Summer 6-14-2019

Observed and Projected Snowmelt Runoff in the Upper Rio Grande in a Changing Climate

Nels R. Bjarke
University of New Mexico - Main Campus

Nels Bjarke

Candidate

Earth and Planetary Science

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. David Gutzler, Chairperson

Dr. Peter Fawcett

Dr. Joseph Galewsky

OBSERVED AND PROJECTED SNOWMELT RUNOFF IN THE UPPER RIO GRANDE IN A CHANGING CLIMATE

by

NELS BJARKE

B.S., Earth and Planetary Science, University of New Mexico, 2014

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Earth and Planetary Sciences

The University of New Mexico
Albuquerque, New Mexico

July, 2019

ACKNOWLEDGEMENTS

This material is based upon work supported by the USGS South Central Climate Adaptation Science Center, under award number 2A16-0424. We thank the U.S. Bureau of Reclamation for making their "Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections" available to us.

(archive at https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/)

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Observed and Projected Snowmelt Runoff in the Rio Grande Headwaters in a Changing Climate

by Nels Bjarke

B.S., Earth and Planetary Science, University of New Mexico, 2014

ABSTRACT

As climate has warmed over the past half century, the strength of the covariance between interannual snowpack and streamflow anomalies in the Rio Grande headwaters has decreased. This change has caused an amplification of errors in seasonal streamflow forecasts using traditional statistical forecasting methods, based on the diminishing correlation between peak snow water equivalent (SWE) and subsequent snowmelt runoff. Therefore, at a time when water resources in south-western North America are becoming scarcer, water supply forecasters need to develop prediction schemes that account for the dynamic nature of the relationship between precipitation, temperature, snowpack and streamflow. We quantify temporal changes in statistical predictive models of streamflow in the upper Rio Grande basin using observed data, and interpret the results in terms of processes that control runoff season discharge. We then compare these observed changes to corresponding statistics in downscaled global climate models (GCMs), to gain insight into which GCMs most appropriately replicate the dynamics of interannual streamflow variability represented by the hydro-climate parameters in the headwaters of the Rio Grande. We quantify how the correlations among temperature, precipitation, SWE, and

streamflow have changed over the last half century within the local climatic and hydrological system. We then assess different long-term GCM-based streamflow projections by their ability to reproduce observed relationships between climate and streamflow, and thereby better constrain projections of future flows as climate warms in the 21st century. In the Rio Grande system, we find that spring season precipitation increasingly contributes to the variability of runoff generation as the contribution of snowpack declines.

TABLE OF CONTENTS

1. Introduction

Regional climate trends in southwestern North America are observed and projected to follow the global trend of warming temperatures, which will have a significant impact on the hydrological system that supplies water to millions who live along the Rio Grande and beyond. As the climate warms, snowpack extent and volume across western North America are projected to decrease in magnitude (Mote et al., 2006), which has been shown in the observed record (Chavarria & Gutzler, 2018) in conjunction with a shift towards earlier maximum snowmelt runoff (Cayan et al., 2001; Stewart, 2009). In addition to these effects, warmer temperatures necessarily imply that more of the annual precipitation in the southwest will fall as rain rather than snow (Knowles, 2006; Barnett et al., 2008). This shift can be problematic for water supply forecasters and water managers, as historically, snow volume has been highly correlated with streamflow while liquid precipitation only offers a moderate correlation (Chavarria & Gutzler, 2018). Within the context of the Rio Grande watershed, a decrease in long-term seasonal forecast certainty and water availability, combined with increasing water demands, could have serious economic consequences for those who depend on the annual water supply.

Despite the projected decrease in snowmelt derived runoff, a significant long-term decrease in the overall streamflow of the Rio Grande has not been observed in recent decades. This is primarily due to a masking effect from an increase in the input from spring precipitation as snowpack declines (Chavarria & Gutzler, 2018). Unfortunately for water managers, as spring precipitation becomes more important to streamflow forecasting, predictions of runoff volumes become less certain due to the high variability and low predictability of the precipitation input. Compounding this uncertainty is the uncertain role that temperature will play as a predictor of streamflow (Lehner et al., 2017b).

Though the short observational record introduces significant analysis uncertainty associated with interannual variability (Deser et al., 2012), there are physical processes that support the conclusion that snowpack has more variable contribution to streamflow under regional warming conditions. Earlier peak snowpack timing (Mote et al., 2006), shorter ablation season (Hurd & Coonrod, 2012), and increased sublimation off the snowpack are possible contributors to the reduction of the natural snowpack reservoir in the Rio Grande headwaters. Dust on snow processes have also been attributed to the decline in the duration of snowpack in the southwest US (Painter et al., 2007; Livneh et al., 2015). All of these physical processes are captured implicitly in statistical streamflow forecast methods, so it is not the aim of this study to directly address the specific driver(s) of the decline in the strength of the snowpack-runoff relationship. Instead, we seek to form a broader assessment of the impact that an observed warming trend is having on the seasonal predictability of the Rio Grande snowmelt driven runoff system within the observed record and compare that assessment to published model projections.

Previous studies on the snowmelt driven river systems in the southwest have applied temperature as both a direct predictive tool for seasonal streamflow forecasts (Lehner et al., 2017b) and a first-order predictor for projected trends in streamflow decline through the 21st century (Udall & Overpeck, 2017). This study critically examines the direct application of temperature as a seasonal predictor of streamflow by using calculations of the evolving correlations between streamflow, snowpack, spring precipitation and temperature in the Rio Grande headwater region within the last half-century. In particular, we anchor out analysis of runoff using snowpack, which is known from previous work to be the first-order generator of streamflow in the major rivers of southwest North America (Garen, 1992). We test the efficacy of temperature, snowpack, and spring precipitation as prediction parameters for seasonal streamflow using a statistical framework to show how the hydrological system is changing within the context of a warming climate.

Our goal is to constrain uncertainty in projections of future flows as climate warms in the 21st century. We document the multidecadal changes in covariate relationships between interannual fluctuations of climate variables and streamflow in historical observations. We then assess model performance based on the ability of individual models to reproduce trends in changing correlations between snowpack, temperature, spring precipitation, and streamflow observed in the historical data. Observationally consistent models are defined from this assessment. The subset of model projections that are defined as observationally consistent is shown to exhibit less spread in projected streamflow through the 21st century than the full ensemble of model projections that we consider.

We also interpret our results to assess the changing contribution of climate parameters to runoff season streamflow within in the context of temperature and precipitation in the Rio Grande headwater region. This allows us to determine the cause of deficiencies in seasonal water supply outlooks that rely on stationary relationships between climate parameters and streamflow.

2. Data Sources

2.1 Historical Data

2.1.1 Snowpack. Snowpack is quantified as snow water equivalent (SWE) taken from the National Resources Conservation Service (NRCS) snow course dataset (*https : //wcc.sc.egov.usda.gov/nwcc/rgrpt*?*report* = *snowcourse*). April 1st SWE (SWE_A) values are used in this study to represent maximum snowpack depth in the Rio Grande headwaters, which has been shown to be the case historically (Chavarria & Gutzler, 2018).

The earliest continuous record of SWE from the snow course data begins in 1951, hence the period of record for this analysis begins in 1951. The extended length of the snow-course data set relative to SNOTEL is preferable here for its usefulness in understanding the longer-term dynamics of the hydrological system within the context of anthropogenic climate change. It should be noted that SWE values obtained from snow-course data are lower in magnitude than SWE values obtained from SNOTEL datasets due to lower elevation of snow-course measurement sites (Chavarria & Gutzler, 2018). However, because SNOTEL and snow-course data show similar interannual variability, and this study relies on methods that examine covariate relationships of climate parameters, this discrepancy in snow magnitude should not be not be problematic unless decline in snowcourse SWE_A is unrepresentative of higher elevation snowpack.

As regional temperatures increase, it could prove effective to implement a more dynamic approach to classification of maximum snowpack by considering March SWE as maximum snow depth for later time periods in order to capture the shift towards earlier snowmelt timing (Cayan et al., 2001; Stewart, 2009). However, in this study we found that accounting for the possible shift towards earlier maximum snowpack yielded no significantly different results than considering April SWE as maximum snowpack for the whole observed record within the framework of this analysis. We therefore use 1 April SWE (SWE_A) exclusively in the analysis, and interpret SWE_A as the annual metric of peak snowpack in the Rio Grande headwaters basin.

2.1.2 Streamflow. Daily mean discharge rates measured at the Del Norte streamgauge were taken from the National Water Information System (NWIS) run by the United States Geological Survey (USGS) (*https//waterdata.usgs.gov/nwis/inventory/?site no* = 08220000). The Del Norte streamgauge was chosen in this study due to its location upstream from major population centers or agricultural diversions, which allows us to ignore the anthropogenic effects of water withdrawal or diversion from the Rio Grande (Mix et al., 2012; Chavarria & Gutzler, 2018; Blythe & Schmidt, 2018). Average daily discharge rates from 1951-2015 were converted into total monthly discharge values by first converting the daily mean flow rate to a total volume of water for each day. Then, total volume for each day of a month is added to arrive at a total sum for monthly water volume discharge. This allows us to relate the depth of snow and liquid precipitation to the volume of discharge that flows past the Del Norte gauge.

Monthly total values of discharge are summed into annual runoff season values. The runoff season in this study is defined as April-June, so the total monthly discharges for these months are added together to create a data set of total annual runoff season discharge (Q_{RO}) from 1951-2015. Other classifications of the runoff season were considered such as March-June and April-July, however March 1st snowpack did not fully capture the extent of the maximum annual snowpack and monsoonal rains in mid-late July confounded interpretation of Q_{RO} as derived from snowmelt.

2.1.3 Temperature and Precipitation. In order to capture the temperature and precipitation across the entire Rio Grande Headwater region (Figure 1b), values were obtained from Oregon State's WestMap PRISM dataset over the area of study. PRISM data are particularly useful for this study due to their spatial interpolation of precipitation and temperature using observed point measurements across a high elevation region where a comprehensive data coverage is not readily available (Daly, 2008). Monthly average max temperature and monthly total precipitation are obtained from PRISM and are converted into seasonal values. Winter season is denoted here as December-March (T_{WI},P_{WI}) and the spring season is denoted as April-June (T_{SP},P_{SP}), such that SWE_A represents snowpack following the winter season and prior to the spring season. The spring season is particularly useful for this study as it allows us to observe the increasingly impactful runoff season precipitation (Chavarria & Gutzler, 2018) and the usefulness of temperature as a seasonal predictor for discharge (Lehner et al., 2018).

https : *//cefa.dri.edu/Westmap/Westmap home.php*?*page* = *timeseries.php*

2.2 Climate Model Output.

2.2.1 Historical and Projected GCM Output. Simulated climate data are obtained from the Bureau of Reclamation (BOR) published bias-corrected spatially-disaggregated (BCSD) CMIP5 model ensemble output (Reclamation, 2013). This global climate model (GCM) output set is produced by using statistical spatial disaggregation methods to increase the spatial resolution of output from the CMIP5 ensemble to 1/8th degree square grid cells. Precipitation and temperature output from individual GCMs are bias-corrected in order to allow for comparison to historical values (Reclamation, 2013). Access to this publicly available dataset and explanation of methods used in its generation can be found at: *https : //gdo − dcp.ucllnl.org/downscaled cmip projections/dcpInterface.html*

Spatial extent of the GCM output is chosen by selecting the location of Del Norte streamgauge as the pour point for streamflow, and all grid cells in which precipitation falls into the upstream watershed and thereby contributes to streamflow at Del Norte are considered here. Streamflow is produced in the BOR projections by feeding the precipitation and temperature output from each individual GCM into the Variable Infiltration Capacity (VIC) land-surface model (Reclamation, 2014). Monthly values of projected climate parameters are aggregated into annual values of total seasonal discharge, total seasonal precipitation, and mean seasonal temperature for purposes of comparing historical observations to climate model projection data. April 1st SWE is also used to represent maximum snowpack for the model projections for comparison of the retrospective model simulations to the historical observations.

3. Methods

3.1 Historical Observation Analysis

3.1.1 Correlations. The initial step for understanding how streamflow is modulated by the different climate parameters is to systematically calculate the covariate relationships between interannual fluctuations of SWE, precipitation, temperature, and streamflow. We identify two different time periods for this approach: an early time period (1951-1983) which represents climate minimally impacted by anthropogenic climate change and a late time period (1983-2015) which represents climate more significantly impacted by anthropogenic climate change (Chavarria & Gutzler, 2018). We generate a correlation table for each time period by calculating the Pearson correlation between interannual fluctuations of each pair of parameters for a given period. This allows us to quantify covariation between individual parameters and, by comparing the results from the two time periods, understand how those covariate relationships have changed between early and late periods in the observed record.

3.1.2 Step-Wise Linear Regression Models. We implement statistical models for the two different time periods using a step-wise approach for the purpose of assessing the predictive utility of individual parameters to account for interannual variability of streamflow, and to clarify mutual correlation effects on the interannual variability of streamflow. SWE_A is used as the first order predictor for all cases in this method, as it is shown to have the highest direct correlation with runoff season discharge for all observational times periods (Garen, 1992) and is directly physically related to the subsequent Q_{RO}. Second and third order predictors are added to the models in varied sequences in order to understand how much interannual streamflow variability can be attributed to parameters individually, after the contribution of snowpack has been accounted for. This method is performed as follows:

1. Each model is trained on total runoff season discharge for each year in the given time period using a linear regression with SWE_A for each year to produce derived linear coefficients.

2. SWE_A values for each year are multiplied by the linear regression coefficients derived in step 1 to produce a linear hindcast of Q_{RO} based only on SWE_A values.

3. A vector of residuals is produced by subtracting the hindcasted streamflow produced in step 2 from the observed Q_{RO} for each year.

4. The next set of linear models is trained on the vector of residuals produced in step 3 using a linear regression with either P_{SP} or T_{SP}.

5. Either P_{SP} or T_{SP} values (whichever is used in step 4) is multiplied by the linear coefficients derived in step 4 to produce a linear hindcast of the residuals based on whichever climate parameter was used in this step.

6. Steps 3-5 are repeated for the parameter not selected for use in steps 4 & 5.

A schematic summary of the order of the steps of the regression models is shown below. The third column only shows two steps as a bivariate approach to the addition of spring precipitation and temperature is applied here. The bivariate method is used to observe how the results would change if no order preference is given to either precipitation or temperature. The results of the step-wise linear regression models give us insight into the contribution of predictive climate parameters to Q_{RO} for all time periods considered. We are able to understand the contribution of correlated climate parameters to Q_{RO} by adding parameters to the regression models in different orders and comparing the results of the prediction skill associated with different ordered models.

1. Step1:	$Q = aSWE_A + b$	$Q = aSWE_A + b$	$Q = aSWE_A + b$
2. Step2:	$Q1_1 = cT_{SP} + d$	$Q1_2 = cP_{SP} + d$	$Q1_3 = cP_{SP} + dT_{SP} + e$
3. Step3:	$Q2_1 = eP_{SP} + f$	$Q2_2 = eT_{SP} + f$	

We assess the overall performance of each statistical model and the increase in predictive skill associated with the addition of each parameter. The entire timespan (1951-2015) will be evaluated using all model formats, in addition to regression based only on the early (1951-1983) and late (1983-2015) time periods. We also examine several overlapping 30-year periods that progress from the early to late time period to compare how the predictive skill associated with the addition of each parameter progresses through the entire observational time period.

The models described above allow us to determine the skill attributable to individual predictive parameters in seasonal streamflow hindcasts for a given (dependent) time period, but do not allow us to assess how well the models will perform when applied to different (independent) time periods. To examine how well the models perform when applied to independent time periods, we will apply the derived parameter coefficients from one time period to the data from the alternate time period and compare the results of the linear prediction of Q_{RO} to the observed historical data. This will allow us to determine biases that are present from derived parameters in the statistical models and allow us to observe shifts in parameter behavior through time.

We use several metrics to analyze model skill. Root-mean-square error (RMSE) diagnoses the average annual error in streamflow prediction and the absolute error reduced with the addition of each parameter in each step. We calculate the RMSE reduction associated with the addition of each parameter after the initial SWE regression in order to compare the error reduction across all statistical models.

3.2 BOR BCSD GCM Output Analysis.

3.2.1 Model Projections. We first examine the BCSD output of the all GCM projections in order to observe ensemble trends and model spread. We subdivide the output into the four representative concentration pathways (RCPs) used in the CMIP5 model ensemble. Winter (Dec-

Mar) and spring (Apr-Jun) temperature and precipitation, April 1st SWE, and total runoff season (Apr-Jun) discharge are all examined for each RCP. A moving-window 30-year average for each parameter is applied to all members of each RCP ensemble to observe trends in mean climatology (IPCC, 2013). We use the ensemble mean change in each parameter from the observational time period (1960-1989) to the late 21st century (2050-2079) as metric to determine long-term changes within the entire ensemble for each RCP forcing.

The bias-correction applied to each GCM simulation fits the cumulative distribution function (CDF) of each parameter for each month produced by the simulation to the CDF of the same parameter in an observational dataset for each month (Reclamation, 2013). The bias correction is applied to both the retrospective simulation period (1950-1999) and the projected simulation period (2000-2099). Therefore, the spread in the projections within each RCP ensemble is associated with model spread and not a significant change in the interannual variability.

3.2.2 Step-Wise Linear Regression Models. The same step-wise regression techniques developed for historical observations are applied to the BCSD CMIP5 model outputs for four 30-year time periods through the end of the 21st century (1960-1989,1990-2019,2020-2049,2050-2079). We compare the results of regression models applied to each member of the BCSD CMIP5 ensemble to results of regression models applied to historical observations by comparing the fraction of interannual streamflow variability associated with each climatological parameter in both historical and future epochs.

4. Changing Snowpack-Streamflow Relationships

4.1 Parameter Correlations. Pearson correlations between parameters (Table 1) reveal, as expected, a significant linear relationship between SWE_A and Q_{RO} in both early and late time periods analyzed. However, there is a dramatic decline in the strength of this correlation between these two periods coincident with the onset of significant warming trends over the region (Figure 2). SWE_A accounts for 79% of the interannual variability of runoff season discharge for years 1951-1983, which decreases to 45% for the years 1983-2015. Despite this significant decrease in correlation and a downward trend in headwater snowpack over the same time period, there is no significant downward trend in total runoff season discharge in the historical observations at Del Norte (Chavarria & Gutzler, 2018).

Conversely, the strength of the correlation between annual anomalies of P_{SP} and Q_{RO} increases through the span of 1951-2015. Early time period observations show that P_{SP} accounts for 11% of the interannual variability of Q_{RO}, which doubles to 22% for the late time period. Earlier work has noted this observed trend in the Rio Grande headwaters (Chavarria & Gutzler, 2018). We investigate further the significance of spring precipitation in the hydrological system and its relationship with spring temperature, a parameter that has been given significant attention in previous endeavors to understand climate change impacts on streamflow (Vano et al., 2014; Udall & Overpeck, 2017; Lehner et al., 2017b).

A well understood feature of regional climatology, shown in Table 1, is the strong (negative) correlation between fluctuations of precipitation and temperature. We observe this relationship in the Rio Grande headwater region, which complicates the interpretation of either individual parameter as the cause of streamflow variability within a single linear regression. The structure of the step-wise regression models used in this study allows us to understand the contribution of P_{SP} and T_{SP} individually to the interannual variability of Q_{RO} in the observed record.

4.2 Regression Results. The significant changes in the direct contribution of predictive parameters to the interannual variability of Q_{RO} are coincident with the onset of observable warming trends in the Rio Grande headwater region (Figure 2). In the early period of the historical record (1951-1983), we observe that SWE_A accounts for a large fraction (79%) of the interannual variability of streamflow at the Del Norte stream gauge (Table 2d). The addition of P_{SP} and/or T_{SP} in subsequent steps of the step-wise regression yields minimal and non-significant error reduction in all three of the statistical model structures. Neither P_{SP} or T_{SP} is significantly contributing directly to total runoff season discharge variability independent of SWE_A for this time period.

In the later time period, SWE_A has much less predictive power. Less than half (45%) of the interannual variability of Q_{RO} can be accounted for with only SWE_A as a predictor in the linear regression (Table 2e). When added as a second predictor in the step-wise structure, both P_{SP} and T_{SP} terms significantly reduce the error of the regression model (Step 2 in Table 2e, middle and right columns). P_{SP} added to the model structure as a third order predictor, after SWE_A and T_{SP}, still reduces the error of the model significantly (Step 3 in Table 2e, middle column). However, when spring temperature is added as a third order predictor, after SWE_A and P_{SP}, there is no significant error reduction in the model (Step 3 in Table 2e, right column).

For both time periods, a third model structure that applies a bivariate approach to the addition of spring precipitation and temperature (not shown) yields nearly identical results to the model structure that adds predictive parameters in the stepwise order $SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$. The bivariate approach gives no priority to the weight of either P_{SP} or T_{SP} in the regression scheme. Nearly identical results of the bivariate approach and the $SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$ ordered regression is consistent with the results of the three step regressions that imply an increasingly important role of P_{SP} on the interannual variability of Q_{RO}, more so than the addition of T_{SP}.

4.3 Full Ensemble BCSD GCM Projections. We examine the climatic changes and uncertainty of all climate parameters considered in the historical observations in the Rio Grande headwaters basin associated with each ensemble of CMIP5 simulation, separately for each emission scenarios. We identify the climatic changes by examining the difference in the ensemble mean for each parameter between 30-year periods in the latter half of the 21st century (2050-2079), and an earlier epoch during the historical time period (1960-1989). We begin the analysis of the BCSD GCM simulations in 1960 to avoid any model spin-up effects.

When examining the ensemble mean of the entire ensemble for each emission scenario of the BCSD GCM output, we see that there is significant dependence on emission scenario through the end of the 21st century (Table 3a). For simplification of discussion, we will specifically examine the results for the lowest emission scenario (RCP2.6) and the highest emission scenario (RCP8.5) as the two end members of this analysis, with the two middle emission scenarios (RCP4.5 and RCP6.0) producing ensemble means that lie between the end member ensemble means.

We note that the 30-year means of all climate variables, with the exception of temperature, show significant sampling uncertainty when considering each entire emission scenario ensemble due to natural variability. The following is a summary of the results of Table 3a for the differences in the high and low emission scenarios:

1. SWE_A decreases slightly for the RCP2.6 ensemble means, and declines significantly for the RCP8.5 ensemble means.

2. The RCP2.6 ensemble mean P_{SP} increases slightly, but will be considered as no significant change due to the uncertainty associated with the mean shift ($\Delta = 0.14cm$ & $\sigma = 0.35cm$). The RCP8.5 ensemble mean P_{SP} decreases significantly (~ 5%) through the end of the 21st century.

3. The T_{SP} RCP2.6 and RCP8.5 ensemble means increase by ~ 1°C and ~ 2°C respectively by the latter half of the 21st century.

4. The RCP2.6 ensemble mean Q_{RO} increases slightly and the RCP8.5 ensemble mean Q_{RO} decreases slightly. However, significant uncertainty associated with all of the emission scenario ensembles reduces our confidence to assert significant change in Q_{RO} for any emission scenario.

We examine the dependence of a mean shift in Q_{RO} on changes in mean P_{SP}, SWE_A, and T_{SP} from the retrospective simulation period (1960-1989) to the late 21st century (2050-2079) for each simulation in the entire ensemble (Figures 4a-4c). Consistent with the contribution of climate parameters to Q_{RO} variability in the observational data, mean changes in simulated Q_{RO} are highly correlated with changes in mean P_{SP} and SWE_A, but have no significant correlation with T_{SP} for the entire ensemble. There is emission scenario dependence on the location of the centroid of each subset of simulations, but the linear dependence of changes in mean Q_{RO} on changes in mean SWE_A and P_{SP} is consistent for all RCP subsets. Equally, none of the RCP subsets reveal any linear dependence of mean changes Q_{RO} on mean changes of $T_{SP.}$

5. Selection of Observationally Consistent Models

5.1 Selection Metrics. The comparison of parameter contribution to Q_{RO} in the BCSD models within the observational time period allows us to determine which models are effectively reproducing the trends in parameter contribution to Q_{RO} in observational data over the historical period.

In order to select Observationally Consistent Models (OCMs), models that most effectively simulate the evolving climate-hydrology relationship with reference to observations, we generate metrics based on results from the step-wise linear regression models applied to the observational data. By selecting models that are consistent with observational trends in parameter contribution to streamflow variability in the historical time period, we are potentially able to reduce uncertainty in projections of streamflow through the end of the 21st century that arises from significant model spread in the entire ensemble.

The selection of OCMs is based on a set of criteria derived from observed trends in parameter skill that result from the step-wise regression models applied to the historical data. We define the criteria for the selection of OCMs as follows:

Criterion 1: Is a majority of the interannual variability of Q_{RO} attributed to SWE_A during the early period (1960-1989) of the retrospective simulation? A fraction of the interannual variability of Q_{RO} attributed to SWE_A ($r^2 > 0.6$) determined by $r^2(Q_{RO},SWE_A)$ resultant from the first step of the step-wise regression models during the early time period (1960-1989) of the retrospective simulation.

Historical $r^2(Q_{RO},SWE_A) = 0.79$

Criterion 2: Does the fraction of interannual variability of Q_{RO} attributed to SWE_A decrease between the two periods of the retrospective simulation? A decrease of ($r^2 > 0.1$) in the fraction of interannual variability of Q_{RO} attributed to SWE_A in the historical time period of the model projections determined by the $\Delta r^2(Q_{RO},SWE_A)$ resultant from the first step of the step-wise regression models between the early time period (1960-1989) to the later time period (1990-2019).

Historical $\Delta r^2(Q_{RO},SWE_A) = -0.34$

Criterion 3: Does P_{SP} contribute a significant fraction to the interannual variability of Q_{RO} during the late period (1990-2019) of the retrospective simulation? A fraction of the interannual variability of Q_{RO} attributed to P_{SP}($r^2 > 0.1$) determined by $r^2(Q^1{}_{RO},P_{SP})$ resultant from the second step of the step-wise regression models during the later time period (1990-2019) during the historical time period.

Historical $r^2(Q^1{}_{RO},P_{SP}) = 0.21$

Criterion 4: Does the seasonal predictability of Q_{RO} decrease between the two periods of the retrospective simulation? A decrease of ($r^2 > 0$) in the fraction of the interannual variability of Q_{RO} attributed to all three parameters using the bivariate approach to the addition of P_{SP} and T_{SP} from the early period (1960-1989) to the late period (1990-2019) of the retrospective simulations.

Historical $\Delta r^2(Q_{RO},(SWE_A,(P_{SP},T_{SP}))) = -0.13$

We note that the selection of OCMs is based on the result of applying the step-wise regression models to individual BCSD simulations, not ensemble averages. We select individual GCM simulations as observationally consistent, and therefore there is no discrimination between different RCPs or GCMs when identifying a simulation as an OCM.

Simulations from the same GCM might be observationally consistent for just a subset of the multiple RCPs for which that GCM was run due to natural variability. Equally, a subset of simulations from multiple GCMs for the same RCP can be classified as observationally consistent, while the same subset of simulations will not necessarily be classified as observationally consistent for a different RCP.

Figures 3a and 3b, along with Table 4, show the results of selecting GCM simulations as OCMs based on the criteria based on the results of step-wise regression models applied to the historical observations. From the entire ensemble of 97 simulations, 14 simulations are selected as observationally consistent based on the criteria proposed. Not all of the realizations from a single GCM are selected as observationally consistent, due to the simulation of natural variability in the GCM compared to the natural variability in the observational data. Simulations that pass most, but not all, of the criteria are highlighted in yellow in Table 4. These simulations are not considered as OCMs in the analysis that compares the OCM subset to the non-OCM subset, but are noted as a recognition of the somewhat arbitrary nature of setting a threshold to determine OCMs.

5.2 Observationally Consistent Models. To examine the ability of each BCSD GCM simulation to reproduce the climatic shifts observed in the historical observations, we apply the same step-wise regression methodology applied to the observational data to all simulations. Models are selected to be OCMs based on the four criteria proposed in section 5.1.

We examine the OCM subset of the entire ensemble of BCSD GCMs to the models not classified as OCMs to determine differences in ensemble means and variance. Due to the small population (and associated sampling uncertainty) of OCMs within each emission scenario, it is useful to compare the OCMs for all emission scenarios to the non-OCM subset for all emission scenarios. In order to make this comparison, we compare the two subsets of GCMs using the mid-century (2020-2049) projections and retrospective simulations of the historical period (1960-1989) which will allow us to avoid the differences in diverging GCM projections (shown in Table 3a) that arise from differing emission scenarios in later decades (IPCC, 2013).

The following is a summary of the results of Table 3c and Figure 5, the differences in the OCM and non-OCM subsets from the entire BCSD GCM ensemble through the mid-century:

1. Both the OCM subset (14 simulations) and the non-OCM subset (83 simulations) show similar width of distributions for the mean changes in all climate parameters except Q_{RO}, as represented by the standard deviation (σ) of the ensemble mean for each subset.

2. Both the OCM and non-OCM subsets simulate decreased ensemble mean SWE_A and P_{SP} through the mid-century. However, the simulated decrease in the ensemble mean for the OCM subset is greater in magnitude for both parameters.

3. The OCM subset simulates slightly less ensemble mean temperature increase through the mid-century compared to the non-OCM ensemble mean.

4. The OCM subset simulates a significant decrease in Q_{RO} for the ensemble mean through the midcentury, with a significantly smaller distribution than the non-OCM subset, which projects an ensemble mean increase in Q_{RO} through the mid-century.

For reference, we also compare the OCM subset of each emission scenario to their respective emission scenario ensemble with the acknowledgement that sampling uncertainty of the OCM subset is large compared to each emission scenario ensemble (Table 3b). For this comparison, we use the latter half of the 21st century (2050-2079) and the retrospective simulations of the historical period (1960-1989) because we individually consider the diverging emission scenario ensembles.

The change in each climate parameter within the OCMs through the latter half of the 21st century is dependent on the emission scenario ensemble, as expected from the same analysis applied to the entire emission ensemble (Table 3b). Compared to each entire emission scenario ensemble, the OCM ensemble means indicate significant decreases in Q_{RO}, P_{SP} and SWE_A through the end of the

21st century. The simulations have much larger distributions due to the sampling uncertainty associated with small sample sizes of the OCM subsets for each emission scenario.

We note that the comparison of changes in 30-year means of climate parameters (P_{SP}, SWE_A, and T_{SP}) and Q_{RO} through the end of the twenty-first century in Fig. 4 is not a product of step-wise regression analysis. The linear relationships between P_{SP}/SWE_A and Q_{RO} are consistent for both the OCM-subset ensemble of simulations and the entire ensemble for each RCP. However, despite there being no significant linear relationship between changes in T_{SP} and Q_{RO} based on all simulations for any of the RCPs, there <u>is</u> a significant, but weak, negative correlation between T_{SP} and Q_{RO} for the OCM-subset of simulations (Fig. 4c). This result is consistent with the analysis of hydroclimate observations that shows a strong negative correlation between regional spring temperature and precipitation anomalies (Table 1). The linear relationship between changes in T_{SP} and Q_{RO} for the OCM-subset through the end of the 21st century is a result of the same negative correlation between T_{SP} and P_{SP} and strong linear relationship between changes in P_{SP} and Q_{RO} for the same time period.

6. Discussion

6.1 Trends in the Historical Observations

6.1.1 SWE. The result of the decreased skill of SWE_A as a first-order predictor for Q_{RO} for the Rio Grande headwaters within the observational record supports previous results that indicate the inappropriateness of the stationarity assumption in historical regression techniques used for seasonal streamflow prediction, within the context of a regional warming trend (Garen, 1992; Milly et al., 2008.; Lehner et al., 2017a). This result alone has significant implications for operational streamflow forecasters that still primarily rely on regression techniques for the prediction of seasonal runoff (Garen, 1992). The decline in seasonal forecast skill associated with snowpack measurements will likely reduce the ability of regional water managers in the southwest to effectively plan for water management in snowmelt-dominated rivers (Chavarria & Gutzler, 2018).

6.1.2 Spring Precipitation and Temperature. The independent contribution of P_{SP} and T_{SP} to the interannual variability of streamflow in the early time period (1951-1983) of this analysis is essentially zero. During that 33-year period, SWE_A is such a dominant contributor to streamflow, that there is little error to reduce in the prediction scheme with the addition of P_{SP} and T_{SP}. With the onset of regional warming trends and the reduction of the contribution of SWE_A to the interannual variability of streamflow, we observe the increase in the independent predictive power of P_{SP} for Q_{RO}. We also observe an increase in the predictive power of T_{SP} as a second-order predictor of Q_{RO}, but the addition of both P_{SP} and T_{SP} as third-order predictors yields results that clarifies the effect of the strong negative correlation between the two parameters.

The error reduction of the statistical model with the addition of P_{SP} as both a second order and third order predictor is significant in the later time period. The addition of T_{SP} as a second order predictor of Q_{RO} is significant during the same time period but is not significant as a third order predictor of Q_{RO}. This reveals that when P_{SP} is accounted for in the statistical model fit to observations, T_{SP} yields no additional predictive power, but when T_{SP} is added to the model first, P_{SP} is still able to offer predictive power for a significant fraction of the interannual variability of Q_{RO}.

We interpret these results to show the increasing importance of spring precipitation as a contributor to the interannual variability of runoff season discharge, within the context of significant regional snowpack decline. There is little evidence of direct contribution of springtime temperatures on Q_{RO} through evapotranspiration processes, independent of precipitation variability. For small-scale, small-magnitude climate change, separating the effects of precipitation and temperature on streamflow can inform interpretation of climate model output (Vano et al., 2014). However, our results indicate that for larger, basin-scale climate change, precipitation and temperature cannot be treated as independent variables affecting streamflow separately (Lehner et al., 2017b; Udall & Overpeck, 2017).

The minimal contribution of temperature to Q_{RO} shown in this study challenges recent studies that have attributed declines in future streamflow in the southwestern US to direct temperature effects (Udall & Overpeck, 2017.; Lehner et al., 2017b). Applying temperature directly as a linear predictor for decline in streamflow ignores the strong negative correlation between temperature and precipitation in the region. Though there is likely skill associated with the addition of temperature as a predictor of streamflow without the addition of precipitation within a seasonal forecast framework, our results suggest that the skill associated with temperature is primarily due to snowpack decline, with residual apparent temperature effects due to the interannual correlation between temperature and precipitation.

Future seasonal forecasts that would apply temperature as direct predictor for streamflow would also rely on the stationarity of the relationship between temperature and precipitation. Our analysis find no empirical evidence that ΔQ_{RO} scales linearly with the ΔT_{SP} in the Rio Grande headwaters as implied by previous studies of the impacts of regional warming on water resources in the southwest US.

6.1.3 Seasonal Predictability. The observed decrease in seasonal forecast skill of runoff-season streamflow associated with snowpack and the increasingly important contribution of spring precipitation presents a serious problem for water managers. Methods that would increase seasonal predictability of precipitation, particularly in the spring/summer seasons, will play a critical role in the future of seasonal streamflow forecasting. There is a significant body of work that has focused on the predictability of winter precipitation in North America, primarily using El Nino Southern-Oscillation signals (Ropelewski & Halpert, 1986; Gershunov & Barnett, 1998; Deser et al., 2018), but there have not been significant advances in prediction skill for spring/summer precipitation for southwestern North America. The application of soil moisture patterns for the monthly-seasonal scale forecasts of precipitation shows some promise (Liu 2003), but there has been limited realization of increased prediction skill using soil moisture for operational use.

6.2 GCM Projections

6.2.1 Temperature. Projections of temperature through the end of the 21st century contain the least amount of uncertainty of all climate parameters that are considered in this study (IPCC, 2013; Udall & Overpeck, 2017). Ensemble means of temperatures show steady increase through the first half of the 21st century for all emission scenarios for both spring and winter temperatures. The ensemble means of different emission scenarios begin to diverge from one another at the mid-century mark, a result that is consistent with the different carbon emission goals (RCPs) diverging at the same time (IPCC, 2013). The ’business as usual’ RCP8.5 emission scenario is closest to the path of current global emissions. The RCP8.5 emission scenario ensemble shows the most significant changes in the ensemble mean of other climate model parameters by the end of the 21st century and yields the most significant change in the hydrological system.

6.2.2 Winter Precipitation and SWE. Winter precipitation shows a slight to moderate increase through the end of the 21st century for the ensemble mean of all emission scenarios in these CMIP5 simulations, while SWE_A shows nearly no change in the ensemble mean for the RCP2.6 emission scenario and a decrease for the RCP8.5 emission scenario. The combination of significantly warmer T_{WI} in combination with increased P_{WI} for RCP2.6 emission scenario yields a similar mean value of SWE_A, with a smaller fraction of the total winter precipitation falling as snow. For the RCP8.5 emission scenario, the ensemble mean reduction of SWE_A would indicate that an even smaller fraction of the winter precipitation is falling as snow despite the input of increased P_{WI}. This effect, in conjunction with the previously stated physical drivers for the change in the timing and duration of the snowpack, will likely complicate further the future application of statistically driven seasonal streamflow forecasts.

6.2.3 Spring Precipitation. Consistent with observations of P_{SP} in the observational record, spring precipitation shows high interannual variability for all model projections through the 21st century. Only the RCP8.5 ensemble mean shows a significant downward trend in P_{SP} through the end of the 21st century (Table 3a), with all other emission scenario ensembles showing no significant trend. With the shift towards an increased significance of the contribution of P_{SP} to the interannual variability of Q_{RO} within the context of warming trends in the observed record, trends in the projection of mean P_{SP} significantly impact trends of projected mean Q_{RO} through the 21st century.

6.2.4 Runoff. The ensemble mean projections of Q_{RO} through the end of the 21st century are consistent with the trends of the correlation of Q_{RO} with the climate parameters in the observed record. The dominant driver of changes in ensemble mean Q_{RO} is significant trends in mean P_{SP} for each emission scenario. RCP8.5 is the only emission scenario that shows a significant downward trend in mean Q_{RO}, which coincides with a negative trend in P_{SP} and SWE_A through end of the 21st century. While snowpack is the primary physical driver of Q_{RO}, we observe a stronger positive correlation between changes in mean P_{SP} and Q_{RO}, than SWE_A and Q_{RO}. Consistent with the results of this analysis applied to the observational data, there is no significant evidence that the increase in mean T_{SP} will directly contribute to loss of streamflow through physical processes such as evapotranspiration through the 21st century.

6.3 Observationally Consistent Models.

The statistically driven framework for selecting OCMs relies on metrics that are susceptible to the uncertainty associated with natural variability. However, the large spread of projections motivates a procedure to select models that best reproduce historical trends in the regional hydroclimate relationships. We propose a procedure here based on strong trends in $\mathrm{SWE_A}$-$\mathrm{Q_{RO}}$ covariance, in addition to trends in trends in the contribution of $\mathrm{P_{SP}}$ to interannual $\mathrm{Q_{RO}}$ variability. We interpret these trends as a forced climate signal that models should reproduce if we are to have high confidence in future streamflow projections.

There is no expectation that the retrospective simulations considered here will match the natural variability in the observed record, but the BCSD models have been shown to effectively simulate climate interactions on both annual and seasonal scales (Wood et al., 2004). Though there is uncertainty associated with the downscaling method used to convert from GCM scale projections to projections useful for regional analysis, there is no expectation that any downscaling method should impact the interannual variability of projections averaged seasonally and over a large area (Gutmann et al., 2014).

The downscaling technique used could significantly impact the expression of extreme events (ie. droughts, flooding) (Timmermans et al., 2018), but the spatial and temporal averaging applied to the observational data and projected model output used in this study should alleviate the problems associated with simulating extreme short-term weather.

The statistical method applied here proposes a tool for assessing confidence in model-projected Q_{RO} trends. We evaluate each simulations ability to simulate historical impacts of regional warming on climate-hydrology interaction. Identifying models that successfully reproduce trends in the contribution of climate parameters to streamflow in the retrospective simulation allows for the reduction of uncertainty associated with significant model spread in future years. The OCMs selected in this application reduce the spread of models for the entire ensemble and each emission scenario.

Of notable interest is the elimination of a majority of the model projections that produce increased mean streamflow through the end of the 21st century compared to the observational time period.

Reduction of confidence in the BCSD GCM projections that simulate increased mean streamflow through the end of the 21st century is of great value for long-term water management policy in the southwest.

7. Conclusion

7.1 Historical Observations Increase in runoff season streamflow forecast uncertainty due to the decrease in skill associated with snowpack during the last half-century poses a significant problem for the management of water resources. Furthermore, we have shown that a larger fraction of the variability of modern runoff season discharge in the Rio Grande is attributed to fluctuations in spring precipitation. It will be increasingly important for operational forecasters to develop methods for strengthening spring seasonal precipitation forecasts, both for the intrinsic importance of precipitation and as a component of surface water supply outlooks.

Previous endeavors to implement seasonal temperature directly into seasonal streamflow forecasting in the southwest rely on the stationarity of the correlation between temperature and precipitation, and do not address precipitation as the direct physical contributor to streamflow variability. In addition, long-term forecasting of streamflow decline in the southwest that relies on linear scaling of streamflow with temperature ignores the strength of the regional scale correlation between temperature and precipitation. Our results suggest that long-term trends in streamflow are closely tied to trends in precipitation. The contribution of precipitation to interannual and longer fluctuations in streamflow cannot be disregarded or parameterized in terms of temperature. Regional analysis of the contribution of precipitation or temperature to the variability of streamflow in the southwest must not ignore the strength of the relationship between the two parameters contained in the observational record.

7.2 GCM Projections The ensemble of BCSD GCM projections of streamflow and precipitation for the Rio Grande headwater region contains significant uncertainty through the end of the 21st century, primarily due to the spread of simulations. We have proposed here a method of selecting observationally consistent models that simulate the observed shift in hydro-climate parameter correlations derived from the observational record during the retrospective simulation period. By selecting OCMs, we are able to reduce uncertainty associated with model spread through the mid-21st century. Specifically, we find reduced confidence in simulations that produce increased mean runoff season discharge through mid-century.

The ensemble mean of the subset of OCMs differentiated by emission scenario produce significantly less runoff season discharge than the entire ensemble mean for the each emission scenario through the end of the 21st century. However, small sample sizes of the OCM subsets considered here limits the reduction in uncertainty associated with simulation spread that we can realize. Further development of selecting observationally consistent models to reduce late-21st century uncertainty in streamflow projections would require a larger set of streamflow simulations derived from GCMs.

8. TABLES AND

FIGURES

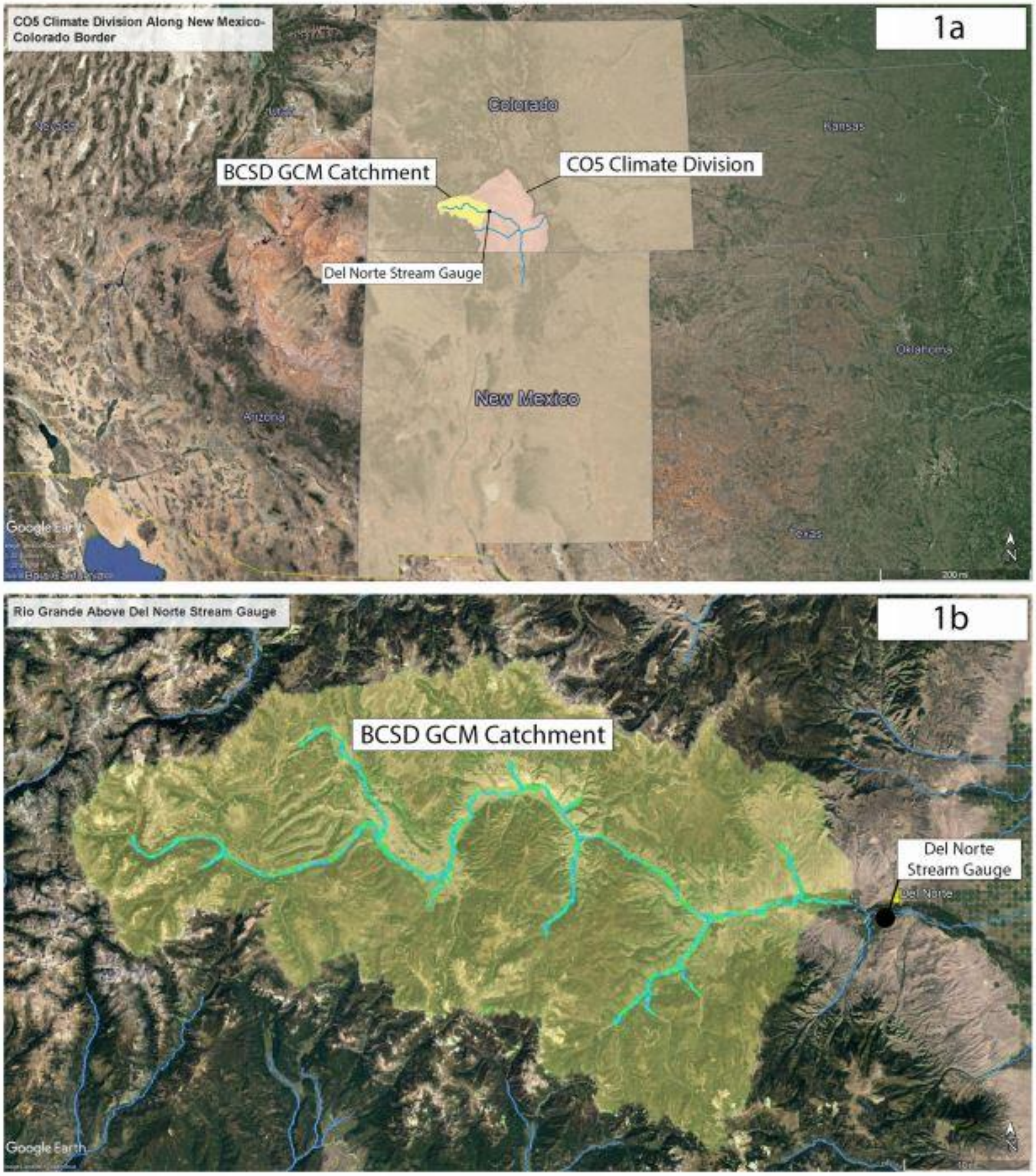

Figure 1a. Map of the CO5 climate division along the New Mexico-Colorado border (light pink) with the location of the Rio Grande main stem (blue). The yellow region overprinted within CO5 defines the area used to determine the precipitation and snow-pack that is routed as streamflow in the VIC model as part of the BCSD GCM projections published by the BOR. **Figure 1b.** A magnified view of the BCSD GCM catchment (yellow) above the Del Norte stream gauge that contributes to the streamflow that passes through the gauge. This is a subsection of the entire CO5 Rio Grande Headwater climate division that is used for the historical observation of precipitation and temperature derived from the PRISM dataset. The main stem of the Rio Grande is highlighted in blue.

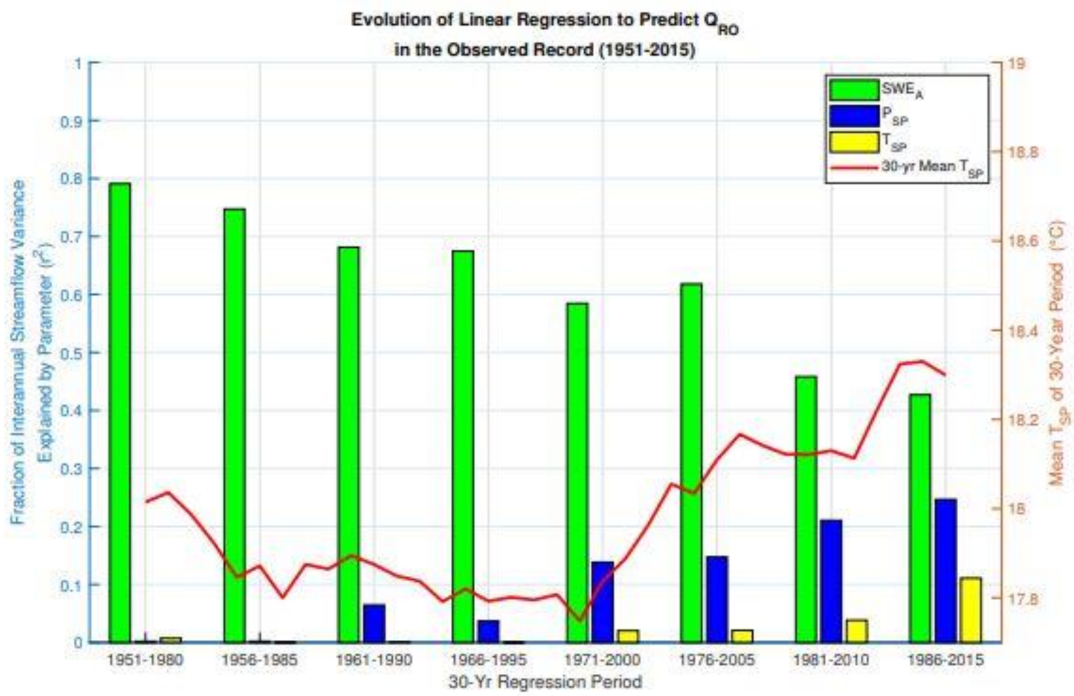

Figure 2 Fraction of interannual variability of Q_{RO} (r^2) (left-hand ordinate) associated with an individual parameter for successive 30-year regression windows. The green bar shows the fraction of interannual variability of Q_{RO} attributed to SWE_A as a first order predictive parameter in the step-wise regression model structure. The blue and yellow bars show the fraction of interannual variability of Q_{RO} attributed to P_{SP} and T_{SP} respectively as second order predictive parameters in the step-wise regressions (step 2 in Table 1). A moving 30-year window with annual resolution is used here to show 30-year mean T_{SP} (shown by the red line with the right-hand ordinate).

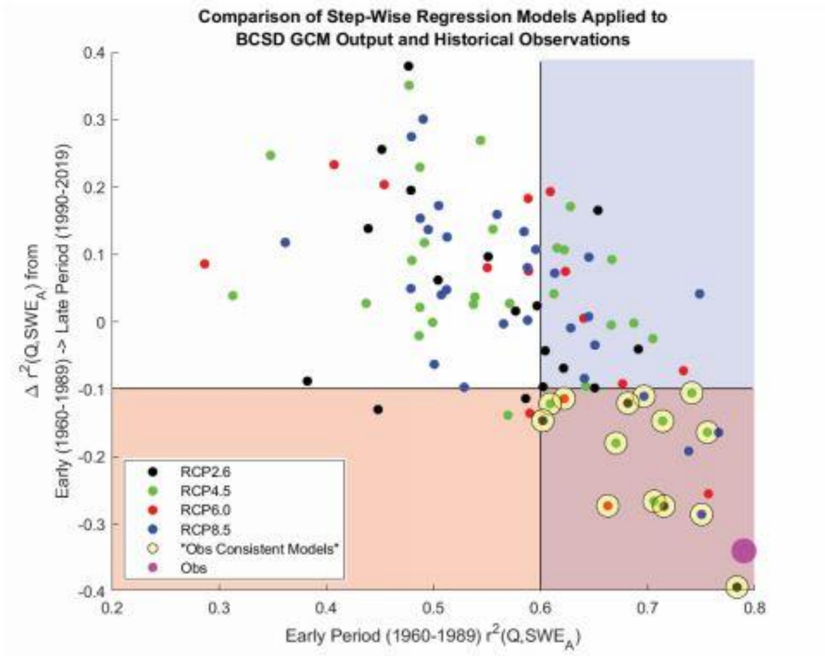

Figure 3a The horizontal axis shows the fraction of interannual variability of Q_{RO} attributed to SWE_A from the early time period (1960-1989) of the retrospective BCSD GCM output for each simulation. The blue shaded area indicates the region that satisfies OCM Criterion 1 ($r^2(Q_{RO}, SWE_A) > 0.6$). The vertical axis shows the change in the fraction interannual variability of Q_{RO} attributed to SWE_A as a first order step-wise predictive parameter from the early time period (1960-1989) to the late time period (1990-2019). The pink shaded area indicates the region that satisfies OCM Criterion 2 ($\Delta r^2(Q_{RO}, SWE_A) < -0.1$). Each colored data point represents a single simulation from the BCSD GCM ensemble and color differentiation represents different emission scenarios. The large purple data point shows the same criteria for the observational record. Yellow outlines for data points are model runs that we identify as the OCMs, selected by satisfaction of all four criteria listed in Sec 5.1. **Figure 3b** Like Figure 3a, but the horizontal axis is the late period (1990-2019) fraction of interannual variability of Q_{RO} attributed to P_{SP} as a second order step-wise predictive parameter. The pink shaded area indicates the region that satisfies OCM Criterion 3 ($r^2(Q^1_{RO}, P_{SP}) > 0.1$). The vertical axis shows the change in the total fraction of interannual variability of Q_{RO} attributed to SWE_A, P_{SP}, and T_{SP} using a bivariate step-wise approach from the early time period (1960-1989) to the late time period (1990-2019). The blue shaded area indicates the region that satisfies OCM Criterion 4 ($r^2(Q_{RO}, (SWE_A, (P_{SP}, T_{SP}))) < 0$).

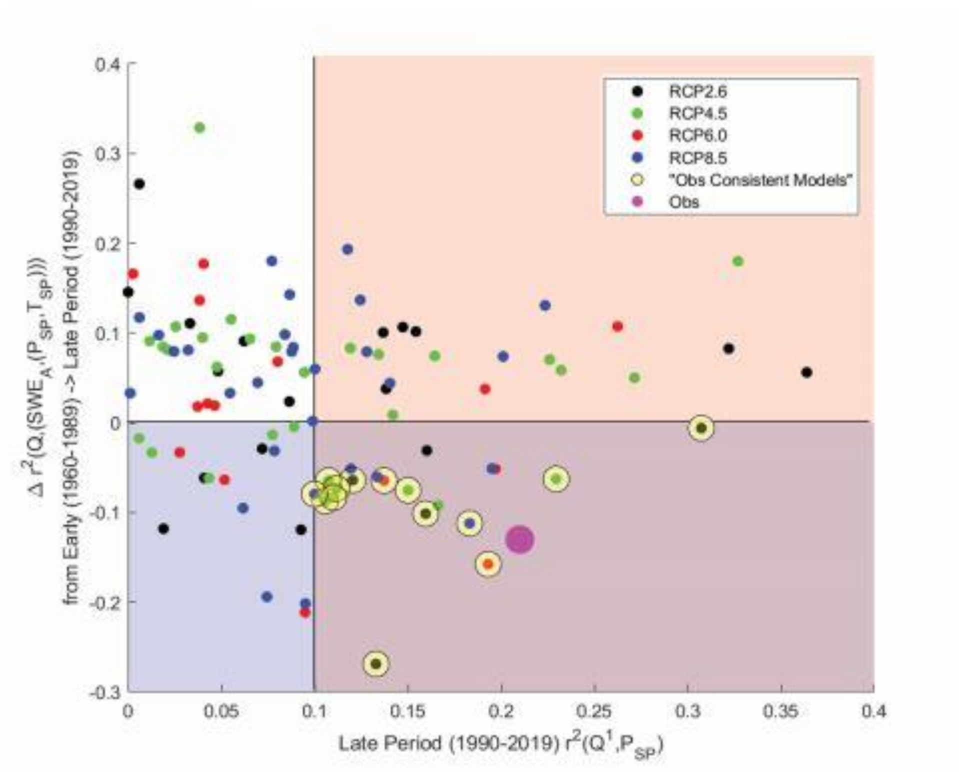

Figure 3b

628

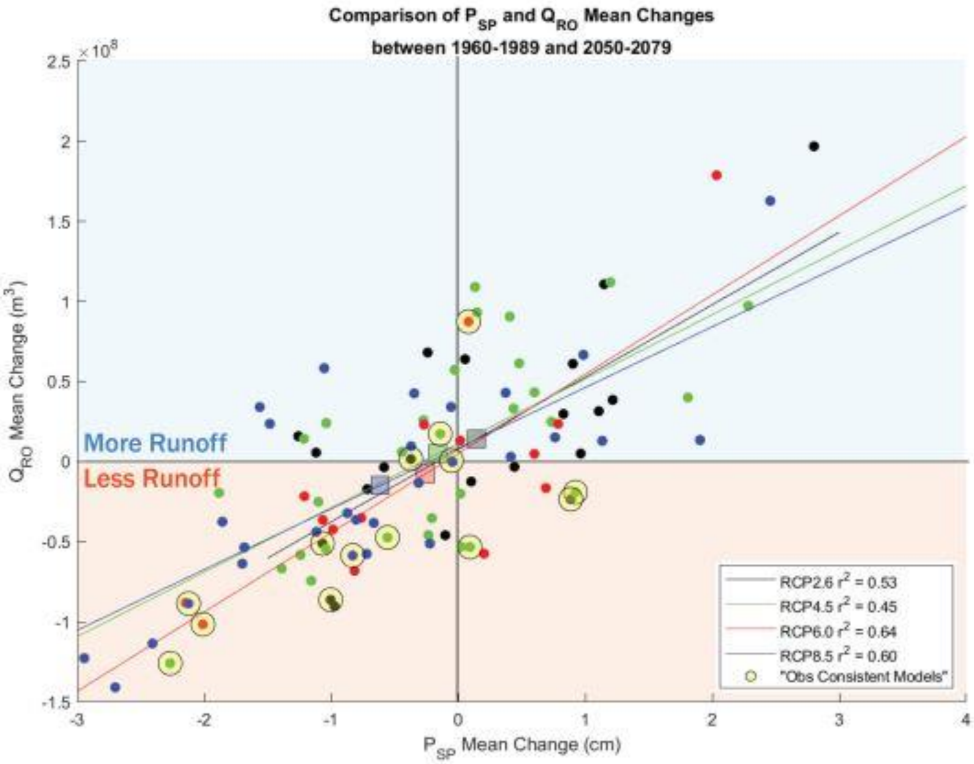

Figure 4a. The horizontal axis represents the $\Delta(Mean\ SWE_A)$ from the early retrospective forecast time period (1960-1989) to the late part of the 21st century (2050-2079) for each BCSD GCM simulation. The vertical axis represents the $\Delta(Mean\ Q_{RO})$ between the same two time periods. Each data point represents a single model simulation, with color differentiation between emissions scenarios. Each colored line is the product of a single linear regression based on all simulations forced by a particular emissions scenario. The centroid of the regression for each scenario is shown by a colored square. Yellow outlines over data points are simulations that are identified as the OCMs selected by the metrics described. The light blue shaded region indicates an increase in mean Q_{RO} between the late part of the 21st century (2050-2079) and the retrospective forecast period (1960-1989), while the light orange region indicates an decrease in mean Q_{RO} between the same periods. **Figure 4b.** Like Figure 4a, but the horizontal axis represents the $\Delta(Mean\ P_{SP})$ between the same time periods described above. **Figure 4c.** Like Figure 4a, but the horizontal axis here represents the $\Delta(Mean\ T_{SP})$ for the same time periods described above. Non-significant regressions (determined by a significance level of $\alpha = 0.05$) are not shown.

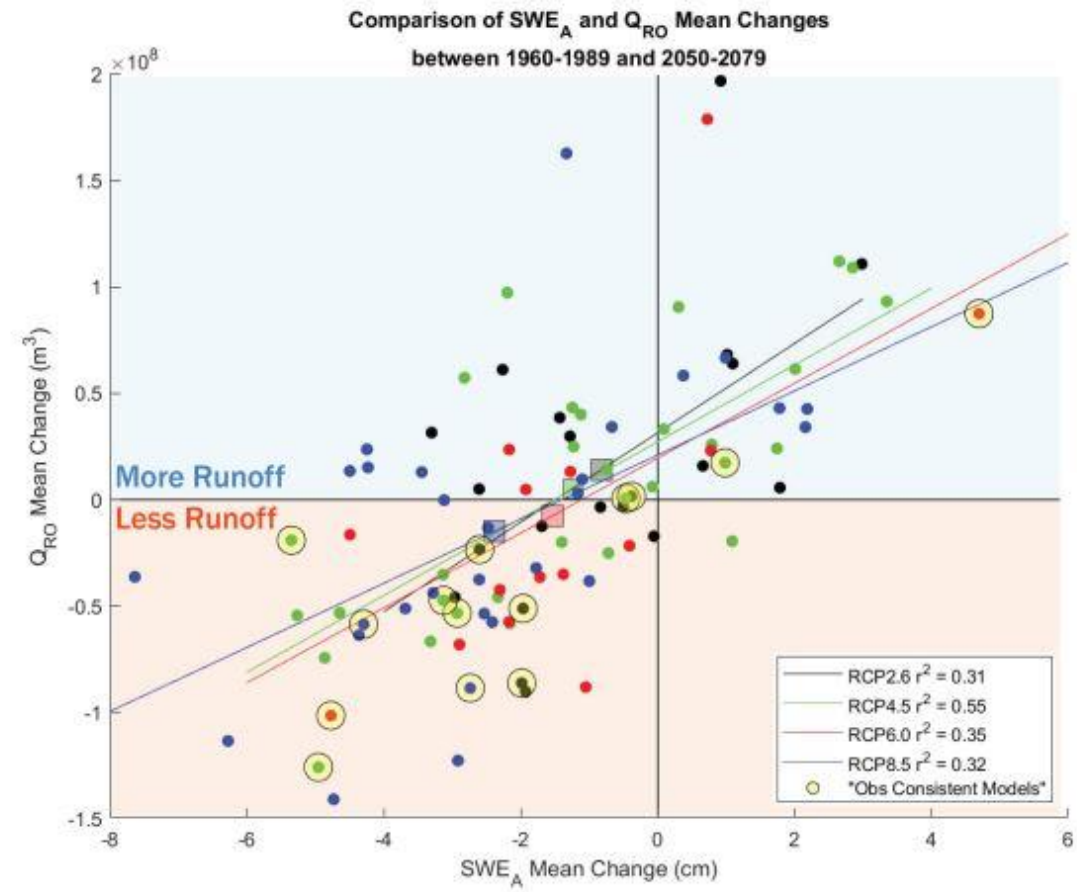

Figure 4b

630

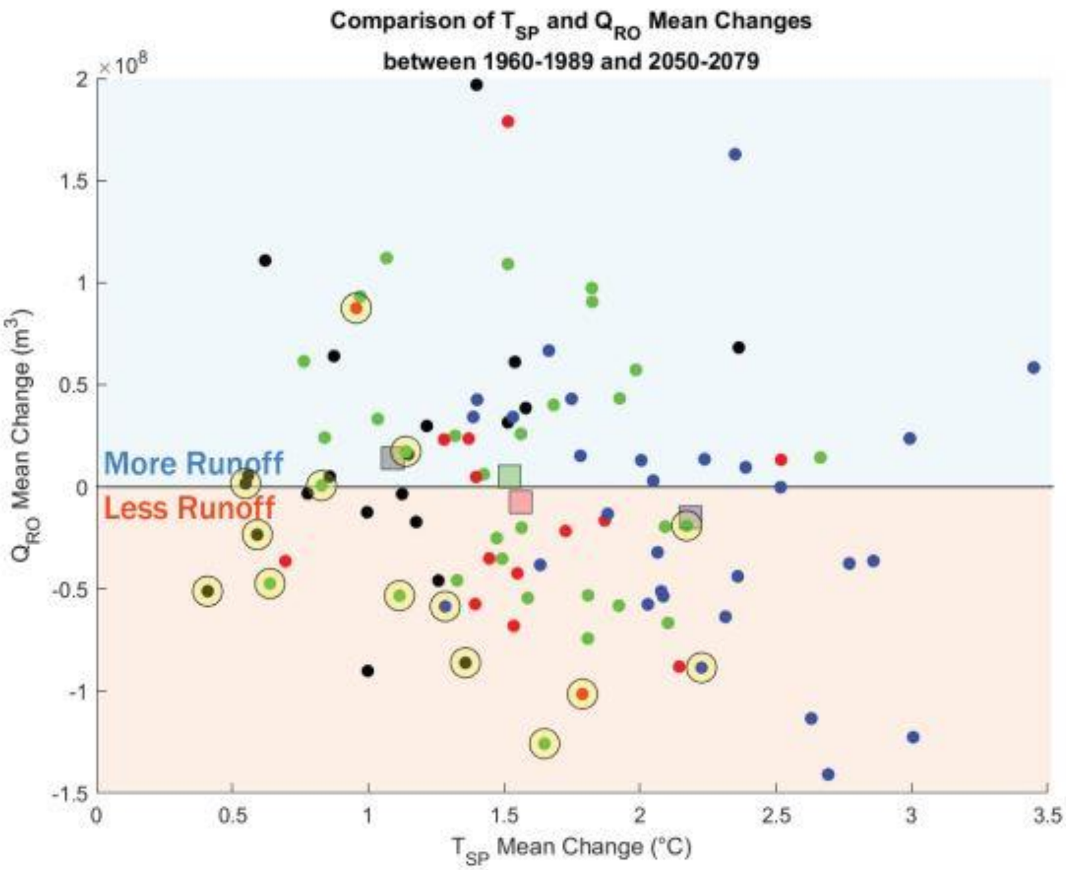

Figure 4c

631

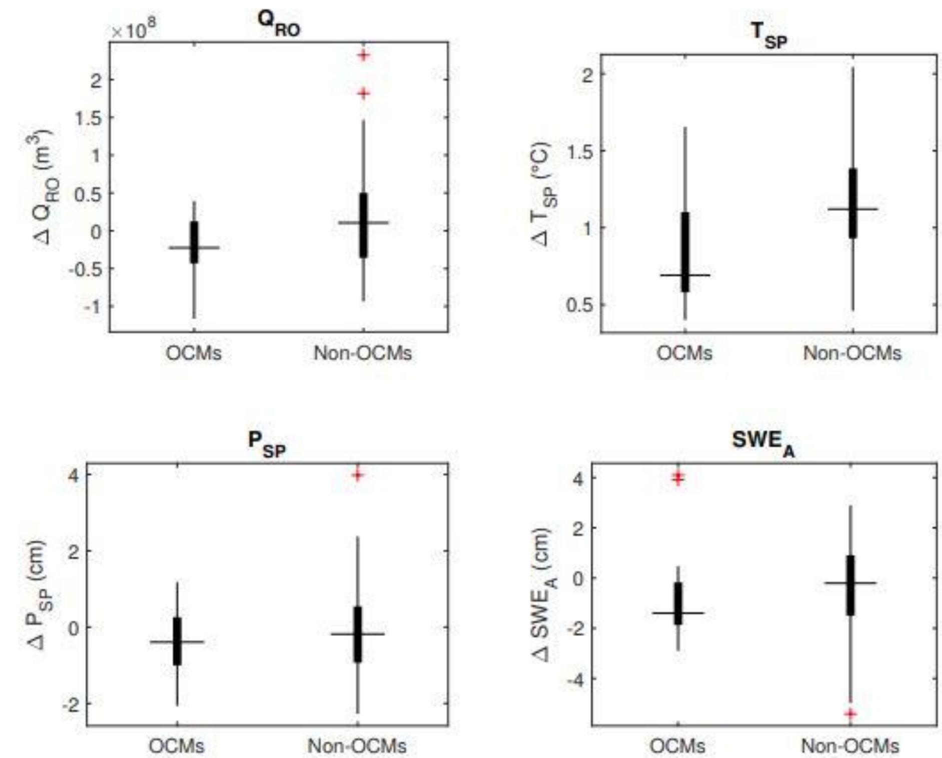

Figure 5. Distributions of the change in climate parameters through mid-century (2020-2049) compared to the retrospective simulation period (1960-1989) for both the OCM and non-OCM subsets of the entire BCSD GCM ensemble for Q_{RO},T_{SP},P_{SP}, and SWE_A. The central line of each box and whisker plot indicates the median of distribution, the boxes indicate the 25th-75th percentile, and the whiskers indicate the 10th-90th percentile of the distribution of models.Red pluses indicate outliers. This figure uses the same subsets of OCMs and non-OCMs as Table 3c.

Table 1: Pearson Correlation for Early Time Period (1951-1983) and Late Time Period (1983-2015) Parameters

.	P_{WI}	P_{SP}	T_{WI}	T_{SP}	Q_{RO}	SWE_A
P_{WI}	1	-0.06	-0.43	-0.24	0.48	0.80
P_{SP}	0.39	1	0.37	-0.68	0.47	0.02
T_{WI}	-0.47	-0.03	1	0.02	0.00	-0.49
T_{SP}	-0.37	-0.58	0.30	1	-0.49	-0.36
Q_{RO}	0.78	0.33	-0.67	-0.40	1	0.67
SWE_A	0.83	0.30	-0.67	-0.36	0.89	1

Table 1 contains the Pearson correlations between climate parameters for both time periods considered in this study. The lower-left half of the table (blue) contains the correlations for the early time period (1951-1983) and the upper-right half of the table (orange) contains the correlations for the late time period (1983-2015).

Table 2a. RMSE ($*10^8 m^3$) for Early Period (1951-1983)Step-wise Model

.	$SWE_A \rightarrow (P_{SP}, T_{SP})$	$SWE_A \rightarrow T_{SP} \rightarrow P_{SP}$	$SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$
Step 1	6.75	6.75	6.75
Step 2	6.64	6.65	6.70
Step 3	.	6.64	6.66

Table 2b. RMSE ($*10^8 m^3$) for Late Period (1983-2015)Step-wise Model

.	$SWE_A \rightarrow (P_{SP}, T_{SP})$	$SWE_A \rightarrow T_{SP} \rightarrow P_{SP}$	$SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$
Step 1	11.43	11.43	11.43
Step 2	8.87	10.79	8.99
Step 3	.	9.81	8.92

Table 2c. RMSE ($*10^8 m^3$) for Entire Period (1951-2015) Step-wise Model

.	$SWE_A \rightarrow (P_{SP}, T_{SP})$	$SWE_A \rightarrow T_{SP} \rightarrow P_{SP}$	$SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$
Step 1	14.60	14.60	14.60
Step 2	12.86	14.19	12.84
Step 3	.	13.35	12.82

Table 2d. r^2 for Early Period (1951-1983) Step-wise Model

.	$SWE_A \rightarrow (P_{SP}, T_{SP})$	$SWE_A \rightarrow T_{SP} \rightarrow P_{SP}$	$SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$
Step 1	0.79	0.79	0.79
Step 2	0.80	0.80	0.79
Step 3	.	0.80	0.80

Table 2e. r^2 for Late Period (1983-2015) Step-wise Model

.	$SWE_A \rightarrow (P_{SP}, T_{SP})$	$SWE_A \rightarrow T_{SP} \rightarrow P_{SP}$	$SWE_A \rightarrow P_{SP} \rightarrow T_{SP}$
Step 1	0.45	0.45	0.45
Step 2	0.67	0.51	0.66
Step 3	.	0.59	0.67

Tables 2a-2c show the root-mean-square error (RMSE) of each step-wise model associated with the addition of parameters for each step. The columns are differentiated by different step ordering and the rows differentiate each step. For example, the 2nd column contains the RMSE of the step-wise model with the addition of climate parameters ordered $SWE_A \rightarrow T_{SP} \rightarrow P_{SP}$. The first row is the RMSE of the linear regression using SWE_A as predictor of Q_{RO}. The second row shows the RMSE of the linear regression model with the addition of T_{SP} as a predictor of the residuals from the previous step. Likewise, the third row shows the RMSE of the linear regression model with the addition of P_{SP} as a predictor of the residuals from the second step. Tables 2d-2e follow the same structure as the previous tables, but show the fraction of interannual variability of Q_{RO} that is attributed to the addition of the climate parameters in each step.

Table 3a. Full Ensemble Projections of Changes in Climate Paramaters								
	RCP 2.6 (21)		RCP 4.5 (31)		RCP 6.0 (16)		RCP 8.5 (29)	
	Δ	σ	Δ	σ	Δ	σ	Δ	σ
SWE_A (cm)	-0.82	0.75	-1.23	0.66	-1.54	0.76	-2.38	0.70
P_{SP} (cm)	0.14	0.35	-0.16	0.34	-0.26	0.38	-0.62	0.33
T_{SP} (C°)	1.09	0.07	1.52	0.06	1.56	0.07	2.19	0.06
Q_{RO} (m^3) $*10^6$	14.09	18.06	4.94	16.62	-7.54	20.26	-14.87	16.89

Table 3a. Δ values are determine by subtracting the ensemble mean value from 2050-2079 projected output from 1960-1989 retrospective simulation output for each emission scenario. σ is the standard deviation of the distribution of the mean changes for individual simulations for the period 2050-2079 for each ensemble.

Table 3b. OCMs Projections of Changes in Climate Paramaters								
	RCP 2.6 (4)		RCP 4.5 (6)		RCP 6.0 (3)		RCP 8.5 (2)	
	Δ	σ	Δ	σ	Δ	σ	Δ	σ
SWE_A (cm)	-1.73	2.75	-2.71	2.93	-0.25	3.41	-3.51	5.09
P_{SP} (cm)	-0.39	1.48	-0.32	1.12	-0.82	1.45	-1.48	2.17
T_{SP} (C°)	0.73	0.43	1.29	0.60	1.18	0.63	1.76	1.01
Q_{RO} (m^3) $*10^6$	-39.80	60.40	-37.50	56.10	-6.82	64.00	-73.50	84.50

Table 3b. Δ values are determine by subtracting the OCM ensemble mean value from 2050-2079 projected output from 1960-1989 retrospective simulation output for each emission scenario. σ is the standard deviation of the distribution of the mean changes for individual simulations for the period 2050-2079 for each OCM ensemble. Numbers in parenthesis next to emission scenarios indicates how many simulations are considered as OCMs within the emission scenario ensemble.

Table 3c. Comparison of OCM and Non-OCM Subsets Through Mid Century				
	OCMs (15)		Non-OCMs (82)	
	Δ	σ	Δ	σ
SWE_A (cm)	-0.66	2.11	-0.36	1.77
P_{SP} (cm)	-0.55	0.94	-0.12	1.10
T_{SP} (C°)	0.83	0.37	1.15	0.32
Q_{RO} (m^3) $*10^7$	-1.82	4.16	1.42	6.20

Table 3c. Δ values are determine by subtracting ensemble mean value from 2020-2049 projected output from 1960-1989 retrospective simulation output for both OCM and non-OCM subsets. σ is the standard deviation of the distribution of the mean changes for individual simulations for the period 2020-2049 for each OCM ensemble. Numbers next to each subset title indicates the number of simulations that are considered for each subset. All RCPs are considered here due to lack of RCP divergence through mid-century.

Table 4. OCM Criteria for all BCSD GCM Simulations				
	Criterion 1	Criterion 2	Criterion 3	Criterion 4
	$r^2(Q_{RO},SWE_A) > 0.6$ for early period (1960-1989)	$\Delta r^2(Q_{RO},SWE_A) < -0.1$ from early (1960-1989) to late (1990-2019) period	$r^2(Q_{RO}',P_{SP}) > 0.1$ for late period (1990-2019)	$\Delta r^2(Q_{RO},(SWE_A,(P_{SP},T_{SP}))) < 0$ from early (1960-1989) to late (1990-2019) period
access1-0.1.rcp45	0.69	0.00	0.01	-0.02
access1-0.1.rcp85	0.56	0.16	0.00	0.12
bcc-csm1-1.1.rcp26	0.72	-0.27	0.31	-0.01
bcc-csm1-1.1.rcp45	0.67	-0.18	0.15	-0.08
bcc-csm1-1.1.rcp60	0.68	-0.09	0.19	0.04
bcc-csm1-1.1.rcp85	0.60	0.10	0.12	0.20
bcc-csm1-1-m.1.rcp45	0.54	0.27	0.02	0.08
bcc-csm1-1-m.1.rcp85	0.48	0.27	0.04	0.08
canesm2.1.rcp26	0.44	0.14	0.14	0.10
canesm2.1.rcp45	0.57	0.03	0.09	0.00
canesm2.1.rcp85	0.59	0.00	0.10	0.00
ccsm4.1.rcp26	0.48	0.38	0.01	0.27
ccsm4.1.rcp45	0.63	0.17	0.02	0.08
ccsm4.1.rcp60	0.55	0.08	0.05	0.02
ccsm4.1.rcp85	0.49	0.15	0.09	0.10
cesm1-bgc.1.rcp45	0.49	0.12	0.09	0.06
cesm1-bgc.1.rcp85	0.48	0.05	0.06	-0.10
cesm1-cam5.1.rcp26	0.65	-0.10	0.02	-0.12
cesm1-cam5.1.rcp45	0.48	0.35	0.04	0.33
cesm1-cam5.1.rcp60	0.45	0.20	0.04	0.18
cesm1-cam5.1.rcp85	0.65	-0.03	0.14	0.04
cmcc-cm.1.rcp45	0.74	-0.11	0.11	-0.06
cmcc-cm.1.rcp85	0.70	-0.11	0.10	-0.08
cnrm-cm5.1.rcp45	0.54	0.03	0.12	0.08
cnrm-cm5.1.rcp85	0.51	0.04	0.10	0.06
csiro-mk3-6-0.1.rcp26	0.65	0.17	0.00	0.15
csiro-mk3-6-0.1.rcp45	0.62	0.11	0.01	0.09
csiro-mk3-6-0.1.rcp60	0.59	0.18	0.00	0.17
csiro-mk3-6-0.1.rcp85	0.75	0.04	0.00	0.04
fgoals-g2.1.rcp26	0.59	-0.11	0.16	-0.03
fgoals-g2.1.rcp45	0.49	0.23	0.04	0.09
fgoals-g2.1.rcp85	0.61	0.08	0.07	0.05
fio-esm.1.rcp26	0.78	-0.39	0.13	-0.27
fio-esm.1.rcp45	0.71	-0.27	0.23	-0.06
fio-esm.1.rcp60	0.76	-0.26	0.09	-0.21
fio-esm.1.rcp85	0.75	-0.29	0.19	-0.11
gfdl-cm3.1.rcp26	0.58	0.02	0.07	-0.03
gfdl-cm3.1.rcp45	0.35	0.25	0.06	0.12
gfdl-cm3.1.rcp60	0.62	0.07	0.03	-0.03
gfdl-cm3.1.rcp85	0.64	0.01	0.08	-0.03
gfdl-esm2g.1.rcp26	0.45	0.26	0.06	0.09
gfdl-esm2g.1.rcp45	0.49	0.02	0.23	0.07
gfdl-esm2g.1.rcp60	0.61	0.19	0.04	0.14
gfdl-esm2g.1.rcp85	0.49	0.30	0.03	0.08
gfdl-esm2m.1.rcp26	0.38	-0.09	0.32	0.08
gfdl-esm2m.1.rcp45	0.31	0.04	0.33	0.18
gfdl-esm2m.1.rcp60	0.29	0.09	0.26	0.11
gfdl-esm2m.1.rcp85	0.36	0.12	0.22	0.13
giss-e2-h-cc.1.rcp45	0.67	0.00	0.01	-0.03
giss-e2-r.1.rcp26	0.62	-0.07	0.04	-0.06
giss-e2-r.1.rcp45	0.76	-0.16	0.11	-0.09
giss-e2-r.1.rcp60	0.66	-0.27	0.19	-0.16

giss-e2-r.1.rcp85	0.53	-0.10	0.20	0.07
giss-e2-r-cc.1.rcp45	0.50	0.00	0.13	0.08
hadgem2-ao.1.rcp26	0.55	0.10	0.05	0.06
hadgem2-ao.1.rcp45	0.62	0.11	0.05	0.06
hadgem2-ao.1.rcp60	0.64	0.01	0.04	0.02
hadgem2-ao.1.rcp85	0.51	0.17	0.01	0.10
hadgem2-cc.1.rcp45	0.49	-0.02	0.23	0.06
hadgem2-cc.1.rcp85	0.63	-0.01	0.12	-0.06
hadgem2-es.1.rcp26	0.68	-0.12	0.12	-0.06
hadgem2-es.1.rcp45	0.56	0.14	0.03	0.11
hadgem2-es.1.rcp60	0.59	0.08	0.04	0.02
hadgem2-es.1.rcp85	0.50	0.13	0.09	0.14
inmcm4.1.rcp45	0.61	0.04	0.08	-0.01
inmcm4.1.rcp85	0.77	-0.17	0.08	-0.19
ipsl-cm5a-mr.1.rcp26	0.60	-0.04	0.14	0.04
ipsl-cm5a-mr.1.rcp45	0.67	0.09	0.08	0.08
ipsl-cm5a-mr.1.rcp60	0.62	-0.11	0.14	-0.06
ipsl-cm5a-mr.1.rcp85	0.74	-0.19	0.09	-0.20
ipsl-cm5b-lr.1.rcp45	0.54	0.04	0.16	0.07
ipsl-cm5b-lr.1.rcp85	0.51	0.13	0.12	0.14
miroc5.1.rcp26	0.60	-0.10	0.09	-0.12
miroc5.1.rcp45	0.71	-0.15	0.11	-0.07
miroc5.1.rcp60	0.59	-0.14	0.20	-0.05
miroc5.1.rcp85	0.58	0.14	0.08	0.18
miroc-esm.1.rcp26	0.45	-0.13	0.36	0.06
miroc-esm.1.rcp45	0.44	0.03	0.27	0.05
miroc-esm.1.rcp60	0.41	0.23	0.08	0.07
miroc-esm.1.rcp85	0.50	-0.06	0.19	-0.05
miroc-esm-chem.1.rcp26	0.50	0.06	0.15	0.11
miroc-esm-chem.1.rcp45	0.71	-0.03	0.04	-0.06
miroc-esm-chem.1.rcp60	0.65	0.10	0.09	0.08
miroc-esm-chem.1.rcp60	0.65	0.09	0.09	0.08
mpi-esm-lr.1.rcp26	0.48	0.19	0.03	0.11
mpi-esm-lr.1.rcp45	0.48	0.09	0.07	0.09
mpi-esm-lr.1.rcp85	0.51	0.05	0.09	0.08
mpi-esm-mr.1.rcp26	0.60	0.02	0.15	0.10
mpi-esm-mr.1.rcp45	0.57	-0.14	0.17	-0.09
mpi-esm-mr.1.rcp85	0.59	0.08	0.13	0.08
mri-cgcm3.1.rcp26	0.60	-0.15	0.16	-0.10
mri-cgcm3.1.rcp45	0.61	-0.12	0.11	-0.08
mri-cgcm3.1.rcp85	0.64	-0.08	0.13	-0.06
noresm1-m.1.rcp26	0.69	-0.04	0.09	0.02
noresm1-m.1.rcp45	0.64	-0.10	0.14	0.01
noresm1-m.1.rcp60	0.73	-0.07	0.05	-0.06
noresm1-m.1.rcp85	0.57	-0.01	0.06	0.03

9. REFERENCE LIST

Barnett, T., Pierce, D. W., Hidalgo, H., Bonfils, C., Santer, B., Das, T., Govindasamy, B., Wood, A., Nozawa, T., Mirin, A. A., Cayan, D. , & Dettinger, M. (2008). Human-Induced Changes in the Hydrology of the Western United States. *Science*, 319. 1080-3. 10.1126/science.1152538.

Blythe, T. L., & Schmidt, J. C. (2018). Estimating the natural flow regime of rivers with longstanding development: The Northern branch of the Rio Grande. *Water Resources Research*, 54, 1212 1236. https://doi.org/10.1002/2017WR021919.

Cayan, D. R., Dettinger M.D., Kammerdiener S.A., Caprio J.M., & Peterson D.H. (2001). Changes in the Onset of Spring in the Western United States. *Bulletin of the American Meteorological Society* 82:399-415.

Chavarria, S.B. & Gutzler, D.S. (2018). Observed Changes in Climate and Streamflow in the Upper Rio Grande Basin. *Journal of the American Water Resources Association* 54(3): 644 659. https://doi.org/10.1111/1752-1688.12640

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., & Pasteris, P.A. (2008). Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*. https://doi.org/10.1002/joc.1688.

Deser, C., Phillips A.S., Tomas R.A., Okumura Y.M., Alexander M.A., Capotondi A., Scott J.D., Kwon Y., & Ohba M. (2012). ENSO and Pacific Decadal Variability in the Community Climate

System Model Version 4. *Journal of Climate*, 25, 26222651. https://doi.org/10.1175/JCLI-D-11-00301.1

Deser, C., Simpson I.R., Phillips A.S., & McKinnon K.A. (2018). How Well Do We Know ENSOs Climate Impacts over North America, and How Do We Evaluate Models Accordingly?. *Journal of Climate*, 31, 49915014. https://doi.org/10.1175/JCLI-D-17-0783.1

Garen, D.C. (1992). Improved techniques in regression-based streamflow volume forecasting. *Journal of Water Resources Planning and Management*, 118(6):654-671.

Gershunov, A. & Barnett T.P. (1998) Interdecadal Modulation of ENSO Teleconnections. Bull. Amer. Meteor. Soc., 79, 27152726. https://doi.org/10.1175/1520-0477(1998)079

Gutmann, E., Pruitt, T., Clark, M.P., Brekke, L., Arnold, J.R., Raff, D.A., & Rasmussen, R.M. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research*, 50, 7167 7186. doi:10.1002/2014WR015559.

Hurd, B.H., & Coonrod, J. (2012). Hydro-Economic Consequences of Climate Change in the Upper Rio Grande. *Climate Research*, 53: 10318. https://doi.org/10.3354/cr01092.

IPCC, (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D.

Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp. doi:10.1017/CBO9781107415324.

Knowles, N., Dettinger M.D., & Cayan D.R. (2006). Trends in snowfall versus rainfall in the western United States. *Journal of Climate*, 19:4545-4559.

Lehner, F., Wahl E.R., Wood A.W., Blatchford D.B., & Llewellyn D. (2017a). Assessing recent declines in Upper Rio Grande runoff efficiency from a paleoclimate perspective. *Geophysical Research Letters*, 44. doi:10.1002/2017GL073253.

Lehner, F., Wood A.W., Llewellyn D., Blatchford D.B., Goodbody A.G., & Pappenberger F. (2017b).

Mitigating the impacts of climate nonstationarity on seasonal streamflow predictability in the U.S. Southwest. *Geophysical Research Letters*, 44. doi:10.1002/2017GL076043.

Liu, Y. (2003). Spatial patterns of soil moisture connected to monthly-seasonal precipitation variability in a monsoon region. J. *Geophysical Research Letters*, 108: 1-14

Livneh, B., Deems J.S., Buma B., Barsugli J.J., Schneider D., Molotch N.P., Wolter K., & Wessman C.A. (2015). Catchment Response to Bark Beetle Outbreak and DustonSnow in the Colorado Rocky Mountains. *Journal of Hydrology*, 523: 196210. https://doi.org/10.1016/j.jhydrol.2015.01.039.

Milly, P.C.D., Betancourt J., Falkenmark M., Hirsch R.M., Kundzewicz Z.W., Lettenmaier D.P., & Stouffer R.J. (2008). Stationarity is Dead: Whither Water Management. *Science*, 319:573-574.

Mix, K., Lopes V.L., & Rast W. (2012). Environmental drivers of streamflow change in the Upper Rio 478 Grande. *Water Resources Management*, 26:253272. https://doi.org/10.1007/s11269-011-9916-9.

Mote, P.W. (2006). Climate-Driven Variability and Trends in Mountain Snowpack in Western North America. *Journal of Climate*, 19, 62096220. https://doi.org/10.1175/JCLI3971.1

Pagano, T., Garen D., & Sorooshian S. (2004). Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 19222002. *Journal of Hydrometeorology*, 5, 896909. https://doi.org/10.1175/15257541

Painter, T.H., Barrett, A.P., Landry, C.C., Neff, J.C., Cassidy, M.P., Lawrence, C.R., McBride, K.E., & Farmer, G.L. (2007). Impact of disturbed desert soils on duration of mountain snow cover. *Geophysical Research Letters*, 34, L12502. doi:10.1029/2007GL030284.

Rango, A. (2006). Snow: The real water supply for the Rio Grande basin. *New Mexico Journal of Science*, 44:99-118.

Ropelewski, C.F. & Halpert M.S. (1986). North American Precipitation and Temperature Patterns Associated with the El Nio/Southern Oscillation (ENSO). *Monthly Weather Review*, 114, 23522362. https://doi.org/10.1175/1520-0493

Reclamation (2013). Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and

Summary of User Needs, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado. 47pp.

Reclamation (2014). Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Hydrology Projections, Comparison with preceding Information, and Summary of User Needs, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado. 110 pp.

Stewart, I.T. (2009). Changes in snowpack and snowmelt runoff for key mountain regions. *Hydrologic Processes*, 23:7894. doi: 10.1002/hyp.7128.

Udall, B. & Overpeck, J. (2017). The twentyfirst century Colorado River hot drought and implications for the future, *Water Resources Research*, 53, 2404 2418. doi:10.1002/2016WR019638.

Vano, J.A., Udall B., Cayan D.R., Overpeck J.T., Brekke L.D., Das T., Hartmann H.C., Hidalgo H.G., Hoerling M., McCabe G.J., & Morino K. (2014). Understanding uncertainties in future Colorado River streamflow. *Bulletin of the American Meteorological Society*, 95:59-78. doi:10.1175/BAMS-D12-00228.1.

Wood, A., Leung, L., Sridhar, V., & Lettenmaier, D.P. (2004). Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Climatic Change*, 62. 189-216.10.1023/B:CLIM.0000013685.99609.9e.