

Interactive comment on “High-resolution projections of surface water availability for Tasmania, Australia” by J. C. Bennett et al.

J. C. Bennett et al.

james.bennett@csiro.au

Received and published: 13 April 2012

Responses to Anonymous Referee #2

Thanks to anonymous referee #2 for this considered review. We respond to the main points in the review (numbered, and in double quotation marks) as follows:

1. Comment: "The authors need to explain clearly (in a paragraph or two) as to how they can force it with just SST as forcing data from 6 GCMs."

Response: We have expanded the description of CCAM, as suggested, in section 3.1. CCAM is a global atmospheric model, and requires no atmospheric forcing from a GCM. In our study, CCAM simulates the atmosphere from sea-surface temperatures (SSTs), sea ice and radiative forcings. The SST and sea ice inputs come from GCMs.

C861

Radiative forcings (including greenhouse gas concentrations) come from the A2 emissions scenario. Two stages of downscaling are carried out for our study, with stage 2 nested in stage 1. Stage 1 is forced by bias-corrected GCM SST inputs and GCM sea ice, as well as radiative forcing. The second stage used the same GCM SST/sea ice forcing and radiative forcing. In addition, Stage 2 was forced by atmospheric variables from stage 1. In this way CCAM is forced only by SSTs and sea-ice concentration from GCMs, because stage 2 is nested inside stage 1.

2. Comment: "The authors need to explain how they and the readers can be sure that the bias correction for historical data for SST (1961-2000/7) is valid in the future, especially when you are applying the same correction for 2070-2099 (about 100 years ahead in time). The authors need to do a validation test using part of the observed SST data (say 1961-1980) where they develop the correction coefficients/factors and then show that they work for the other part of historical SST data (say 1981-2000/7). This should be shown in the paper so that we can have some confidence in the results for the future."

Response: We have carried out split-sample cross validation on SST bias-correction, as suggested, and described these in section 3.1:

'Correcting GCM SST biases improved the simulation of seasonal precipitation patterns (Ashfaq et al., 2010) and precipitable water (Held and Soden, 2006) in GCM experiments, and improved rainfall simulations by CCAM in experiments over tropical regions (Nguyen et al., 2011). GCM SST biases can exceed $\pm 3^{\circ}\text{C}$ for the GCMs used in this study (Randall et al., 2007), with the largest biases occurring at high latitudes and near the west coasts of South America and Africa. We found the bias-correction removed SST biases very effectively in split-sample cross-validation tests (not shown). In almost all regions cross-validated biases were less than $\pm 0.5^{\circ}\text{C}$ for all GCMs. No cross-validated SST biases were greater than $\pm 1^{\circ}\text{C}$. This suggests that the SST bias-correction is likely to be reliable for the period 2010-2100.'

C862

3. Comment: "I have major issues with bias correcting the outputs from RCM downscaling (rainfall and APET). It is not standard way to bias correct the outputs (there are a number of statistical downscaling papers which all bias correct the inputs but not the outputs). The authors can go one step further and just bias correct the RCM runoff so that they do not need to use any hydrological model?"

Response: The reviewer correctly notes that not all studies bias-correct outputs from downscaled climate projections before using them in hydrological models, and that biases can be corrected before downscaling of climate variables only. However, we disagree that this means our method is not the 'standard way', for two reasons: 1) there are a number of studies that do correct regional climate model (RCM) outputs after downscaling and then use them to force hydrological models, and we have cited several of these (Wood et al., 2004; Kilsby et al., 2007; Akhtar et al., 2009); 2) the diversity of published downscaling and bias-correction methods means there is not really a 'standard way'.

We do not agree that bias-correcting RCM outputs for use in hydrological models constitutes a 'major issue' with our study, as 1) the number of published studies that bias-correct RCM outputs for use in hydrological models (already noted) are a scientific precedent and a strong basis for our method, and 2) our method produces realistic river flows for the large majority of the 86 catchments we tested.

It is possible to simply bias-correct runoff outputs from RCMs, as the reviewer describes. We have not pursued this method because of the magnitudes of RCM runoff biases. Biases of runoff from RCMs/GCMs are typically much larger than rainfall biases (in the case of CCAM, annual runoff biases over Tasmania exceed 600% in some regions). This means that bias-corrections of runoff are likely to be less stable than bias-corrections of rainfall, and are likely to be less reliable for 2010-2100.

4. Comment: "The authors have stated that they are under predicting rainfall/runoff using the bias corrected rain and APET. This is to be expected as the high runoff events

C863

in a catchment are associated with majority of the grid cells within the catchment getting high rainfall on the same day or two which cannot be captured properly when you do cell by cell bias quantile-quantile mapping (it is done on ranked values and so timing is not considered)."

Response: The reviewer is correct in pointing to the timing of events as the cause of underestimation in a majority of catchments, and this was acknowledged in our manuscript. This is possible for any bias-correction method that does not correct multi-day rain or flow events as a single event (e.g. Ines and Hansen (2006), Piani et al.(2010)). As a consequence larger streamflows tend to be underestimated. For example, we find streamflow bias of -11.5% for the median catchment simulated with the SIMHYD model. However, spatial relationships between cells do not appear to contribute to the underestimation of high runoff. As briefly mentioned in the manuscript, we tested the tendency of catchments to under/overestimate runoff (some of these tests are described by Bennett et al. (2011)). We did not include these results in the paper to keep the paper as brief as possible. The tests showed that the underestimation of runoff is probably not caused by spatial factors, but rather by the underestimation of multi-day rainfall events by the bias-corrected CCAM outputs at individual grid cells.

5. Comment: "The authors need to do a validation test using part of the observed rainfall and APET data (say 1961-1980) where they develop the correction coefficients/factors and then show that they work fine for the other part of historical rainfall and APET data (say 1981-2000/7)... The authors say in the discussion very briefly that when they tested different periods for calibration and validation, they got very different results."

We have included an additional analyses cross-validating the quantile mapping using rainfall as an example (method described in section 3.2.1; results described in section 4.1):

'3.2.1 Split-sample cross-validation of quantile mapping

C864

When bias-corrections are applied to future projections, they are often implicitly assumed to be consistent through time (Ines and Hansen, 2006; Wood et al., 2004). However, there is some evidence that bias-corrections can vary with the choice of training period (Piani et al., 2010b; Li et al., 2010). We test this assumption using split-sample cross-validation for quantile mapping of rainfall. We carry out two sets of cross-validation:

- 1) We train the quantile mapping for the period 1962-1984 and validate against the period 1985-2007.
- 2) To minimise the effects of longer-term (decadal or greater) oscillations or trends in either the observed or simulated rainfalls, we train the quantile mapping on odd years for 1962-2007 (1963, 1965, ..., 2005, 2007) and validate against even years for 1962-2007 (1962, 1964, ..., 2004, 2006).

4.1 Cross-validation of quantile mapping

Fig 1 shows cross-validation biases for mean annual rainfall for the GFDL-CM2.0 simulation. Rainfall biases of the GFDL-CM2.0 simulation shown in Fig 1 are very similar for the other RCM simulations (not shown). When the quantile mapping is trained on odd years (1963, ..., 2007), cross-validation biases are less than $\pm 10\%$ almost everywhere (Fig 1a). Cross-validation biases are much smaller than biases of uncorrected RCM rainfalls (Fig 1c). Biases of uncorrected RCM rainfalls are larger than $\pm 30\%$ for much of Tasmania, and exceed 150% in some cells. The relatively small cross-validation biases suggest that the quantile mapping is effective outside the training period. Accordingly, the quantile mapping is likely to be reliable for the period 2010-2100. A longer training period (e.g. 30+ years) may have produced even small cross-validation biases. We had 47 years of synchronous simulations and observations, which allowed 23 years to train the quantile mapping for the cross-validation tests (assuming an equally long validation period). 23 years may be an insufficiently long period to sample the natural variance in rainfall.

C865

When the quantile mapping is trained on 1962-1984 (Fig 1b), cross-validation biases are larger ($\pm 10\%$ to $\pm 25\%$), and occur over larger areas than when the quantile mapping is trained on odd years (1963, ..., 2007) (Fig 1a). This discrepancy in performance is caused by changes in Tasmanian rainfall between 1962-1984 and 1985-2007. The period 1962-1984 experienced higher rainfalls over much of Tasmania than 1985-2007, particularly in eastern Tasmania. RCM simulations are not synchronised with observations, and the change in observed rainfall from 1962-1984 to 1985-2007 is not present in the RCM simulations. When the quantile mapping is trained to match the higher rainfalls in of 1962-1984, it consequently overestimates rainfall during the drier period 1985-2007. (The exception to this pattern is over the south-western mountains, which experienced very wet years in 1994 and 1996; here the bias-correction underestimates rainfalls in the validation period.) These long-term changes in Tasmanian rainfall do not have the same effect on the quantile mapping when it is trained on odd years over a period that includes both wet and dry periods, resulting in lower cross-validation biases (Fig 1a). This illustrates the importance of sampling long periods to generate temporally stable cumulative frequency distributions for quantile mapping.'

Please note that the caption of the figure is too long to fit in the dialogue box on the author comments upload webpage. The full figure caption is:

Figure 1 Effects of changing the training period on quantile mapping performance. Plots show biases in mean annual rainfall for the 10 km² GFDL-CM2.0 simulation. (a) Biases of bias-corrected RCM rainfall where quantile mapping is trained with the period 1962-1984; biases shown are calculated for 1985-2007. (b) Biases of bias-corrected RCM rainfall where quantile mapping is trained with odd years occurring in the period 1962-2007 (1963, 1965, ..., 2005, 2007); biases shown are calculated for even years occurring during the period 1962-2007 (1962, 1964, ..., 2004, 2006). (c) Biases of uncorrected RCM rainfalls for 1961-1990. Rainfall biases for the other five RCM simulations are very similar to the GFDL-CM2.0 simulation illustrated in this figure.

C866

6. Comment: "The authors seem to put one reason for any issues with the runoff projection – problem/deficiencies in the hydrological models. This is not true, for example, the deficiencies in low streamflow that you are getting is not because of the problems with the hydrological models BUT because of the objective function that you have used to calibrate the hydrological models."

The reviewer correctly points out that a different calibration method or objective function could have been used to improve hydrological performance at low flows, and this was not discussed in our manuscript. We have restated the last paragraph of section 4.2.2 to emphasise that our findings are specific to the models as calibrated for our study, and that a different calibration method or different hydrological models may give a contrary result. We have also acknowledged that high flows could have been better represented in certain catchments with different calibration methods or hydrological models in the last paragraph of section 4.2.3. We have acknowledged that some the bias-corrected RCM inputs contribute to poor low-flow replication. The general comment about low-flow replication being a common problem in hydrological models has been removed.

Changes not requested by referee:

To keep the paper as brief as possible we have abbreviated the abstract and some elements of the introduction and discussion in an attempt to compensate for additional text included to address both anonymous reviewers' comments. These abbreviations do not substantively alter the content or arguments in the paper.

References:

Akhtar, M., Ahmad, N., and Booij, M. J.: Use of regional climate model simulations as input for hydrological models for the Hindukush-Karakorum-Himalaya region, *Hydrology and Earth System Sciences*, 13, 1075–1089, doi: 10.5194/hess-13-1075-2009, 2009.

Ashfaq, M., Skinner, C. B., and Diffenbaugh, N. S.: Influence of SST biases on future

C867

climate change projections, *Climate Dynamics*, 36, 1303-1319, doi: 10.1007/s00382-010-0875-2, 2010.

Bennett, J. C., Grose, M. R., Post, D. A., Ling, F. L. N., Corney, S. P., and Bindoff, N. L.: Performance of quantile-quantile bias-correction for use in hydroclimatological projections, MODSIM2011 International Congress on Modelling and Simulation, Perth, December, 2011.

Held, I. M., and Soden, B. J.: Robust responses of the hydrological cycle to global warming, *Journal of Climate*, 19, 5686-5699, 2006.

Ines, A. V. M., and Hansen, J. W.: Bias correction of daily GCM rainfall for crop simulation studies, *Agricultural and Forest Meteorology*, 138, 44-53, doi: 10.1016/j.agrformet.2006.03.009, 2006.

Kilsby, C. G., Tellier, S. S., Fowler, H. J., and Howles, T. R.: Hydrological impacts of climate change on the Tejo and Guadiana rivers, *Hydrology and Earth System Sciences*, 11, 1175-1189, doi: 10.5194/hess-11-1175-2007, 2007.

Li, H., Sheffield, J., and Wood, E. F.: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, *Journal of Geophysical Research*, 115, D10101, doi: 10.1029/2009JD012882, 2010.

Nguyen, K. C., Katzfey, J. J., and McGregor, J. L.: Global 60 km simulations with CCAM: evaluation over the tropics, *Climate Dynamics*, 0930-7575, doi: 10.1007/s00382-011-1197-8, 2011.

Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, *Theoretical and Applied Climatology*, 99, 187-192, doi: 10.1007/s00704-009-0134-9, 2010.

Wood, A., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate outputs, *Climatic Change*,

C868

C869

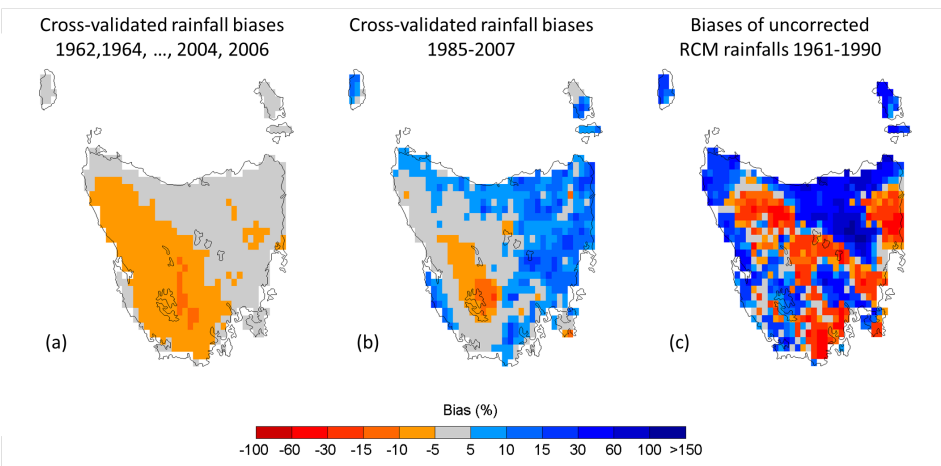


Fig. 1. Effects of changing the training period on quantile mapping performance. Plots show biases in mean annual rainfall for the 10 km² GFDL-CM2.0 simulation.

C870