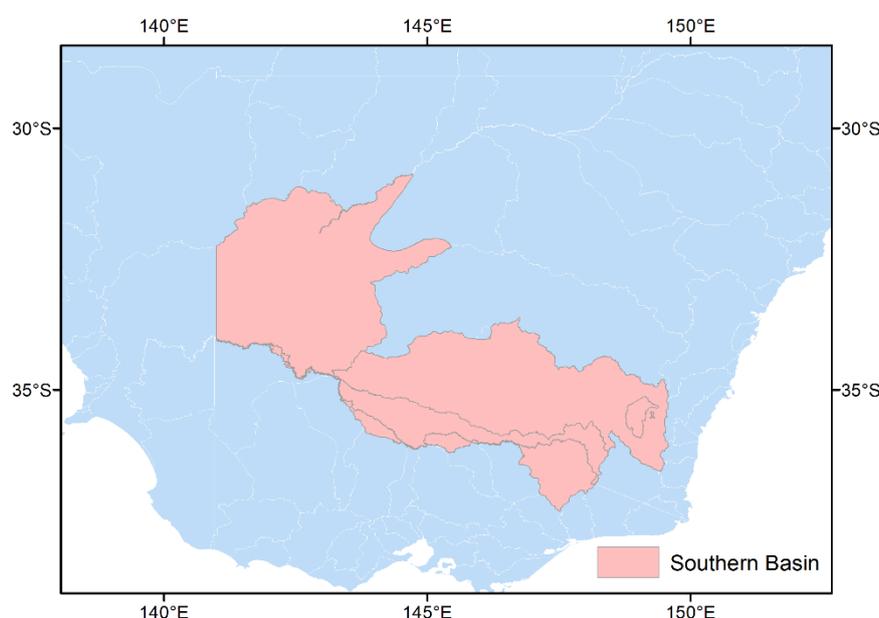
The scaling approach should be applied to different RCMs rather than an ensemble mean given the likely variation between models. Due to the nonlinear nature of hydrological transformation, this approach will ensure that model uncertainty in hydrological estimates is suitably accounted for.

1 Introduction

The department has adopted a risk-based methodology to account for climate variability and change in developing its regional water strategies. The method involves the use of stochastically generated long-term sequences of climate data to characterise the current climate, and the application of scaling factors to the stochastic data to generate future climate projections.

The stochastic modelling uses historical (observed and reconstructed) records of daily rainfall, evapotranspiration and temperature to generate synthetic data for 10,000 years that reflect variability over the instrumental record from 1889 to 2018. The stochastic sequences provide insights into natural climate variability beyond the available observations. The scaling factors for a future climate are derived based on projections from the NSW and Australian Regional Climate Modelling (NARCLiM) project and applied to the stochastic.

The stochastic data generation methodology has been applied to multiple basins across New South Wales using a multisite stochastic data generator conditioned on the Interdecadal Pacific Oscillation (IPO) documented in Leonard and Westra (2020). A similar method is planned for data generation in the New South Wales southern basin region, shown in Figure 1. An independent expert panel review of the department's climate risk method recommended that ongoing improvement of the stochastic generation methodology be given high priority. The panel recommended further work to understand the implications of the influence of multiple climate drivers, and the existence of non-stationarity in the instrumental record of the southern New South Wales region on stochastic time series generation. The work presented in this report addresses these questions.



The southern basin region consists of water resource plan areas of the New South Wales Murray and Lower Darling, Murrumbidgee, and in parts, the Victorian Murray areas.

Figure 1 Southern basin of New South Wales

The regional climate of south-east Australia is highly variable, and there is a strong a priori reason to suspect that key meteorological variables are likely to already be experiencing change in the southern system. Literature notes warming signals in mean temperatures that are distinguishable from the background interannual and low-frequency variability in this region (Ukkola et al. 2019). These temperature changes are consistent with expected changes due to global warming (Karoly and Braganza 2005; Jones 2012). The regional rainfall patterns exhibit substantial variability (CSIRO 2012; Hope et al. 2017) and are potentially affected by both natural and anthropogenic influences on the climate system. These influences on the changing patterns of rainfall in this region have been the subject of much focused research, especially since the Millennium Drought.

Section 2 of the report briefly reviews literature on the influences of regional and large-scale climate drivers on precipitation¹ patterns in south-east Australia, and the historical and expected changes in these influences, to assist in interpreting non-stationarity results in subsequent parts of this report. Section 3 summarises literature on the historical trends in climatic variables in this region and Section 4 summarises the future climate projections.

Subsequently, a study using data from representative catchments in southern New South Wales and Victoria is undertaken to assess the observational record. The assessment involves an analysis of trends in the observed data from the pilot sites and split sample calibration and validation of stochastic replicates as recommended by the expert review panel. The results of this study are presented in Section 5 and are summarised in Section 6. Based on the review of literature and the results of the pilot study, options and recommendations are provided for stochastic data generation to characterise 'historical' and 'current' climates in south-east Australia in Section 7.

¹ Note that the report uses rainfall and precipitation synonymously. Where there is reference to literature of atmospheric processes, the cited literature may use precipitation more broadly than just referring to rainfall.

2 Large-scale patterns of variability and change relevant to south-east Australia

The precipitation in south-east Australia exhibits substantial interannual and intra-seasonal variability, influenced by large-scale patterns of global ocean–atmosphere variability. Different moisture systems contribute to precipitation in this region. These include low pressure systems that bring in moisture from the regional oceans (the Pacific, Indian and Southern Oceans), north-western cloud bands that originate in the Indian Ocean, eastern coastal troughs from northern Australia, and east coast low pressure systems (CSIRO 2012; Hope et al. 2017; Dowdy et al. 2019). The region exhibits a cool season precipitation regime, which is dominated by moisture contributions from southern westerlies. The circulation patterns in the southern oceans affect the location of subtropical ridges in the Australian mid-latitudes, which in turn influence the penetration of low-pressure systems from the south into the continent during the cool season.

Thus, the precipitation in this region is influenced by the oceanic and atmospheric patterns in the Pacific, Indian and Southern Oceans, which collectively ‘modulate’ weather patterns by imparting long-term (generally interannual) persistence and mediate some (particularly circulation-related) aspects of the anthropogenic climate change signal on local weather patterns. The indicators of these patterns, typically referred to as climate ‘drivers’, are the El Niño Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD), Southern Annular Mode (SAM), and the IPO. These drivers influence the regional rainfall in south-east Australia in different seasons and at different scales (CSIRO 2012; Hope et al. 2017).

The current understanding of these influences is summarised in the following subsections. The summary is based on technical reports by the South Eastern Australian Climate Initiative (SEACI) (CSIRO 2012), the Climate Change in Australia initiative (CSIRO and BOM 2015; Timbal et al. 2015) and the Victorian Climate Initiative (VicCI) (DELWP 2016; Hope et al. 2017), and the Victorian Climate Projections 2019 (VCP19) (Clarke et al. 2019a).

2.1 Influence of large-scale drivers: ENSO, IOD, SAM and IPO

ENSO and IOD are dipole modes of ocean–atmosphere variability in the tropical Pacific and Indian Oceans. Positive values of IOD refer to higher sea surface temperatures in the western equatorial Indian Ocean. El Niño is the ENSO mode associated with higher sea surface temperatures in the eastern equatorial Pacific Ocean. Positive IOD and El Niño are associated with higher sea surface temperatures and tropical convective centres that are located farther from the Australian continent, leading to lower rainfall in Australia. ENSO and IOD are understood to influence the rainfall in south-east Australia primarily during winter and spring (CSIRO 2012; Hope et al. 2017).

The SAM is a mode of mid- and high-latitude climate variability associated with north–south shift of the atmospheric mass between the polar region and the mid-latitudes. A positive value of SAM indicates a shift of mid-latitude storm tracks towards the South Pole. The SAM influences the rainfall over south-east Australia differently in different seasons. During winter, the shift of mid-

latitude storm tracks towards the South Pole that occurs with a positive SAM is associated with reduced rainfall over south-east Australia. During the summer, a positive value of SAM is associated with increased onshore transport of tropical moisture in eastern Australia and a subsequent increase in the warm season rainfall.

Thus, in general, positive SAM, positive IOD and El Niño events result in lower cool season rainfall in south-east Australia. The relationships are further complicated due to interactions between these modes. On decadal time scales, a pattern of Pacific climate variability – the IPO – affects the interannual variations associated with ENSO and IOD. When the IPO is in the negative phase (cold phase), variability of ENSO and IOD are weakened, and the coupling between them is also weakened. During this phase, the impacts of ENSO and IOD on eastern Australian rainfall is strengthened. The variability in SAM is also related to the ENSO, mainly during the warm season (Hope et al. 2017).

2.2 Influence of east coast lows on the south-east coast

In addition to the large-scale drivers outlined above, the precipitation in the south-east coast is influenced by low pressure systems known as east coast lows (ECLs) (Dowdy et al. 2019). ECLs are cyclonic systems that occur near the south-east coast due to both mid-latitude and tropical influences. These systems can occur during any time of the year, but are more common and intense during the cooler months (Dowdy et al. 2013). ECL-related rainfall events exhibit a spatial contrast since the rainfall from these events primarily occurs on the eastern coastal regions rather than inland areas because of the Great Dividing Range. These systems may also cause snowfall in the mountainous regions of south-east Australia (Fiddes et al. 2015). ECLs are associated with large rainfall events and multiple climate hazards on the south-east coast. These impacts have resulted in focused research in recent years to study the characteristics of ECLs, the influence of climate drivers on ECLs, and their expected changes into the future.

Studies have characterised ECLs in observations using multiple climatic features and report a large interannual variability in the number of ECL systems in the historical record (Di Luca et al. 2015; Pepler et al. 2015). The relation between ECLs and large-scale climate drivers (ENSO, IOD and SAM) is reported to be generally weak, but some studies report mixed findings that indicate that some types of ECLs may potentially be related to SAM (Dowdy et al. 2013; Pepler et al. 2015). A warming climate is expected to result in fewer ECL-related rainfall events during the cooler months, while there are large uncertainties in the expected ECL changes during the warmer months. The historical record also shows a decline in the number of ECLs, but the trend is not significant (Dowdy et al. 2013; Pepler et al. 2015).

Studies have examined the skill of regional climate models in simulating ECLs in historical and future scenarios. Regional simulations reproduce the climatology of ECLs (Di Luca et al. 2016), and future regional climate projections from NARClIM exhibit a decline in ECL frequency during the cooler months. The decreasing signal is reported to be robust in the different ensemble members and is consistent with the decreasing trend in the historical record and expected changes into the future (Pepler et al. 2016).

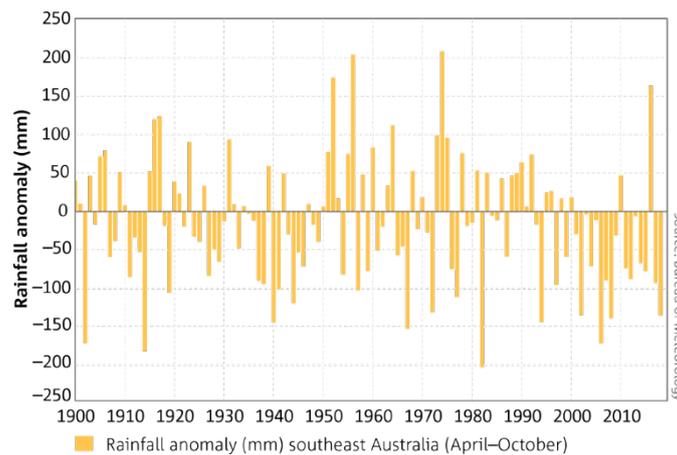
3 Climatic trends in the historical record of south-east Australia

In this section, publications documenting the historical trends in precipitation, evaporation and temperature in south-east Australia are reviewed. This region broadly covers the mainland Australian region south of 33°S and east of 135°E and encompasses most of the southern basin that is the focus of this report. Although there are some differences in the region referred to as 'south-east' in the literature, these latitude–longitude ranges are broadly consistent.

3.1 Rainfall

There is a well-documented decreasing trend in mean rainfall during the cool season, primarily during autumn, in south-east Australia. This trend is reported in studies since the mid-2000s; the reduction in cool season rainfall post mid-1990s influenced the long-term and medium-term mean precipitation trends in this region (Gallant et al. 2007). The reported trends in the attributes of precipitation are summarised in Table 1.

Further research explored the climatic features that led to the reduction in precipitation. The cool season decline was a notable feature of the Millennium Drought that this region experienced from 1997 to 2009; the signal has persisted post the drought. Figure 2 shows the cool season (April to October) rainfall anomalies in south-east Australia from 1900 to 2018 from the 2018 State of the Climate report (BOM and CSIRO 2019). The report documents that the southern half of Australia (south of 26°S) had below-average cool season rainfall in 17 of the 20 years from 1999. The recent years with above-average rainfall (2010 and 2016) were generally associated with drivers of higher-than-usual rainfall across Australia (strong negative IOD in 2016; La Niña in 2010) (BOM and CSIRO 2019).



Anomalies are calculated with respect to 1961 to 1990 averages.
Source: BOM and CSIRO 2019.

Figure 2 Anomalies of April–October rainfall for south-east Australia (south of 33°S, east of 135°E inclusive)

The current understanding indicates that changes in large-scale atmospheric features influence the observed trends. ENSO and IOD primarily influence the rainfall in south-east Australia during winter and spring and they do not have a major impact on the autumn rainfall. Hence, these climate drivers are not thought to be the primary cause of the declining autumn rainfall (Timbal and Hendon 2011). The general consensus from SEACI and VicCI (CSIRO 2012; CSIRO and BOM 2015; Hope et al. 2017) suggests that the decline in cool season rainfall is associated with the expansion of the tropics, and increasing intensity of the subtropical high located over the continent. Tropical expansion causes the mid-latitude storm tracks responsible for most of the cool season rainfall in south-east Australia to move further south. The tropical expansion and changes in the intensity of the subtropical high are also associated with positive trends in SAM, which indicate a poleward shift of the mid-latitude westerlies.

These climatic changes and the precipitation trends in the region are at least partly attributable to global warming (CSIRO 2012; CSIRO and BOM 2015; Hope et al. 2017). Experiments using climate models indicate that expanding tropics can only be reproduced when global atmospheric changes in greenhouse gases, aerosols and depletion of stratospheric ozone are incorporated into global model simulations (Nguyen et al. 2018). Literature has studied the regional variation in Hadley circulation by dividing the globe into three 'sectors' based on centres of upper level tropical divergence (the upward branch of the Hadley circulation) (Nguyen et al. 2018). There is enhanced expansion of the Hadley circulation in the Asia–Pacific (Australian) sector compared to the African and South American sectors. This regional expansion is linked to the negative phase of the IPO, and may reduce when the IPO changes phase (Nguyen et al. 2018). Therefore, both natural variability and climate warming are understood to contribute to the declining rainfall trends in south-east Australia (Hope et al. 2017).

While tropical expansion and trends in SAM appear to be the drivers of the cool season precipitation trend in this region, the influence of other drivers, such as the ENSO and IOD, are relevant with respect to the changes during warm season. A La Niña mode concurrent with a negative phase of the IOD is associated with the rainfall events during spring of 2010–11 that marked the end of the Millennium Drought.

Table 1 Summary of reported precipitation trends in the historical record of south-east Australia

Authors	Dataset	Period	Region	Index type	Findings
Gallant et al. (2007)	Stations (95 across Australia)	1910–2005 and 1950–2005	Six regions in Australia	<ul style="list-style-type: none"> Total rain Rain days (threshold = 1 mm) Mean rain per rain day Extreme intensity, frequency, and proportion of the total (95th and 99th percentile thresholds) 	<ul style="list-style-type: none"> For the south-east region, long-term (1910–2005) results show decreasing trends in autumn in total rainfall and extreme intensity. In the medium term (1950–2005), autumn rainfall indices show decreasing trends in most characteristics of rainfall, except the extreme proportion indices which showed significant increasing trends.
Alexander et al. (2007)	0.25° x 0.25° gridded data	1910–2005 and 1950–2005	Whole country	<ul style="list-style-type: none"> Annual and seasonal precipitation Extremes (Max 1-day, Max 5-day) Number of heavy and very heavy precipitation days Consecutive wet and dry days Annual proportions from extremes (Note: detailed results from these indices are not included in the publication) 	<ul style="list-style-type: none"> Decreasing trends in autumn precipitation in south-east Australia during 1950–2005. Trends in extremes are correlated with the trends in mean, in general, across the whole country.
Taschetto and England (2009)	Gridded BOM rainfall data at 0.5° resolution	1970–2005	Whole country	<ul style="list-style-type: none"> Annual and seasonal total rainfall Frequency of moderate (up to 1 SD from the mean), heavy (from 1 to 2 SD from mean), very heavy (more than 2 SD from mean) rainfall events 	<ul style="list-style-type: none"> Decreasing trends in rainfall over Victoria, southern South Australia and southern New South Wales during summer and autumn (stronger in autumn). Frequency of very heavy rainfall events during MAM also show a decline.

Authors	Dataset	Period	Region	Index type	Findings
Risbey et al. (2013)	Average rainfall from 8 stations in Mallee region in Victoria	1956–2009	Mallee region in Victoria	<ul style="list-style-type: none"> Total rainfall during the cool season (Apr to Oct) 	<ul style="list-style-type: none"> Decreasing trends in cool season rainfall. Reported to be primarily associated with a decline in rainfall from cut-off lows.
Theobald et al. (2016)	Station data	1958–2012	Snowy Mountains	<ul style="list-style-type: none"> Annual, cool season (Apr to Oct) and warm season (Nov to Mar) rainfall total and frequency of wet days ($P > 1$ mm) and heavy precipitation days ($P > 10$ mm) Frequency and intensity of extreme precipitation ($P > 90$th percentile threshold) 	<ul style="list-style-type: none"> Annual decreasing trends in the frequency of $P > 10$ mm events, but an increase in the total precipitation the events generate. The increase in precipitation from these events occurs during the warm season. Annual intensity of extreme precipitation shows an increasing trend.
Ukkola et al. (2019)	Area average records from BOM for 6 regions across Australia	1900–2018	Whole country	<ul style="list-style-type: none"> Annual and seasonal totals 	<ul style="list-style-type: none"> No significant trends in regional mean rainfall in south-east Australia.

3.2 Pan evaporation

The non-stationarity in historical evapotranspiration records is not well established. The available literature has used pan evaporation records to study trends post-1975. Earlier studies reported a decreasing trend in pan evaporation over the period 1975–2002 across the country, potentially caused by reduced atmospheric demand associated with decreasing wind speed or radiation (Roderick and Farquhar 2004). This trend is reduced or reversed (that is, to increasing trends or no trends) in more recent analyses, respectively using data up to 2016 (Stephens et al. 2018) and 2018 (Ukkola et al. 2019), due to temperature driven changes in vapour pressure deficit. These reported trends are summarised in Table 3.

Table 2 Summary of reported evapotranspiration trends in the historical record of south-east Australia

Authors	Dataset	Period	Region	Index type	Findings
Roderick and Farquhar (2004)	30 BOM sites across the country	1970–2002	Entire country	Pan evaporation	Decreasing trends over Australia primarily due to a decline in atmospheric demand, associated with declines in surface radiation or wind speed
Stephens et al. (2018)	41 BOM sites across the country	1975–2016	Entire country	Pan evaporation	The declining trends detected earlier have reduced or become neutral in south-east Australia. The changes in trends from previous reports are due to increases in temperature driven vapour pressure deficits
Ukkola et al. (2019)	Area average monthly observations from BOM	1975–2018	Whole country	Pan evaporation	No significant trends in south-east Australia

3.3 Temperature

Literature reports a warming signal in temperature records in south-east Australia, consistent with the expected changes due to global warming. Karoly and Braganza (2005) studied the trends in minimum (Tmin), maximum (Tmax) and mean daily temperatures in south-east Australia over a 50-year period from 1954 to 2003, after removing the rainfall-related component of temperature variations. The study reports a clear anthropogenic warming signal in observed temperature trends in south-east Australia. Jones (2012) noted discrete step changes in minimum and maximum temperatures in south-east Australia and attributed these changes to episodes of anthropogenic regional warming. The study reports that the temperature data of south-east Australia was stationary from 1910 to 1967, followed by a series of step changes, 0.7 °C in Tmin in 1968, 0.5 °C in Tmax in 1973 and 0.8 °C in Tmax in 1997, attributed to anthropogenic warming. The reported trends in temperature records in south-east Australia are summarised in Table 2, which shows increasing trends in mean, minimum and maximum temperatures over the region.

Thus, there is a well-established warming signal in the temperature records in south-east Australia, especially post-1960. The Climate Change in Australia initiative reports an increase of 0.8 °C in the Murray basin cluster (Timbal et al. 2015), which consists of most of the southern basin, over the 1910–2013 time period assuming a linear trend. The trends are higher for temperature minimums (total increase of 1 °C) than for maximums (total increase of 0.7 °C) The mean temperature over Australia has increased by just over 1 °C during the period 1910–2018 and the Victorian mean historical changes over 1910–2018 reported by VCP19 (Clarke et al. 2019a) is also just above 1 °C.

Table 3 Summary of reported temperature trends in the historical record of south-east Australia

Authors	Dataset	Period	Region	Index type	Findings
Ashcroft et al. (2012)	Daily data from 103 stations across the country; monthly area-averaged anomalies for states and Northern Territory	1860–1909, 1910–1959, 1960–2011 and 1860–2011	South-east Australia	<ul style="list-style-type: none"> Annual, DJF and JJA means of Tmax, Tmin Diurnal temperature range (DTR) 	Positive trends in annual, DJF and JJA Tmax and Tmin. Stronger trends in Tmax (1.12 °C) and Tmin (0.93 °C) post-1960
Jones (2012)	Homogenised temperature data on a 0.25° grid	1910–2010	South-east Australia	<ul style="list-style-type: none"> Annual means of Tmin and Tmax 	Positive step changes in Tmin and Tmax post-1967
Ukkola et al. (2019)	Area average records from BOM for 6 regions across Australia	1910–2018	Whole country	<ul style="list-style-type: none"> Annual and seasonal mean temperature 	The mean air temperature trends during all seasons are positive and statistically significant in south-east Australia. The strongest trends are in summer (DJF, 0.015 °C/yr) and autumn (MAM, 0.012 °C/yr)

4 Future climate projections for south-east Australia

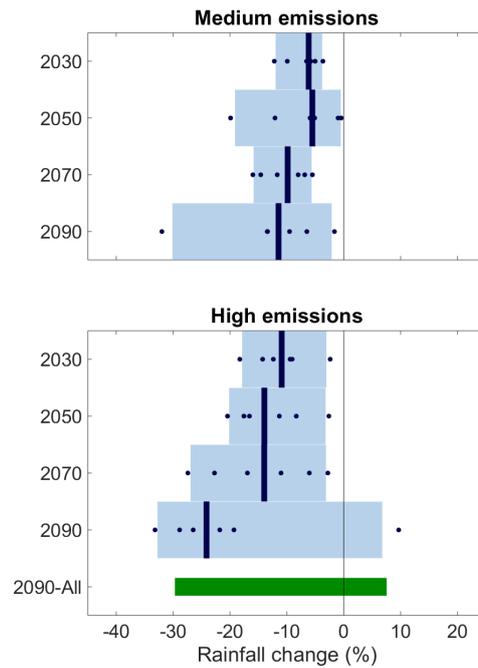
A summary of the future climate projections for this region is included as an alternative line of evidence to understand climatic non-stationarity. Climate projections for south-east Australia indicate future precipitation and temperature changes consistent with the trends in the current observational record; that is, warmer and drier in general, especially in the cool season. From climate model projections, positive values of the SAM and an increase in the number of positive IOD events are likely in the future, bringing drier conditions to south-east Australia during the cool season. The changes in ENSO and its interactions with the SAM and IOD into the future are currently unknown (CSIRO 2012; Clarke et al. 2019a).

Future projections for the Murray basin cluster are documented as part of the Climate Change in Australia initiative (Timbal et al. 2015). This cluster includes most of the southern basin region that is the focus of this report. Climate projections indicate that increases of 0.6–1.3 °C are expected in the Murray basin cluster in the near term (2020–2039) with respect to a baseline of 1986–2005. The report considers the physical understanding of climatic relationships and results from downscaling future projections to conclude that there is high confidence that cool season rainfall will decline in future, but the magnitude of decline is very uncertain (Timbal et al. 2015). It is reported with high confidence that natural climate variability will remain the major driver of rainfall differences in the near term (2030) from the climate of 1986–2005.

Climate projections from VCP19 include regional projections for the Ovens and Murray catchments (Clarke et al. 2019b) that are also used in the pilot study documented in Section 5 of this report. The key changes projected for these catchments are as follows.

- Daily minimum and maximum temperatures are projected to continue to increase. An increase in Tmax of 1.0–1.9 °C (since 1990) is expected by the 2030s.
- Rainfall is projected to be very variable but will continue to decline in winter and spring (medium to high confidence) and autumn (low confidence).
- Intensification is expected of 1-in-20-year maximum daily extreme rainfall events.

The VCP19 report projects changes under medium (RCP 4.5) and high (RCP 8.5) emission scenarios with respect to a baseline climate of 1986–2005. The changes are projected for 20-year periods up to 2090. The first period spans 2020 to 2039. Thus, 'current' climate at the time of writing this report (2023) sits within the period of these projections. The projected rainfall changes in the Ovens and Murray catchments reported by VCP19 are shown in Figure 3.



Bars show the 10th to 90th percentile range. Blue bars: results from the new downscaled modelling. Dark vertical line: median. Dark blue dots: individual models. Green bar (at bottom): results from all available modelling (high resolution and GCM) for comparison, at high emissions scenario, at 2090. Source: Clarke et al. 2019b.

Figure 3 Projected changes (compared to 1986–2005) in annual mean rainfall in the Ovens and Murray catchments for medium emissions (top) and high emissions (bottom)

The annual mean rainfall during 2020–2039 is projected to decrease with the largest declines during spring. The projected median change in annual rainfall is –6% (range of –12% to –4%) under the medium emissions scenario, and –11% (range of –18% to –3%) under the high emissions scenario, and further declines are anticipated in subsequent decades (Figure 3). The annual maximum temperatures are expected to increase by 1.1 °C under the medium emissions scenario and 1.4 °C under the high emissions scenario during 2020–2039; the projected changes in minimum temperatures are smaller in magnitude for the Ovens and Murray catchments. The projected annual changes in pan evaporation during 2020–2039 are positive, with a median change of 8% under the medium emissions scenario and 10.8% under the high emissions scenario.

5 Pilot study: assessment of the historical record

The review of literature indicates the presence of non-stationarity in the historical record of south-east Australia, with some references attributing key aspects of change to anthropogenic climate change. A pilot study using data from a few representative catchments is undertaken to assess the presence of significant trends in the attributes of climatic variables and the implications for stochastic time series generation in the southern New South Wales region. As part of the assessment, a trend analysis is performed of the historical record of the pilot catchments and split sample testing of the stochastic model, according to the recommendations of the independent review panel.

The focus of this pilot study was to analyse key time series provided by department that are used as inputs for hydrological modelling. As a result, interpretation of the pilot results needs to consider the following caveats.

- The homogeneity of the data has not been reviewed in this analysis. For this reason, the results presented below can be used to assess non-stationarity of the analysed data but cannot be interpreted as a climate change attribution study.
- The analysis uses data from a relatively small number of stations relative to other peer-reviewed assessments of trends in key weather variables in south-east Australia, and thus is not suitable to assess large-scale drivers of change (for example, the effects of changes to circulation patterns) that present themselves when evaluating trends over larger geographical areas.

For these reasons, results from this section should be considered in the context of the wider literature, as summarised in Sections 3 and 4, rather than viewed in isolation.

5.1 Data

Data from pilot sites located in the Upper Murray, Ovens and Snowy catchments, is shown in Figure 4.

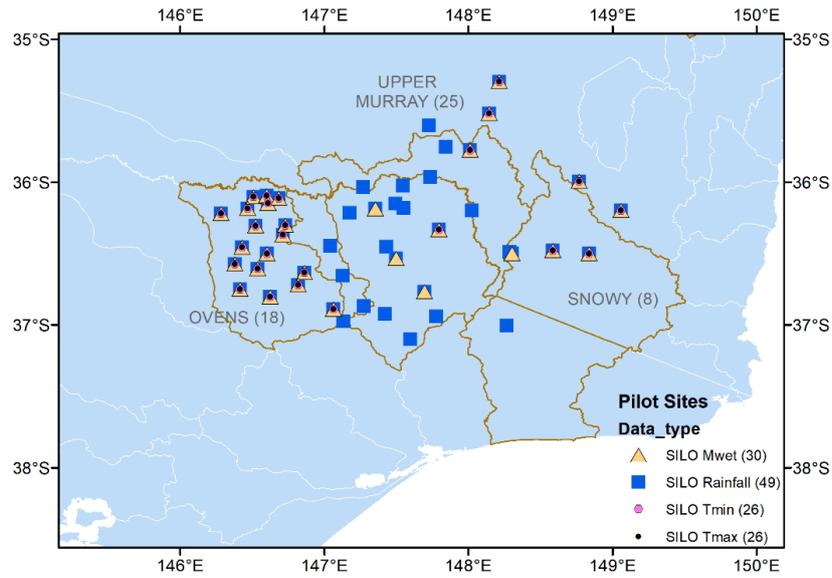


Figure 4 Location of rainfall, evapotranspiration and temperature pilot sites

Table 4 provides a summary of the data from the pilot sites which consists of precipitation, evapotranspiration (Mwet – the tag refers to Morton wet formulation), and minimum/maximum daily temperature (Tmin, Tmax) time series. All data were sourced from the SILO database. The time series span 130 years from 1 January 1889 to 31 December 2018 and there were no missing values (owing to predetermined infilling methods used to construct the data).

Table 4 Summary of number of different observation time series by variable type

Variable type	Pilot basin – Upper Murray	Pilot basin – Ovens	Pilot basin – Snowy
SILO Rain	25	18	6
SILO Mwet	5	18	7
SILO Tmin/Tmax	0	18	8

As described above, the analysis of trends using raw station observations of Tmax and Tmin is not advisable because signals of changes in instrument, methods of data collection, and station location may exist in the data (Trewin 2013). Such analyses are generally performed using homogenised temperature datasets. In the process of homogenisation, raw temperature data from multiple sites are examined visually and statistically to create homogenised datasets that minimise discrepancies across time (Ashcroft et al. 2012; Trewin 2013; Trewin et al. 2020).

The Australian Climate Observations Reference Network – Surface Air Temperature dataset version 2 (ACORN–SAT v2) (Trewin et al. 2020) is one such high-quality dataset, prepared and made available by the Bureau of Meteorology, Australia. This data has been used in literature to study the changes in temperature in Australia (van Wijngaarden and Mouraviev 2016; Allen et al. 2019), and thus analyses of data provided by the Department of Planning and Environment for the pilot sites are supplemented by an analysis of ACORN–SAT v2 data at key locations in the study area.

The temperature stations from the pilot sites are located in the Ovens and Snowy catchments. Three ACORN–SAT v2 stations are located in the vicinity of these sites, as shown in Figure 5.

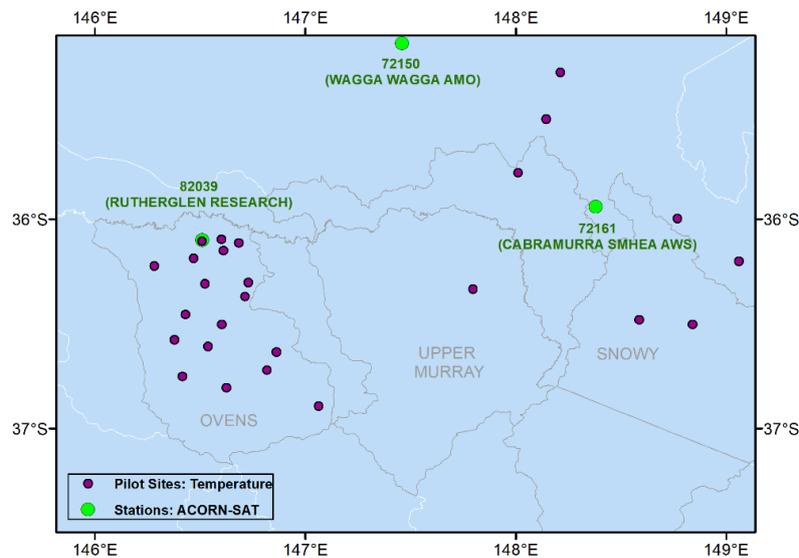
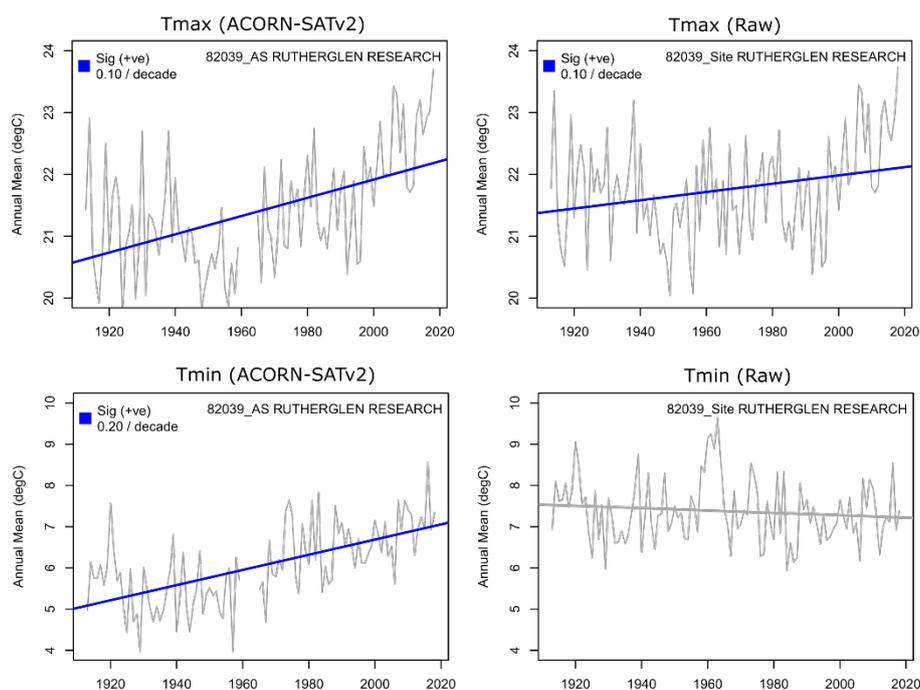


Figure 5 Location of pilot temperature sites and 3 ACORN–SAT v2 stations in the vicinity

Homogenised data at pilot site 82039 (Rutherglen Research) is available from the ACORN–SAT v2 dataset, so that the differences between the data using the 2 methods can be compared. Figure 6 shows the differences in long-term trends calculated from the raw station data and the homogenised ACORN–SAT v2 data at the same site using 101 years' data between 1913 and 2018. There are major differences in the trends estimated from the 2 data sources. Such differences also exist in the short-term trends estimated from the 2 data sources at this site (Appendix A).

These differences indicate that the site-level temperature observations from the pilot sites are unsuitable for the analysis of trends. While there is some debate on the homogenisation of temperature data (Marohasy and Abbot 2016), the use of homogenised data for assessment of trends is the existing globally accepted standard of analysis (Hewaratchi et al. 2017; Squintu et al. 2019; Vincent et al. 2020). Therefore, a trend analysis is performed using the homogenised data from the 3 ACORN–SAT v2 stations located close to the pilot region (shown in Figure 5) as part of this study. The length of homogenised data records at each of these sites is listed in Table 5.



Blue trendlines indicate the presence of significant trends (at 5% level) using the Mann Kendall trend test. The data from 1960 to 1964 are missing at this site in the ACORN–SAT v2 data.

Figure 6 Annual mean Tmax and Tmin at site 82039 (Rutherglen Research) from raw station data and homogenised ACORN–SAT v2 data with linear trendlines

Table 5 ACORN–SAT v2 stations used for analysis

Variable type	Station number	Period of record
Tmax, Tmin	82039	8 November 1912 to 31 May 2019
Tmax, Tmin	72150	1 January 1910 to 31 May 2019
Tmax, Tmin	72161	1 January 1962 to 31 May 2019

5.2 Methodology

5.2.1 Trend detection

The statistical significance of temporal trends in various attributes of the time series are assessed using a non-parametric Mann Kendall test at two-sided 5% significance level. The Mann Kendall test is a well-established technique employed in studies for assessment of hydroclimatic time series (Lavender and Abbs 2013; Theobald et al. 2016; Ukkola et al. 2019). The magnitude of significant trends is quantified using least squares regression, similar to the analysis performed by Theobald et al. (2016).

While this approach can assess whether individual sites exhibit statistically significant trends, in multisite analyses there is often a non-negligible probability of detecting one or more individual sites with significant trends even under the null hypothesis of no trends (for example at the 5% significance level one would expect an average of 5 out of every 100 sites to experience statistically significant trends under the null hypothesis that there is no trend). As such, a field

significance test is used to determine whether the number of stations experiencing statistically significant trends is more than would be expected under the null hypothesis. The field significance of the trends are assessed in this study using a bootstrap resampling procedure (see, for example, Do et al. (2017)). The bootstrap procedure uses resampled data to obtain an estimate of the 95th percentile value of the percentage of significant sites that may occur due to chance. If the percentage of sites exhibiting significant trends in the historical record exceeds this estimate, then the trends are considered field significant.

The methodology used for the analysis of trends consists of the following steps:

1. The significance of site-level trends is estimated using the Mann Kendall test. The proportion of sites that exhibit significant positive and negative trends in the historical record are calculated.
2. The entire dataset is randomly resampled in time while preserving the spatial structure. The new resampled data therefore contains a new sequence of years (for example, {1967, 1954 2003, 1895, 1920...}). The site-level significant trends in the resampled dataset are estimated using the Mann Kendall test and the proportion of sites that exhibit positive and negative trends are calculated as done in step 1.
3. A bootstrapping procedure is used to repeat step 2 1,000 times. The samples are used to create a distribution of percentage of significant sites that may occur in the region due to chance. If the proportion of significant sites in the historical data (step 1) is higher than the 95th percentile value of the proportion of significant sites that may occur due to chance, the historical trend is field significant.

The attributes of rainfall, evaporation and temperature used for the trend analyses are listed in Table 6. These attributes are selected to comprise 'hydrologically relevant' features of the respective variables, which may have a bearing on hydrological response of the respective catchments. The trend analysis is performed using the entire dataset of 130 years (1889 to 2018) as well as a recent subset of the dataset, to enable comparison with existing literature.

Table 6 Attributes of hydroclimatic variables and the periods used for trend analyses

Variable	Attribute	Definition	Analysis period
Rainfall	Total	Total annual and seasonal (DJF, MAM, JJA and SON) rainfall (mm)	1889 to 2018 and 1950 to 2018
Rainfall	Wet day rainfall	Mean annual and seasonal wet day ($P \geq 1$ mm) rainfall (mm/day)	1889 to 2018 and 1950 to 2018
Rainfall	Number of wet days	Annual and seasonal number of wet days ($P \geq 1$ mm) (days)	1889 to 2018 and 1950 to 2018
Rainfall	Heavy day rainfall	Annual and seasonal heavy day ($P \geq 10$ mm) rainfall (mm/day)	1889 to 2018 and 1950 to 2018
Rainfall	Number of heavy rainfall days	Mean annual and seasonal heavy day ($P \geq 10$ mm) rainfall (mm/day)	1889 to 2018 and 1950 to 2018
Rainfall	Mean dry spell duration	Annual mean number of consecutive days with rainfall less than 1 mm (days)	1889 to 2018 and 1950 to 2018
Rainfall	Maximum dry spell duration	Annual maximum number of consecutive days with rainfall less than 1 mm (days)	1889 to 2018 and 1950 to 2018
Rainfall	Extreme intensity	Annual mean rainfall during days with rainfall greater than the 95th percentile (mm/day)	1889 to 2018 and 1950 to 2018
Rainfall	Extreme frequency	Annual mean number of days with rainfall greater than the 95th percentile (days)	1889 to 2018 and 1950 to 2018
Evapotranspiration	Total evapotranspiration	Total annual and seasonal (DJF, MAM, JJA and SON) evapotranspiration (mm)	1889 to 2018 and 1975 to 2018
Temperature	Mean temperature	Annual and seasonal (DJF, MAM, JJA and SON) mean daily minimum temperature Annual and seasonal (DJF, MAM, JJA and SON) mean daily maximum temperature	1913 to 2018 and 1960 to 2018

5.2.2 Split sample stochastic simulations

The potential implications of non-stationarity on the results of stochastic analyses using split sample tests are assessed. This is achieved using a split sample methodology, in which the stochastic generation model is calibrated against one part of the time series (usually the earlier part of the record) and then validated against the other part of the record. This split sample approach provides an analogy to possible issues that could arise by calibrating a stochastic model against the full historical record and assuming it is representative of current or future conditions.

The split sample stochastic simulations are performed using the precipitation and evapotranspiration (M_{wet}) data from the pilot sites to assess the ability of a stochastic model calibrated against the earlier part of the record to capture statistics corresponding to the later part of the record. The split sample tests are designed based on the recommendations of the

independent review panel and consist of '1990 reference' and 'drought reference' experiments. These are defined as follows:

- 1990 reference – The experiment uses data up to 1990 to calibrate the stochastic model and data after 1990 to validate stochastic simulations.
- drought reference – The experiment includes data up to the end of the Millennium Drought (year 2009) to calibrate the model, and the remaining period data to validate the simulations.

The calibration-validation tests are performed using the full record (1889 to 2018, 130 years) as well as a shorter period of data (1950 to 2018, 69 years) to assess the performance of the stochastic model while calibrated using different record lengths. Note that the year 1950 used here is selected arbitrarily to consider the potential of a using a shorter baseline (for example, corresponding to the NARClIM 1.5 baseline starting in 1950). In total the experiment suite consists of the 4 experiments shown in Figure 7.

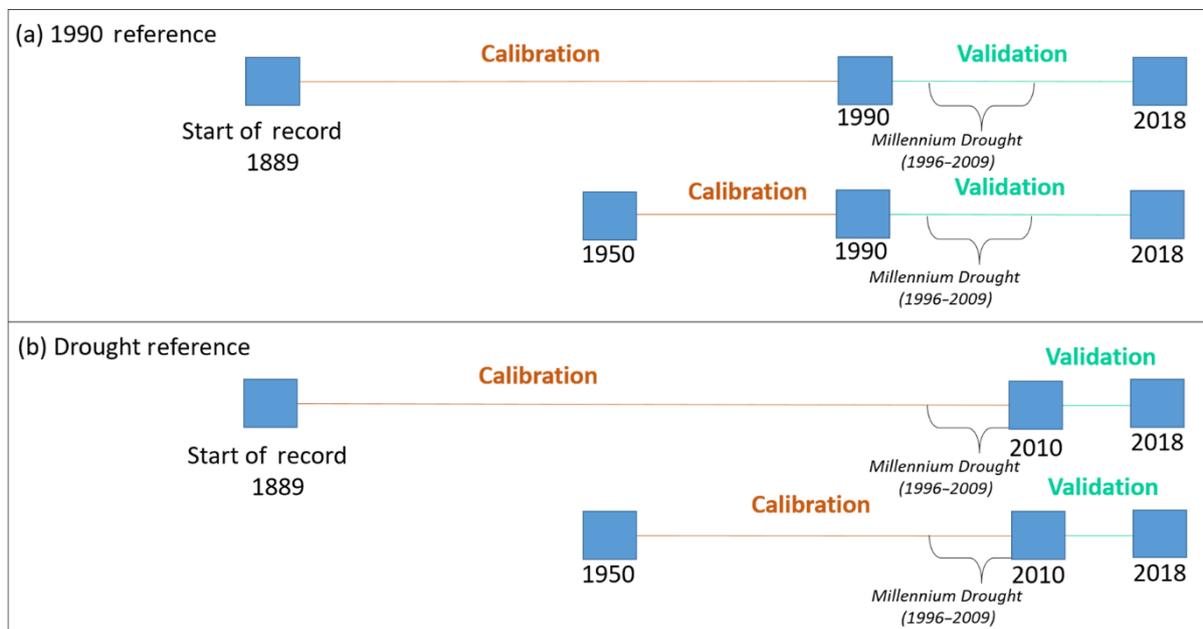


Figure 7 Split sample experiments

The stochastic model used for the experiments is based on the latent variable model formulation documented in Bennett et al. (2018). The model implemented for this pilot study is not conditioned on the IPO and it is a single-site version of the spatial field rainfall model used for stochastic time series generation in the northern New South Wales basins (Leonard and Westra 2020). The model is calibrated site-wise using observations from the calibration period and used to generate stochastic time series. Time series are generated for 100 replicates of data length corresponding to the validation period.

The simulated data are compared to observations during both the calibration and validation periods for assessment of the split sample experiments. The comparison is based on the attributes of hydroclimatic time series that show major trends in the historical record identified using non-parametric trend testing at the pilot sites. The mean values of the attributes from the simulated data are compared to the mean values of observations from the calibration and validation periods at site level. The results are presented using histograms of the differences between the observations and simulations across all the pilot sites.

5.3 Results of non-parametric trend testing

5.3.1 Trends in rainfall attributes

Table 7 summarises the number of sites that exhibit significant trends in rainfall attributes during the different periods of analysis. The trends in the attributes of rainfall that are significant in both the long-term (1889 to 2018) and short-term (1950 to 2018) analyses are:

- a decreasing trend in cool season totals
- a decreasing trend in the number of wet days annually and during the cooler seasons of the year
- an increasing trend in annual extreme rainfall intensity.

Table 7 Number of sites that exhibit significant trends in annual and seasonal attributes of rainfall

Attribute	Trend	1889 to	1950 to								
		2018 ANN	2018 DJF	2018 MAM	2018 JJA	2018 SON	2018 ANN	2018 DJF	2018 MAM	2018 JJA	2018 SON
Total rainfall (mm)	Pos	5 (10%)	11 (22%) *	2 (4%)	2 (4%)	3 (6%)	1 (2%)	1 (2%)	0 (0%)	1 (2%)	0 (0%)
Total rainfall (mm)	Neg	8 (16%)	0 (0%)	6 (12%)	15 (31%) *	1 (2%)	13 (27%) *	0 (0%)	13 (27%) *	8 (16%) *	24 (49%) *
Total rainfall (mm)	None	36 (73%)	38 (78%)	41 (84%)	32 (65%)	45 (92%)	35 (71%)	48 (98%)	36 (73%)	40 (82%)	25 (51%)
Mean wet day (P >= 1 mm) rainfall (mm/day)	Pos	21 (43%) *	19 (39%) *	5 (10%)	8 (16%)	22 (45%) *	5 (10%)	6 (12%)	1 (2%)	2 (4%)	3 (6%)
Mean wet day (P >= 1 mm) rainfall (mm/day)	Neg	7 (14%)	0 (0%)	7 (14%) *	10 (20%) *	2 (4%)	7 (14%)	0 (0%)	4 (8%)	8 (16%) *	4 (8%)
Mean wet day (P >= 1 mm) rainfall (mm/day)	None	21 (43%)	30 (61%)	37 (76%)	31 (63%)	25 (51%)	37 (76%)	43 (88%)	44 (90%)	39 (80%)	42 (86%)
Number of wet (P >= 1 mm) days	Pos	7 (14%)	10 (20%) *	1 (2%)	3 (6%)	3 (6%)	0 (0%)	6 (12%)	0 (0%)	1 (2%)	0 (0%)
Number of wet (P >= 1 mm) days	Neg	19 (39%) *	3 (6%)	16 (33%) *	19 (39%) *	11 (22%) *	24 (49%) *	0 (0%)	26 (53%) *	10 (20%) *	29 (59%) *
Number of wet (P >= 1 mm) days	None	23 (47%)	36 (73%)	32 (65%)	27 (55%)	35 (71%)	25 (51%)	43 (88%)	23 (47%)	38 (78%)	20 (41%)
Mean heavy day (P >= 10 mm) rainfall (mm/day)	Pos	17 (35%) *	9 (18%) *	2 (4%)	7 (14%) *	27 (55%) *	7 (14%) *	4 (8%)	2 (4%)	2 (4%)	5 (10%)
Mean heavy day (P >= 10 mm) rainfall (mm/day)	Neg	2 (4%)	0 (0%)	0 (0%)	3 (6%)	0 (0%)	1 (2%)	0 (0%)	1 (2%)	1 (2%)	1 (2%)
Mean heavy day (P >= 10 mm) rainfall (mm/day)	None	30 (61%)	40 (82%)	47 (96%)	39 (80%)	22 (45%)	41 (84%)	45 (92%)	46 (94%)	46 (94%)	43 (88%)

Attribute	Trend	1889 to 2018 ANN	1889 to 2018 DJF	1889 to 2018 MAM	1889 to 2018 JJA	1889 to 2018 SON	1950 to 2018 ANN	1950 to 2018 DJF	1950 to 2018 MAM	1950 to 2018 JJA	1950 to 2018 SON
Number of heavy rainfall (P >= 10 mm) days	Pos	4 (8%)	15 (31%) *	1 (2%)	2 (4%)	6 (12%)	0 (0%)	1 (2%)	0 (0%)	1 (2%)	0 (0%)
Number of heavy rainfall (P >= 10 mm) days	Neg	6 (12%)	0 (0%)	5 (10%)	11 (22%) *	1 (2%)	15 (31%) *	0 (0%)	7 (14%) *	6 (12%) *	20 (41%) *
Number of heavy rainfall (P >= 10 mm) days	None	39 (80%)	34 (69%)	43 (88%)	36 (73%)	42 (86%)	34 (69%)	48 (98%)	42 (86%)	42 (86%)	29 (59%)
Mean extreme (P > 95th percentile) rainfall (mm/day)	Pos	15 (31%) *	n/a	n/a	n/a	n/a	9 (18%) *	n/a	n/a	n/a	n/a
Mean extreme (P > 95th percentile) rainfall (mm/day)	Neg	1 (2%)	n/a	n/a	n/a	n/a	1 (2%)	n/a	n/a	n/a	n/a
Mean extreme (P > 95th percentile) rainfall (mm/day)	None	33 (67%)	n/a	n/a	n/a	n/a	39 (80%)	n/a	n/a	n/a	n/a
Frequency of extreme (P > 95th percentile) rainfall days	Pos	13 (27%) *	n/a	n/a	n/a	n/a	3 (6%)	n/a	n/a	n/a	n/a
Frequency of extreme (P > 95th percentile) rainfall days	Neg	3 (6%)	n/a	n/a	n/a	n/a	2 (4%)	n/a	n/a	n/a	n/a

Attribute	Trend	1889 to 2018 ANN	1889 to 2018 DJF	1889 to 2018 MAM	1889 to 2018 JJA	1889 to 2018 SON	1950 to 2018 ANN	1950 to 2018 DJF	1950 to 2018 MAM	1950 to 2018 JJA	1950 to 2018 SON
Frequency of extreme (P > 95th percentile) rainfall days	None	33 (67%)	n/a	n/a	n/a	n/a	44 (90%)	n/a	n/a	n/a	n/a
Maximum dry (P < 1 mm) spell duration (days)	Pos	0 (0%)	n/a	n/a	n/a	n/a	13 (27%) *	n/a	n/a	n/a	n/a
Maximum dry (P < 1 mm) spell duration (days)	Neg	31 (63%) *	n/a	n/a	n/a	n/a	1 (2%)	n/a	n/a	n/a	n/a
Maximum dry (P < 1 mm) spell duration (days)	None	18 (37%)	n/a	n/a	n/a	n/a	35 (71%)	n/a	n/a	n/a	n/a
Average dry (P < 1 mm) spell duration (days)	Pos	6 (12%)	n/a	n/a	n/a	n/a	0 (0%)	n/a	n/a	n/a	n/a
Average dry (P < 1 mm) spell duration (days)	Neg	13 (27%) *	n/a	n/a	n/a	n/a	2 (4%)	n/a	n/a	n/a	n/a
Average dry (P < 1 mm) spell duration (days)	None	30 (61%)	n/a	n/a	n/a	n/a	47 (96%)	n/a	n/a	n/a	n/a

Pos = significant positive trend; Neg = significant negative trend; None = no significant trends; n/a = not applicable.

*Asterisked, shaded cells indicate trends that are field significant (at 5% level).

The values in brackets indicate the number of sites as a percentage of the total number of sites (49 sites).

Based on these results, the significant trends in the various rainfall attributes are summarised as follows.

- **Annual and seasonal totals:** The long-term (1889 to 2018) analysis shows negative trends in the JJA total rainfall and an increase in the DJF total. In the short-term (1950 to 2018) analysis, the decreasing trends are significant at a larger number of stations, and the declines occur in MAM, JJA and SON seasons. The highest number of stations show statistically significant declines in SON (49%).
- **Mean wet day rainfall:** In the long-term analysis, positive trends in the mean wet day rainfall are field significant at the annual scale. The positive trends occur in DJF and SON; negative trends occur in the cool season (MAM and JJA). In the short-term analysis, only the negative trends in JJA are significant.
- **Number of wet days:** In both the short-term and long-term analyses, negative trends are significant at the annual scale. In the long term, positive trends occur in DJF; negative trends during the MAM, JJA and SON seasons. In the short-term analysis, only the negative trends during MAM, JJA and SON are significant. The negative trends during MAM and SON occur at more stations.
- **Mean heavy day rainfall days and the number of heavy rainfall days:** The trends in heavy rainfall intensity and frequency vary between the analyses performed at long and short timescales. In the long-term analysis, the prominent signal is an increase in mean heavy day rainfall during all seasons except MAM. In the short-term analysis, the prominent signal is a decrease in the number of heavy rainfall events in all seasons except DJF.
- **Intensity and frequency of extreme rainfall days:** Both long-term and short-term analyses show an increase in the intensity of extreme rainfall days. The increase in frequency of extreme days is only significant in the long-term analysis.
- **Dry spell durations:** Long-term analysis shows declining trends in the mean dry spell duration and maximum dry spell duration. The short-term analysis shows an increasing trend in maximum dry spell duration.

No comparison is provided here of the trends at the pilot sites with literature that examined trends in precipitation characteristics using gridded datasets primarily for the purpose of assessing the fidelity of global climate models in capturing these trends (for example, Alexander and Arblaster 2009, 2017). The spatial scale of these studies is very different from the station-level analyses of trends at the pilot sites; hence comparison proves difficult, and potentially misleading. However, in general, the declining trends in cool season rainfall appear to be sufficiently widespread in spatial scale to be apparent in literature documenting analysis at both larger and finer spatial scales (Nicholls 2010; Theobald et al. 2016).

Table 8 compares the trends in rainfall attributes to the trends reported in literature for the same region (summarised in Section 3). The decreasing trend in cool season precipitation in the pilot sites is consistent with this reported decline in this region. The decreasing trends in the frequency of wet days during the cool season are also broadly consistent with the trends documented in the literature. The increasing trend in the annual intensity of extreme precipitation in the pilot sites does not appear to be as widespread spatially. The increasing trend in extreme intensity is consistent with literature focused on near catchments (Theobald et al. 2016), but inconsistent with literature documenting larger spatial scale analyses (Gallant et al. 2007).

Table 8 Trends in rainfall attributes in the pilot sites and the changes reported in literature

Attribute	Study or publication	Period	Annual	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
Total rainfall	Pilot study long	1889–2018	N	I	I	I	-	-	-	WD	WD	WD	N	N	N
Total rainfall	Gallant et al. (2007) ^a	1910–2005	N	N	N	N	WD	WD	WD	N	N	N	N	N	N
Total rainfall	Pilot study short	1950–2018	WD	N	N	N	WD	WD	WD	D	D	D	WD	WD	WD
Total rainfall	Gallant et al. (2007) ^a	1950–2005	N	N	N	N	WD	WD	WD	N	N	N	N	N	N
Total rainfall	Risbey et al. (2013)	1956–2009	n/a	n/a	n/a	n/a	n/a	WD	WD	WD	WD	WD	WD	WD	n/a
Total rainfall	Taschetto and England (2009) ^b	1970–2005	WD	N	N	N	WD	WD	WD	N	N	N	N	N	N
Mean wet day rainfall	Pilot study long	1889–2018	WI	WI	WI	WI	D	D	D	I or D	I or D	I or D	WI	WI	WI
Mean wet day rainfall	Gallant et al. (2007) ^a	1910–2005	N	N	N	N	N	N	N	N	N	N	N	N	N
Mean wet day rainfall	Pilot study short	1950–2010	N	N	N	N	N	N	N	D	D	D	N	N	N
Mean wet day rainfall	Gallant et al. (2007) ^a	1950–2005	N	N	N	N	WD	WD	WD	N	N	N	N	N	N
Number wet days	Pilot study long	1889–2018	WD	I	I	I	WD	WD	WD	WD	WD	WD	I	I	I
Number wet days	Gallant et al. (2007) ^a	1910–2005	N	N	N	N	N	N	N	N	N	N	N	N	N
Number wet days	Pilot study short	1950–2018	WD	N	N	N	WD	WD	WD	D	D	D	WD	WD	WD
Number wet days	Gallant et al. (2007) ^a	1950–2005	N	N	N	N	WD	WD	WD	N	N	N	N	N	N
Number wet days	Taschetto and England (2009) ^{b,c}	1970–2005	WD	N	N	N	WD	WD	WD	N	N	N	N	N	N
Number wet days	Theobald et al. (2016) ^d	1958–2012	N	N	N	N	N	WD	WD	WD	WD	WD	WD	WD	N
Heavy day rainfall	Pilot study long	1889–2018	WI	I	I	I	N	N	N	I	I	I	WI	WI	WI
Heavy day rainfall	Pilot study short	1950–2018	I	N	N	N	N	N	N	N	N	N	N	N	N
Number of heavy rainfall days	Pilot study long	1889–2018	N	WI	WI	WI	N	N	N	D	D	D	N	N	N
Number of heavy rainfall days	Pilot study short	1950–2018	WD	N	N	N	D	D	D	D	D	D	WD	WD	WD
Number of heavy rainfall days	Taschetto and England (2009) ^{b,e}	1970–2005	N	N	N	N	WD	WD	WD	N	N	N	N	N	N
Number of heavy rainfall days	Theobald et al. (2016)	1958–2012	N	N	N	N	N	WD	WD	WD	WD	WD	WD	WD	N

Attribute	Study or publication	Period	Annual	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
Mean dry spell duration*	Pilot study long	1889–2018	WD	n/a											
Mean dry spell duration*	Pilot study short	1950–2018	N	n/a											
Maximum dry spell duration*	Pilot study long	1889–2018	WD	n/a											
Maximum dry spell duration*	Pilot study short	1950–2018	WD	n/a											
Extreme rainfall intensity	Pilot study long	1889–2018	WI	n/a											
Extreme rainfall intensity	Gallant et al. (2007) ^{a,f}	1910–2005	N	n/a											
Extreme rainfall intensity	Pilot study short	1950–2018	I	n/a											
Extreme rainfall intensity	Gallant et al. (2007) ^{a,f}	1950–2018	N	n/a											
Extreme rainfall intensity	Theobald et al. (2016)	1958–2012	WI	n/a											
Extreme rainfall frequency	Pilot study long	1889–2018	WI	n/a											
Extreme rainfall frequency	Gallant et al. (2007) ^{a,f}	1910–2005	N	n/a											
Extreme rainfall frequency	Pilot study short	1950–2018	N	n/a											
Extreme rainfall frequency	Gallant et al. (2007) ^{a,f}	1950–2005	N	n/a											
Extreme rainfall frequency	Theobald et al. (2016)	1958–2012	N	n/a											

D = decrease; I = increase; N = no significant trends; n/a = not applicable, since analyses for the attribute or season are not available from the source; WI = widespread increase (trends that are present at more than 25% of the sites)

WD = widespread decrease (trends that are present at more than 25% of the sites).

^a Used a larger south-east Australia region for their analysis

^b Used gridded data; the comparison is based on approximately locating the pilot region from their figures

^c Analysed moderate rainfall events as events within 1 standard deviation of the mean

^d Analysed changes in the mean rainfall in the Snowy region

^e Analysed the number of heavy (1 to 2 standard deviations from mean) and very heavy (more than 2 standard deviations from mean) rainfall events

^f Defined extreme rainfall as the 95th percentile of rainfall

Having reviewed the statistical significance of trends, now consider the examination of the magnitude of trends for cases that are field significant in the pilot catchments and consistent with literature. The magnitude of trends is estimated site-wise using linear least squares regression at all 49 rainfall sites. The median absolute value of the trends and the median percentage of trends at the 49 sites are presented in Table 9.

Table 9 Median magnitude of trends at 49 pilot sites

Attribute	Period	Median trend – Annual	Median trend – DJF	Median trend – MAM	Median trend – JJA	Median trend – SON
Total rainfall in mm/decade (%/decade)	1889 to 2018	x	3.4 mm (2%)	x	-2.4 mm (-1%)	X
Total rainfall in mm/decade (%/decade)	1950 to 2018	-16.8 mm (-2.1%)	x	-10.7 mm (-5.5%)	-3.7 mm (-1.5%)	-8.1 mm (-3.7%)
Number of wet days in days/decade (%/decade)	1889 to 2018	-0.6 mm (-0.6%)	0.1 mm (0.7%)	-0.2 mm (-0.9%)	-0.3 mm (-1%)	-0.2 mm (-0.9%)
Number of wet days in days/decade (%/decade)	1950 to 2018	-1.9 mm (-1.8%)	x	-0.9 mm (-4.4%)	-0.2 mm (-0.7%)	-0.9 mm (-3.6%)
Extreme rainfall intensity in mm/day/decade (%/decade)	1889 to 2018	0.2 mm (0.5%)	n/a	n/a	n/a	n/a
Extreme rainfall intensity in mm/day/decade (%/decade)	1950 to 2018	0.3 mm (0.7%)	n/a	n/a	n/a	n/a

n/a = not applicable; x = trends that are not field significant.

Absolute trends per decade are presented first, followed by percentage change per decade in parentheses.

The rainfall totals and the number of wet days in MAM and SON exhibit the largest trends, especially in the short-term analysis. The spatial pattern of the magnitude of the short-term trends in the number of wet days and total seasonal rainfall during MAM, JJA and SON are shown in Figures 8 and 9. The decreasing trends in the number of wet days are present at a number of sites across the study region during MAM and SON. The decreasing trends in seasonal totals are more widespread during SON.

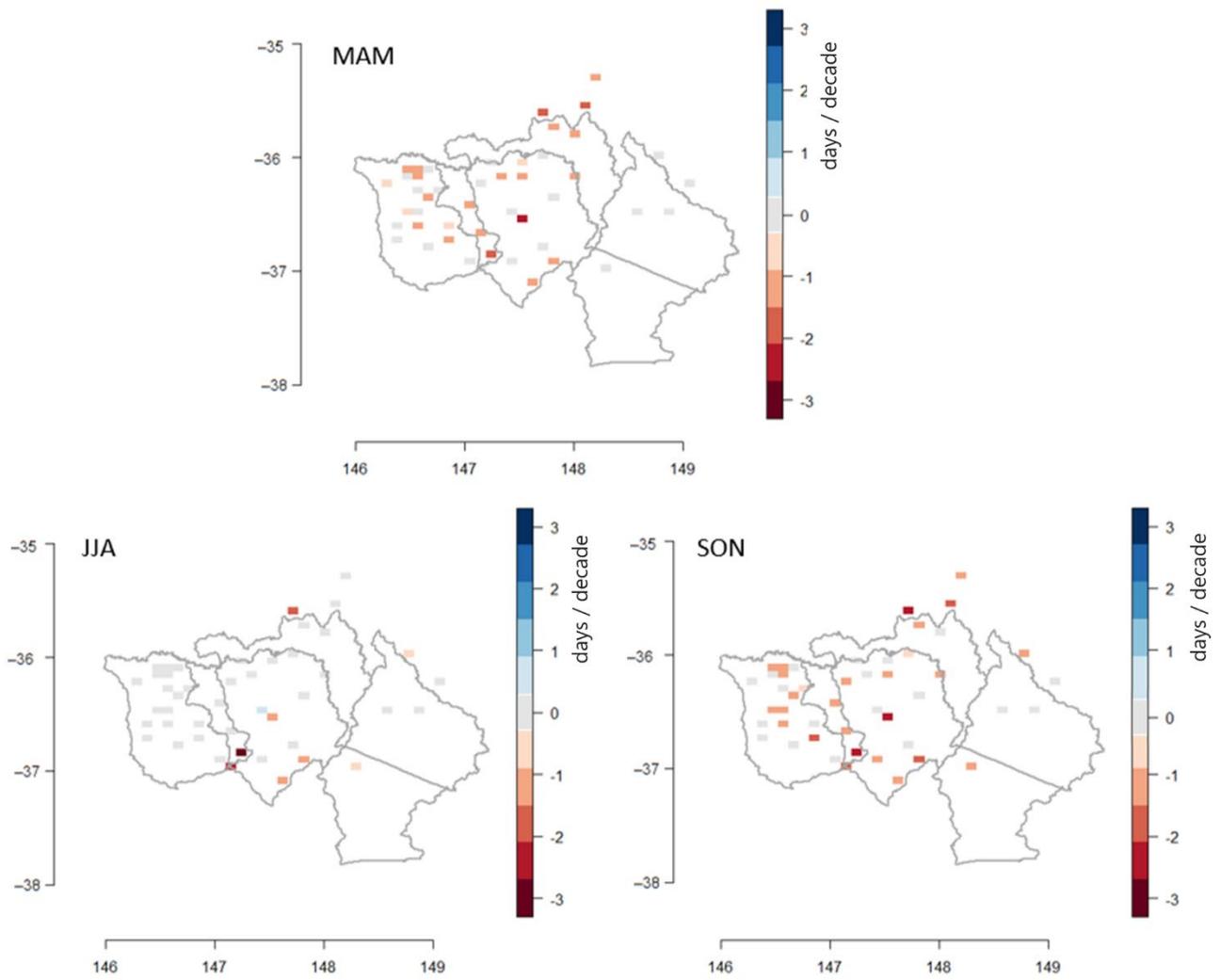


Figure 8 Short-term (1950 to 2018) seasonal trends in number of wet days (in days/decade)

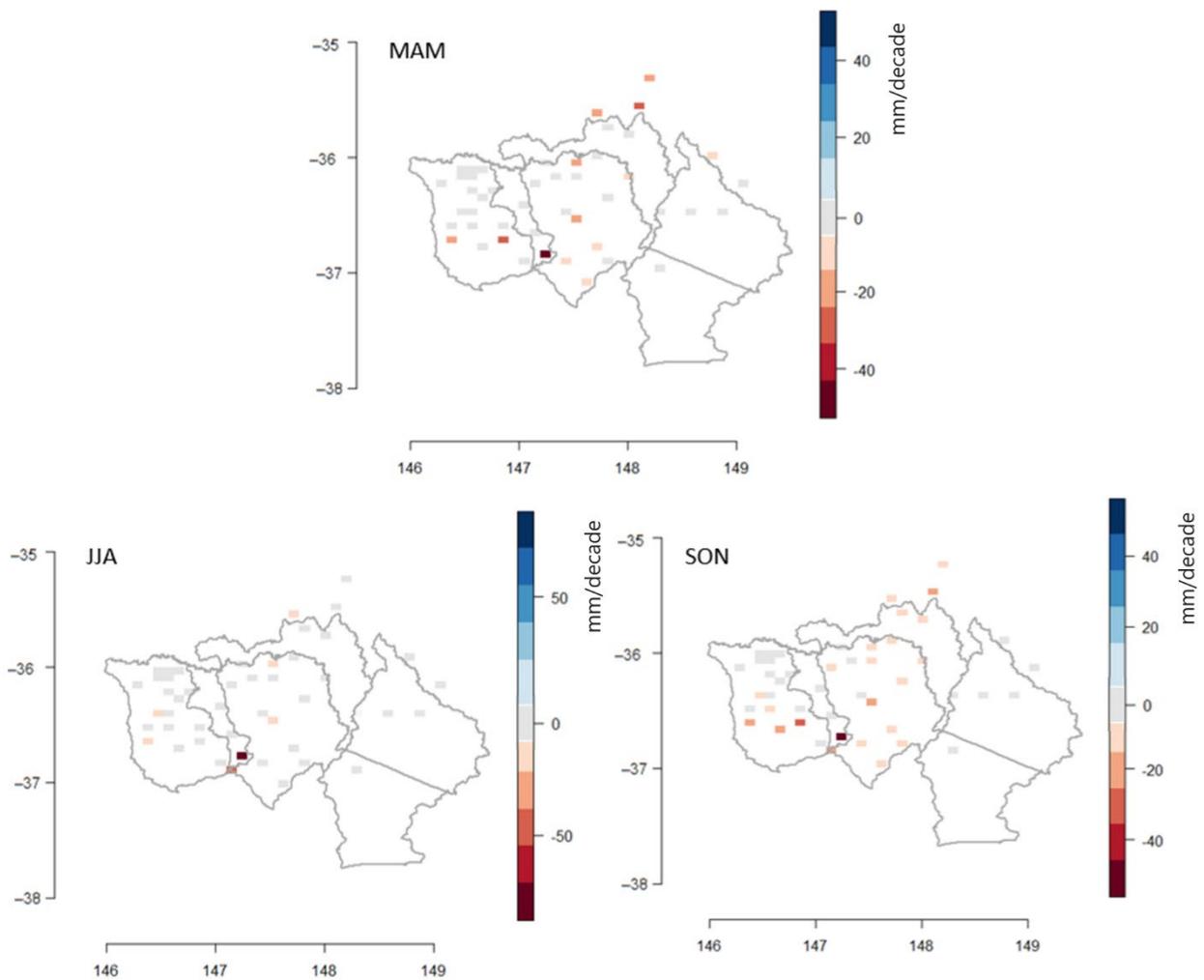


Figure 9 Short-term (1950 to 2018) seasonal trends in total rainfall (in mm/decade)

5.3.2 Trends in evapotranspiration

The trends in evapotranspiration (Mwet) are examined at the pilot sites over the full period of record (1889–2018) as well as a short-term period (1975–2018), consistent with the pan evaporation trends reported in literature. Table 10 shows the number of sites that exhibit significant trends in evapotranspiration during the 2 time periods of analysis. The long-term (1889–2018) analysis shows an increasing trend in annual total Mwet. At seasonal scale, the increase occurs during MAM, JJA and SON. The trend in JJA is widespread with 97% of the sites exhibiting an increasing trend. In the short-term (1975–2018) analysis, the annual and seasonal trends are not field significant.

Table 10 Number of sites that exhibit significant trends in annual and seasonal total evapotranspiration (Total Mwet in mm)

Period	Trend	Annual	DJF	MAM	JJA	SON
1889–2018	Pos	8 (27%) *	2 (7%)	8 (27%) *	29 (97%) *	9 (30%) *
1889–2018	Neg	1 (3%)	3 (10%)	0 (0%)	0 (0%)	0 (0%)
1889–2018	None	21 (70%)	25 (83%)	22 (73%)	1 (3%)	21 (70%)
1975–2018	Pos	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
1975–2018	Neg	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
1975–2018	None	30 (100%)	30 (100%)	30 (100%)	30 (100%)	27 (90%)

Pos = significant positive trend; Neg = significant negative trend; None = no significant trends.

* Asterisked, shaded cells indicate trends that are field significant (at 5%).

The value in brackets indicates the number of sites as a percentage of the total number of sites (30 sites).

Comparing the evapotranspiration (Mwet) trends from the pilot sites with reported trends in pan evaporation in Table 11. It is to be noted that pan evaporation records are different from the Morton wet evapotranspiration data from the pilot sites; however, this comparison is included here in the absence of trend studies in literature using estimated Morton wet evapotranspiration data. Literature reports decreasing trends in pan evaporation during the 1975–2002 period, possibly due to decreasing wind speeds or atmospheric demand (Roderick and Farquhar 2004). The trends are reported to be reversed or insignificant in south-east Australia when recent observations up to 2016 (Stephens et al. 2018) and 2018 (Ukkola et al. 2019) are used for analyses. The insignificant trends in the short-term pilot analysis are consistent with the results of Ukkola et al. (2019).

Table 11 Trends in total evaporation in the pilot sites and the changes reported in literature

Study	Period	Annual
Pilot study long	1889–2018	WI
Pilot study short	1975–2018	N
Roderick and Farquhar (2004)	1975–2002	WD
Stephens et al. (2018)	1975–2016	I or N
Ukkola et al. (2019)	1975–2018	N

I = increase; N = no significant trends; ; WD = widespread decrease (trends that are present at more than 25% of the sites); WI = widespread increase (trends that are present at more than 25% of the sites).

The literature uses pan evaporation records, which are different from the estimated Morton wet evaporation data used in the pilot study.

5.3.3 Trends in temperature

The trends in minimum and maximum temperatures from 3 ACORN–SAT v2 stations close to the pilot sites are presented here. The long-term (1913–2018) trends are calculated using available data from 2 of the stations; the short-term (1960–2018) trends are calculated using data from all 3 stations. Since the analysis uses only 3 sites, a field significance test of the results is not performed, and instead report trends at individual sites.

All the temperature sites show significant increasing trends. Tables 12 and 13 present the magnitude of significant trends in Tmax and Tmin at the site level and the magnitude of trends reported in literature in this region. The ACORN–SAT v2 sites near the pilot region show significant increasing trends in Tmax and Tmin, consistent with literature. The magnitudes of trends are also generally consistent. The short-term (1960–2018) increasing trends in Tmax and Tmin are slightly higher at these stations compared to the regional trends reported by Ashcroft et al. (2012).

Table 12 The magnitude of trends in Tmax at the ACORN–SAT v2 sites and the trends reported in literature (°C/decade)

Study or publication	Period	Annual	DJF	MAM	JJA	SON
Ashcroft et al. (2012)	1860–2011	0.04	0.03	n/a	0.08	n/a
Jones (2012)	1910–2010	0.07	n/a	n/a	n/a	n/a
Ukkola et al. (2019) *	1910–2018	0.11	0.15	0.12	0.09	0.10
Pilot study long, station 82039	1913–2018	0.15	0.10	0.22	0.11	0.12
Pilot study long, station 72150	1913–2018	0.07	0.10	0.12	N	N
Ashcroft et al. (2012)	1960–2011	0.22	0.21	n/a	0.22	n/a
Pilot study short, station 82039	1960–2018	0.35	0.31	0.31	0.26	0.45
Pilot study short, station 72150	1960–2018	0.29	0.30	0.25	0.17	0.45
Pilot study short, station 72161	1960–2018	0.37	0.40	0.26	0.33	0.52

* Studied trends in mean temperatures as the average of Tmax and Tmin.

N = no significant trends; I = increase; n/a = not applicable, i.e. analyses for the attribute or season are not available from source.

Table 13 The magnitude of trends in Tmin at the ACORN–SAT v2 sites and the trends reported in literature (°C/decade)

Study/publication	Period	ANN	DJF	MAM	JJA	SON
Ashcroft et al. (2012)	1860–2011	0.07	0.10	n/a	0.05	n/a
Jones (2012)	1910–2010	0.11	n/a	n/a	n/a	n/a
Ukkola et al. (2019) *	1910–2018	0.11	0.15	0.12	0.09	0.10
Pilot study long, station 82039	1913–2018	0.18	0.29	0.14	0.12	0.17
Pilot study long, station 72150	1913–2018	0.25	0.26	0.29	0.19	0.20
Ashcroft et al. (2012)	1960–2011	0.18	0.23	n/a	0.18	n/a
Pilot study short, station 82039	1960–2018	0.22	0.40	-	-	0.19
Pilot study short, station 72150	1960–2018	0.32	0.42	0.24	0.25	0.36
Pilot study short, station 72161	1960–2018	0.21	0.30	0.13	0.10	0.32

* Studied trends in mean temperatures as the average of Tmax and Tmin.

– dash indicates sites where trend is not reported due to missing data or questions regarding data quality

n/a = not applicable, i.e. analyses for the attribute or season are not available from source.

5.4 Results of split sample testing

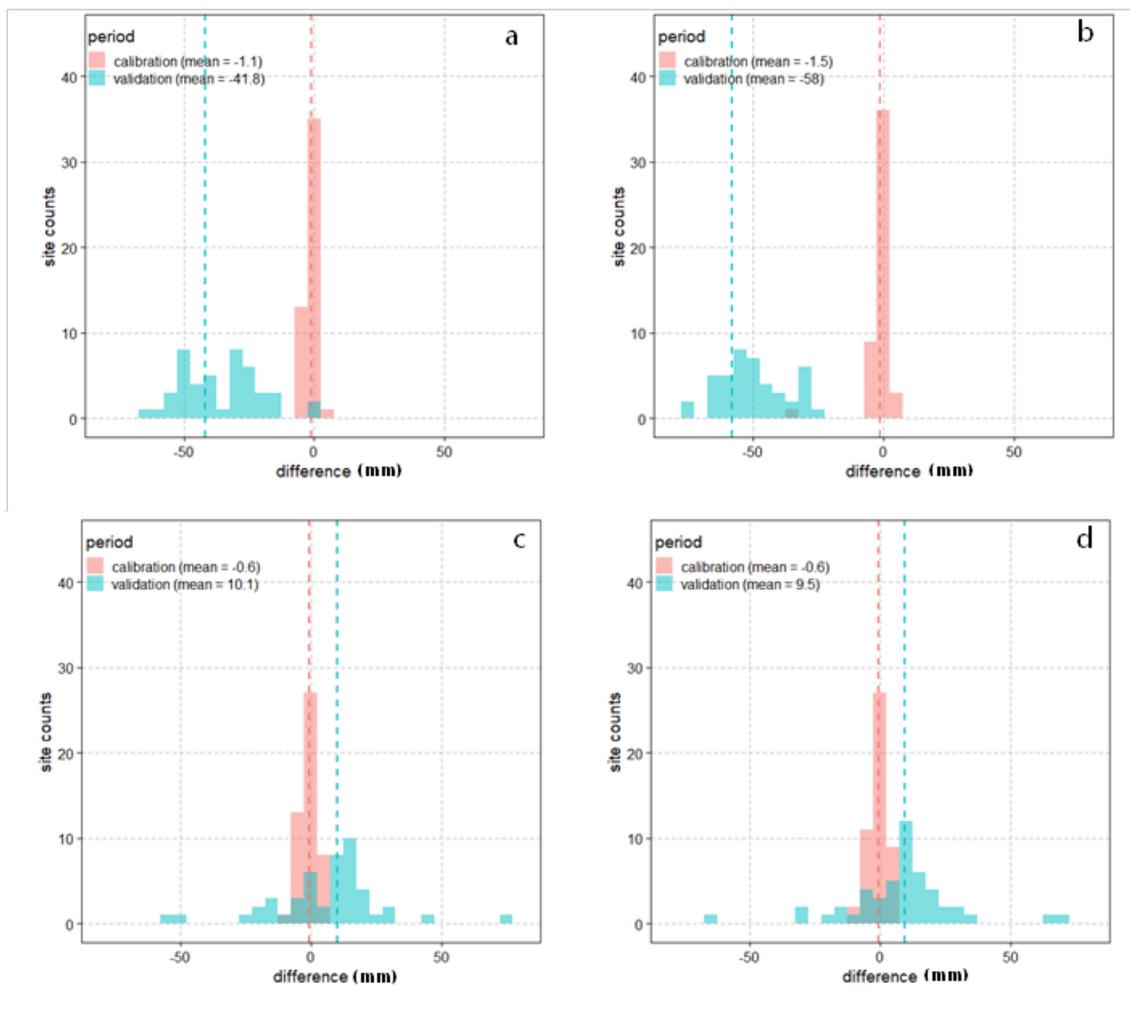
Four split sample simulations are performed using the historical data from the pilot sites to assess the ability of the stochastic model to capture the attributes of rainfall and evapotranspiration in recent observations, following the methodology outlined in Section 5.2.2. The split sample tests consist of the 1990 reference and drought reference experiments. The mean attributes of the simulated data are compared with the mean attributes of observations during the calibration and validation time periods in the subsections below.

5.4.1 Rainfall

The trend analysis showed significant trends in cool season rainfall totals, annual/cool season number of wet days, and annual extreme rainfall intensity at the pilot sites. Here these attributes are examined in the simulated data with respect to the mean values in observations during the calibration and validation periods. The results of the experiments for some key attributes of rainfall are examined in this section. The figures showing the split sample results for other attributes are included in Appendix A for brevity.

Figure 10 shows histograms of the differences in the mean MAM rainfall totals during the calibration and validation time periods from the MAM seasonal rainfall totals in the simulated time series. The 1990 reference split sample tests are shown in Figure 10(a) and (b). When stochastic models are calibrated using observed data up to 1989, the simulated data show substantial differences (42–58 mm) when compared to the post-1990 validation period. When the Millennium Drought is included in the calibration period, the MAM rainfall during the validation period is closer to both the calibration period and the stochastically simulated data, as shown in Figure 10(c) and (d). For example, the simulations

exhibit biases of 42 mm (+24%) with respect to the validation period in one of the 1990 reference experiments; the biases reduce to 10 mm (+4%) when the drought is included in the calibration period. The histograms of total rainfall during JJA and SON are available in Appendix A. In each of these experiments, the seasonal rainfall totals during the calibration period are close to the simulated data, as expected. If the seasonal rainfall totals during the validation periods are different from that in the calibration period, the mean values during the validation period are different from the simulated data as well. The mean monthly rainfall is used for the calibration of the stochastic model, and the simulated time series closely replicate this statistic during the calibration period.

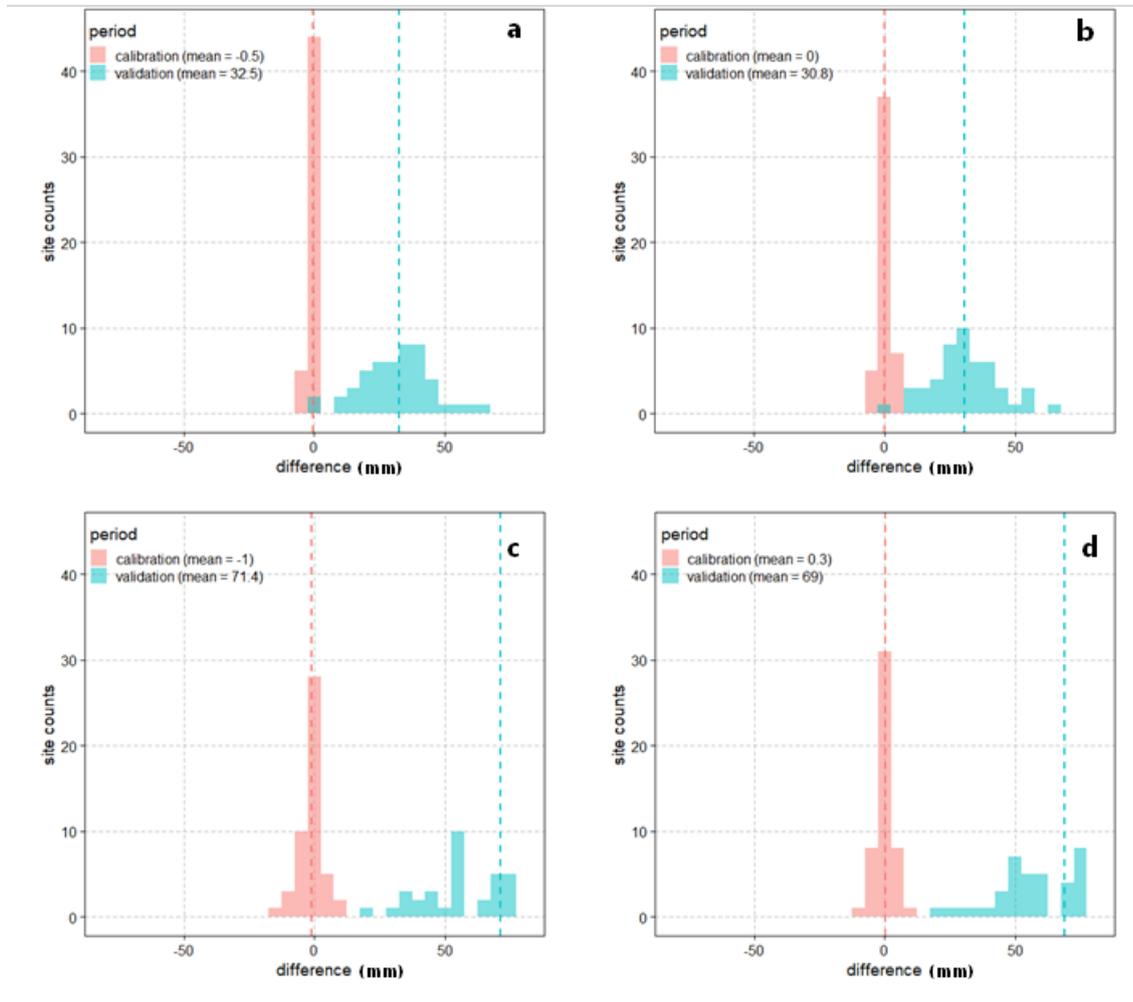


(a) 1990-2018 reference split sample test using a long calibration period (1889–1989), (b) 1990-2018 reference split sample test using a short calibration period (1950-1989), (c) drought reference (2010–2018) split sample test using a long calibration period (1889–1989), and (d) drought reference (2010–2018) split sample test using a short calibration period (1950-1989). The dashed vertical lines mark the means of the respective histograms. The magnitudes of the means are shown in the plot legends.

Figure 10 Histograms of the mean differences in total rainfall in MAM during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites

The results of the split sample tests in the MAM seasonal rainfall totals are consistent with the significant decreasing trends detected from the analysis of trends at the pilot sites. The mean total rainfall figures during MAM, JJA and SON (Appendix A) are lower during the recent validation periods compared to the earlier calibration periods used for the simulations, consistent with the negative trends detected in these seasons.

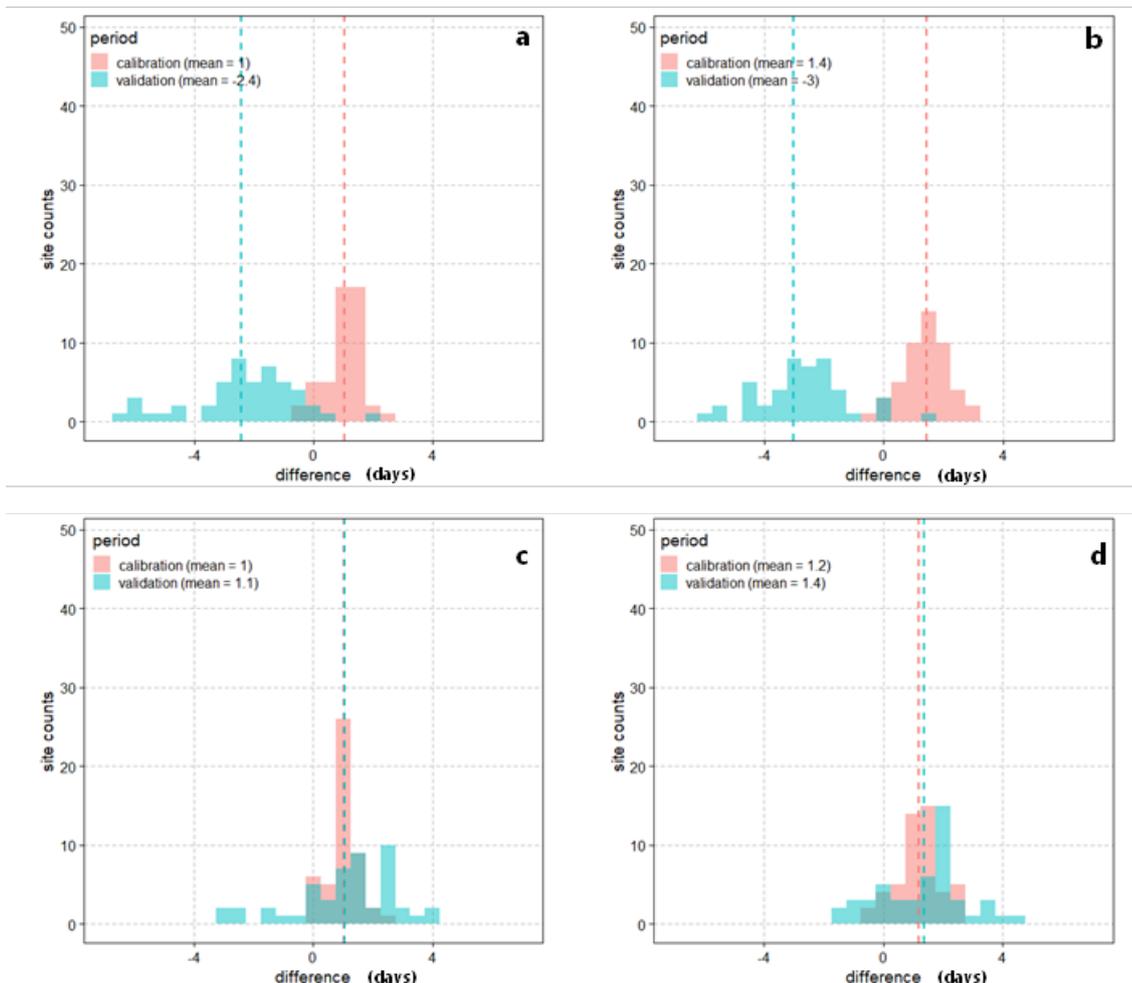
While the inclusion of the drought in the calibration period reduces the biases in MAM total rainfall in the simulations, there are increases in biases in some other statistics. The histograms of the total rainfall during DJF are shown in Figure 11. The simulations underestimate the rainfall during DJF in the 1990 reference experiments (-32mm , -16%). After the inclusion of data up to 2009 in the calibration, the biases increase (-72 mm , -29%). This change is generally consistent with the long-term positive trend in total DJF rainfall; however, the large positive difference during the recent validation period ($+69\text{--}71\text{ mm}$) could be due to statistically insignificant changes in the most recent records.



(a) 1990–2018 reference split sample test using a long calibration period (1889–1989), (b) 1990–2018 reference split sample test using a short calibration period (1950–1989), (c) drought reference (2010–2018) split sample test using a long calibration period (1889–1989), and (d) drought reference (2010–2018) split sample test using a short calibration period (1950–1989). The dashed vertical lines mark the means of the respective histograms. The magnitudes of the means are shown in the plot legends.

Figure 11 Histograms of the mean differences in total rainfall in DJF during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites

Figure 12 shows the histograms of the differences in the number of wet days during MAM during the calibration and validation periods from the stochastically generated data. When compared to the data during the calibration period, the stochastic simulations underestimate the number of wet days in all 4 experiments. The number of wet days is not a statistic that is used in the calibration of the model, and the simulated time series show a slight bias in this statistic with respect to the calibration period data (1–1.4 days per season).



(a) 1990–2018 reference split sample test using a long calibration period (1889–1989), (b) 1990–2018 reference split sample test using a short calibration period (1950–1989), (c) drought reference (2010–2018) split sample test using a long calibration period, and (d) drought reference (2010–2018) split sample test using a short calibration period (1950–1989). The dashed vertical lines mark the means of the respective histograms. The magnitudes of the means are shown in the plot legends.

Figure 12 Histograms of the mean differences in the number of wet days in MAM during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites

In the 1990 reference experiments, the mean number of wet days during the validation period is lower than that during the calibration period, consistent with the negative trends detected in this attribute (Table 8). The mean number of wet days during the validation period is lower than that in the simulations by 2.4–3 days per season. When the Millennium Drought is included in the calibration period, the number of wet days during the calibration and validation time periods are very similar. Hence in the drought reference split sample tests, the simulations underestimate the number of wet days with respect to both the calibration and validation period statistics.

The histograms of the differences in the number of wet days during the other seasons (JJA, SON and DJF) are available in Appendix A. Like the results during MAM, the mean number of wet days from the stochastic simulations is biased lower compared to the calibration period (by 0.8 to 2.7 days per season) for all seasons and both periods of calibration. The differences of the mean statistics during the validation period vary. The mean number of wet days in JJA and SON is lower during the validation periods compared to the calibration periods, consistent with the trends detected in the pilot sites. In contrast, the number of wet days in DJF during the validation period is higher, especially during the 2010–2018 validation period (+4.1 to +4.3 days per season). While the long-term trend in the number

of wet days in DJF are positive at the pilot sites, there are no significant short-term trends during this season; therefore, this positive signal in the number of wet days in DJF during the most recent validation period could be due to statistically insignificant trends in the recent record.

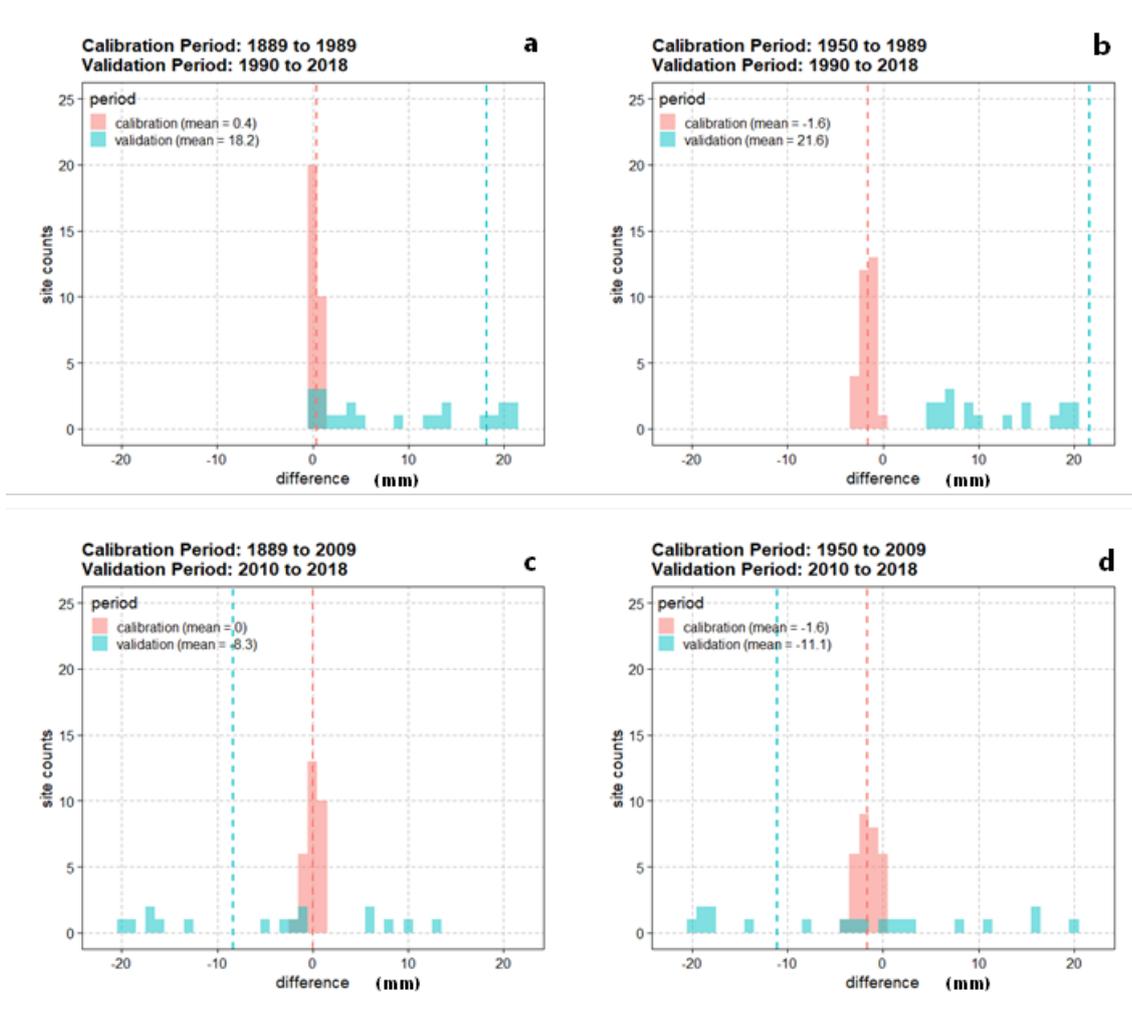
The histograms of annual extreme rainfall intensity are available in Appendix A. The annual mean extreme rainfall intensity during the calibration period is close to that in the simulated data. During the validation period, the rainfall intensity is higher than the simulations, consistent with the positive trends at the pilot sites.

To summarise:

- The Millennium Drought is a high-leverage event; the statistics of the simulated rainfall can vary significantly depending on whether the drought is included in the calibration period or validation period.
- The inclusion of the Millennium Drought improves some statistics during the validation period but leads to a deterioration in others: The simulations show biases with respect to the validation period data, and the sign and magnitude of biases vary with season and period. In the 1990 reference split sample test, both the validation period mean MAM rainfall and the MAM number of wet days are biased higher in the simulations. The statistics match better in the drought reference tests once the Millennium Drought is included in the calibration period. However, other statistics, such as the total rainfall during DJF and annual extreme rainfall intensity, show larger biases in the drought reference experiments.

5.4.2 Evapotranspiration

The mean annual total evapotranspiration from the stochastic simulations is compared with the mean annual evapotranspiration during the calibration and validation periods in Figure 13. The simulations match the annual evapotranspiration values during the calibration period in all the simulations. In the 1990 reference split sample tests, the mean evapotranspiration during the validation period is higher than the simulated totals (and calibration period) by 18–20 mm. When the model is calibrated using data up to 2009, the annual totals from the simulations are closer to the annual totals during the validation period (mean difference: 8–11 mm). The split sample tests using shorter and longer calibration periods do not show any major influences on the differences between the simulations and the validation period.



(a) 1990-2018 reference split sample test using a long calibration period (1889-1989), (b) 1990-2018 reference split sample test using a short calibration period (1950-1989), (c) drought reference (2010-2018) split sample test using a long calibration period (1889-1989), and (d) drought reference (2010-2018) split sample test using a short calibration period (1950-1989). The dashed vertical lines mark the means of the respective histograms. The magnitudes of the means are shown in the plot legends.

Figure 13 Histograms of the mean differences in total annual evapotranspiration during the calibration and validation time periods from stochastic simulations (in mm) at 30 pilot sites

6 Summary of key findings from assessment of non-stationarity

Based on an analysis of multiple lines of evidence, comprising a review of available literature and the assessment of pilot sites in the Ovens, Upper Murray and Snowy catchments using both Mann Kendall and split sample tests, there is evidence of non-stationarity in the record for rainfall and temperature. The non-stationarity in evapotranspiration is not as well established, likely in part due to the different processes driving the different 'types' of evapotranspiration (for example, pan versus Morton). Specific findings varied depending on the dataset used (for example, gridded versus point; whether the data had been homogenised or not; the spatial domain of the data; the temporal period of analysis), but the broad conclusions are:

- **Temperature records are non-stationary** – Literature documents temperature increases in this region, especially post-1960 (Jones 2012; CSIRO and BOM 2015). The Climate Change in Australia initiative reports an increase of 0.8 °C in the Murray basin cluster (which contains the pilot catchments) over 1910–2013 assuming a linear trend, with higher trends for temperature minimums than for maximums; this is broadly consistent with a mean temperature increase over Australia by just over 1 °C during the slightly longer period 1910–2018. Climate projections indicate that increases of 0.6–1.3 °C are expected in the Murray basin cluster in the near term (2020–2039) with respect to a baseline of 1986–2005. The homogenised temperature sites near the pilot region analysed here also show statistically significant increases. The increases amount to 0.8–1.5 °C in maximum temperatures and 1.9–2.6 °C in minimum temperatures during the period 1913–2018.
- **Cool season and annual rainfall totals are non-stationary** – Literature documents declines in cool season (April to October) rainfall by 10–20% in south-east Australia since the mid-1990s, predominantly in autumn and early winter. The trend assessment using data from the pilot sites is in agreement with literature and shows short-term decreases in autumn, winter and spring rainfall totals. The short-term (1950–2018) trends are strongest in autumn – the median decline in autumn rainfall at the pilot sites amounts to 5.5%/decade for the period 1950–2018; the median declines in annual total rainfall at the pilot sites amount to 2.1%/decade for the same period.
- **There are trends in multiple attributes of rainfall** – Literature reports decreasing trends in the number of wet days during the cool season; the signal also exists in the data from the pilot sites. In addition, the pilot sites show a short-term decline in spring (SON) rainfall. This result is not consistent with other large-scale studies in this region but is consistent with regional studies in nearby catchments. There is an increasing trend in annual extreme rainfall intensity at the pilot sites. There is less consensus in literature on extreme rainfall intensity; however, the trend in the pilot sites is consistent with the reported trend in a nearby catchment.
- **Non-stationarity in the evapotranspiration is not as well established** – Literature documents negative trends in pan evaporation over the period 1975–2002, whereas studies using more recent data report insignificant/increasing trends. There are long-term increases in annual Morton wet evapotranspiration at the pilot sites; the short-term trends are not statistically significant. The pilot study results are not directly comparable with available literature based on pan evaporation data, and so some uncertainty remains regarding the non-stationarity of evapotranspiration data.
- **Multiple climate drivers influence the rainfall in south-east Australia** – Literature documents that the declining cool season rainfall is associated with an expansion of the tropics, increasing intensity of the subtropical ridge over the continent and positive trends in the SAM. Literature indicates that these changes in large-scale patterns during the cool season are at least partly attributable to climate change (CSIRO 2012; Hope et al. 2017). Other climate drivers, notably ENSO

and IOD, influence the interannual variability regional rainfall, primarily affecting rainfall in winter and during the warm season (CSIRO and BOM 2015; Hope et al. 2017).

- **The period used for calibration influences the statistics of the simulated stochastic data** – The split sample tests show that the Millennium Drought is a ‘high-leverage’ event, in the sense that statistics of the simulated rainfall can vary significantly depending on whether the drought is included in the calibration period or validation period. When the drought is included in the calibration period, the match between the simulations and observed data post-drought is dependent upon the statistic and season under consideration. The inclusion of the drought in the calibration period brings the autumn rainfall and number of wet days in the simulated data close to recent (2010–2018) observations, but results in larger deviations in simulated summer rainfall, number of wet days in summer and extreme rainfall intensity. For example, when the drought is excluded from calibration, the simulations show mean biases in the MAM rainfall during the validation period of +42 mm (+24%). When the drought is included in calibration the bias reduces; the simulations show biases in total MAM rainfall during the validation period of –10 mm (–4%); however, in the same tests, the bias in DJF rainfall increases from a mean of –32 mm (–16%) to –71 mm (–29%) with the inclusion of the drought in the calibration period.

Although the pilot analysis does not seek to attribute trends to a specific driver (whether it be natural climate variability, anthropogenic climate change or other potential drivers of change), several studies have attributed the temperature shifts/trends and cool season rainfall trends to anthropogenic influences (CSIRO 2012; Jones 2012; CSIRO and BOM 2015, Hope et al. 2017). The SEACI synthesis report states that ‘the decline in rainfall across south-eastern Australia was at least partly attributable to climate change’ (CSIRO 2012, p. 4). The Climate Change in Australia report notes that the literature on rainfall attribution in southern Australia is based on ‘implied attribution’, using inferred causality from the attribution of large-scale drivers. The report states that more formal attribution to definitive causes is yet to be established due to a range of uncertainties, but at the same time notes that the ‘drying across southern Australia cannot be explained by natural variability alone’ (CSIRO and BOM 2015, p. 45).

Future regional projections indicate a warmer and drier climate with respect to a baseline of 1986–2005. In the Ovens and Murray catchments, VCP19 projections (Clarke et al. 2019b) report annual rainfall decreases during the 2020–2039 period. The projected median change is –6% (range of –12% to –4%) under the medium emissions scenario, and –11% (range of –18% to –3%) under the high emissions scenario, with further declines anticipated in subsequent decades. Climate projections for Murray basin cluster indicate increases in potential evapotranspiration in all seasons; VCP19 projects 8–10% increases in pan evaporation by the 2030s (i.e. 2020–2039 time-slice centred on 2030). The projected changes for the Murray basin cluster are based on Morton’s wet potential evapotranspiration, whereas the pan evaporation projections from VCP19 is based on pan evaporation modelled directly by the climate model.

7 Options for stochastic time series generation in the southern basin

Due to reported trends in the historical record for the southern basin, which is at least partly attributable to anthropogenic influences, there is a need to consider the implications of non-stationarity on the methodology used for the stochastic risk assessment. The working definition of non-stationarity used here is that the parameters of the stochastic model (and thus the statistics associated with the stochastically generated data) vary with time. This time variation means that stochastic results need to be reported relative to a specific period. The following terminology is used in this report.

- **Historical record.** This is used synonymously with the instrumental record (although noting that paleoclimatic records of the IPO are used to inform the stochastic sequences in previous studies). For the purposes of this report, the historical record is usually interpreted to mean the period 1890–2018, as this is the period of infilled data available from the SILO database.
- **Baseline period.** The term baseline is usually reported relative to a climatological baseline. The specific baseline period will depend on the climate models used and various other factors. The World Meteorological Organization (WMO) suggests defining climate baselines of at least 30 years in duration (WMO 1989). Further studies document that a baseline of 40 years is required for hydrological investigations in south-east Australia to capture the significant year-to-year variability associated with key processes (Potter et al. 2016).
- **Current climate.** This is interpreted here to mean the risk at the current time (for example, the year 2023, which is the date this report was written) or for a window centered on the current time.
- **Future climate.** This can refer to any future period for which projections are available and is usually reported as a static estimate using some window, typically 20–30 years. Consistent with WMO applications, hydrological applications usually benefit from longer windows to increase the signal-to-noise ratio; however, there may be significant changes over this period (that is, the projected data within a 30-year window may not be stationary). It is noted that near-term climate projections have started encompassing the current climate; for example, VCP19 uses a future window of 2020–2039 and thus encompasses the current year.

An example can help understanding of these terms: the temperature in Victoria is reported to have increased 0.5 °C over the historical record from 1910 to 1995, and the current (2018) temperature is about 0.6 °C warmer than the 1986 to 2005 baseline (Clarke et al. 2019a). Future temperature projections suggest increases of between 0.5 °C and 1.3 °C for the period 2020 to 2039, relative to the same baseline.

The stochastic method used for the northern basins is based on the historical record and simulated to be conditional to the IPO, with the implicit assumption that the climate is stationary over this period. These simulations are intended to capture variability in the baseline historical record, which was composed of (SILO infilled) climatic measurements from the period 1890–2018 as well as IPO distribution dwell times informed by paleoclimatic records. The Department of Planning and Environment developed additional methodology to amalgamate the variability from the stochastic replicates with climatic signals informed by regional climate models (RCMs). This method applies NARCLiM-based scaling factors to the stochastic model outputs.

Trade-offs are associated with modifying this methodology to incorporate climate non-stationarity, balancing the complexity of the climate signal, the strength of various non-stationarity elements as compared to the underlying variability, and the variety of methods that may be available to explicitly

accommodate known trends. Here a range of options available to account for the presence of non-stationarity in stochastic time series generation are outlined. The options vary in complexity.

7.1 Use the entire historical record to calibrate the stochastic model

Using the entire historical record to calibrate the stochastic model the simplest option, forming a naive default. The advantage of this option is that it makes full use of the available historical data (1889–2018), which improves the statistical precision of parameters used in the stochastic model. The disadvantage is that significant long-term shifts in the historical record (as opposed to ‘natural’ variability including interdecadal variations associated with the IPO) may lead to generated sequences that are not representative of current climate.

Moreover, assessments of future risk based on climate model results are invariably derived with respect to a historical baseline, and the extent to which the full (1889–2018) historical record is reflective of any given climatological baseline used for climate projections is unclear and is likely to depend on the specific climate variable, as well as the statistics for those variables (given the different findings from including/excluding the Millennium Drought depending on whether one was looking at MAM seasonal total, MAM number of wet days, JJA seasonal total or extreme rainfall intensity).

7.2 Use a climatological baseline to calibrate the stochastic model

This option involves using a shorter climatological baseline for calibration of the model, which can then be scaled using climate model outputs to derive both current and future estimates of variability. The baseline could be chosen to be consistent with the NARcliM scaling factors that are used for future projections by the department. The benefit of this approach is that it maintains basic consistency with NARcliM by having an identical baseline. The NARcliM 1.0 baseline is 1990–2009, and the NARcliM 1.5 baseline is 1950–2005. The method is thus suitable for NARcliM scaling-based current and future climate applications.

The trade-off is that the initial years of data (say, prior to 1950 or 1990 depending on the version of NARcliM) are no longer utilised and there is a corresponding loss of precision in the parameters used as the basis of the stochastic model. The significant interannual and interdecadal variability in particular means that depending on whether one includes (say) the Millennium Drought as part of the baseline will make a significant difference on the simulated results.

It is noted that this baseline will reflect neither current nor future climate, and thus climate model outputs will be required to adjust the stochastic sequences to reflect these periods. This requires decisions on the climatological baselines and, if the NARcliM outputs are to be used, the specific version of NARcliM to be used for analysis. In particular, the NARcliM 1.0 baseline of 20 years is, for the reasons described above, too short to capture the historical climate.

As already discussed, the WMO recommends a minimum baseline length of 30 years (WMO 1989). Guidelines developed under VicCI recommend a baseline of at least 40 years, given the high interannual variability in rainfall and runoff in this region (Potter et al. 2016). Another disadvantage of the method is that the use of simple scaling to represent future climate would not reflect changes in statistics such as the number of wet days or extremes.

Scaling of the baseline climate using climate models may present significant challenges. Multiplying rainfall by seasonal change factors will mean that the number of wet days will stay constant and the extremes will change in the same manner as seasonal rainfall totals, which is inconsistent with the findings of the non-stationarity analysis presented above. It may be that these other statistics are of 'second order' importance compared to seasonal and annual totals, depending on the hydrological application (for example, long-term catchment yield assessments), but this would need to be tested. Alternatively, more complex scaling options, such as a quantile-based approach, could be used to jointly change the seasonal totals, wet days and extremes in a manner that reflects the climate projections.

Several further variants could be considered that involve using a hybrid baseline, but they present additional challenges that are likely to overcomplicate the methodology. The variants are detailed in Section 7.3.

7.3 Use a hybrid baseline of the historical climate to calibrate the stochastic model

Another option is to utilise a hybrid baseline period. For example, some statistics may have negligible or no shift between a baseline and a longer historical period. Therefore, additional precision would be achieved by calibrating selected statistics to a longer record, while reserving only some key statistics (such as the mean value) to the baseline period. The challenge with this approach is that the term baseline is no longer informative, since it is not clear what period is being represented in the calibration.

Another option is to scale the historical record up to the values observed in the baseline period. This is one of the approaches to address non-stationarity recommended by the expert review panel. The benefit of this approach is that it provides a mechanism for utilising the entire record. A challenge with this approach is that it requires the scaling of the prior historical record to be representative. This procedure is unlikely to be straightforward, because:

- Multiple statistics are changing in different ways, so scaling by seasonal or annual totals may not lead to appropriate adjustments of the other statistics. In particular, the pilot study identified negative trends in cool season totals, negative trends in number of wet days and positive trends in extreme intensity, which would require care in how the historical observations are mapped.
- Historical non-stationarity is likely to encompass a combination of natural climate variability and anthropogenic climate change, so one would need to assess the magnitude of historical change *specifically associated with anthropogenic climate change* to scale the historical values appropriately. If this is not done, then aspects of the signal associated with natural variability may also inadvertently be removed (at least in part), which is unlikely to be desirable.

7.4 Use the inverse method to generate a current or future climate

This approach could be used to target statistic values representative of a particular period, such as the current or future climate. Unlike traditional stochastic generation methods in which the stochastic generator parameters are estimated to achieve the best performance over a calibration dataset, the 'inverse' method defines a set of 'target' statistics (for example, total annual rainfall, seasonal rainfall, number of wet days, intermittency, extremes) and then calibrates the stochastic generator against those statistics. The advantage of this method is that it can target the most recent values corresponding to current climate, whereas other methods (described in Section 7.2) would use a baseline of historical climate spanning the selected years, which may not reflect current climate.

The key challenge of this method is defining the target statistics reflecting current or future climate. This is not a trivial exercise and may require examining historical climate trends in combination with climate modelling evidence. Moreover, this approach has not yet been implemented in large-scale real-world analyses, and thus would require a period of method testing prior to widespread implementation.

7.5 Use weather typing to capture specific mechanisms of generation of precipitation

This method involves stratifying the historical record by synoptic types (such as ECLs) to provide a scaling method that can explicitly account for mechanisms with strong expected change. The methodology would involve additional complication introduced by the multiple weather types. It may also double count the effects of scaling when deployed alongside methods that are needed to scale other components. For example, if ECLs were introduced as a scaling category with their own relationships to future climate, other scaling relationships (for example, shift in seasonality, shift in extremes) would need to be developed for the specific case of non-ECL weather events. This method is likely to be highly prospective, given the potential of double counting the effects of synoptic meteorology if they are also explicitly included in climate model projections.

The trade-offs involved in the 5 options detailed in Sections 7.1–7.5 are summarised in Table 14. The discussion of method advantages and disadvantages is predicated on the assumption that the historical record is non-stationary and that at least part of the non-stationarity is attributable to anthropogenic climate change.

Table 14 Summary of options for stochastic time series generation in the southern basin, with advantages/disadvantages described in the context of a non-stationarity climate signal

Option	Advantages	Disadvantages	Application
1. Entire record	Long record, historical enables increased precision of parameter estimation.	Reflective of neither the current climate nor a climatological baseline	Reflection of long-term historical climate
2. Historical baseline	Matches NARClIM baseline and thus can be used as the basis for current and future climate assessments.	Need to choose which NARClIM baseline, noting that NARClIM 1.0 is too short relative to WMO and VicCI recommendations. Not reflective of current climate and thus requires some level of subsequent processing. Simple seasonal scaling would not reflect changes in wet days and other statistics, although there may be alternative (e.g. quantile-based) methods that could address this issue.	NARClIM-derived current and future climate applications
3. Hybrid historical baseline	Uses full historical record. Possible extra precision in some attributes compared to baseline only.	Advantage of extra data before baseline is unclear. There are multiple complicating factors associated with implementation of this method, including complex variations with key attributes, and the need to separate the natural and anthropogenic components of any historical trends prior to adjustments.	Adjustments would enable stochastic data to reflect current climate
4. Inverse method	Can be implemented to match current and future climate. Allows changes in all key attributes.	Difficulties Method development and testing required.	Could be designed to reflect current and future climate
5. Weather typing	Allows for specific mechanisms (ECLs).	Additional complication introduced by multiple types. Scaling is complicated by types, and there exists a significant possibility of double counting .	Specific allowance for future changes to weather types rather than simple scaling

Another point raised by the independent review panel is the role of multiple climate drivers in the southern region. Evidence in literature indicates a strong trend in the SAM during the cool season, while the warm season is influenced by ENSO and IOD. Note that the previous stochastic generation methodology based on climatic relationship to the IPO relied on partitioning the climate under the

assumption of stationarity in the IPO. This is conceptually different from assuming a trend signal in the relevant climate drivers, since it does not require projection of the trend associated with the climate driver. Also, in locations where the partition between IPO positive and negative shows negligible difference, the partition collapses back to the underlying marginal model (that is, the model does not become 'worse' by inclusion of the IPO in areas where the IPO is not important).

Although there is the potential to develop a non-stationary stochastic model that directly conditions on trend properties (such as SAM), this is inadvisable. Apart from the many technical methodological challenges of developing the non-stationary model and conditioning it on an RCM, any application of RCM-derived scaling factors of meteorological variables (such as temperature, rainfall and evapotranspiration) would then need to exclude the role of that driver as part of the analysis to avoid double counting this effect (given that the intent of the stochastic conditioning is to have the influence of the relevant driver already included in the analysis). Therefore, this is not presented as a viable option.

7.6 Recommendation for stochastic generation

Among the options for stochastic generation, methods that apply a 'standard' calibration to a selected baseline are preferable given the degree of expertise and modelling required for options such as synoptic typing and the inverse method.

The selection of the calibration period is important in that it should be as long as possible while also providing consistency with the method of assessment of climate projections. The NARCIIM 1.5 baseline is 1950–2005, which has a suitable minimum baseline length for hydrological studies, given typical recommendations of 30 years (WMO 1989) or 40 years (Potter et al. 2016). Restricting calibration of the stochastic model to the period 1950–2005 implies a loss of precision in parameter estimates, but the need for compatibility with climate model scaling is greater.

The complexity of change in climate attributes gives rise to the possibility that scaling by the application of change factors to seasonal or annual totals may not lead to appropriate adjustments of the other statistics, for example, a negative trend in the number of wet days. For this reason, quantile scaling methods should be applied in preference to simple scaling.

The scaling approach should be applied to different RCMs rather than an ensemble mean given the likely variation between models. Due to the nonlinear nature of hydrological transformation, this approach will ensure that model uncertainty in hydrological estimates is suitably accounted for.

8 References

- Alexander LV and Arblaster JM (2009) 'Assessing trends in observed and modelled climate extremes over Australia in relation to future projections', *International Journal of Climatology*, 29(3):417–435, doi:10.1002/joc.1730.
- Alexander LV and Arblaster JM (2017) 'Historical and projected trends in temperature and precipitation extremes in Australia in observations and CMIP5', *Weather and Climate Extremes*, 15:34–56, doi:10.1016/j.wace.2017.02.001.
- Alexander LV, Hope P, Collins D, Trewin B, Lynch A and Nicholls N (2007) 'Trends in Australia's climate means and extremes: a global context', *Australian Meteorological Magazine*, 56(1):1–18.
- Allen KJ, Anchukaitis KJ, Grose MG, Lee G, Cook ER, Risbey JS, O'Kane TJ, Monselesan D, O'Grady A, Larsen S and Baker PJ (2019) 'Tree-ring reconstructions of cool season temperature for far southeastern Australia, 1731–2007', *Climate Dynamics*, 53(1–2):569–583, doi:10.1007/s00382-018-04602-2.
- Ashcroft L, Karoly DJ and Gergis J (2012) 'Temperature variations in southeastern Australia, 1860–2011', *Australian Meteorological and Oceanographic Journal*, 62:227–245, doi:10.22499/2.6204.004.
- Bennett B, Thyer M, Leonard M, Lambert M and Bates B (2018) 'A comprehensive and systematic evaluation framework for a parsimonious daily rainfall field model', *Journal of Hydrology*, 556:1123–1138, doi:10.1016/j.jhydrol.2016.12.043.
- BOM (Bureau of Meteorology) and CSIRO (2019) *State of the climate 2018*, Commonwealth of Australia, Canberra.
- Clarke J, Grose M, Thatcher M, Hernaman V, Heady C, Round V, Rafter T, Trenham C and Wilson L (2019a) *Victorian climate projections 2019: technical report*, CSIRO, Melbourne.
- Clarke J, Grose M, Thatcher M, Round V and Heady C (2019b) *Ovens Murray climate projections 2019*, CSIRO, Melbourne.
- CSIRO (2012) *Climate and water availability in south-eastern Australia: a synthesis of findings from Phase 2 of the South Eastern Australian Climate Initiative (SEACI)* (Vol. 41), CSIRO, Australia.
- CSIRO and BOM (2015) *Climate change in Australia: projections for Australia's NRM regions, Technical Report*, CSIRO and BOM, Australia.
- DELWP (Department of Environment, Land, Water and Planning) (2016) *Guidelines for assessing the impact of climate change on water supplies in Victoria*, DELWP, Melbourne.
- Do HX, Westra S and Leonard M (2017) 'A global-scale investigation of trends in annual maximum streamflow', *Journal of Hydrology*, 552:28–43, doi:10.1016/j.jhydrol.2017.06.015.
- Dowdy AJ, Mills GA, Timbal B, Griffiths M and Wang Y (2013) 'Understanding rainfall projections in relation to extratropical cyclones in eastern Australia', *Australian Meteorological and Oceanographic Journal*, 63(3):355–364, doi:10.22499/2.6303.001.

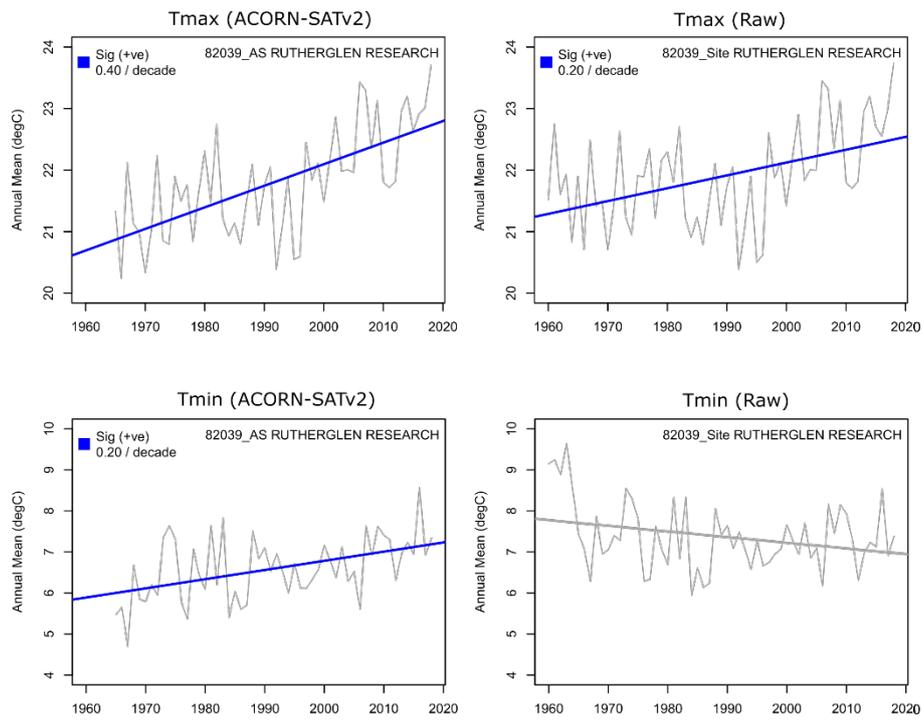
- Dowdy AJ, Pepler A, Di Luca A, Cavicchia L, Mills G, Evans JP, Louis S, McInnes KL, Walsh K (2019) 'Review of Australian east coast low pressure systems and associated extremes', *Climate Dynamics*, 53(7–8):4887–4910, doi:10.1007/s00382-019-04836-8.
- Fiddes SL, Pezza AB and Barras V (2015) 'Synoptic climatology of extreme precipitation in alpine Australia', *International Journal of Climatology*, 35(2):172–188, doi:10.1002/joc.3970.
- Gallant AJE, Hennessy KJ and Risbey J (2007) 'Trends in rainfall indices for six Australian regions: 1910–2005', *Australian Meteorological Magazine*, 223–239, <https://www.researchgate.net/publication/242601092>.
- Hewaratchi AP, Li Y, Lund R and Rennie J (2017) 'Homogenization of daily temperature data', *Journal of Climate*, 30(3):985–999, doi:10.1175/JCLI-D-16-0139.1.
- Hope P, Timbal B, Hendon H, Ekström M and Potter N (2017) *A synthesis of findings from the Victorian Climate Initiative (VicCI)*, <http://www.bom.gov.au/research/projects/vicci/>.
- Jones RN (2012) 'Detecting and attributing nonlinear anthropogenic regional warming in southeastern Australia', *Journal of Geophysical Research: Atmospheres*, 117(D4), doi:10.1029/2011JD016328.
- Karoly DJ and Braganza K (2005) 'A new approach to detection of anthropogenic temperature changes in the Australian region', *Meteorology and Atmospheric Physics*, 89(1–4):57–67, doi:10.1007/s00703-005-0121-3.
- Lavender SL and Abbs DJ (2013) 'Trends in Australian rainfall: contribution of tropical cyclones and closed lows', *Climate Dynamics*, 40(1–2):317–326, doi:10.1007/s00382-012-1566-y.
- Leonard M and Westra S (2020). *Methodology report for multisite rainfall and evaporation data generation of the northern basin*, report prepared for Department of Planning and Environment by the University of Adelaide.
- Di Luca A, Evans JP, Pepler A, Alexander L and Argüeso D (2015) 'Resolution sensitivity of cyclone climatology over eastern Australia using six reanalysis products', *Journal of Climate*, 28(24):9530–9549, doi:10.1175/JCLI-D-14-00645.1.
- Di Luca A, Evans JP, Pepler A, Alexander L and Argüeso D (2016) 'Evaluating the representation of Australian east coast lows in a regional climate model ensemble', *Journal of Southern Hemisphere Earth Systems Science*, 66, 108–124, doi:10.22499/3.6602.003.
- Marohasy JJ and Abbot JW (2016) 'Southeast Australian maximum temperature trends, 1887–2013: an evidence-based reappraisal', in Easterbrook D (ed) *Evidence-based climate science*, 2nd edn, (pp. 83–99), Elsevier, doi:10.1016/B978-0-12-804588-6.00005-7.
- Nguyen H, Hendon HH, Lim EP, Boschhat G, Maloney E and Timbal B (2018) 'Variability of the extent of the Hadley circulation in the southern hemisphere: a regional perspective', *Climate Dynamics*, 50(1–2):129–142, doi:10.1007/s00382-017-3592-2.
- Nicholls N (2010) 'Local and remote causes of the southern Australian autumn-winter rainfall decline, 1958–2007', *Climate Dynamics*, 34(6):835–845, doi:10.1007/s00382-009-0527-6.
- Pepler AS, Di Luca A, Ji F, Alexander LV, Evans JP and Sherwood SC (2015) 'Impact of identification method on the inferred characteristics and variability of Australian east coast lows', *Monthly Weather Review*, 143(3):864–877, doi:10.1175/MWR-D-14-00188.1.

- Pepler AS, Di Luca A, Ji F, Alexander LV, Evans JP and Sherwood SC (2016) 'Projected changes in east Australian midlatitude cyclones during the 21st century', *Geophysical Research Letters*, 43(1):334–340, doi:10.1002/2015GL067267.
- Potter N, Chiew F, Zheng H, Ekström M and Zhang L (2016) *Hydroclimate projections for Victoria at 2040 and 2065*, CSIRO, doi:10.4225/08/5892224490e71.
- Risbey JS, McIntosh PC and Pook MJ (2013) 'Synoptic components of rainfall variability and trends in southeast Australia', *International Journal of Climatology*, 33(11):2459–2472, doi:10.1002/joc.3597.
- Roderick ML and Farquhar GD (2004) 'Changes in Australian pan evaporation from 1970 to 2002', *International Journal of Climatology*, 24(9):1077–1090, doi:10.1002/joc.1061.
- Squintu AA, van der Schrier G, Brugnara Y and Klein Tank A (2019) 'Homogenization of daily temperature series in the European Climate Assessment & Dataset', *International Journal of Climatology*, 39(3):1243–1261, doi:10.1002/joc.5874.
- Stephens CM, McVicar TR, Johnson FM and Marshall LA (2018). 'Revisiting pan evaporation trends in Australia a decade on', *Geophysical Research Letters*, 45(20):11,164–11,172, doi:10.1029/2018GL079332.
- Taschetto AS and England MH (2009) 'An analysis of late twentieth century trends in Australian rainfall', *International Journal of Climatology*, 29(6):791–807, doi:10.1002/joc.1736.
- Theobald A, McGowan H and Speirs J (2016) 'Trends in synoptic circulation and precipitation in the Snowy Mountains region, Australia, in the period 1958–2012', *Atmospheric Research*, 169:434–448, doi:10.1016/j.atmosres.2015.05.007.
- Timbal B and Hendon H (2011) 'The role of tropical modes of variability in recent rainfall deficits across the Murray-Darling Basin', *Water Resources Research*, 47(12), doi:10.1029/2010WR009834.
- Timbal B, Abbs D, Bhend J, Chiew F, Church J, Ekström M, Kirono D, Lenton A, Lucas C, McInnes K, Moise A, Monselesan D, Mpelasoka F, Webb L and Whetton P (2015) 'Murray basin cluster report', *Climate change in Australia projections for Australia's natural resource management regions: cluster reports*, (eds) Ekström M, Whetton P, Gerbring C, Grose M, Webb L and Risbey J, CSIRO and BOM.
- Trewin B (2013). 'A daily homogenized temperature data set for Australia', *International Journal of Climatology*, 33(6):1510–1529, doi:10.1002/joc.3530.
- Trewin B, Braganza K, Fawcett R, Grainger S, Jovanovic B, Jones D, Martin D, Smalley R and Webb V (2020) 'An updated long-term homogenized daily temperature data set for Australia', *Geoscience Data Journal*, 7(2):149–169, doi:10.1002/gdj3.95.
- Ukkola AM, Roderick ML, Barker A and Pitman AJ (2019) 'Exploring the stationarity of Australian temperature, precipitation and pan evaporation records over the last century', *Environmental Research Letters*, 14:124035, doi:10.1088/1748-9326/ab545c.
- Vincent LA, Hartwell MM and Wang XL (2020). 'A third generation of homogenized temperature for trend analysis and monitoring changes in Canada's climate', *Atmosphere-Ocean*, 58(3):173–191, doi:10.1080/07055900.2020.1765728.

van Wijngaarden WA and Mouraviev A (2016) 'Seasonal and annual trends in Australian minimum/maximum daily temperatures', *The Open Atmospheric Science Journal*, 10:39–55, doi:10.2174/1874282301610010039.

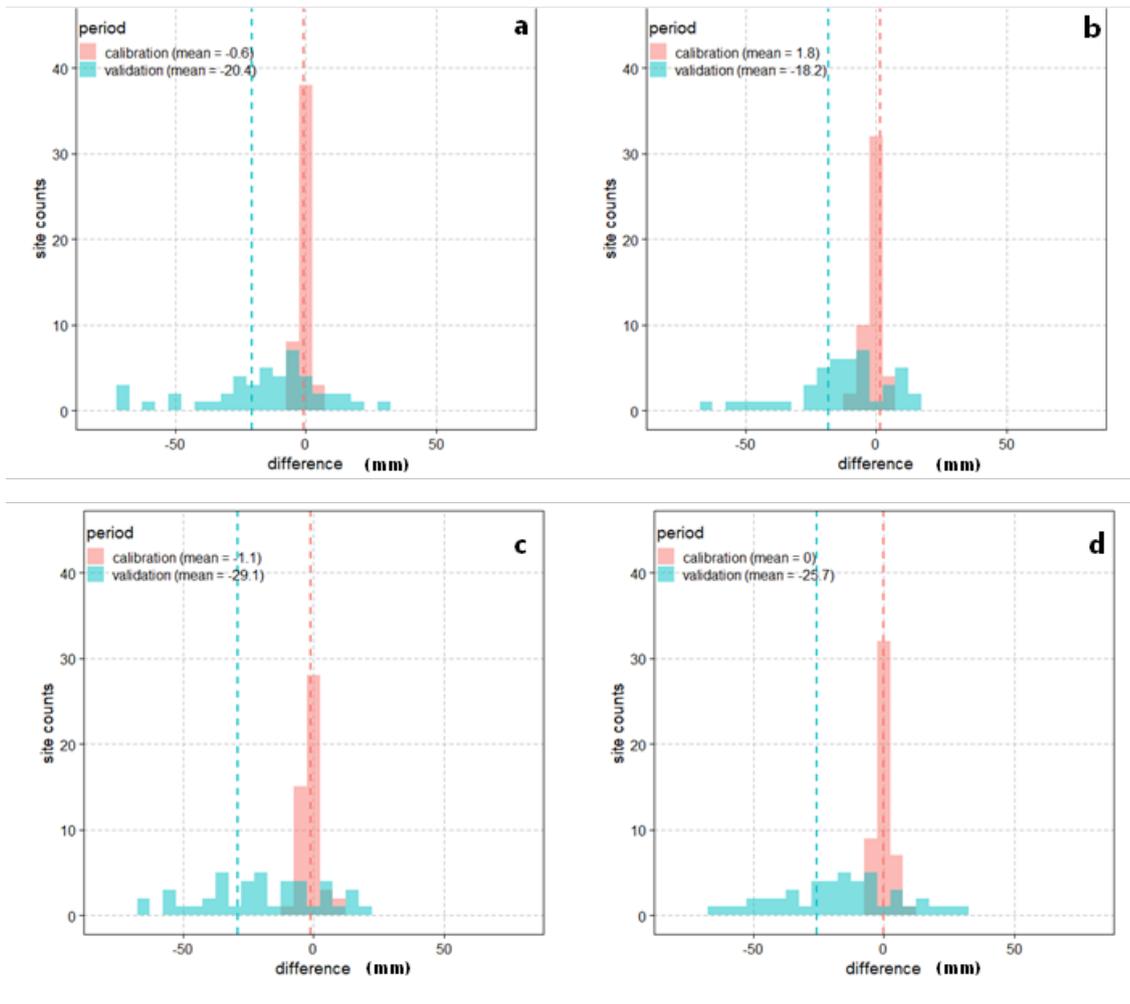
WMO (World Meteorological Organization) (1989) *Calculation of monthly and annual 30-year standard normals*, WMO, Geneva, <https://library.wmo.int/idurl/4/43912>.

Appendix A – Supplementary material



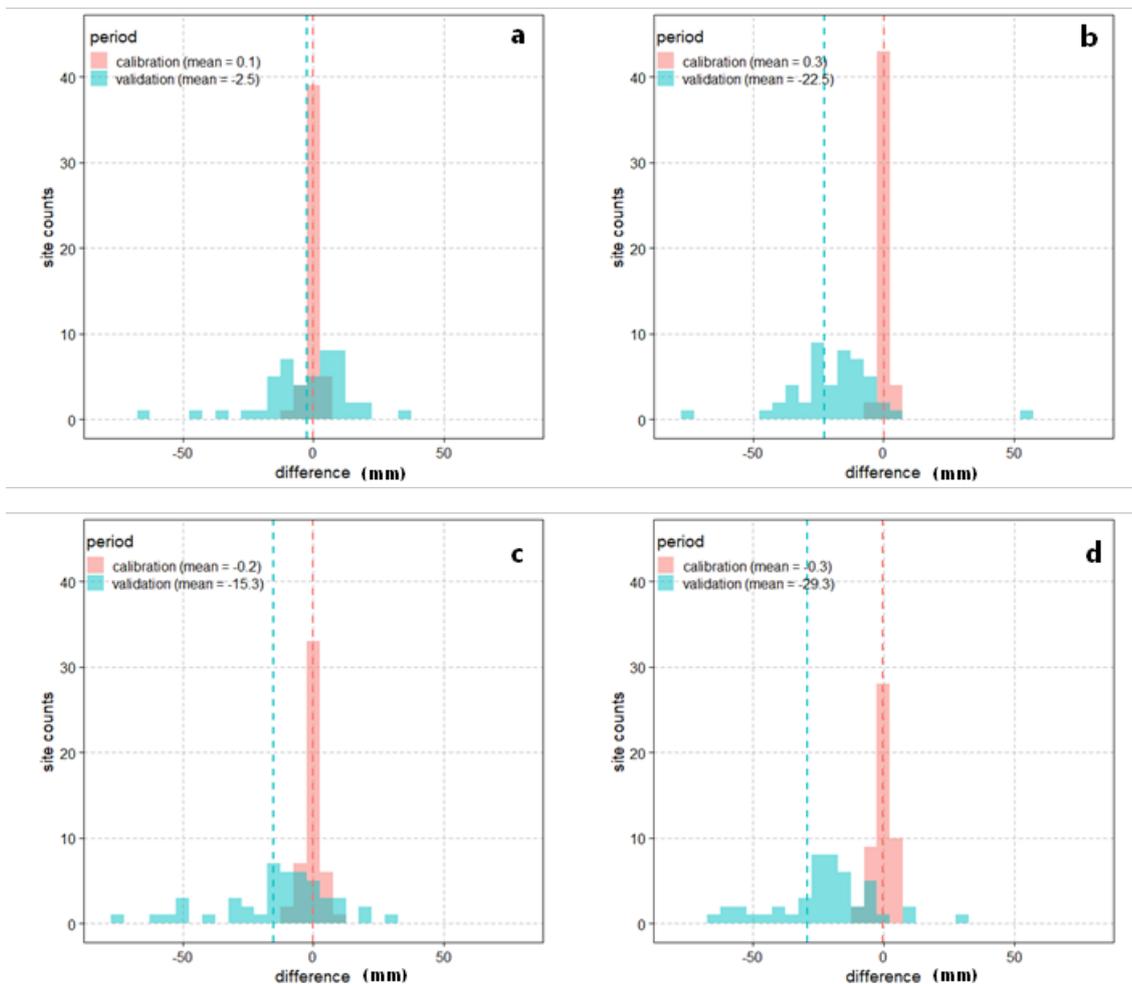
Blue trendlines indicate the presence of significant trends (at 5% level) using the Mann Kendall trend test. The data from 1960 to 1964 is missing in the ACORN-SAT v2 data.

Figure A.1 Annual mean short-term (post-1960) Tmax and Tmin at site 82039 (Rutherglen Research) from raw station data and homogenised ACORN-SAT v2 data with linear trendlines



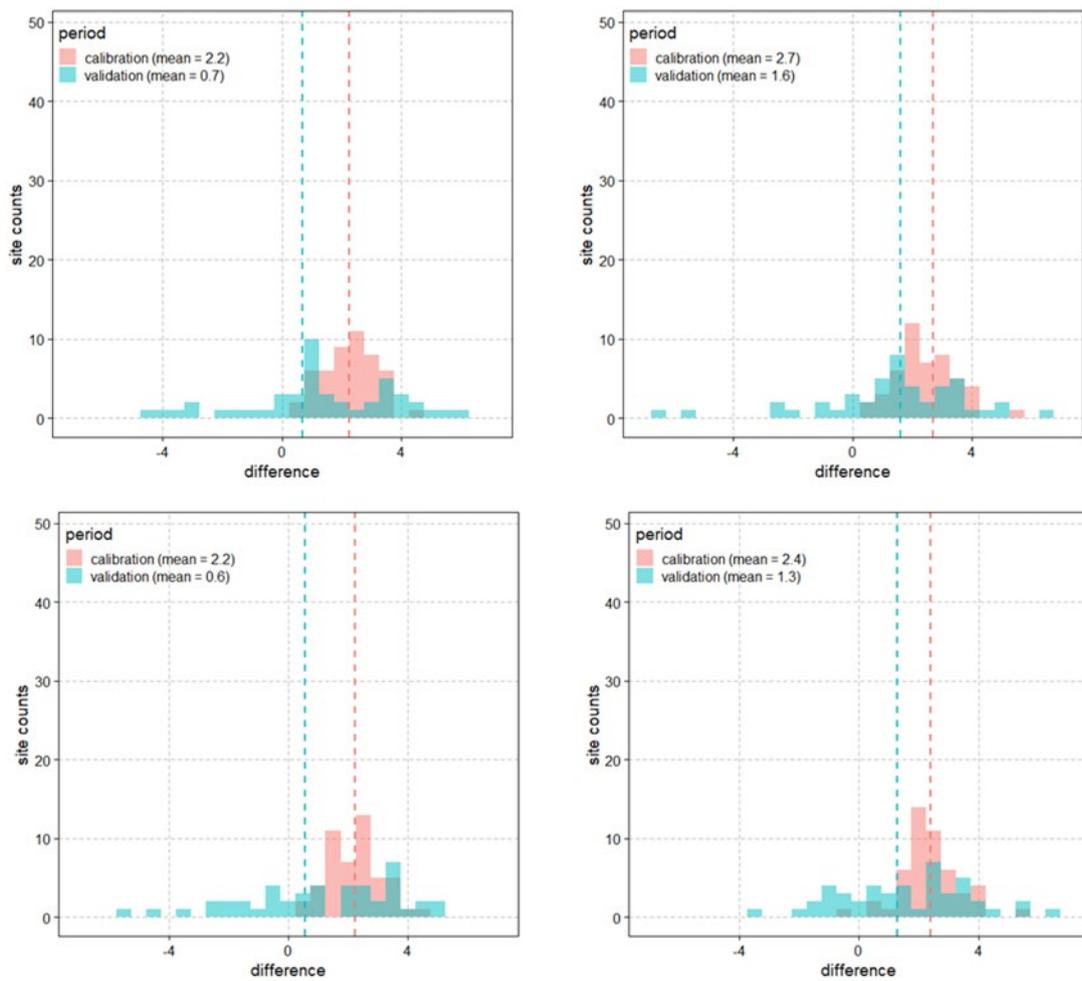
(a) 1990 reference split sample test using a long calibration period, (b) 1990 reference split sample test using a short calibration period, (c) drought reference split sample test using a long calibration period, and (d) drought reference split sample test using a short calibration period. The dashed vertical lines mark the means of the respective histograms.

Figure A.2 Histograms of the mean differences in total rainfall in JJA during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites



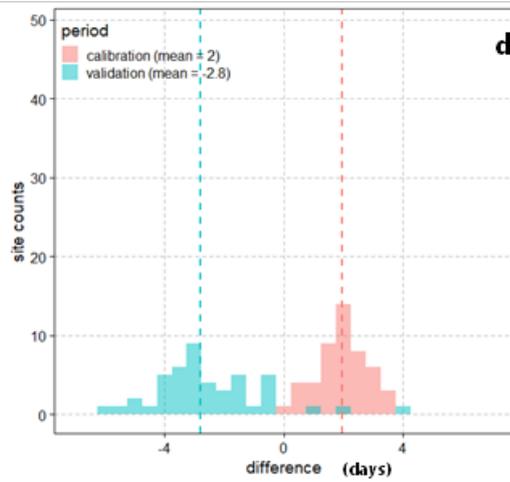
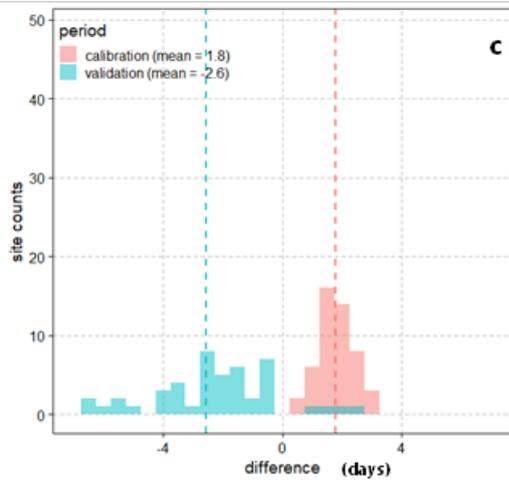
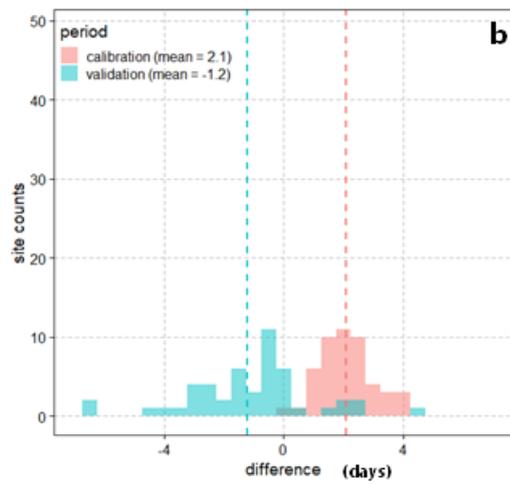
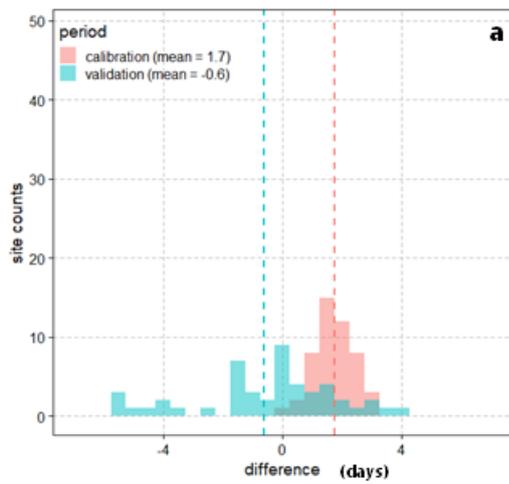
(a) 1990 reference split sample test using a long calibration period, (b) 1990 reference split sample test using a short calibration period, (c) drought reference split sample test using a long calibration period, and (d) drought reference split sample test using a short calibration period. The dashed vertical lines mark the means of the respective histograms.

Figure A.3 Histograms of the mean differences in total rainfall in SON during the calibration and validation time periods from stochastic simulations (in mm) at 49 pilot sites



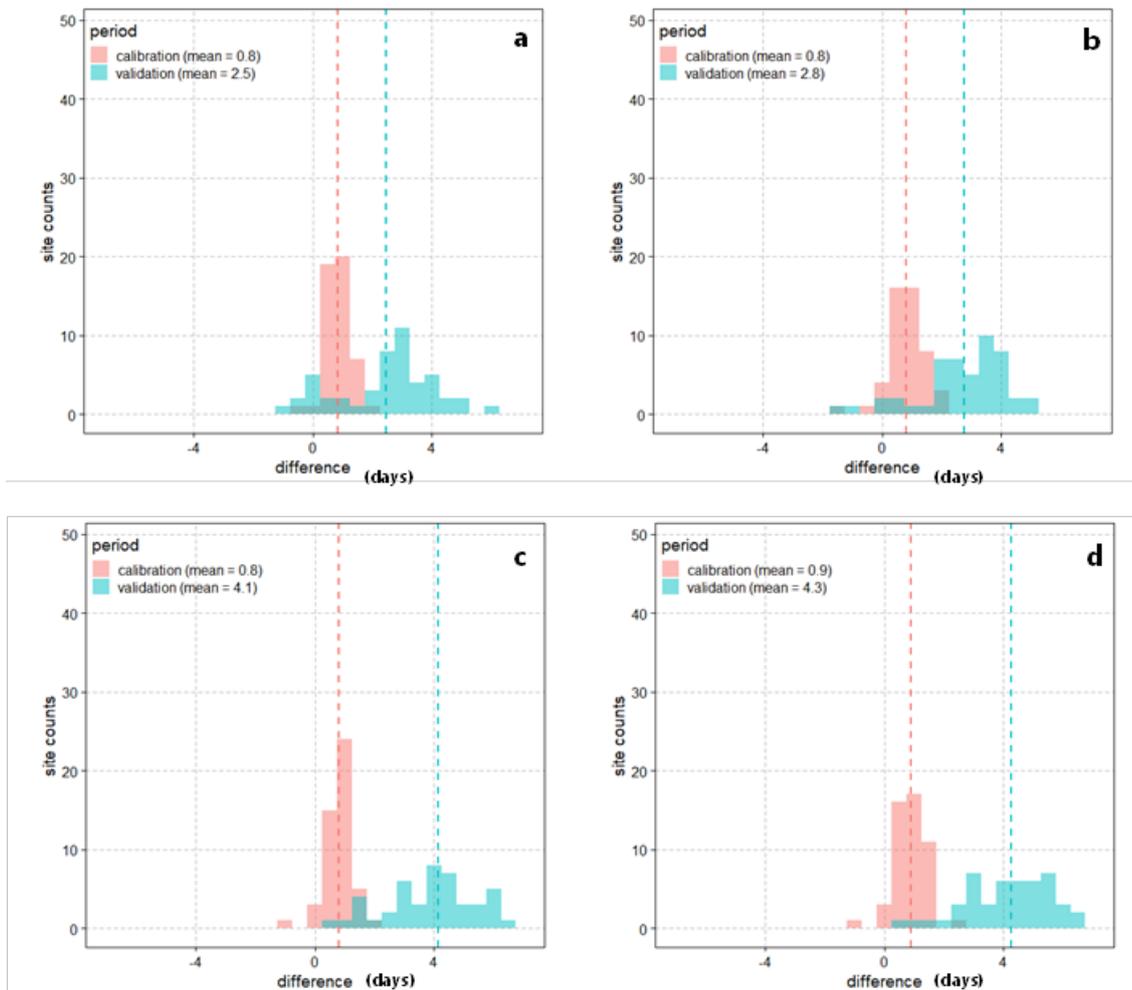
(a) 1990 reference split sample test using a long calibration period, (b) 1990 reference split sample test using a short calibration period, (c) drought reference split sample test using a long calibration period, and (d) drought reference split sample test using a short calibration period. The dashed vertical lines mark the means of the respective histograms.

Figure A.4 Histograms of the mean differences in number of wet days in JJA during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites



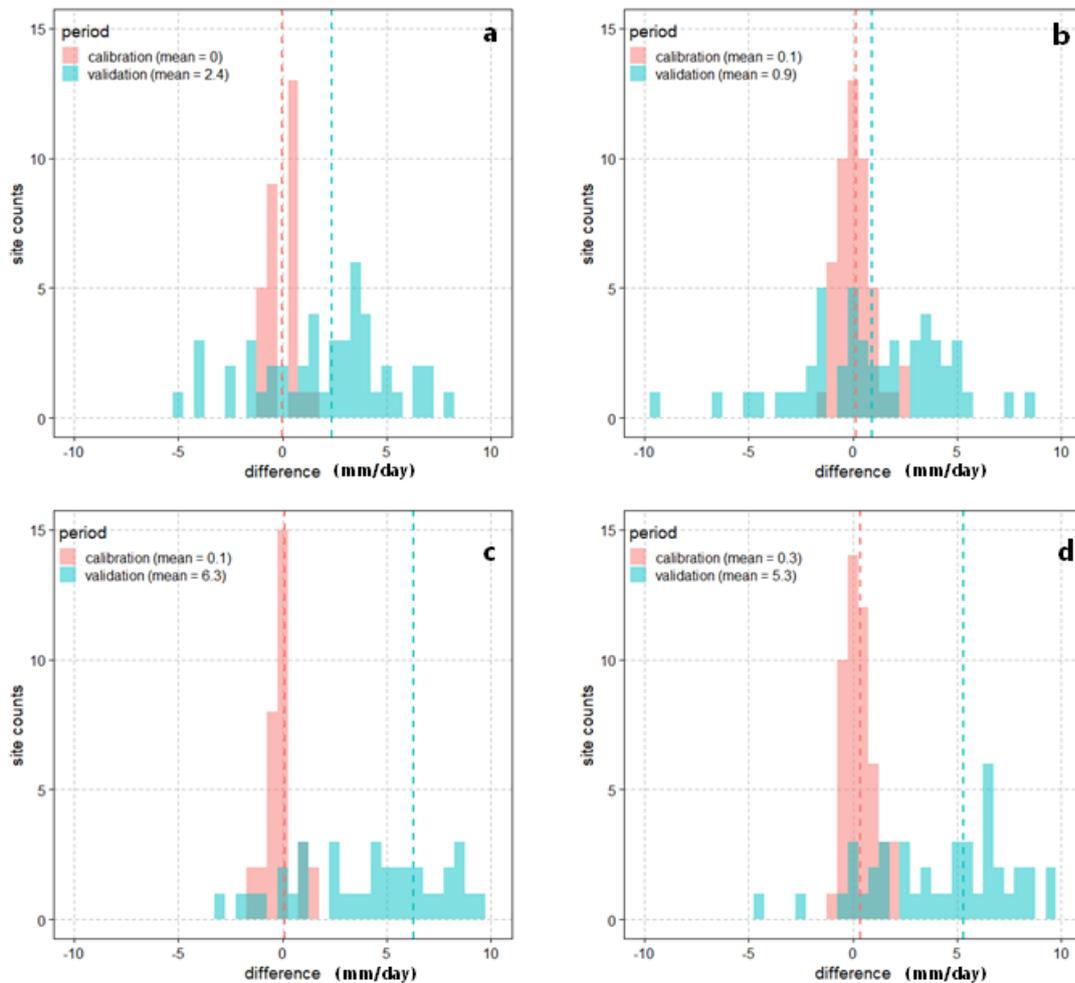
(a) 1990 reference split sample test using a long calibration period, (b) 1990 reference split sample test using a short calibration period, (c) drought reference split sample test using a long calibration period, and (d) drought reference split sample test using a short calibration period. The dashed vertical lines mark the means of the respective histograms.

Figure A.5 Histograms of the mean differences in number of wet days in SON during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites



(a) 1990 reference split sample test using a long calibration period, (b) 1990 reference split sample test using a short calibration period, (c) drought reference split sample test using a long calibration period, and (d) drought reference split sample test using a short calibration period. The dashed vertical lines mark the means of the respective histograms.

Figure A.6 Histograms of the mean differences in number of wet days in DJF during the calibration and validation time periods from stochastic simulations (in days) at 49 pilot sites



(a) 1990 reference split sample test using a long calibration period, (b) 1990 reference split sample test using a short calibration period, (c) drought reference split sample test using a long calibration period, and (d) drought reference split sample test using a short calibration period. The dashed vertical lines mark the means of the respective histograms.

Figure A.7 Histograms of the mean differences in annual extreme rainfall intensity during the calibration and validation time periods from stochastic simulations (in mm/day) at 49 pilot sites