Supplementary Material

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)

Content

1 Overlap of tree ring chronologies

While there exists overlap between some of the original tree ring chronologies used to develop the reconstruction for Otowi Bridge streamflow (Upper Rio Grande River basin, or URG) and URG precipitation, we note that this overlap is very low (Table S1), and there is no overlap with the URG temperature reconstruction. Specifically, out of 13 streamflow reconstructions used as predictors in our URG precipitation reconstruction (cf. Diaz and Wahl, 2015), only 4 share chronologies with the Otowi Bridge streamflow reconstruction (http://www.treeflow.info/content/rio-grande-owoti-new-mexico-update, cf. Woodhouse et al. 2012); and out of these the Otowi Bridge streamflow reconstruction shares just one chronology with three of the URG precipitation predictors, and two chronologies with one of them (Table S1). Critically, for the URG precipitation reconstruction, the Otowi Bridge streamflow reconstruction itself was not used, unlike in Diaz and Wahl (2015). We note further that not only do these few overlapping chronologies enter into the reconstruction algorithm as part of a much larger group of chronologies encompassed by the full set of employed streamflow predictors, the principal components of these streamflow reconstructions were then filtered and used as the actual predictors in the spatial regression algorithm. This will likely further dilute any influence of the common chronologies identified here between the Otowi Bridge streamflow reconstruction and the URG precipitation reconstruction.

Table S1. Tree ring chronologies used in the Otowi Bridge streamflow reconstruction (cf. Woodhouse et al. 2012) and their overlap with the chronologies used in the Upper Rio Grande River basin (URG) temperature reconstruction (Wahl and Smerdon 2012; hereafter WS12), and the chronologies used in the streamflow reconstructions that were employed as predictor data in the URG precipitation reconstruction. Chronologies that are shared between two reconstructions are marked with '1', all others are marked with '0'.

		Temperature Reconstruction	**Streamflow reconstructions (of a total of 13) employed in precipitation reconstruction**			
Chronologies used in **Otowi Bridge reconstruction** of Upper Rio Grande River (with ITRDB #)		WS12 (94 total predictors)	San Juan River near Bluff, UT	Colorado River near Cisco, UT	Upper Colorado River at Lee's Ferry, CO	Clear Creek near Golden, CO
Gould Reservoir, CO	CO605	0	0	0	0	0
Slickrock, CO	CO626	0	1	0	0	0
Unaweep, CO	CO629	0	0	0	0	0
Wild Rose, CO	CO631	0	1	1	1	0
Chokechery Canyon, CO	CO637	0	0	0	0	1
Mesa Alta – pine, NM	NM584	0	0	0	0	0
Altamont Triangle (CO)	NA	0	0	0	0	0
Total of 7 chronologies used in Otowi Bridge reconstruction that are also used in other reconstruction		0	2	1	1	1

2 Validation of reconstructions

While no entirely new reconstruction has been used in the study here, we still briefly investigate how the reconstructions compare to the observational datasets during their common period. Note, that while the streamflow reconstruction (Woodhouse et al. 2012) has been calibrated against the same observational data we use, the reconstructions of temperature (Wahl and Smerdon 2012) and precipitation (Diaz and Wahl 2015) were calibrated against observational data sets other than the ones used in this study.

Table S2 shows high r^2 (all $p < 0.001$) and low relative root mean square errors of 11%-30% for the different reconstructions. The runoff ratio reconstruction performs least well in terms of r^2, although it still resolves nearly 50% of the variance of its target. Since the precipitation reconstruction is different from that reported by Diaz and Wahl (2015) due to the removal of the Otowi Bridge reconstruction as a predictor, noted above, we also provide validation statistics for the individual grid cells we used in Table S3 (in that case comparing the reconstructions to the original instrumental data set employed in the reconstruction process).

Table S2. Validation statistics of the individual reconstructions used in the study. Relative root mean square error (rRMSE) is simply RMSE relative to the mean of the observations during the common period. Temperature was reconstructed as anomalies, which prohibits the calculation of rRMSE. P-values are calculated for r, taking into account autocorrelation.

Variable	**Common period**	**r^2**	**p**	**RMSE**	**rRMSE**
Streamflow	1943-2012	0.64	< 0.001	268 KAF	29%
Precipitation	1943-1977	0.61	< 0.001	1,560 KAF	11%
Runoff ratio	1943-1977	0.42	< 0.001	0.017	30%
Temperature	1895-1980	0.49	< 0.001	0.44 °C	*n.a.*

Table S3. Coefficient of Efficiency (CE) statistics in validation for the modified precipitation reconstruction by *Diaz and Wahl (2015)*. Bold values are the grid boxes used to construct the precipitation reconstruction for the Upper Rio Grande basin.

Lat/Lon	-107.25	-106.75	-106.25	-105.75	-105.25	-104.75
35.75	0.57089	0.5434207	0.4002189	0.4539471	0.1674572	0.01598923
36.25	0.5384434	**0.5395357**	**0.4570064**	**0.4461445**	**0.1561925**	0.01186647
36.75	0.549256	**0.5259534**	**0.5070044**	**0.4825245**	**0.35267**	0.2204372
37.25	0.5535897	**0.5206735**	**0.5180529**	**0.4250602**	**0.4220072**	0.4122258
37.75	**0.4810739**	**0.5191891**	**0.3894573**	**0.2590211**	0.226543	0.362338
38.25	0.3921034	**0.4134108**	**0.3878251**	0.3749382	0.228881	0.3603675

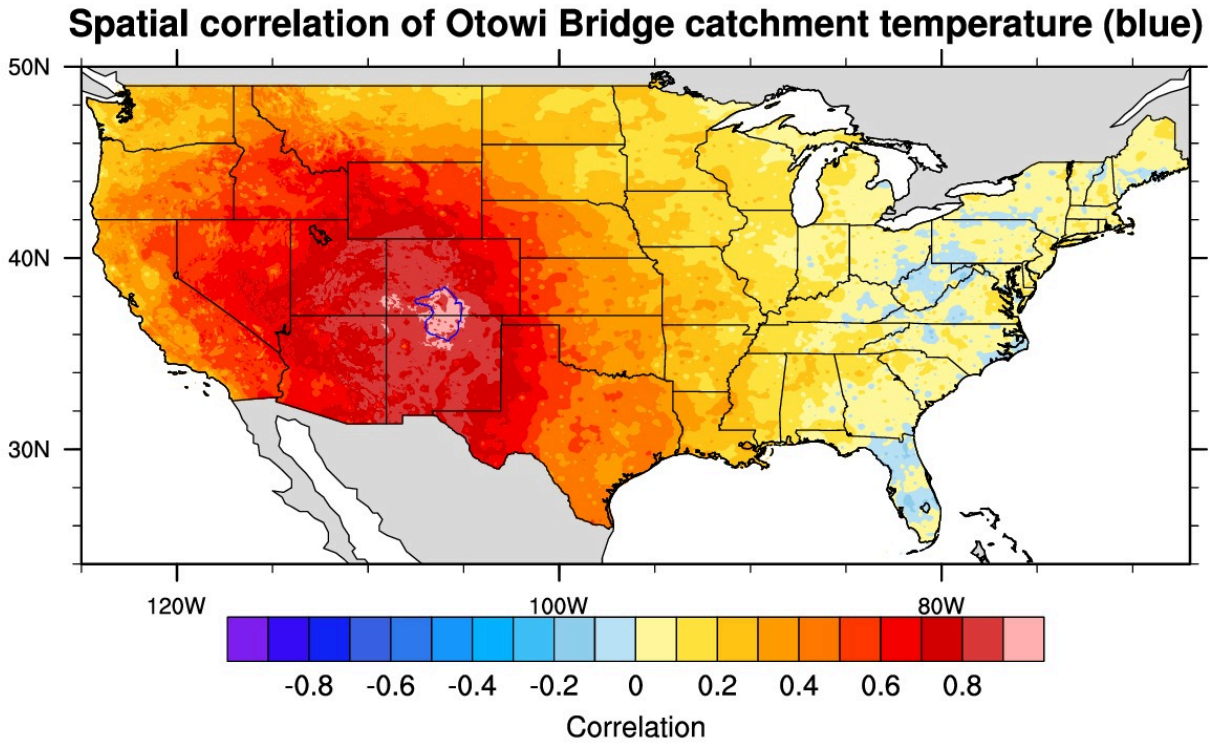

Figure S 1. Correlation of annual mean temperature averaged over the Otowi Bridge catchment (blue contours) with annual mean temperature over the contiguous US, based on PRISM data from 1895-2015. See Wahl and Smerdon (2012), Fig. S6, for comparable grid coverage of the annual temperature reconstruction.

3 Reconstruction uncertainties

The streamflow reconstruction for Otowi Bridge did not report time-variant uncertainties (http://www.treeflow.info/content/rio-grande-owoti-new-mexico-update, cf. Woodhouse et al. 2012). Instead, we generate an ensemble of streamflow reconstructions, using a Monte Carlo framework informed by the residual between the reconstruction and the observation according to the following procedure:

1. Transform observation and reconstruction into a more Gaussian space by taking the square-root.
2. Restore the reconstruction's variance to match the variance of the observed record during the period of overlap (1943-2012); this guards against potentially damped variance in the reconstruction, since the standard deviation of the reconstruction was found to be about 15% lower than in observations during the period of overlap.
3. Calculate the residual between the variance-adjusted reconstruction and observations during the period of overlap (hereafter "residual"). No significant autocorrelation was found in the residual, and the residual is effectively unbiased (mean ~ 0).
4. Model 1,000 time series of the residual the same length as the reconstruction with the same auto-correlation as the residual, using Hosking's algorithm (Hosking 1984).
5. Add one of the modeled residual time series to the variance-adjusted reconstruction. This represents one alternative pseudo-reconstruction, constrained by the reconstruction uncertainty in relation to the instrumental data. Repeat 1,000 times to generate an ensemble. We note that this procedure is based on an assumption of unbiased residuals, which is the case here (see point 3).
6. Transform the time series back into the original space.

Note that this is a conservative uncertainty estimation, since each of the newly generated streamflow reconstructions is not constrained by recalibrating to observations. Strictly speaking, this is not an ensemble of reconstructions, but one single reconstruction with uncertainties. This distinction is important, since the approach laid out here typically produces a larger uncertainty range than in the case where one actually reconstructs (including calibration) the time series multiple times to generate a true ensemble of reconstructions.

114

115 To be consistent, we calculate the uncertainty for the precipitation reconstruction in the identical
116 way as for the streamflow reconstruction. As expected, the resulting uncertainty range is
117 somewhat larger than the originally reported uncertainty range (Diaz and Wahl 2015). For the
118 reconstruction of runoff ratio, we then simply calculate runoff ratio for the 1,000x1,000
119 combinations of streamflow and precipitation. For the reconstruction of temperature we used the
120 reported uncertainties, which are based on a 1,000 member probabilistic ensemble, providing
121 estimated Monte Carlo distributions of reconstruction skill (Wahl and Smerdon 2012).

122

123

4 Probability of recent trends being unprecedented

Using the ensembles of reconstructions for the different variables (see Section 3), we estimate the probability that the most recent 30-year trend (1986-2015) is unprecedented in context of the past 440+ years. Specifically, we calculate all possible negative 30-year trends across the ensembles and rank them. We then determine the percentile of those 30-year trends that corresponds to the observed 30-year trend. Table S4 shows these probabilities, where a value of, e.g., 95% indicates that 95% of all negative 30-year trends across the ensembles are weaker than the observed trend.

Table S4. Observed 1986-2015 trends in water year streamflow, precipitation, and runoff ratio, as well as the probability of those trends being the strongest in context of the paleoclimate record (1571-2015; 445 years).

	1986-2015 trend (30 years^{-1})	Probability of being the strongest trend since 1571
Streamflow	–925.4 KAF	99.1%
Precipitation	–5,066.3 KAF	97.9%
Runoff ratio	–0.037	97.8%

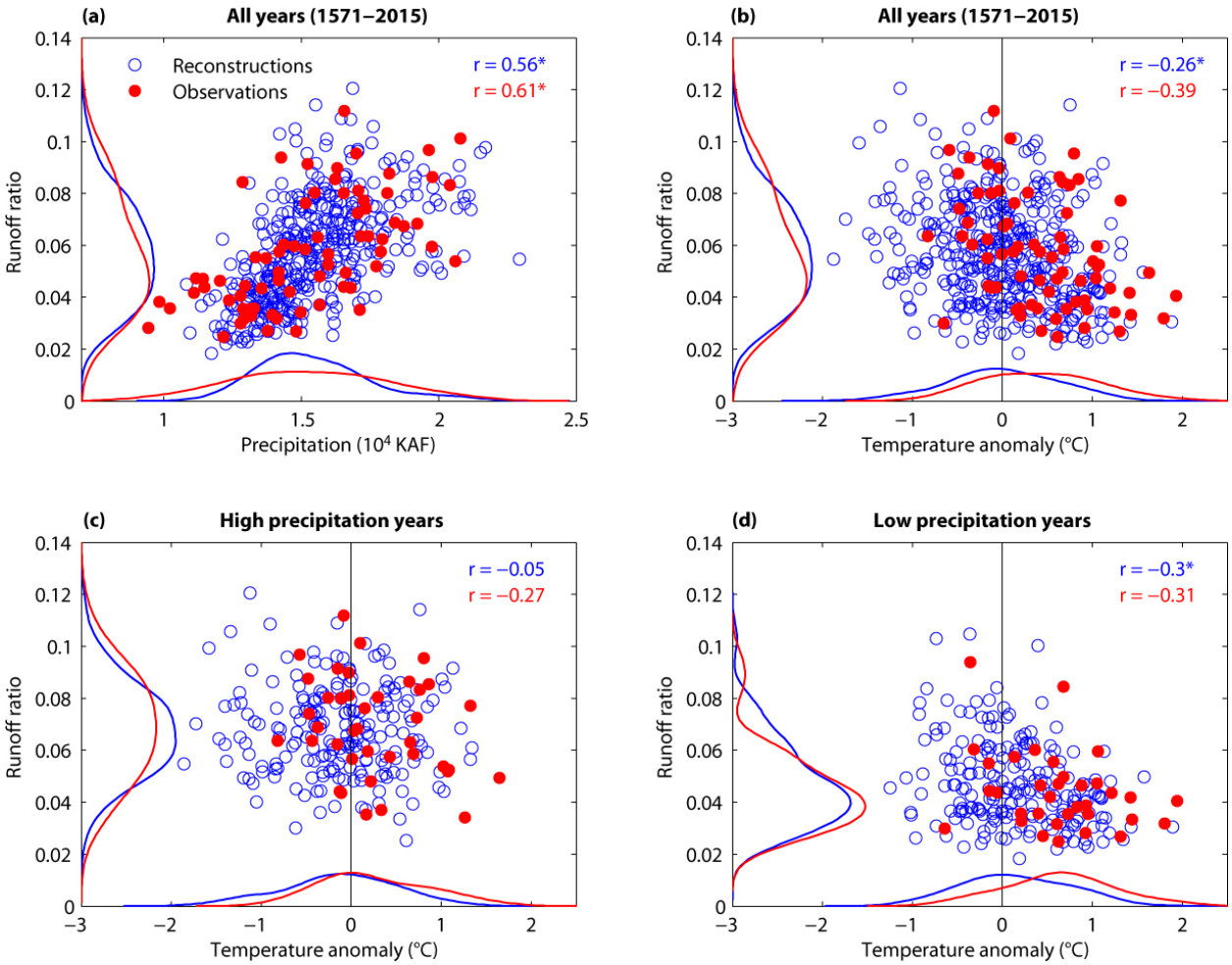

Figure S 2. Scatter plots of **(a)** runoff ratio versus precipitation, and **(b-d)** runoff ratio versus temperature anomaly (relative to median of 1571-1977) from reconstructions and observations. Panels **(c-d)** show the same data as panel **(b)**, but split up in years in which precipitation was **(c)** above and **(d)** below the median of the reconstruction. Correlations are given in the top right corner, asterisks indicate significance with 95% confidence taking autocorrelation into account. Normalized probability density functions are given on the respective axis; the amplitude scaling is arbitrary and not related to units of the magnitude axis it is plotted on.

5 Sensitivity to temperature reconstruction

Figure S3 shows the same information as Figure 2a, but comparing with two other temperature reconstructions: Figure S3a uses the same annual mean temperature as in the main text (Wahl and Smerdon 2012), meaning this panel is identical to Figure 2a, except that the time period is different to facilitate comparison with the other temperature reconstruction (1600-1977, rather than 1571-1977). Figure S3b uses April-September temperature, extracted at 35°N, 110°W from a 5°x5° reconstruction (Briffa et al. 1992). Figure S3c uses annual maximum monthly mean temperature for the Southern Colorado Plateau (Salzer and Kipfmueller 2005).

The general relationship identified in Figure 2 and the main text holds for Figure S3 as well, namely that very low runoff ratios (<10th percentile) are more likely during dry and warm conditions than during dry and cold conditions. The main difference between Figure S3a and Figures S3b and S3c is the suggested frequency of low runoff ratios (<30th percentile) during dry years: Figure S3a shows 30% in cold years and 53% in warm years, Figure S3b shows 43% in cold summers and 40% in warm summers, and Figure S3c shows 38% in years with a cold annual maximum monthly temperature and 45% in years with a warm annual maximum monthly temperature.

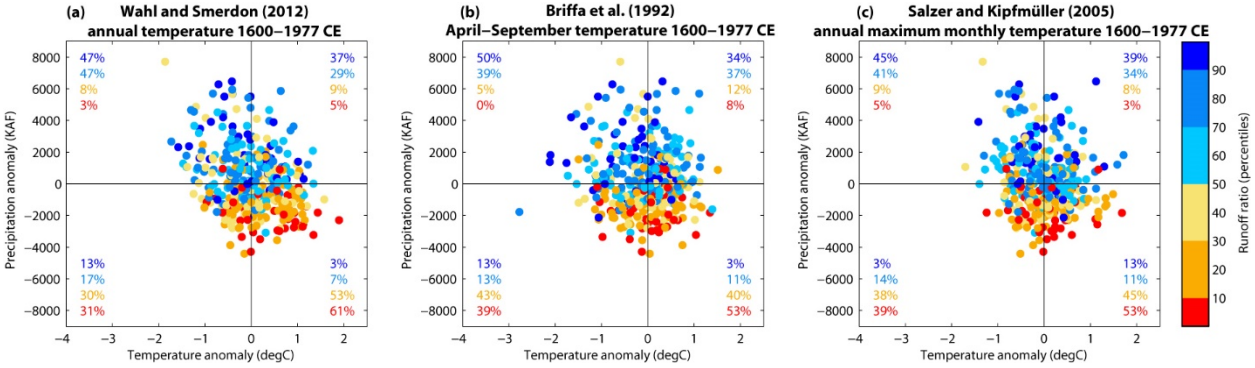

Figure S 3. Runoff ratio at Otowi Bridge (colors) as a function of water year precipitation and (a) annual mean temperature from Wahl and Smerdon (2012), (b) April-September temperature from Briffa et al. (1992), and (c) annual maximum monthly mean temperature from Salzer and Kipfmueller (2005).

6 Accounting for correlation between precipitation and temperature

The correlation between precipitation and temperature on interannual basis is of concern as some of the reported higher probabilities for very low runoff ratio years might arise simply from the fact that warm years tend to coincide with dry years, so that there are overall simply more warm and dry years than cold and dry years. This co-variation of precipitation and other variables has also been discussed in context of streamflow elasticity (Andreassian et al. 2016). To account for this empirically, we provide an alternative version of Figure 2 from the main text (Figure S4): the data are identical, but the percentages shown now correspond to the fractions of very low/low/high/very high runoff ratio years of total years *within* each quadrant. It can be see that very low runoff ratio years are 1.8 (reconstruction), 1.4 (observations), and 2.1 (CESM) times more likely during warm and dry years than during cold and dry years.

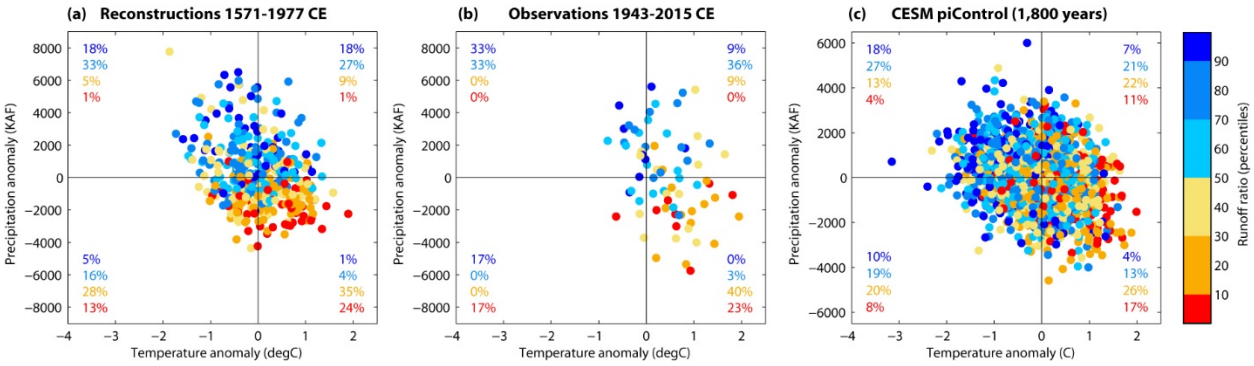

Figure S 4. Same as Figure 2 of the main text, except the percentage numbers are relative to the total number of years within each quadrant.

7 References

Andreassian, V., L. Coron, J. Lerat, and N. Le Moine, 2016: Climate elasticity of streamflow revisited - An elasticity index based on long-term hydrometeorological records. *Hydrol. Earth Syst. Sci.*, **20**, 4503–4524, doi:10.5194/hess-20-4503-2016.

Briffa, K. R., P. D. Jones, and F. H. Schweingruber, 1992: Tree-Ring Density Reconstructions of Summer Temperature Patterns across Western North America since 1600. *J. Clim.*, **5**, 735–754, doi:10.1175/1520-0442(1992)005<0735:TRDROS>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.

Hosking, J. R. M., 1984: Modeling persistence in hydrological time series using fractional differencing. *Water Resour. Res.*, **20**, 1898–1908, doi:10.1029/WR020i012p01898.

Salzer, M. W., and K. F. Kipfmueller, 2005: Reconstructed temperature and precipitation on a millennial timescale from tree-rings in the southern Colorado Plateau, U.S.A. *Clim. Change*, **70**, 465–487, doi:10.1007/s10584-005-5922-3. http://link.springer.com/10.1007/s10584-005-5922-3 (Accessed September 3, 2014).

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., 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.