Evaluation report for multisite rainfall, evapotranspiration and temperature data generation of the southern region

Michael Leonard, Duc Cong Hiep Nguyen 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: PUB24/329

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

List of figures

List of tables

Executive summary

The New South Wales Department of Climate Change Energy, the Environment and Water (the department) has adopted a risk-based methodology to account for climate variability and change in developing its regional water strategies. This report concerns the second of 3 phases investigating the southern region catchments (Snowy, Murray, Upper Murrumbidgee, Murrumbidgee, northern Victoria, South Australia) in the Murray–Darling Basin. A first phase reviewed literature on physical mechanisms influencing the regional climate of south-east Australia and conducted a pilot study on non-stationarity to inform the selection of methods for characterising future climate risk. The current phase documented here is the generation of stochastic rainfall, evapotranspiration and temperature to form an underlying dataset that will be modified in the third phase according to information from the initial pilot study and based on projections from the NSW and Australian Regional Climate Modelling (NARCliM) project.

The method used in the southern region is similar to previous work generating stochastic data for the northern and Macquarie regions (Leonard et al. 2019, 2020). Specifically, the method involves the use of stochastically generated long-term sequences of climate data that characterise the current climate. The sequences are subsequently routed through rainfall-runoff models to develop risk-based estimates of flow to inform economic and environmental assessment. The stochastic model is calibrated against historical (observed/reconstructed) records of daily rainfall, evapotranspiration and temperature from 1889 to 2018 to then generate synthetic data for 10,000 years that are reflective of variability over the historical period. The stochastic sequences provide insights into natural climate variability beyond the available observations.

The southern region has 370 rainfall, 414 evapotranspiration and 97 temperature sites. Due to having minimum and maximum daily temperature data (rather than average daily), a basic extension was made to the stochastic model to ensure that temperature minimum time series do not exceed temperature maximum time series. A further point of methodological development was the use of rainfall-runoff transformation as an integrative measure of the quality of the climatic data. This approach is implemented separately from the standard assessment of climatic data reported using traffic lights. In response to the evaluation against runoff data, additional post-processing is implemented to improve the performance of simulated runoff, albeit at the expense of metrics focused solely on the climatic inputs.

The simulated climatic data are systematically evaluated according to a variety of statistics (for example, multiannual totals, monthly distribution, extreme values), where rules are established a priori to assign labels ‘Good’, ‘Fair’ or ‘Poor’ to a given statistic (Leonard et al. 2020). The labels are then applied to a set of metrics across all sites to reflect the overall performance category. Results are presented in Table 1. Although the labelling is standardised, the evaluation nonetheless requires interpretation to account for whether departures stem from the model or from the data and whether they are significant. Like previous regions, the model typically performs well for annual and monthly totals, with a known limitation of Poor performance in the variability of the number of wet days between months (Leonard et al. 2020) and where Fair performance is typically seen when evaluating extremes (Leonard et al. 2020).

Table 1 Overall performance summary for simulated rainfall, evapotranspiration and temperature data

Element	Statistic	Overall performance category
Rainfall	Annual total rain 1 year	Good
Rainfall	Annual total rain 2 year	Good
Rainfall	Annual total rain 5 year	Good
Rainfall	Annual total rain 10 year	Good
Rainfall	Mean monthly rain totals	Good
Rainfall	Standard deviation monthly rain totals	Good
Rainfall	Mean monthly wet proportion	Good
Rainfall	Standard deviation monthly wet proportion	Poor
Rainfall	1-day rainfall extremes	Fair
Rainfall	2-day rainfall extremes	Fair
Rainfall	3-day rainfall extremes	Fair
Evapotranspiration	Annual total evapotranspiration 1 year	Good
Evapotranspiration	Annual total evapotranspiration 2 year	Good
Evapotranspiration	Annual total evapotranspiration 5 year	Fair†
Evapotranspiration	Annual total evapotranspiration 10 year	Fair to Good
Evapotranspiration	Mean monthly evapotranspiration totals	Good
Evapotranspiration	Standard deviation monthly evapotranspiration totals	Good
Temperature	Annual mean temperature 1 year	Good
Temperature	Annual mean temperature 2 year	Good
Temperature	Annual mean temperature 5 year	Fair
Temperature	Annual mean temperature 10 year	Poor
Temperature	Mean monthly temperature mean	Good
Temperature	Standard deviation monthly temperature mean	Good
Temperature	1-day temperature extremes	Poor†
Temperature	2-day temperature extremes	Fair
Temperature	3-day temperature extremes	Fair

†Denotes statistic that includes obvious discrepancies at some sites due to infilling of the historical record (see also Leonard et al. 2019).

The department provided runoff models for 27 catchments across the region, which collectively required 128 climatic time series as inputs and produced outputs at 88 streamflow locations. The stochastic streamflow time series are compared to 'virtual observed' streamflow, which is the output generated when the 1890–2018 observed climatic data are input to the rainfall-runoff model (Bennett et al. 2018), enabling a consistent like-for-like comparison on the effect of the runoff transformation. Post-processing was applied at some sites based on inspection of the flow distributions. The performance of the distribution of annual flow is assessed across 88 sites in terms of the error (relative difference between observed and simulated). The error of the mean is –0.8% (2.3% when using absolute values of difference) and of the standard deviation is –1.35% (4.5% for absolute values of difference). Additional analyses were conducted for the daily distribution of flow and for the minimum annual flow for consecutive years.

1 Introduction

The New South Wales Government has committed to delivering regional water strategies for the state that reflect the influence of climatic variability on water resources. To this end, 10,000-year time series of stochastic data have been generated for 370 rainfall, 414 evapotranspiration and 97 temperature sites. The generated data assume climatic stationarity and mimic interannual and interdecadal oscillations in the historical record according to the Interdecadal Pacific Oscillation (IPO). The generated data do not account for projections of future climate.

This document summarises the development of stochastic data for the southern basin region, including evaluation of subsequently generated streamflow for selected catchments. The sub-regions of the southern basin region are labelled Upper Murrumbidgee, Murrumbidgee, Snowy, Murray, Victoria and South Australia, although the boundaries do not affect any aspect of the model calibration and simulation. Some sites sit outside the regional boundary (for example, Broken Hill and the Adelaide Plains), but are associated with this dataset for purposes determined by the needs of existing hydrological models. This document follows a similar methodology to data generated for previous regions; therefore, the methodology report (Leonard et al. 2020) is useful to provide background information on the modelling framework. The methodology report is also parallel to a pilot study of climatic influences in the southern region (Devanand et al. 2020) and a precursor to future reports on amalgamating stochastic replicates with information based on NARCliM climate projections.

The generated time series will characterise key statistical properties of rainfall, evapotranspiration and temperature as necessary for hydrological response and water planning. To statistically evaluate the performance of the model, the 10,000-year simulations are partitioned according to 77 replicates of 129 years to match the length of the historical record for direct comparison of statistical quantities. Performance summaries are provided for metrics (described below) using a traffic light system (Good, Fair, Poor) based on predefined rules rather than ad hoc assessment. Performance summaries are provided for the following metrics:

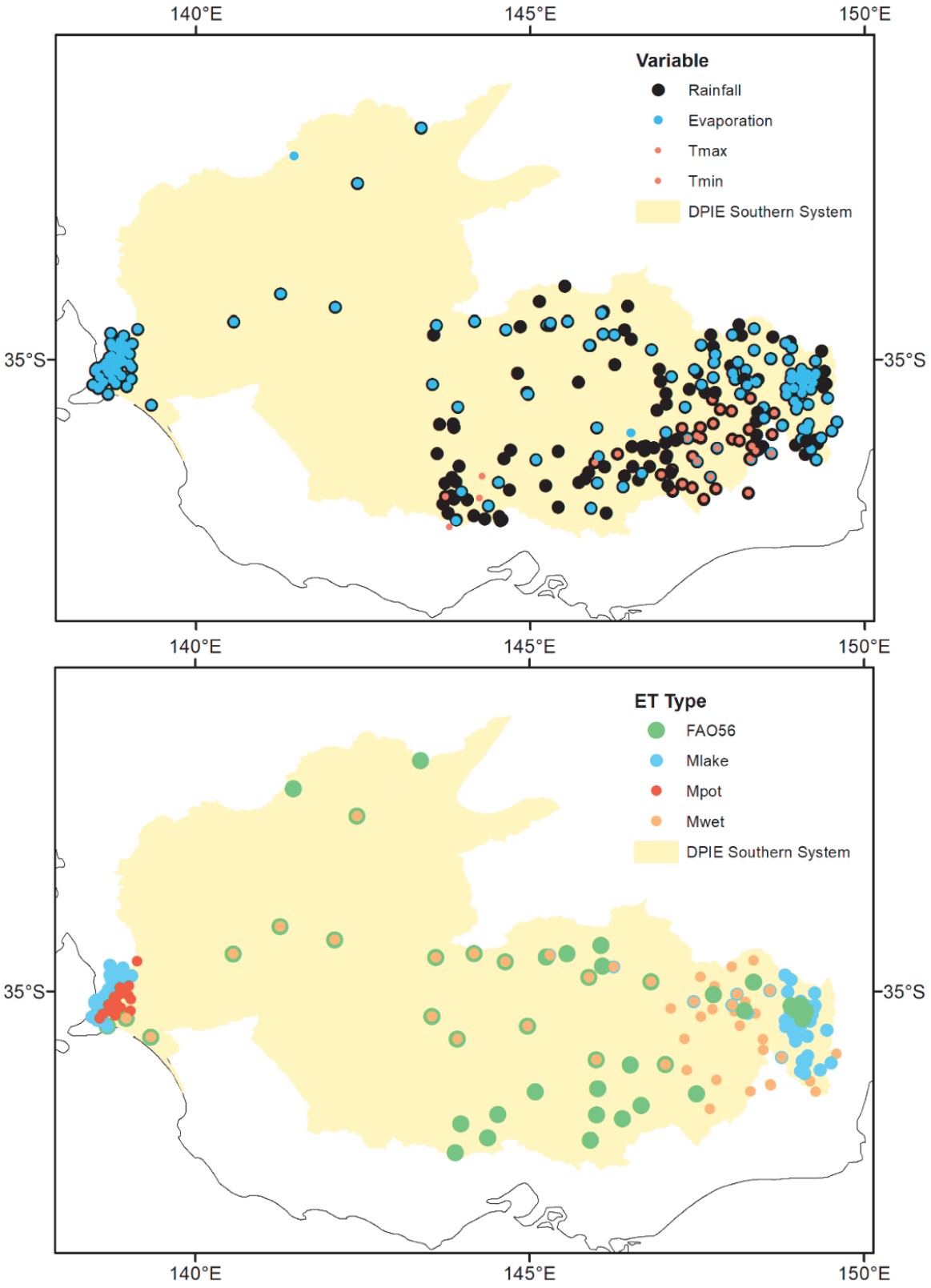
- 1-year, 2-year, 5-year and 10-year totals to assess the variability of rainfall depth, evapotranspiration and temperature over multiple timescales from interannual to decadal,
- monthly mean and standard deviation of rainfall depth and evapotranspiration to assess seasonal performance and intra-annual variability,
- monthly mean and standard deviation of the proportion of wet days to assess the number and distribution of rain days,
- 1-day, 2-day and 3-day annual rainfall maximums to assess the intensity, frequency and duration of extreme rainfall events,

Additional plots are generated for analysis and discussion, in particular the flow response for catchments in the southern region. Whereas other regions were post-processed to ensure exact matching of the monthly rainfall and evapotranspiration means, discrepancies in simulated flow are nonetheless possible. Discrepancies can occur because the climatic data has fluxes across every timescale – from daily to decadal – and while there are myriad statistics for checking, they are not exhaustive and cannot always identify the essential features that determine the streamflow response in a specific catchment. The rainfall-runoff process is complex and nonlinear, so unbiased climatic inputs do not guarantee that corresponding streamflow will be unbiased. Given the priority is to have matching streamflow distributions, post-processing was applied to the

climatic data to give a better match to streamflow. This post-processing method is simpler than the method used in the western region (Leonard et al. 2021), which sought a rule-based method for 'improving' rainfall that was ignorant of any streamflow performance. The approach adopted here effectively calibrates the input rainfall quantiles against the streamflow quantiles. This document summarises a range of performance metrics associated with the generated time series, using metrics consistent with those used to evaluate other regions.

2 Analysis of observed data

Time series of 10,000 years have been generated for the southern region river catchments. There are 881 time series in total for the southern region river catchments, made up of 370 rainfall time series, 414 evapotranspiration time series and 97 temperature time series at the locations shown in Figure 1 (top panel). The sites are concentrated around the upper reaches of the Murray and Murrumbidgee, with an additional cluster in Adelaide. The lower panel of Figure 1 shows the composition of different types of evapotranspiration data. The tag 'Mwet' refers to the Morton wet formulation, 'Mpot' refers to the Morton potential evapotranspiration formulation, 'FAO56' refers to the reference crop formulation and 'Mlake' refers to the Morton lake evapotranspiration formulation.



Multiple time series are co-located at some sites.

Figure 1 Location of rainfall, evapotranspiration and temperature sites (top panel) and break-down of evapotranspiration sites (bottom panel)

Table 2 provides a summary of the data used in the southern region. As with other regions, the majority of the data were sourced from the SILO database, with the addition of rainfall and potential evapotranspiration data used within IQQM modelling available at a number of sites. The tag 'IQQM' refers to the data from IQQM modelling. The temperature data are paired at most sites as daily time series of the minimum and maximum temperatures. All sites spanned 129 years, from 1 January 1890 to 31 December 2018, and there were no 'missing' values (owing to predetermined infilling methods used to construct the input data), though extended periods of infilling with median values are evident with evapotranspiration data at some sites.

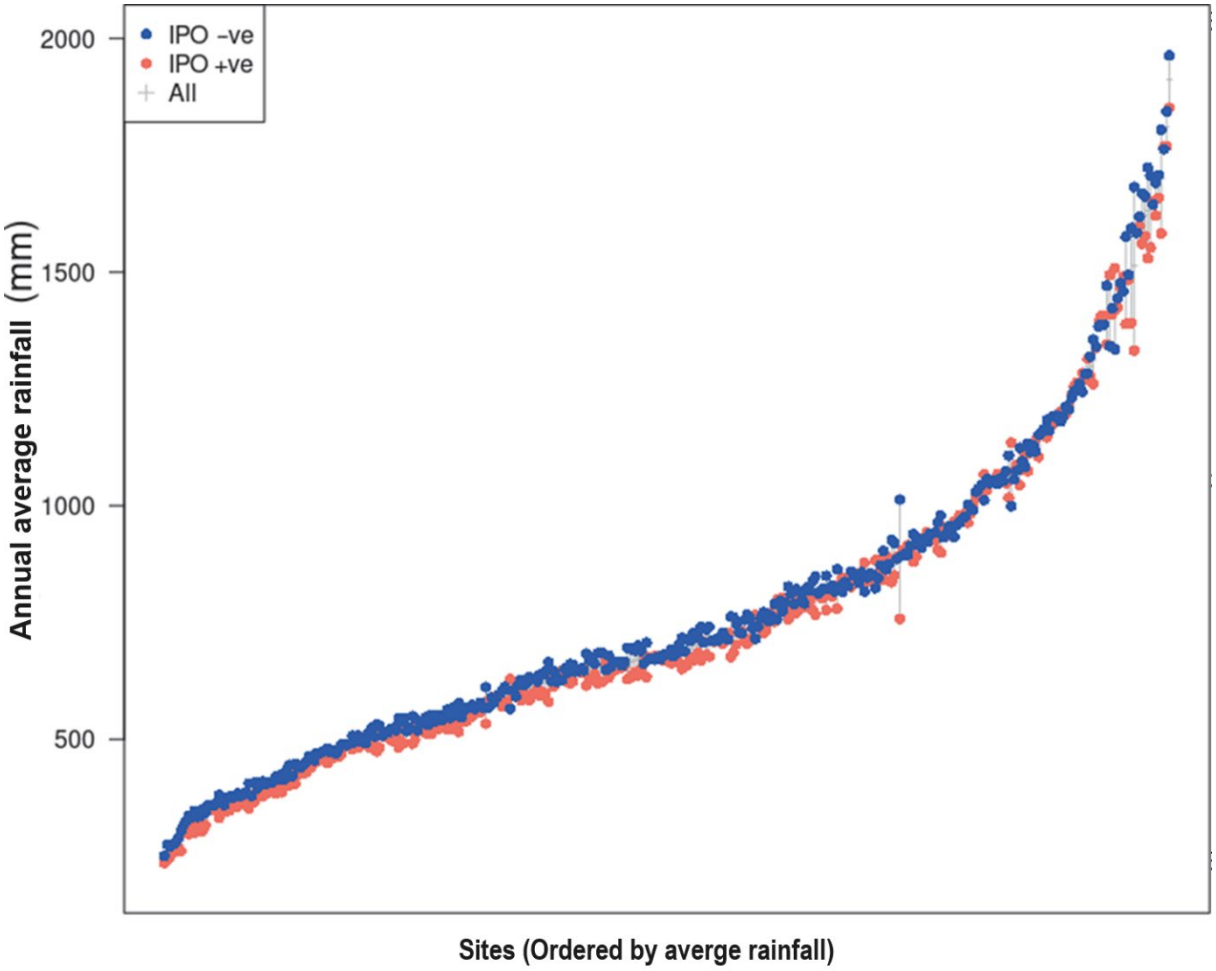
Table 2 Number of observation time series by region and variable type

Variable type	Snowy	Murrumbidgee	Upper Murrumbidgee	Murray	SA	Victoria
SILO_Rain	33	59	55	74	46	89
IQQM_Rain	0	0	14	0	0	0
SILO_Mwet	10	44	32	29	33	12
SILO_Mlake	8	40	21	25	33	11
SILO_FAO56	0	12	19	50	0	11
SILO_Mpot	0	0	0	0	17	0
IQQM_PET	0	0	7	0	0	0
SILO_Tmax	20	0	0	25	0	7
SILO_Tmin	20	0	0	25	0	0
Total rainfall	33	59	69	74	46	89
Total evapotranspiration	18	96	79	104	83	34
Total temperature	40	0	0	50	0	7

Calibration and simulation of stochastic data in previous regions used the IPO to partition the data into positive and negative states. Separate models are fitted to each state and then combined into a final time series. Table 3 summarises the difference this partitioning makes for rainfall in prior modelled regions in New South Wales alongside the southern region. It is clear from Table 3 that the influence of the IPO is much lower in the southern region. The same methodology is nonetheless retained; it simply means that the parameters fitted to each state will converge towards similar values when there is negligible difference between the states and the effect of the IPO diminishes accordingly. Figure 2 visualises the difference in IPO state at each rainfall site. Similar analyses were conducted for evapotranspiration and temperature data but differences were negligible.

Table 3 Summary of average difference in annual rainfall between IPO states and corresponding percentage difference relative to the mean annual rainfall

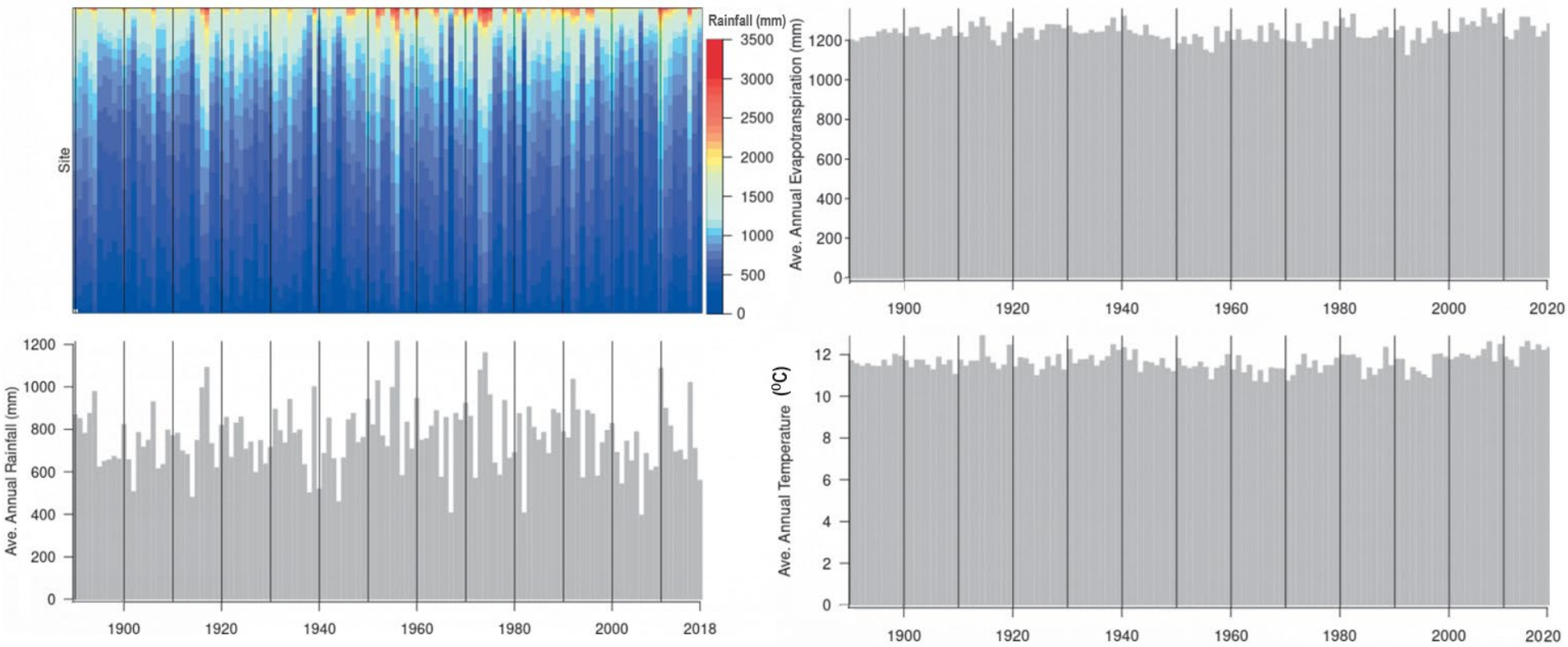
Region	mm	% difference
Far North Coast	135	8
North Coast	143	9
Border	47	7
Gwydir	66	9
Namoi	74	11
Macquarie	75	11
Western	55	9
Southern	32	3



Each vertical bar represents 1 site. The IPO negative state (blue symbols) coincides with ~30 mm more rainfall per year than the IPO positive state (red symbols).

Figure 2 Mean annual rainfall at sites in the southern region reported by IPO phase

The left panel of Figure 3 shows the time series of annual rainfall per site (top row) and average per year (bottom row). As noted for other regions, rainfall exhibits a highly variable characteristic from year to year. The rainfall sites are coloured in ascending order from average driest to average wettest site and show high variability. For example, 1975 was a wet year with majority of the sites (>85%) receiving more than 1000 mm (that is, the 'turquoise' to 'red' colouring), while in some other years, such as 2006, most sites (>85%) received less than 500 mm ('blue' colouring). The right panel of Figure 3 shows average annual evapotranspiration (top) and average annual temperature (bottom). These climatic variables are shown for completeness, and it is evident they are far more consistent from year to year than rainfall.



The top-left panel shows the annual total for each site where the sites have been sorted by their long-term average so that the wettest site appears at the top of the figure and the driest site appears at the bottom. The other 3 panels show a bar plot of the arithmetic average across all sites per year for rainfall, evapotranspiration and temperature, respectively. Black lines show 10-year periods.

Figure 3 Time series of annual total rainfall (left), evapotranspiration (top-right) and temperature (bottom-right)

Figure 4 provides one method for comparing recent conditions to previous conditions from the historical record, by marking the lowest (red), second lowest (blue) and third lowest (green) rainfall totals on record for 1-year, 2-year, 5-year and 10-year periods. From this figure, the 1997–2009 Millennium Drought is the driest period on record for this region, also containing the driest 1-, 2- and 5-year periods on record.

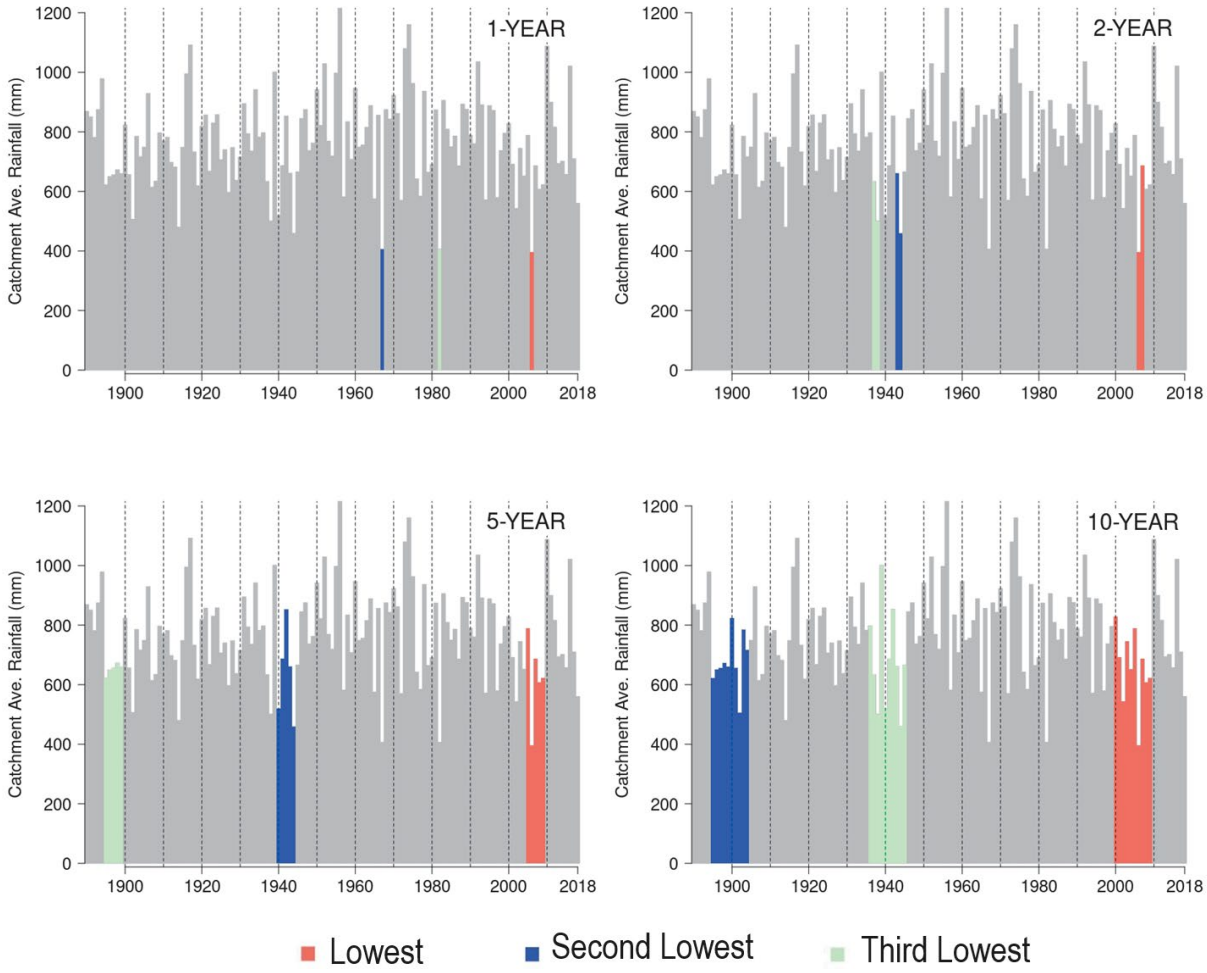


Figure 4 Summary of (multi-) annual rainfall totals for lowest periods on record for 1-year, 2-year, 5-year and 10-year periods

Figure 5 shows the spatial gradient of annual average rainfall, evapotranspiration and temperature across the catchment. The gradient of the Great Dividing Range is clear with the highest rainfall totals, lowest evapotranspiration and lowest temperatures in the Snowy Mountains region, with lower rainfall, higher evapotranspiration and higher temperature moving through inland Victoria. The difference in annual totals is a factor of three for rainfall and a factor of two for evapotranspiration and temperature respectively. The Adelaide region shows a mild gradient due to the Mt Lofty Ranges, with lower rainfall and higher evapotranspiration on the eastern side of the ranges that drain into the Murray River.

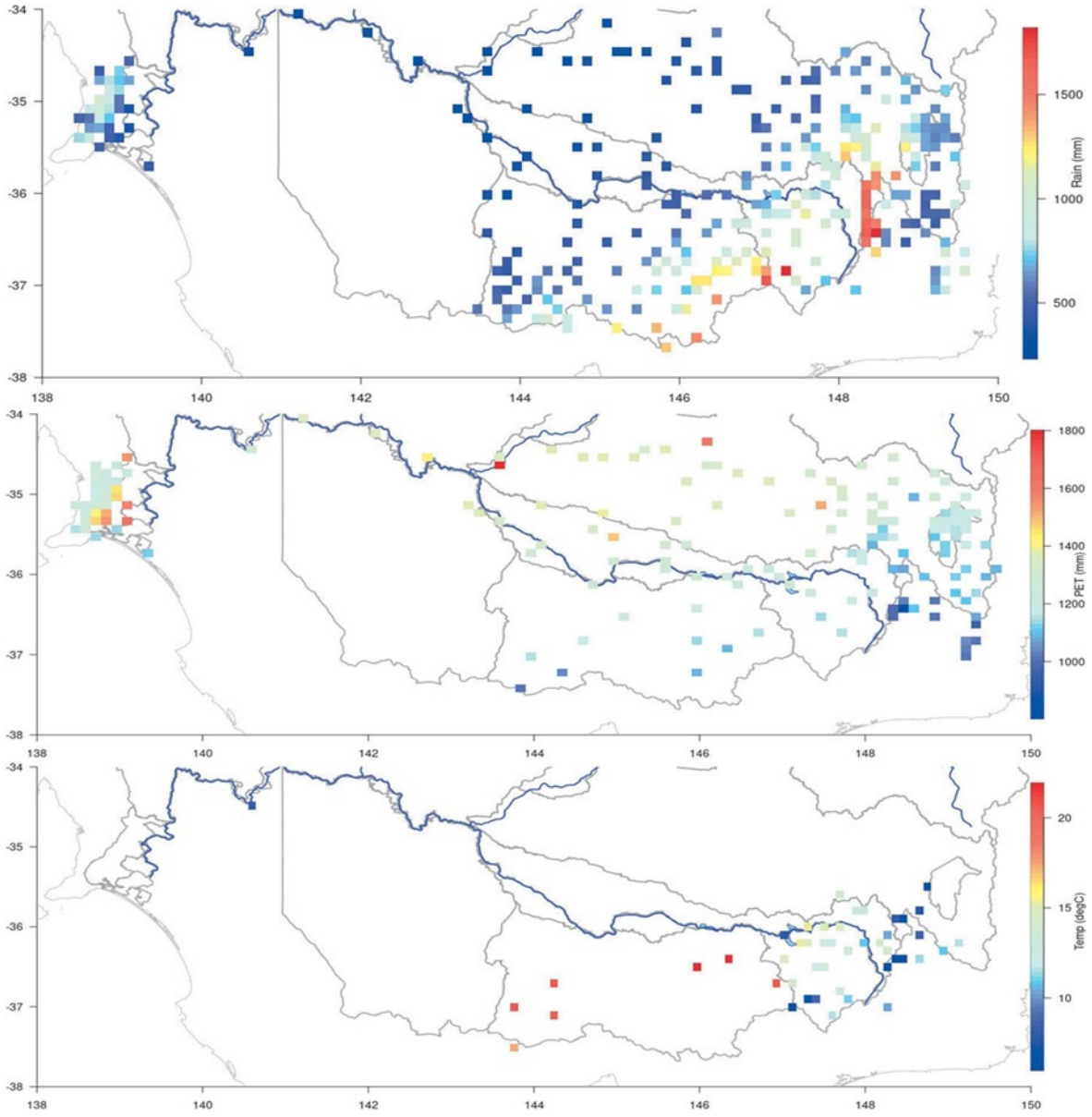


Figure 5 Spatial gradient of annual average totals of rainfall (top), potential evapotranspiration (middle) and temperature (bottom)

3 Evaluation of stochastic data

Sections 3.1 to 3.9 provide traffic light performance summaries aggregated across all sites in the region. Representative plots and discussion are provided for each statistic. Section 3.10 discusses the performance of streamflow generated from the stochastic time series for selected catchments.

3.1 Distribution of (multi-) annual rainfall totals

Table 4 shows the summary performance for 1-year, 2-year, 5-year and 10-year rainfall totals across the 370 sites in the southern basin region. The performance is generally Good. Figure 6 shows an example of a 'Fair/Poor' site for 5-year and 10-year totals, which is the worst performing site in the simulation. These metrics are labelled with worse performance primarily due to the discrepancy between 0.2 and 0.5 quantiles, but it can be seen there is overall less variability in the simulations than the observations, including in the upper tail.

Table 4 Performance summary of (multi-) annual rainfall totals for 370 rainfall sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Annual total rain 1 year	Good	100	0	0
Annual total rain 2 year	Good	99	1	0
Annual total rain 5 year	Good	95	4	1
Annual total rain 10 year	Good	92	5	3

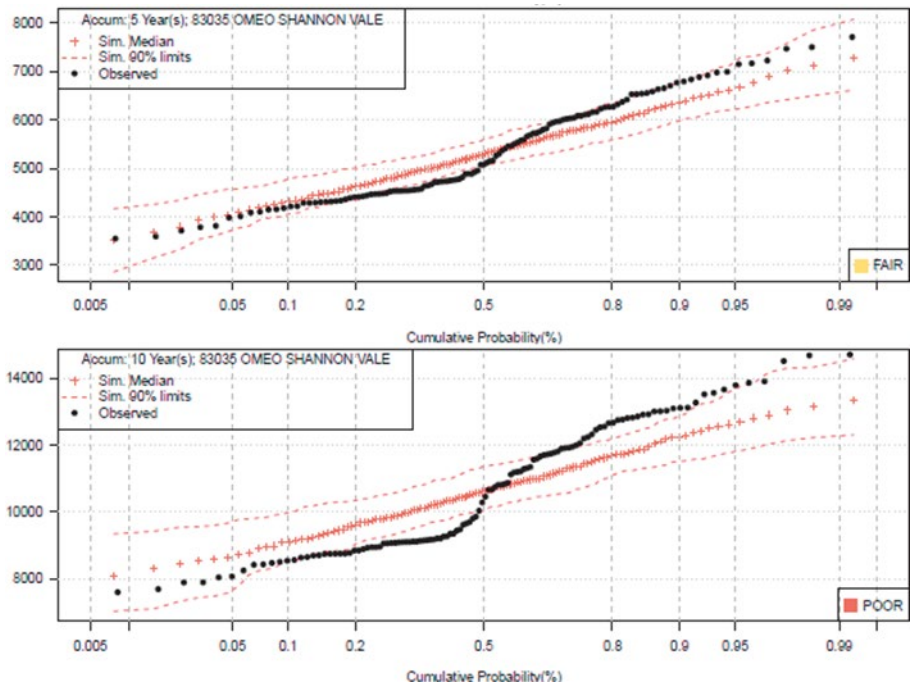


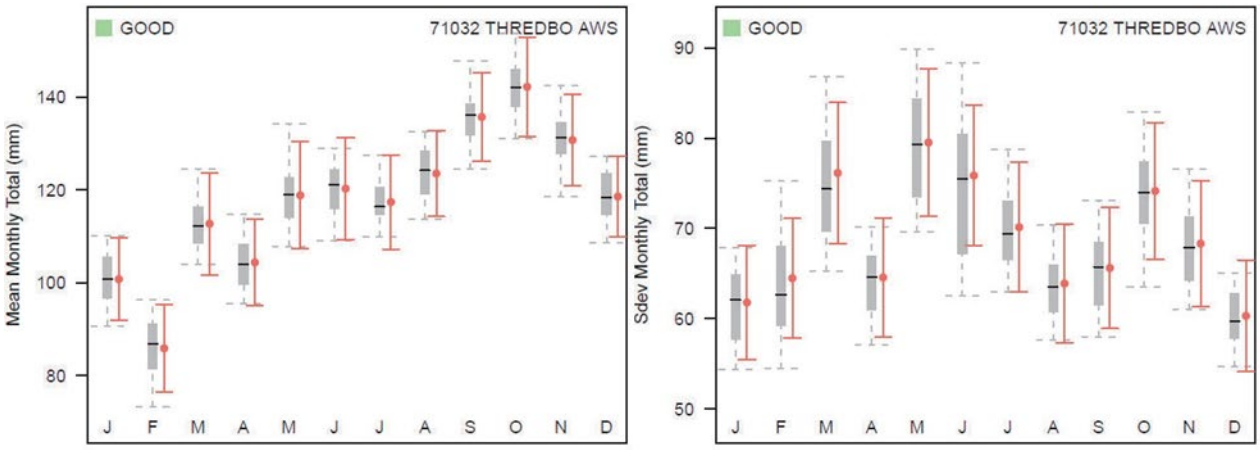
Figure 6 Distribution of multiannual total rainfall at the site showing the worst performance in the region – Fair/Poor performance for 5-year and 10-year totals

3.2 Distribution of monthly rainfall totals

Table 5 shows the summary performance for the mean and standard deviation of monthly rainfall totals across the 370 sites in the southern region. The distribution of monthly totals matches well in terms of mean and standard deviation. Figure 7 shows a representative location.

Table 5 Performance summary of the mean and standard deviation of monthly rainfall totals in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Mean monthly rain totals	Good	100	0	0
Standard deviation monthly rain totals	Good	100	0	0



Simulated in grey, observed in red.

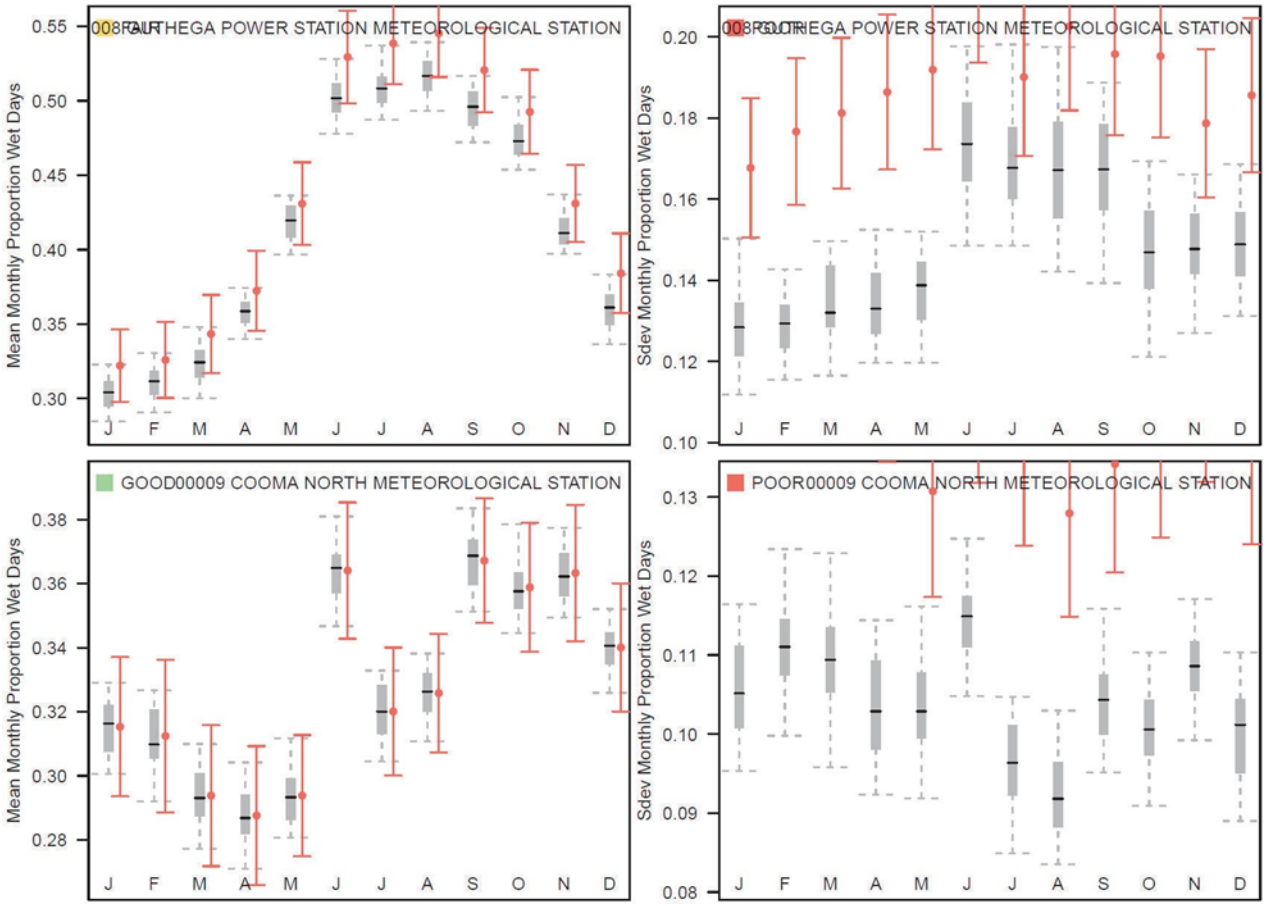
Figure 7 Distribution of mean and standard deviation of monthly total rainfall at a site showing Good performance

3.3 Distribution of proportion of wet days

Table 6 shows the summary performance for the mean and standard deviation of the proportion of wet days for each month across the 370 rainfall sites in the southern region. Consistent with the results from other regions, this statistic does not perform as well as others, and is difficult to control in the model. The means show Good performance at all but one site (shown in Figure 8). The standard deviation is Poor at all sites, as with previous regions (Leonard et al. 2019, 2020, 2021). Figure 8 demonstrates the Poor performance for monthly standard deviation of wet day proportions, which is underestimated by the simulation at all sites. For a month of 30 days, a discrepancy of 3–4% amounts to about 1 day per month. Given that the annual and monthly totals are generally well matched, a corollary of the slight underestimation in the variability of the number of wet days is a slight overestimation in the variability of rainfall amounts per wet day.

Table 6 Performance summary of the mean and standard deviation of monthly proportion of wet days in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Mean monthly wet proportion	Good	99	1	0
Standard deviation monthly wet proportion	Poor	0	0	100



Simulated in grey, observed in red.

Figure 8 Distribution of mean (left) and standard deviation (right) of proportion of wet days for each month at the worst performing site (top row) and a representative site (bottom row)

3.4 Distribution of rainfall extremes

Table 7 shows the summary performance for the 1-day, 2-day and 3-day annual maximums, showing Overall Fair performance. Figure 9 shows four sites with a range of performance classifications. Inspecting these sites, the range of classifications is largely due to characteristically tight simulation bounds in the lower tail with wider simulation bounds in the upper tail.

Table 7 Performance summary of the 1-day, 2-day and 3-day annual maximums for 370 rainfall sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
1-day rainfall extremes	Fair	36	64	0
2-day rainfall extremes	Fair	44	56	0
3-day rainfall extremes	Fair	36	63	1

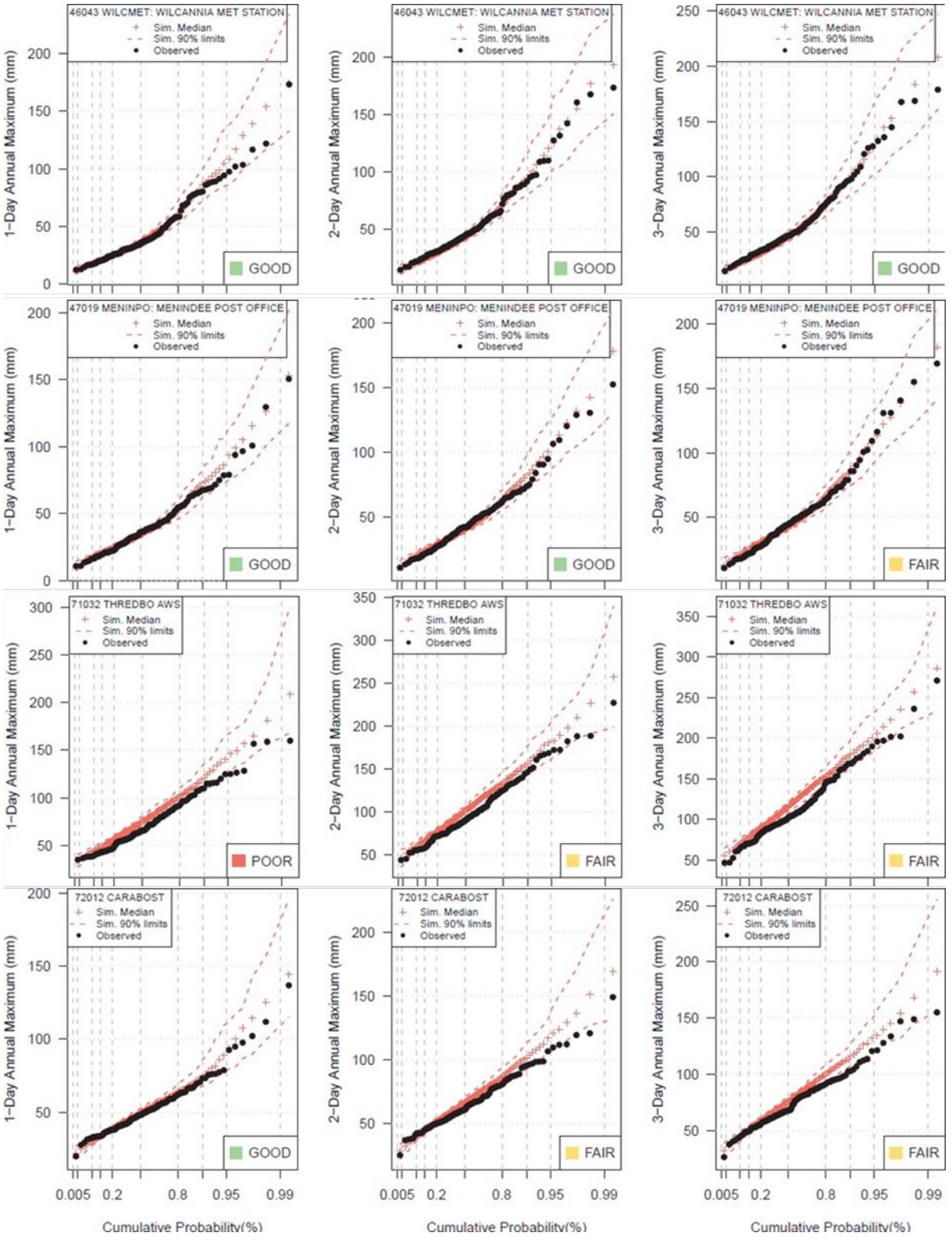


Figure 9 Distribution of 1-day, 2-day and 3-day annual maximums for rainfall at 4 representative sites in the region

3.5 Distribution of (multi-) annual evapotranspiration totals

Table 8 shows the summary performance for 1-year, 2-year, 5-year and 10-year evapotranspiration totals across the 276 Mwet, Mpot, FAO56 and IQQM evapotranspiration sites in the southern region. The Mlake time series are calculated from the simulated Mwet time series via a trivial regression, and hence are not included in this summary. The Fair and Poor performance, particularly at longer timescales, is due to an artefact in the observed data, in which extended periods have been infilled with median values. Figure 10 presents an example of the impact of the infilling, showing Fair performance for 5-year totals and Poor performance for 10-year totals.

Table 8 Performance summary of annual and multi-year evapotranspiration totals for 276 evapotranspiration sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Annual total evapotranspiration 1 year	Good	94	6	0
Annual total evapotranspiration 2 year	Good	93	7	0
Annual total evapotranspiration 5 year	Fair	42	53	5
Annual total evapotranspiration 10 year	Fair	9	50	41

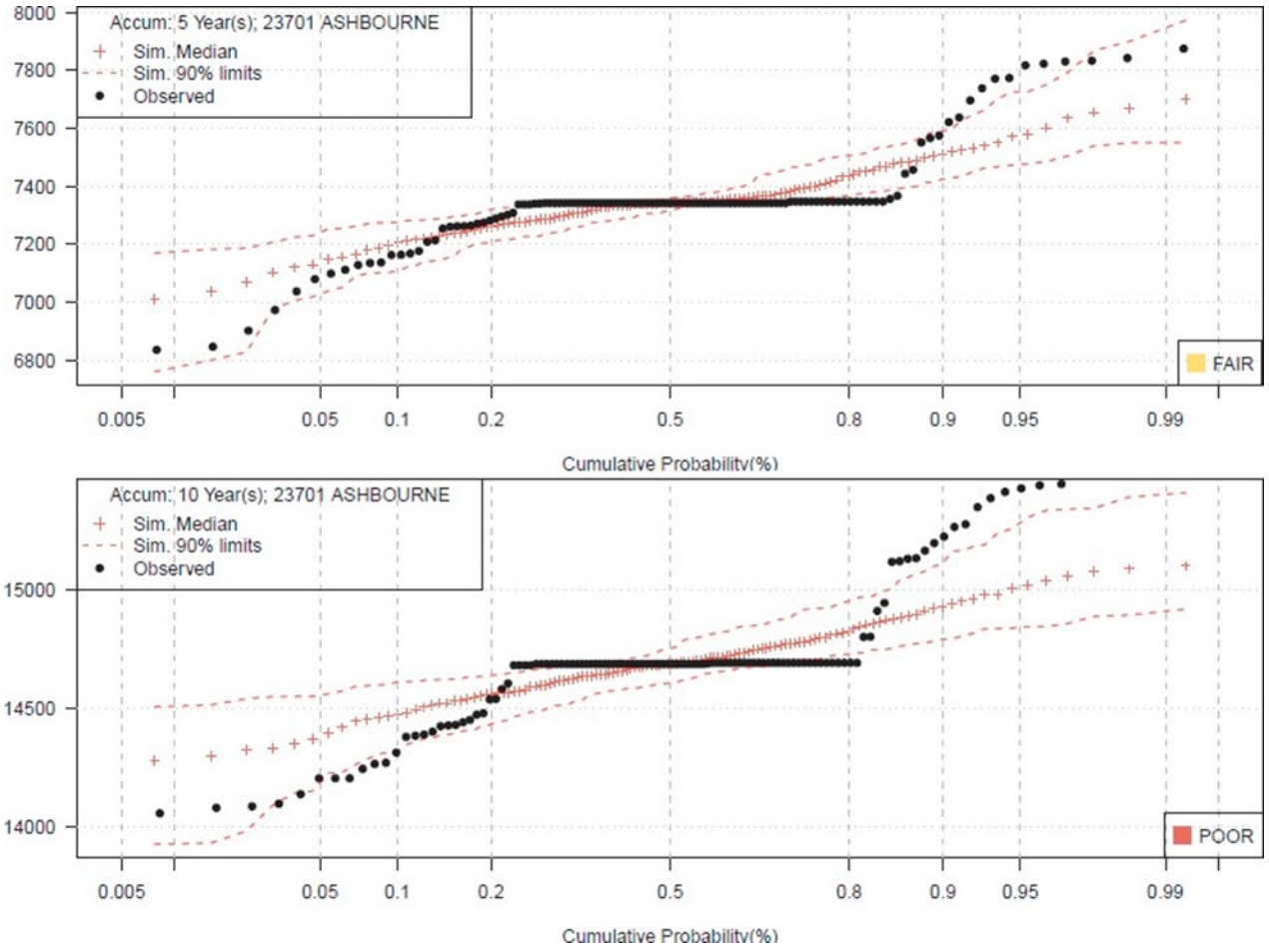


Figure 10 Distribution of annual total evapotranspiration at one of the worst performing sites for all durations

3.6 Distribution of monthly evapotranspiration totals

Table 9 shows the summary performance for the mean and standard deviation of monthly evapotranspiration totals across the 276 Mwet, Mpot, FAO56 and IQQM evapotranspiration sites in the southern basin region. Good performance is expected for this statistic. Figure 11 shows the performance at a typical simulated site. Figure 11 (left panel) shows that the distribution of the

mean monthly total is very narrow for each month in both the observed and simulated records. Figure 11 (right panel) shows that the standard deviation of monthly totals is matched well.

Table 9 Performance summary of the mean and standard deviation of monthly evapotranspiration totals for 276 (Mwet, Mpot, FAO56 & IQQM) sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Mean monthly evapotranspiration totals	Good	100	0	0
Standard deviation monthly evapotranspiration totals	Good	100	0	0

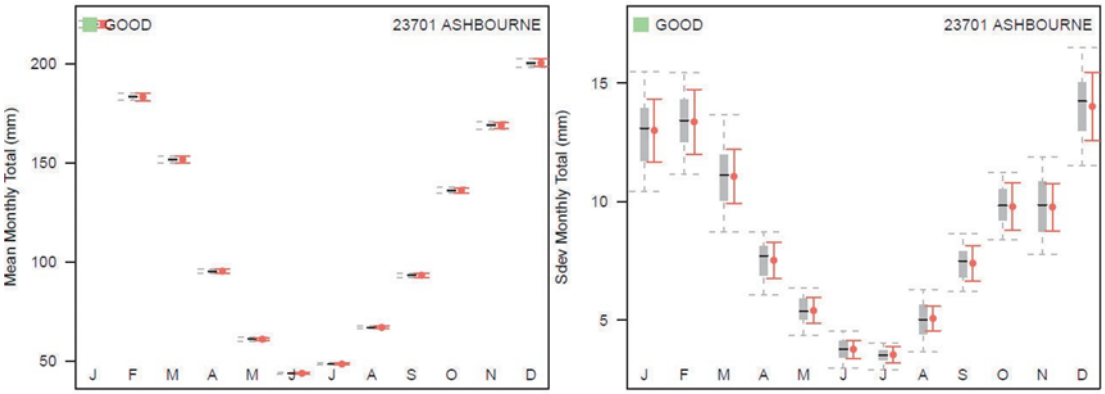


Figure 11 Distribution of mean and standard deviation of monthly total evapotranspiration at a representative site

3.7 Distribution of (multi-) annual temperature means

Table 10 shows the summary performance for 1-year, 2-year, 5-year and 10-year annual mean temperatures across the 97 Tmin and Tmax sites in the southern basin region. While there is Overall Good performance at the 1-year and 2-year scale, at longer timescales the simulations do not show enough variability. Figure 12 shows four durations for one of the worst performing sites in the region, where the variability is difficult to match. There are several aspects to consider in relation to this site:

- The site has a significant drop between the 0.2 and 0.5 quantiles that is difficult for the model to match.
- Post-processing steps at an annual scale have limited the variability of the data, which has affected the longer timescale averages.
- There is variability at longer timescales of the observed data that is not able to be captured well by the model.

Table 10 Performance summary of (multi-) annual temperature means for 97 temperature sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Annual mean temperature 1 year	Good	98	2	0
Annual mean temperature 2 year	Good	85	15	0
Annual mean temperature 5 year	Fair	24	55	22
Annual mean temperature 10 year	Poor	1	33	66

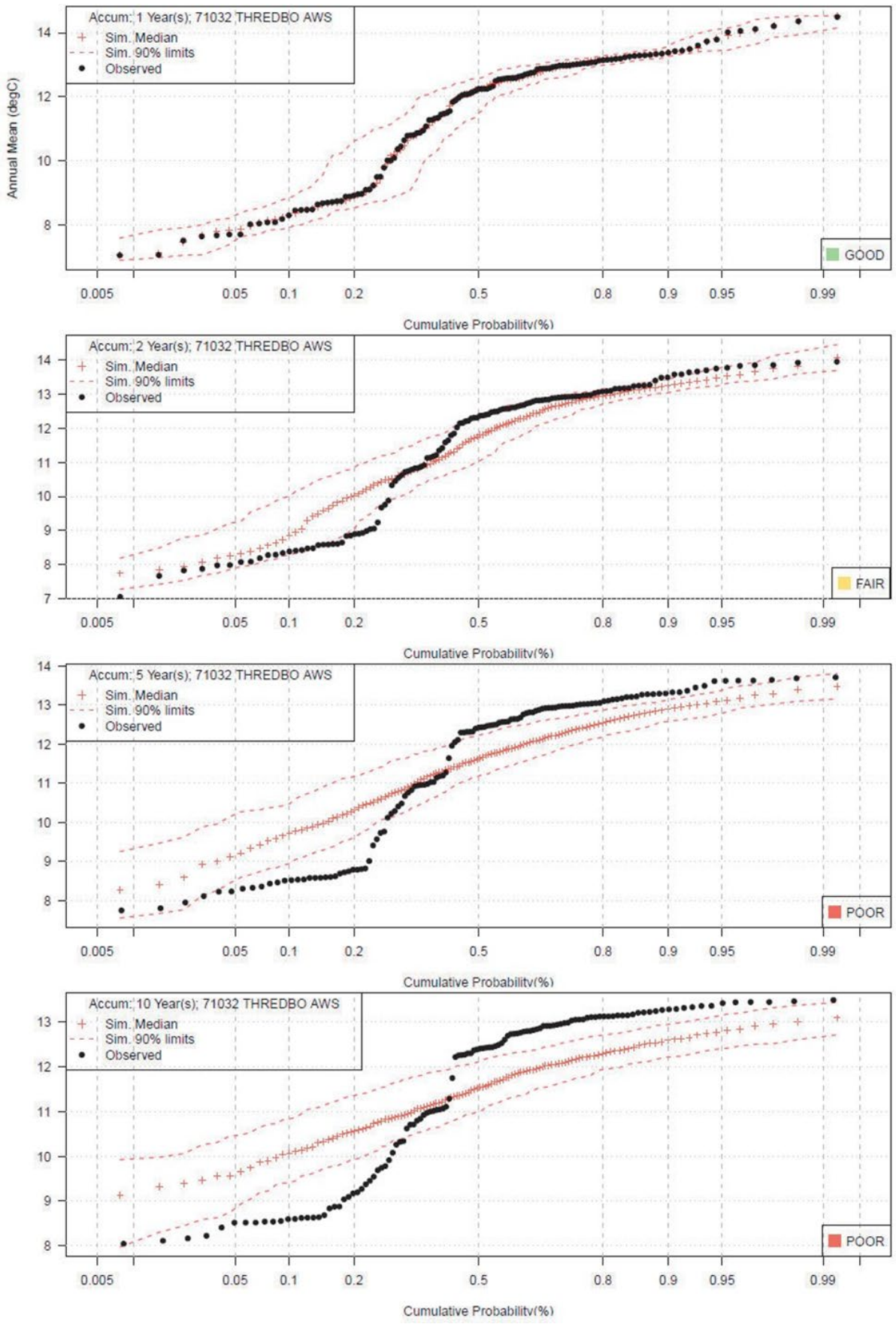


Figure 12 Distribution of annual temperature means at one of the worst performing sites for all durations

3.8 Distribution of monthly temperature means

Table 11 shows the summary performance for the mean and standard deviation of monthly temperature means across the 97 Tmin and Tmax sites in the southern region. Given the monthly totals are post-processed to have an identical mean, a good match is expected for this statistic. Figure 13 shows the performance at a typical simulated site. Figure 13 (left panel) shows that the distribution of the mean monthly mean is very narrow for each month in both the observed and simulated records. Figure 13 (right panel) shows that the standard deviation of monthly means is matched well, except for February.

Table 11 Performance summary of the mean and standard deviation of monthly means for 97 temperature sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
Mean monthly temperature mean	Good	99	1	0
Standard deviation monthly temperature mean	Good	96	4	0

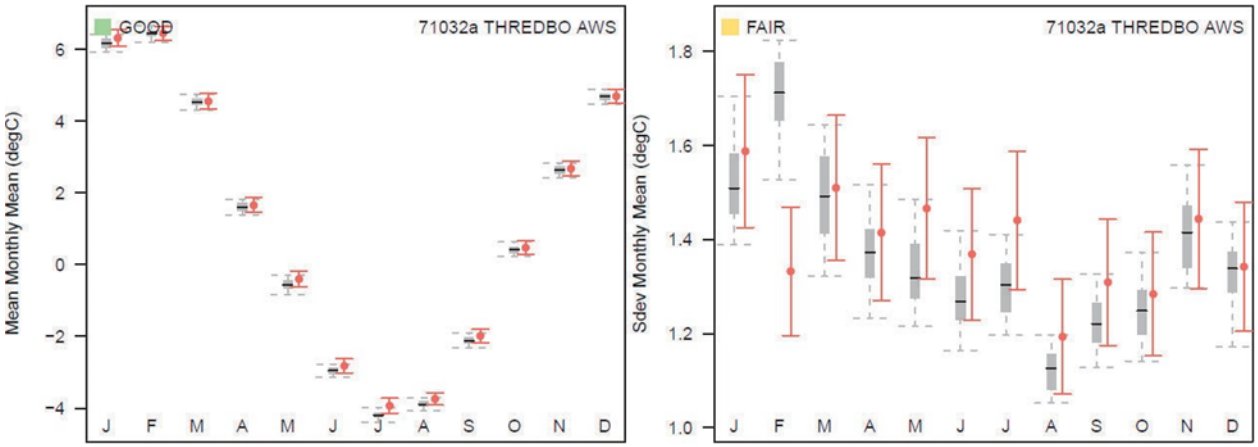


Figure 13 Distribution of mean and standard deviation of monthly temperature means at a representative site

3.9 Distribution of temperature extremes

Table 12 shows the summary performance for the 1-day, 2-day and 3-day annual maximums of the maximum temperature (minimum temperatures are out of scope), showing Overall Poor performance of 1-day extremes but Overall Fair performance for the 2-day and 3-day extremes. Figure 14 shows four representative sites. Especially for 1-day extremes with Poor performance, the simulated extremes tend to overestimate the observations. This metric is not fitted as part of the calibration and it seems to be related to parameters in the month of February causing an over-simulation. This artefact can be remedied with quantile correction, but it has not been applied in the present dataset.

Table 12 Performance summary of the 1-day, 2-day and 3-day annual maximums for 97 temperature sites in the southern region

Statistic	Overall performance category	(%) Good performance	(%) Fair performance	(%) Poor performance
1-day temperature extremes	Poor	0	25	75
2-day temperature extremes	Fair	1	84	15
3-day temperature extremes	Fair	33	59	8

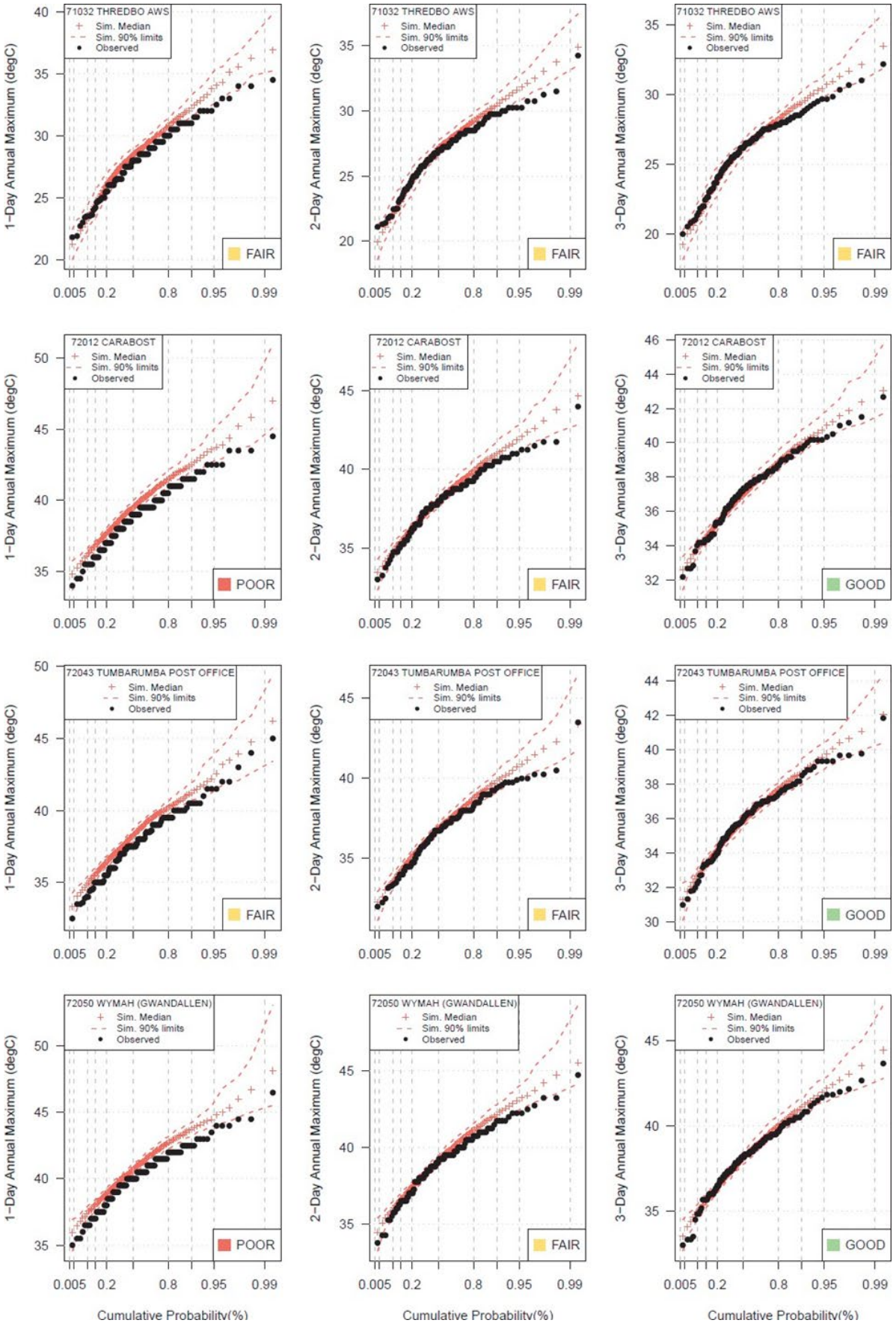
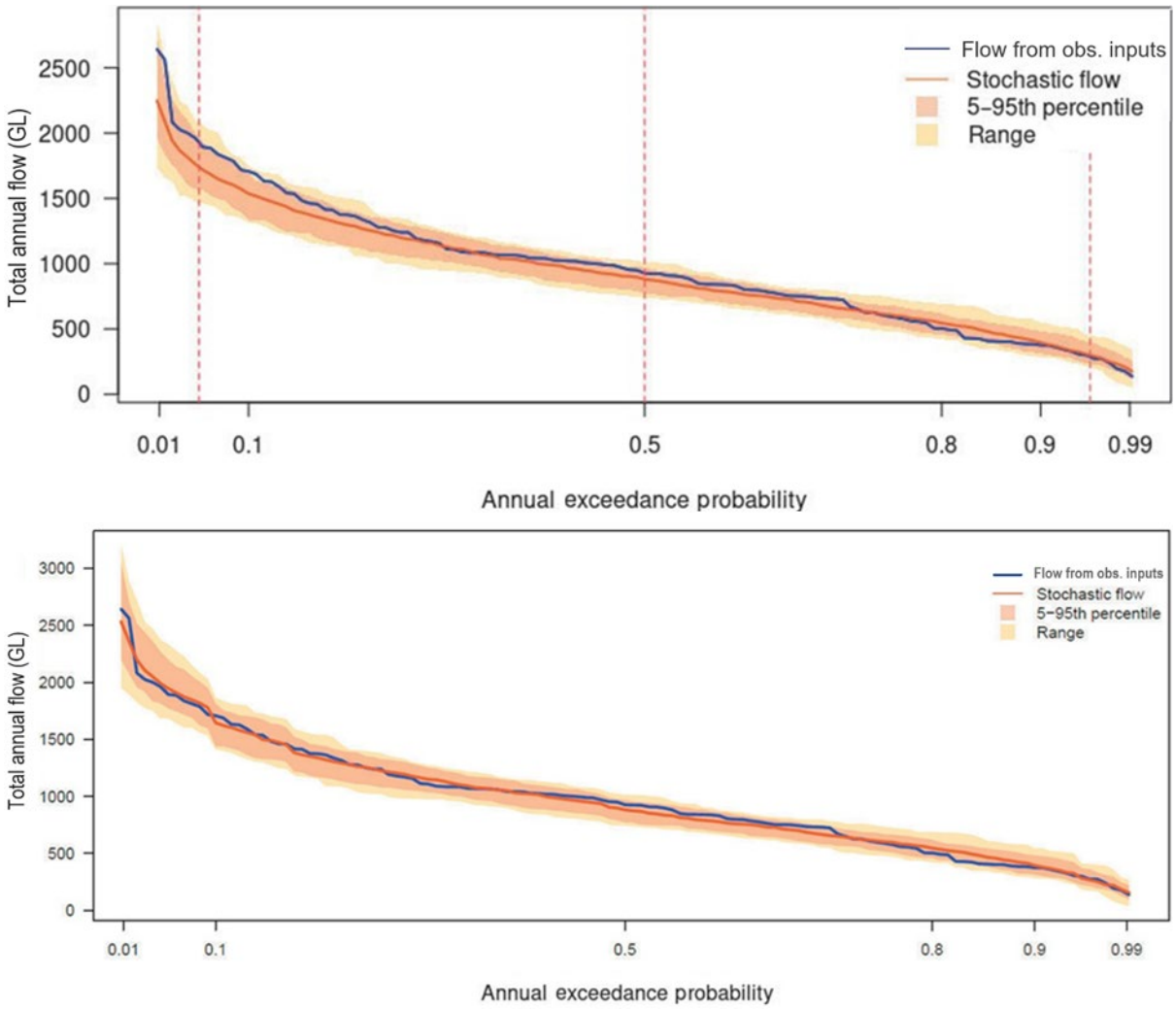


Figure 14 Distribution of 1-day, 2-day and 3-day annual maximums at 4 representative sites

3.10 Comparison to streamflow metrics

The department provided a set of 27 parameter files corresponding to Sacramento models of catchments in the southern region. The models were used as is for transforming climatic inputs to streamflow, yielding 88 output time series. While not all time series correspond to the main outflow node of a catchment, all reaches were used for the purposes of assessing streamflow performance. The hydrological models utilise 128 climatic time series as inputs, which is about 16% of all the rainfall and evapotranspiration sites. The observed climatic inputs and the simulated inputs were both transformed to streamflow to enable comparison of performance metrics closer to water management decisions. Having the hydrological model in common and not comparing to actual streamflow observations means there is a consistent comparison that does not introduce confounding factors.

Figure 15 provides an example where the flow time series has been generated for catchment 401211a for both the observed climatic inputs (blue line) and the simulated replicates (shaded orange) prior to streamflow correction. It is clear from the top panel of Figure 15 that there is a bias in the upper tail (and by corollary, in the mean and standard deviation), where the simulated streamflow is lower than the observed. The bottom panel of Figure 15 shows the simulated flows after post-processing. Importantly, the bias in the upper tail was not due to biases in the annual or monthly averages of rainfall or evapotranspiration inputs. Preliminary analyses have identified that it is due to a complicated relationship between the rainfall extremes and antecedent rainfall across multiple sites (and hence beyond a direct fix to the cause of the issue).



The top panel shows the original transformation having a significant discrepancy in the upper tail with the simulations consistently below the observed values despite being within the 90% limits. The bottom panel shows the shifted streamflows in the upper tail from post-processing.

Figure 15 Distribution of total annual streamflow transformed from climatic inputs for the 401211a catchment

Catchment models integrate fluxes of water and energy over time and can therefore pick up on features of rainfall/evapotranspiration not immediately evident when comparing climatic metrics. Discrepancies in streamflow indicate there are differences in the climatic inputs that are not otherwise apparent from the climatic metrics, and the differences have potentially been amplified by the catchment model. Whether or not the model provides realistic or appropriate amplification is beyond the scope of this project. The comparison to streamflow does not diagnose the attributes of the rainfall that cause the discrepancy.

Thus, a simple remedy was applied that seeks to provide unbiased streamflow without identifying the root cause of any discrepancy. The approach is to post-process rainfall by a factor proportional to the discrepancy in streamflow at that same quantile, which is not a perfect assumption given the considerable integration of the climatic inputs in time, but nonetheless effective given the coarse correspondence between high/low rainfall and high/low streamflow. This method is effective only for those catchments included in the calibration and may not hold for new catchments or catchment models. It is an area of research that requires attention, so that methods

can be developed that effectively use streamflow-based information in the development of stochastic models such that the end-of-system response has a good match.

Figure 16 and Figure 17 provide examples for a representative site of two other flow statistics used to assess streamflow performance. The statistics are, respectively, the daily flow duration curve and the minimum annual total flow from multiple consecutive years. These statistics were assessed visually but did not form part of the method of post-processing, which prioritised the distribution of annual streamflow.

Figure 18 provides a summary of the bias in the annual streamflow distribution at all 88 sites according to the mean, standard deviation, 5%, 50% (median), and 95% quantiles. The error of the mean averaged across all sites is –0.8% (2.3% when using absolute values) and of the standard deviation is –1.35% (4.5% for absolute values). The bias in the annual streamflow distribution is relatively low after applying the post-processing.

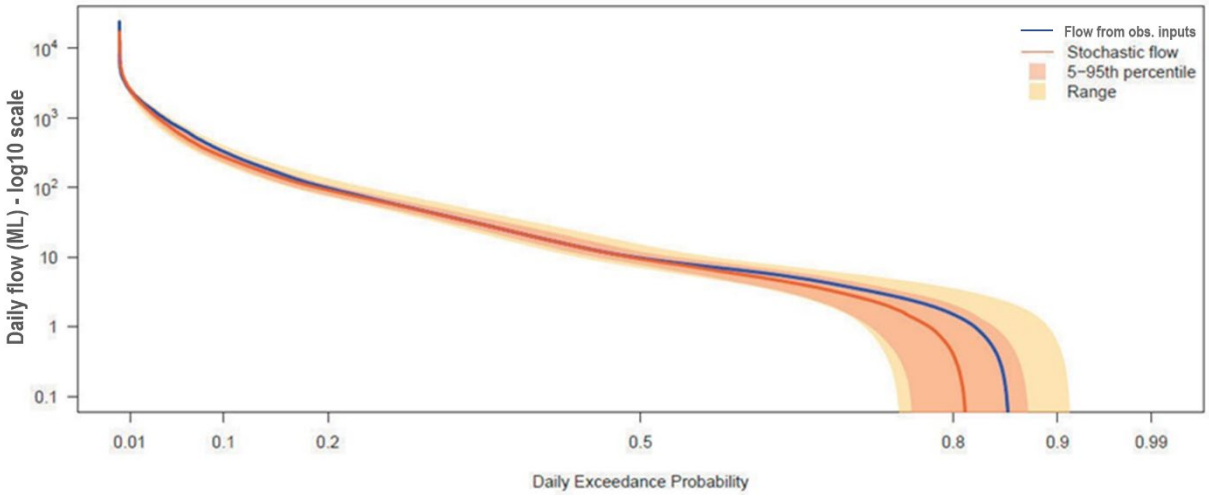


Figure 16 Distribution of the daily flow duration curve, site 405226

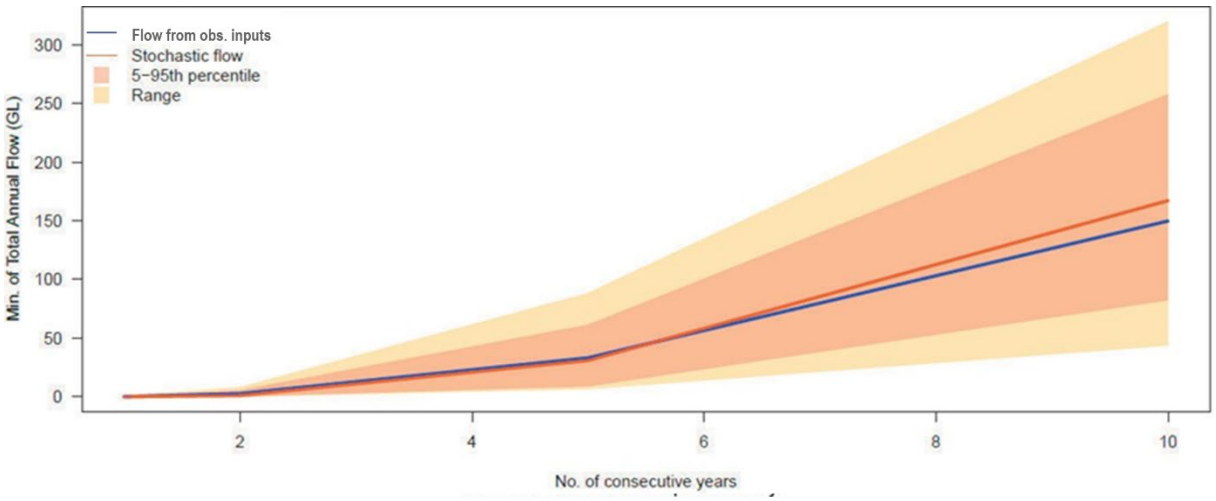
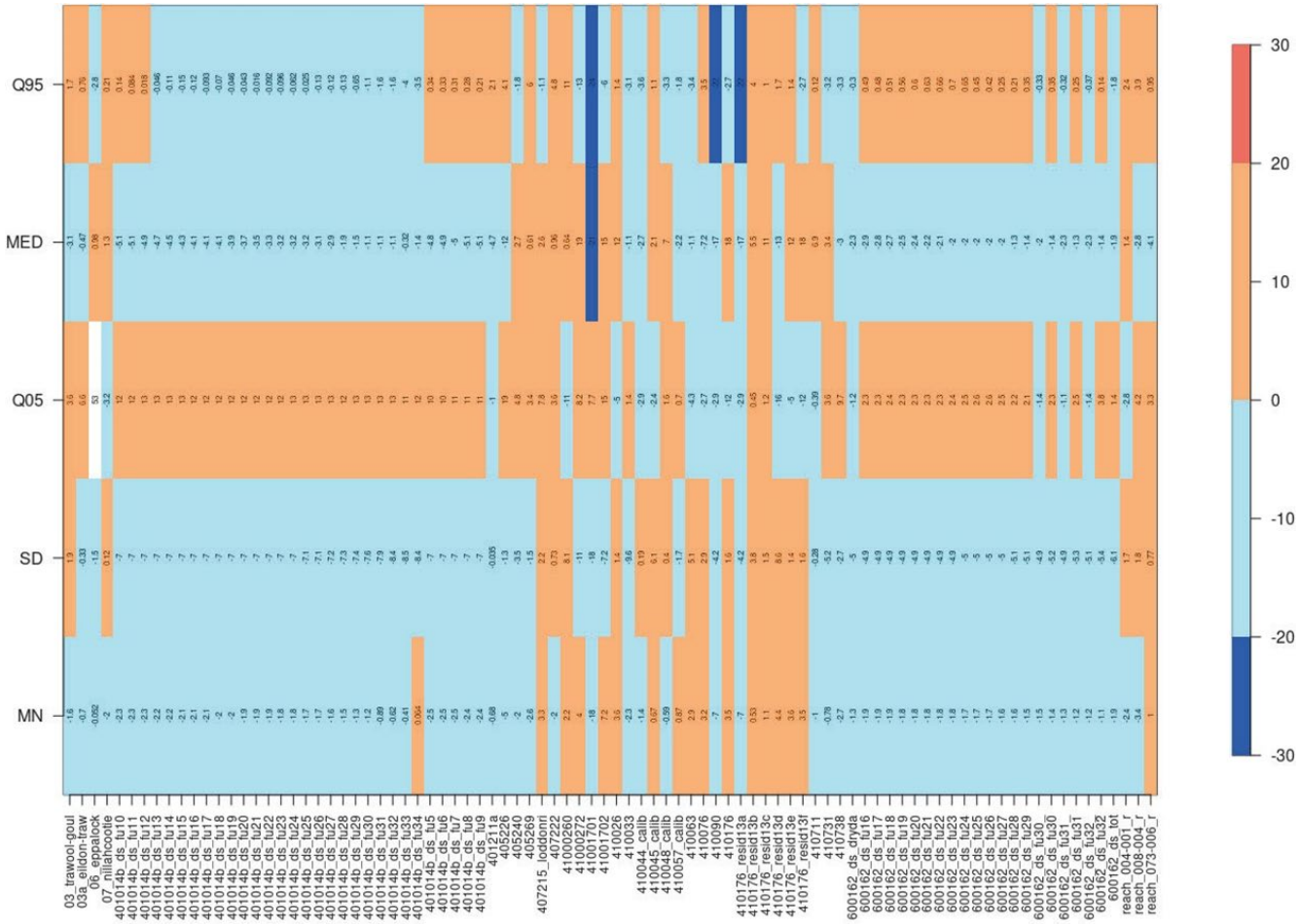


Figure 17 Distribution of the minimum of total annual streamflow across multiple consecutive years, site 405226



Error is 100 x (sim. – obs.) / obs. Units: %.

Figure 18 Error in selected metrics (mean, standard deviation and quantiles 0.05, 0.50 (median), 0.95) of the annual streamflow distribution at 88 model output locations

4 References

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.

Devanand A, Leonard M and Westra S (2020) *Implications of non-stationarity for stochastic time series generation in the southern basin*, Pilot study undertaken by Adelaide University.

Leonard M, Westra S and Bennett B (2019) *Multisite rainfall and evaporation data generation for the Macquarie water infrastructure project*, University of Adelaide.

Leonard M, Westra S and Bennett B (2020) *Methodology report for multisite rainfall and evapotranspiration data generation of the northern basin*, Report prepared for Department of Planning, Industry and Environment by the University of Adelaide.

Leonard M, Westra S and Bennett B (2021) *Methodology report for multisite rainfall and evapotranspiration data generation of the northern basin: Annex F western region*, Report prepared for Department of Planning, Industry and Environment by the University of Adelaide.