The authors document the development of a national set of hydrological projections for Australia. They provide good motivation for the development of the product/service and provide extensive description and evaluation of the output. My main comment is around the use of the 3 bias correction methods to expand the ensemble. It's not clear to me why they looked at 3 methods and what it adds. E.g. at one point (for temperature) the authors describe how "the three bias corrections methods are almost indistinguishable…" (pg 12, line 9, Figure 6). Also at pg 16, line 9 (re Figure 12), the authors comment: "The first observation implies that in general, the choice of GCM contributes more to the ensemble spread than the application of the BC techniques". I think some more discussion and justification for the inclusion of all 3 bias correction methods is required.

It wasn't apparent how the various bias correction techniques would affect the projections when they were first chosen. We decided on using three techniques due to their availability: QME was developed at the Bureau, we had established connections with UNSW to obtain the MRNBC code and the ISIMIP2b method is freely available and was specifically designed to be used for impact studies. A thorough evaluation and "ranking" of the BC methods is available (Vogel et al., 2023). This study found that some methods may be more appropriate depending on the application. For instance, the QME may perform better at capturing the extremes, while the MRNBC was found to perform better overall, especially for the hydrological output.

Reviewer three asked a similar question, so for completeness we will include our response to them here:

- " Page 16 Line 9-10: Do the results imply that perhaps only one best performing bias correction technique is needed for your application? Perhaps more GCMs and/or RCMs should be included to better gauge the uncertainty of the future projections.

That may indeed be the case. For instance we found that QME performed best when measured against extremes, while the MRNBC performed the "best" overall when ranked across a range of metrics (Vogel et al., 2023). It would be ideal to include a full suite of CMIP models, or at least, a subset based on benchmarking criteria. That opportunity was not available to us for NHP (as explained in Section 2). In any case, the frequentist approach may not be the best way in which to represent uncertainty, particularly for impact assessment studies (Shepherd, 2021).

There are emerging opportunities however for the Australian Climate Service (ACS) to produce similar data sets based on a carefully selected subset of CMIP6 models (Grose et al., 2023). One of their key findings is that: "The projections cannot be considered a probabilistic or balanced estimate of uncertainty given the limited ensemble size and underlying epistemic uncertainties. The ensemble can however be used in a 'climate futures' or 'storyline' approach to illustrate plausible future climates that broadly span the range of possibilities suggested by CMIP6, while producing added value at the regional scale." The assessment reports produced for the NHP (https://awo.bom.gov.au/about/overview/assessment-reports) were also structured with a storyline approach, which we document in Section 8.3."

We have modified the text (now page 12; lines 29-32) to:

" The three bias correction methods follow very similar trajectories and only slight differences are discernible for the GFDL-ESM2M model. Since temperature is a smooth field (both spatially and temporally) and all the bias correction methods are variants of quantile matching, this indicates that

at least for the mean and large geographical aggregation, the bias correction imparts less uncertainty than GCM selection in the projections."

5 Other points:

Page 5, line 1-2: I respect that pragmatic choices need to be made but I think the authors should make some comment around the validity of RCP8.5 for future risk assessment e.g. https://www.nature.com/articles/d41586-020-00177-3

We tried to address this with the following sentence:

10 "These emissions pathways were chosen to provide a high (RCP8.5) and moderate (RCP4.5) set of temperature projections, noting that the set of modelled greenhouse gas emission pathways provided in CMIP5 have relatively minimal deviation before 2050 and the observed climate change trends for CO_2 emissions and temperature in recent decades indicate that the high emissions pathway (RCP 8.5) has been followed more closely than other emissions pathways (e.g., RCP 2.6)
15 (Schwalm et al., 2020; Stocker et al., 2013)."

We have added: "However, we also note that RCP 8.5 may be at the high end of emission scenarios for future risk assessment (Peters and Hausfather, 2020).

Page 5, line 24: Why is this sentence in bold?

20 This has been fixed.

Figure 2: Perhaps it's obvious, but I think somewhere in the caption the authors should write that the darker shaded bars indicate selected models

The caption has been modified to: "*The ranking of surface temperature (tas) for the GCMs used in this study (darker shading) with regard to the CCiA ensemble (lighter shading). The horizontal bars*
25 *indicate the change signal (the difference of the nationally-averaged quantity from the climatology for the period 1976-2005). Four 30-year periods are shown centred on 2030, 2050, 2070 and 2085, for RCP8.5."*

Page 6, line 1: The use of 1976-2005 historical period is a departure from CCiA. Can the authors
30 comment why they chose that period?

We chose that end year (2005) based on the year the CMIP5 historical data finished (like CCIA). Additionally, we chose a reference time period of 30 years since our end goal was to produce hydrological projections, where we wanted to try and capture some long period hydrological features. For instance, a 20-yr reference period may not be of sufficient length to capture drought
35 periods.

Page 7, line 10: Why not include the CCAM simulations to get a better picture of the spread relative to CCiA?

This analysis was to show where our selection of GCMs sit within the CMIP5 ensemble. Since CCAM is a stretched grid model forced by a GCM (see (Thatcher and McGregor, 2009) it carries much information from the forcing GCM. It is true that in the configuration used for NHP (using bias-corrected sea surface temperature), that the interior of CCAM is allowed to evolve freely (i.e., it can develop a climate quite distinct from the forcing GCM) it still contains information from that GCM. The intention was to show the spread of the selected GCMs within the CMIP5 ensemble, particularly given that the selected GCMS were a subset of those recommended in CCIA. The spread resulting from the inclusion of the CCAM is explained in further analyses, particularly that of Section 7.

Page 7, line 21: GCM/RCM not just GCM

GCM changed to "GCM and/or RCM simulations…"

Page 7, line 23: This sentence doesn't make sense. I think the "are" before precipitation should be replaced with a colon or dash.

Thank you for spotting this. The "are" before precipitation has been replaced with a colon.

Page 7, line 26: Does this need to be updated to AGCD?

At the time of producing preparing the observational data, the data was known as AWAP. It is now known as AGCD (v1.0) / AWAP. We have added a footnote: " The interpolation used to produce the AWAP analysis is currently being updated by the Bureau of Meteorology and is now known as Australian Gridded Climate Data (AGCD) / AWAP; v 1.0.0. See (Evans et al., 2020) for details."

Page 9, line 15: This sentence doesn't make sense and needs to be rewritten.

Sentence now reads: "Non-parametric QM techniques map the simulated quantiles of the cumulative distribution function (CDF) to the observed CDF quantiles without any underlying assumptions that the variable can be modelled by a mathematical distribution, whereas for parametric QM, distributions are fitted to the variables before application of the QM".

Page 12, line 26: As I understand it, the purpose of an ensemble member is to add new information. How are the 3 bias correction methods adding new information? Can the authors comment on this? (see my main comment)

Figure 6, 7: Perhaps I missed it but why are CCAM-MRNBC and CCA-QME not included?

This was due to time constraints. The ISIMIP2b method was the first to be implemented during the NHP and was already available when the CCAM output was also available, while the other two methods were still under development. We have added a footnote (footnote number 5) in Section 4 (Page 10, line 1):

" Due to time constraints, the only bias correction algorithm applied to the CCAM output was the ISIMIP2b method."

Fig 8: Although it's mentioned in the text, I think it would be helpful to add a sentence in the caption about why there are data gaps in the maps.

The following has been added to the caption: "Data sparse regions have been masked (see Section 3.1)".

Page 16, line 5: NRMs (or NRM regions) not NRMS

NRMS changed to NRMs.

Figure 12: It seems pointless to label each of the 4 plots with "Southern and South Western Flatlands" – this could just be written in the caption. I think having the variable (e.g. precipitation, soil moisture etc.) clearly visible at the top of each plot would be helpful.

5   We have removed the NRM description (Southern and South Western Flatlands), however, we have not included the variable in the title but rather in the figure caption to be consistent with the other plots.

Figure 13: It's very difficult to interpret these plots, can the resolution be sharpened?

We appreciate that there is a lot of information in these plot. We considered only showing one
10   variable and including the other variables in the supplementary material; however, it is important to show the seasonal characteristics for all the hydrological variables in the main manuscript. Our other option would be to have each panel as a separate plot, however, that would increase the figure numbers (and we had reviewer comments about trying to reduce the number of figures).

We have included a high resolution eps file to examine the finer details of the plots (e.g. examining
15   specific bias correction methods and/or GCM). The explanation in the text is mainly concerned with the ensemble mean interpretation, which is illustrated in the low-resolution (png format) version.

Page 17, line 13/14: You've written "antecedent conditions soil moisture conditions". I assume you mean "antecedent soil moisture conditions".

Yes, modified accordingly.

20   Page 22, line 8: It seems strange that 'personnel issues' is listed here but not earlier in the manuscript.

Reviewer 1 had a comment about this phrase and it has now been modified (following the recommendations of reviewer 1) to:

" Due to the large spatial domain, only the output of four GCMs was able to be bias-corrected…"

25

References:

Evans, A., Jones, D., Smalley, R. and Lellyett, S.: An enhanced gridded rainfall dataset scheme for Australia. [online] Available from: http://www.bom.gov.au/research/research-reports.shtml, 2020.

Peters, G. P. and Hausfather, Z.: Emissions - the "business as usual" story is misleading, Nature,
30   577(January), 618–620, 2020.

Schwalm, C. R., Glendon, S. and Duffy, P. B.: RCP8.5 tracks cumulative CO2 emissions, Proc. Natl. Acad. Sci. U. S. A., 117(33), doi:10.1073/PNAS.2007117117, 2020.

Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M. M. B., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P. M.: Climate change 2013 the physical science basis: Working Group I
35   contribution to the fifth assessment report of the intergovernmental panel on climate change., 2013.

Thatcher, M. and McGregor, J. L.: Using a scale-selective filter for dynamical downscaling with the conformal cubic atmospheric model, Mon. Weather Rev., 137(6), 1742–1752, doi:10.1175/2008MWR2599.1, 2009.

Vogel, E., Johnson, F., Marshall, L., Bende-Michl, U., Wilson, L., Peter, J. R., Wasko, C., Srikanthan, S.,

Sharples, W., Dowdy, A., Hope, P., Khan, Z., Mehrotra, R., Sharma, A., Matic, V., Oke, A., Turner, M., Thomas, S., Donnelly, C. and Duong, V. C.: An evaluation framework for downscaling and bias correction in climate change impact studies, J. Hydrol., 622, 129693, doi:10.1016/J.JHYDROL.2023.129693, 2023.

5