Hydroclimatic projections for the Murray-Darling Basin based on an ensemble derived from Intergovernmental Panel on Climate Change AR4 climate models

Fubao Sun,[1] Michael L. Roderick,[1,2] Wee Ho Lim,[1] and Graham D. Farquhar[1]

Received 31 July 2010; revised 11 November 2010; accepted 13 January 2011; published 1 April 2011.

[1] We assess hydroclimatic projections for the Murray-Darling Basin (MDB) using an ensemble of 39 Intergovernmental Panel on Climate Change AR4 climate model runs based on the A1B emissions scenario. The raw model output for precipitation, P, was adjusted using a quantile-based bias correction approach. We found that the projected change, ΔP, between two 30 year periods (2070–2099 less 1970–1999) was little affected by bias correction. The range for ΔP among models was large (~±150 mm yr^{-1}) with all–model run and all-model ensemble averages (4.9 and −8.1 mm yr^{-1}) near zero, against a background climatological P of ~500 mm yr^{-1}. We found that the time series of actually observed annual P over the MDB was indistinguishable from that generated by a purely random process. Importantly, nearly all the model runs showed similar behavior. We used these facts to develop a new approach to understanding variability in projections of ΔP. By plotting ΔP versus the variance of the time series, we could easily identify model runs with projections for ΔP that were beyond the bounds expected from purely random variations. For the MDB, we anticipate that a purely random process could lead to differences of ±57 mm yr^{-1} (95% confidence) between successive 30 year periods. This is equivalent to ±11% of the climatological P and translates into variations in runoff of around ±29%. This sets a baseline for gauging modeled and/or observed changes.

Citation: Sun, F., M. L. Roderick, W. H. Lim, and G. D. Farquhar (2011), Hydroclimatic projections for the Murray-Darling Basin based on an ensemble derived from Intergovernmental Panel on Climate Change AR4 climate models, *Water Resour. Res.*, *47*, W00G02, doi:10.1029/2010WR009829.

1. Introduction

[2] The Murray-Darling Basin (MDB) located in south-east Australia (Figure 1) is Australia's food bowl, with almost 40% of Australia's agricultural production. The region supports extensive grazing, dryland cropping, and, most importantly, a variety of irrigated crops. Acute water shortages in the basin in recent years as a result of drought and overallocation have focused attention on the long-term sustainability of activities within the basin [e.g., *Murray-Darling Basin Commission*, 2009; *Potter et al.*, 2010; *Maxino et al.*, 2008]. Superimposed on that are concerns about the possible impact of climate change on water availability in the future.

[3] One key part of the information base used to evaluate likely future conditions is the projections from state-of-the-art coupled atmosphere-ocean general circulation models. The most recent compilation of model simulations has been made available through the World Climate Research Programme's Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel data set [*Meehl et al.*, 2007]. The same database was used to prepare the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [*IPCC*, 2007] (hereafter referred to as IPCC AR4 models) and is widely used to assess the hydrologic impact of climate change [e.g., *Groves et al.*, 2008; *Chiew et al.*, 2009].

[4] Climate scientists usually examine the statistical properties of ensembles that are based on raw model output [e.g., *Tebaldi and Knutti*, 2007; *Knutti et al.*, 2010]. The underlying idea is that each model run is a physically plausible representation of the future climate. However, such climate models have a coarse spatial resolution (greater than two or three hundreds of kilometers) and known problems of bias [e.g., *Fowler et al.*, 2007; *Groves et al.*, 2008; *Wilby and Harris*, 2006].

[5] For precipitation in particular, different models or, on occasion, different runs of the same model give very different regional-scale simulations for the historical period and projections for the future as documented in the global water atlas [*Lim and Roderick*, 2009] and elsewhere [*Johnson and Sharma*, 2009]. An interesting point here is that in terms of globally integrated precipitation, there is little practical difference between various simulations (i.e., for the historical period) and projections (i.e., for the future) of climate models [*Lim and Roderick*, 2009]. The large differences in model simulations and model projections occur at regional scales.

[6] Two basic approaches have been used to make regional-scale projections. The first, here called the ranking

[1] Research School of Biology, Australian National University, Canberra, ACT, Australia.
[2] Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia.

Copyright 2011 by the American Geophysical Union.
0043-1397/11/2010WR009829

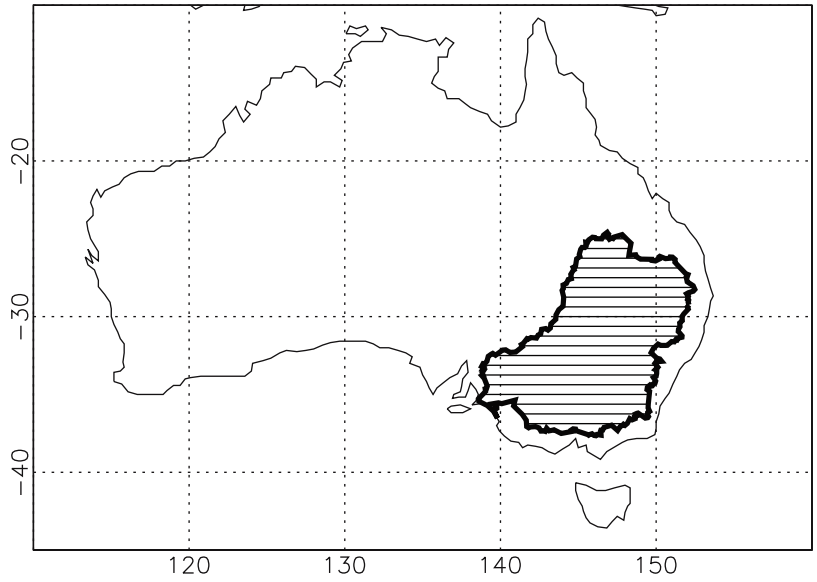


Figure 1. Location of the Murray-Darling Basin (MDB, shaded area).

approach, is based on the idea that some models give better representations of the recent past in a given region. The ranking is based on comparing the model output with observations over the historical period. Previous research on precipitation scenarios for the MDB have generally followed this approach [e.g., *Maxino et al.*, 2008; *Chiew et al.*, 2009; *Smith and Chandler*, 2010]. The ranking approach can have some interesting consequences. For example, as summarized by *Smith and Chandler* [2010, Table 2], the model rankings for precipitation simulations over the MDB are different from the ranking for the entire Australian continent. Taking this to the limit, it might turn out that the most highly ranked model for a particular purpose would vary from one region to the next.

[7] The second, here called the bias correction approach, is based on the idea that the statistical properties of the model output can be adjusted to be identical with observations over the historical period. This approach has been widely used in climate change impact studies [*Wood et al.*, 2004; *Maurer and Hidalgo*, 2008]. The simulation from each individual model run is adjusted so that the overall mean and the variance match observations for the historical period.

[8] In comparing the two approaches, each model will have the same ranking after bias correction, and each model will therefore contribute equally to the ensemble. In contrast, the whole aim of model ranking is to change the relative weights and, in some cases, even remove models (i.e., zero weight) from the ensemble. This could have a major impact on the projected hydroclimatic changes. In that respect, the aim of this paper is to develop an understanding of the properties of the individual model runs and the statistical properties of the ensemble of the simulations, projections, and projected changes. To do that we use data from the global water atlas [*Lim and Roderick*, 2009].

2. Data and Methods

2.1. Climate Model Database

[9] The global water atlas [*Lim and Roderick*, 2009] is based on monthly climate model output for precipitation (P), evaporation (E), and land area fraction available in the multimodel climate data archive for the IPCC AR4 models [*Meehl et al.*, 2007]. In preparing the atlas, there was no a priori model selection; that is, all models having available data (for P and E) were used. In total, the output for 39 paired model runs from 20 different climate models (Table 1) was available for the historical period known as the 20C3M scenario (climate in the 20th century) and for one future scenario, the A1B scenario [*IPCC*, 2000]. The A1B scenario (the 750 ppm stabilization scenario) assumes midrange emissions for 2000–2099. The 39 individual model runs are here called the all–model run ensemble. Multiple runs were available for eight of the models (Table 1), with each run representing a different set of initial conditions [e.g., *Rotstayn et al.*, 2007, p. 5]. Hence, those multiple runs can be used to examine the sensitivity to initial conditions.

Table 1. Summary of the Climate Model Output Showing Number of Monthly Runs Available for Each Model-Scenario Combination

	Model and Country	20C3M	A1B
Model 1	BCCR-BCM2.0, Norway	1	1
Model 2	CGCM3.1(t63), Canada	1	1
Model 3	CNRM-CM3, France	1	1
Model 4	CSIRO-Mk3.0, Australia	3	1
Model 5	CSIRO-Mk3.5, Australia	1	1
Model 6	GFDL-CM2.0, USA	3	1
Model 7	GISS-AOM, USA	2	2
Model 8	GISS-EH, USA	5	3
Model 9	GISS-ER, USA	9	2
Model 10	INGV-ECHAM4, Europe, ECMWF	1	1
Model 11	INM-CM3.0, Russia	1	1
Model 12	IPSL-CM4, France	1	1
Model 13	MIROC3.2_HIRES, Japan	1	1
Model 14	MIROC3.2_MEDRES, Japan	2	2
Model 15	MIUB-ECHO-G, Germany/Korea	5	3
Model 16	MPI-ECHAM5, Germany	4	4
Model 17	NCAR-CCSM3.0, USA	8	7
Model 18	NCAR-PCM1, USA	4	4
Model 19	UKMO-HADCM3, UK	2	1
Model 20	UKMO-HADGEM1, UK	2	1
Total		57	39

Output from the 20 models (where multiple runs were averaged) are here called the all-model ensemble.

[10] The model output was resampled into a common geographic grid of dimensions 2.5° × 2.5° (~270 km × 270 km). The monthly model output was aggregated into annual time series, and projected changes were calculated from the difference in P, E, and $P - E$ between the end of the 20th (1970–1999) and 21st (2070–2099) centuries. A geographic mask defining the MDB (Figure 1) was used in conjunction with the model-specific land area fraction to extract model output for the MDB region.

[11] Initial calculations revealed a problem with the estimates of E. When integrated over the MDB, the 30 year averages in E were often greater than P in many models. Our investigation found that the problem was caused by the way the climate model output is archived. The climate models presumably calculate P and E separately for the land and ocean fractions of each grid box and add them, adjusting for land area fraction as appropriate to obtain the total P and E for each grid box. However, the separate grid box level estimates for the land and ocean components are not archived. When reconstructing the estimates, the base level assumption is that the land and ocean E and P scale directly with the respective area fractions. This ignores the fact that over the long term, E does not exceed P over land but can over the ocean (or lake) where water is always available for evaporation. The worst-case scenario occurs along dry coastal regions where E from the ocean will be much higher than E from land. This is very clear in global maps made using the raw outputs of P and E (e.g., for their difference, see *Held and Soden* [2006, Figure 7]) where there are clear discontinuities along arid coastlines such as around much of Australia or in the Middle East.

[12] We tried various schemes, and the final approach to reconstructing the climate model output for E is described in detail in Appendix A. In brief, as a "work-around," after calculating E in each grid box within the MDB, we tested whether the 100 year average E was greater than P. If so, then the 100 year average E was reset to be equal to the 100 year average P. It should be noted that by following this procedure, we may still have E greater than P in any given year or even in a 30 year average. Hence, our approach is not ideal, but the alternative, of ignoring the problem, led to unphysical, and unrealistic, results for the MDB water balance. This problem is a regional one and could be resolved if the land and ocean components of P, but especially E, were archived separately for each grid box in the climate model output.

2.2. Precipitation Observations

[13] We obtained the time series of observed annual precipitation for the MDB from the Bureau of Meteorology of Australia (1900–2008, http://www.bom.gov.au). To aid with the bias correction, we also obtained the monthly precipitation data from the Global Precipitation Climatology Center (GPCC) database [*Rudolf and Schneider*, 2005; *Rudolf et al.*, 2010]. This data set is developed from rain gauge–based precipitation data interpolated on a 2.5° × 2.5° grid from 1901 to 2007. We compared the 107 year time series of annual precipitation for the MDB from the two sources and found that they were, for all practical purposes, identical (linearly regressed with the slope 1.02, intercept −6.03 mm yr^{-1}, and determination coefficient 0.996). Hence, the GPCC monthly precipitation data set was used to undertake the bias correction of precipitation data as described in section 2.3.

2.3. The Bias Correction Method

[14] Traditional quantile-based mapping bias correction approaches adjust the mean and variance of a model simulation to agree with the statistical properties of the observations. Specifically, the cumulative distribution function (CDF) of the model output is adjusted to agree with the CDF of the observations [*Wood et al.*, 2004; *Maurer and Hidalgo*, 2008]. In detail, for a given grid box and month, one first locates the percentile value for the model simulation and then replaces the simulated monthly precipitation with the observed monthly precipitation from the same percentile in the (observed) CDF. This is the bias-corrected output. The remaining challenge is how to adjust the model projections. Recently, *Li et al.* [2010, Figure 3] proposed an approach where the difference over time in the model output (CDF model projection minus CDF model simulation) is preserved in the projection. For the future projection, one first constructs the CDF for the model projection, simulation, and the observations and then locates the corresponding percentile values in the three CDFs. The bias-corrected output is calculated by adding the difference between the model projection and simulation to the observation at the same percentile. Hence, projected changes in the model output should be preserved.

[15] Following the *Li et al.* [2010] method, we used monthly P estimates (GPCC, 1901–2007) to perform the bias corrections. We then aggregated to annual data and calculated P for the relevant 1970–1999 and 2070–2099 periods. Note that on completion of the procedure, the (monthly) variance in each individual model run for the historical period will, by construction, equal the (monthly) variance in observations. However, after the bias correction, the variance of the annual time series of each model run will not necessarily be equal to the variance in the (annual) observations because the bias correction method used monthly data [*Sun et al.*, 2010].

3. Results

3.1. Hydrologic Balance of Raw Model Outputs

[16] The hydrologic summaries of the raw outputs from the 39 climate model runs for the MDB are described in Table 2. In the 39 model runs examined, P for the 1970–1999 period varied from 230.1 to 994.4 mm yr^{-1}. The all–model run average was 578.4 (±176.2 mm yr^{-1}) (plus or minus standard deviation, denoted by σ) compared with the observed value of 517.4 mm yr^{-1}. Of the 39 model runs, 22 showed increases in P to the end of the 21st century while 17 showed decreases. The average change for the all–model run ensemble was for a very small increase in MDB annual precipitation of 4.9 mm yr^{-1} by the end of the 21st century. However, the range is large, with some model runs projecting a drop in annual precipitation of as much as 150 mm (e.g., CSIRO-Mk3.0, CSIRO-Mk3.5, and MPI-ECHAM5 Run2), while other model runs simulate increases of about the same amount (e.g., MIROC3.2_MEDRES).

[17] The results for the all-model ensemble for the 1970–1999 period were virtually identical, with a large range

Table 2. Summary of Raw Hydrologic Outputs for the MDB (mm yr^{-1})

	1970–1999 (20C3M)			2070–2099 (A1B)			Δ (A1B-20C3M)		
	P	E	$P-E$	P	E	$P-E$	ΔP	ΔE	$\Delta(P-E)$
BCCR-BCM2.0 Run1	867.1	719.5	147.6	908.2	766.1	142.1	41.1	46.6	−5.5
CGCM3.1(t63) Run1	468.9	462.9	5.9	534.9	526.3	8.6	66.0	63.4	2.6
CNRM-CM3 Run1	478.5	458.9	19.6	473.0	459.9	13.1	−5.6	0.9	−6.5
CSIRO-Mk3.0 Run1	555.6	547.8	7.8	461.9	471.2	−9.3	−93.7	−76.6	−17.1
CSIRO-Mk3.5 Run1	458.9	452.4	6.5	349.1	352.0	−2.9	−109.8	−100.5	−9.4
GFDL-CM2.0 Run1	360.0	351.4	8.6	304.9	302.4	2.5	−55.1	−49.0	−6.1
GISS-AOM Run1	284.1	282.9	1.3	227.9	230.1	−2.1	−56.2	−52.8	−3.4
GISS-AOM Run2	254.7	256.2	−1.6	232.9	234.8	−1.9	−21.8	−21.4	−0.4
GISS-EH Run1	984.2	897.6	86.6	1062.8	965.1	97.7	78.5	67.4	11.1
GISS-EH Run2	994.4	897.6	96.8	1048.6	946.5	102.1	54.2	48.9	5.3
GISS-EH Run3	974.2	893.6	80.6	1022.2	933.9	88.4	48.0	40.3	7.7
GISS-ER Run6	641.7	623.4	18.3	681.3	659.4	21.8	39.5	36.0	3.5
GISS-ER Run8	643.6	622.8	20.8	654.2	638.9	15.2	10.6	16.2	−5.6
INGV-ECHAM4 Run1	690.5	635.2	55.3	637.1	588.4	48.7	−53.4	−46.8	−6.6
INM-CM3.0 Run1	517.2	480.4	36.9	517.9	500.5	17.4	0.7	20.1	−19.5
IPSL-CM4 Run1	230.1	215.7	14.4	166.4	172.9	−6.5	−63.7	−42.8	−20.9
MIROC3.2_HIRES Run1	631.3	620.3	11.0	665.1	626.3	38.8	33.9	6.0	27.9
MIROC3.2_MEDRES Run2	680.0	675.5	4.6	835.4	809.3	26.1	155.4	133.8	21.6
MIROC3.2_MEDRES Run3	701.5	696.5	5.0	852.7	815.2	37.6	151.2	118.7	32.5
MIUB-ECHO-G Run1	540.4	535.9	4.5	598.0	591.7	6.3	57.7	55.9	1.8
MIUB-ECHO-G Run2	523.3	522.9	0.4	598.6	586.6	12.0	75.3	63.7	11.6
MIUB-ECHO-G Run3	532.8	527.4	5.4	626.0	614.4	11.6	93.2	87.0	6.2
MPI-ECHAM5 Run1	387.2	381.9	5.3	332.8	331.8	1.0	−54.4	−50.1	−4.3
MPI-ECHAM5 Run2	489.4	482.6	6.8	331.0	328.9	2.1	−158.5	−153.7	−4.8
MPI-ECHAM5 Run3	419.1	408.0	11.1	336.8	334.0	2.8	−82.3	−74.0	−8.3
MPI-ECHAM5 Run4	357.4	356.8	0.7	378.7	372.3	6.4	21.2	15.6	5.7
NCAR-CCSM3.0 Run1	643.8	494.0	149.8	676.4	530.1	146.4	32.7	36.1	−3.5
NCAR-CCSM3.0 Run2	658.9	503.9	155.0	667.7	520.3	147.4	8.9	16.4	−7.5
NCAR-CCSM3.0 Run3	646.2	489.8	156.4	665.7	517.6	148.1	19.5	27.8	−8.3
NCAR-CCSM3.0 Run5	634.5	487.4	147.0	666.6	523.2	143.4	32.1	35.7	−3.6
NCAR-CCSM3.0 Run6	628.6	484.9	143.7	660.9	515.1	145.7	32.3	30.2	2.1
NCAR-CCSM3.0 Run7	617.3	477.2	140.2	675.5	526.7	148.8	58.1	49.5	8.6
NCAR-CCSM3.0 Run9	613.8	474.5	139.4	678.7	529.3	149.5	64.9	54.8	10.1
NCAR-PCM1 Run1	569.0	569.1	−0.1	543.8	547.5	−3.7	−25.2	−21.7	−3.6
NCAR-PCM1 Run2	589.6	579.2	10.4	569.7	572.2	−2.5	−20.0	−7.0	−12.9
NCAR-PCM1 Run3	595.2	585.9	9.3	572.6	572.0	0.5	−22.6	−13.9	−8.8
NCAR-PCM1 Run4	579.8	574.8	5.0	561.8	567.4	−5.6	−18.0	−7.3	−10.6
UKMO-HADCM3 Run1	525.0	519.8	5.3	457.8	452.1	5.6	−67.3	−67.6	0.3
UKMO-HADGEM1 Run1	590.2	572.1	18.1	514.0	496.0	18.1	−76.2	−76.1	−0.1
Observation	517.4								
The All–Model Run Ensemble									
Mean across all runs	578.4	533.8	44.6	583.3	539.2	44.1	4.9	5.4	−0.5
σ	176.2	154.0	57.8	217.7	188.2	58.8	68.3	60.9	11.4
Min	230.1	215.7	−1.6	166.4	172.9	−9.3	−158.5	−153.7	−20.9
Max	994.4	897.6	156.4	1062.8	965.1	149.5	155.4	133.8	32.5
Number of runs showing increases							22	23	16
Number of runs showing decreases							17	16	23
The All-Model Ensemble									
Mean across all models	556.2	525.6	30.6	548.1	519.2	28.9	−8.1	−6.4	−1.7
σ	179.3	155.6	45.1	221.1	190.9	46.3	68.2	60.6	12.6
Min	230.1	215.7	−0.1	166.4	172.9	−9.3	−109.8	−100.5	−20.9
Max	984.3	896.3	147.6	1044.5	948.5	147.0	153.3	126.2	27.9
Number of models showing increases							9	10	6
Number of models showing decreases							11	10	14

(230.1–984.3 mm yr^{-1}) and very similar mean and standard deviation (556.2 ± 179.3 mm yr^{-1}). The projected future change, averaged over all models, was also very small (−8.1 mm yr^{-1}) within a large range (−109.8–153.3 mm yr^{-1}).

[18] The evaporation results are not as reliable for the reasons outlined previously, and even after our adjustment, there are small negative values of $P - E$ for a 30 year period in some model runs. With that caveat in mind, the all–model run runoff ($P - E$) for the period 1970–1999 varied from −1.6 to 156.4 mm yr^{-1} with an average of 44.6 mm yr^{-1} compared with the observed runoff of ≈27 mm yr^{-1} (M. L. Roderick and G. D. Farquhar, A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties, submitted to *Water Resources Research*, 2010). Of the 39 model runs examined, 16 showed increases in $P - E$ to the end of the 21st century while 23 showed decreases. The average change for the all–model run ensemble was for a very small decrease in MDB annual runoff of −0.5 mm yr^{-1} by the end of the 21st century. However, again we emphasize that the range (−20.9–32.5 mm yr^{-1})

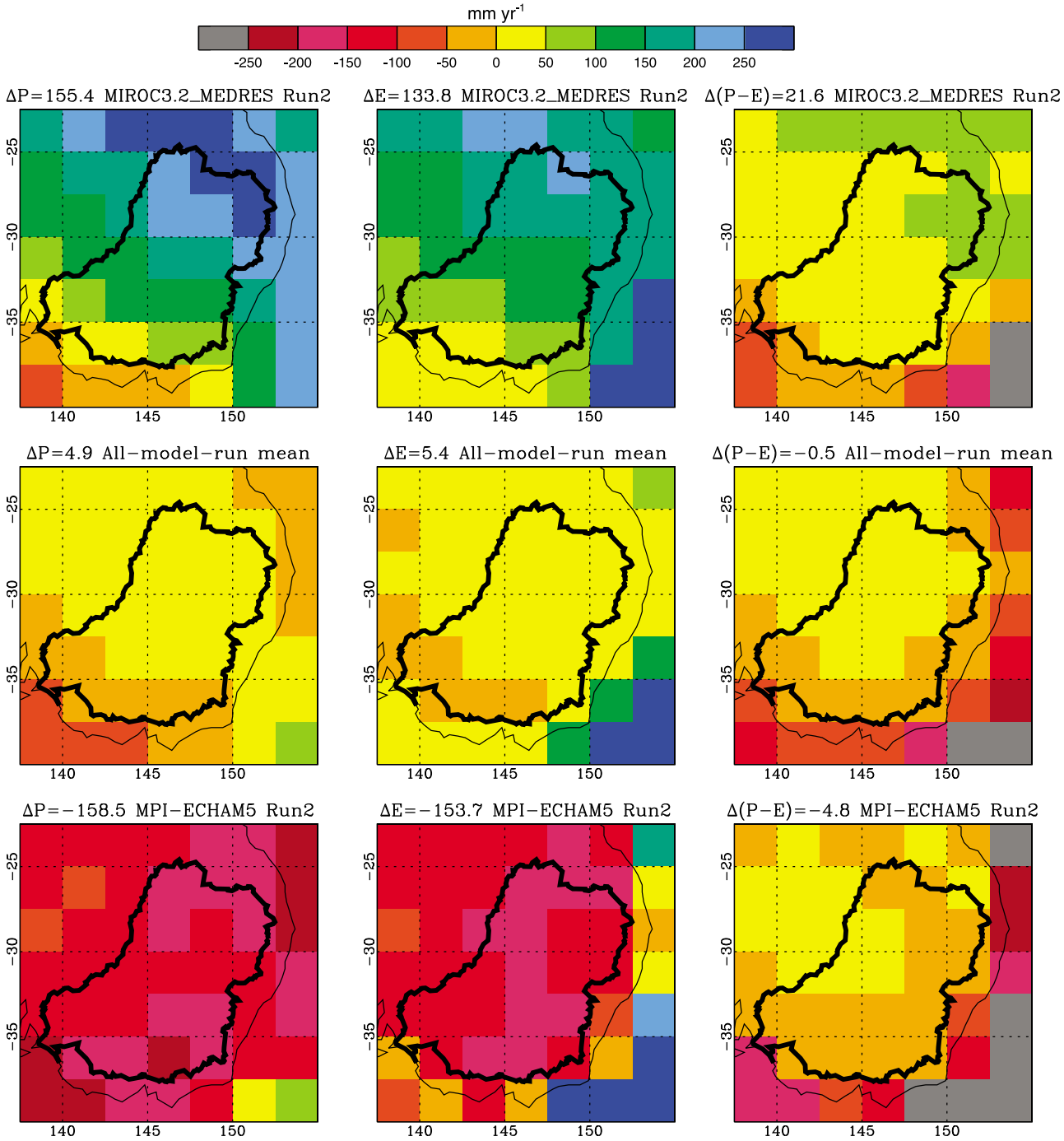


Figure 2. Changes of precipitation ΔP, evaporation ΔE, and runoff $\Delta(P - E)$ (2070–2099 less 1970–1999) for the MDB in the (top) wettest and (bottom) driest projections alongside the mean of the (middle) all–model run ensemble. The change integrated over the MDB (mm yr^{-1}) is shown at the top of each plot.

is large relative to the average change. The overall results for the all-model ensemble were more or less the same (Table 2).

[19] The time series (P, E) for all 39 model runs are shown in Figure S1 in Text S1 in the auxiliary material and document the diversity of simulations and projections described above.[1] That diversity is further explored in Figure 2, where we show changes in the grid box level hydrologic balance for the model run with the largest projected increase in P (MIROC3.2_MEDRES Run2, ΔP = 155.4 mm yr^{-1}) as well as the largest projected decrease (MPI-ECHAM5 Run2, $\Delta P = -158.5$ mm yr^{-1}) along with the mean of the all–model run ensemble. In summary, the raw outputs of the IPCC AR4 models show a large range of simulations and projections for the MDB. The overall conclusion about the hydrologic changes projected for the MDB (small change in ensemble mean but with large variation

[1]Auxiliary materials are available in the HTML. doi:10.1029/2010WR009829.

among individual runs/models) is the same as found for the whole of Australia [*Lim and Roderick*, 2009].

[20] Given the above noted land-ocean boundary problems in using simulations or projections for E (and $P - E$), we focus on P, the most important of the hydrologic variables in the MDB, in the remainder of this paper. In terms of P, the generic feature of the projections is that the average across all model runs or all models shows little change (4.9 and −8.1 mm yr^{-1}) within a very large range (up to ±150 mm yr^{-1}) by the end of the 21st century.

3.2. Precipitation Ensemble Using Raw Model Output

[21] The P ensemble (Figure 3) has been constructed using the raw model output for all 39 runs (Figure 3a). Note that the mean of the all–model run ensemble is virtually constant from year to year and is more or less the same as the mean of the all-model ensemble (Figures 3b and 3c and Table 2). The mean of the (all–model run or all-model) ensemble is larger than observed, and this is adjusted later using the bias correction approach described in section 3.3.

[22] Earlier work reported that the lag one (year) autocorrelation was ≈0 for both (annual) precipitation and runoff in all subbasins within the MDB [*Potter et al.*, 2008]. Here we extend that result with the finding that the autocorrelation in (annual) precipitation is also ≈0 for lags up to 36 years in both observations and climate model output for the MDB (Figure 3d). The implication is that over the last century, the MDB annual precipitation time series closely approximates a purely random time series (of a given variance). Hence, precipitation in a given year is (statistically) independent of that in the previous year(s). What is especially important is that (nearly) all of the 39 model runs also share the same basic characteristic (Figure 3d). Hence, from a statistical viewpoint, the model runs examined here also approximate purely random time series.

3.3. Precipitation Ensemble Using Bias-Corrected Output

[23] The bias evident in the raw output time series (Figure 3b and Figure S1 in Text S1) was removed using the quantile-based method [*Wood et al.*, 2004; *Li et al.*, 2010]. The before and after results for each of the 39 model runs are summarized in Table 3 and are shown in Figure S2 in Text S1. As anticipated, the variability in model simulations (for 1970–1999) is vastly reduced in the bias-corrected output (minimum 435.0, mean 486.6, and maximum 546.6 mm yr^{-1}) compared to the raw model output (minimum 230.1, mean 578.4, and maximum 994.4 mm yr^{-1}). The same holds for the 2070–2099 period. The statistical properties of the ensemble after bias correction are depicted in Figure 4. The overall reduction in the spread of model simulations and projections is clearly evident by comparing Figures 4a, 4b, and 4c with Figures 3a, 3b, and 3c. Importantly, the autocorrelation analysis shows the same basic pattern as found in the raw model output (compare Figures 4d and 3d).

[24] Despite the changes induced by the bias removal procedure, the statistics of the projected difference, ΔP, were more or less unchanged (Table 3). For example, in the all–model run ensemble, the statistics of ΔP calculated from raw model output (mean 4.9, σ = 68.3, minimum −158.5, and maximum 155.4 mm yr^{-1}) are for all practical purposes identical with that calculated after the bias correction procedure (mean 5.5, σ = 69.2, minimum −156.1, and maximum 158.1 mm yr^{-1}) (Table 3). The same holds for the all-model ensemble.

[25] Another key feature of the results is that the ensemble (either the all–model run or all-model) mean is more or less constant over time both before (Figure 3c) and after bias correction (Figure 4c). This is a very interesting result. A literal interpretation, of relevance only to the bias-corrected output, is that the ensemble average across a number of model runs (e.g., 39) in a given year is a good approximation to the corresponding average of the 30 year time series for a given model run.

[26] The interannual variance of each individual model run was, in some instances, changed markedly by the bias correction procedure (Table 4). However, the statistical properties of the all–model run ensemble of the variances in the simulations, projections, and projected changes were all, more or less, unchanged (Table 4).

[27] In summary, the bias correction scheme forced the ensemble to more or less replicate the statistical properties of the observations (Figure S2 in Text S1). The correction scheme did slightly alter the projections of ΔP for individual model runs (Table 3). However, those changes were small, and the statistical properties of the projections of ΔP (Table 3) were virtually unchanged by the bias correction procedure. We examine the variability in projections of ΔP further in section 3.4.

3.4. Understanding Variability in Future Projections of ΔP

[28] The finding, via the autocorrelation analysis, that the observed, simulated, and projected P time series are from a statistical viewpoint more or less random, has important implications. For example, as a base case, assume that the P time series were to remain a stationary time series into the foreseeable future. On that basis, we could anticipate differences between 30 year averages of P, for example, between the 1970–1999 and 2070–2099 periods, just by chance. How large could the differences be?

[29] To investigate the differences, we calculated the variance of the observed MDB annual time series. Then, by assuming a purely random process with that variance, we numerically generated a time series of 130 random numbers. Averages were taken over the first 30 numbers and the last 30, and the difference was taken. We repeated that process 10,000 times to simulate a statistical distribution for ΔP. In the initial calculations we used a normal distribution to generate the random number sequence, and the resulting distribution of ΔP was also normal. Subsequent tests with other assumed distributions for the random time series (e.g., uniform, gamma) showed that regardless of that choice, the resulting distribution of ΔP was always very close to normal. The results for the MDB (observed variance (1900–1999) 12,510 (mm yr^{-1})^2) were for a mean ΔP of 0 (as per the assumption) and for a variance of the ΔP distribution of 826 (mm yr^{-1})^2. The equivalent 2σ bound (95% confidence interval) was ∼±57 mm yr^{-1}, and the overall bounds (assessed at the 0.01% percentiles) were ∼±110 mm yr^{-1}. The resulting interpretation is that under the stationary

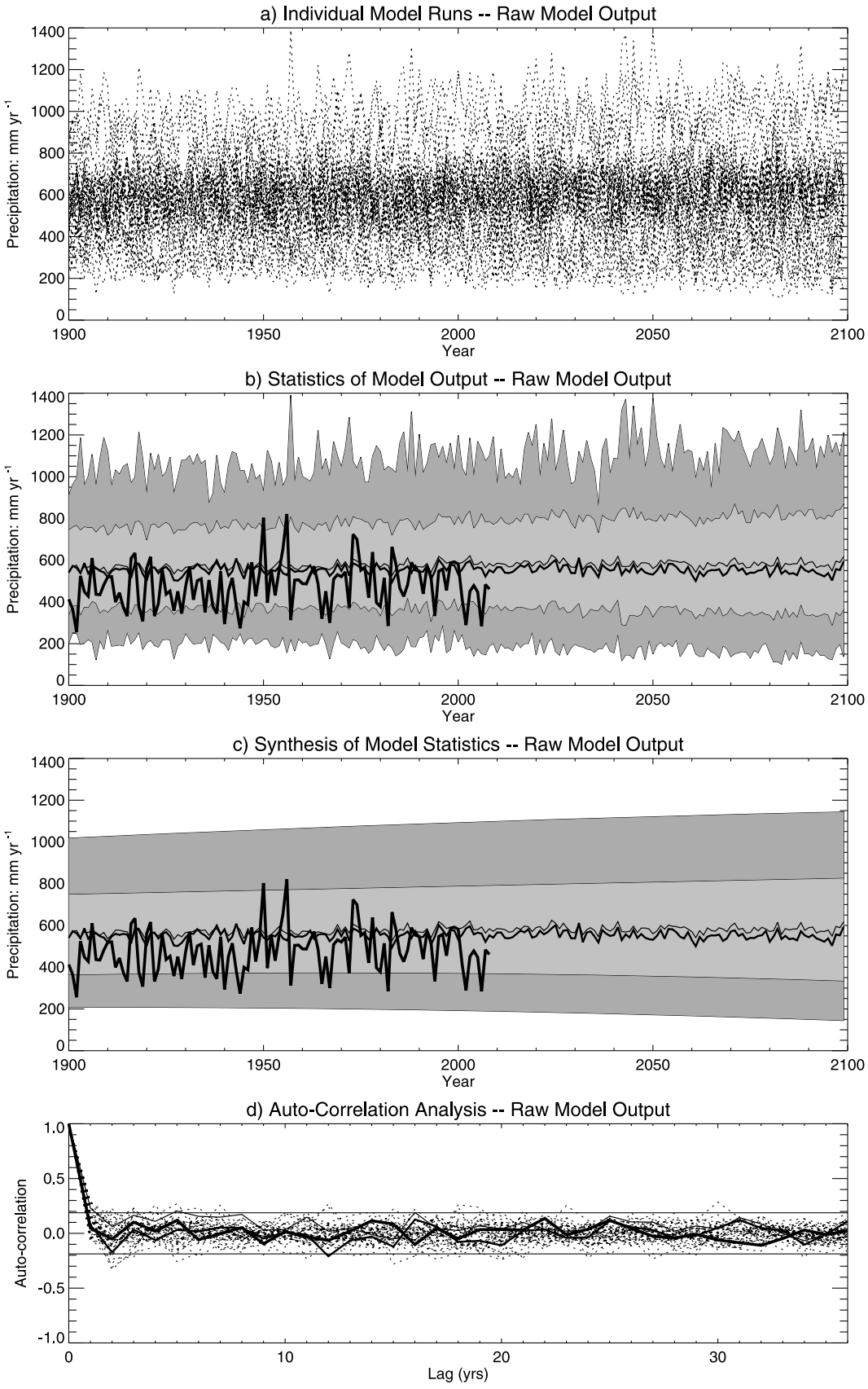

a) Individual Model Runs -- Raw Model Output
b) Statistics of Model Output -- Raw Model Output
c) Synthesis of Model Statistics -- Raw Model Output
d) Auto-Correlation Analysis -- Raw Model Output
Precipitation: mm yr-1
Year
Auto-correlation
Lag (yrs)
0
200
400
600
800
1000
1200
1400
1900
1950
2000
2050
2100
1.0
0.5
0.0
-0.5
-1.0
10
20
30

Figure 3

Table 3. MDB Precipitation Summary Before and After Bias Correction (mm yr^{-1})

	Raw Output			Bias-Corrected Output		
	1970–1999	2070–2099	Change	1970–1999	2070–2099	Change
BCCR-BCM2.0 Run1	867.1	908.2	41.1	497.2	544.7	47.5
CGCM3.1(t63) Run1	468.9	534.9	66.0	486.5	555.5	68.9
CNRM-CM3 Run1	478.5	473.0	−5.6	486.8	484.9	−2.0
CSIRO-Mk3.0 Run1	555.6	461.9	−93.7	515.1	420.9	−94.2
CSIRO-Mk3.5 Run1	458.9	349.1	−109.8	480.0	371.5	−108.5
GFDL-CM2.0 Run1	360.0	304.9	−55.1	499.4	440.2	−59.2
GISS-AOM Run1	284.1	227.9	−56.2	482.8	413.6	−69.3
GISS-AOM Run2	254.7	232.9	−21.8	451.6	428.7	−22.9
GISS-EH Run1	984.2	1062.8	78.5	506.6	582.4	75.7
GISS-EH Run2	994.4	1048.6	54.2	524.7	574.5	49.8
GISS-EH Run3	974.2	1022.2	48.0	497.8	548.8	51.0
GISS-ER Run6	641.7	681.3	39.5	469.3	502.7	33.4
GISS-ER Run8	643.6	654.2	10.6	455.0	461.1	6.1
INGV-ECHAM4 Run1	690.5	637.1	−53.4	506.2	462.6	−43.6
INM-CM3.0 Run1	517.2	517.9	0.7	446.2	458.0	11.8
IPSL-CM4 Run1	230.1	166.4	−63.7	464.4	393.9	−70.5
MIROC3.2_HIRES Run1	631.3	665.1	33.9	435.0	463.3	28.4
MIROC3.2_MEDRES Run2	680.0	835.4	155.4	457.8	615.9	158.1
MIROC3.2_MEDRES Run3	701.5	852.7	151.2	480.2	637.3	157.2
MIUB-ECHO-G Run1	540.4	598.0	57.7	490.2	543.6	53.4
MIUB-ECHO-G Run2	523.3	598.6	75.3	472.9	558.1	85.2
MIUB-ECHO-G Run3	532.8	626.0	93.2	505.7	601.6	95.8
MPI-ECHAM5 Run1	387.2	332.8	−54.4	478.8	419.2	−59.5
MPI-ECHAM5 Run2	489.4	331.0	−158.5	546.6	390.5	−156.1
MPI-ECHAM5 Run3	419.1	336.8	−82.3	505.7	419.1	−86.5
MPI-ECHAM5 Run4	357.4	378.7	21.2	470.0	495.6	25.6
NCAR-CCSM3.0 Run1	643.8	676.4	32.7	493.7	524.6	30.9
NCAR-CCSM3.0 Run2	658.9	667.7	8.9	501.4	514.6	13.2
NCAR-CCSM3.0 Run3	646.2	665.7	19.5	493.9	517.2	23.4
NCAR-CCSM3.0 Run5	634.5	666.6	32.1	493.0	522.1	29.1
NCAR-CCSM3.0 Run6	628.6	660.9	32.3	479.4	514.7	35.3
NCAR-CCSM3.0 Run7	617.3	675.5	58.1	490.0	546.9	57.0
NCAR-CCSM3.0 Run9	613.8	678.7	64.9	479.5	542.8	63.3
NCAR-PCM1 Run1	569.0	543.8	−25.2	489.9	469.2	−20.7
NCAR-PCM1 Run2	589.6	569.7	−20.0	503.6	480.4	−23.2
NCAR-PCM1 Run3	595.2	572.6	−22.6	510.7	490.4	−20.3
NCAR-PCM1 Run4	579.8	561.8	−18.0	508.7	495.5	−13.2
UKMO-HADCM3 Run1	525.0	457.8	−67.3	456.8	394.2	−62.6
UKMO-HADGEM1 Run1	590.2	514.0	−76.2	462.9	388.9	−74.0
		The All–Model Run Ensemble				
Mean across all runs	578.4	583.3	4.9	486.6	492.0	5.5
σ	176.2	217.7	68.3	22.9	68.3	69.2
Min	230.1	166.4	−158.5	435.0	371.5	−156.1
Max	994.4	1062.8	155.4	546.6	637.3	158.1
Number of runs showing increases				22		
Number of runs showing decreases				17		
		The All-Model Ensemble				
Mean across all models	556.2	548.1	−8.1	481.4	474.3	−7.1
σ	179.3	221.1	68.2	22.4	70.8	69.1
Min	230.1	166.4	−109.8	435.0	371.5	−108.5
Max	984.3	1044.5	153.3	515.1	626.6	157.6
Number of models showing increases				9		
Number of models showing decreases				11		

assumption employed here, we could expect 95% of all changes in P over successive 30 year periods in the MDB to be within ~±57 mm yr^{-1}, with a further 5% up to the outer bounds (~±110 mm yr^{-1}) of the distribution. Changes beyond that would immediately identify that the time series cannot be stationary.

3.5. Variability in Future Projection of ΔP in Climate Models Over the MDB

[30] Following the above example for the MDB, we can immediately see that the magnitude of the bounds (e.g., 2σ for 95% confidence or 0.01% for outer limits) must scale

Figure 3. Precipitation time series (1900–2099) for the MDB based on raw model output. (a) The 39 individual model runs. (b) Statistical summary of model output. The shaded background denotes the minimum, maximum, and 1σ range. The solid lines denote observations (1900–2008) (thick) and the mean of the all-model (medium) and all–model run (thin) ensembles. (c) Synthesis of model statistics. (d) Autocorrelation analysis for the 39 model runs (dotted lines) with the solid lines as per Figure 3b. The 95% confidence level (two straight thin lines) is also shown [*Brockwell and Davis*, 1987].

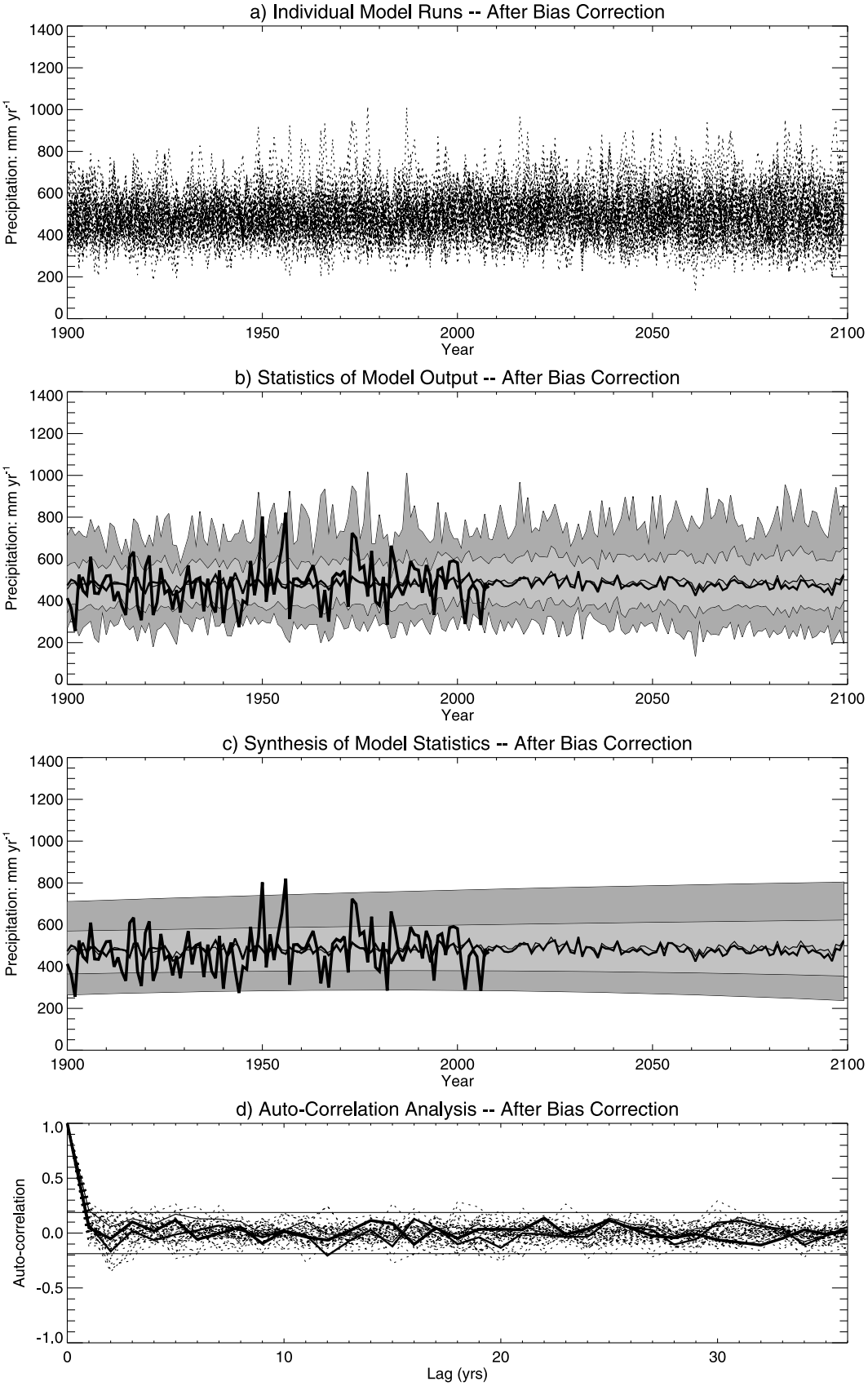


Figure 4. Analogous to Figure 3 after bias correction.

Table 4. Summary of Variance of Annual Precipitation Time Series Over the MDB From the Raw Output and After Bias Correction ($(\text{mm yr}^{-1})^2$)

	Raw Output			Bias-Corrected Output		
	1900–1999	2000–2099	Change	1900–1999	2000–2099	Change
BCCR-BCM2.0 Run1	12213	19076	6863	7535	12154	4619
CGCM3.1(t63) Run1	6835	8746	1911	9696	12714	3019
CNRM-CM3 Run1	20517	19790	−727	23995	25307	1312
CSIRO-Mk3.0 Run1	15569	24060	8491	16726	26275	9550
CSIRO-Mk3.5 Run1	22779	13510	−9269	25098	15314	−9784
GFDL-CM2.0 Run1	9106	12107	3001	10110	14030	3919
GISS-AOM Run1	3527	2519	−1009	9575	7226	−2349
GISS-AOM Run2	3818	3680	−139	10466	11265	799
GISS-EH Run1	12047	15465	3419	8991	12129	3139
GISS-EH Run2	14346	13674	−672	11350	10919	−430
GISS-EH Run3	14771	14653	−119	11800	11879	79
GISS-ER Run6	7512	9814	2302	8970	10275	1305
GISS-ER Run8	8800	8747	−53	10762	10025	−737
INGV-ECHAM4 Run1	17059	10054	−7005	11522	7668	−3855
INM-CM3.0 Run1	8326	8792	466	12547	13576	1029
IPSL-CM4 Run1	1974	1783	−190	7263	6435	−828
MIROC3.2_HIRES Run1	10995	11790	795	8798	10574	1776
MIROC3.2_MEDRES Run2	11951	16948	4998	12211	17244	5033
MIROC3.2_MEDRES Run3	13582	16610	3028	12773	16984	4211
MIUB-ECHO-G Run1	6310	6367	58	9429	9027	−402
MIUB-ECHO-G Run2	5189	8991	3801	8307	13235	4928
MIUB-ECHO-G Run3	5806	6581	775	9082	10511	1429
MPI-ECHAM5 Run1	11710	19287	7577	13795	23682	9887
MPI-ECHAM5 Run2	22704	22333	−370	21299	21338	39
MPI-ECHAM5 Run3	18662	17906	−756	20410	20388	−22
MPI-ECHAM5 Run4	10011	17208	7197	13211	23566	10355
NCAR-CCSM3.0 Run1	5726	8950	3223	5458	8697	3238
NCAR-CCSM3.0 Run2	8900	9286	386	8198	9644	1446
NCAR-CCSM3.0 Run3	6115	8774	2659	5431	8600	3168
NCAR-CCSM3.0 Run5	8302	8591	290	8034	8237	203
NCAR-CCSM3.0 Run6	9465	7429	−2035	9441	6886	−2555
NCAR-CCSM3.0 Run7	5872	7113	1241	6154	7282	1128
NCAR-CCSM3.0 Run9	8157	8754	598 7	699	8712	1013
NCAR-PCM1 Run1	8563	9344	781	9519	11331	1811
NCAR-PCM1 Run2	10822	9495	−1327	11322	10584	−738
NCAR-PCM1 Run3	11340	8286	−3054	12585	10107	−2477
NCAR-PCM1 Run4	8915	8120	−795	10605	10271	−334
UKMO-HADCM3 Run1	13808	13657	−151	12468	13476	1008
UKMO-HADGEM1 Run1	16938	12792	−4146	10640	8293	−2348
		The All–Model Run Ensemble				
Mean across all runs	10061	10721	660	10610	11820	1211
σ	4638	4803	2886	3426	4689	3125
Min	1974	1783	−7005	5431	6435	−3855
Max	22704	22333	7577	21299	23682	10355
Number of runs showing increases				27		
Number of runs showing decreases				12		
Observation	12510					

with the variance of the original time series. This has profound consequences for interpreting changes in the individual model runs. In particular, visual inspection of the raw output time series for each of the 39 model runs (Figure S1 in Text S1) shows that the variance in the model runs can be as much as double (e.g., CSIRO-Mk3.5 Run1 and MPI-ECHAM5 Run2) or as little as a quarter (e.g., IPSL-CM4 Run1 and GISS-AOM Run1 and Run2) of the observed variance. To test further, we calculated the variance of the annual time series for each of the 39 model runs in both the 20th and 21st centuries (Table 4). The results show the large range in the variance of the raw model output relative to observations over the 20th century.

[31] Using those insights, we prepared a plot showing the projected change ΔP (2070–2099 less 1970–1999) for each of the 39 model runs versus the variance of that model run for the 1900–1999 period. Overlaid on that plot are calculations (per the above description) of the statistical bounds ($\pm 1\sigma$, $\pm 2\sigma$, $\pm 3\sigma$, maximum, and minimum) assuming a stationary distribution (Figure 5). Key features, some alluded to previously, emerge immediately. For example, the results for model 12, (one run of) model 16, (two runs of) model 14, and (two runs of) model 15 all fall outside the outer bounds for a random stationary process. The statistical characteristics of those time series have changed substantially, and those projections cannot be considered stationary.

[32] A closer examination of the above noted model runs is warranted (see Figure S1 in Text S1). The time series for model 12 (IPSL-CM4) shows a long steady downward decline in the mean with little change in variability around that trend. This model run violated the stationary assumption because the mean changed. However, the variability in the model run is much smaller than observed over the historical period, and the model simulation is not convincing.

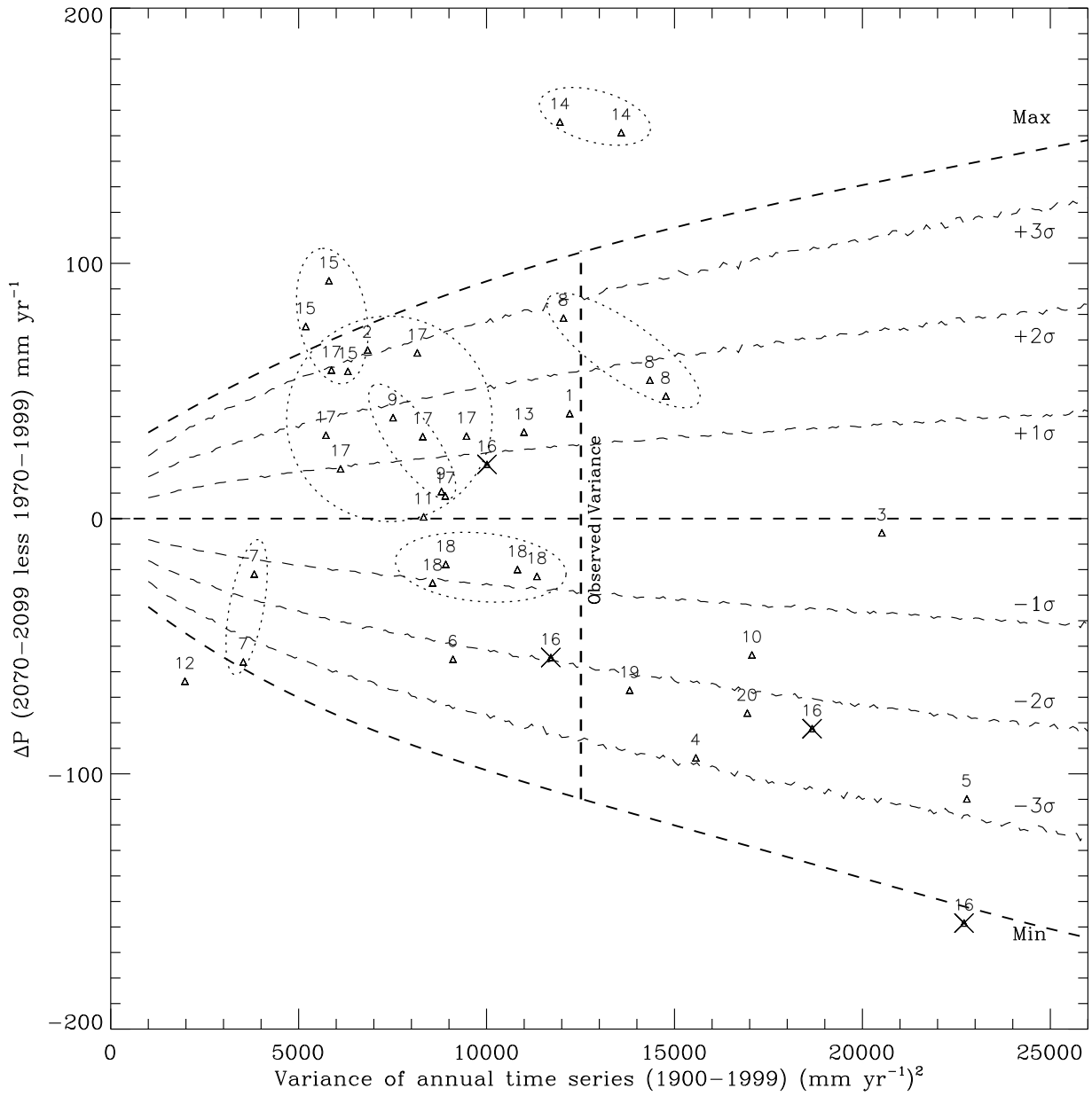


Figure 5. Projected change of precipitation ΔP over the MDB (2070–2099 less 1970–1999) for the 39 model runs versus the variance of the annual time series (1900–1999) for each model run. All calculations use raw model output. Each triangle denotes one model run with numbering of models per Table 1. The variance of the observed time series (12510 (mm yr^{-1})^2) (Table 4) is denoted by the vertical line. The curves ($\pm1\sigma$, $\pm2\sigma$, $\pm3\sigma$) were generated numerically assuming a random stationary time series. Max is defined as the 99.99th percentile and Min is defined as the 0.01st percentile. The dotted ellipse highlights models with multiple runs, with the exception that multiple runs for model 16 are marked by a cross.

The time series for both runs of model 12 (MIROC3.2_MEDRES) show that the stationary assumption was violated because of a marked upward trend in both mean and variance. The time series for (two runs of) model 15 (MIUB-ECHO-G) show that the stationary assumption was violated because of a steady increase in the mean with little change in variability. The contrast here is interesting; each model run had readily understood reasons for violating the stationary assumption. However, the reasons were different. The remaining (one run of) model 16 (MPI-ECHAM5 Run2) is an enigma. This model run, projected the largest of all decreases in ΔP, and examination of the time series (Figure S1 in Text S1) shows a marked decline in the mean with perhaps a slight increase in variability around the mean. The enigma is that each of the four model runs gave markedly different results. In contrast, multiple runs from the other seven models tend to cluster together (Figure 5).

[33] The 2σ bounds approximate the 95% confidence interval. Using that as a guide, we can also identify many other model runs that are unlikely to be stationary, including models 2, 4, 5, and 6, (one run of) model 7, (one run of) model 8, (one run of) model 15, (two runs of) model 16, (two runs of) model 17, and models 19 and 20. In summary, the projected change ΔP falls outside the bounds of an

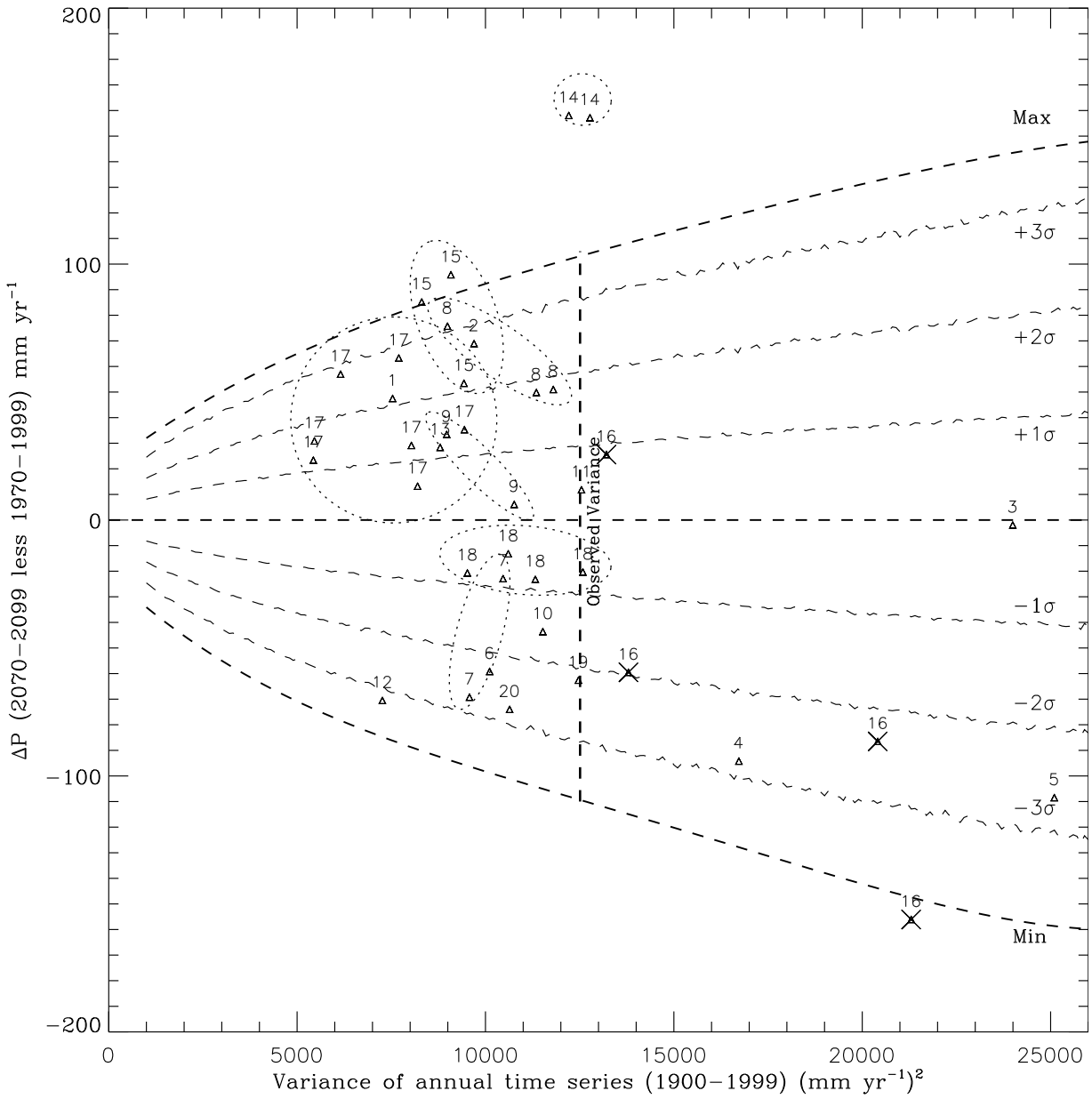


Figure 6. Analogous to Figure 5 after bias correction.

assumed stationary process in six model runs and outside the $\pm 2\sigma$ range in a further 13 model runs. The remaining 20 model runs fall within $\pm 2\sigma$ bounds. Given the previous results for the autocorrelation analysis, it would be difficult to distinguish those time series from one generated by a purely random process.

[34] The single runs of the two CSIRO models present an interesting case study because both models projected large decreases in P that approximate the $\pm 3\sigma$ level, implying that the projected time series has changed substantially. Results for model 5 (CSIRO-Mk3.5) show a large decrease in variability about the mean (Figure S1 in Text S1 and Table 4). However, model 4 (CSIRO-Mk3.0) shows the opposite with a large increase in variability about the mean (Figure S1 in Text S1 and Table 4). The central difficulty here is that both models have a variance during the historical period that is larger than observed. With only a single run available, it is difficult to come to any firm conclusions. More recent research with a slightly different variant of the Mk3.0 model has shown quite large differences in ΔP between multiple runs [*Rotstayn et al.*, 2007, Figure 19].

[35] The bias correction procedure changed the variance of the annual time series in the climate model output (Table 4). We also used the bias-corrected model outputs to prepare Figure 6 analogous to Figure 5. While many of the details of the plot are different (compare Figure 5), the overall pattern and the general conclusions drawn above remain.

4. Discussion

[36] A previous study on the MDB evaluating the capacity of different IPCC AR4 climate models to simulate temperature and precipitation found the IPSL-CM4 and CSIRO-Mk3.0 to be the best overall models [*Maxino et al.*,

2008]. Of those, the IPSL-CM4 model was ranked the best for precipitation [*Maxino et al.*, 2008, Table II]. We used the difference-variance framework to investigate the P annual time series in the (single) IPSL-CM4 model run in considerable detail. We found that this particular time series was not very convincing. The mean P was less than the half the observed value, and more importantly, there was little year-to-year variation in P: the variance was around 1/6 of the observed value (Table 4 and Figure S1 in Text S1). The totally different conclusions highlighted above arise because the earlier work considered monthly, i.e., intra-annual, P, while we considered annual P, i.e., the interannual variation.

[37] Of the 20 models examined here, 8 had multiple runs available. Of those, the results for 7 models show that while there was some variation between different runs, the overall projections still tended to converge in a "region" of the difference-variance plot (Figure 5). There was an exception: multiple runs from the MPI-ECHAM5 model (model 16 in Figures 5 and 6) diverged. That exception raises an important point. Would any of the 12 models having a single run behave the same as the MPI-ECHAM5 model if multiple runs were submitted? Of course, we do not know the answer. In that respect, we believe that there is, at least at the moment, some reason to be cautious about overinterpreting the results from single runs of a climate model.

5. Summary and Conclusions

[38] The results presented here are based on 39 model runs from 20 different IPCC AR4 climate models (Table 1). For each of the 39 model runs, the main results are all derived from annual P time series for the historical period (1900–1999) and for a future (2000–2099) that follows the IPCC A1B emissions scenario. Other emission scenarios could have been used, but here we pursued an understanding of the statistical nature of the ensembles and of the simulations and projections for P. Our main focus was the projected change ΔP (defined as the difference 2070–2099 less 1970–1999).

[39] Of the 20 models, 12 contributed a single run, while two or more runs were available from the other 8 models. The model population was extremely diverse, with 7 (of the 39) runs being contributed by a single model. It was thus possible that a simple average over all runs would be biased toward those models contributing the most runs. To investigate this possibility, we created two P ensembles. The first, called the all–model run ensemble, included all 39 model runs. The second was formed by averaging the multiple runs (where necessary) across each model to create a 20 member all-model ensemble. Despite the heterogeneous nature, the statistical properties of the simulations and projections for P were, for all practical purposes, identical for both ensembles (Table 2). On that basis, we only summarize results from the all–model run ensemble.

[40] The range in P among the raw model output was large (Figure 3a), with obvious bias relative to the observed time series (Figures 3b and 3c). After bias correction, the climatological range was much reduced (Figures 4b and 4c). Finally, while the bias correction did change the simulation (1970–1999) and projection (2070–2099) for P, it did not materially alter the projected change ΔP (Table 3).

[41] The main findings (Figures 5 and 6) are derived from the autocorrelation analysis (Figures 3d and 4d). We found that the observed annual P time series for the MDB could be considered to be a purely random time series with no time dependence at any of the (time) lags considered (Figure 3d). Just as importantly, we found that (nearly) all of the annual time series from the 39 model runs also shared the same basic characteristic. We emphasize that these results held in both observations and models, both before (Figure 3d) and after (Figure 4d) bias correction. The consequences are important.

[42] First, by assuming that the MDB annual P time series remains stationary into the future, one can estimate the probabilistic variations in ΔP that result solely from random fluctuations. This provides an extremely useful base case against which to assess the projections for ΔP by climate models. Second, it is important to consider the variance of the model time series when comparing different model estimates of ΔP because the magnitude of random fluctuations scales with the variance. In particular, the variance in individual model runs for the historical period was up to twice, or as little as 1/6, the observed value. Hence, it is virtually impossible to interpret differences in ΔP between different model runs without considering the variance of each model run. To address this situation, we developed the "difference-variance" plot (Figures 5 and 6). This enabled us to rapidly identify those model runs showing large changes in ΔP relative to their variance.

[43] What does all this imply for projections of ΔP in the MDB? In terms of the all–model run ensemble, there was a large range ($\sim\pm 150$ mm yr^{-1}) in projected change ΔP (Table 2 and Figures 5 and 6). When the projections of ΔP are averaged over all model runs, the result is, more or less, zero change (Table 2). One contribution arising from this work is the "base case" scenario. For the MDB, we anticipate that a purely random process could lead to differences of ± 57 mm yr^{-1} (95% confidence) between successive 30 year periods. This is equivalent to $\pm 11\%$ of the climatological P and with all else constant, translates into variations in runoff of around $\pm 29\%$ (~ 7.7 mm yr^{-1} on a catchment-wide basis and equivalent to ~ 7700 GL yr^{-1}) (Roderick and Farquhar, submitted manuscript, 2010). This sets a baseline for gauging modeled and/or observed changes.

Appendix A: Adjustment for Evaporation in Mixed Land-Ocean (Water) Grid Boxes

[44] For mixed grid boxes with land and ocean or lake, the 100 year average (e.g., 1900–1999, 2000–2099) of $P - E$ in most climate models initially was negative. As a consequence, the evaporation estimates for the land component of mixed grid boxes required adjustment for both periods of 1900–1999 and 2000–2099 to make the results physically realistic. The procedure adopted is described below.

[45] Generally, for every grid box,

$$\begin{aligned} E &= \alpha_f \times E_L + \left(1 - \alpha_f\right) \times E_O \\ P &= \alpha_f \times P_L + \left(1 - \alpha_f\right) \times P_O, \end{aligned} \tag{A1}$$

where E, E_L, and E_O are annual evaporation for the whole grid box, land area (whose fractional area is α_f), and ocean area of the grid box, respectively; P, P_L, and P_O are annual precipitation with the same meaning. Here, we assume, with the same respective meanings, $P = P_L = P_O$.

[46] In the initial step, we set $E = E_L = E_O$. For the grid boxes where $\alpha_f > 0$ and $E_L > P_L$, we set

$$E_L' = E_L \frac{\overline{P_L}}{\overline{E_L}}, \tag{A2}$$

where E_L' is the adjusted land evaporation for the grid box and $\overline{P_L}$ and $\overline{E_L}$ are the 100 year average annual land evaporation and precipitation for the grid box, respectively. For those grid boxes, the ocean evaporation will be

$$E_O' = \frac{E - \alpha_f \times E_L'}{1 - \alpha_f}. \tag{A3}$$

Then for the whole adjusted period (1900–1999) we have

$$\overline{E_L'} = \overline{E_L \frac{\overline{P_L}}{\overline{E_L}}} = \overline{E_L} \frac{\overline{P_L}}{\overline{E_L}} = \overline{P_L} = \overline{P}. \tag{A4}$$

Note that for a different period, e.g., 1970–1999, the averaged and adjusted evaporation $\overline{E_L'}$ is not necessarily equal to $\overline{P_L}$.

[47] **Acknowledgments.** We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI), and the WCRP's Working Group on Coupled Modeling (WGCM) for their roles in making available the WCRP CMIP3 multimodel data set. Support of this data set is provided by the Office of Science, U.S. Department of Energy. This research was supported by the Murray-Darling Basin Authority (contract MD1318) and by the Australian Research Council (DP0879763).

References

Brockwell, P. J., and R. A. Davis (1987), *Time Series: Theory and Methods*, Springer, New York.

Chiew, F. H. S., J. Teng, J. Vaze, D. A. Post, J. M. Perraud, D. G. C. Kirono, and N. R. Viney (2009), Estimating climate change impact on runoff across southeast Australia: Method, results, and implications of the modeling method, *Water Resour. Res.*, *45*, W10414, doi:10.1029/2008WR007338.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007), Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, *27*, 1547–1578, doi:10.1002/joc.1556.

Groves, D. G., D. Yates, and C. Tebaldi (2008), Developing and applying uncertain global climate change projections for regional water management planning, *Water Resour. Res.*, *44*, W12413, doi:10.1029/2008WR006964.

Held, I., and B. Soden (2006), Robust responses of the hydrological cycle to global warming, *J. Clim.*, *19*, 5686–5699.

Intergovernmental Panel on Climate Change (IPCC) (2000), Special report on emission scenarios, The Hague, Netherlands. (Available at www.ipcc.ch)

Intergovernmental Panel on Climate Change (IPCC) (2007), *Climate Change* 2007*: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., Cambridge Univ. Press, Cambridge, U. K.

Johnson, F., and A. Sharma (2009), Measurement of GCM skill in simulating variables relevant for hydroclimatological assessments, *J. Clim.*, *22*, 4373–4382, doi:10.1175/2009JCLI2681.1.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G.A. Meehl (2010), Challenges in combining projections from multiple models, *J. Clim.*, *23*, 2739–2758.

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

Lim, W. H., and M. L. Roderick (2009), *An Atlas of the Global Water Cycle: Based on the IPCC AR*4 *models*, ANU E Press, Canberra. (Available at http://epress.anu.edu.au/global_water_cycle_citation.html)

Maurer, E. P., and H. G. Hidalgo (2008), Utility of daily vs. monthly large scale climate data: An intercomparison of two statistical downscaling methods, *Hydrol. Earth Syst. Sci.*, *12*, 551–563.

Maxino, C. C., B. J. McAvaney, A. J. Pitman, and S. E. Perkins (2008), Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation, *Int. J. Climatol.*, *28*, 1097–1112.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multimodel data set: A new era in climate change research, *Bull. Am. Meteorol. Sci.*, *88*, 1383–1394, doi:10.1175/BAMS-88-9-1383.

Murray-Darling Basin Commission (2009), Annual report 2008–2009, *Publ.* 41/09, Canberra. (Available at http://www.mdba.gov.au/MDBA-Annual-Report/index.html)

Potter, N. J., F. H. S. Chiew, A. J. Frost, R. Srikanthan, T. A. McMahon, M. C. Peel, and J. M. Austin (2008), Characterisation of recent rainfall and runoff in the Murray-Darling Basin, a report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project, 40 pp., Commonw. Sci. and Ind. Res. Organ., Clayton, Victoria, Australia.

Potter, N. J., F. H. S. Chiew, and A. J. Frost (2010), An assessment of the severity of recent reductions in rainfall and runoff in the Murray-Darling Basin, *J. Hydrol.*, *381*, 52–64, doi:10.16/j.jhydrol.2009.11.025.

Rotstayn, L. D., et al. (2007), Have Australian rainfall and cloudiness increased due to the remote effects of Asian anthropogenic aerosols?, *J. Geophys. Res.*, *112*, D09202, doi:10.1029/2006JD007712.

Rudolf, B., and U. Schneider (2005), Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations, paper presented at 2nd Workshop, Int. Precip. Working Group, Monterey, Calif.

Rudolf, B., A. Becker, U. Schneider, A. Meyer-Christoffer, and M. Ziese (2010), GPCC status report July 2010 (On the most recent gridded global data set issued in fall 2010 by the Global Precipitation Climatology Centre (GPCC)), Global Precip. Climatol. Cent., Offenbach, Germany. (Available at http://gpcc.dwd.de)

Smith, I., and E. Chandler (2010), Refining rainfall projections for the Murray Darling Basin of south-east Australia—The effect of sampling model results based on performance, *Clim. Change*, *102*, 377–393, doi:10.1007/s10584-009-9757-1.

Sun, F., M. L. Roderick, G. D. Farquhar, W. H. Lim, Y. Zhang, N. Bennett, and S. H. Roxburgh (2010), Partitioning the variance between space and time, *Geophys. Res. Lett.*, *37*, L12704, doi:10.1029/2010GL043323.

Tebaldi, C., and R. Knutti (2007), The use of the multimodel ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. A*, *365*, 2053–2075, doi:10.1098/rsta.2007.2076.

Wilby, R. L., and I. Harris (2006), A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, *Water Resour. Res.*, *42*, W02419, doi:10.1029/2005WR004065.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004), Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Clim. Change*, *62*, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.

G. D. Farquhar, W. H. Lim, M. L. Roderick, and F. Sun, Research School of Biology, Australian National University, Canberra, ACT 0200, Australia. (michael.roderick@anu.edu.au)