



The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin

B. L. Harding¹, A. W. Wood², and J. R. Prairie³

¹AMEC Environment & Infrastructure, Boulder, CO, USA

²NOAA, National Weather Service, Portland, OR, USA

³Bureau of Reclamation, Boulder, CO, USA

Correspondence to: B. L. Harding (ben.harding@amec.com)

Received: 4 December 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 17 January 2012

Revised: 28 September 2012 – Accepted: 1 October 2012 – Published: 5 November 2012

Abstract. The impact of projected 21st century climate conditions on streamflow in the Upper Colorado River Basin was estimated using a multi-model ensemble approach wherein the downscaled outputs of 112 future climate projections from 16 global climate models (GCMs) were used to drive a macroscale hydrology model. By the middle of the century, the impacts on streamflow range, over the entire ensemble, from a decrease of approximately 30 % to an increase of approximately the same magnitude. Although prior studies and associated media coverage have focused heavily on the likelihood of a drier future for the Colorado River Basin, approximately 25 to 35 % of the ensemble of runs, by 2099 and 2039, respectively, result in no change or increases in streamflow. The broad range of projected impacts is primarily the result of uncertainty in projections of future precipitation, and a relatively small part of the variability of precipitation across the projections can be attributed to the effect of emissions pathways. The simulated evolution of future temperature is strongly influenced by emissions, but temperature has a smaller influence than precipitation on flow. Period change statistics (i.e., the change in flow from one 30-yr period to another) vary as much within a model ensemble as between models and emissions pathways. Even by the end of the current century, the variability across the projections is much greater than changes in the ensemble mean. The relatively large ensemble analysis described herein provides perspective on earlier studies that have used fewer scenarios, and suggests that impact analyses relying on one or a few climate scenarios are unacceptably influenced by the choice of projections.

1 Introduction

The US Secretary of the Interior, acting through the Bureau of Reclamation (Reclamation), is responsible for the operation of eight major water storage reservoirs on the Colorado River and its upper tributaries, including Glen Canyon Dam and Hoover Dam. The Boulder Canyon Project Act of 1928 designated the Secretary as water master, responsible for distributing all Colorado River water below Hoover Dam in conformance with federal law, water delivery contracts and the international agreement with Mexico. Through operation of federal facilities and administration of water deliveries, Reclamation contributes to the management of water supplies for more than 30 million people and nearly 4 million acres of agricultural land. Reclamation is continuously conducting long- and short-term water resources analyses to support planning and operations in the Colorado River Basin. In January 2010, Reclamation initiated the *Colorado River Basin Water Supply and Demand Study* (Reclamation, 2011a), which is being conducted with water management agencies representing the seven Colorado River Basin states (Arizona, California, Colorado, Nevada, New Mexico, Utah and Wyoming.) The work described in this paper was conducted in support of the study.

Hydrology and climate change in the Colorado River Basin

An extensive record of streamflows and reconstructed natural flows exists in the Colorado River Basin, but the value of that record as the sole basis for estimating future conditions has come into question because of the prospect that

anthropogenic climate change will change the mean and variability of streamflows and evapotranspiration (Milly et al., 2008). Global climate model (GCM) projections of future climate over a multi-decadal time frame indicate that the Colorado River Basin will become warmer. Projections of future precipitation are more complex, with the multi-model average of projections showing little change in annual precipitation in the water-producing regions of the basin, but generally showing a seasonal shift in the temporal pattern of precipitation. Changes in temperature and precipitation will influence hydrologic processes on the land surface, which in turn will cause changes in streamflows (Hayhoe et al., 2004; Barnett et al., 2005; Maurer, 2007) but the magnitude and the sign of these changes is uncertain.

The hydroclimatic heterogeneity of the Colorado River Basin complicates understanding its hydrology and hydrologic response to projected changes in climate. The latitude of the water-producing regions in the basin lies at the northern boundary of the area in the American Southwest in which earlier studies have projected declines in precipitation and runoff (Milly et al., 2005; Seager et al., 2007; Seager and Vecchi, 2010). To the north of this boundary, in contrast, studies have tended to predict increases in rainfall and runoff. If the future boundary instead falls south of the Colorado River Basin headwaters (water producing regions), streamflows in the Upper Colorado River Basin will be reduced less than projected, or may increase. Given the large uncertainty over future climate evolution at the scale of this transition zone, advances in climate science (perhaps including higher-resolution earth system models) will be required before these consequences can be projected with much confidence (Seager and Vecchi, 2010). The topography of the basin also influences the interaction between climate and hydrology – the majority of precipitation in the basin falls in a very small fraction of its area (Reclamation, 2011b), mostly in the form of snow that is stored over seasonal time scales. Snow accumulation and ablation are the most significant processes affecting the timing of streamflow generation in the basin, and the dynamics of the snow cycle will be significantly impacted by rising temperatures regardless of changes in precipitation. Our comprehension of variations in snow ablation is currently grappling with quantifying uncertainties due to sporadic deposition of dust on the surface of the snowpack, which accelerates snowmelt and affects basin runoff efficiency (Painter et al., 2007, 2010). The diversity of topography in the basin is another complicating factor. Outside of the cold, high-elevation areas, evapotranspiration is the dominant process affecting the water budget. Basin efficiencies (outflows divided by precipitation) range from about 30 % for sub-basins at higher average elevations to virtually zero in lower-elevation sub-basins (Reclamation, 2011b). The average runoff efficiency of the entire basin is approximately 15 %. These factors, among others, contribute to the difficulty of quantifying and simulating the hydrology and hydrologic sensitivity of the basin to climate forcings.

Studies of the impact of climate change on the hydrology of the basin began about 30 yr ago. Early, scenario-based assessments of the regional impact of climate change indicated that a warming climate would likely lead to streamflow reductions of 30 % or more in the Colorado River Basin (Stockton and Boggess, 1979; Revelle and Waggoner, 1983). Recent regional-scale analyses based on the GCM outputs also promoted an unequivocal consensus that the southwestern United States would become drier. Milly et al. (2005) evaluated projected change in runoff simulated by GCMs. Seager et al. (2007) and Seager and Vecchi (2010) inferred changes in runoff from changes in atmospheric transport of moisture simulated by GCMs. Seager and Vecchi (2010) analyze precipitation minus evaporation ($P - E$) as a proxy for runoff in southwestern North America, which extends from the high plains to the Pacific Ocean and from the latitude of the Oregon-California border to southern Mexico. They project that the region will dry in the coming century, largely driven by a decline in winter precipitation caused by a pole-ward shift of the winter Pacific storm track. In contrast, the Reclamation Study found that the ensemble of downscaled climate projections trends toward increases in wintertime precipitation for sub-basins in the Colorado River Basin above Lees Ferry (Reclamation, 2011b, Appendix B7).

Studies that utilized hydrologic models to translate projections of temperature and precipitation from climate models into streamflows show more equivocal results than the large-scale studies based on GCM atmospheric output fields. Nash and Gleick (1991) used a conceptual hydrologic model to evaluate the effect on Colorado River flow at Lees Ferry, AZ (a major stream gage just above the bottom of the Upper Colorado River Basin; see Fig. 1), of changes in precipitation and temperature in several constructed scenarios and as projected by four global climate models (GCMs). Three of the four GCM projections led to decreases in streamflow (as much as 24 %) while the impact from the fourth projection was to leave streamflow virtually unchanged. When evaluating impact on smaller basins within the Upper Colorado River Basin using a second hydrologic model, Nash and Gleick found that higher-elevation basins might be more likely to respond to some projections of future climate with an increase in flow. As in many studies of climate change impacts in the western US, all of the scenarios and projections indicated that peak runoff would occur earlier in the year.

Christensen et al. (2004; hereafter C04) used an ensemble of three projections of future climate from a single GCM based on the “business as usual” emissions pathway. These projections were statistically downscaled into 1/8-th degree, daily forcings using the bias corrected and spatial-disaggregation (BCSD) method outlined in Wood et al. (2002, 2004) (hereafter W02-04) and were subsequently translated into streamflows using the Variable Infiltration Capacity (VIC) hydrology model (Liang et al., 1994, 1996). Projected climate in that study exhibited reductions in precipitation and increases in temperature, leading to reductions

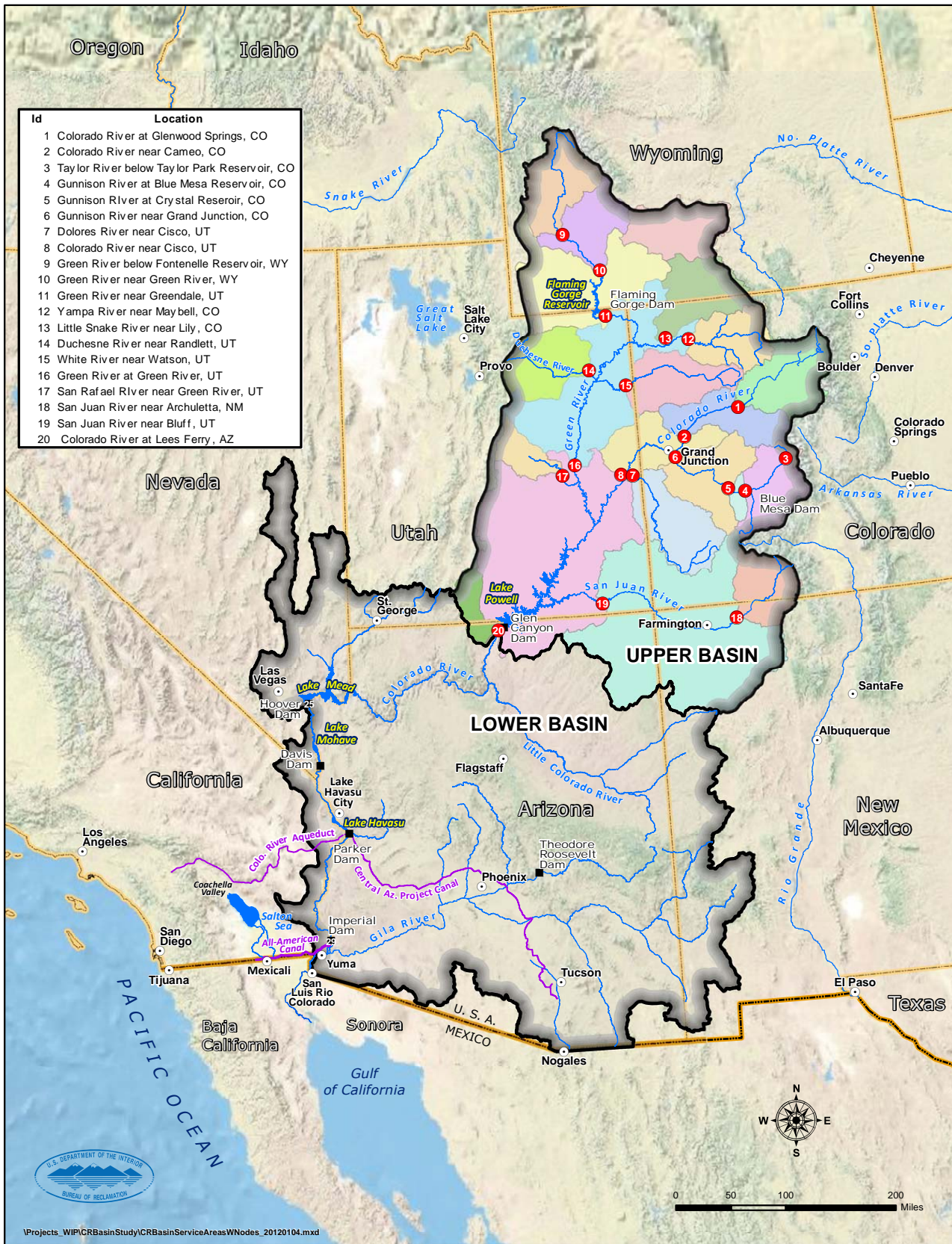


Fig. 1. Colorado River Basin and CRSS inflow points above Lees Ferry, Arizona.

in annual runoff at Imperial Dam, near the bottom of the Colorado River Basin, of 17 % over the period 2070–2098. The more detailed spatial and process land surface representation of C04 indicated that the sensitivity of the Colorado River Basin to changes in precipitation was roughly twice what had been indicated by Nash and Gleick, but confirmed the earlier conclusion that peak runoff would occur earlier in the year.

Using the same technique as C04 but expanding the range of climate projections to the ensemble described in Maurer (2007), and updating the hydrology model calibration to that used in Wood and Lettenmaier (2006), Christensen and Lettenmaier (2007; hereafter CL07) used 22 climate projections from 11 GCMs and two emissions pathways (A2 and B1; Nakicenovic et al., 2000) to assess climate impacts on hydrology and water resources in the Colorado River Basin. With a larger number of projections came a larger range of projected changes in flow at Imperial Dam ranging from an increase (five projections) of as much as 23 % to a decrease (seventeen projections) of as much as 36 % for both emissions pathways for the period 2070–2099.

The prior studies also show that the range of uncertainty in the future flow projections has increased with increasing numbers of projections evaluated. This observation is the primary motivation for this paper. Using a larger ensemble of projections than those featured in prior studies, we explore the relative contributions of uncertainty between different emissions pathways, ensemble members and climate models, across several periods of the 21st century. This effort does not describe in detail the findings of the overall climate change impact assessment study because they have been reported in Reclamation (2011b). Instead, we focus on results for the Colorado River at Lees Ferry, AZ, and contrast the full-ensemble results with those of earlier studies. Our intent is to provide insight for the impact assessment community into the consequences of model, ensemble and pathway choice for studies of this nature.

2 Methods

The impact assessment approach used in Reclamation (2011b) followed that of C04, CL07 and other recent GCM-based, multi-model hydrology studies of the last decade. In brief, an ensemble of downscaled and bias-corrected climate projections was used to force a distributed, macro-scale hydrology model which generated estimates of runoff and other hydrology variables. The projections of future climate were produced by 16 GCMs forced by IPCC Special Report on Emissions Scenarios (SRES; Nakicenovic et al., 2000) emission pathways B1, A1B and A2. Simulated runoff was routed to 20 streamflow locations within the Upper Colorado River Basin, including the Lees Ferry location. The Reclamation (2011b) work differed from CL07 primarily by using a much larger ensemble of projections (112 versus 22; including 21 of the CL07 projections). Other

differences included using a slightly longer baseline period of natural flows and historical weather (56 yr versus 50 yr), using a slightly different method to disaggregate monthly projections to a daily time step, encompassing a smaller geographic scope (the Upper Colorado River Basin versus the entire Colorado River Basin), and omitting analysis of consequent water resources impacts in the basin. This section describes the flow and meteorological datasets involved (Sect. 2.1) and the hydrology model and its calibration (Sect. 2.2).

2.1 Data

2.1.1 Historical natural flows

Natural flow represents flow that would have occurred at the location had historical depletions and reservoir regulation not been present. Reclamation developed natural flows for 29 locations required by the Colorado River Simulation System (CRSS), a water resources management model used by Reclamation for long-term planning studies (see Prairie and Callejo, 2005). Monthly incremental natural flows at 20 locations upstream of Lees Ferry (see Fig. 1) were obtained from Reclamation (2009) and, following the CRSS approach, were summed to obtain the total Lees Ferry natural flow.

2.1.2 Historical daily meteorology

A daily meteorological climatology that includes precipitation, maximum temperature, minimum temperature and wind speed for the period from 1949 through 2005, developed as described in Maurer et al. (2002), formed the historical climatology and forcing dataset used in this study. The data are aligned spatially to match the NOAA/NASA Land Data Assimilation System (LDAS; Mitchell et al., 2004) grid, which has a spatial resolution of 1/8-th degree latitude by longitude and covers a domain from 25° N to 53° N and 67° W to 125° W, which includes the continental United States as well as part of Canada and Mexico.

2.1.3 Simulated historical and projected climate

Simulated monthly average precipitation and monthly average temperature projections spanning the period 1950 through 2099 were obtained from the Bias Corrected and Downscaled WCRP CMIP3 Climate Projections website (WCRP, 2009; <http://gdo-dcp.ucllnl.org/>). At the time of this study, the archive contained 112 projections of monthly temperature and precipitation, aligned spatially with the LDAS grid, with each projection consisting of an overlap period of 1950 through 1999 and a projection period of 2000 through 2099. These projections come from 16 GCMs and three SRES (Special Report on Emissions Scenarios) emissions pathways as shown in Table 1. The emissions pathway columns indicate the number of realizations that were available for each GCM.

Table 1. Downscaled WCRP CMIP3 GCMs and emissions pathways.

	Modeling group, country	IPCC model I.D.	Emissions pathway		
			A2	A1b	B1
1	Bjerknes Centre for Climate Research	BCCR-BCM2.0	1	1	1
2	Canadian Centre for Climate Modelling and Analysis	CGCM3.1 (T47)	5	5	5
3	Météo-France/Centre National de Recherches Météorologiques, France	CNRM-CM3	1	1	1
4	CSIRO Atmospheric Research, Australia	CSIRO-Mk3.0	1	1	1
5	US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM2.0	1	1	1
6	US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM2.1	1	1	1
7	NASA/Goddard Institute for Space Studies, USA	GISS-ER	1	2	1
8	Institute for Numerical Mathematics, Russia	INM-CM3.0	1	1	1
9	Institut Pierre Simon Laplace, France	IPSL-CM4	1	1	1
10	Center for Climate System Research (The University of Tokyo), National Institute Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC), Japan	MIROC3.2 (medres) for	3	3	3
11	Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA	ECHO-G	3	3	3
12	Max Planck Institute for Meteorology, Germany	ECHAM5/MPI-OM	3	3	3
13	Meteorological Research Institute, Japan	MRI-CGCM2.3.2	5	5	5
14	National Center for Atmospheric Research, USA	PCM	4	6	7
15	National Center for Atmospheric Research, USA	CCSM3	4	4	2
16	Hadley Centre for Climate Prediction and Research/Met Office,	UK UKMO-HadCM3	1	1	1

The monthly climate datasets were produced using the statistical bias-correction and spatial disaggregation (BCSD) method described in W02-04. The method was first implemented for downscaling general circulation model seasonal climate predictions to support hydrologic forecasting (W02; Wood et al., 2005; Wood and Lettenmaier, 2006) and adapted for downscaling future climate scenario model output (W04; also C04; Van Rheen et al., 2004; Payne et al., 2004). The BCSD method has since been employed in a number of more recent climate change impact analyses, in regions such as the western US (CL07; Barnett et al., 2008; Maurer, 2007), the continental US (Maurer et al., 2002; Maurer and Hidalgo, 2008), and Central America (Maurer et al., 2009) among other locations.

In brief, the BCSD approach involves three steps: (1) bias-correction affects a quantile-mapping adjustment of monthly climate-model-scale precipitation and temperature outputs, a step which aligns the monthly climatologies of the climate model variables during a historical period (e.g., 1950–1999) with an observed climatology for the same period and spatial scale; (2) spatial-disaggregation from the climate-model scale to the fine scale is accomplished by applying the interpolated bias-corrected variable anomalies from the coarse scale to a fine scale climatology, using multiplicative anomalies for precipitation and additive anomalies for temperature; and (3) temporal disaggregation from monthly to a finer time step (e.g., daily) via a resampling and adjustment of historical weather patterns from the hydrology model forcing climatology. This development of the monthly-scale scenarios is detailed in Maurer et al. (2007b). The final disaggregation step differed in some regards from the original W02-04 studies, instead following the Maurer et al. (2007b) implementation.

A more detailed discussion of the resampling and disaggregation is provided in Appendix A.

2.1.4 BCSD downscaling considerations

The BCSD downscaling approach has been found in prior studies (such as those referenced above) to be generally successful in translating monthly scale climate signals from the GCM output to the fine resolution at the monthly scale. W02 showed that using BCSD to downscale retrospective monthly climate-model-scale observed precipitation and temperature fields reproduced the monthly mean and variance of hydrologic simulations (for the Ohio River Basin). For water resources-oriented studies that involve monthly and coarser time scale analyses of hydrologic changes resulting from warming and moisture changes, BCSD has provided many useful insights. The BCSD approach of W02-04 is not adequate for every type of climate change study, however. The approach is not suitable for certain studies in which changes in sub-monthly (e.g., daily) meteorological quantities are important (e.g., assessing changes in extreme precipitation or changes in temperature extrema). Also, the BCSD approach may perform poorly where daily meteorological climatologies are highly skewed or exhibit threshold behavior. For example, in the US Southwest, the intermittency and resulting skewness of precipitation can lead in BCSD to pathological combinations of sample and monthly scaling factors that produce unrealistically high daily values – a difficulty that is less pronounced in rainier climates. Adaptations such as sample substitution described in W02-04 were developed to ameliorate these difficulties, but they are not a comprehensive solution. Other approaches such as constructed analogues (compared with BCSD in Maurer and Hidalgo, 2008)

may be more appropriate in such hydroclimatic regimes. Finally, climate model sequences may not be realistic, which undermines the utility of BCSD and other downscaling approaches that do not modify this climate characteristic.

2.2 Hydrology model

2.2.1 Description

The daily disaggregated projections were used to force a hydrology model of the Upper Colorado River Basin implemented using the Variable Infiltration Capacity (VIC) Model (Liang et al., 1994, 1996). The VIC model is a distributed (gridded) macro-scale (regional-scale) physical hydrology model that simulates the water balance around each grid cell. VIC produces a time series of runoff, baseflow, evapotranspiration, soil moisture, and snow water equivalent for each grid cell. Following completion of simulation of the full forcing period, the runoffs from all of the grid cells in the model are routed to points of interest (Lohmann et al., 1998a). Distinguishing characteristics of the VIC model are described at length elsewhere (e.g., Nijssen et al., 2001; Wood et al., 1992).

VIC has several applications to climate change studies in basins around the world (Wood et al., 1992; Liang et al., 1996; Lohmann et al., 1998a,b). The VIC model has been used to support assessments of the impact of climate change in many river basins in the western United States, including California's Central Valley (Van Rheen et al., 2004; Maurer et al., 2007b; Anderson et al., 2008; Reclamation, 2008), the Colorado River Basin (C04, CL07) and the Columbia-Snake River Basin (Payne et al., 2004). VIC is considered well adapted to application to the UCRB because it allows for a relatively detailed representation of the land surface, considering both the grid scale and sub-grid variability, and because it has physically based models of snow dynamics and evapotranspiration. These capabilities address many of the complexities of the basin hydrology described above.

2.2.2 Model calibration

A VIC model is specified by a set of global parameters, options and variables and by reference to a set of gridded parameters and a set of gridded forcings. The most important of the global parameters and options determine the number of soil layers, the time step and the duration of the simulation, and control the simulation approach. For this application, the model was configured to represent three soil layers using a daily time step at the 1/8-th degree resolution matching the forcings and routing model. Each VIC grid cell is characterized with parameters describing vegetation and soil. A calibrated set of model parameters for the Colorado River Basin used by C04, updated by Wood and Lettenmaier (2006) and then used in CL07 was applied for this study. These parameters differed from the Maurer et al. (2002) specifications

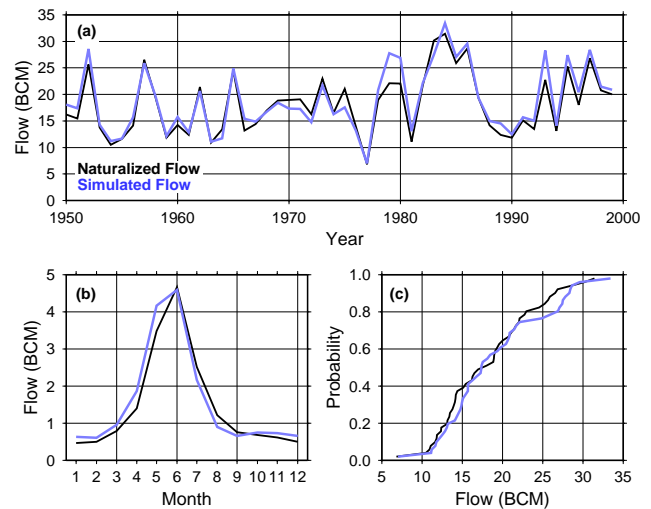


Fig. 2. Comparison of simulated and observed annual streamflow for the Colorado River at Lees Ferry, AZ. (a) Time series, (b) monthly hydrographs, and (c) ECDFs. BCM is billion (10^9) m^3 .

mainly as a result of calibration, and by the inclusion of average July air temperatures required to implement a condition removing canopy vegetation above tree line elevations. Documentation regarding the VIC parameters and the tree line adjustment is given at the VIC model website (<http://www.hydro.washington.edu/Lettenmaier/Models/VIC/>).

The VIC model provided a good fit between simulated and historical natural flows for gage locations covering large basin areas such as the Colorado River at the Lees Ferry gage. The duration of the simulation (i.e., 1950–2099) and the initial value of soil moisture in the simulation were the only modifications to the prior parameter sets. Here, soil moisture values were initialized using the average simulated values for 1970 through 1999. The calibrated VIC model reproduced annual average flow volumes at Lees Ferry with a bias of approximately 4 %, a coefficient of determination of 0.92 and a Nash-Sutcliffe efficiency of 0.91. The model, on average, tended to simulate slightly earlier runoff and slightly lower peak monthly flows compared to historical natural flows. Figure 2a shows that the model bias appears to be concentrated in a few relatively wet years following dry years. Figure 2b shows that the model exhibits a wet bias prior to June and a dry bias in July, August and September. Figure 2c, showing the empirical cumulative distribution function (ECDF) of the streamflows, shows that the model is relatively unbiased between about the 40th percentile and the 75th percentile of flows, and generally shows a positive bias outside that range. Figure 2c also indicates that the model over-simulates the highest values.

The simulation results of Fig. 2 cannot be compared directly with those reported in C04 because C04 did not report results at Lees Ferry. Although the details of the monthly hydrograph in Fig. 2 differ slightly from the hydrograph at

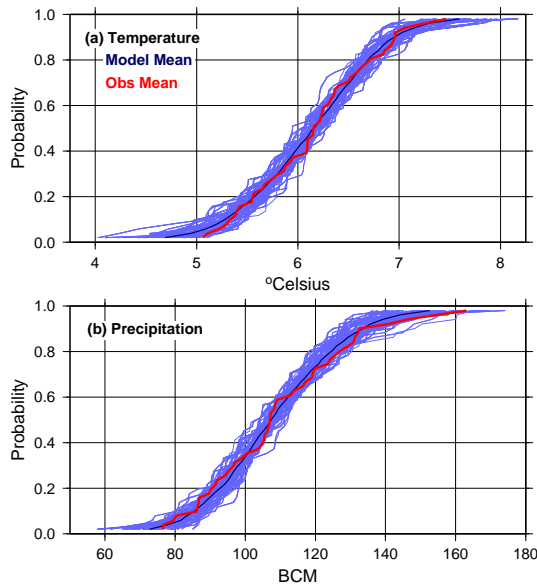


Fig. 3. ECDFs of annual average temperature and total precipitation averaged over the Upper Colorado River Basin above Lees Ferry, Arizona for the period 1950–1999. The downscaled forcing ensemble from 112 GCM projections (light blue), and their mean, is compared with observed (a) temperature and (b) precipitation calculated from the Maurer et al. (2002) dataset.

Imperial Dam, AZ, reported in C04, the quality of calibration in the two studies appears similar. Possible reasons for the differences include the following: (a) the modeling reported herein used version 4.0.7 of the VIC model, whereas C04 used versions 4.0.3 and 4.0.4; (b) the calibration and validation periods varied, with this work evaluating statistics over the period 1950–2005 while C04 used a shorter period; and (c) the difference in reporting locations.

3 Results and discussion

In this section, we present results from the climate change impact assessment for the Colorado River at the Lees Ferry, AZ gage, and refer the reader to a wider range of results and analyses in Reclamation (2011b). The single-gage results also serve to illustrate characteristics of the uncertainty in the analysis.

3.1 Downscaled climate, past and future

As noted earlier, the model forcings derived from 112 climate projections are bias-corrected so that their monthly climatologies are consistent with historical forcing climatologies. This property is assessed in Reclamation (2011b), and we evaluate (Fig. 3) the consistency of the resulting empirical cumulative distribution functions (ECDFs) of annual total precipitation and annual average temperature for the 112 simulations of historical climate (GCM output over the

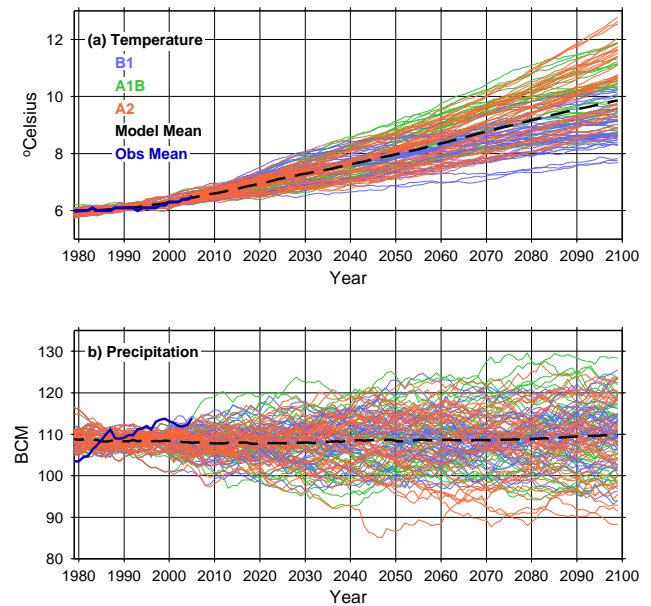


Fig. 4. Projected evolution of climate in the Upper Colorado River Basin above Lees Ferry, Arizona. Shown are 30-yr averages (trailing) of climate for (a) temperature and (b) precipitation. Precipitation is reported as the total volume falling in the basin.

overlap period, 1950 through 1999), relative to the historical forcing distribution to which they were downscaled. Although the downscaled and historical annual climatologies are similar, residual bias remains in the projected annual values. This result is due to the application of bias correction on a monthly basis, without further adjustments to correct annual biases that would result from different persistence behavior in historical and GCM anomalies, non-climatological spatial GCM-scale climate patterns and other factors.

As prior studies have shown, the large ensemble of GCM projections indicates that temperatures have generally been increasing in the recent historical period and will continue to do so, but future precipitation projections contain large uncertainty. Figure 4 shows the evolution of projected temperature and precipitation for the 112 downscaled projections, averaged spatially over the drainage area of the Colorado River above Lees Ferry, AZ, and temporally over trailing 30 yr periods (for this reason the time series start in 1979). The mean of all 112 projections and the mean of the historical values (from the forcing climatology) during the overlap period from 1979 through 1999 are also shown. The color separation of projection trajectories by emissions pathway indicate that the A2 emissions pathways have a weak tendency toward being drier and warmer late in the 21st century than those from the other pathways.

Reclamation (2011c) reports that BCSO downscaled future precipitation projections from the LLNL archive (which are used here) exhibit a “relative wettening” in precipitation changes from current climate, relative to those from the

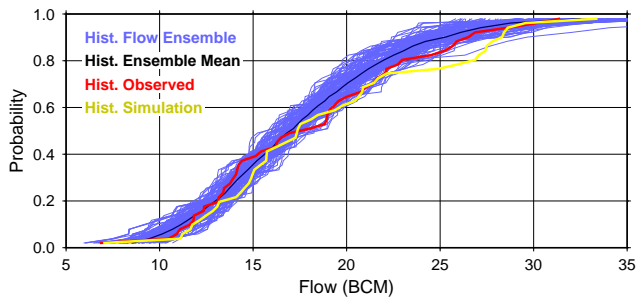


Fig. 5. ECDFs of downscaled simulated flows on the Colorado River at Lees Ferry, AZ, 1950 through 1999, compared with observed and historical simulated flows.

underlying (raw) CMIP3 GCM output from which they are derived. By 2069, the difference affects approximately two-thirds of model projections, reaches a maximum at about the 75th percentile and is greater in dryer areas than in wetter areas. In the snow-producing regions of the Upper Colorado River Basin, the wetting effect approaches approximately a 2% maximum, but allocating the differences depicted in Reclamation (2010c) to emphasize runoff-generating regions and across the distribution of projections results in an aggregate basin mean wetting effect of approximately 0.7%. The effect results from the BCSD quantile-mapping approach, and may be a realistic result of mapping projected precipitation between GCM distributions and observed distributions (generally a transform from a more normal distribution to a more skewed distribution). The hydrologic implications of this difference are discussed in Sect. 4.3.

3.2 Downscaled streamflow, past and future

Ideally, the simulated flow climatology driven by the downscaled GCM-based forcings during the historical period should be consistent with the historical flow climatology driven by historical forcings. Figure 5 compares the ECDFs of annual streamflows at Lees Ferry, Arizona, for the historical 1950–1999 period from all 112 GCM-based simulations to the observation-based natural flows and the model-simulated historical flows.

The GCM-based flow ensemble for the historical period shows a dry bias above the 40th percentile relative to the observation-based natural flows and a larger dry bias relative to the simulated historical flows. Neither the GCM-based annual precipitation nor temperature exhibited biases in directions consistent with this flow bias (i.e., drier and/or warmer); thus the lower GCM-based flows result from some other aspect of the projections. As noted earlier, the bias likely results from either the seasonal pattern, spatial pattern, the inter-annual sequence of precipitation and temperature, or some combination of these factors. Relatively high flow values in the Colorado River Basin follow from the sequencing of multiple months of anomalous climate, having either high

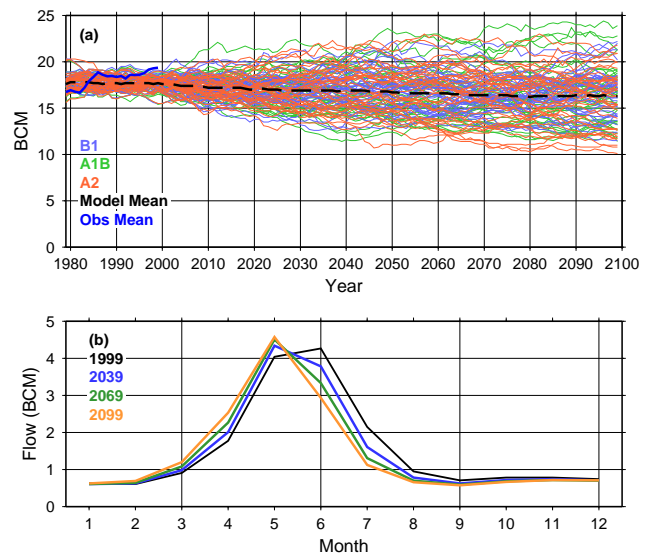


Fig. 6. (a) Simulated 30-yr average (trailing) streamflows of the Colorado River at Lees Ferry, AZ, 1979 through 2099. (b) The mean monthly average streamflows for the three future projection periods, compared with the historical 30-yr period flow ending in 1999.

precipitation or cold temperatures, often in combination, and such multivariate, temporal (and possibly spatial) structure may be poorly represented in a GCM's climate system. Rearranging the baseline climate data in a sequence similar to that simulated by a GCM has been shown to introduce a dry bias (Joe Barsugli, personal communication, 2010).

The hydrologic consequences of the projected climate changes are illustrated in Fig. 6, in which Fig. 6a depicts the projected evolution of 30-yr mean streamflow at Lees Ferry, Arizona, for all 112 projections. The mean of all projections and the observed natural flows are also shown. For clarity, the flows simulated using the hydrology model forced by the historical climate from Maurer et al. (2002) are omitted, as the difference between the 30-yr mean values of the two sets of flows is slight. The streamflow projection average shows a slight downward trend (reaching -8% by the end of the century) that is enveloped by broad uncertainty. Consistent with the forcing projections, the largest declines are associated with the A2 emissions pathway, though this ensemble feature is not prominent until the final decades of the century. The signature of warming effects on western US snowmelt-dominated streamflow is clearly present: namely, the mean monthly hydrograph in Fig. 6b shows a mean shift in peak runoff toward earlier in the year and a small decrease in annual volume that grows progressively larger during the 21st century.

Using the entire ensemble of projections, an analysis of cross-correlations between average streamflow, precipitation and temperature changes for three future 30-year periods indicates that streamflow changes are almost entirely linked to precipitation changes (e.g., with correlation values greater

Table 2. Cross-correlations of average projected streamflow, precipitation and temperature for one historical and three future 30-yr periods.

Cross-correlation between	Simulation period			
	1970–1999	2010–2039	2040–2069	2070–2099
Streamflow and precipitation	0.91	0.95	0.95	0.94
Streamflow and precipitation after removal of temperature influence	0.91	0.96	0.95	0.93
Streamflow and temperature	−0.19	−0.47	−0.51	−0.60
Streamflow and temperature after removal of precipitation influence	−0.21	−0.63	−0.51	−0.45
Precipitation and temperature	−0.11	−0.27	−0.36	−0.49

than 0.9). The partial cross-correlations (i.e., from correlating residuals from each variable's linear regressions on temperature) are nearly identical (Table 2) to the full correlations. The full and partial correlations of streamflow with temperature are negative and moderate, indicating that higher temperature alone draws the system toward lower flows. The strengthening relationship between projected temperature and precipitation throughout the 21st century may indicate a linkage between the two variables in the climate system, but is not investigated here.

3.3 Uncertainty in projections of climate and streamflow

The large ensemble assessed in this study offers an opportunity to evaluate the wide variation (hence uncertainty) in results shown in Fig. 6a across the spectrum of the available projections. Of particular interest are the *relative changes* derived from the different projections (e.g., future climate relative to a control or historical climate baseline). In this study, relative change is calculated for each projection separately and reported as a percent difference in the mean flow of the future 30-yr period relative to the projection's historical 30-yr period mean flow (for 1970–1999). Note that the baseline for each projection is not the observed or simulated historical flow (i.e., forced with observations), but the downscaled simulated historical flow for the same projection, a choice that avoids including biases or statistical artifacts of the downscaling in the relative change calculation.

Figure 7a shows the ECDF of the changes for the 30-yr periods ending in 2039, 2069 and 2099, and Table 3 summarizes statistics describing the mean and spread of each distribution of differences. For each choice of emissions pathway or future period, the range of projected changes is large ($\pm 30\%$). This range is a result of uncertainty derived from model choice and ensemble member variability, and is far larger than the uncertainty arising from the choice of analysis period. Figure 7b shows the empirical distributions of projected changes separately for each SRES emissions pathway. Below the 40th percentile, the difference in flow changes between pathways is larger than the differences across future periods (Fig. 7a) – i.e., about $\pm 10\%$ – but still small compared to the standard deviation of the ensemble for each

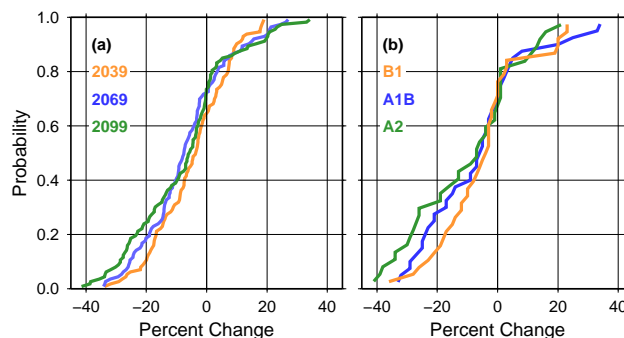


Fig. 7. (a) Distribution of projected changes in average streamflow for 30-yr periods ending in 2039, 2069 and 2099, relative to simulated historical (1970–1999) average flow from each downscaled projection. Positive values are flow increases. (b) Distribution of projected relative flow changes for 2070–2099 for each SRES emissions pathway.

emissions pathway. The A2 projections produce the largest decreases and smallest increases of the three emissions pathways, as illustrated in Fig. 6a. Figure 7 also shows that approximately one-quarter to one-third of the projections suggest no change in runoff or a wetter future for Colorado River flow (by 2099 and 2039, respectively). If the “relative wettening” effect (Reclamation, 2011c) described earlier (0.7%) is removed from these results by using an estimated precipitation elasticity of streamflow for the Lees Ferry streamflow of 2.5 (see Sankarasubramanian et al., 2001), the resulting 1.7% drying of flow changes reduces the fraction of future projections of Colorado River at Lees Ferry flows that are wetter by 2069 from slightly less than 30% to approximately 20%.

The contributions and significance of different uncertainties illustrated in Fig. 7 can be quantified in various ways. It is worth observing that the largest difference between the median projected runoff changes of different periods and of different emissions pathways are about 4 and 3%, respectively, whereas the smallest standard deviation of projected changes for each period and emissions pathway are 12 and 14%, respectively (Table 3). Given the substantial uncertainties contributed by internal model variability, GCM choice, and possibly other factors, illustrated in Fig. 7a, can even the

Table 3. Average projected percent change in streamflow for 30-yr periods ending in 2039, 2069 and 2099, and for different SRES emissions pathways for the latter period, calculated for different percentiles and statistics of the ensemble distribution. Positive values are flow increases, in percent.

	Projection period			Emissions pathway, 2070–2099 period		
	2010–2039	2040–2069	2070–2099	B1	A1B	A2
maximum	19	27	34	23	34	21
90 %	10	12	14	2	20	13
75 %	4	1	1	0	1	1
50 %	–4	–8	–6	–4	–6	–7
25 %	–14	–15	–19	–14	–21	–27
10 %	–20	–25	–29	–23	–29	–35
minimum	–34	–34	–41	–36	–33	–41
mean	–4.7	–6.9	–7.6	–5.3	–7.1	–10.6
std. dev.	12.0	13.8	16.0	16.7	17.2	14.0

largest ensemble quantify a change signal with confidence? Deser et al. (2012) used a t-test approach on a large number of ensembles sampled from 40 runs of a single climate model to estimate whether future ensemble mean changes are significantly different from zero in the context of ensemble uncertainty. The study found that relatively large ensembles (i.e., dozens) of runs from one model were necessary to calculate a statistically significant future change until the middle of the century. Following Deser et al. (2012), we estimate statistical significance of the projected 30-yr mean flow changes for the 112 member multi-model ensemble at 2039, 2069 and 2099. All three changes were statistically significant at $p=0.05$. A Welch's t-test (for comparing samples unequal in size and variance, using the Welch-Satterthwaite equation for degrees of freedom) indicates that the emissions pathway flow changes are not significant from each other at $p=0.10$ or higher at any time period. Note, because individual projections for runoff are relatively uncorrelated with other projections, they are treated as being independent within each sample.

Alternatively, a two-way analysis-of-variance (ANOVA) can assess the contributions of emissions pathway and period to future streamflow changes by factoring 336 change results from 3 periods each with 36 A2, 37 B1 and 39 A1B (total 112) projections. ANOVA results generally support the t-test results, finding that for runoff neither period nor emissions pathways are significant factors at $p=0.05$, but period does barely achieve significance at $p=0.10$. Also, neither factor is significant for precipitation, but both are significant well below $p=0.05$ for temperature. In this usage, however, ANOVA assumptions are violated to the degree that the sample distributions depart from Gaussian and their variances are not equal; thus it may be wise to consider the results judiciously. It may also be arguable whether models are truly independent given the sharing of physics components and other characteristics (Pennell and Reichler, 2011). Further

evaluating whether the choice of model is a significant additional factor is undermined by small sample sizes (a minimum of 9, e.g., from 3 pathways and 3 periods, for GCMs with one run) and even larger differences between sample sizes and variances.

Table 1 shows that the numbers of runs contributed by the GCMs are not equal; just seven of the models provide 75 % of the runs. Should planners be concerned about this unequal weighting? We constructed one hundred 48-member projection ensembles by re-sampling the 112-member ensemble such that each combination of GCM and emissions pathway was represented by one run randomly selected from all of the runs available for that combination in the 112-member ensemble. This strategy weighted equally all of the GCMs within each projection ensemble, and the resulting distributions were compared to the 112-member ensemble using the Kolmogorov-Smirnov test. In 99 of 100 cases, the two distributions were not significantly different in the three future periods (with p-values for rejecting difference over 0.90). In the remaining single case (for the 2039 period only), a difference of the full and equally weighted distributions could not be rejected at a p-value of 0.2. We find that the arbitrary weighting of models arising from the unequal ensemble members for each model does not significantly alter our findings.

The contrasting influences of GCM choice, emissions pathway and future period on projected changes are illustrated further in Fig. 8, which shows the time evolution of 30-yr average flows for all runs of the 16 GCMs separately. The substantial variation in flow trends between models depicts the strong dependence of the signal of the relative change from present to future on the model or set of models selected as a basis for analysis. For instance, the IPSL CM4 model produces slight upward trends in flow until the middle of the 21st century and downward trends thereafter, whereas the MIUB ECHO G model produces progressively decreasing flows for nearly all ensemble members. The separation

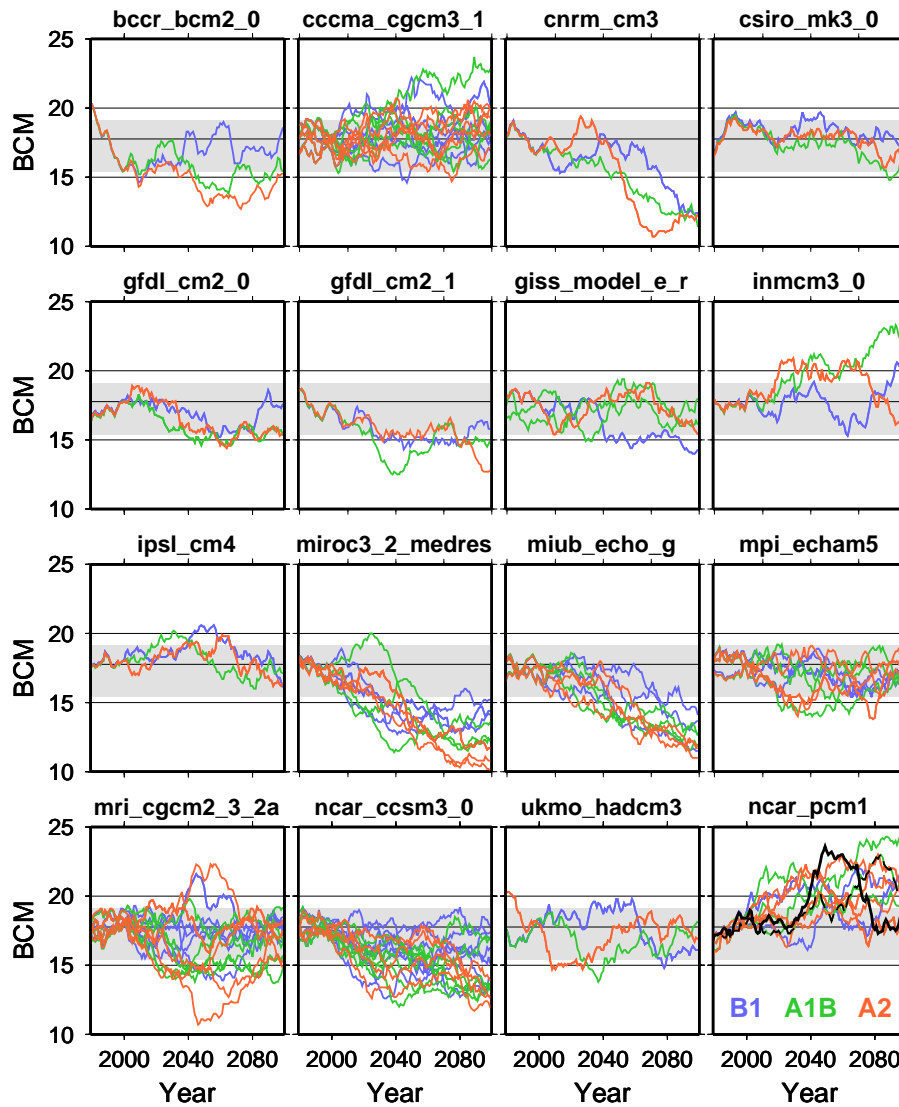


Fig. 8. Time evolution of projected 30-yr averages (trailing) of annual Colorado River streamflow at Lees Ferry, AZ. Projections are color-coded by emissions pathway (see bottom right panel), with the exception of two runs that are highlighted for the NCAR PCM1 model (see text).

of these results further by emissions pathway indicates that the sensitivity of the models to the emissions pathway varies widely between models as well. Figures 9 and 10 show the corresponding time evolution of 30-yr average total precipitation and temperature, respectively. The notion that the A2 pathways will produce a stronger change for any given period than the more benign B1 pathways is evident for streamflow and precipitation for only a few models, e.g., the MIUB ECHO G model. While this notion arguably describes future temperature projections, it cannot be extended to regional precipitation or streamflow projections.

A significant source of the large uncertainty in projected relative changes in precipitation and streamflow is the variation in period and phase of simulated low-frequency (e.g.,

decadal) variability among GCMs. This finding is highlighted for the NCAR_PCM1 model in the bottom right panel of Figs. 8–10, which illustrates low-frequency phase differences for two runs of the A1B emissions pathway. If used to characterize trends for streamflow for the second half of the 21st century, the two runs would represent the end members of the distribution of trends during that period for all projections shown in Fig. 8, despite being generated from the same model when forced with the same emissions trajectory. Similarly, the four A2 pathway runs for the MRI_CGCM2_3_2A model span a range from –30 to +30 % at the mid-century. All of the model runs exhibit this phase variation to varying degrees, which confounds interpretation of the climate change impacts for any particular

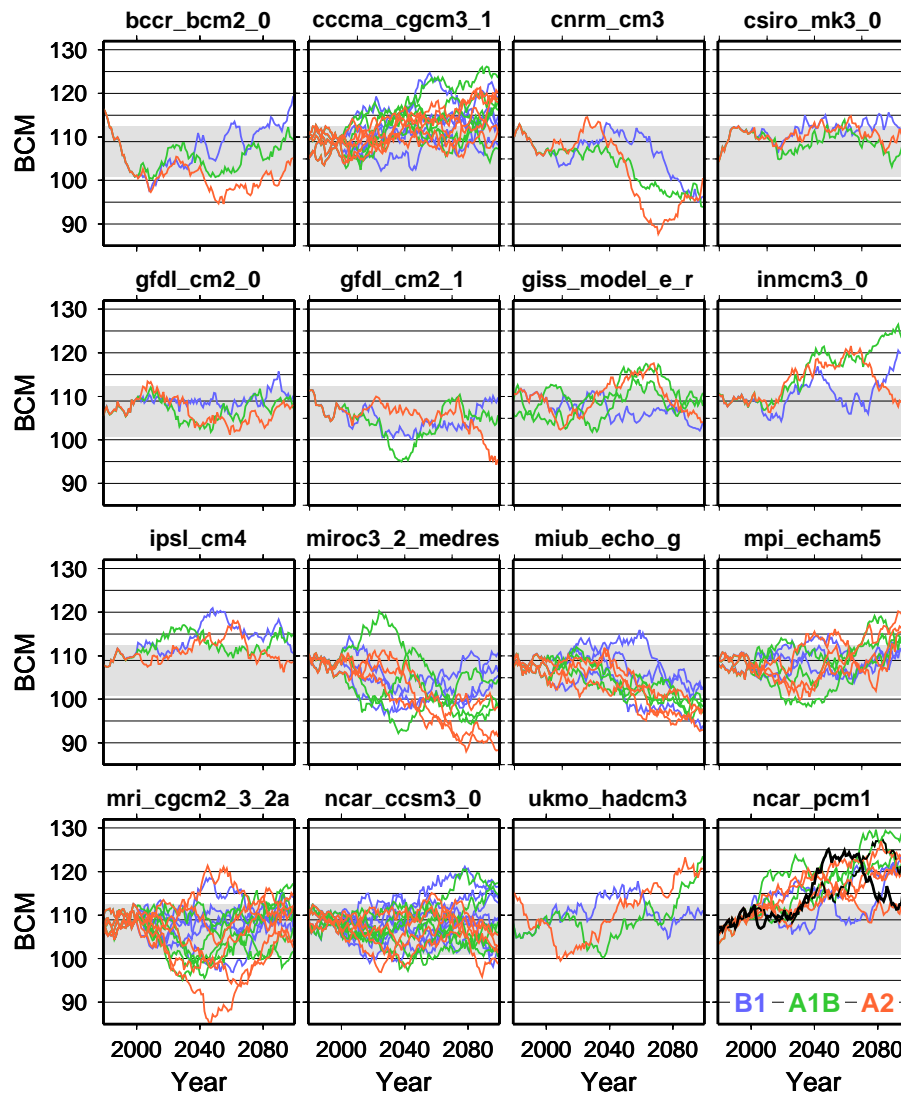


Fig. 9. Time evolution of projected 30-yr average (trailing) of total annual precipitation upstream of Lees Ferry, AZ. Projections are color-coded by emissions pathway (see bottom right panel), with the exception of two runs that are highlighted for the NCAR PCM1 model (see text).

future period, and particularly when a small subset of projections forms the basis of the analysis. The analysis and illustration of the influence of internal model variability on temporal features of GCM-based regional streamflow projections is analogous (and related) to the illustration of the influence of model variability on spatial features of projected climate in Deser et al. (2012; see Fig. 8 from that work). Deser et al. (2012) are broadly supported by climate research toward characterizing projection uncertainty (e.g., Hawkins and Sutton, 2011) which understands that GCMs may accurately simulate low-frequency variability but do not reproduce the observed phasing of this variability for the 20th and 21st century periods in CMIP3-style coupled model simulations. The simulated low-frequency phase is largely random, adding noise to projected trends or outcomes. This internal

model variability problem similarly undermines initialized decadal predictions from climate models (e.g., Meehl et al., 2009).

3.4 Comparison with previous work

The large ensemble used in this study provides perspective on prior published climate change impact assessments. For instance, CL07 estimated future runoff using a similar approach to this study – i.e., BCSO downscaling of GCM outputs and VIC model hydrologic simulation – but used only 22 climate projections, which were taken from 11 GCMs forced with the B1 and A2 emissions pathways. The correlation of projected change results for individual GCMs in CL07 and this study is high, approximately 96 %, due to the

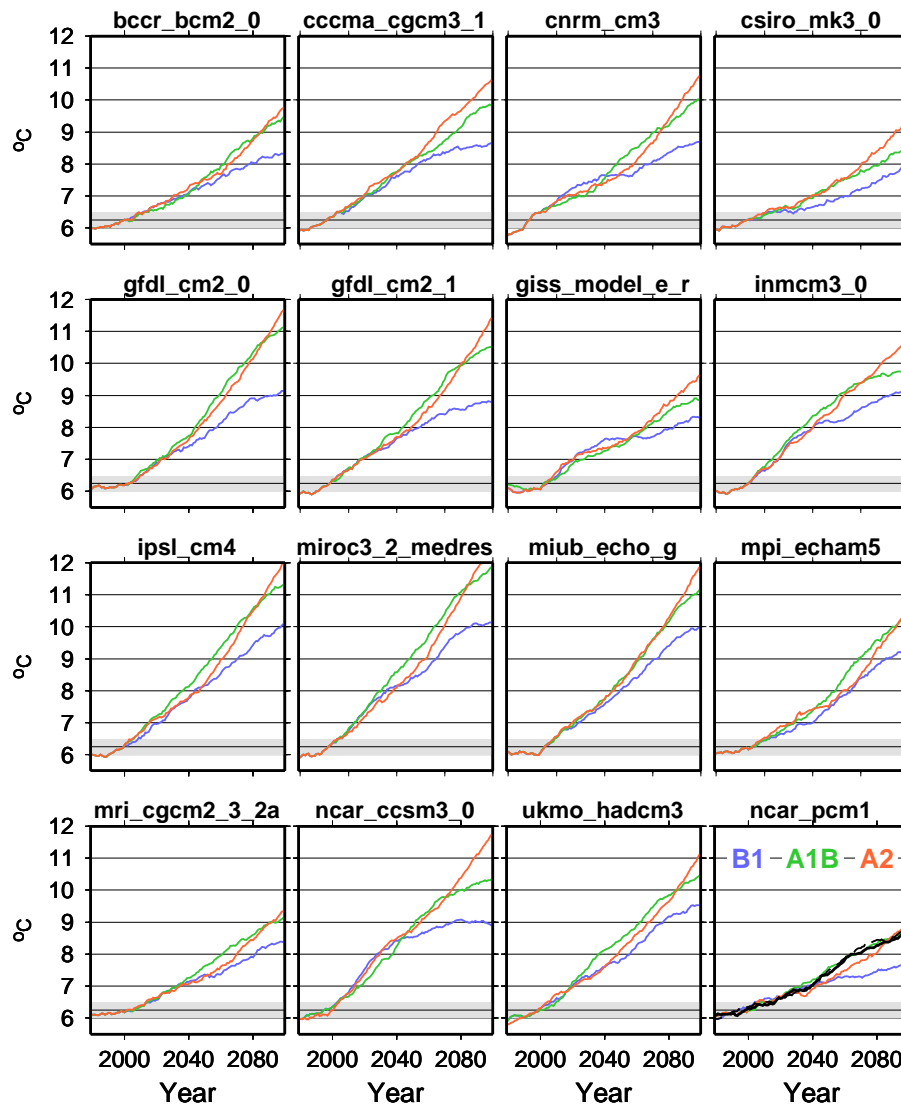


Fig. 10. Time evolution of projected 30-yr averages (trailing) of temperature upstream of Lees Ferry, AZ. Projections are color-coded by emissions pathway (see bottom right panel), with the exception of two runs that are highlighted for the NCAR PCM1 model (see text).

similarity of methods, even though CL07 reported impacts on streamflow at a different location (Imperial Dam versus Lees Ferry).

In Fig. 11, we illustrate the difference in projected change for Lees Ferry streamflow between the CL07 “subset” and the larger ensemble by highlighting the results from this study for the GCM runs that were also included in CL07. For the 30-yr periods ending in 2069 and 2099, the difference is minimal: the CL07 subset represents the range and central tendency of the larger distribution. For the 30-yr period ending in 2039, however, the CL07 projections produce a wetter estimate of the mean and the inter-quartile range of projected impacts. While the selected projections reproduce the ensemble mean for 2069 and 2099, their distribution is uneven, e.g., the projections selected for 2099 are heavily weighted

in the inter-quartile range, and this streamflow bias can propagate bias into sectoral impact studies, e.g., water management modeling. Figure 8 illustrates that projected changes based only on one GCM (such as the NCAR PCM1 used in C04 and related studies) would impose even greater sampling error onto findings, relative to a much larger projection ensemble. CL07 did not report enough information to evaluate the significance of their reported mean changes in flow, so we evaluated the CL07 GCM-emissions pathway combinations ensemble using a matching sample of projections from this study (with the exception of one run). In contrast to the projected ensemble mean changes based on the 112-member ensemble, and using the two-sample, unequal variance t-tests (as before), only the CL07 A2 2099 and B1 2069 ensemble mean changes were significant at $p=0.05$.

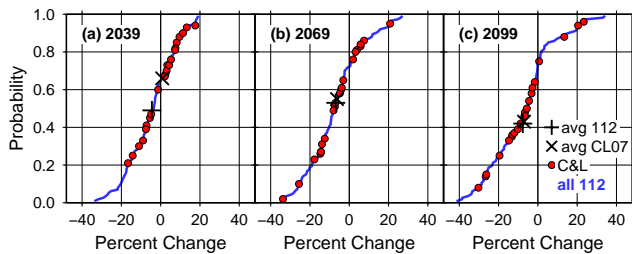


Fig. 11. Distributions of projected changes in 30-yr averages (trailing) of annual streamflow from GCMs included in CL07, compared to the distribution for all CMIP3 projections in the LLNL archive (present study).

In another study, Milly, et al. (2005) reported a reduction in runoff in the Upper Colorado River Basin in 23 of 24 pairs of model runs (96 %), with an ensemble median runoff decrease of between 10 and 25 % by 2060, as shown in Fig. 12 (Backlund et al., 2008; Milly et al., 2005). Seager et al. (2007) and Seager and Vecchi (2010) also suggest a broad consensus among climate models that changes in atmospheric circulation will lead to additional drying in the American Southwest region that includes the Colorado River Basin, with consequent impacts on water resources. In contrast, our study suggests less certainty about the water resources consequences, in that, depending on the time frame, from about one-fourth to about one-third of the available climate projections lead to no change or to an *increase* in streamflow at Lees Ferry (just above the outlet of the Upper Colorado River Basin.) CL07 also shows a higher likelihood of wetter conditions than were suggested by any of the three large-scale studies that were based directly on analysis of GCM outputs. The different findings likely do not stem entirely from the BCSD wettening effects noted earlier, because that effect is not large enough, but are also the result of different scale and runoff process representations between the large-scale studies versus the current study and others such as C04 and CL07. The statistical downscaling and hydrologic modeling studies include a far more explicit treatment of the spatial and seasonal heterogeneity in runoff generation in this study (Backlund et al., 2008; CL07), which leads to a stronger and more realistic weighting of climate changes in the mountainous runoff generation areas on the basin's northern and eastern boundaries, and during winter, when runoff-producing snowpack accumulates. This view supports the finding of Wilby and Harris (2006) that, in climate change impact analyses, the downscaling process and hydrology model structure and scale are the second and third largest sources of uncertainty, respectively.

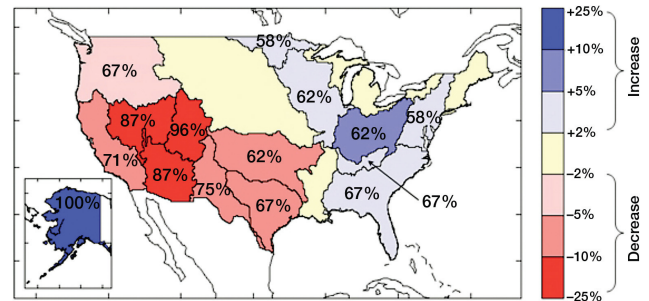


Fig. 12. Median changes in runoff interpolated to USGS water resources regions from Milly et al. (2005) from 24 pairs of GCM simulations for 2041–2060 relative to 1901–1970. Percentages are fraction of 24 runs for which differences had same sign as the 24-run median. Results re-plotted from Milly et al. (2005) by P. C. D. Milly, USGS. From Backlund et al. (2008).

4 Discussion and conclusions

This study presented and dissected results of an impact assessment of projected climate change on streamflows in the Upper Colorado River Basin. The assessment applied high-resolution hydrology modeling to the broadest range of climate projections used for any hydrology or water resources impact study published for the region to date. The objective of the analysis was to provide insight for water resources managers into the substantial uncertainty inherent in GCM-based impact studies.

The impact assessment findings generally agree with the earlier work that used similar methods with smaller GCM ensembles (e.g., CL07 and C04). For example, a largely temperature-driven shift to earlier runoff is a robust projection in this snowmelt system. Such a timing shift will reduce water availability for agriculture and other economic uses of water where those uses do not have access to sufficient reservoir regulation, and will also affect ecological conditions. This assessment of future Colorado River runoff changes, however, is more equivocal than earlier large-scale studies (based on analysis of runoff or moisture transport in GCM outputs), or the often-cited central tendency of CL07: even adjusting for the possible effect of the BCSD “wettening,” approximately 20 % of climate projections lead to estimates of future conditions by 2069 where average annual flows in the UCRB are unchanged or increase. The contrasting finding arises from this study's more comprehensive accounting of the spatial and temperature variability in key hydrologic processes that are driven by downscaled climate changes.

We also confirm that precipitation projection uncertainty overwhelms temperature uncertainty in influencing future streamflow uncertainty, and that, all things being equal, temperature increases are correlated with streamflow reductions in the Colorado River Basin. Because precipitation change consequences are so large, the effect of “wettening” (Reclamation, 2010) or other potential downscaling artifacts

(however small) warrant investigation, and not least because BCSD downscaled projections such as those of this study have been widely used for impact assessment in the water resources field. In the meantime, impact assessments based on the BCSD downscaled projections should recognize this potential bias.

The streamflow projection uncertainty analysis found that the effect of different emissions pathways (i.e., B1 or A2) and future period (e.g., 2010–2039) on projected streamflow changes was small relative to the effect of GCM and GCM ensemble member choice, even by the end of the 21st century. In general, we confirm the assertion of Wilby and Harris (2006) that the uncertainties in GCM climate sensitivities have a substantially larger effect on streamflow projection than emission uncertainty. This future climate uncertainty has a spatial component that is particularly relevant in the Colorado River Basin because most of the flow in the basin originates in small regions of high, mountainous terrain that are on the northern border of the region (Reclamation, 2011b; Seager and Vecchi, 2010; C04). Depending on the climate projection, these regions may lie beneath an intensifying future winter Pacific storm track, or they may not – which translates directly into future streamflow uncertainty that varies not only in magnitude but also in direction (e.g., including a range of streamflow outcomes from increases to decreases of 30% by mid-century). Reducing the spatial component of climate change uncertainty may require higher resolution climate models that better represent the mountainous terrain within the basin, but the simulation of other fundamental climate system phenomena must also advance. For example, climate teleconnections influence the transport of moisture to those high-runoff regions, yet the climate of the tropical Pacific that drives those teleconnections is not well simulated by the current generation of climate models (Seager and Vecchi, 2010).

Future precipitation (hence streamflow uncertainty) also has a temporal component. GCMs differ substantially as to streamflow trend amplitude, direction and particularly the phasing of low-frequency (decadal) variability. The projected trajectories of streamflow show that this unforced variability can be the largest source of uncertainty at certain future time periods. The low frequency temporal variability also means that, for any given projection, the estimated change in streamflow between one period (say, “the present”) and another (e.g., 2040–2069) is dominated by a random or chaotic component. This variability may reflect imperfections in our state of knowledge about climate system sensitivity and/or the chaotic behavior of the weather and climate system. Regardless of the cause, it means that sampling error is a real danger when interpreting change signals taken from a small numbers of projection analyses. For the Colorado River Basin and other river systems, selecting a priori a subset of projections without a strong physically based rationale can introduce considerable bias in an impact assessment. Studies based on a small sample of projections can reduce risk

of improperly characterizing noise as signal by assessing the significance of projected change results, but such studies will nonetheless tend to underestimate the range of possible future conditions.

Projection uncertainty engenders considerable disagreement about how to apply climate projections to impact assessment. Some research suggests that the range of impact estimates based on a large ensemble of projections is a minimum bound and is practically irreducible, at least in the foreseeable future (Stainforth et al., 2007; Wilby, 2010). If so, impact assessments must evaluate what some have termed an “ensemble of opportunity” – i.e., an ensemble of all available projections (which we term simply a “large” ensemble) – from which it may be possible to develop probabilistic estimates of impacts (Tebaldi and Knutti, 2007) and a skillful ensemble mean (Gleckler et al., 2008). This approach envisions an end-to-end evaluation of the full ensemble in sectoral applications. Statistical approaches offer a path for downscaling a large ensemble cheaply, but statistical downscaling is not an ideal solution for many applications – for instance, it is difficult to implement in multi-variate contexts and raises obvious concerns due to the use of stationary relationships between GCM and observed climatologies. And even impact assessments based on statistical downscaling may assess a small projection ensemble to simplify follow-on modeling efforts (e.g., for hydrological, ecological, or water resources impacts).

Small projection ensemble experiments (using one to several GCM runs or limited GCM projection ensemble members) are ubiquitous in downscaling studies based on meso- or finer scale regional climate models (RCMs). In light of the uncertainties described in the previous section, one cannot but question the application of small projection samples to inform (at least ostensibly) stakeholder planning decisions. Stakeholders in application sectors such as agriculture, wine-making, skiing, and ecological assessment are understandably attracted to a few attributes of dynamical downscaling – e.g., high spatial-resolution and multi-variate scenario outputs, and their accounting for small-scale climate feedbacks. Because the fine-scale results reflect large-scale projection forcing, however, there is still a need to assess the significance of these results in the context of broader climate projection uncertainty. Meso-scale models have undeniable value both in applications (such as numerical weather prediction) and research (such as climate and weather system diagnosis). Sensitivity analysis may currently be the most appropriate paradigm for dynamical downscaling use in stakeholder-focused climate change impact assessments (rather than serving as “the downscaling tool” in a small-sample end-to-end assessments). Ultimately, whether the analytical approaches used involve dynamical or statistical downscaling, reconciling the implications of projection uncertainty with the stakeholder needs for detailed information remains a fundamental and pressing challenge.

What are individual planning groups to do when the cost of evaluating a large ensemble exceeds available resources? Two possibilities are the following. The first is for an external, centralized group – perhaps a national laboratory or agency – to do the work via an efficient, continental-scale project: that is, to develop downscaled datasets based on a comprehensive ensemble of available projections, running sufficient process models to produce meteorological, hydrological, ecological and other variables of interest for planning purposes at appropriate time and space scales. Planners could access this large ensemble to define specific impact scenarios or use the full ensemble to characterize probabilistic risk. An example of this type of centralized effort is the hydrologic outputs from Reclamation's regional water supply assessment (Reclamation, 2011c), leveraging the data platform of Maurer et al. (2007a) at <http://gdo-dcp.ucllnl.org/>. Though costly, such a database would likely require less total investment than many independent evaluations for particular locales, and far less than the consequences of poorly informed planning decisions. A second approach is to mimic the information in the large ensemble using a subset ensemble for detailed process study, while attempting to design the smaller ensemble to represent the signal and uncertainty of the full ensemble. This “scenario” approach is popular with water utilities, but contains challenges such as the need to select or create scenarios that represent signals and uncertainty consistently from one time period or variable or region to the next. Reclamation, for example, reduced the 112 projections in the CMIP3 archive to 5 representative scenarios for a specific future period (Reclamation, 2010).

Impact assessment effort may also be reduced by culling the large ensemble of projections of GCMs deemed to produce poor simulations of observed climate. Indeed, it is reasonable to ask whether models such as INMCM3.0 and MIUB_ECHO_G have differences in quality in the region of interest, given their far different results for streamflow impacts. Culling, model selection and/or weighting have been explored (e.g., Tebaldi et al., 2004; Pierce et al., 2009; Santer et al., 2009; Brekke et al., 2008) but without yielding a strong consensus that would motivate standardized approaches for practical application. Due to uncertainty such as low-frequency variability and the difficulty of measuring model quality, culling or weighting may not reduce the projection range for variables such as precipitation. For impact assessment, the best strategy given the current suite of GCMs may be to cull the GCM pool stringently using some notion of “a good GCM,” and then to require from each included GCM an ensemble of sufficient members to extract signal from noise. Knowledgeable culling is another aspect of impact assessment that likely exceeds the resources of individual utilities (and indeed many researchers); thus this strategy also argues for a centralized effort that could provide not only downscaled projection data but also insight into their fitness for different regional applications. Further research in this area is clearly needed.

Most water sector planning decisions cannot address the range of hydrologic outcomes from the large projection ensembles without regrets; thus water planning efforts have repeatedly identified reducing future streamflow uncertainty as a critical need (e.g., Barsugli et al., 2009). Until the projection uncertainty can be reduced and better understood, long-term planning decisions will continue to rely heavily on the subjective perceptions and risk appetites of water resources managers, and their stakeholders, investors and financiers.

Appendix A

Creating historical and projected meteorological forcings

This section gives additional detail on the creation of the downscaled climate forcings used as hydrological model inputs. The 112 monthly time step projections were disaggregated to a daily time step by a resampling approach similar to that used in the original W02-04 studies, but differing in some regards, following the Maurer (2007b) implementation. The steps and the differences are as follows.

1. Daily time step patterns for precipitation and temperature are selected by month from past historical months, using the same month selection sequences for both precipitation and temperature as in W02. Also in keeping with W02, the same month-year daily pattern selection was applied to all fine resolution grid cells in a hydrologic domain (in this case, a river basin such as the Colorado River Basin) to preserve the spatial coherence of the daily patterns across domains that, when used for hydrologic analysis, would require a realistic degree of spatial synchronization.
2. The month selection is conditioned using a “4-square approach” that randomly selects a historical month from one of four climate-type bins – dry-cool, dry-warm, wet-cool, wet-warm – depending on the bin into which the climate model variable anomalies fell. For example, daily patterns applied to monthly wet-warm climate model anomalies were selected from month-year combinations (e.g., January 1982) which were wet and warm in the observed climatology. The observed climatology is defined for the years 1950–1999, based on Maurer et al. (2002) forcings. In the original W02-04 approach, only two climate-type bins (wet and dry) were used.
3. The month-long daily patterns of precipitation and temperature minima and maxima (T_{min} and T_{max} , respectively) are then scaled (in the case of precipitation) or shifted (in the case of temperature) so that their monthly aggregations equate to the bias-corrected, downscaled 1/8-th degree monthly values, again as in W02, where more detail is given.

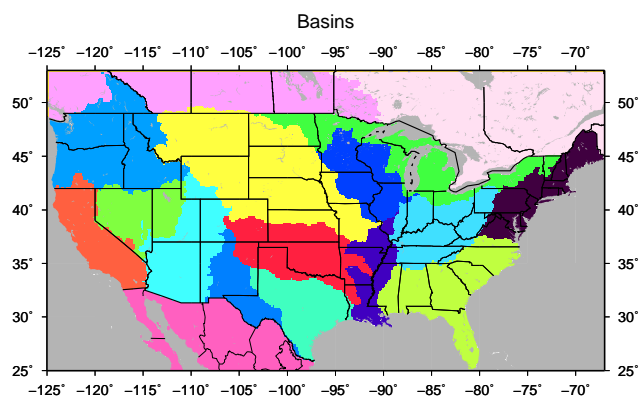


Fig. A1. River basin delineations used in archiving the forcing datasets. Regions include: 1 – Northwest and Columbia; 2 – California; 3 – Great Basin; 4 – Colorado River; 5 – Rio Grande; 6 – Missouri River; 7 – Arkansas-Red; 8 – South Central (Gulf); 9 – Great Lakes Drainage; 10 – Upper Mississippi; 11 – Lower Mississippi; 12 – Ohio; 13 – North East; 14 – South East; 15 – Eastern Canada; 16 – Western Canada; 17 – Mexico.

4. In climatologically dry locations and times of year, the resampling approach can result in “pathological” pattern selections, even accounting for monthly scale climate characteristics (e.g., wet/warm). It is possible to select a daily pattern that does not include a single day of precipitation, or contains only one or two, that cannot be scaled to produce the target monthly values from the corrected, downscaled climate simulation without generating unrealistically high values of daily precipitation. In these cases, which occur for a single grid cell and a single month, a replacement month-year selection is made that does not give a pathological result. The drawback of the substitution is that it undermines the spatial pattern preservation at the daily scale and can desynchronize hydrologic responses that are routed down-slope in a catchment to produce streamflow. For example, the substitute month may have rainfall at the end of the month versus rainfall earlier in the month for the original selection, across a part of the domain. Viewed from a continental, hydrologic perspective, such substitutions occur most frequently in dry locations and in dry seasons, which have less hydrologic significance than wetter situations because they do not generate significant runoff.

5. The rules for the sample substitution differ between the original W02-04 implementation and the adaptation used for this dataset. The original implementation applied a minimum monthly total precipitation limit (e.g., 4 mm), and if the selection produced a scalar above a threshold (e.g., 2, meaning the sample precipitation would be doubled), the sample had to have no fewer than a specified number of wet days (e.g., 6). Different values were applied for the three settings for

different river basins, based on an objective of limiting substitution frequencies to approximately 5 % or lower. The new implementation did not evaluate wet days per month, but limited sample selection to months having greater than 2 mm of precipitation and producing scaling factors of less than 35.

Data and code availability: the gridded daily time-series meteorological datasets are currently housed at the DOE National Energy Research Computing Center (NERSC), and are stored in a binary format used for input by the VIC model. The datasets and the disaggregation code used for this research (from W02 and Maurer et al., 2007b) can be obtained by contacting Andy Wood (Andy.Wood@noaa.gov). The forcings have been archived by river basin, as depicted in Fig. A1, but smaller domains can be requested.

Acknowledgements. This work was primarily funded by Reclamation, Lower Colorado Region. The climate change daily forcing datasets were produced and contributed by co-author Wood using resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.3TIER, Inc. (www.3tier.com) is also acknowledged for accommodating part of A. W. Wood’s effort during his employment at 3TIER. AMEC is acknowledged for accommodating part of B. L. Harding’s effort. We thank Ken Nowak for assistance related to sampling the projection ensembles. We also thank two reviewers and our editor for their insightful suggestions.

This paper describes the analysis and views of the authors and does not represent any official position, view or policy of the authors’ employing agencies or company.

Edited by: H. Cloke

References

- Anderson, J., Chung, F., Anderson, M., Brekke, L., Easton, D., Ejeta, M., Peterson, R., and Snyder, R.: Progress on Incorporating Climate Change into Management of California’s Water Resources, in: Climatic Change, Springer, The Netherlands, 89, Supplement 1, 91–108, 2008.
- Backlund, P., Janetos, A., Schimel, D., Hatfield, J., Boote, K., Fay, P., Hahn, L., Izaurralde, C., Kimball, B. A., Mader, T., Morgan, J., Ort, D., Polley, W., Thomson, A., Wolfe, D., Ryan, M. G., Archer, S. R., Birdsey, R., Dahm, C., Heath, L., Hicke, J., Hollinger, D., Huxman, T., Okin, G., Oren, R., Randerson, J., Schlesinger, W., Lettenmaier, D., Major, D., Poff, L., Running, S., Hansen, L., Inouye, D., Kelly, B. P., Meyerson, L., Peterson, B., and Shaw, R.: The effects of climate change on agriculture, land resources, water resources, and biodiversity in the United States, A Report by the US Climate Change Science Program and the Subcommittee on Global Change Research, US Department of Agriculture, Washington, D.C., USA, 362 pp., 2008.
- Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, *Nature*, 438, 303–309, 2005.

- Barnett, T., Pierce, D. W., Hidalgo, H., Bonfils, C., Santer, B. D., Das, T., Bala, G., Wood, A. W., Nazawa, T., Mirin, A., Cayan, D., and Dettinger, M.: Human-induced changes in the hydrology of the western United States, *Science*, 319, 1080–1083, doi:10.1126/science.1152538, 2008.
- Barsugli, J., Anderson, C., Smith, J., and Vogel, J.: Options for Improving Climate Modeling to Assist Water Utility Planning for Climate Change, Water Utility Climate Alliance, San Francisco, California, 2009.
- Brekke, L. D., Dettinger, M. D., Maurer, E. P., and Anderson, M.: Significance of model credibility in estimating climate projection distributions for regional hydroclimatological risk assessments, *Climatic Change*, 89, 371–394, 2008.
- Christensen, N. S. and Lettenmaier, D. P.: A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin, *Hydrol. Earth Syst. Sci.*, 11, 1417–1434, doi:10.5194/hess-11-1417-2007, 2007.
- Christensen, N. S., Wood, A. W., Voisin, N., and Lettenmaier, D.: The Effects of Climate Change on the Hydrology and Water Resources of the Colorado River basin, *Climatic Change*, 62, 337–363, 2004.
- Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, *Clim. Dynam.*, 38, 527–546, doi:10.1007/s00382-010-0977-x, 2012.
- Gleckler, P. J., Taylor, K. E., and Doutraiaux, C.: Performance metrics for climate models, *J. Geophys. Res.*, 113, D06104, doi:10.1029/2007JD008972, 2008.
- Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dynam.*, 37, 407–418, doi:10.1007/s00382-010-0810-6, 2011.
- Hayhoe, K., Cayan, D., Field, C., Frumhoff, P. C., Maurer, E. P., Miller, N. L., Moser, S. C., Schneider, S. H., Cahill, K. N., Cleveland, E. E., Dale, L., Drapek, R., Hanemann, R. M., Kalkstein, L. S., Lenihan, J., Lunch, C. K., Neilson, R. P., Sheridan, S. C., and Verville, J. H.: Emissions pathways, climate change, and impacts on California, *P. Natl. Acad. Sci.*, 101, 12422–12427, 2004.
- Liang, S., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, 99, 14415–14428, 1994.
- Liang, S., Lettenmaier, D. P., and Wood, E. F.: One-Dimensional Statistical Dynamic Representation of Subgrid Spatial Variability of Precipitation in the Two-Layer Variable Infiltration Capacity Model, *J. Geophys. Res.*, 101, 21403–21422, 1996.
- Lohmann, D., Raschke, E., Nijssen, B., and Lettenmaier, D. P.: Regional Scale Hydrology: I. Formulation of the VIC-2L Model Coupled to a Routing Model, *Hydrol. Sci.*, 43, 131–142, 1998a.
- Lohmann, D., Raschke, E., Nijssen, B., and Lettenmaier, D. P.: Regional Scale Hydrology: II. Application of the VIC-2L Model to the Weser River, Germany, *Hydrol. Sci.*, 43, 143–158, 1998b.
- Maurer, E. P.: Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California under two emissions scenarios, *Climatic Change*, 82, 309–325, 2007.
- Maurer, E. P. and Hidalgo, H. G.: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods, *Hydrol. Earth Syst. Sci.*, 12, 551–563, doi:10.5194/hess-12-551-2008, 2008.
- Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., and Nijssen, B.: A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States, *J. Climate*, 15, 3237–3251, 2002.
- Maurer, E. P., Brekke, L., Pruitt, T., and Duffy, P. B.: Fine-resolution climate projections enhance regional climate change impact studies, *EOS Trans. AGU*, 88, 504, 2007a.
- Maurer, E. P., Stewart, I. T., Bonfils, C., Duffy, P. B., and Cayan, D.: Detection, attribution, and sensitivity of trends toward earlier streamflow in the Sierra Nevada, *J. Geophys. Res.*, 112, D11118, doi:10.1029/2006JD008088, 2007b.
- Maurer, E. P., Adam, J. C., and Wood, A. W.: Climate model based consensus on the hydrologic impacts of climate change to the Rio Lempa basin of Central America, *Hydrol. Earth Syst. Sci.*, 13, 183–194, doi:10.5194/hess-13-183-2009, 2009.
- Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T.: Decadal prediction: Can it be skillful?, *B. Am. Meteorol. Soc.*, 90, 1467–1485, doi:10.1175/2009BAMS2778.1, 2009.
- Milly, P. C. D., Dunne, K. A., and Vecchia, A. V.: Global pattern of trends in streamflow and water availability in a changing climate, *Nature*, 438, 347–350, 2005.
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R., Kundzewicz, Z., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, *Science*, 319, 573–574, 2008.
- Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., Tarpley, J. D., Lettenmaier, D. P., Marshall, C. H., Entin, J. K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, B. H., and Bailey, A. A.: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, 109, D07S90, doi:10.1029/2003JD003823, 2004.
- Nakicenovic, N., Davidson, O., Davis, G., Grübler, A., Kram, T., Lebre La Rovere, E., Metz, B., Morita, T., Pepper, W., Pitcher, H., Sankovski, A., Shukla, P., Swart, R., Watson, R., and Dadi, Z. (Eds.): Emissions Scenarios, A Special Report of Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, 599 pp., 2000.
- Nash, L. L. and Gleick, P. H.: The Sensitivity of Streamflow in the Colorado River basin to Climatic Changes, *J. Hydrol.*, 125, 221–241, 1991.
- Nijssen, B., O'Donnel, G. M., and Lettenmaier, D. P.: Predicting the Discharge of Global Rivers, *J. Climate*, 14, 3307–3323, 2001.
- Painter, T. H., Barrett, A. P., Landry, C. C., Neff, J. C., Cassidy, M. P., Lawrence, C. R., McBride, K. E., and Farmer, G. L.: Impact of disturbed desert soils on duration of mountain snow cover, *Geophys. Res. Lett.*, 34, L12502, doi:10.1029/2007GL030284, 2007.
- Painter, T. H., Deems, J. S., Belnap, J., Hamlet, A. F., Landry, C. C., and Udall, B.: Response of Colorado River runoff to dust radiative forcing in snow, *P. Natl. Acad. Sci.*, 107, 17125–17130, 2010.

- Payne, J. T., Wood, A. W., Hamlet, A. F., Palmer, R. N., and Lettenmaier, D. P.: Mitigating the effects of climate change on the water resources of the Columbia River basin, *Climatic Change*, 62, 233–256, 2004.
- Pennell, C. and Reichler, T.: On the Effective Number of Climate Models, *J. Climate*, 24, 2358–2367, doi:10.1175/2010JCLI3814.1, 2011.
- Pierce, D. W., Barnett, T. P., Santer, B. D., and Gleckler, P. J.: Selecting global climate models for regional climate change studies, *P. Nat. Acad. Sci. USA*, 106, 8441–8446, 2009.
- Prairie, J. and Callejo, R.: Natural Flow and Salt Computation Methods, Calendar Years 1971–1995, US Department of the Interior, Bureau of Reclamation, Salt Lake City, Utah, 122 pp., 2005.
- Reclamation: Sensitivity of Future CVP/SWP Operations to Potential Climate Change and Associated Sea Level Rise, Appendix R in Biological Assessment on the Continued Long-term Operations of the Central Valley Project and the State Water Project, prepared by Bureau of Reclamation, US Department of the Interior, Sacramento, California, 134 pp., 2008.
- Reclamation: Colorado River Basin Natural Flow and Salt Data, <http://www.usbr.gov/lc/region/g4000/NaturalFlow/index.html>, last access: 2 April 2009.
- Reclamation: Climate Change and Hydrology Scenarios for Oklahoma Yield Studies, Bureau of Reclamation, US Department of the Interior, Denver, Colorado, 2010.
- Reclamation: Colorado River Basin Water Supply and Demand Study, Interim Report No. 1, Status Report, Bureau of Reclamation, US Department of the Interior, Boulder City, Nevada, 2011a.
- Reclamation: Colorado River Basin Water Supply and Demand Study, Interim Report No. 1, Technical Report B – Water Supply Assessment, Bureau of Reclamation, US Department of the Interior, Boulder City, Nevada, 2011b.
- Reclamation: West-Wide Climate Risk Assessments: Bias-Corrected and Spatially Downscaled Surface Water Projections, Bureau of Reclamation, US Department of the Interior, Denver, Colorado, 2011c.
- Revelle, R. R. and Waggoner, P. E.: Effects of a Carbon Dioxide Induced Climatic Change on Water Supplies in the Western United States, in: *Changing Climate*, National Academy of Sciences Press, Washington, D.C., 1983.
- Sankarasubramanian, A., Vogel, R. M., and Limbrunner, J. F.: Climate elasticity of streamflow in the United States, *Water Resour. Res.*, 37, 1771–1781, 2001.
- Santer, B. D., Taylor, K. E., Gleckler, P. J., Bonfils, C., Barnett, T. P., Pierce, D. W., Wigley, T. M. L., Mears, C., Wentz, F. J., Brüggemann, W., Gillett, N. P., Klein, S. A., Solomon, S. P., Stott, A., and Wehner, M. F.: Incorporating model quality information in climate change detection and attribution studies, *P. Natl. Acad. Sci.*, 106, 14778–14783 2009.
- Seager, R. M. and Vecchi, G. A.: Greenhouse warming and the 21st Century hydroclimate of southwestern North America, *P. Natl. Acad. Sci.*, 107, 21277–21282, 2010.
- Seager, R., Ting, M., Held, I., Kushnir, Y., Lu, J., Vecchi, G., Huang, H., Harnik, N., Leetmaa, A., Lau, N., Li, C., Velez, J., and Naik, N.: Model Projections of an Imminent Transition to a More Arid Climate in Southwestern North America, *Science*, 10, 1126, doi:0.1126/science.1139601, 2007.
- Stainforth, D. A., Downing, T. E., Washington, R., Lopez, A., and New, M.: Issues in the interpretation of climate model ensembles to inform decisions *Philos. T. Roy. Soc. A*, 365, 2163–2177, 2007.
- Stockton, C. W. and Boggess, W. R.: Geohydrological implications of climate change on water resource development, Final Report, US Army Coastal Engineering Research Center, Fort Belvoir, Virginia, p. 224, 1979.
- Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, *Philos. T. Roy. Soc. A*, 365, 2053–2075, doi:10.1098/rsta.2007.2076, 2007.
- Tebaldi, C., Mearns, L. O., Hychka, D., and Smith, R. L.: Regional probabilities of precipitation change; A Bayesian analysis of multimodel simulations, *Geophys. Res. Lett.*, 31, L24213, doi:10.1029/2004GL021276, 2004.
- Van Rheezen, N. T., Wood, A. W., Palmer, R. N., and Lettenmaier, D. P.: Potential implications of PCM climate change scenarios for Sacramento-San Joaquin River Basin hydrology and water resources, *Climatic Change*, 62, 257–281, 2004.
- Wilby, R. L.: Evaluating climate model outputs for hydrological applications, *Hydrolog. Sci. J.*, 55, 1090–1093, 2010.
- Wilby, R. L. and Harris, I.: 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, 2006.
- Wood, E. F., Lettenmaier, D. P., and Zartarian, V. G.: A Land-Surface Hydrology Parameterization with Subgrid Variability for General Circulation Models, *J. Geophys. Res.*, 97, 2717–2728, 1992.
- Wood, A. W. and Lettenmaier, D. P.: A testbed for new seasonal hydrologic forecasting approaches in the western U.S., *B. Am. Meteorol. Soc.*, 87, 1699–1712, doi:10.1175/BAMS-87-12-1699, 2006.
- Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern United States, *J. Geophys. Res.-Atmos.*, 107, 4429, doi:10.1029/2001jd000659, 2002.
- Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change*, 15, 189–216, 2004.
- Wood, A. W., Kumar, A., and Lettenmaier, D. P.: A retrospective assessment of climate model-based ensemble hydrologic forecasting in the western U.S., *J. Geophys. Res.*, 110, D04105, doi:10.1029/2004JD004508, 2005.
- WCRP – World Climate Research Programme’s: Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset, Archive of downscaled climate projections; served at: http://gdo-dcp.ucllnl.org/downscaled_cmip3_projections/, last access: 22 April 2009.