The following supplement accompanies the article

Future impacts of climate change on streamflows across Victoria, Australia: making use of statistical downscaling

S. Fiddes*, B. Timbal

*Corresponding author: sonya.fiddes@climate-energy-college.org

Climate Research 71: 219–236 (2016)

A. GCM bias correction

Previous studies on historical runs of the CMIP5 ensemble show that for rainfall, both the mean and variance over the Australian region are often underestimated (Irving et al. 2012). This dry bias is carried though via the SDM, due to sampling issues of low frequency-high rainfall events (Timbal et al. 2009).

Initial testing of the reconstructed historical streamflow data (before calibration of the downscaled CMIP5 data was performed) clearly shows this bias was passed onto the streamflow reconstructions. For this reason, a calibration step to remove systematic biases was required and applied initially for the 1950-2005 period, and then to the 2006-2100 period, assuming the relationship of the model outputs to the observations remains the same throughout the 21st century.

For the rainfall data, correction of both the monthly mean and the variance was needed. This was done as per the formula below, where P = precipitation, $\bar{P}$ = is the long-term mean precipitation (with a different value for each calendar month), σ = the standard deviation. Subscript c = the calibrated precipitation, 1 being the first stage, 2 the second, m = uncorrected model data, a = AWAP data.

(1) $$P_{c_1} = P_m \times \frac{\overline{P_a}}{\overline{P_m}}$$

(2) $$P_{c_2} = \left[P_{c_1} - \overline{P_{c_1}}\right] \times \frac{\sigma_a}{\sigma_m} + \overline{P_{c_1}}$$

This method “normalised” the monthly mean and annual cycle by multiplying the model timeseries by the ratio of the monthly AWAP to model mean (1). It then subtracts the new annual cycle from the new timeseries and multiplies the anomalies by the ratio of the AWAP to model standard deviation in order to correct the variance (2). Finally it adds the annual cycle back in. This was performed for each catchment and GCM individually, for each month of the year.

Due to very small mean biases for temperature, the calibration for the model temperature was simply to inflate the variance. Here, T = maximum temperature and $\bar{T}$ = monthly mean maximum temperature. The method used was essentially the same as correcting the precipitation variance, as seen by (3), where the anomalies were corrected without causing significant change to the mean.

(3) $$T_c = \left[T_m - \overline{T_m}\right] \times \frac{\sigma_a}{\sigma_m} + \overline{T_m}$$

The GCM data was then passed through the streamflow reconstruction method, and corrected for instances where streamflow was reported as negative (possible due to the simplicity of the model used). Here, negative streamflows were set to zero, and, to prevent artificially increasing the total streamflow, the remaining streamflow was then normalised against the observations, following the same method as that applied to rainfall above (equations 1 and 2).

The same methods were then applied to the future projection, using the existing ratios (observations to historical GCM runs) of the mean and standard deviation in both rainfall and temperature; therefore assuming that the biases between observations and GCM downscaled data remain constant in future climate simulations.

B. GCM Ranking

The final stage of the GCM evaluation was to rank the downscaled and calibrated models according to their performance in reconstructing the streamflow mean and variance to determine if some models were performing poorly. Downscaled models were ranked on the two measures, whereby the closer the average reconstructed values were to 100% the better the model performance. To combine these two into an overall ranking, the rank of the mean and variance was added up for each model and ordered from smallest to largest. These rankings can be seen in Table S1.

Figure 4a in main text indicates that no one model performs particularly badly and most show outliers overestimating the mean by up to 40%. A similar result is found for the variance (Figure 4b), although for a negative bias. Two models, CCSM4 and MPI-ESM-MR were ranked in the top 5 for both measures, whilst MIROC5 was in the bottom 5 for both (Table S1).

Previous evaluations of model performance in simulating weather and climate phenomena over Australia most pertinent to this study are provided in Table S1. The M-Statistic, a skill score developed by Watterson et al. (1996), is presented for southeastern Australia's temperature and Australia's rainfall (Watterson et al. 2013). In addition the ability for the model to reproduce the latitude of the sup-tropical ridge across eastern Australia (Grose et al. 2015) is presented.

When comparing the top and bottom models to previous analyses focusing on direct climate model outputs for rainfall and temperature (Watterson et al. 2013) or the models' ability to reproduce the sub-tropical ridge (Grose et al. 2015) there is little agreement. This is not unexpected as the statistical downscaling relies on large-scale predictors, many in the lower troposphere and hence is bypassing the surface variables which may show larger biases due to systematic errors in model physical parameterisations rather than model dynamics. Furthermore, the applied calibration has also improved the original model output.

Testing the impact of the worse performing models on the overall results, it was found that removing the bottom 5 ranked models from our analysis made little difference to the reconstructed mean and the variance, while calculating the average of only the top 5 ranked models improved the mean and variance by less than 1%. These small changes while sub-sampling the ensemble of 22 downscaled models suggested that there is little expected improvement in the reliability of the future projections by removing models from this analysis.

Nevertheless, future projections across all models are compared to the limited sub-sampling of the bottom and top 5 ranked models in order to evaluate if that impacts the mean climate change response. These sub-sampled models are found to project similar changes (within approximately 5%) in streamflow for both RCP pathways (Table S2). The overall impact of selecting models on the annual streamflow projections is small and not consistent, therefore suggesting no particular bias toward a drier or wetter future if only a small sub-sample of the models are considered. This again implies that there is limited impact in selecting models based on its performance at reproducing the current climate, as the statistical downscaling reduces the largest model biases.

Table S1: Results of the models' evaluation relevant to this study: skill score (M-statistic) for temperature in southeast Australia and rainfall across Australia (Watterson et al. 2013), evaluation of the model ability to reproduce the sub-tropical ridge (Grose et al., 2015); model rankings in this study with respect to the mean and variance of reconstructed streamflow, and the two combined.. Scores in bold blue (red) indicate the top 5 (bottom 5) model performances.

Model	M-Stat Temp.South Aus	M-Stat Precip Aus	STR eval	Streamflow Rank		
				Mean	Variance	All
ACCESS1-0	**575**	552	2	8	16	13
ACCESS1-.3	492	544	2	**19**	6	10
bcc-csm1-1-m	**573**	525	2	12	**19**	**19**
BNU-ESM	**388**	451	**3**	**22**	15	**18**
CanESM2	**542**	492	2	9	**3**	**5**
CCSM4	519	**379**	2	**2**	**4**	**2**
CMCC-CMS	471	564	**1**	13	**18**	17
CNRM-CM5	**587**	**602**	**1**	**20**	13	16
CSIRO-mk3.6	**431**	482	**3**	15	7	8
GFDL-ESM2G	**383**	472	2	16	14	15
GFDL-ESM2M	458	469	2	**4**	17	12
HadGEM2-CC	533	541	**1**	14	**1**	**3**
IPSL-CM5A-LR	**395**	**403**	**3**	**1**	11	6
IPSL-CM5A-MR	477	404	2	17	**20**	**21**
IPSL-CM5B-LR	**424**	**596**	**3**	**5**	12	9
MIROC5	488	432	2	**21**	**22**	**22**
MIROC5-ESM	434	**342**	**3**	**18**	9	14
MIROC5-ESM-CHEM	450	**333**	**3**	7	**5**	**4**
MPI-ESM-LR	**542**	**593**	2	10	8	7
MPI-ESM-MR	513	**640**	**1**	**3**	**2**	**1**
MRI-CGCM3	511	**599**	**1**	6	**21**	**20**
NorESM1-M	480	**347**	2	11	10	11

Table S2: Downscaled climate model mean changes for annual Victorian streamflows using the R_1 regression. Changes are provided for the 2035-2065 and 2070-2100 periods in percent change from the reconstructed1975-2005 reference period, following the RCP 4.5 (top) and RCP 8.5 (bottom). Changes are shown using the full statistical reconstructions including monthly temperature for the full ensemble mean and for the five highest ranked and five lowest ranked models mean based on several model evaluation tests performed.

2050				
Ranking method	***RCP 4.5***		***RCP 8.5***	
Full Ensemble	-16.2		-21.4	
	Top 5	**Bottom 5**	**Top 5**	**Bottom 5**
Streamflow	-10.8	-16.5	-20.3	-25.5
M-Stat – Temp	-15.3	-21.9	-24.6	-18.9
M-Stat – Precip	-15.2	-12.3	-21.3	-14.2
STR	-14.0	-17.2	-23.2	-15.6
2085				
Ranking method	***RCP 4.5***		***RCP 8.5***	
Full Ensemble	-26.1		-44.8	
	Top 5	**Bottom 5**	**Top 5**	**Bottom 5**
Streamflow	-24.1	-26.7	-40.3	-49.2
M-Stat – Temp	-27.9	-25.0	-44.3	-50.6
M-Stat – Precip	-25.4	-15.9	-41.4	-31.2
STR	-28.8	-21.1	-44.1	-47.3

References:

Grose M, Timbal B, Wilson L, Bathols J, Kent D (2015) The subtropical ridge in CMIP5 models, and implications for projections of rainfall in southeast Australia. Aust Meteorol Oceanogr J 65: 90–106

Irving DB, Whetton P, Moise AF (2012) Climate projections for Australia: a first glance at CMIP5. Austr Met and Oceon Jour 62: 211–225

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

Watterson IG (1996) Nondimensional measures of climate model performance. Int J Climatol 16: 379–391

Watterson IG, Hirst AC, Rotstayn LD (2013) A skill score based evaluation of simulated Australian climate. Aust Met Ocean J 63: 181–190