RECLAMATION
Managing Water in the West

Runoff Efficiency and Seasonal Streamflow Predictability in the U.S. Southwest

Research and Development Office
Science and Technology Program
(Final Report) ST-2015-8730-01

U.S. Department of the Interior
Bureau of Reclamation
Research and Development Office

October 2018

Mission Statements

The U.S. Department of the Interior protects America's natural resources and heritage, honors our cultures and tribal communities, and supplies the energy to power our future.

The Department of the Interior protects and manages the Nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors its trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated island communities.

<table>
<tr><td colspan="2">REPORT DOCUMENTATION PAGE</td><td>Form Approved
OMB No. 0704-0188</td></tr>
<tr><td>T1. REPORT DATE:
OCTOBER 2018</td><td>T2. REPORT TYPE:
RESEARCH</td><td>T3. DATES COVERED</td></tr>
<tr><td colspan="2" rowspan="3">T4. TITLE AND SUBTITLE
Runoff Efficiency and Seasonal Streamflow Predictability in the U.S. Southwest</td><td>5a. CONTRACT NUMBER</td></tr>
<tr><td>5b. GRANT NUMBER
R15AC00049</td></tr>
<tr><td>5c. PROGRAM ELEMENT NUMBER
1541 (S&T)</td></tr>
<tr><td colspan="2" rowspan="3">6. AUTHOR(S)
Dagmar Llewellyn, 505-462-3594, dllewellyn@usbr.gov
Andy Wood, 303-497-8257, andywood@ucar.edu
Flavio Lehner, 303-497-8395, flehner@ucar.edu</td><td>5d. PROJECT NUMBER
ST-2015-8730-01</td></tr>
<tr><td>5e. TASK NUMBER</td></tr>
<tr><td>5f. WORK UNIT NUMBER
ALB-402</td></tr>
<tr><td colspan="2">7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Albuquerque Area Office
555 Broadway Blvd. NE, Suite 100
Albuquerque, NM 87102

Research Applications Laboratory
National Center for Atmospheric Research
P.O. Box 3000
Boulder, CO 80307-3000</td><td>8. PERFORMING ORGANIZATION REPORT NUMBER</td></tr>
<tr><td colspan="2" rowspan="2">9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Research and Development Office
U.S. Department of the Interior, Bureau of Reclamation,
PO Box 25007, Denver CO 80225-0007</td><td>10. SPONSOR/MONITOR'S ACRONYM(S)
R&D: Research and Development Office
BOR/USBR: Bureau of Reclamation
DOI: Department of the Interior</td></tr>
<tr><td>11. SPONSOR/MONITOR'S REPORT NUMBER(S)
ST-2015-8730-01</td></tr>
<tr><td colspan="3">12. DISTRIBUTION / AVAILABILITY STATEMENT
Final report can be downloaded from Reclamation's website: https://www.usbr.gov/research/</td></tr>
<tr><td colspan="3">13. SUPPLEMENTARY NOTES</td></tr>
<tr><td colspan="3">14. ABSTRACT (Maximum 200 words)
This project involved two primary components: (1) evaluation of runoff efficacy changes using observed and paleo reconstructed flows and (2) exploration of the benefits of including available seasonal temperature forecasts into seasonal water supply forecast models to improve their skill.</td></tr>
<tr><td colspan="3">15. SUBJECT TERMS
Rio Grande, Runoff, Water Supply Forecasting</td></tr>
</table>

<table>
<tr><td colspan="3">16. SECURITY CLASSIFICATION OF:</td><td rowspan="2">17. LIMITATION OF ABSTRACT
U</td><td rowspan="2">18. NUMBER OF PAGES</td><td>19a. NAME OF RESPONSIBLE PERSON
Dagmar Llewellyn</td></tr>
<tr><td>a. REPORT
U</td><td>b. ABSTRACT
U</td><td>c. THIS PAGE
U</td><td>19b. TELEPHONE NUMBER
505-462-3594</td></tr>
</table>

BUREAU OF RECLAMATION

Research and Development Office
Science and Technology Program

(Final Report) ST-2015-8730-01

Runoff Efficiency and Seasonal Streamflow Predictability in the U.S. Southwest

Prepared by: Flavio Lehner
Project Scientist, NCAR

Checked by: Dagmar Llewellyn
Hydrologist, Albuquerque Area Office, UC Region

Peer Review: Kenneth Nowak
Water Availability Research Coordinator, Research and Development Office

Executive Summary

This project involved two primary components: (1) evaluation of runoff efficacy changes using observed and paleo reconstructed flows and (2) exploration of the benefits of including available seasonal temperature forecasts into seasonal streamflow forecast models to improve their skill. In the first investigation, it was found that recent declines in runoff efficiency were in part attributable to anomalously warm temperatures. In the second part of the project, the aforementioned finding was tested for potential improve seasonal streamflow forecasts. Specifically, seasonal temperature forecasts were incorporated into the statistical water supply methodology employed by the Natural Resources Conservation Service (NRCS). The outcome was moderate improvement in seasonal water supply forecast skill. These two components of the project each resulted in a publication to the journal Geophysical Research Letters (GRL).

Contents

Main Report

Appendix A to this report provides a project summary, including the project purpose and research formulation, research phases, peer-reviewed publications and news reports, and next steps.

The outcomes of this project are documented in two peer-reviewed publications, both to the journal Geophysical Research Letters (GRL), which are provided as Appendices B and C to this report. The first (Appendix B) is “Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective” (Lehner, Wahl, Wood, Blatchford, & Llewellyn, 2017). The second (Appendix C) is “Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest” (Lehner, et al., 2017).

References

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

Lehner, F., Wood, A., Llewellyn, D., Blatchford, D., Goodbody, A., & Pappenberger, F. (2017). Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. *Geophysical Research Letters*. doi:10.1002/2017GL076043

Appendix A

Project Summary

Runoff Efficiency and Seasonal Streamflow Predictability in the U. S. Southwest

Appendix A: Project Summary

Bureau of Reclamation Fiscal Year 2015-2017 Science & Technology Program Project #8730

Postdocs Applying Climate Expertise (PACE) Fellowship for Improving Seasonal Forecasting to Support Operational Decision-Making within Reclamation Service Areas

March 30, 2018

PACE Fellow: Flavio Lehner – Postdoctoral Fellow, Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO; expertise in climate variability

PACE Mentors: The mentor team consisted of:

- Andy Wood – Project Scientist III and Scientific Supervisor, Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO; expertise in operational hydroclimate forecasting.
- Douglas Blatchford – Bureau of Reclamation, Lower Colorado Region, Boulder City, NV; expertise in Reclamation river and reservoir operations.
- Dagmar Llewellyn – Bureau of Reclamation, Upper Colorado Region, Albuquerque Area Office, Albuquerque, NM; expertise in Reclamation river and reservoir operations, and use of forecast information to support water management.

Duration: 1st April 2016 to 31st March 2018

Project Objective: To investigate potential methods to improve seasonal streamflow forecasting for Reclamation's service areas in the Upper Colorado and Lower Colorado Regions, including the Colorado River and Upper Rio Grande basins.

Content of this Report

Research Formulation:

The original project proposal to Reclamation's Science & Technology Program (FY-2015 #8730), titled "*A web-based data assimilation framework for improving operational decision-making"*, detailed a range of potential strategies for improving operational decision-making within Reclamation. The selected fellowship candidate, Flavio Lehner, has expertise in climatology and forecasting. Therefore, the project team elected to focus the work under this fellowship on improvements to forecasting of climate and hydrology, with a secondary goal of understanding trends in hydroclimate that might be relevant for Reclamation's management and planning challenges. The scope of the PACE project was designed to complement ongoing collaborative research toward advancing streamflow prediction practices between NCAR's Research Applications Lab and Reclamation (supported under the current and prior Cooperative Agreements between the two institutions, from 2013 to the present).

The Upper Rio Grande basin was selected as the initial focus area for project efforts to improve seasonal streamflow forecasts. This basin has less sophisticated forecasting tools available to it than many other basins in Reclamation's service area, including the Colorado Basin. Also, Reclamation water operations practitioners in this basin described poor forecast skill for this basin, with a tendency for over-forecasting biases during recent decades.

Reclamation's Research Office was notified of the selected fellow, and the intended direction for the fellowship research. Dagmar Llewellyn of the Upper Colorado Region, Albuquerque Area Office was added to the project as a co-Principal Investigator for Reclamation, along with Douglas Blatchford of the Lower Colorado Region.

Research, Phase I: Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective.

Dr. Lehner undertook a detailed review of the recent seasonal streamflow forecast biases in the Upper Rio Grande basin, applying prior expertise with paleo-based and climate-model-based analyses. He hypothesized that the over-forecasting bias resulted from declining runoff efficiency (water year streamflow divided by water year precipitation), in part driven by warming trends.

To evaluate this hypothesis, Dr. Lehner turned to paleoclimate reconstructions, which allowed him to extend his analyses beyond the available instrumental record. The long paleoclimate reconstructions allowed Dr. Lehner to tease out the factors contributing to variations in runoff efficiency in a more robust fashion than would be possible from the short instrumental record. It also enabled him to put the recent downward trend in runoff efficiency in context of the long paleoclimate record and to evaluate the extent to which this recent trend has precedent. For this phase of the research, Dr. Lehner teamed up with Eugene Wahl from the National Oceanic and Atmospheric Administration (NOAA) Paleoclimatology Group, also in Boulder, CO.

The research found that runoff efficiency varies primarily in proportion to precipitation, but that there exists a clear secondary influence of temperature. In years of low precipitation, very low runoff efficiencies are made 2.5–3 times more likely by high temperatures. This temperature sensitivity appears to have strengthened in recent decades, implying future water management vulnerability should recent warming trends in the region continue.

The resulting paper (Lehner et al. 2017a) features a number of scientific novelties, such as the first successful reconstruction of runoff efficiency from prior to the instrumental record, and the first documentation of the influence of temperature on streamflow over such a long period for this region of the world. The paper consequently garnered the attention of a couple of newspapers and blogs, as well as decision makers and researchers in and outside Reclamation (see list below).

<u>Peer-Reviewed Publication (Appendix B)</u>

Lehner, F., E. R. Wahl, A. W. Wood, D. Blatchford, and D. Llewellyn, 2017a: *Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective*. Geophys. Res. Lett., doi:10.1002/2017GL073253.

<u>NCAR/UCAR AtmosNews Article:</u>

https://www2.ucar.edu/atmosnews/news/126957/warmer-temperatures-cause-decline-in-key-runoff-measure

<u>Conference Posters and Presentations</u>

- AGU Fall Meeting 2016 in San Francisco (December 12 2016): "Declining runoff efficiency in the Southwestern US and implications for forecasting and water management" (Talk)
- Southern Nevada Water Authority conference on Colorado River Water Management (May 23 2017): "The influence of temperature on runoff efficiency: Implications for streamflow forecasting" (Talk)
- Community Earth System Model (CESM) Paleoclimate Working Group Meeting (March 2 2017): "Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective" (Talk)

<u>Field Tours</u>

- Yuma Area Office operations, September 13-15 2016

<u>Research Collaborations</u>

- Eugene Wahl, National Oceanic and Atmospheric Administration (NOAA) Paleoclimatology Group, Boulder, CO.

<u>Outreach to Reclamation, the Water Management Community, and the Public</u>

- Presentation to Reclamation Yuma Area Office, Yuma, AZ (September 15 2016)
- Presentation to Lower Colorado Region Water Operations Group, Boulder City, NV (February 8, 2017)
- Presentation to Southern Nevada Water Authority, Las Vegas, NV (February 9, 2017)
- Presentation to National Weather Service Interagency symposium of hydrology and climate, Albuquerque, NM (via webinar; April 18 2017)
- Newspapers and blogs:
    - "*How does global warming affect flows in the Rio Grande?*" (Summit County Citizen's Voice and Coyote Gulch blog, May 11 2017; https://coyotegulch.blog/2017/05/11/how-does-global-warming-affect-flows-in-the-rio-grande/)
    - "*Warmer temperatures drying the Rio Grande*" (Climate Central, May 12 2017; http://www.climatecentral.org/news/warmer-temperatures-drying-rio-grande-21446)

- "*Boulder scientist leads study probing warming impact on runoff*" (Daily Camera – Boulder News, May 17 2017; http://www.dailycamera.com/news/boulder/ci_30994168/boulder-scientist-leads-study-probing-warming-impact-runoff)
- "Calentamiento afecta río Bravo" (El Manana Nuevo Laredo, May 21 2017; http://elmanana.com.mx/noticia/136151/Calentamiento-afecta-rio-Bravo.html)
- "*Warmer temperatures cause decline in key runoff measure*" (Geophysical Research Letters, Editor's Highlight, May 22 2017; https://agupubs.onlinelibrary.wiley.com/hub/article/10.1002/2017GL073253/editor-highlight/)

Research Phase II: Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest.

The downward trend in runoff efficiency over the last 30 years was shown, in the first phase of this research, to be significant. This trend clearly poses a challenge for the current statistical forecasting models that assume a stationary background climate. While prior research has shown limitations in using seasonal precipitation forecasts for streamflow forecasting, the first paper published as part of this fellowship supported the idea that temperature forecasts could be beneficial in correcting for warming trends in regression-based forecasts. Therefore, the team focused the second phase of this research fellowship on an exploration of the benefits of including available seasonal temperature forecasts into seasonal streamflow forecast models to improve their skill. To that end, they used the publicly available seasonal climate forecasts from the North American Multi-Model Ensemble (NMME; 7 models) and later was able to add forecasts from the proprietary European Centre for Medium Range Weather Forecasts (ECMWF; 1 model), leveraging an ongoing collaboration between the ECMWF Head of Forecasting (Florian Pappenberger) and Dr. Wood. Dr. Wood introduced Dr. Lehner to the methods, data and practice of statistical water supply forecasting and to colleagues in the operational forecast center (Natural Resources Conservation Service, NRCS) serving Reclamation in the Upper Rio Grande basin.

Dr. Lehner's experience with climate model dynamics and datasets proved instrumental in analyzing and staging the large ensemble prediction dataset for application to streamflow forecasting. In the ensuing work, the team corroborated that they could indeed improve the streamflow forecast skill for key headwater gages in the Colorado and Rio Grande by around 10% in hindcasts over the period 1987-2016 by adding temperature information to the current operational forecasting approach by the Natural Resources Conservation Service (NRCS). Angus Goodbody, a hydrologic forecaster at NRCS, joined the co-author team, as well as Dr. Pappenberger. The resulting paper on the forecasting work was published in December 2017 (Lehner et al. 2017b) and presented on several occasions, such as the American Geophysical Union (AGU) Fall Meeting in New Orleans.

<u>Peer-Reviewed Publication (Appendix C)</u>

Lehner, F., A. W. Wood, D. Llewellyn, D. B. Blatchford, A. G. Goodbody, and F. Pappenberger, 2017b: Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. *Geophys. Res. Lett.*, doi:10.1002/2017GL076043.

<u>NCAR/UCAR AtmosNews Article:</u>
https://www2.ucar.edu/atmosnews/just-published/131553/taking-temperature-streamflow-forecasts

Conference Posters and Presentations

- AGU Fall Meeting 2017, New Orleans, LA (December 11 2017): “Using temperature forecasts to improve seasonal streamflow forecasts in the Colorado and Rio Grande Basins” (Talk)
- AGU Fall Meeting 2017, New Orleans, LA (December 12 2017): “Projected drought risk in 1.5°C and 2°C warmer climates” (Invited Talk)
- New Mexico Water Resources Research Institute Annual New Mexico Water Conference, Socorro, NM (August 16 2017): “Using Temperature Forecasts to Improve Seasonal Streamflow Forecasts in the Colorado and Rio Grande Basins” (Poster)
- National Center for Environmental Prediction Subseasonal-to-Seasonal Science Meeting, College Park, MD (September 15 2017): “Using Temperature Forecasts to Improve Seasonal Streamflow Forecasts in the Colorado and Rio Grande Basins” (Poster)
- National Center for Atmospheric Research Water System Retreat, Boulder, CO (January 17 2018): “Mitigating the impacts of climate non-stationarity on seasonal streamflow predictability in the US Southwest” (Talk)
- “Surviving Peak Drought and Warming” Conference, University of Arizona, Tucson, AZ (March 29 2018): “Drought in 1.5°C and 2°C warmer climates: uncertainties and implications” (Invited Talk)
- Western Snow Conference, Albuquerque, NM (April 17 2018): “On the use of snow and climate information in statistical streamflow forecasting” (Talk)

Research Collaborations

- Florian Pappenberger, Head of Forecasting, European Center for Medium-Range Weather Forecasts (ECMWF), Reading, United Kingdom
- Angus Goodbody, hydrologic forecaster, Natural Resources Conservation Service, Portland, OR

Field Tours

- Lower Colorado Region Operations, Boulder City, NV, February 7-9, 2017
- Upper Rio Grande Tour – Alamosa Office Operations, Reclamation’s Closed Basin Project, and San Luis Valley irrigation operations, May 1-3, 2017
- “Connecting People to Rivers”, Rio Chama Wild and Scenic Reach, hosted by Rio Grande Restoration, July 14-16, 2017

Outreach to Reclamation, the Water Management Community, and the Public

- Participation in discussions at USGS New Mexico Water Science Center, Upper Rio Grande Focus Area Study annual research symposium (via webinar; December 5 2017)
- Meeting with Colorado Division of Water Resources in Alamosa, CO (May 1 2017)
- Presentation at Reclamation’s Albuquerque Area Office to a group of approximately 30 local water management stakeholders in the Middle Rio Grande valley (via webinar; January 23, 2018)
- Presentation to Rio Grande Compact Engineer Advisor’s Meeting for Compact Year 2017 (March 6, 2018)

Ongoing Research and Collaboration

The team has remained in contact with the NRCS to discuss and design ways to incorporate the research results into NRCS's operational forecasting model. Dr. Lehner has also been in close contact with Carolyn Donnelly, water operations supervisor, and Lucas Barrett, a hydrologic modeler in Reclamation's Albuquerque Area Office, and provided Lucas Barrett with the project's experimental streamflow forecasts to be used in the Upper Rio Grande Water Operations Model (URGWOM) to project the 2018 snowmelt runoff in the Rio Grande Basin. Thus, the project is beginning to directly benefit Reclamation's water operations in the Rio Grande basin, and has the potential to similarly benefit forecasts and water operations in other parts of Reclamation's service area.

Reclamation also cares about the timing of runoff as it occurs in spring of each year. That is, Reclamation water managers would like to know ahead of time in which week the peak streamflow will occur. This interest is motivated by ecological considerations, mostly related to fish spawning and riparian health. Initial work is underway to explore the predictability of peak streamflow for the Upper Colorado and Rio Grande using antecedent hydrologic information, such as snow pack, accumulated precipitation, or soil moisture, as well as seasonal climate information. Only a few streamflow gages offered sufficient correlations between these potential predictors and peak streamflow, and in almost all cases this correlation did not translate into predictive skill in a cross-validation hindcasting framework. Other sources of predictability for peak streamflow timing were discussed (e.g., radiative forcing from dust on snow), but ultimately were not pursued due to lack of clear applications potential and/or overlap with other research groups in the field.

Project supervision and interaction

At the beginning of the project, a repeating monthly call with Andy Wood, as well as Douglas Blatchford and Dagmar Llewellyn from Reclamation was established, leading to a regular exchange of ideas and feedback that was very useful for tracking progress and making adjustments to the research focus, where necessary. Mr. Blatchford and Ms. Llewellyn invited relevant people from Reclamation, the USGS, and other water-management institutions to these calls, thereby allowing the research effort to be optimally aligned with stakeholder needs.

The project supervisors arranged a number of interactions with the decision makers in water operations that were critical for the success of the project. These included a visit to the Lower Colorado Operations Office and the Yuma Area Office, where the exposure of the research led to invited talks in meetings with the Southern Nevada Water Authority, and with Colorado and Rio Grande river stakeholders. Dr. Lehner was also able to give a presentation at the National Center for Environmental Prediction (at the Subseasonal-to-Seasonal, or S2S, Science Meeting).

The team also facilitated insightful site visits to Reclamation projects in Yuma AZ, Boulder City NV, Alamosa, CO, and Taos and Chama, NM. Similarly, Dr. Wood used his many contacts in both the private, federal, and academic sector of water management in the Western US to support the project with critical data and information.

Financial reporting

The total Reclamation funding for this PACE fellowship was $119,980, including funding in Fiscal Year 2015, 2016, and 2017.

Funding available under this PACE Fellowship did not adequately cover some of the publication costs and travel of Dr. Lehner, nor the needed mentoring time for Dr. Lehner by Dr. Wood. To successfully complete this project, Dr. Wood leveraged contributions from his ongoing Reclamation forecasting project, and dedicated time from his project to mentor Dr. Lehner and contribute to the scientific outcomes of this fellowship. Dr. Wood also organized NCAR funds to pay for new computing infrastructure for Dr. Lehner. Dr. Wood facilitated a smooth project progression by managing budget through leveraging synergies with his other projects.

Next Steps

As a result of the research initiated under this PACE fellowship, two projects have been funded under Reclamation's Fiscal Year (FY) 2018 Science & Technology Program. These follow-on projects are described below. Dr. Lehner has been promoted to a Project Scientist I position at NCAR in order to continue his involvement with this research, through these Science & Technology Program projects. Work on this effort is expected to be initiated in May, 2018.

The first of these follow-on projects is three-year project titled *Improving the robustness of southwestern US water supply forecasting in the face of climate trends and variability.* This project will identify parts of the Reclamation management domain that are experiencing climate trends that may be undermining the effectiveness and skill of seasonal water supply forecasts (WSFs) that are used to inform water management decisions. 'Skill' is a multi-faceted description of the quality of a forecast, including components of accuracy, reliability, and precision. It will also identify regions in which seasonal climate (and particularly temperature) forecasts from the National Multi-Model Ensemble (NMME) may be sufficiently skillful to make WSFs responsive to seasonal and interannual climate trends, and thereby enhance their skill. In the Upper Rio Grande basin, where research under Dr. Lehner's PACE fellowship already provided evidence that such changes are underway, the project will refine and demonstrate NMME-based strategies for improving the resilience of the existing operational forecasting methods (statistical prediction and model-based Ensemble Streamflow Prediction, or ESP) at key water management input locations (to the existing RiverWare models), enabling experimental operational scenario analysis. The work will lead to operations-ready methods suitable for adoption in the two major forecasting agencies, the National Weather Service (NWS) and the Natural Resources Conservation Service (NRCS). The Reclamation-wide assessment of hydroclimate sensitivities to trends and NMME climate forecast skill will provide foundational information for potential follow-on decision support studies in basins in Reclamation's service area other than the Upper Rio Grande.

The second of these follow-on projects is also a three-year project, which is titled "*Development of short-range forecasts of weather-driven channel losses and gains to support Reclamation Water Management."* On the field trip to the Reclamation Yuma Area Office (September 13-15 2016), Reclamation water-management staff Edward Virden (Director of Operations) and Hong Nguyen-DeCorse indicated a strong need for improved short-term precipitation forecasts to support water operations in the Yuma area. Although only exploratory work was conducted on the Lower Colorado basin during the PACE fellowship, clear research needs were identified during the site visit. These needs eventually led to a successful proposal to Reclamation's 2018 Science & Technology Program in Fiscal Year 2018 to conduct research on short-term weather and loss-gain modeling in the Lower Colorado Basin. For this project, a collaboration with NOAA's Michael Scheuerer and Tom Hamill, both in Boulder, CO, was established, with the goal of improving the forecast skill of quantitative, short-term precipitation forecasts in the Yuma area.

Appendix B

Peer-Reviewed Publication for Research Phase I:

Lehner, F., E. R. Wahl, A. W. Wood, D. Blatchford, and D. Llewellyn, 2017a: Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective Geophys. Res. Lett., doi:10.1002/2017GL073253

Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective

Flavio Lehner[1], Eugene R. Wahl[2], Andrew W. Wood[1], Douglas B. Blatchford[3], Dagmar Llewellyn[4]

[1]Research Application Laboratory, National Center for Atmospheric Research, Boulder, USA

[2]Paleoclimatology Group, NOAA's National Centers for Environmental Information, Boulder, USA

[3]Lower Colorado Regional Office, Bureau of Reclamation, Boulder City, USA

[4]Albuquerque Area Office, Bureau of Reclamation, Albuquerque, USA

Corresponding author: Flavio Lehner (flehner@ucar.edu)

Key Points:

- The decreasing runoff efficiency trend from 1986-2015 in the Upper Rio Grande River basin is unprecedented in the last 445 years
- Very low runoff ratios are 2.5 to 3 times more likely when temperatures are above-normal than when they are below-normal
- The trend arises primarily from natural variability but runoff sensitivity to temperature implies further declines should warming continue

Abstract

Recent decades have seen strong trends in hydroclimate over the American Southwest, with major river basins such as the Rio Grande exhibiting intermittent drought and declining runoff efficiencies. The extent to which these observed trends are exceptional has implications for current water management and seasonal streamflow forecasting practices. We present a new reconstruction of runoff ratio for the Upper Rio Grande basin back to 1571 CE, which provides evidence that the declining trend in runoff ratio from the 1980s to present-day is unprecedented in context of the last 445 years. Though runoff ratio is found to vary primarily in proportion to precipitation, the reconstructions suggest a secondary influence of temperature. In years of low precipitation, very low runoff ratios are made 2.5-3 times more likely by high temperatures. This temperature sensitivity appears to have strengthened in recent decades, implying future water management vulnerability should recent warming trends in the region continue.

1 Introduction

Streamflow in most watersheds in the American Southwest is driven primarily by winter precipitation, with a secondary contribution from summer precipitation (Serreze et al. 1999). Much of the winter precipitation falls as snow in the mountains and runs off in spring and early summer, and peak snowmelt-driven streamflows typically occur between March and July. The influence of summer precipitation increases to the south due to the increased influence of the North American monsoon (Woodhouse et al. 2013), but the headwater regions of rivers such as the Colorado and Rio Grande are dominated by winter precipitation. Seasonal outlooks for runoff volume driven by spring snowmelt, termed water supply forecasts (WSFs), leverage the relationship between winter precipitation and summer streamflow by using predictors such as observed winter snow water equivalent (SWE) and accumulated precipitation to forecast spring runoff. Forecasts have traditionally been made beginning in January of the same year (Pagano et al. 2014). The skill of these WSFs at longer lead times depends on both the strength and stability of the relationship between these predictors and the coming spring runoff. The runoff ratio, or the fraction of runoff generated by a given amount of precipitation, can serve as a simple metric illustrating the efficiency of this translation. Hence, decadal variations in runoff ratio would indicate non-stationarity in this translation, which in turn can alter the forecast skill. In the context of WSFs, relevant runoff ratio calculations might include the spring streamflow volume divided by winter precipitation up to the start or end of the forecast period. In the context of assessing hydroclimate variability more generally, and as necessitated by the temporal resolution of currently available paleoclimate reconstructions, total water year (October-September) streamflow and precipitation might be used.

In the American Southwest, and specifically the Upper Rio Grande River basin (URG), annual runoff ratios are sensitive to a number of factors. The relative contributions of winter, spring, and summer precipitation to the water year (WY) total precipitation are important because summer precipitation typically does not contribute to streamflow as much as winter precipitation (Hamlet et al. 2005), in part due to the higher evaporative losses in summer. In WSFs, for example, primarily the winter precipitation is used as a predictor, while spring and summer precipitation variability after the forecast date contributes to the forecasting uncertainty (Pagano et al. 2004; Rosenberg, et al. 2011), especially since winter and summer precipitation in the American Southwest are not necessarily correlated on interannual time scales (Griffin et al. 2013; Coats et al. 2015). Spring temperatures and wind speeds, which control evaporative loss, also influence the magnitude and timing of peak SWE in the headwaters (Dettinger and Cayan 1995). Human influences can strongly modify natural streamflows; among these, groundwater pumping (Alley et al. 2002) is less easily corrected for than other impairments such as reservoir storage operations and measured diversions and return flows. Finally, recent research has suggested that dust loading on snowpack can induce earlier melt and reduced runoff volumes (Painter et al. 2010).

Unexpected seasonal, interannual, or decadal variations in any of these factors can lead to WSF biases. In the URG, water resources managers have noted systematic over-forecasting biases in the recent decade. The 2000s and 2010s exhibited intermittent drought conditions, thus the forecast model's calibration over a longer period that is relatively wetter on average (i.e., including the wetter decades of the 1980s and 1990s) is likely to be a partial cause of this bias. Indeed, runoff ratios have been declining since the mid-1980s in the adjacent Upper Colorado River basin (Woodhouse et al. 2016) and similar trends exist in the URG (Figure 1). Because the recent decades have also been marked by substantial upward temperature trends, the question of whether runoff ratio declines can be linked to temperature increases, and thus potentially to anthropogenic global warming, have gained the attention of the water management community (Reclamation 2016; Udall and Overpeck 2017).

Assessing the long-term significance of the recent runoff efficiency changes is hindered, however, by relatively short periods of observational records for streamflow, precipitation, and SWE in many watersheds of the American West, which limit the data available for training statistical forecast models. This obstacle motivates the development of reconstructions of streamflow, precipitation, temperature, and their relationships that extend beyond the instrumental period and thus place recent variations in runoff ratio and associated forecast biases in the URG into a longer-term context. There have been extensive efforts to understand hydrologic variability and improve seasonal forecasting in the Colorado River basin (Franz et al. 2003), but less attention has been paid to the URG. Notably, an estimated 5 million people depend on Rio Grande River water, which is shared between the US and Mexico, making it one of the most allocated rivers in the world (Dahm et al. 2005).

Here we use existing and new reconstructions of annual streamflow and precipitation to extend the record of runoff ratio of the URG back to 1571 of the Common Era (CE). We use these records to assess the extent to which observed changes in WY runoff ratio have precedent over the past 445 years. The close correspondence between WY runoff ratio and seasonal runoff ratios, as discussed above, makes this analysis relevant for water resource management. In addition, we use temperature reconstructions and a climate model simulation to investigate the role of temperature and large-scale circulation patterns in influencing periods of high and low runoff ratio.

2 Materials and Methods

2.1 Observational data sets

We use naturalized monthly Rio Grande River streamflow at Otowi Bridge (USGS 08313000; commonly referred to as Otowi Index Supply) from 1942-2015 obtained from the State of New Mexico (Nabil Shafike, personal communication). The naturalization does not include potential impairments from groundwater pumping, the influence of which on streamflow is not currently well constrained. For precipitation and surface air temperature, we use the Parameter Elevation Regression on Independent Slopes Model (PRISM) data set from 1895-2015 (Daly et al. 2008) and spatially average each field across the surface drainage area corresponding to the the Otowi Bridge gauge, defined by the hydrologic unit code (HUC6) regions 130100 and 130201. For precipitation, we multiply the average value with the drainage area of this mask to convert it to units of volume.

2.2 Paleoclimate reconstructions

We use existing tree ring-based reconstructions of water year (October-September) streamflow at the Otowi Bridge gauge, as well as water year precipitation and annual mean temperature over the associated drainage basin, covering a common period (1571-1977). The streamflow reconstruction uses moisture-sensitive tree-ring species, which reflect a combination of winter precipitation and summer evapotranspiration and thus capture key features of streamflow variability (Woodhouse et al. 2012). It was calibrated against naturalized flows in the 20th century and covers the period 1450-2012 CE (updated version; http://www.treeflow.info/content/rio-grande-owoti-new-mexico-update).

The precipitation reconstruction is a modified version of the 0.5° x 0.5° western US precipitation reconstruction by Diaz and Wahl (2015), covering the period 1571-1977. The precipitation reconstruction relies on tree ring-based streamflow reconstructions, but, crucially for the study here, the streamflow reconstruction from Otowi Bridge has been excluded in the construction of this modified version. Thus, the streamflow and precipitation reconstructions used here are largely independent, with very few shared original chronologies (Table S1 and Supplementary Material Section 1). To estimate precipitation in the Rio Grande basin upstream of Otowi Bridge, we spatially averaged the reconstructed precipitation over the aforementioned Otowi drainage region and multiplied it by the drainage area to obtain units of volume.

For annual mean temperature we extract a 5° x 5° grid cell centered at 37.5 °N, 107.5 °W (which corresponds roughly to the Rio Grande headwaters) from the reconstruction by Wahl and Smerdon 2012. The coarse spatial resolution of this reconstruction does not weaken the analysis here because the length scale of high spatial correlation ($r>0.8$) of the URG annual mean temperature encompasses the size of the selected grid cell in observations (Figure S1). While the choice of annual mean is motivated by the available reconstruction data, we note that annual mean and the more critical melt season (Mar-Aug) mean temperature in the URG basin are highly correlated ($r=0.74$ in observations 1895-2015). For the determination of reconstruction uncertainties see Supplementary Material Section 3.

2.3 Model simulation

We use an 1,800-year long preindustrial control simulation (piControl) from the Community Earth System Model (CESM), which is described in detail by Kay et al. (2015). CESM is a fully-coupled Earth System Model with components of atmosphere, ocean, sea ice, and land surface (Hurrell et al. 2013). In the configuration here, all components are run at ~1°x1° horizontal resolution. The forcing represents perpetual 1850 CE conditions for atmospheric composition, orbital parameters, and land cover.

We extract streamflow from this simulation by extracting the routed runoff at the Otowi Bridge location from the 0.5° x 0.5° River Transport Model embedded in CESM. We recognize that the CESM runoff and routing schemes are coarse and contain climatological biases at the watershed scale, but we expect that they will sufficiently discriminate high and low flows driven by large scale climate variations to be useful in the context of this study. CESM precipitation and surface air temperature are then extracted by mapping the Otowi Bridge drainage area onto the CESM grid.

3 Results

3.1 Hydroclimate over the past four centuries

At Otowi Bridge, streamflow has varied on interannual to decadal time scales, with pronounced periods of low flow as identified and discussed in Woodhouse et al. (2012). Figures 1a and 1b show the reconstructed and observed time series of WY precipitation and streamflow for Otowi Bridge, Figure 1c shows the runoff ratio resulting from dividing streamflow by precipitation, and Figure 1d shows annual mean temperature. Beyond the decadal time scale, however, no prolonged periods of high or low flow were recorded in the reconstruction, consistent with other Southwestern US findings that multi-decadal drought conditions were more prevalent in the first half of the last millennium (Cook et al. 2004; Meko et al. 2007; Coats et al. 2016), although there is a 16th century megadrought that ended just before our reconstructions begin (Stahle et al. 2000). Comparing the last four centuries of reconstructed streamflow to the recent decades of measured streamflow at Otowi Bridge clearly indicates that the observed annual high values of the 1980s and the low value of 2002 are exceptional, but not unprecedented. The highest value in the observations is 2,074 KAF (1,000 acre feet; in 1985) and lowest value is 235 KAF (in 2002), whereas the highest reconstructed value is 2,123 KAF (in 1720) and the lowest value is 216 KAF (in 1685). Due to uncertainties in the reconstruction (Woodhouse et al. 2006), which are likely larger than the margin between the observed and reconstructed highest and lowest flows, it is uncertain but conceivable that these recent extrema are the highest and lowest flows in more than 400 years.

The 10-year smoothed time series (thick line in Figure 1a) clearly shows the 1980s to be the decade of highest flow over the whole time period, while the early 2000s tie within uncertainties with the 1580s and 1770s for the decade of lowest flow. Most importantly, the short sequencing of the exceptionally high- and low-flow decades within the last 30 years results in this period showing the strongest 30-year streamflow trends of the entire period 1571-2015 (histogram in Figure 1a, at 99.1% probability; see Supplementary Material, Section 4).

Precipitation is strongly correlated with streamflow (r=0.75 in reconstructions 1571-1942, r=0.89 in reconstructions 1943-1977, r=0.77 in observations 1943-1977, r=0.79 in observations 1943-2015) and largely drove the extreme streamflow periods in both reconstructions and observations (Figure 1b). Precipitation is also strongly correlated (r=0.73, 1571-1977) with a reconstruction of April snow water equivalent in the Rio Grande headwaters (Pederson et al. 2011), suggesting that winter-spring precipitation can explain at least 50% of water year precipitation variability. Similar to streamflow, the 1980s stand out as an exceptional decade with precipitation values almost consistently above the long-term average derived from the reconstruction. Consequently, this wet decade and the subsequent decline into the generally drier 2000s also produced the strongest 30-year precipitation trend of the entire period (histogram in Figure 1b, at 97.9% probability). Interestingly, the 1990s were exceptionally wet as well, but had lesser impact on the streamflow record than the 1980s (compare Figure 1a and 1b), leaving room for additional explanatory factors, as discussed later.

Due to the strong influence of precipitation on streamflow and runoff ratio in these arid regions (Vano et al. 2012), the reconstructed time series of runoff ratio features many of the same high and low value periods as the precipitation reconstruction (Figure 1c). Again, the 1980s show exceptionally high runoff ratios and the decline into the early 2000s also marks the strongest 30-year trend in the entire period (histogram in Figure 1c, at 97.8% probability). However, there are a few periods, including the 1990s and the mid-nineteenth century, in which the relationship between precipitation and runoff ratio appears to be weaker.

Compared to precipitation and streamflow, reconstructed temperature in the URG shows distinct multi-decadal (relatively lower frequency) variations (Figure 1d). A roughly century-long cold period between 1600 and 1700 was followed by a similarly long period of above-long-term mean temperatures, followed by a sharp decrease and then gradual rise of temperature until present-day. The highest reconstructed temperatures occurred in the late 18th century and rival the observed high values of the 20th and 21st century, although the past 15 years are clearly the warmest period of such length over the last 440+ years (cf. Wahl and Smerdon 2012). Unlike reconstructed streamflow, precipitation, and runoff ratio, observed 30-year temperature trends fall well within the distribution of the reconstruction.

3.2 Role of temperature

While precipitation is the main driver of interannual streamflow variations in the URG, temperature also influences streamflow and hence runoff ratio. To investigate the role of temperature in interannual variations of runoff ratio, we plot runoff ratio (in percentile units) as a function of precipitation and temperature anomalies (Figure 2; all time series are relative to their median due to the non-Gaussian distribution of precipitation and runoff ratio, see Figure S2a). First, the figure illustrates the weak, but statistically significant negative correlation between temperature and precipitation ($r=-0.28$ in reconstructions 1571-1942, $r=-0.39$ in reconstructions 1943-1977, $r=-0.37$ in observations 1943-1977, $r=-0.30$ in observations 1943-2015) that is typical for this region (Trenberth and Shea 2005; correlation coefficients between −0.30 and −0.50 based on reanalysis data). Second, the stratification of high and low runoff ratio years according to associated precipitation anomalies clearly shows that positive precipitation anomalies are an important prerequisite for high runoff ratios with 76% of the years of high (> 70th percentile) and 88% of the very high (> 90th percentile) runoff ratios coinciding with positive precipitation anomalies in reconstructions (upper two quadrants in Figure 2a). In turn, 81% of the low (< 30th percentile) and 95% of the very low (< 10th percentile) runoff ratio years coincide with negative precipitation anomalies. Third, and most importantly for this study, a further stratification according to temperature shows that when precipitation is below the median, low and very low runoff ratios are 1.7 and 2.5 times as likely to occur, respectively, in warm years (51% and 68%; bottom right quadrant in Figure 2a) than in cold years (30% and 27%; bottom left quadrant in Figure 2a). Also, there exists a significant correlation between runoff ratio and temperature that is almost entirely driven by the relationship of the two variables in dry years, with no significant correlation in wet years (Figure S2b-d). Repeating the analysis with reconstructions of summer and annual maximum monthly temperature does not alter these conclusions (Supplementary Material Section 5 and Figure S3).

The relationships found in the reconstructions are also clearly visible in the shorter (73 years) observational record (Figure 2b), which exhibits strong warming during the recent decades. In fact, 86% and 88% of all low and very low runoff ratio years, respectively, were dry and warm, while 0% and 13% of the low and very low runoff ratio years, respectively, were dry and cold. Notwithstanding the uncertainties due the small observational sample, the recent warm decades appear to have been an important factor in very low runoff ratio years.

Turning to the CESM simulation, we find the model generally reproduces the sensitivities of runoff ratio that are found in the reconstructions and observations (Figure 2c): 59% of high runoff ratio years and 65% of very high runoff ratio years occur in wet years (above-median precipitation; top two quadrants in Figure 2c). In turn, 59% of low and 67% of very low runoff ratio years occur in dry years (below-median precipitation; bottom two quadrants in Figure 2c). Further, the apparent importance of high temperatures for the occurrence of very low runoff ratio years is found in CESM as well: 50% of the very low runoff ratio years occur in a dry and warm year, while only 17% occur in a dry and cold year, making it approximately 3 times as likely to have a very low runoff ratio year if temperatures are above normal rather than below normal (Figure 2c). Notable differences of the model output from the reconstructions and observations are the percentages of very high runoff ratio years when it is dry and cold, and the opposite very low runoff ratio years when it is wet and warm (Figure 2c, bottom-left and upper-right quadrants, respectively).

Due to the negative correlation between precipitation and temperature there exists a natural tendency for dry years to coincide with warm years. To account for this, we investigated the likelihood for very low/low/high/very high runoff ratios conditional on the background climate of the respective year and find the results reported above to be robust (Supplementary Material Section 6 and Figure S4).

3.3 Circulation composites

To investigate the large scale atmospheric circulation patterns potentially associated with certain cases of very high and low runoff ratios in the reconstruction, we search for analogous situations in CESM and create composite maps. Here, we focus on the following four situations (using the very low/low/high/very high categories defined above):

- A: Years with very high runoff ratio, high precipitation, and low temperature.
- B1: Years with very high runoff ratio, below-median precipitation, and below-median temperature.
- B2: Years with very low runoff ratio, below-median precipitation, and below-median temperature.
- C: Years with very low runoff ratio, low precipitation, and high temperature.

To construct the composites, we extract sea level pressure (SLP), precipitation, and temperature during the years that fulfill the above criteria from the CESM simulation and average them (Figure 3). Naturally, not all four situations occur with equal frequency in the 1,800 model years analyzed; all four composites combined cover 9.3% of the 1,800 total model years.

Situation A features a deep Aleutian Low over the North Pacific in both the cold (October-March) and warm (April-September) seasons, leading to a strong North-South temperature gradient across North America and high precipitation totals over much of the contiguous US (hereafter "US"; Figure 3a-b). Both cold and warm season responses are robust over much of North America and the North Pacific (no stippling in Figure 3). Situation A in the cold season is reminiscent of the canonical El Niño response over the North Pacific-North America region. Indeed, 57% of all situations A coincide with a winter (Dec-Feb) in which the Nino 3.4 index (sea surface temperatures averaged over 170-120 °W, –5-5 °N) exceeds 1 standard deviation.

Situation B1, in which very high runoff ratios occur with below-median temperatures but in conjunction with below-median precipitation, shows a sharp contrast between cold and warm season in terms of circulation and precipitation (Figure 3c-d). The cold season features a wave train across the Pacific and North America, somewhat resembling the surface signature of the Pacific North American pattern. A deep Aleutian Low channels cold air from the Bering Sea to the US, while northern Canada receives positive temperatures anomalies due to the southerly flow on the east side of the Aleutian Low (Figure 3c). Together with another low pressure anomaly over the US East Coast, these two SLP anomalies cause substantial positive precipitation anomalies across large parts of the US. In the warm season, the low pressure anomaly over the Northeast Pacific is weaker, and the SLP anomaly on the US East coast moves further inland (Figure 3d). The resulting flow across the central US is predominantly northerly, causing dry and cold conditions and counteracting the moisture influx into the Southwestern US that is typical for the North American summer monsoon. The contrasting precipitation totals from the cold and warm season result in a net negative water year precipitation anomaly, but due to the high accumulation in the cold season and the relatively cold warm season, runoff ratios remain very high.

Situation B2, in turn, during which very low runoff ratios occur during years of below-median precipitation and temperature, features a positive SLP anomaly over much of the northeastern Pacific in the cold season, diverting incoming storms from the Pacific to Canada and steering cold Arctic air across most of North America (Figure 3e). In the warm season, the positive SLP anomaly over the North Pacific is weaker and no clear circulation patterns are established over the US (Figure 3f), although temperatures are slightly elevated and precipitation is slightly reduced over much of the western US. As a net result, this situation is mainly dominated by the cold season precipitation deficit, which together with slightly above-average temperatures during the warm season, appears to be sufficient to drive very low runoff ratios.

Finally, situation C, during which some of the lowest runoff ratios of all CESM years occur is in many ways the reversal of situation A, with high temperatures and low precipitation over most of the US in both seasons (Figure 3g-h). The key features of the cold season composite are negative SLP anomalies over the Gulf of Alaska and a blocking high over the US West coast, steering storms into the northern half of the US West coast, while leaving the southern half of the coast and the central US dry (Figure 3g). While resembling La Niña, only 38% of the winters in this composite show a Nino 3.4 index < -1 standard deviation. In the warm season, the blocking high over the Pacific persists and a thermal low sets in over the central US, creating very warm and dry conditions, and further decreasing streamflow relative to the precipitation decrease, thus causing anomalously low runoff ratios (Figure 3h).

All of these situations resemble viable climatological circulation patterns and can arise from unforced climate variability, as demonstrated by the use of a control simulation, which were found to contain substantial multi-decadal variability of large scale circulation patterns (Deser et al. 2012a). Our results therefore suggest that decadal variations in the frequency of these circulation patterns (for example associated with the relative frequency of El Niño and La Niña events in recent decades; Meehl et al. 2009) might not, or only to a small degree, be associated with externally forced climate change, e.g., from increasing greenhouse gas concentrations. Due to the short observational record and the small signal-to-noise ratio of forced sea level pressure trends in simulations (Deser et al. 2012b), detection and attribution of anthropogenically forced changes in observed circulation patterns and hydroclimate over the American Southwest remains an active area of research (Prein et al. 2016).

4 Summary and conclusions

In summary, paleoclimate reconstructions suggest that both the high and low annual runoff ratios of the most recent decades in the URG were extreme in context of the last 440 years. As a consequence, the 30-year declining trend in runoff ratio from the mid-1980s to present-day appears to be unprecedented, and is problematic for current statistical seasonal streamflow forecasting approaches that assume hydroclimatic stationarity. Although decadal-scale trends in runoff ratio are driven primarily by precipitation variations, the paleoclimate record also reveals an important role for temperature in creating some of the lowest runoff ratio years in the last four centuries. Supported by a long climate model simulation, we estimate that in years with below-median precipitation, very low (< 10th percentile) runoff ratios are 2.5-3 times more likely if temperatures are warmer than normal (above-median).

If recent warming trends continue, our findings suggest a further decline in runoff ratios in the URG and other Southwestern US basins. Nevertheless, the paleoclimate record and associated circulation composites indicate that low and high runoff ratios of almost equal magnitude as observed in recent decades are possible in the absence of any significant greenhouse gas forcing trend. In this light, careful detection and attribution is warranted when diagnosing underlying causes of recent hydroclimate trends in the Southwestern US.

Acknowledgments

We thank Andrew Newman, Keith Musselmann, Kevin Sampson, and Nabil Shafike for discussion or technical support, as well as two anonymous reviewers for their very constructive feedback. Part of this work was inspired by a workshop held at the Lamont-Doherty Earth Observatory of Columbia University on June 1-3, 2016, titled "Comparing data and model estimates of hydroclimate variability and change over the Common Era". We acknowledge the efforts of all those who contributed to producing the CESM LE and the paleoclimate reconstructions. The CESM simulation is available on the Earth System Grid (www.earthsystemgrid.org), the temperature reconstruction is available at NOAA Paleoclimatology (www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets), the streamflow reconstruction is available at www.treeflow.info, the precipitation reconstruction and the Otowi Index Supply are available from the corresponding author. The National Center for Atmospheric Research is sponsored by the National Science Foundation. F. L. is supported by a Postdoc Applying Climate Expertise (PACE) fellowship cosponsored by NOAA and the Bureau of Reclamation, A. W. W. is supported by Reclamation under Cooperative Agreement #R11AC80816, and by the US Army Corps of Engineers (USACE) Climate Preparedness and Resilience Program. F.L. designed the study, conducted the analysis, and led the writing; E. R. W. created all new reconstructions; E. R. W. and A. W. W. helped design the study; all authors contributed to the writing.

References

Alley, W. M., R. W. Healy, J. W. LaBaugh, and T. E. Reilly, 2002: Flow and storage in groundwater systems. *Science*, **296**, 1985–1990, doi:10.1126/science.1067123.

Coats, S., J. E. Smerdon, R. Seager, D. Griffin, and B. I. Cook, 2015: Winter-to-summer precipitation phasing in southwestern North America: A multicentury perspective from paleoclimatic model-data comparisons. *J. Geophys. Res. Atmos.*, **120**, 8052–8064, doi:10.1002/2015JD023085.

——, J. E. Smerdon, K. B. Karnauskas, and R. Seager, 2016: The improbable but unexceptional occurrence of megadrought clustering in the American West during the Medieval Climate Anomaly. *Environ. Res. Lett.*, **11**, 74025, doi:10.1088/1748-9326/11/7/074025. http://stacks.iop.org/1748-9326/11/i=7/a=074025?key=crossref.9488431852c8066bb4124d4718e10863.

Cook, E. R., C. A. Woodhouse, C. M. Eakin, D. M. Meko, and D. W. Stahle, 2004: Long-term aridity changes in the western United States. *Science*, **306**, 1015–1018, doi:10.1126/science.1102586. http://www.ncbi.nlm.nih.gov/pubmed/15472040.

Dahm, C. N., R. J. Edwards, and F. P. Gelwick, 2005: Gulf Coast Rivers of the Southewestern United States. *Rivers of North America*, 180–228.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, doi:10.1002/joc.1688.

Deser, C., and Coauthors, 2012a: ENSO and pacific decadal variability in the community climate system model version 4. *J. Clim.*, **25**, 2622–2651, doi:10.1175/JCLI-D-11-00301.1.

——, A. Phillips, V. Bourdette, and H. Teng, 2012b: Uncertainty in climate change projections: The role of internal variability. *Clim. Dyn.*, **38**, 527–546, doi:10.1007/s00382-010-0977-x. http://link.springer.com/10.1007/s00382-010-0977-x (Accessed July 21, 2014).

Dettinger, M. D., and D. R. Cayan, 1995: Large-scale atmospheric forcing of recent trends towards early snowmelt runoff in California. *J. Clim.*, **8**, 606–623, doi:10.1175/1520-0442(1995)008<0606:LSAFOR>2.0.CO;2.

Diaz, H. F., and E. R. Wahl, 2015: Recent California water year precipitation deficits: A 440-year perspective. *J. Clim.*, **28**, 4637–4652, doi:10.1175/JCLI-D-14-00774.1.

Franz, K. J., H. C. Hartmann, S. Sorooshian, and R. Bales, 2003: Verification of National Weather Service Ensemble Streamflow Predictions for Water Supply Forecasting in the Colorado River Basin. *J. Hydrometeorol.*, **4**, 1105–1118, doi:10.1175/1525-7541(2003)004<1105:VONWSE>2.0.CO;2.

Griffin, D., and Coauthors, 2013: North American monsoon precipitation reconstructed from tree-ring latewood. *Geophys. Res. Lett.*, **40**, 954–958, doi:10.1002/grl.50184.

Hamlet, A. F., P. W. Mote, M. P. Clark, and D. P. Lettenmaier, 2005: Effects of temperature and precipitation variability on snowpack trends in the sestern United States. *J. Clim.*, **18**, 4545–4561, doi:10.1175/JCLI3538.1. http://dx.doi.org/10.1175/JCLI3538.1.

Hurrell, J. W., and Coauthors, 2013: The Community Earth System Model: A Framework for Collaborative Research. *Bull. Am. Meteorol. Soc.*, **94**, 1339–1360, doi:10.1175/BAMS-D-12-00121.1. http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-12-00121.1 (Accessed August 17, 2014).

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bull. Am. Meteorol. Soc.*, **published**.

Meehl, G. A., A. Hu, and B. D. Santer, 2009: The mid-1970s climate shift in the pacific and the relative roles of forced versus inherent decadal variability. *J. Clim.*, **22**, 780–792, doi:10.1175/2008JCLI2552.1.

Meko, D. M., C. A. Woodhouse, C. A. Baisan, T. Knight, J. J. Lucas, M. K. Hughes, and M. W. Salzer, 2007: Medieval drought in the upper Colorado River Basin. *Geophys. Res. Lett.*, **34**, doi:10.1029/2007GL029988.

Pagano, T., D. Garen, and S. Sorooshian, 2004: Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 1922–2002. *J. Hydrometeorol.*, **5**, 896–909, doi:10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2.

——, A. Wood, K. Werner, and R. Tama-Sweet, 2014: Western U.S. water supply forecasting: A tradition evolves. *Eos (Washington. DC).*, **95**, 28–29, doi:10.1002/2014EO030007.

Painter, T. H., J. S. Deems, J. Belnap, A. F. Hamlet, C. C. Landry, and B. Udall, 2010: Response of Colorado River runoff to dust radiative forcing in snow. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 17125–17130, doi:10.1073/pnas.0913139107. http://apps.webofknowledge.com.ezp.sub.su.se/full_record.do?product=WOS&search_mode=GeneralSearch&qid=78&SID=N2NEdqxYRhcZduK347X&page=1&doc=1.

Pederson, G. T., and Coauthors, 2011: The unusual nature of recent snowpack declines in the North American cordillera. *Science*, **333**, 332–335, doi:10.1126/science.1201570.

Prein, A. F., G. J. Holland, R. M. Rasmussen, M. P. Clark, and M. R. Tye, 2016: Running dry: The U.S. Southwest's drift into a drier climate state. *Geophys. Res. Lett.*, **43**, 1272–1279, doi:10.1002/2015GL066727.

Reclamation, B. of, 2016: *Climate Change Adaptation Strategy*. 33 pp. https://www.usbr.gov/climate/docs/2016ClimateStrategy.pdf.

Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty, 1999: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resour. Res.*, **35**, 2145–2160, doi:10.1029/1999WR900090.

Stahle, D. W., E. R. Cook, M. K. Cleaveland, M. D. Therrell, D. M. Meko, H. D. Grissino-Mayer, E. Watson, and B. H. Luckman, 2000: Tree-ring data document 16th century megadrought over North America. *Eos (Washington. DC).*, **81**, doi:10.1029/00EO00076.

Trenberth, K. E., and D. J. Shea, 2005: Relationships between precipitation and surface temperature. *Geophys. Res. Lett.*, **32**, 1–4, doi:10.1029/2005GL022760.

Udall, B., and J. Overpeck, 2017: The 21 st Century Colorado River Hot Drought and Implications for the Future. *Water Resour. Res.*, doi:10.1002/2016WR019638.

Vano, J. A., T. Das, and D. P. Lettenmaier, 2012: Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature. *J. Hydrometeorol.*, **13**, 932–949, doi:10.1175/JHM-D-11-069.1. http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-11-069.1.

Wahl, E. R., and J. E. Smerdon, 2012: Comparative performance of paleoclimate field and index reconstructions derived from climate proxies and noise-only predictors. *Geophys. Res. Lett.*,

39, doi:10.1029/2012GL051086.

Woodhouse, C. A., S. T. Gray, and D. M. Meko, 2006: Updated streamflow reconstructions for the Upper Colorado River Basin. *Water Resour. Res.*, **42**, doi:10.1029/2005WR004455.

Woodhouse, C. A., D. W. Stahle, and J. V Diaz, 2012: Rio Grande and Rio Conchos water supply variability over the past 500 years. *Clim. Res.*, **51**, 147–158, doi:Doi 10.3354/Cr01059.

Woodhouse, C. A., D. M. Meko, D. Griffin, and C. L. Castro, 2013: Tree rings and multiseason drought variability in the lower Rio Grande Basin, USA. *Water Resour. Res.*, **49**, 844–850, doi:10.1002/wrcr.20098.

Woodhouse, C. A., G. T. Pederson, K. Morino, S. A. McAfee, and G. J. McCabe, 2016: Increasing influence of air temperature on upper Colorado River streamflow. *Geophys. Res. Lett.*, **43**, 2174–2181, doi:10.1002/2015GL067613.

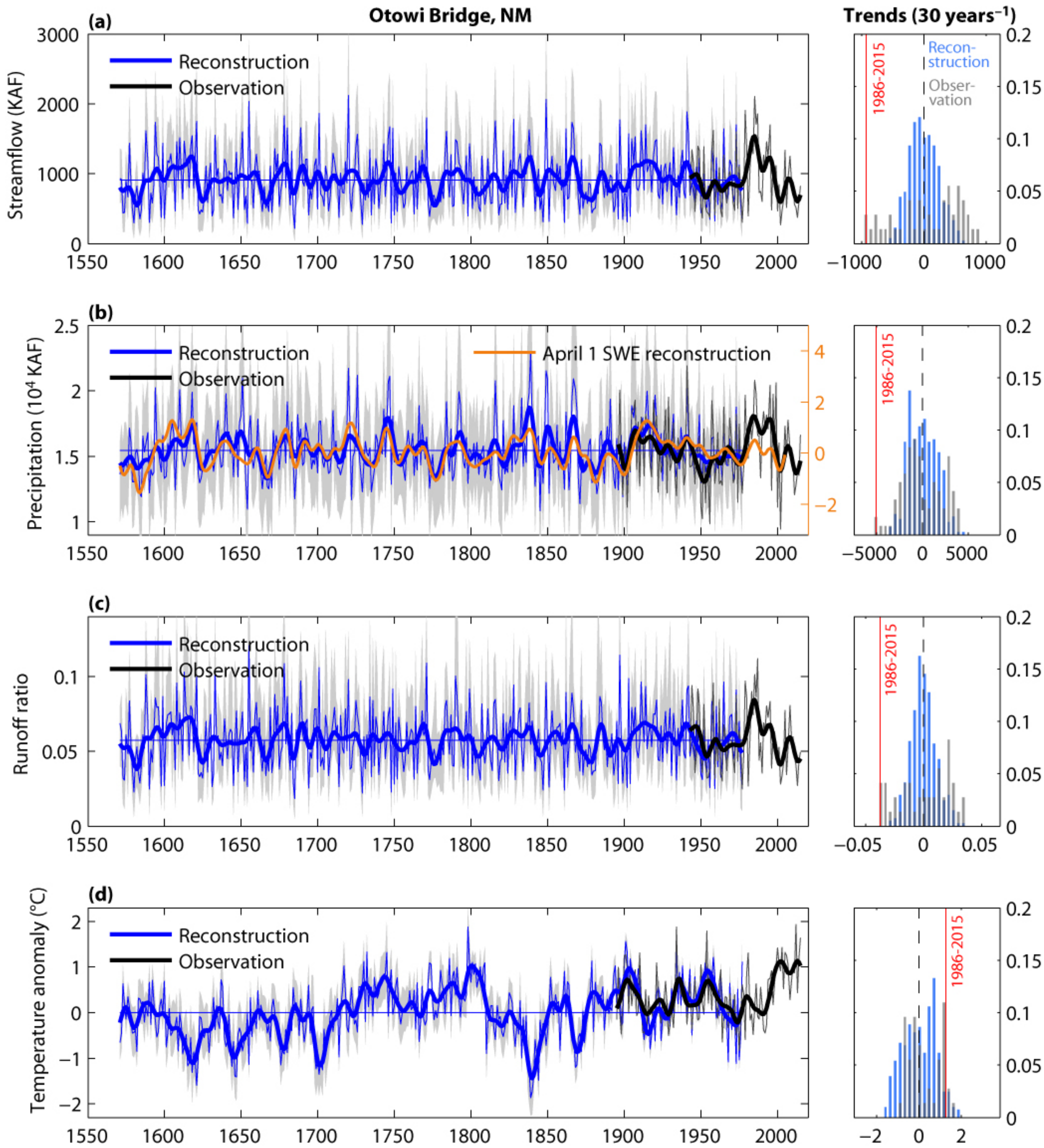


Figure 1. Time series of reconstructed (blue) and observed (black) **(a)** streamflow at Otowi Bridge, **(b)** precipitation upstream of Otowi Bridge and normalized snow water equivalent (SWE; orange), **(c)** runoff ratio for Otowi Bridge, **(d)** average surface air temperature upstream of Otowi Bridge. Thin lines are water year totals: except temperature, which are annual means; and SWE, which is April 1. Thick lines are smoothed with a 10-year Fourier low pass filter. Blue

horizontal lines give the reconstruction mean 1571-1977. Thin gray shading indicates 5-95% reconstruction uncertainty. Right column shows normalized histograms of all 30-year trends of the water year/annual mean data. Red vertical line indicates the most recent 30-year trend 1986-2015. See Section 2 for data sources and details.

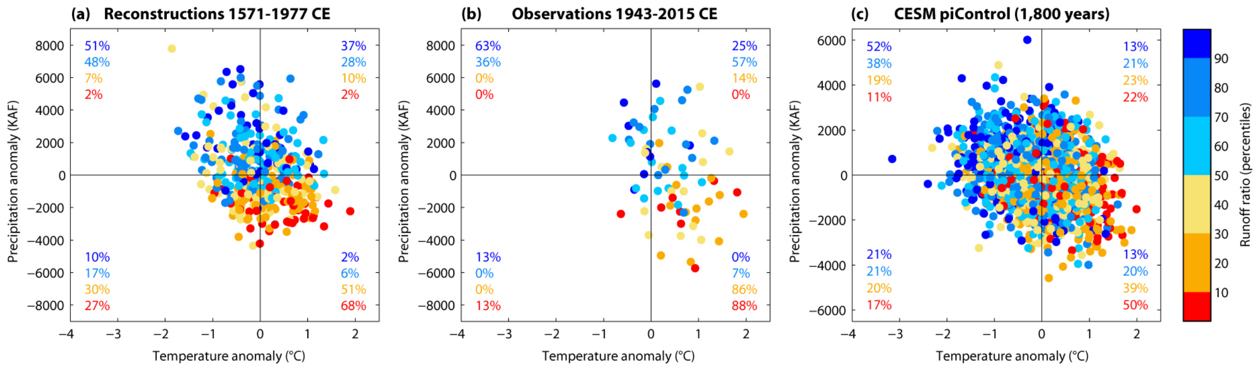
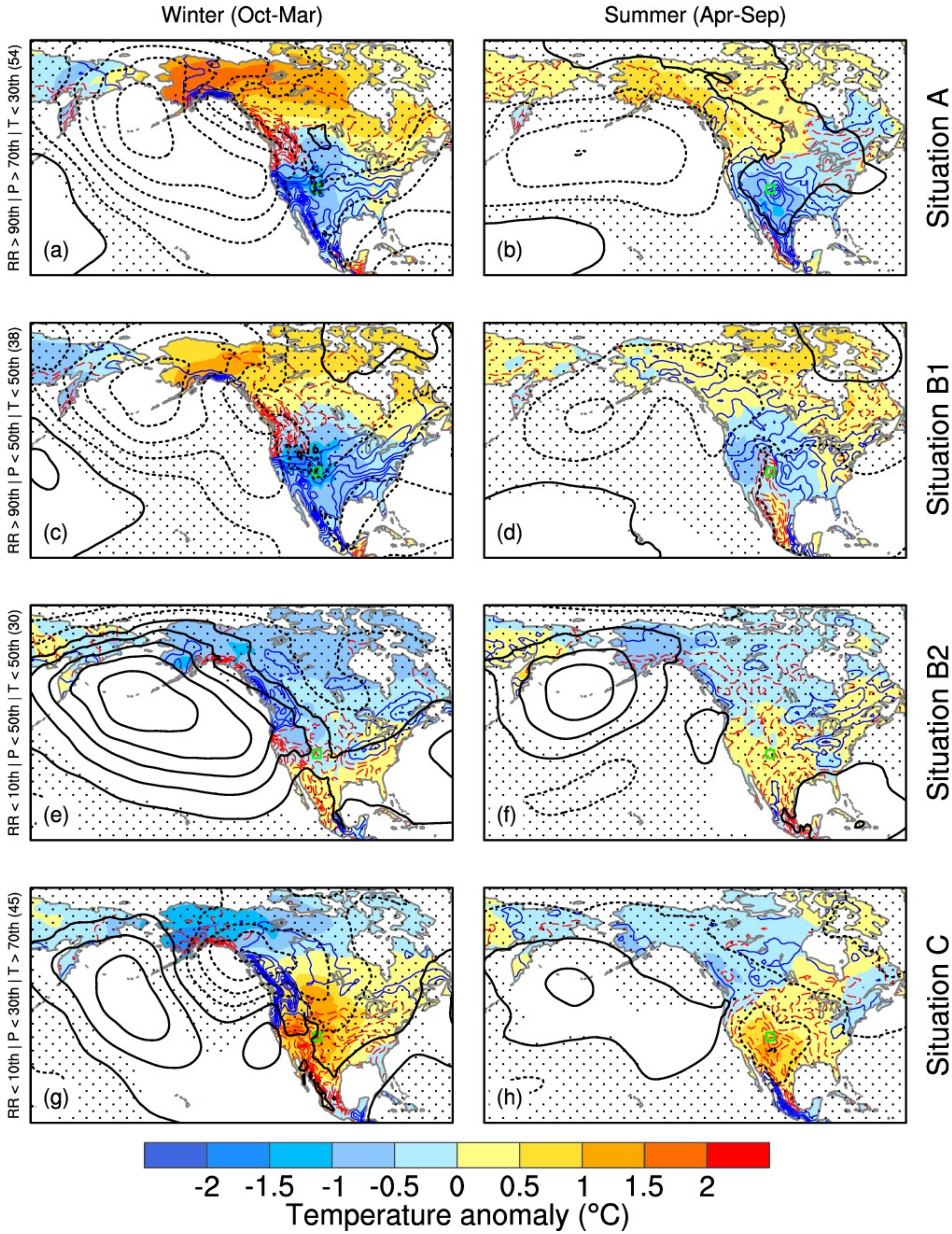


Figure 2. Runoff ratio at Otowi Bridge (colors) as a function of water year precipitation and annual mean temperature from **(a)** reconstructions, **(b)** observations, and **(c)** CESM control simulation (1,800 years total). Time series are relative to their median; in the case of observations, relative to the median of the reconstructions. Colored numbers give the percentage of very low (< 10^{th} percentile), low (< 30^{th}), high (> 70^{th}), and very high (> 90^{th}) runoff ratio years that fall within a given quadrant of precipitation and temperature anomalies.

485



Winter (Oct-Mar)
Summer (Apr-Sep)
RR > 90th | P > 70th | T < 30th (54)
RR > 90th | P < 50th | T < 50th (38)
RR < 10th | P < 50th | T < 50th (30)
RR < 10th | P < 30th | T > 70th (45)
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
Situation A
Situation B1
Situation B2
Situation C
-2 -1.5 -1 -0.5 0 0.5 1 1.5 2
Temperature anomaly (°C)

486

Figure 3. Composite situations from CESM control simulation of temperature (shading), precipitation (blue and red contours; increment of 0.1 mm/day, starting at ±0.05 mm/day), and sea level pressure (black contours; increment of 0.5 hPa, starting at ±0.25 hPa) anomalies for years with (a-b) very high runoff ratio (RR) while precipitation (P) is high and temperature (T) is low, (c-d) very high RR while both P and T are below median, (e-f) very low RR while both P and T are below median, and (g-h) very low RR while P is low and T is high. Left column shows cool season (Oct-Mar) means, right column warm season (Apr-Sep) means. Negative anomalies are given as dashed contours. Stippling (sea level pressure) indicates non-significant difference at 95% probability level. The number of years forming each composite situation is given in brackets. The area of the Upper Rio Grande basin is indicated by the green square.

Appendix C

Peer-Reviewed Publication for Research Phase II:

Lehner, F., A. W. Wood, D. Llewellyn, D. B. Blatchford, A. G. Goodbody, and F. Pappenberger, 2017b: Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. Geophys. Res. Lett., doi:10.1002/2017GL076043.

Mitigating the impacts of climate non-stationarity on seasonal streamflow predictability in the US Southwest

Flavio Lehner[1], Andrew W. Wood[1], Dagmar Llewellyn[2], Douglas B. Blatchford[3], Angus G. Goodbody[4], Florian Pappenberger[5]

[1] Research Applications Laboratory, National Center for Atmospheric Research, Boulder, USA
[2] Bureau of Reclamation, Albuquerque Area Office, Albuquerque, USA
[3] Bureau of Reclamation, Lower Colorado Regional Office, Boulder City, USA
[4] National Water and Climate Center, Natural Resources Conservation Service, Portland, USA
[5] Forecast Department, European Centre for Medium-Range Weather Forecasts, Reading, UK

Corresponding author: Flavio Lehner (flehner@ucar.edu)

revised for *Geophysical Research Letters*
November 2017

Key Points:

- Seasonal temperature forecasts from climate prediction models are skillful over the headwaters of the Colorado and Rio Grande river basins
- Adding temperature information to current operational seasonal streamflow forecasts in snowmelt-driven basins improves forecast skill
- Temperature forecasts help mitigate impacts of non-stationarity on US Southwest streamflow predictability under increasing temperatures

Abstract

Seasonal streamflow predictions provide a critical management tool for water managers in the American Southwest. In recent decades, persistent prediction errors for spring and summer runoff volumes have been observed in a number of watersheds in the American Southwest. While mostly driven by decadal precipitation trends, these errors also relate to the influence of increasing temperature on streamflow in these basins. Here we show that incorporating seasonal temperature forecasts from operational global climate prediction models into streamflow forecasting models adds prediction skill for watersheds in the headwaters of the Colorado and Rio Grande River basins. Current dynamical seasonal temperature forecasts now show sufficient skill to reduce streamflow forecast errors in snowmelt-driven regions. Such predictions can increase the resilience of streamflow forecasting and water management systems in the face of continuing warming as well as decadal-scale temperature variability, and thus help to mitigate the impacts of climate non-stationarity on streamflow predictability.

1 Introduction

With growing populations and rising temperatures, the pressure on water resources in the southwestern United States (US) is increasing and expected to continue to do so over the next decades (Reclamation 2016). Water resources in California, Nevada, Arizona, Utah, Colorado, New Mexico, and Texas are currently almost entirely allocated for agricultural, industrial and municipal uses and are heavily managed, with seasonal streamflow forecasts being a key tool used to inform this management. Seasonal streamflow forecasts for a range of lead times are among the most economically valuable streamflow predictions made in the US and around the world, given their significance for water management (Hamlet et al. 2002; Raff et al. 2013).

Seasonal streamflow forecasts in the Upper Rio Grande river basin, for example, are used to predict the annual water delivery requirements between Colorado, New Mexico, and Texas under an interstate river allocation agreement, the Rio Grande Compact, to plan for water storage and to inform associated reservoir management decisions. The forecasts in combination with those decisions enable projections of the water supplies that will be available to farmers, which in turn can influence cropping decisions. In addition, supplemental water supply to the Upper Rio Grande basin is imported each year from the Colorado River system through trans-basin diversions. Forecasts of the water available for diversion are used to estimate the portion of the imported water that will need to be purchased by the Federal government to support the needs of endangered species, as well as for planning of drinking water operations in major municipalities. On the much larger Colorado River system, as well, water supply forecasts issued in spring are essential to make reservoir storage and release decisions that help avoid shortage conditions in Lake Mead and Lake Powell, and that determine water and hydropower allocations affecting 7 southwestern US states. These decisions influence water and energy costs for major American cities such as Los Angeles, Las Vegas and Phoenix, and major irrigation regions such as California’s Imperial Valley and Arizona’s Welton Mohawk Irrigation and Drainage District.

Although it is difficult to quantify the value of seasonal forecasts or the marginal value of forecast improvements, the value of the water managed using such forecasts rises well into the billions of dollars each year (Hamlet et al. 2002; Pierce 2010). In comparison, the costs of enhancements to operational water supply forecasting are small, especially when they represent an extension of the current approaches, similar to the cost-benefit ratio of improved flood forecasting (Pappenberger et al. 2015). In recent decades the western US has seen strong hydroclimatic trends and decadal variability, leading to variable streamflow forecasting skill and a likelihood of sub-optimal management decisions (Pagano and Garen 2005). To better grapple with water resource management challenges arising from hydroclimate non-stationarity and increasing water demands, improved efficiency in water management practices is critically needed (Milly et al. 2008; Lins and Cohn 2011; Steinschneider and Brown 2012).

Operational seasonal streamflow forecasts in snowmelt driven basins commonly derive skill from the stability of relationships between winter precipitation and snow water equivalent (SWE) with spring to summer melt season runoff (e.g., April-July streamflow). In some cases, but less commonly, additional predictability is found in observations of prior streamflow, soil moisture, and in climate indices such as El Niño-Southern Oscillation (Wood et al. 2005; Koster et al. 2010; Shukla and Lettenmaier 2011; Kalra et al. 2013; Harpold et al. 2017; Bell et al. 2017). The simplest operational form of seasonal streamflow prediction relies on statistical models that quantify these relationships, such as principal component regression (PCR) models trained on observed in situ data records of ~30 years (Garen 1992). These ‘water supply forecasts’ (WSFs) have traditionally been made beginning in January of the same year with updates on the first day of each month to incorporate new precipitation and SWE observations (Pagano et al. 2014b). Operational forecasts are published by regional River Forecasting Centers and the US Department of Agriculture National Resources Conservation Service (NRCS). A second common form of seasonal streamflow prediction involves the use of dynamic watershed models to predict future watershed states and fluxes (Day 1985; Pagano et al. 2014a).

The skill of statistical WSFs varies with lead time and also on decadal time scales, with basins such as the Upper Colorado River (UC) and Upper Rio Grande (URG) showing declining skill since the 1980s (Pagano et al. 2004). While extensive research has been conducted on how to improve seasonal streamflow forecasting systems (Moradkhani et al. 2004; Wood and Lettenmaier 2006, 2008; Crochemore et al. 2016; Mendoza et al. 2017), the reasons for decadal variations in skill of a fixed forecasting system remain relatively elusive. Pagano and Garen (2005) argue that these skill variations originate primarily from interannual to decadal climate variations, rather than basin-specific processes or human interference. As such, successful prediction of interannual to decadal climate variability has the potential to stabilize streamflow forecasting skill.

Besides decadal climate variability, southwestern US water resources are also sensitive to the influence of anthropogenically-forced climate change, be it via temperature, precipitation, or atmospheric circulation changes (Lettenmaier and Gan 1990; Christensen et al. 2004; Barnett et al. 2005; Mote et al. 2005). For semi-arid and snowmelt driven basins such as the UC and URG, numerous studies have indicated that increasing temperature decreases streamflow (Christensen et al. 2004; Nowak et al. 2012; Vano et al. 2012; Woodhouse et al. 2016; Griffin and Friedman 2017; Udall and Overpeck 2017; Lehner et al. 2017). Specifically, runoff efficiency – a metric indicating the fraction of precipitation that ends up as streamflow – is more likely to be low when temperatures are above average (Nowak et al. 2012; Lehner et al. 2017). As a consequence, the relationship between winter moisture accumulation (precipitation and SWE) and summer streamflow is evidently non-stationary and can be influenced by temperature.

The influence of temperature on runoff efficiency is problematic for WSFs in light of their underlying stationarity assumptions with regard to the background climate during the forecast period. Statistical models using observed accumulated precipitation and SWE at the start of the forecast without additional temperature information for the forecast period would under-predict streamflow for cool forecast periods and over-predict streamflow for warm forecast periods, in part because they do not include the information of the secular warming trend and associated evaporation losses over the entire period.

Here we investigate (1) recent hydroclimate trends and streamflow forecast errors in the study region, the URG and parts of the UC, (2) the seasonal predictability of temperature over this region, and (3) whether including predicted temperatures in WSFs improves seasonal streamflow forecasting skill. To that end, we generate WSFs via the current operational strategy, termed ‘baseline forecast’, as well as WSFs that include seasonal temperature forecasts as a predictor, termed ‘temperature-aided forecast’. The comparison of the two approaches enables us to assess the potential to improve streamflow forecasting skill by including temperature forecasts, as well as the sufficiency of current operational temperature forecasts for this purpose. Section 2 introduces the data and methods used, Section 3 presents the results, and Section 4 discusses their wider implications.

2 Data and methods

2.1 Streamflow, precipitation, snow water equivalent, and temperature datasets

Estimates of naturalized monthly streamflow at a number of gages across the UC and URG are obtained from the NRCS; the gages are marked with circles in Fig. 1a and are listed in Table S1. For each gage and year from 1987 to 2016, the total streamflow for the respective forecasting "target period" (e.g., Apr-July cumulative flow) is calculated. Observations of water year-to-date cumulative precipitation and instantaneous SWE at the 1st of Jan, Feb, Mar, Apr, and May are extracted from the same snow telemetry monitoring (SNOTEL) stations as used in the operational forecasting by NRCS, but only if they cover the entire hindcasting period 1987-2016 (triangles in Fig. 1a; see also Table S1); this is to ensure consistency and reproducibility across the hindcasting period. The year 1987 is chosen as a start year because it offers continuous streamflow and SNOTEL measurements across all gages considered here. Monthly mean temperature is taken from the Parameter Elevation Regression on Independent Slopes Model (PRISM) data set (Daly et al. 2008) averaged over the box indicated in Fig. 1 (35.5-39.5°N, 108.5-105.0°W). Precipitation used to calculate runoff efficiency in Fig. 1b is taken from PRISM as well, summed up over the watersheds upstream of Rio Grande at Otowi Bridge, San Juan at Bluff, and Gunnison at Grand Junction.

2.2 Seasonal temperature forecasts

Seasonal temperature forecasts are derived from 8 initialized coupled climate models that produce seasonal climate forecasts (Table S2): the North American Multimodel Ensemble (NMME; Kirtman et al. (2014)), which comprises of 7 models, and the System 4 seasonal forecasting model from the European Center for Medium-Range Weather Forecast (ECMWF; Molteni et al. (2011)). In their current configuration, these models issue forecasts each month for lead times of up to 12 months with various numbers of ensemble members (10-51). Since we are interested in extracting the seasonally predictable signal, we use each model's ensemble mean (rather than all its individual ensemble members) of monthly mean 2-m temperature hindcasts issued from January 1987 to May 2016, averaged over the area indicated in Fig. 1a. We then use an equal-weights multi-model mean across the 8 models, since we found this method to perform, in terms of correlation with observed temperature, as well as or better than other weighting schemes in cross-validation across issue dates and lead times of interest (we tested a performance-weighted multi-model mean and an equal-weights mean of the overall three best models CFSv2, NASA, and ECMWF; not shown). For each streamflow forecast issue month (1st January, 1st February, etc), temperature is averaged from that issue month until the end of the main runoff period (July). Alternatives to this choice were tested, such as using spring (March-May) average temperature or the average over the next or the next two months after issue date, but were found to be inferior (not shown).

2.3 Streamflow forecasting procedure

The marginal benefit of including seasonal temperature information in WSFs can be evaluated through benchmarking the performance of enhanced WSF models against models based on the current operational forecast practice. We mimic the operational forecasting procedure of the NRCS's operational WSF by using SNOTEL data in a principal component regression (PCR) trained on 30 years (1987-2016) of observed naturalized streamflow of the respective target period (Garen 1992), hereafter 'baseline forecast'. Before use in the PCR, all predictors are standardized (subtraction of mean and division by standard deviation) and streamflow is seminormalized via a square root transformation, as is consistent with NRCS practice. The number of principal components (PCs) retained is determined through an iterative process as described in Garen (1992). Specifically, individual PCs are used in a linear regression and the significance of the regression coefficients is determined via a *t*-test; only PCs are retained that result in significant regression coefficients and that show a physically plausible relationship with streamflow (i.e., positive coefficients, indicating that high precipitation and SWE typically leads to high streamflow and vice versa). In our case, one PC is retained for all streamflow gages, consistent with Harpold et al. (2017) who also duplicated the NRCS's WSF. For each forecast issue date, forecasts are cross-validated by training the model on 29 of the 30 years and forecast the remaining (out-of-sample) year, loop through all 30 years to evaluate performance. Note that our baseline forecast likely differs slightly from the officially published NRCS forecast over the past decades, since those may also include additional but non-continuous snow course information and/or newer SNOTEL data. As discussed above, for consistency across watersheds, we only use datasets of consistent record length (1987-2016).

We then reforecast the same time period using the same information, but add the ensemble mean temperature anomaly of the 8 seasonal forecasting models as an additional predictor to the PCR (hereafter 'temperature-aided forecast'). For a given year and forecast issue date (e.g., January 1, February 1, March 1, April 1, and May 1 1987), the mean temperature prediction from the forecast issue date to the end of July is averaged over the box indicated in Fig. 1a. For all gages, the regression coefficients derived from the PCR are such that precipitation and SWE always exhibit a positive relationship with streamflow, and temperature always a negative one, indicating a physically plausible interaction of precipitation, SWE, and temperature in describing streamflow. The same rules for PC retention are applied and one PC was retained in all cases.

2.4 Skill metrics

Prediction skill for the baseline and temperature-aided streamflow forecast is calculated via a leave-one-out cross validation from 1987 to 2016. Each year between 1987 and 2016 is hindcasted with a principal component regression model that has been calibrated on the remaining 29 years of data, and the resulting time series of 30 streamflow predictions are verified against the corresponding observations.

We quantify forecast skill using the following metrics: (i) correlation, (ii) relative root mean squared error (rRMSE, in %), (iii) the Brier Skill Score (BSS) for streamflow < 33^{rd} percentile, and (iv) Continuous Ranked Probability Skill Score (CRPSS; Hersbach (2000)). Correlation and rRMSE describe how well the model predicts the variability and the absolute values, respectively, of the observed time series. The third metric provides insight into the ability of the model to predict dry conditions relevant to droughts in the US Southwest, and the fourth metric, which measures the ability of the forecast model to correctly predict the cumulative distribution function of the observed streamflow data, is used to quantify probabilistic prediction skill.

Since the skill metrics BSS and CRPSS rely on probabilistic forecasts, we derive exceedance probabilities from the standard error of the forecasts, consistent with NRCS' approach (Garen

1992). Both BSS and CRPSS are expressed as skill relative to a certain reference forecast (typically persistence or climatology). Here, we express them relative to the 'baseline forecast' to emphasize the improvement relative to the current operational approach.

3 Results

3.1 Hydroclimate trends and streamflow forecast errors

Recent hydroclimate trends in the UC and URG headwaters are illustrated by plotting the runoff efficiency as a function of temperature anomalies for streamflow gages at the outflow of the headwaters of the Gunnison, San Juan, and Rio Grande (Fig. 1b; these three gages are representative of the dynamics at other gages, see Fig. S1). A clear temperature sensitivity exists, leading to relatively reduced streamflow under positive temperature anomalies. Even in the absence of a strong precipitation trend, higher temperatures are shifting the partitioning of precipitation from snow to rain, a phenomenon that is detectable at virtually all SNOTEL stations in the region (Fig. 1c), thereby changing the peaks and timing of both snowmelt and runoff. Higher temperatures also allow for more evaporative loss between when the snow falls and when the water arrives at the streamflow gages downstream, which is a key hydrologic dynamic leading to forecast errors.

Relatively persistent forecast errors are confirmed by the forecast record in the UC and URG: streamflow gage records in these two basins show a tendency to be under-predicted in the 1980s and 1990s and over-predicted in the 2000s and 2010s (Fig. 1d and Fig. S1). While these forecast errors are in part related to unusually wet springs and summers in the 1980-90s and unusually dry springs and summers in the 2000-10s, there exists evidence that streamflow in recent years was lower than expected from precipitation deficits alone (Woodhouse et al. 2016; Lehner et al. 2017), pointing to a simultaneous influence of temperature on streamflow and thus on forecast error. This theory is further corroborated by a significant correlation of streamflow forecast error with both anomalous precipitation and temperature after the forecast issue date (Fig. S2). This relationship holds even when the natural correlation between precipitation and temperature is accounted for, a result consistent with earlier studies (Harding et al. 2012).

3.2 Temperature forecast skill

While uncertainty in multi-decadal projections of precipitation in the US Southwest remains high, climate models such as those included in the 5th phase of the Coupled Model Intercomparison Project (CMIP5) project future temperature increases (Fig. 2a) with far more certainty (van Oldenborgh et al. 2013). Similarly, dynamical seasonal climate prediction models, such as the 8 models from the NMME and ECMWF are more skillful in predicting temperature than precipitation (Becker et al. 2014; Slater et al. 2016). The ensemble mean across these 8 seasonal forecasting models captures the observed warming trend of recent decades as well as part of the interannual variability of spring-to-summer temperature over the UC and URG headwaters region at lead times of up to 5 months, showing significant correlations ranging between 0.65 and 0.75 (Fig. 2a,b). The combination of these two results leads to a usable temperature forecast skill in the context of streamflow prediction in this region. The ECMWF model is the best-performing individual model overall, although not necessarily for every lead times and not necessarily when compared to the multi-model mean across all 8 models.

3.3 Improved streamflow forecast skill

We find that augmenting the baseline forecasting approach through the use of temperature predictors adds prediction skill across the majority of streamflow gages and issue dates in the study region, which is representative of snowmelt-influenced watersheds in many parts of the western US. These benefits are illustrated through the skill difference between the baseline and temperature-aided forecasts for all skill metrics considered (Fig. 3). The median relative improvement across gages and skill metrics is between 1% and 5% with some spread across gages. The vast majority of these improvements are statistically significant in light of sampling uncertainty (see Section 3.4). However, the probabilistic skill for drought conditions (BSS) is improved less consistently than the other skill metrics. All four skill metrics indicate larger improvements for later issue dates, which likely results from a combination of better temperature forecast skill at shorter lead times and the potential for stronger temperature anomaly signals due to a shorter averaging period (e.g., May-July versus January-July).

When considering the median skill across gages within each basin, improvements tend to be larger in the Rio Grande and San Juan than in the Gunnison. The variations of forecast improvements across gages reflects the different temperature sensitivity of catchment hydrology in different locations. The sensitivity of spring runoff to temperature is affected by factors such as the basin distribution of elevation and aspect, vegetation and land cover (Male and Gray 1981), making it difficult to disentangle the reasons for an individual forecast's improvement using a statistical model only. No relationship between magnitude of skill improvement and basin elevation is found (not shown).

We also calculate the theoretical skill improvement resulting from using the actually observed temperature and found it overall to be only marginally higher than with the temperature-aided forecast based on predicted temperature (Fig. 3b,c). This indicates that the majority of the temperature information that adds skill to WSF can indeed be extracted from seasonal prediction models. Since temperature in this region over the period 1987-2016 shows a strong positive trend, the question arises how much of the added skill is attributable to the trend alone. Using the observed linear temperature trend from 1987 to 2016 as a predictor in the WSF model (thereby excluding any interannual variability that might be predictable by seasonal prediction models), we show that the trend alone adds most of the skill that originates from the seasonally forecasted temperatures (Figure 3b and 3c).

Finally, we repeat the forecasting using the temperature forecasts from the ECMWF model only, since it is the best-performing individual model (Fig. 2b), and from the 7 NMME models only (i.e., without ECMWF). Interestingly, we found the streamflow forecasting skill to be roughly equal in all three cases (Fig. S3). This suggests that temperature forecasts from ECMWF model contain about as much information, with regard to streamflow forecasting, as the 7 NMME models combined.

3.4 Robustness of forecast skill improvements

The skill is improved for the majority of the total of 100 possible forecasts (20 gages x 5 issue dates). For correlation, rRMSE, and CRPSS, 99% of forecasts are improved, with 98-100% of those significantly. For BSS, only 62% of all forecasts are improved, 95% of those significantly (see also Table S3). Significance is established through a Monte Carlo approach in which all forecasts and the associated skill score calculations are repeated 1,000 times on 30-year samples constructed from bootstrapping the original 30 years with replacement. If the 95th percentile of this distribution of skill scores shows an improvement, the skill improvement is considered significant at the 95% confidence level.

4 Discussion and conclusions

The skill improvement demonstrated here for seasonal streamflow forecasts in the Upper Rio Grande and Upper Colorado River basins can be of significant value to State and Federal water managers, which, in turn, can benefit water users throughout these basins (Carolyn Donnelly and Craig Cotton, personal communication). Despite its limited spatial extent, the study here is of relevance for other snow-melt driven basins across the US and the world, since streamflow forecast skill in such basins is often driven by the same temperature-sensitive processes.

We show that current seasonal climate prediction models are skillful in forecasting both the long-term trends and interannual variability of seasonal temperatures for this region. This temperature information adds skill to existing 'water supply forecasts' (WSFs), mitigating some of the forecast errors introduced through climate non-stationarity, and moving the WSFs closer to their maximally expected forecast skill based on relationships between observed snow, precipitation, and temperature. Additional predictability might be available once seasonal precipitation forecasts become more skillful.

For the statistical WSFs shown here, the proposed extension involves accessing and incorporating temperature predictions into existing statistical forecasting models. Conventional forecasting approaches based on hydrologic models (such as Ensemble Streamflow Prediction, or ESP, a popular operational method that is not discussed in this paper) are also commonly dependent on climate stationarity assumptions and thus are also likely to benefit from additional temperature forecast information. Fortunately, many techniques for inclusion of conditional climate information have been described in the literature over the last several decades for both statistical and model-based forecasting (e.g., Werner et al. 2004; Beckers et al. 2016; Mendoza et al. 2017; see also the special issue of Wetterhall et al. 2017), including examples of using NMME and ECMWF (Yuan et al. 2013; Mo and Lettenmaier 2014; Thober et al. 2015; Crochemore et al. 2016). It may be impossible to protect or increase streamflow prediction skill in all locations in the face of a non-stationary climate, but expanding the use of model-based seasonal climate predictions, and particularly temperature forecasts, appears to be one pragmatic strategy for hydroclimates that are similar to the US Southwest.

Despite the evidence of forecast skill improvement through inclusion of temperature, this study does not support detailed conclusions regarding the hydrologic processes that underpin changes in prediction skill, as the temperature influence on streamflow can be dampened or amplified due to other effects and non-linear interactions (e.g., related to groundwater use or vegetation alterations). Our focus on minimally impaired gages in headwater locations aims to circumvent this issue, but we cannot exclude all possibilities of processes amplifying or canceling each other. Similarly, using low-dimensional statistical models only, we are unable to disentangle why certain gages show greater improvement than others. Process-based observation and modeling studies tackling this question may therefore be a valuable next step for the hydrologic forecasting community.

Acknowledgments

We thank Pablo A. Mendoza and Carolyn Donnelly for constructive discussion, two anonymous reviewers for helpful feedback, Kevin Sampson for technical assistance, and the different modeling centers for making available the seasonal climate forecasts. NMME data is available publicly from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center. ECWMF System 4 data is not available anymore, but its successor System 5 (not used here) has been released in November 2017 and is publicly available. The National Center for Atmospheric Research is sponsored by the National Science Foundation. F.L. is supported by a Postdoc Applying Climate Expertise (PACE) fellowship cosponsored by the NOAA Climate Program Office and the Bureau of Reclamation, organized by the Cooperative Programs for the Advancement of Earth System Science. A.W.W. is supported by Reclamation under Cooperative Agreement #R11AC80816 and by the U.S. Army Corps of Engineers (USACE) Climate Preparedness and Resilience Program.

References

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier, 2005: Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, **438**, 303–309, doi:10.1038/nature04141.

Becker, E., H. Van den Dool, and Q. Zhang, 2014: Predictability and forecast skill in NMME. *J. Clim.*, **27**, 5891–5906, doi:10.1175/JCLI-D-13-00597.1.

Beckers, J. V. L., A. H. Weerts, E. Tijdeman, and E. Welles, 2016: ENSO-conditioned weather resampling method for seasonal ensemble streamflow prediction. *Hydrol. Earth Syst. Sci.*, **20**, 3277–3287, doi:10.5194/hess-20-3277-2016.

Bell, V. A., H. N. Davies, A. L. Kay, A. Brookshaw, and A. A. Scaife, 2017: A national-scale seasonal hydrological forecast system: development and evaluation over Britain. *Hydrol. Earth Syst. Sci. Discuss.*, 1–19, doi:10.5194/hess-2017-154.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Clim. Change*, **62**, 337–363, doi:10.1023/B:CLIM.0000013684.13621.1f.

Crochemore, L., M. H. Ramos, and F. Pappenberger, 2016: Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. *Hydrol. Earth Syst. Sci.*, **20**, 3601–3618, doi:10.5194/hess-20-3601-2016.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, doi:10.1002/joc.1688.

Day, G. N., 1985: Extended Streamflow Forecasting Using NWSRFS. *J. Water Resour. Plan. Manag.*, **111**, 157–170, doi:10.1061/(ASCE)0733-9496(1985)111:2(157).

Garen, D. C., 1992: Improved Techniques in Regression-Based Streamflow Volume Forecasting. *J. Water Resour. Plan. Manag.*, **118**, 654–670, doi:10.1061/(ASCE)0733-9496(1992)118:6(654).

Griffin, E. R., and J. M. Friedman, 2017: Decreased Runoff Response to Precipitation, Little Missouri River Basin, Northern Great Plains, USA. *JAWRA J. Am. Water Resour. Assoc.*, **80526**, 1–17, doi:10.1111/1752-1688.12517.

Hamlet, A. F., D. Huppert, and D. P. Lettenmaier, 2002: Economic Value of Long-Lead Streamflow Forecasts for Columbia River Hydropower. *J. Water Resour. Plan. Manag.*, **128**, 91–101, doi:10.1061/(ASCE)0733-9496(2002)128:2(91).

Harding, B. L., A. W. Wood, and J. R. Prairie, 2012: The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin. *Hydrol. Earth Syst. Sci.*, **16**, 3989–4007, doi:10.5194/hess-16-3989-2012.

Harpold, A. A., K. Sutcliffe, J. Clayton, A. Goodbody, and S. Vazquez, 2017: Does Including Soil Moisture Observations Improve Operational Streamflow Forecasts in Snow-Dominated

Watersheds? *J. Am. Water Resour. Assoc.*, **53**, 179–196, doi:10.1111/1752-1688.12490.

Hersbach, H., 2000: Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather Forecast.*, **15**, 559–570, doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Kalra, A., W. P. Miller, K. W. Lamb, S. Ahmad, and T. Piechota, 2013: Using large-scale climatic patterns for improving long lead time streamflow forecasts for Gunnison and San Juan River Basins. *Hydrol. Process.*, **27**, 1543–1559, doi:10.1002/hyp.9236.

Kirtman, B. P., and Coauthors, 2014: The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Am. Meteorol. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.

Koster, R. D., S. P. P. Mahanama, B. Livneh, D. P. Lettenmaier, and R. H. Reichle, 2010: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nat. Geosci.*, **3**, 613–616, doi:10.1038/ngeo944.

Lehner, F., E. R. Wahl, A. W. Wood, D. Blatchford, and D. Llewellyn, 2017: Assessing recent declines in Upper Rio Grande River runoff efficiency from a paleoclimate perspective. *Geophys. Res. Lett.*, doi:10.1002/2017GL073253.

Lettenmaier, D. P., and T. Y. Gan, 1990: Hydrologic sensitivities of the Sacramento-San Joaquin River Basin, California, to global warming. *Water Resour. Res.*, **26**, 69–86, doi:10.1029/WR026i001p00069.

Lins, H. F., and T. A. Cohn, 2011: Stationarity: Wanted dead or alive? *J. Am. Water Resour. Assoc.*, **47**, 475–480, doi:10.1111/j.1752-1688.2011.00542.x.

Male, D. H., and D. M. Gray, eds., 1981: *Handbook of Snow: Principles, Processes, Management and Use*. The Blackburn Press, Caldwell, New Jersey, USA, 776 pp.

Mendoza, P. A., A. W. Wood, E. Clark, E. Rothwell, M. P. Clark, B. Nijssen, L. D. Brekke, and J. R. Arnold, 2017: An intercomparison of approaches for improving predictability in operational seasonal streamflow forecasting. *Hydrol. Earth Syst. Sci. Discuss.*, 1–37, doi:10.5194/hess-2017-60.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer, 2008: Climate change. Stationarity is dead: whither water management? *Science*, **319**, 573–574, doi:10.1126/science.1151915.

Mo, K. C., and D. P. Lettenmaier, 2014: Hydrologic Prediction over the Conterminous United States Using the National Multi-Model Ensemble. *J. Hydrometeorol.*, **15**, 1457–1472, doi:10.1175/JHM-D-13-0197.1. http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-13-0197.1%5Cnhttp://journals.ametsoc.org/doi/pdf/10.1175/JHM-D-13-0197.1.

Molteni, F., and Coauthors, 2011: The new ECMWF seasonal forecast system (System 4). *Tech. Memo. ECMWF*, 49.

Moradkhani, H., K. Hsu, H. V. Gupta, and S. Sorooshian, 2004: Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *J. Hydrol.*, **295**, 246–262, doi:10.1016/j.jhydrol.2004.03.027.

Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining mountain snowpack in western north America. *Bull. Am. Meteorol. Soc.*, **86**, 39–49, doi:10.1175/BAMS-86-1-39.

Nowak, K., M. Hoerling, B. Rajagopalan, and E. Zagona, 2012: Colorado River Basin Hydroclimatic Variability. *J. Clim.*, **25**, 4389–4403, doi:10.1175/JCLI-D-11-00406.1.

van Oldenborgh, G., M. Collins, J. Arblaster, J. Christensen, J. Marotzke, S. Power, M. Rummukainen, and T. Zhou, 2013: Atlas of Global and Regional Climate Projections. *Clim. Chang. 2013 Phys. Sci. Basis. Contrib. Work. Gr. I to Fifth Assess. Rep. Intergov. Panel Clim. Chang.*, 1311–1394, doi:10.1017/CBO9781107415324.029.

Pagano, T., D. Garen, and S. Sorooshian, 2004: Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 1922–2002. *J. Hydrometeorol.*, **5**, 896–909, doi:10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2.

——, A. Wood, K. Werner, and R. Tama-Sweet, 2014a: Western U.S. water supply forecasting: A tradition evolves. *Eos (Washington. DC).*, **95**, 28–29, doi:10.1002/2014EO030007.

Pagano, T. C., and D. Garen, 2005: The recent increase in Western US streamflow variability and persistence. *J. Hydrometeorol.*, **6**, 173–179, doi:10.1175/JHM410.1.

——, and Coauthors, 2014b: Challenges of Operational River Forecasting. *J. Hydrometeorol.*, 140516115449007, doi:10.1175/JHM-D-13-0188.1. http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-13-0188.1.

Pappenberger, F., H. L. Cloke, D. J. Parker, F. Wetterhall, D. S. Richardson, and J. Thielen, 2015: The monetary benefit of early flood warnings in Europe. *Environ. Sci. Policy*, **51**, 278–291, doi:10.1016/j.envsci.2015.04.016.

Pierce, J. A. S., 2010: *A Measure of Snow: Case Studies of the Snow Survey and Water Supply Forecasting Program*. Salt Lake City, UT, 109 pp. https://www.wcc.nrcs.usda.gov/ftpref/downloads/factpub/MeasureofSnowFullReport.pdf.

Raff, D., L. Brekke, K. Werner, A. Wood, and K. White, 2013: *Short-Term Water Management Decisions: User Needs for Improved Climate, Weather, and Hydrologic Information*. 231 pp. http://www.ccawwg.us/.

Reclamation, B. of, 2016: *Climate Change Adaptation Strategy*. 33 pp. https://www.usbr.gov/climate/docs/2016ClimateStrategy.pdf.

Shukla, S., and D. P. Lettenmaier, 2011: Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci.*, **15**, 3529–3538, doi:10.5194/hess-15-3529-2011.

Slater, L. J., G. Villarini, and A. A. Bradley, 2016: Evaluation of the skill of North-American Multi-Model Ensemble (NMME) Global Climate Models in predicting average and extreme precipitation and temperature over the continental USA. *Climate Dynamics*.

Steinschneider, S., and C. Brown, 2012: Dynamic reservoir management with real-option risk hedging as a robust adaptation to nonstationary climate. *Water Resour. Res.*, **48**, 1–16, doi:10.1029/2011WR011540.

Thober, S., R. Kumar, J. Sheffield, J. Mai, D. Schäfer, and L. Samaniego, 2015: Seasonal Soil Moisture Drought Prediction over Europe Using the North American Multi-Model Ensemble (NMME). *J. Hydrometeorol.*, **16**, 2329–2344, doi:10.1175/JHM-D-15-0053.1. http://journals.ametsoc.org/doi/10.1175/JHM-D-15-0053.1.

Udall, B., and J. Overpeck, 2017: The 21 st Century Colorado River Hot Drought and Implications for the Future. *Water Resour. Res.*, doi:10.1002/2016WR019638.

Vano, J. A., T. Das, and D. P. Lettenmaier, 2012: Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature. *J. Hydrometeorol.*, **13**, 932–949, doi:10.1175/JHM-D-11-069.1.

Werner, K., D. Brandon, M. Clark, and S. Gangopadhyay, 2004: Climate Index Weighting Schemes for NWS ESP-Based Seasonal Volume Forecasts. *J. Hydrometeorol.*, **5**, 1076–1090, doi:10.1175/JHM-381.1.

Wetterhall, F., I. G. Pechlivanidis, M.-H. Ramos, A. W. Wood, Q. J. Wang, E. Zehe, and U. Ehret, 2017: Special issue: Sub-seasonal to seasonal hydrological forecasting. *Hydrol. Earth Syst. Sci.*,.

Wood, A. W., and D. P. Lettenmaier, 2006: A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bull. Am. Meteorol. Soc.*, **87**, 1699–1712, doi:10.1175/BAMS-87-12-1699.

——, and ——, 2008: An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.*, **35**, doi:10.1029/2008GL034648.

——, A. Kumar, and D. P. Lettenmaier, 2005: A retrospective assessment of National Centers for Environmental prediction climate model-based ensemble hydrologic forecasting in the western United States. *J. Geophys. Res. D Atmos.*, **110**, 1–16, doi:10.1029/2004JD004508.

Woodhouse, C. A., G. T. Pederson, K. Morino, S. A. McAfee, and G. J. McCabe, 2016: Increasing influence of air temperature on upper Colorado River streamflow. *Geophys. Res. Lett.*, **43**, 2174–2181, doi:10.1002/2015GL067613.

Yuan, X., E. F. Wood, J. K. Roundy, and M. Pan, 2013: CFSv2-Based seasonal hydroclimatic forecasts over the conterminous United States. *J. Clim.*, **26**, 4828–4847, doi:10.1175/JCLI-D-12-00683.1.

Figures

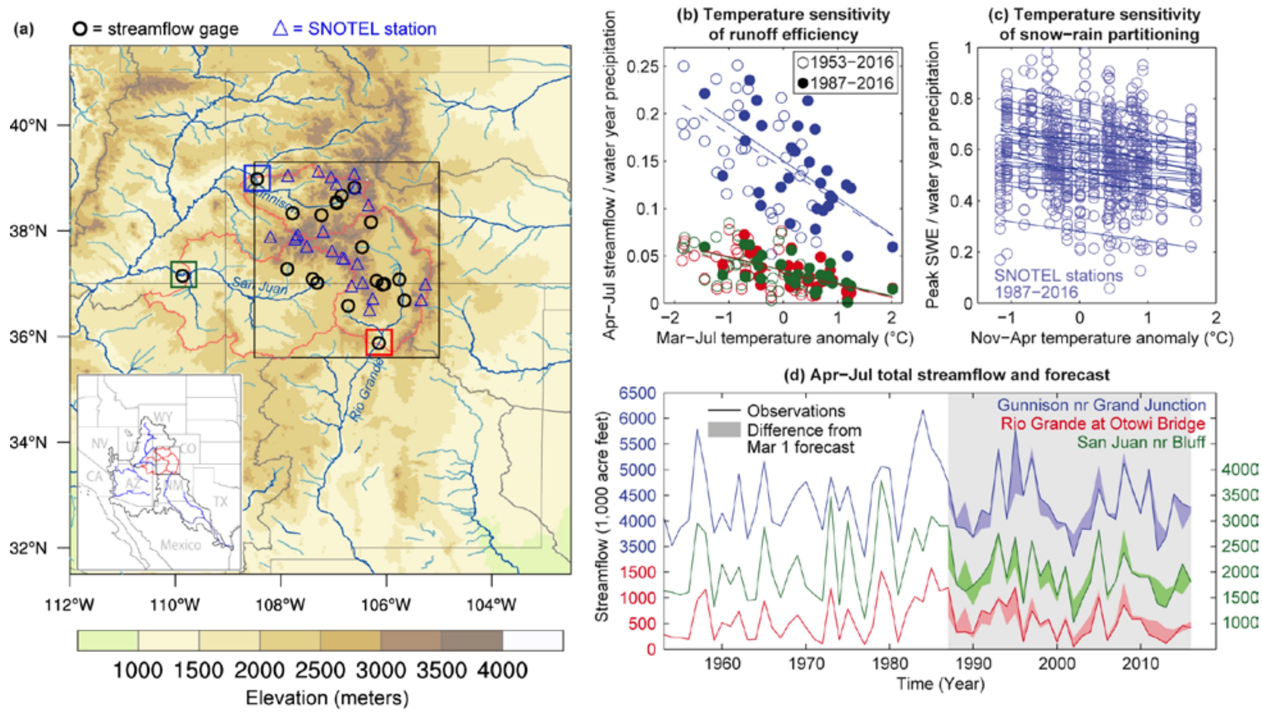


Figure 1: (a) Map showing the main rivers, basins, (circles) streamflow gages, and (triangles) SNOTEL stations analyzed in this study. (b) Runoff efficiency – spring-summer streamflow divided by water year precipitation – for 3 selected gages marked with colored boxes in (a). (c) Snow-rain partitioning – peak snow water equivalent (SWE) divided by water year precipitation – as a function of winter-spring temperature for all SNOTEL stations analyzed in this study (each linear trend line is for one SNOTEL station). (d) Observed and forecasted streamflow for the 3 selected gages; solid lines are the observed streamflow, while colored shading indicates the difference between the observed and forecasted streamflow, i.e., the larger the shading the larger the forecast error; gray shading indicates time period analyzed in this study. See text for more details on datasets.

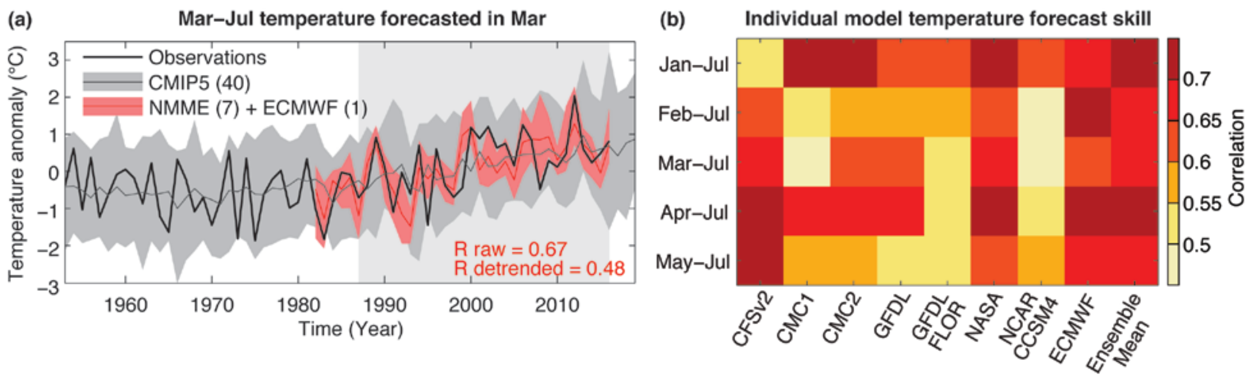


Figure 2: (a) March-July mean temperature anomalies relative to 1982-2016 from observations, 40 CMIP5 models, and seasonal prediction models (NMME+ECMWF), averaged over the box indicated in Fig. 1a. The red line is the mean across NMME-ECMWF models, the gray line is the mean across CMIP5 models, and the black line is observations. Shading indicates the 5-95% range. (b) Correlation between observed and forecasted temperature for different temperature targets and seasonal prediction models for 1982-2016. Forecasts are initialized at the start of each predicted period. All correlations are significant at 95% confidence.

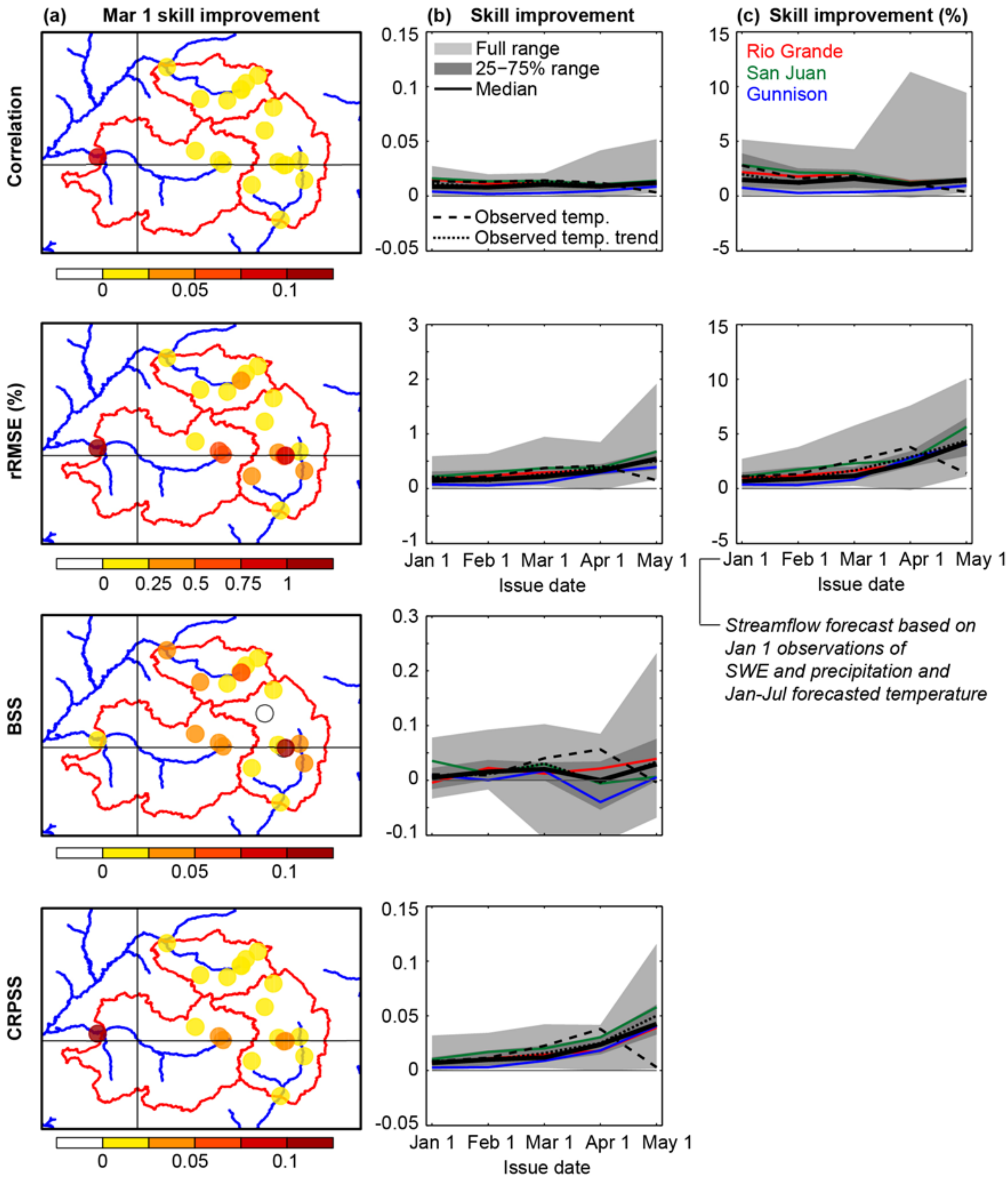


Figure 3: (a) Absolute skill improvement of the temperature-aided forecast relative to the baseline forecast at individual gages for issue date 1st March as an illustrative example. (b) Absolute skill improvement for all gages as a function of issue date. (c) Relative skill improvement for all gages as a function of issue date. Solid lines are the median across (black) all gages and (colors) the three basins. Dashed line is the median across all gages when using

observed temperature, mimicking the hypothetical case where the future temperature is known at the time of forecast issue, and dotted line is the median when using only the linear trend of observed temperature.