4-26-2020

Increased drought severity tracks warming in the United States' largest river basin

Justin T. Martin

Gregory T. Pederson

Connie A. Woodhouse

Edward R. Cook

Gregory J. McCabe

See next page for additional authors

Authors

Justin T. Martin, Gregory T. Pederson, Connie A. Woodhouse, Edward R. Cook, Gregory J. McCabe, Kevin J. Anchukaitis, Erika K. Wise, Patrick J. Erger, Larry Dolan, Marketa McGuire, Subhrendu Gangopadhyay, Katherine J. Chase, Jeremy S. Litell, Stephen T. Gray, Scott St. George, Jonathan M. Friedman, David J. Sauchyn, Jeannine-Marie St-Jacques, and John King

Increased drought severity tracks warming in the United States' largest river basin

Justin T. Martin[a,1], Gregory T. Pederson[a], Connie A. Woodhouse[b,c], Edward R. Cook[d], Gregory J. McCabe[e], Kevin J. Anchukaitis[b,c], Erika K. Wise[f], Patrick J. Erger[g], Larry Dolan[h,2], Marketa McGuire[i], Subhrendu Gangopadhyay[i], Katherine J. Chase[j], Jeremy S. Littell[k], Stephen T. Gray[k], Scott St. George[l], Jonathan M. Friedman[m], David J. Sauchyn[n], Jeannine-Marie St-Jacques[o], and John King[p]

[a]Northern Rocky Mountain Science Center, US Geological Survey, Bozeman, MT 59717; [b]School of Geography and Development, University of Arizona, Tucson, AZ 85721; [c]Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ 85721; [d]Lamont-Doherty Earth Observatory, Palisades, NY 10964; [e]Integrated Modeling and Prediction Division, Water Mission Area, US Geological Survey, Denver, CO 80225; [f]Department of Geography, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599; [g]Missouri Basin Region, US Bureau of Reclamation, Billings, MT 59107; [h]Montana Department of Natural Resources and Conservation, Helena, MT 59601; [i]Technical Service Center, US Bureau of Reclamation, Denver, CO 80225; [j]Wyoming-Montana Water Science Center, US Geological Survey, Helena, MT 59601; [k]Alaska Climate Adaptation Science Center, US Geological Survey, Anchorage, AK 99503; [l]Department of Geography, Environment and Society, University of Minnesota, Minneapolis, MN 55455; [m]Fort Collins Science Center, US Geological Survey, Ft. Collins, CO 80526; [n]Prairie Adaptation Research Collaborative, University of Regina, Regina, SK S4S 0A2, Canada; [o]Department of Geography, Planning and Environment, Concordia University, Montreal, QC H3G 1M8, Canada; and [p]Lone Pine Research, Bozeman, MT 59715

Edited by Cathy Whitlock, Montana State University, Bozeman, MT, and approved March 28, 2020 (received for review September 17, 2019)

Across the Upper Missouri River Basin, the recent drought of 2000 to 2010, known as the "turn-of-the-century drought," was likely more severe than any in the instrumental record including the Dust Bowl drought. However, until now, adequate proxy records needed to better understand this event with regard to long-term variability have been lacking. Here we examine 1,200 y of streamflow from a network of 17 new tree-ring–based reconstructions for gages across the upper Missouri basin and an independent reconstruction of warm-season regional temperature in order to place the recent drought in a long-term climate context. We find that temperature has increasingly influenced the severity of drought events by decreasing runoff efficiency in the basin since the late 20th century (1980s) onward. The occurrence of extreme heat, higher evapotranspiration, and associated low-flow conditions across the basin has increased substantially over the 20th and 21st centuries, and recent warming aligns with increasing drought severities that rival or exceed any estimated over the last 12 centuries. Future warming is anticipated to cause increasingly severe droughts by enhancing water deficits that could prove challenging for water management.

drought severity | streamflow | temperature | precipitation | water resources

In much of the western United States (hereafter "the West"), water demand (i.e., the combination of atmospheric demands, ecological requirements, and consumptive use) is approaching or has exceeded supply, making the threat of future drought an increasing concern for water managers (1–5). Prolonged drought can disrupt agricultural systems and economies (6–9), challenge river system control and navigation (10, 11), and complicate management of sensitive ecological resources (12, 13). Recently, ample evidence has emerged to suggest that the severity of several regional 21st-century droughts has exceeded the severity of historical drought events; these recent extreme droughts include the 2011 to 2016 California drought (14, 15) and the 2000 to 2015 (16, 17) drought in the Colorado River basin.

Conspicuously absent thus far from investigations of recent droughts has been the Missouri River, the longest river in North America draining the largest independent river basin in the United States (18). Similar to California (14) and the Upper Colorado River Basin (16, 17), parts of the early 21st century have been remarkably dry across the Upper Missouri River Basin (UMRB) (19). In fact, our assessment of streamflow for the UMRB suggests that the widespread drought period of 2000 to 2010, termed the "turn-of-the-century drought" by Cook et al. (19), was a period of observationally unprecedented and sustained hydrologic drought likely surpassing even the drought of the Dust Bowl period.

Northern Hemisphere summer temperatures are now likely higher than they have been in the last 1,200 y (20), and the unique combination of recent anomalously high temperatures (20) and severe droughts across much of the West (14, 16, 17) has led numerous researchers to revisit the role of temperature in changing the timing and efficiency of runoff in the new millennium (16, 21–24). Evidence suggests that across much of the West atmospheric moisture demands due to warming are reducing

Significance

Recent decades have seen droughts across multiple US river basins that are unprecedented over the last century and potentially longer. Understanding the drivers of drought in a long-term context requires extending instrumental data with paleoclimatic data. Here, a network of new millennial-length streamflow reconstructions and a regional temperature reconstruction from tree rings place 20th and early 21st century drought severity in the Upper Missouri River basin into a long-term context. Across the headwaters of the United States' largest river basin, we estimated region-wide, decadal-scale drought severity during the "turn-of-the-century drought" *ca.* 2000 to 2010 was potentially unprecedented over the last millennium. Warming temperatures have likely increasingly influenced streamflow by decreasing runoff efficiency since at least the late 20th century.

Author contributions: J.T.M., G.T.P., C.A.W., E.R.C., and E.K.W. designed research; J.T.M., G.T.P., G.J.M., K.J.A., and M.M. performed research; J.T.M., G.T.P., G.J.M., and K.J.A. analyzed data; J.T.M., G.T.P., C.A.W., G.J.M., K.J.A., E.K.W., S.G., J.S.L., S.T.G., S.S.G., J.M.F., D.J.S., and J.-M.S.-J. wrote the paper; and P.J.E., L.D., S.G., K.J.C., J.S.L., S.T.G., J.M.F., D.J.S., J.-M.S.-J., and J.K. contributed data.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: The Upper Missouri Basin runoff-season temperature reconstruction is available from the National Oceanic and Atmospheric Administration National Centers of Environmental Information (https://www.ncdc.noaa.gov/paleo/study/29413). The tree-ring chronologies used in the temperature reconstruction are available online from the PAGES 2K version 2 consortium (https://doi.org/10.1038/sdata.2017.88).

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: justinmartin@usgs.gov.

[2]Retired.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1916208117/-/DCSupplemental.

First published May 11, 2020.

the effectiveness of precipitation in generating streamflow and ultimately surface-water supplies (16, 22–29).

The waters of the Upper Missouri River originate predominantly in the Rocky Mountains of Montana, Wyoming, and Colorado, where high-elevation catchments capture and store large volumes of water as winter snowpack that are later released as spring and early summer snowmelt (9). This mountain water is an important component of the total annual flow of the Missouri, accounting for roughly 30% of the annual discharge delivered to the Mississippi River on average, but ranging between 14% to more than 50% from year to year, most of which is delivered during the critical warm-season months (May through September) (9, 30). Across much of the UMRB, cool-season (October through May) precipitation stored as winter snowpack has historically been the primary driver of streamflow, with observed April 1 snow-water equivalent (SWE) usually accounting for at least half of the variability in observed streamflow from the primary headwaters regions (9). However, since the 1950s, warming spring temperatures have increasingly driven regional snowpack declines that have intensified since the 1980s (31–33). By 2006, these declines amounted to a low snowpack anomaly of unusual severity relative to the last 800 y and spanned the snow-dominated watersheds of the interior West (32). A recent reassessment of snowpack declines across the West by Mote et al. (33) suggests continued temperature-driven snowpack declines through 2016 totaling a volumetric storage loss of between 25 and 50 km^3, which is comparable to the storage capacity of Lake Mead (~36 km^3), the United States' largest reservoir.

Here we examine the extended record (*ca*. 800 to 2010 CE) of streamflow and the influence of temperature on drought through the Medieval Climate Anomaly, with a focus on the recent turn-of-the-century drought in the UMRB. The role of increasing temperature on streamflow and basin-wide drought is examined in the UMRB over the last 1,200 y by analyzing a basin-wide composite streamflow record developed from a network of 17 tree-ring–based reconstructions of streamflow for major gages in the UMRB (Fig. 1) (34) and an independent runoff-season (March through August) regional temperature reconstruction. We also explore the hydrologic implications (e.g., drought severity and spatial extent) and climatic drivers (temperature and precipitation) of the observed changes in streamflow across the UMRB and characterize shifts in the likelihood of extreme flow levels and reductions in runoff efficiency across the basin.

Results and Discussion

The Turn-of-the-Century Drought in a Long-Term Context. Persistent streamflow deficits during the turn-of-the-century drought were greater than those observed at any other time since widespread gaging of streamflow began across the UMRB in the early 20th century (Figs. 2 *A* and *C* and 3*A*). However, a more robust understanding of how the turn-of-the-century drought compares to past droughts in the UMRB requires the multicentury perspective provided by paleoclimate and paleohydrologic data. Such datasets provide numerous historical events for comparison and have documented the occurrence of very severe drought events in the American Southwest (35–38), California (39), and the Southern Great Plains (40, 41) during the last millennium. In these other regions, several drought events of the last millennium, often identified as "megadroughts" (19, 39, 40), are unrivaled in recent times in terms of severity and/or duration. Until recently (34), no comparable proxy records of hydrologic drought existed for the UMRB.

Using a 1,200-y, basin-wide Upper Missouri River streamflow reconstruction (34), we place the severity and duration of the turn-of-the-century drought in the context of long-term hydroclimatic variability (Fig. 2). The reconstruction skillfully captures the observed variation of streamflow across the basin and was specifically designed for the assessment of drought conditions over time (34). Combining this record (800 to 1929) with the naturalized flow records (1930 to 2010) from across the UMRB, we developed a representative estimate of basin-wide average streamflow spanning 800 to 2010 CE (Fig. 2*A* and *SI Appendix*, section S2 and Figs. S3–S6). We then developed a record of drought deficits focused on decadal-scale variability from the basin-wide flow estimate by first defining drought events as any sequence of two or more years in which the 10-y cubic smoothing spline (42) of reconstructed or observed streamflow anomalies was negative. Drought severity was then quantified as the magnitude of flow deficits over the period of each drought event determined by the value of the spline itself in units of SDs (*z*-scores) (Fig. 2*C*) (*SI Appendix*, section S2 and Figs. S3–S6). The duration of each drought event was determined as the number of years the smoothed streamflow anomaly remained negative.

In terms of the most severe flow deficits, the driest years of the turn-of-the-century drought in the UMRB appear unmatched over the last 1,200 y (Fig. 2*C*). Only a single event in the late 13th century rivaled the greatest deficits of this most recent event; however, the lowest point in the spline of streamflow during the

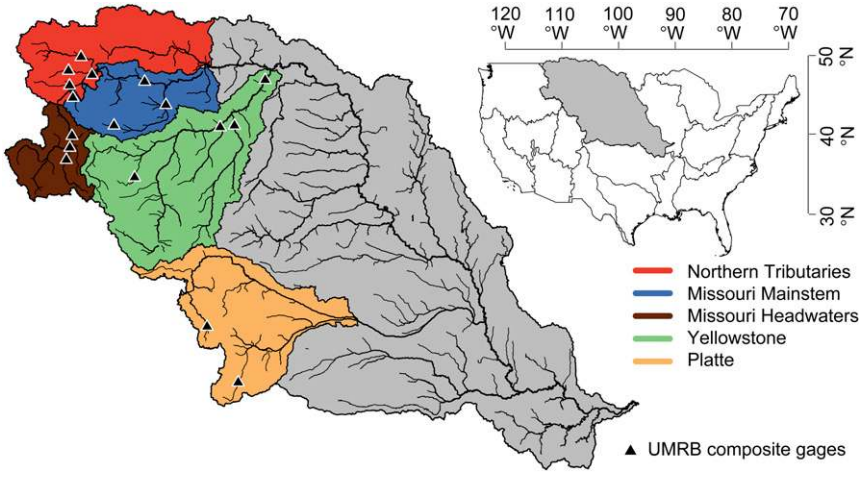

Fig. 1. The Missouri River Basin and its subregions. The location of the Missouri River Basin within the continental United States (gray watershed, upper right) and the location of the five hydrologically distinct subregions (colored watersheds) that define the UMRB. Reconstructed gages used to develop the estimate of basin-wide mean annual streamflow are shown as triangles.

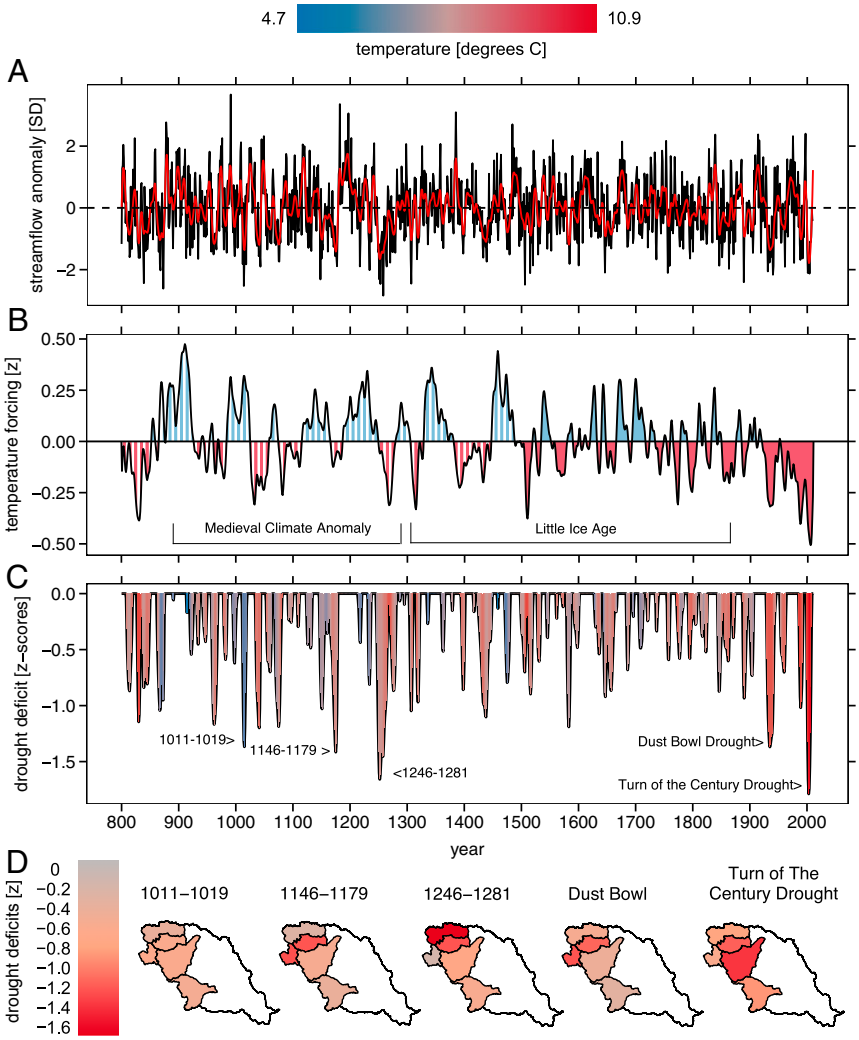

Fig. 2. Streamflow, temperature forcing of streamflow, drought severity, and spatial distributions of the five most severe droughts in the UMRB over the last 1,200 y. The time series of reconstructed streamflow (*A*), the 10-y spline of temperature forcing of streamflow (*B*), and decadal-scale drought deficits (*C*) for the UMRB. The red line in *A* shows the 10-y cubic smoothing spline of streamflow and the dashed horizontal line shows the long-term mean. Color in *B* denotes positive (blue) and negative (red) forcings. White hatching in *B* denotes the period of high uncertainty in the UMRB regional temperature reconstruction. Color in *C* denotes temperature in the basin over the periods of the various drought events. *D* shows the spatial distributions of maximum flow deficits during the five major droughts annotated in *C*; 1930 to 2010 are instrumental data and 800 to 1929 are reconstructed.

13th century drought was 0.13 SDs (*s*) higher than that of the turn-of-the-century drought. The robustness of this finding was tested in multiple ways including by using various spline lengths from 5 to 15 y to quantify drought deficits and by comparing the intensity [cumulative deficit/duration (43)] of drought deficits determined from the unsmoothed streamflow record over time. In all cases except for the highest degree of smoothing (>12 y), the deficits during the driest years of the turn-of-the-century drought exceed those of all earlier droughts in the record (*SI Appendix*, section S2 and Figs. S3–S6).

In terms of duration, however, the 13th century drought was over three times the length of the turn-of-the-century drought, overlapping with the start of the severe and sustained "Great Drought" period in the American Southwest (35, 36) that has been implicated in the abandonment of Anasazi settlements in that region (44). Its length firmly places it in the league of other "megadroughts" reported in North America that are unprecedented in modern times in terms of duration. However, in the six centuries that followed, coincident with the period of the Little Ice Age (*ca*. 1300s to late 1800s), drought severity in the UMRB was relatively mild. Only a single event in the late 1500s and consistent with the timing of the 16th-century North American megadrought reported by Stahle et al. (41) rivaled the flow deficits common during the period of the Medieval Climate Anomaly. This long hiatus in drought severity was abruptly ended by the onset of the Dust Bowl drought in the 1930s, which produced the fourth-lowest streamflow departure in the 1,200-y record. This severe and sustained drought was followed 70 y later by the largest decadal-scale flow deficits on record during the turn-of-the-century drought. Thus, in terms of drought events with decadal persistence in streamflow, two of the four most severe droughts in the last 1,200 y appear to have occurred within the last century.

To contextualize the spatial extent and magnitude of the turn-of-the-century drought relative to the four other most severe events in the 1,200-y record, we compared the maximum negative departure of the 10-y spline within each subregion during each drought event (Fig. 2*D*). The maximum deficit during the turn-of-the-century drought was more than 1 *s* below the long-term mean flow level in all subbasins of the UMRB with

the most severe flow deficits centered on the Yellowstone basin and the Missouri Mainstem region. Drought severity during the driest decade of the 13th century drought was positioned over the northern tier of the UMRB while the maximum severity of both the 12th-century drought and the Dust Bowl drought was focused over the Missouri Headwaters and Mainstem regions. The fifth-most-severe drought, which occurred during the early 11th century, appears to be the only event to reach peak flow deficits during below-average temperatures (Fig. 2*C*). The spatial distribution of this event was characterized by more moderate flow deficits evident across the entire basin (Fig. 2*D*).

Twentieth- and 21st-Century Streamflow and Climate Relationships. We explored the 20th- and 21st-century climatic drivers (temperature and precipitation) of the observed changes in UMRB streamflow primarily using naturalized streamflow records compiled from 31 gage records representing nearly every major subbasin in the UMRB. These "naturalized" records represent instrumental-period measurements of streamflow with human influences such as upstream withdrawals, diversions, and reservoir operations removed. Using hierarchical clustering of the streamflow data, five hydrologically distinct subregions within the UMRB are evident: the Northern Tributaries, the Missouri Mainstem region, the Missouri Headwaters, the Yellowstone River, and the Platte River (Fig. 1) (34). We then generated composite records of streamflow for each of these subregions to assess the basin-specific climatic influences on streamflow by averaging standardized reconstructed flow (1900 to 1929) joined to observed (1930 to 2010) flow from those constituent gage records that covered a common period of 1900 through the turn-of-the-century drought (1900 to 2010) (17 records total; Fig. 1). These streamflow records were then compared with subbasin average temperature and precipitation records derived from the 4-km × 4-km gridded Precipitation-elevation Regression on Independent Slopes Model (PRISM) (45) and the Vose et al. (46) climate datasets to examine long-term relationships between hydroclimate and streamflow in the UMRB. Both datasets were used for comparison in these analyses since uncertainties exist in all gridded climate datasets that may affect the results. We also considered that some evidence from the adjacent Columbia River basin and western portions of the UMRB suggests possible underestimation of mountain precipitation in gridded climate datasets during earlier parts of the record (47). Additional analyses using modified precipitation datasets were carried out to estimate how potential underestimation of precipitation could affect our results. See *SI Appendix*, section S3 and Figs. S7–S12 for additional explanation of these uncertainties and supporting analyses. Importantly, the analyses of 20th- and 21st-century climate and streamflow relationships were found to be largely insensitive to the choice of climate dataset used.

The relative influence of temperature on streamflow in each of the five subregions of the UMRB after accounting for the influence of observed precipitation is shown in Fig. 3. Temperature negatively influences streamflow in the UMRB in general, explaining roughly 6% of the variability in instrumental streamflow alone over the 20th and 21st centuries after accounting for precipitation, which explains ∼45%. This estimate is similar to estimates of the influence of temperature on Upper Colorado streamflow (∼6%) (15, 23). In the UMRB, a defined shift in the relative influence of temperature on the generation of streamflow is evident since 1984 (Fig. 3 *A* and *B*). This is consistent with the findings of numerous studies that have identified a distinct shift in the behavior of various biophysical systems ranging from plant phenology to snowmelt timing across North America and beyond that is centered on the mid-1980s (31, 48–51). The year 1984 marks a clear and substantial shift in the negative temperature influence on streamflow relative to preceding decades of the instrumental record and is consistent across all major subbasins of the UMRB (Fig. 3*A*). Prior to 1984, observed streamflow was higher than expected given observed precipitation due to the occurrence of relatively cooler temperatures over this period. However, after 1984, observed precipitation translated to lower than expected streamflow due to warmer temperatures. Consequently, average natural flows across the UMRB declined to levels not seen since the Dust Bowl era by the late 1980s and exceeded those flow reductions by the early 2000s (Fig. 3*A*). The most likely timing of this shift in the UMRB was determined using a series of different length (5 to 30 y) moving-window *t* tests on the time series of the precipitation-adjusted temperature forcing of streamflow. This produced nested probabilities of the timing of a significant shift in the temperature influence occurring on both short and long timescales, which peak at 1984 ($P < 0.001$ on average).

The timing of reduced streamflow and the shifting influence of temperature on water supplies in the UMRB mirrors the snowpack declines in the headwaters of the UMRB from the mid-1970s to late 1980s, which have been attributed in large part to rising spring temperatures (31, 33). Declines in both SWE and snow fraction (ratio of snow to total precipitation) then intensified into the early 2000s (9), coincident with the strongest negative temperature forcing and lowest UMRB streamflow (Fig. 3 *A* and *B*) of the turn-of-the-century drought.

Impacts and Potential Mechanisms Underlying the Observed Changes in the Influence of Temperature on UMRB Streamflow. The 20th- and 21st-century temperature forcing and drought severity records display close synchrony between the influence of warming on streamflow and increasing hydrologic drought severity, suggesting a strong mechanistic link between the two (Fig. 2 *B* and *C*). We investigated both the impacts and potential mechanisms of this linkage over the period of available precipitation data (since 1900) and the highest-quality portion of the reconstructed temperature and streamflow records (since 1800). This allowed us to describe the changes in the likelihood of extreme hydroclimatic conditions within the basin over the last two centuries and the efficiency of streamflow generation relative to changes in temperature since 1900.

The relationship between temperature and streamflow extremes over time was examined by quantifying the likelihood of their cooccurrence by tallying years where standardized values of both temperature and streamflow fall 1 *s* or greater from their respective long-term means over three time periods. The time periods assessed were 1800 to 2010, 1900 to 2010, and 1984 to 2010, when the temperature–streamflow relationship changed distinctly in the climate–streamflow analysis (Fig. 3). A greater occurrence of hot–dry extremes in the UMRB since 1900 is evident with ∼81% of extreme years falling into this category, while ∼59% of extreme years would be classified as hot–dry years over the period since 1800 (Fig. 4*A*). Restricting the extreme year analysis to only the events since 1900, we find that 53% of extreme years since 1900 were coeval hot–dry years. Since 1984, every extreme year has been a hot–dry year (Fig. 4*B*), representing a substantially greater likelihood of hot–dry extremes across the basin since 1984 relative to the period of 1900 to 2010. The combination of elevated air temperature and low streamflow presents a dual challenge for water managers in the UMRB where both agricultural irrigation demands and in-stream water quality for aquatic species are top management priorities (52). High temperatures and low flows simultaneously increase heat stress and evaporative demand on crops while reducing available water for irrigation. Likewise, low-flow conditions restrict available habitat and exacerbate the risk of excessive water temperatures for aquatic species during anomalously warm years (53).

Severe drought in the UMRB and elsewhere in the West is primarily the result of regional precipitation deficits with evidence pointing to an increasing temperature influence (15, 22).

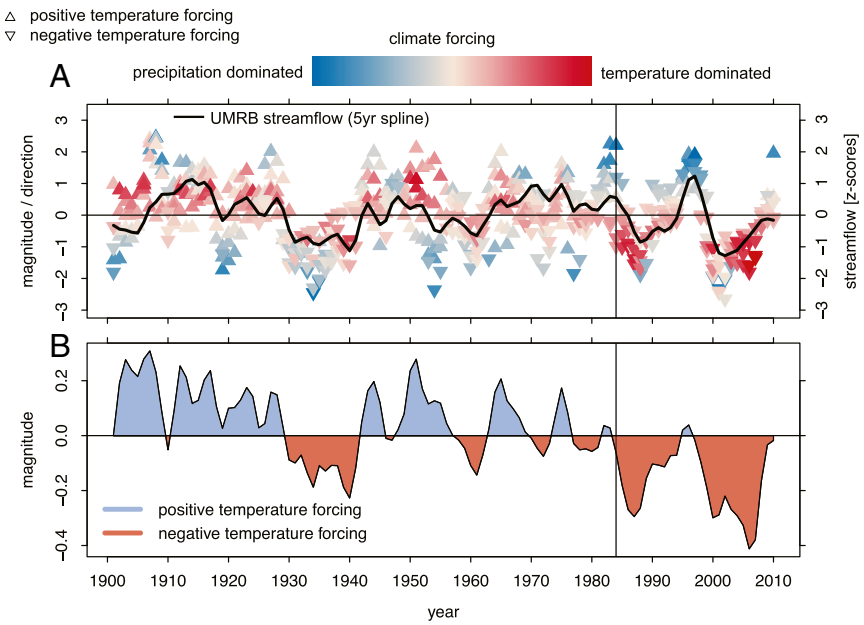

Fig. 3. Climate forcing of streamflow. (*A*) The relative forcing of precipitation and temperature (arrows) on basin-wide mean annual streamflow (black line) in the UMRB. Colored arrows show the individual relative forcing estimates for each subregion of the UMRB for each year, $n = 550$. Colors denote which climate variable was more dominant in the combined forcing of streamflow relative to its long-term average influence. The *y* axis shows the relative magnitude and direction of that combined forcing. The direction of the arrows shows the direction of forcing of the temperature component of that combined forcing (up = supporting and down = suppressing streamflow). All data shown are derived from the 5-y cubic smoothing splines of streamflow and climate data. (*B*) The relative forcing of temperature on streamflow is determined as the temperature anomaly times its multiple-regression coefficient for predicting streamflow along with precipitation. The black line denotes the mean temperature forcing of temperature on streamflow for all subregions of the UMRB; 1930 to 2010 are instrumental streamflow data and 1900 to 1929 are reconstructed. Climate data are from PRISM.

Since the beginning of the 20th century, UMRB runoff-season temperature has warmed considerably (1.4 ± 0.6 °C) and significantly ($P < 0.001$) (*SI Appendix*, Figs. S2 and S10), and increasing temperature can reduce streamflow by removing water from the landscape via evapotranspiration (ET) before it reaches rivers and streams. On the other hand, warmer air temperatures and meteorological drought are also physically linked because a lack of moisture at the surface limits the conversion of downwelling long-wave radiation to latent heat (evaporation) and increases sensible heating (54). This complimentary relationship (55, 56) between temperature as both a driver of potential ET, and an indicator of actual ET, limits the mechanistic inference that can be gleaned from observed relationships between temperature and streamflow alone. Rather, the role of temperature in streamflow generation may more clearly be inferred directly from the relationship between precipitation and streamflow. This is because, over the long term, precipitation not realized as streamflow leaves the landscape primarily as ET, and potential ET is strongly controlled by temperature (57).

We examined the relationship between UMRB precipitation and streamflow directly by estimating the runoff efficiency (RE) (defined here as the difference between streamflow and precipitation anomalies) (28) across the basin from 1900 to 2010. We found significant evidence for a reduction in RE over the 20th and 21st centuries across the Upper Missouri Basin (Fig. 5 *A*–*C*). In an analysis of water-year flows for each gage record where combined reconstructed and observed records are complete from 1900 to 2010, warmer temperatures and reduced RE are apparent and significantly different ($P < 0.001$) across the UMRB. These changes are clear when comparing time periods before and after 1984 (Fig. 5 *A* and *B*), as well as when comparing the notable historical droughts of record, with the turn-of-the-century drought exhibiting significantly ($P < 0.001$) lower REs than even the Dust Bowl drought (Fig. 5*C*). We also carried out an analogous comparison of RE using a monthly water balance model (58), allowing us to directly link estimated changes in ET with changes in RE over time within a framework where closure of the regional water balance is constrained by the physics of the model (Fig. 5 *D* and *E*) (*SI Appendix*, section S4 and Figs. S13 and S14). The 20th- and 21st-century records of temperature, precipitation, ET, and RE from this exercise suggest that the increase in temperature observed over the period is likely responsible for the decrease in RE by way of an increase in ET relative to precipitation (Fig. 5 *D* and *E* and *SI Appendix*, Table S1). This in turn points to a likely mechanism for explaining increased drought severity across the UMRB in recent decades (Figs. 2*C* and 3*A*). For additional estimates of UMRB RE based on different climate datasets see *SI Appendix*, sections S3 and S4 and Figs. S9, S12, and S14. However, results were found to be largely unaffected by the choice of climate dataset or analytical method.

Conclusions

The 1,200-y drought history of the UMRB suggests that, while the turn-of-the-century drought was shorter in duration than numerous earlier events, it may have exceeded the droughts of both the recent and distant past in terms of severity during the driest years of the drought. Similarly, the late 20th and early 21st century in the basin have been characterized by an increasing frequency of coeval hot–dry years that challenge both supply and demand of surface water resources during a period when numerous persistent low streamflow events have resulted in a general drying of the basin relative to the early and mid-20th century.

In consideration of these results, it is important to note that irreducible uncertainties exist in gridded regional climate datasets over complex terrain (*SI Appendix*, section S3). Additionally, there are inherent limitations to empirical, observational assessments and modeling exercises, such as those employed here,

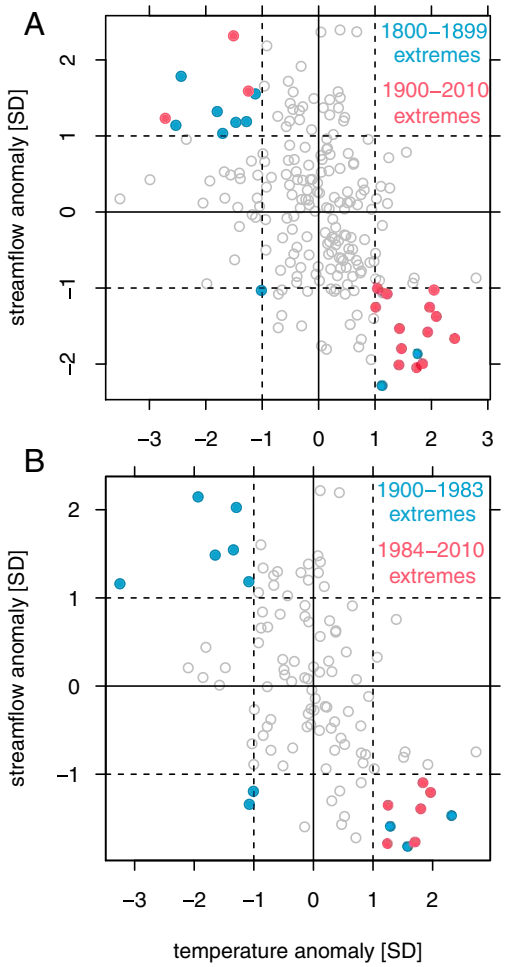

Fig. 4. Extremes in the temperature–streamflow relationship. (*A*) The relationship between temperature and streamflow anomalies for the period 1800 to 2010 (all points, $n = 210$). Colored points show those years in which both temperature and streamflow values were greater than one SD (s) from their mean levels (extremes). Red points show the extremes recorded from 1900 to 2010. (*B*) The temperature–streamflow anomaly relationship from 1900 to 2010 (all points, $n = 110$) and extremes (colored points). Red points show the extremes recorded from 1984 to 2010. Here extremes are defined by the variability in the records from 1900 to 2010 compared with that of 1800 to 2010 shown in *A*.

that challenge our ability to accurately and mechanistically define the often-interrelated influences of precipitation and temperature on streamflow and surface water resources in the western United States. For example, uncertainty remains on the exact contribution of temperature versus precipitation to the 20th-century streamflow declines in the Upper Colorado River Basin despite significant research effort (e.g., refs. 15, 23, 51, and 59), although a clearer picture of the growing temperature influence emerges with improved estimates of key water and energy balance drivers developed at geographically relevant scales (29). Nonetheless, the challenges posed by severe drought such as the turn-of-the-century drought in the UMRB only highlight the need for better climate and hydrologic monitoring moving forward. Improved monitoring would allow for advances in hydrological and statistical modeling, further reducing uncertainties around the proximal drivers of extreme drought under warming conditions.

Despite these challenges, the analyses of basin-wide observations from the UMRB presented here suggest that the enhanced drying during the turn-of-the-century drought, and persistent periods of low flow since the mid-1980s, are coincident with observations of warming air temperatures and reduced RE in the basin (i.e., the proportion of precipitation contributing to streamflow) (24, 26, 28), of which the UMRB appears particularly vulnerable (60). Snowpack has historically been the primary driver of streamflow in the UMRB (9), with recent temperature-driven snowpack declines and earlier spring melt-out documented across the Northern Rockies and the West essentially mirroring drought severity in the UMRB (31–33). Thus, it appears likely that the unusual severity of the turn-of-the-century drought reflects multiple complex hydroclimatic influences centered on the vulnerability of this snow-driven water supply to the effects of warming temperatures (16, 22).

Modeling efforts suggest that future reductions in RE can be expected to continue with warming (*SI Appendix*, section S4 and Figs. S13 and S14), as increased temperature contributes directly to the observed changes in precipitation phase (snow to rain) (9), reduction of mountain snowpack development (32), reductions in surface albedo (29), and enhancement of warm-season ET (9, 24, 29). The combination of hydrologic changes, such as reduced RE and an increasing likelihood of hot–dry extreme years, represent significant challenges for water management in the UMRB. Recent trends (9) and projected changes (52) suggest a future that may require the capture and storage of increasingly early snowmelt runoff, with increased risk of either severe flooding or increasingly severe drought for the portion of the Missouri basin lacking significant multiyear storage capacity. Improvements in multiyear to decadal forecasting capabilities made by incorporating temperature information in snowmelt dominated basins (61) combined with implementing subbasin drought plans (52) could result in enhanced infrastructure operation and water allocation during increasingly severe future drought events.

Methods

Naturalized Streamflow and Climate Data. For the purpose of this study, the UMRB is defined as the region ranging east to west from roughly 105°W longitude to the continental divide, and from north to south from the Milk River in Canada to the South Platte River in Colorado (Fig. 1). An initial collection of 31 naturalized streamflow records for key gaging locations across the UMRB were compiled by Martin et al. (34), representing records deemed to reasonably represent natural flow with limited impacts from human activity (62–66). For the analyses in this study, that dataset was then reduced to only those records used in generating the basin-wide composite streamflow reconstruction and contains a total of 17 streamflow records (Fig. 1) (34).

The climate data used were the 4-km × 4-km gridded monthly temperature and precipitation data from the PRISM dataset (45) for water years (October through September) 1900 through 2010. Analyses of 20th- and 21st-century hydroclimate were also carried out using the nCLIMDIV climate dataset (46) for comparison with results based on PRISM. These analyses and results are discussed in *SI Appendix*, section S3 and shown in *SI Appendix*, Figs. S7–S12. It should be noted that uncertainties exist in both natural estimates of streamflow and in gridded climate datasets. These are inherent limitations that result from the difficulties associated with quantifying human modification of gaged streamflow as well as the patterns of weather that occur between station-based measurements in both space and time. Such uncertainties are not explicitly quantified here.

Development of the Basin-Wide Runoff-Season Temperature Reconstruction. We used the North American tree-ring network from the second phase of the PAGES2k project (67) as our initial set of predictors since it was developed specifically for the reconstruction of regional and global temperatures. This dataset excludes tree-ring chronologies with climate relationships dominated by precipitation or moisture sensitivity, ensuring the records used here primarily reflected temperature conditions and do not overlap with the chronologies used in reconstructing streamflow. Additionally, to ensure fidelity to regional temperatures we screened the full North American network for chronologies within 1,000 km of 110W and 46.75N (*SI Appendix*, Fig. S1) that had a positive and significant ($P < 0.10$) relationship with March through August mean temperature. This selection radius was based on the

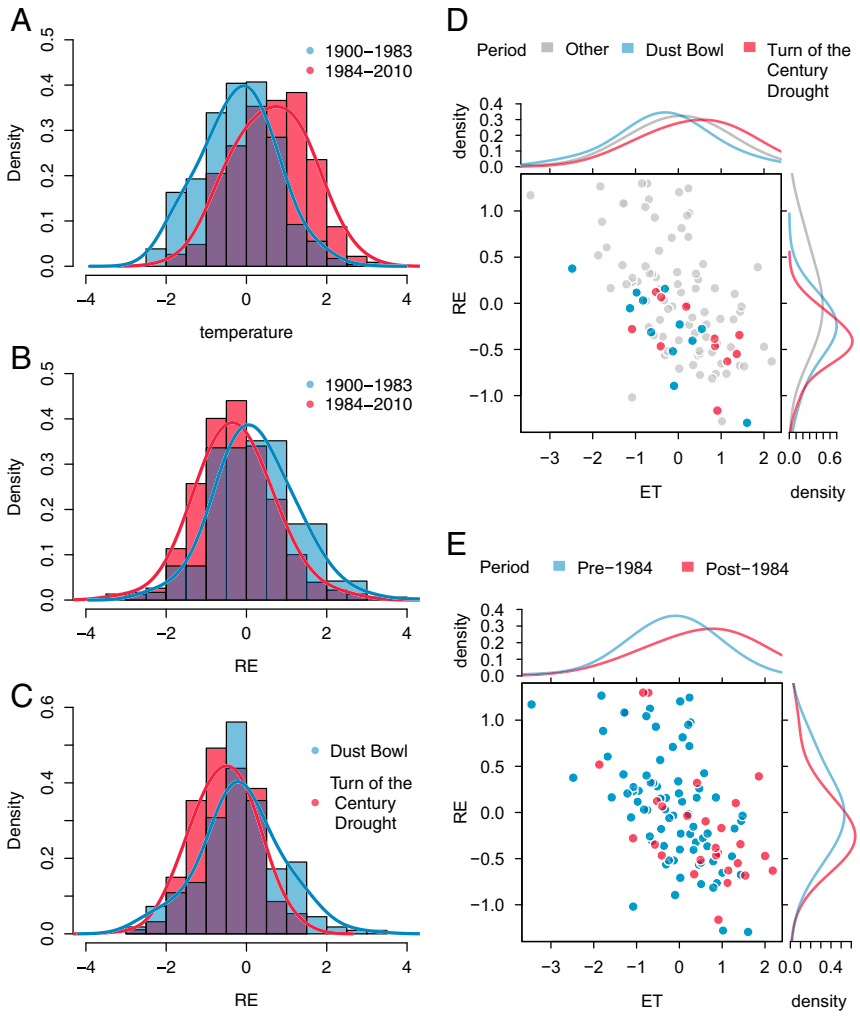

Fig. 5. UMRB basin-wide temperature, ET, and RE. Distributions of (*A*) observed runoff-season (March through August) temperature and (*B*) RE from 1900 to 1983 (blue) and 1984 to 2010 (red), $n = 1870$. *C* shows the distributions of observed RE during the years of the Dust Bowl drought (blue) and turn-of-the-century drought (red), $n = 408$. Lines show the kernel density estimates of the distributions. *D* and *E* show the relationship between aggregated, basin-wide modeled ET and RE with color in *D* denoting values during the two major droughts of record in the UMRB and color in *E* denoting values from before and after 1984. Curves on the top and right axes show the kernel density estimates of the distributions of the values within each time period being compared. *SI Appendix*, Table S1 provides statistics for the time period comparisons based on the full, nonaggregated modeling results.

spatial extent of the Upper Missouri watersheds, previous temperature reconstructions (20, 68), and analyses of observational data (69). We also excluded bristlecone pine chronologies that are known to have a complex and topographically mediated relationship to temperature (70, 71). Applying these criteria resulted in a predictor set of 34 tree-ring chronologies, 10 composed of ring-width measurements and 24 of maximum latewood density. This latter measurement is known to be a better proxy for temperature than ring width alone (20, 72).

PRISM (45) monthly temperature data averaged over the primary snowmelt and ET months of March through August served as the predictand for our reconstruction. This temperature record is the average of the temperature records developed for each of the five UMRB subregions (Fig. 1) and spans the period 1900 to 2014. We used a nested composite-plus-scale method (69, 72, 73) to reconstruct regional mean runoff-season temperature from the network of temperature sensitive tree-ring data (*SI Appendix*, section S1 and Figs. S1 and S2).

Composite Climate Data for Each of the Five Naturalized Flow Regions. To investigate the relationship between climate and streamflow for each region, we identified the major hydrologic unit (HU) level-8 watersheds that made up the primary drainage area for each regional cluster then averaged the PRISM 4-km gridded climate data. Climate data were averaged for each month of each year for each variable across the HU 8 watersheds falling within each cluster. The HU 8 watersheds used to estimate climate in each cluster are identified in *SI Appendix*, Table S2.

Spatial Distributions of the Most Severe Droughts. In order to characterize the geographic distribution of drought during the most severe events in the record, we first calculated decadal flow deficits separately for each of the five clustered subregions. This provided a record of drought severity for each subregion relative to the long-term variation in streamflow for the region itself. We then assessed the level of flow deficits for each subregion over the driest decade within each of the five major droughts evident in the basin-wide drought severity record. (Fig. 2*D*).

Estimating the Relative Forcing of Precipitation and Temperature on Streamflow. To estimate the temperature forcing of streamflow since 800 (Fig. 2*B*), we followed the approach of Pederson et al. (31). Using linear regression, we regressed the time series of water-year (prior October through September) streamflow *z*-scores for each of the five subregions against the time series of mean runoff-season (March through August) temperature *z*-scores for each subregion. We then multiplied the regression coefficients by the time series of temperature *z*-scores to estimate the relative forcing of temperature on streamflow.

To estimate both the precipitation and temperature forcing of streamflow since 1900 (Fig. 3), we used an analogous multiple regression (MLR) approach.

We regressed the streamflow *z*-scores for each UMRB subbasin against the time series of total water-year precipitation and mean runoff-season temperature *z*-scores then multiplied the regression coefficients by the time series of temperature and precipitation *z*-scores to estimate the relative forcing of each variable on streamflow over time. For each year of the common observational period, this quantifies the relative magnitude and direction of the contributions of temperature and precipitation to streamflow volumes. We quantified the magnitude and sign of the combined forcing of both variables as the sum of the relative forcings (Fig. 3*A*, *y* axis). We also show which variable is more anomalous in its forcing of streamflow each year relative to its long-term average influence (absolute value of the dominant forcing – absolute value of the subordinate forcing) using color in Fig. 3*A*. The relative influence of temperature alone on streamflow over time after accounting for the influence of precipitation is shown in Fig. 3*B*. We used a series of moving windows from 5 to 30 y preceding and following each year in the temperature-forcing record to identify the most likely year that the apparent shift to persistent negative temperature forcing of streamflow occurred based on a two-tailed *t* test in which the null hypothesis was that the magnitude of the forcing preceding and following a given year in the record were the same. For this analysis, all data were instrumental in origin except for the 1900 to 1929 streamflow values which were reconstructed from tree rings (34).

It is important to note that the average correlation between temperature and precipitation across the five subregions of the UMRB during the 20th and early 21st century is −0.39. This means that one variable could potentially account for up to roughly 16% of the variability in the other if one variable fully controlled the response of the other. In reality, precipitation can lead to changes in temperature (e.g., sensible versus latent heating) (56) or temperature can drive changes in precipitation (74), and this can happen in various ways that are difficult to quantify. This highlights an important limitation of any MLR analysis in which predictor variables are correlated resulting in a degree of uncertainty that will always exist when trying to quantify the possible effect of a single predictor variable on the response. In this particular case, we based our analyses on the determined relationships between either temperature or precipitation and streamflow, while holding the second predictor variable constant. However, because some information about variability in streamflow is shared by variability in both temperature and precipitation, our estimation of those relationships is somewhat less precise than if temperature and precipitation varied completely independently.

Estimating the Probability of Extremes in the Temperature vs. Streamflow Relationship over Time. We investigated the occurrence of extremes in the relationship between temperature and streamflow by first identifying years in which both temperature and streamflow values were further than 1 *s* from their respective means (hereafter "extreme" years). This established four possible conditions in which extreme years could occur, dry–hot, dry–cold, wet–hot, and wet–cold, in terms of temperature and streamflow, respectively. We then calculated the percentage of extreme years falling into each category, carrying out the calculation for the 211-y period from 1800 to 2010, the period since 1900, and the period since 1984. The percentages for the period since 1900 are in reference to extreme years defined by the SD of the full 211 y, while the percentages for the period since 1984 are in reference to extreme years defined over the period since 1900.

Estimating the Change in RE over Time. Following the approach of Woodhouse and Pederson (28), we estimated RE at every gage in the composite record for every year since 1900 (1900 to 2010) as the difference between standardized streamflow and precipitation. Differences in RE between time periods were assessed using two-tailed *t* tests.

Data Availability. The Upper Missouri Basin naturalized streamflow records and tree-ring–based naturalized streamflow reconstructions used in this study are available online from the US Geological Survey (USGS) (https://doi.org/10.5066/P9FC7ILX). The tree-ring chronologies used in the streamflow reconstructions are available online from National Oceanic and Atmospheric Administration National Centers of Environmental Information (https://www.ncdc.noaa.gov/paleo/study/26831). The Upper Missouri Basin runoff-season temperature reconstruction is available from the National Oceanic and Atmospheric Administration National Centers of Environmental Information (https://www.ncdc.noaa.gov/paleo/study/29413). The tree-ring chronologies used in the temperature reconstruction are available online from the PAGES 2K version 2 consortium (https://doi.org/10.1038/sdata.2017.88).

ACKNOWLEDGMENTS. Research support provided through the NSF Paleo Perspectives on Climate Change (P2C2) Program (Grants 1404188, 1403957, 1403102, and 1401549), NSF Grant 1803995, the NSF Graduate Research Fellowship Program (Grant 1049562), the Graduate Research Internship Program (GRIP), the US Bureau of Reclamation WaterSMART Program (Sustain and Manage America's Resources for Tomorrow), the state of Montana Department of Natural Resources and Conservation, the Lamont-Doherty Earth Observatory (contribution number 8398), the USGS Powell Center for Synthesis and Analysis, the USGS Land Resources Mission Area, and the North Central Climate Adaptation Science Center. Coordination of GRIP at USGS is through the Youth and Education in Science programs within the Office of Science Quality and Integrity. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.

1. US Bureau of Reclamation, Colorado River Basin water supply and demand study: Executive summary. https://www.usbr.gov/watersmart/bsp/docs/finalreport/ColoradoRiver/CRBS_Executive_Summary_FINAL.pdf. Accessed 23 August 2019.
2. US Bureau of Reclamation, St. Mary River and Milk River Basins study summary report. https://www.usbr.gov/watersmart/bsp/docs/finalreport/Milk-StMary/Milk-StMary_SummaryReport.pdf. Accessed 23 August 2019.
3. US Bureau of Reclamation, Republican River Basin study: Final executive summary report. https://www.usbr.gov/watersmart/bsp/docs/finalreport/republican/republican-river-basin-study-executive-summary-report.pdf. Accessed 23 August 2019.
4. US Bureau of Reclamation, Sacramento and San Joaquin Rivers basin study: Basin study report and executive summary. https://www.usbr.gov/watersmart/bsp/docs/finalreport/sacramento-sj/Sacramento_SanJoaquin_SUMMARY.pdf. Accessed 23 August 2019.
5. US Bureau of Reclamation, Klamath River Basin study: Summary report. https://www.usbr.gov/watersmart/bsp/docs/klamath/summaryreport.pdf. Accessed 23 August 2019.
6. C. Lesk, P. Rowhani, N. Ramankutty, Influence of extreme weather disasters on global crop production. *Nature* **529**, 84–87 (2016).
7. J. Lawrimore, S. Stephen, "Climate of 2002 annual review" (National Climate Data Center, Ashville, NC, 2003).
8. J. T. Overpeck, Climate science: The challenge of hot drought. *Nature* **503**, 350–351 (2013).
9. E. K. Wise, C. A. Woodhouse, G. J. McCabe, G. T. Pederson, J. M. St-Jacques, Hydroclimatology of the Missouri River Basin. *J. Hydrometeorol.* **19**, 161–182 (2018).
10. US General Accounting Office, "Water resources: Corps' management of ongoing drought in the Missouri River Basin" (US General Accounting Office, Washington, DC, 1992).
11. V. M. Mehta, N. J. Rosenberg, K. Mendoza, Simulated impacts of three decadal climate variability phenomena on water yields in the Missouri River Basin. *J. Am. Water Resour. Assoc.* **47**, 126–135 (2011).
12. A. J. DeLonay, R. B. Jacobson, D. M. Papoulias, D. G. Simpkins, M. L. Wildhaber, Ecological requirements for pallid sturgeon reproduction and recruitment in the Lower Missouri River: A research synthesis 2005–08. https://digitalcommons.unl.edu/usgspubs/106. Accessed 23 August 2019.
13. D. Garrick, K. Jacobs, G. M. Garfin, Decision making under uncertainty: Shortage, stakeholders and modeling in the Colorado River basin. *J. Am. Water Resour. Assoc.* **44**, 381–398 (2008).
14. D. Griffin, K. J. Anchukaitis, How unusual is the 2012–2014 California drought? *Geophys. Res. Lett.* **41**, 9017–9023 (2014).
15. H. Kwon, U. Lall, A copula-based nonstationary frequency analysis for the 2012–2015 drought in California. *Water Resour. Bull.* **52**, 5662–5675 (2016).
16. B. Udall, J. Overpeck, The twenty-first century Colorado River hot drought and implications for the future. *Water Resour. Res.* **53**, 2404–2418 (2017).
17. J. Prairie, K. Nowak, B. Rajagopalan, U. Lall, T. Fulp, A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data. *Water Resour. Res.* **44**, W06423 (2008).
18. P. R. Seaber, F. Kapinos, G. L. Knapp, Hydrologic unit maps-Water supply paper 2294. https://pubs.usgs.gov/wsp/wsp2294/. Accessed 2 October 2018.
19. E. R. Cook *et al.*, Megadroughts in North America: Placing IPCC projections of hydroclimatic change in a long-term paleoclimate context. *J. Quat. Sci.* **25**, 48–61 (2010).
20. K. J. Anchukaitis *et al.*, Last millennium northern hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quat. Sci. Rev.* **163**, 1–22 (2017).
21. J. L. Weiss, C. L. Castro, J. T. Overpeck, Distinguishing pronounced droughts in the southwestern United States: Seasonality and effects of warmer temperatures. *J. Clim.* **22**, 5918–5932 (2009).
22. A. P. Williams *et al.*, Contributions of anthropogenic warming to California drought during 2012–2014. *Geophys. Res. Lett.* **42**, 6819–6828 (2015).
23. C. A. Woodhouse, G. T. Pederson, K. Morino, S. A. McAfee, G. J. McCabe, Increasing influence of air temperature on upper Colorado River streamflow. *Geophys. Res. Lett.* **43**, 2174–2181 (2016).
24. G. J. McCabe, D. M. Wolock, G. T. Pederson, C. A. Woodhouse, S. McAfee, Evidence that recent warming is reducing upper Colorado River flows. *Earth Interact.* **21**, 1–14 (2017).
25. K. M. Andreadis, D. P. Lettenmaier, Trends in 20th century drought over the continental United States. *Geophys. Res. Lett.* **33**, L10403 (2006).
26. E. R. Griffin, J. M. Friedman, Decreased runoff response to precipitation, Little Missouri River Basin, northern Great Plains, USA. *J. Am. Water Resour. Assoc.* **53**, 576–592 (2017).

27. A. F. Hamlet, P. W. Mote, M. P. Clark, D. P. Lettenmaier, Effects of temperature and precipitation variability on snowpack trends in the western United States. *J. Clim.* **18**, 4545–4561 (2005).
28. C. A. Woodhouse, G. T. Pederson, Investigating runoff efficiency in upper Colorado River streamflow over past centuries. *Water Resour. Res.* **54**, 1–15 (2017).
29. P. C. D. Milly, K. A. Dunne, Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation. *Science* **367**, eaay9187 (2020).
30. P. A. Norton, M. T. Anderson, J. F. Stamm, "Trends in annual, seasonal, and monthly streamflow characteristics at 227 streamgages in the Missouri River Watershed, water years 1960–2011" (Scientific Investigations Report 2014–5053, US Geological Survey, 2014), p. 128.
31. G. T. Pederson, J. L. Betancourt, G. J. McCabe, Regional patterns and proximal causes of the recent snowpack decline in the Rocky Mountains, U.S. *Geophys. Res. Lett.* **40**, 1811–1816 (2013).
32. G. T. Pederson *et al*., The unusual nature of recent snowpack declines in the North American Cordillera. *Science* **543**, 332–336 (2011).
33. P. W. Mote, S. Li, D. P. Lettenmaier, M. Xiao, R. Engel, Dramatic declines in snowpack in the western US. *Clim Atmos Sci* **1**, 2 (2018).
34. J. T. Martin *et al*., 1200 years of Upper Missouri River streamflow reconstructed from tree rings. *Quat. Sci. Rev.* **224**, 1–15 (2019).
35. A. E. Douglass, The secret of the Southwest solved with talkative tree rings. *Natl. Geogr. Mag.*, 736–770 (1929).
36. A. E. Douglass, "Dating Pueblo Bonito and other ruins of the Southwest" (Contributed Technical Papers, Pueblo Bonito Series 1, National Geographic Society, 1935).
37. C. A. Woodhouse, D. M. Meko, G. M. MacDonald, D. W. Stahle, E. R. Cook, A 1,200-year perspective of 21st century drought in southwestern North America. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21283–21288 (2010).
38. C. A. Woodhouse, S. T. Gray, D. M. Meko, Updated streamflow reconstructions for the Upper Colorado River Basin. *Water Resour. Res.* **42**, W05415 (2006).
39. S. Stine, Extreme and persistent drought in California and Patagonia during the Medieval time. *Nature* **369**, 546–549 (1994).
40. D. W. Stahle, F. K. Fye, E. R. Cook, R. D. Griffin, Tree-ring reconstructed megadroughts over North America since A.D. 1300. *Clim. Change* **83**, 133–149 (2007).
41. D. W. Stahle *et al*., Tree-ring data document 16th century megadrought over North America. *Eos* **81**, 121–132 (2000).
42. E. R. Cook, L. A. Kalriukstis, *Methods of Dendrochronology* (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990).
43. US Bureau of Reclamation, Colorado River Basin water supply and demand study: Technical report B-water supply assessment. http://www.usbr.gov/lc/region/programs/crbstudy/finalreport/TechnicalReportB-WaterSupplyAssessment/TR-B_Water_Supply_Assessment_FINAL.pdf. Accessed 11 February 2018.
44. R. L. Axtell *et al*., Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proc. Natl. Acad. Sci. U.S.A.* **99** (suppl. 3), 7275–7279 (2002).
45. C. Daly *et al*., Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.* **28**, 2031–2064 (2008).
46. R. S. Vose *et al*., Improved historical temperature and precipitation time series for U.S. climate divisions. *J. Appl. Meteorol. Climatol.* **53**, 1232–1251 (2014).
47. C. H. Luce, J. T. Abatzoglou, Z. A. Holden, The missing mountain water: Slower westerlies decrease orographic enhancement in the Pacific Northwest USA. *Science* **342**, 1360–1364 (2013).
48. T. R. Ault, A. K. Macalady, G. T. Pederson, J. L. Betancourt, M. D. Schwartz, Northern Hemisphere modes of variability and the timing of spring in western North America. *J. Clim.* **24**, 4003–4014 (2011).
49. G. J. McCabe, J. L. Betancourt, G. T. Pederson, M. D. Schwartz, Variability common to first leaf dates and snowpack in the western conterminous United States. *Earth Interact.* **17**, 1–18 (2013).
50. G. J. McCabe, M. P. Clark, Trends and variability in snowmelt runoff in the Western United States. *J. Hydrometeorol.* **6**, 476–482 (2005).
51. P. C. Reid *et al*., Global impacts of the 1980s regime shift. *Glob. Change Biol.* **22**, 682–703 (2016).
52. US Bureau of Reclamation, "Upper Missouri basin impacts assessment" (US Bureau of Reclamation, Denver, CO, 2018).
53. R. P. Kovach *et al*., An integrated framework for ecological drought across riverscapes of North America. *Bioscience* **69**, 418–431 (2019).
54. M. L. Roderick, F. Sun, W. H. Lim, G. D. Farquhar, A general framework for understanding the response of the water cycle to global warming over land and ocean. *Hydrol. Earth Syst. Sci.* **18**, 1575–1589 (2014).
55. W. Brutsaert, M. B. Parlange, Hydrologic cycle explains the evaporation paradox. *Nature* **396**, 30 (1998).
56. M. Aminzadeh, M. L. Roderick, D. Or, A generalized complementary relationship between actual and potential evaporation defined by a reference surface temperature. *Water Resour. Res.* **52**, 385–406 (2016).
57. C. W. Thornthwaite, An approach toward a rational classification of climate. *Geogr. Rev.* **38**, 55–94 (1948).
58. G. J. McCabe, D. M. Wolock, Century-scale variability in global annual runoff examined using a water balance model. *Int. J. Climatol.* **31**, 1739–1748 (2011).
59. M. Hoerling *et al*., Causes for the century-long decline in Colorado River flow. *J. Clim.* **32**, 8181–8203 (2019).
60. G. J. McCabe, D. M. Wolock, M. Valentin, Warming is driving decreases in snow fractions while runoff efficiency remains mostly unchanged in snow-covered areas of the western United States. *J. Hydrometeorol.* **19**, 803–814 (2018).
61. F. Lehner *et al*., Mitigating the impacts of climate nonstationarity on seasonal streamflow predictability in the U.S. Southwest. *Geophys. Res. Lett.* **44**, 12208–12217 (2017).
62. L. D. Brekke *et al*., "Climate change and water resources management: A federal perspective" (Circular 1331, US Geological Survey, 2010).
63. L. Cary, C. Parrett, "Synthesis of natural flows at selected sites in the Upper Missouri River Basin, Montana, 1928–1989" (Water-Resources Investigations Rep. 95-4261, US Geological Survey, 1996).
64. K. J. Chase, "Streamflow statistics for unregulated and regulated conditions for selected locations on the Upper Yellowstone and Bighorn Rivers, Montana and Wyoming, 1928–2002" (USGS Scientific Investigations Rep. 2014-5115, US Geological Survey, 2014).
65. S. Gangopadhyay, T. Pruitt, "Climate change analysis for the St. Mary and Milk River systems in Montana" (Reclamation Technical Memorandum No. 86-68210–2010-04, 2010).
66. J. R. Slack, A. M. Lumb, J. M. Landwehr, "Hydro-Climatic Data Network (HCDN)—A USGS streamflow data set for the U.S. for the study of climate fluctuations" (US Geological Survey, 1994).
67. PAGES 2k Consortium, A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data* **4**, 1–33 (2017).
68. E. R. Cook *et al*., Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 C.E. *Clim. Dyn.* **41**, 2957–2972 (2013).
69. R. S. Bradley, P. D. Jones, "Little Ice Age" summer temperature variations: Their nature and relevance to recent global warming trends. *Holocene* **3**, 367–376 (1993).
70. A. G. Bunn, M. K. Hughes, M. W. Salzer, Topographically modified tree-ring chronologies as a potential means to improve paleoclimate inference. *Clim. Change* **105**, 627–634 (2011).
71. M. W. Salzer, E. R. Larson, A. G. Bunn, M. K. Hughes, Changing climate response in near-treeline bristlecone pine with elevation and aspect. *Environ. Res. Lett.* **9**, 1–8 (2014).
72. R. Wilson *et al*., Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quat. Sci. Rev.* **134**, 1–18 (2016).
73. J. Esper, E. R. Cook, F. H. Schweingruber, Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* **295**, 2250–2253 (2002).
74. P. Berg, C. Moseley, J. O. Haerter, Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.* **6**, 181–185 (2013).

PNAS
www.pnas.org

Supplementary Information for

Increased drought severity tracks warming in the United States' largest river basin

Justin T. Martin[1], Gregory T. Pederson[1], Connie A. Woodhouse[2,3], Edward R. Cook[4], Gregory J. McCabe[5], Kevin J. Anchukaitis[2,3], Erika K. Wise[6], Patrick Erger[7], Larry Dolan[8], Marketa McGuire[9], Subhrendu Gangopadhyay[9], Katherine Chase[10], Jeremy Littell[11], Stephen Gray[11], Scott St. George[12], Jonathan Friedman[13], Dave Sauchyn[14], Jeannine St. Jacques[15], and John King[16]

[1]U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, USA
[2]School of Geography and Development, University of Arizona, Tucson, AZ, USA
[3]Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ, USA
[4]Lamont-Doherty Earth Observatory, Palisades, New York, USA
[5]U.S. Geological Survey, Water Resources Division, Denver, Colorado, USA
[6]Department of Geography, University of North Carolina, Chapel Hill, NC, USA
[7]U.S. Bureau of Reclamation, Great Plains Regional Office, Billings, MT, USA
[8] Montana Department of Natural Resources and Conservation, Helena, MT, USA
[9] U.S. Bureau of Reclamation, Technical Service Center, Denver, CO, USA
[10]U.S. Geological Survey, Wyoming-Montana Water Science Center, Helena, MT, USA
[11]U.S. Geological Survey, Alaska Climate Adaptation Science Center, Anchorage AK, USA
[12]Department of Geography, Environment and Society, University of Minnesota, Minneapolis, MN, USA
[13]U.S. Geological Survey, Fort Collins Science Center, Ft. Collins, CO, USA
[14]Department of Geography and Environmental Studies, University of Regina, Regina, Saskatchewan, Canada
[15]Department of Geography, Planning and Environment, Concordia University, Montreal, Quebec, Canada
[16]Lone Pine Research, Bozeman, MT, USA

Corresponding author: Justin Martin (justinmartin@usgs.gov)

This PDF file includes:

Supplementary text S1-S4
Figures S1 to S14
Tables S1 and S2
SI References

Supplemental Information Text

S1. Cross-validation of modeled temperature

The nested Composite-Plus-Scale (CPS) method was used to reconstruct regional mean runoff-season (Mar-Aug) temperatures from a network of high-elevation, temperature sensitive ring-width and maximum density (MXD) tree-ring data (Fig. S1 and S2). In this approach, all of the available tree-ring data are normalized and averaged to create a mean composite proxy series. This single series is then scaled to the mean and standard deviation of the instrumental data over the calibration period, and reconstruction skill is evaluated based on the residuals. We used a calibration period of 1930 to 1976 (bounded by the last year of the common period covered by all the available tree-ring chronologies) and a validation period of 1901 to 1930 (bounded by the first year of available PRISM temperature data). In a nested CPS approach, the reconstruction is extended back in time by recalculating the composite and rescaling it to the instrumental temperature data iteratively each time there is a change in sample size. This allows the reconstruction to extend further back in time while also accounting for changes in sample size, reconstruction accuracy, and the resulting change in reconstruction calibration and validation skill.

The reconstruction of regional runoff-season temperatures is skillful back to approximately 1500 CE (Fig. S2). Importantly, both the Reduction of Error (RE) and Coefficient of Efficiency (CE) statistics remain positive through this period. Prior to 1500, the loss of proximal MXD chronologies to our study region results in a temperature reconstruction without sufficient skill for further statistical analyses. The reconstruction, however, is still plotted alongside the streamflow and drought deficits prior to 1500 since the sign of decadal temperature anomalies likely retains enough useful information to illustrate whether a drought event occurred under generally cool or warm conditions.

S2. Assessing the sensitivity of comparing droughts over time to methodological choices

Drought deficits are presented here as a calculation based on standardized streamflow values (*z*-scores) but can be calculated on observed streamflow values as well (units of cubic feet per second [CFS]). In generating the UMRB composite record, taking the average of the former may yield a different result from the latter in the case where flow time series of very different magnitudes are averaged. Calculating drought deficits on observed streamflow values more closely approximates the mass dynamics of a drought estimate that is akin to the total flow volume measured at a single terminal gaging location at the outlet of the basin, which is non-existent in the UMRB. To determine how this affects our estimate of basin-wide deficits, we also quantified deficits based on the UMRB composite record generated by taking the simple average of the streamflow records in CFS units. In this case, the temporal dynamics of individual large rivers can be expected to have a

greater effect on the basin wide average than individual small rivers. Figure S3 depicts this deficit record calculated on streamflow values in CFS. In this alternative approach, the decadal deficits of the Turn-of-the-Century Drought remain substantially more severe than those of other droughts in the record.

Additionally, to contextualize the Turn-of-the-Century drought with paleoclimatic and paleohydrologic records, it was necessary to splice the basin-wide composite streamflow reconstruction (which ends in 1998) to the observed record that runs through 2010. This splice occurred at the year 1930 in order to include the dry years of the 1930s as reflected in the actual gage records. Although the streamflow reconstruction is mean and variance scaled to its respective observational record over the common period, splicing the instrumental and reconstructed records inevitably introduces bias at the transition of record types in that the reconstruction itself will fail to perfectly capture the variation in the observed hydrology. Thus, we attempted to assess potential bias that existed in the reconstruction relative to the observational record.

In this bias assessment, we first quantified the error in the prediction of low flow values in the streamflow reconstruction versus the observed record over the common period for each gage in the composite record. We did so by comparing reconstructed and observed flows from years where flow values were more than 1 s (standard deviation) below the mean flow level. The mean low-flow level in the reconstruction matched that of the observed record very closely over the calibration period such that the mean low-flow level for the reconstruction was 0.006 s lower than the observed flow. To determine the effect this very small underestimation of observed low flows had on the full-length drought deficit record, we explicitly aligned the reconstructed and observed flow records based on their respective low-flow means. This places a greater emphasis on the accuracy of the lowest reconstructed flows in comparing drought years between reconstructed and observed records. Even under this scenario, where the reconstruction mean is matched to the observed over only the driest years of the calibration period, the peak severity of Turn-of-the-Century Drought still appears unprecedented over the last 1200 years (Fig S4).

Also, because we focused on decadal-scale variability in drought severity, we filtered higher frequency variability (< 10 years) from the long-term streamflow record using the cubic smoothing spline of streamflow anomalies with a spline length of 10 years and a frequency response of 0.5. To avoid the potential for end-effects to exaggerate the severity of the Turn-of-the-Century Drought, we padded the end of the time-series with flow values set to the maximum observed flow in the 1200 year record for a number of years equal to half the length of the spline. Thus, the severity of the Turn-of-the-century drought is almost certainly underestimated due to the application of a maximum flow value padding.

Since the choice of filtering method affects both the definition and magnitude of drought events, we explored how calculating drought severity based on splines of

lengths from five to fifteen years affected the comparison of droughts over time. Figure S5 shows the drought severity records for each of these spline lengths, demonstrating that Turn-of-the-Century Drought is more severe than other droughts on record when assessed over periods of various lengths from five to twelve years. At spline lengths of 13 years and beyond, it is likely that the 13th century drought represents a drought event of greater persistent severity than the Turn-of-the-Century Drought, as is to be expected from an event over three times as long. However, because fitting longer splines to the streamflow record requires more maximum flow value padding of the recent years, the underestimation of drought severity during the Turn-of-the-Century Drought also increases when assessed via longer splines. Future assessments of changing drought dynamics through the 21st century in the UMRB will require updated naturalized streamflow records.

Finally, to test another method that integrates information on both drought severity and duration, we also calculated the intensity of each drought in the 1200-year record as the cumulative deficit / length of the drought. We used the same drought definition (two or more consecutive negative-anomaly years in the 10-year spline of streamflow) but calculated the intensity of each drought event directly from the un-smoothed annual flow anomalies. Again, the Turn-of-the-Century drought was ranked as significantly more intense than any other drought event on record (Fig. S6).

S3. Assessing the sensitivity of relationships between temperature, precipitation, and streamflow to uncertainties in gridded climate data

Numerous gridded estimates of precipiation and temperature for the western United States are now available for various time periods since the late 1800s. It is well known that all gridded climate datasets are inherently uncertain owing to issues that may include measurement errors in underlying data (1, 2), inhomogeneities in individual or network records (3–5), and methodological choices specific to the development of each dataset (6). Additionally, the sparsity of current and espescially earlier observations from which gridded datasets are derived is an inherent limitation in all gridded products (7).

While a certain lack of independence between all climate datasets exists, we attempted to determine the effect of our choice of climate data (i.e. PRISM) on the findings of this study to the degree possible by carrying out the hydroclimatic analyses reported herein on homoginized climate datasets of equivalent length to PRISM (8), and on a modified version of the PRISM precipitation dataset that extends back to 1950. Modifications made to the orgininal PRISM dataset are from Luce et al. (9), and were developed to account for potential biases that may, or may not, be present in the precipitation dataset related to reduced orgraphic precipitation due to a reduction in 700 hectopascal (hpa) zonal (U700) windspeeds. These efforts are described in detail below.

In the construction of the PRISM gridded precipitation dataset, the 4x4 km grid cells are first interpolated from observations (i.e met stations) using a local regression function informed by climatological normals for the 1981-2010 period (10). This local regression approach accounts for the influence of lapse rates (and other variables) due to topography on the interpolation of monthly climate values between distant obervation locations. Based on the work by Luce et al., 2013 (9) covering the Pacific Northwest region, we recognize the possibility that the use of 1981-2010 climatology may not accurately reflect the relationship between topography and precipitation prior to the early 1980s in light of evidence suggesting that winter zonal 700 hpa westerly winds over portions of the region have declined since the mid-20th century. This is because during winter the speed of lower tropospheric winds is strongly correlated with precipitation in mountainous regions conducive to orographic enhancement of precipitation. Thus decreasing westerly winds in recent decades may have decreased mountain precipitation, and this trend may not be accurately reflected in gridded precipitation records due to the lack of high-elevation weather stations in the region. However, the likelihood that this situation exists in the UMRB is highly uncertain as the work of Luce et. al., 2013 focused primarily on the Pacific Northwest, not the UMRB. Additonally, hydrologically significant moisture influxes into the the Upper Missouri over both the cool and warm seasons occur from a greater diversity of source regions and stormtracks (11)(e.g. wintertime northwesterly flow patterns, springtime moisture advection off the Gulf of Mexico via the low-level jet and cutoff lows) than the dominant zonal flow pattern responsible for the majority of the moisture delivery across the the Columbia Basin (i.e. the Pacific Nortwest). Climatically important sources of moisture aside, since Luce et al. (9) do show correlations between U700 winds and precipitation in some less rain-shadowed catchments in the far western portion of the UMRB it is worth invetigating the potential influence such a procces could have on our results.

To understand how an underestimation of mountain precipitation in the UMRB prior to the early 1980s could affect our analyses, we used the standardized regression slopes for the effect of U700 windspeed on precipitation from Luce et al. (9) to calculate the ammount of precipitation that could hypothetically be missing from PRISM grids in each year prior to 1981. We were able to perform this analysis from 1950 forward based on the reported decrease in average U700 winds of 0.19 m/s/decade. We estimated a standardized regression slope of $\sim$0.12 fraction precipitation per m/s for the reported stations that fall within the UMRB. We calculated the hypothetical underestimation of precipitation in each year as the total annual precipitation times 0.019, times the number of years before 1981 times 0.12, times the mean of 1981-2010 precipitation, then added that quantity back to the total precipitation for each year (Fig. S7). Because we carried out this exercise on HU 8 watershed-aggregated PRISM grids, all grid cells in the UMRB headwater basins were effectively adjusted wet in the earlier decades (not just those mountain grids with the most significant potential for orographic enhancement). The adjustment to PRISM precipitation likely over-corrects for any actual

underestimation of precipitation that may be present in PRISM data prior to 1981. This essentially biases the precipitation of the 1981 – 2010 period dry relative to earlier periods, creating a situation in which precipitation appears to have more of an increasing influence on streamflow generation than it does if PRISM precipitation is left unaltered.

Even in this rather extreme hypothetical situation, the influence of temperature on UMRB streamflow during the drought periods of the late 1980s and early 2000s has been more negative and of greater magnitude than during previous decades (Fig. S8). As expected, the primary influence of adusting PRISM wet in earlier decades is to assign slightly more importance to the effect of precipitation on recent droughts and slightly less importance on recent pluvials. Nonetheless, a clear shift to an increasingly negative forcing of temperature on streamflow is evident in most sub-basins since the 1980s, and is coincident with very dry conditions in the late 1980s and early 2000s (Fig. S8).

Adjusting PRISM precipitation data for hypothetical precipitation biases also affects our estimates of UMRB runoff efficiency over time. Although doing so shortens the available period of record for the analysis to 1950 – 2010, figure S9 shows that even with higher precipitation estimates prior to 1981, runoff efficiency remains significantly lower during the period from 1984 – 2010 than during the preceding three decades (p<0.001).

An additional uncertainty in PRISM data is the known existence of inhomogeneities in some observational networks and records underlying many gridded products including the PRISM dataset (3, 4). Some level of inhomogeneity over time in portions of the available station data network is likely due to climate-independent changes in station site conditions, instrumentation, and location. These issues can be difficult or impossible to quantify, however, several gridded climate datasets now exist that employ pairwise homogenization techniques to limit the influence of potential inhomogeneities in an effort to make them more suitable for trend and long-term analyses (8, 12, 13). The nCLIMDIV dataset by Vose et al. (10) is directly comparable to the PRISM dataset in terms of temporal coverage (1895 to 2010) and employs such a homogenization technique. Thus, while the nCLIMDIV and PRISM datasets are not independent of each other owing to shared underlying data and certain methodologies, comparison of the two allows an understanding of the potential effects of inhomogeneities in PRISM data on trends reflected by that dataset. Figure S10 shows the comparison of the UMRB aggregated runoff-season temperature and water-year precipitation anomalies from the PRISM and nCLIMDIV datasets. Though the PRISM dataset shows slightly more warming and wetting over the period of record, the differences in trend are not statistically significant, with 95% confidence bounds encapsulating each other at all points in time. Strong fidelity between interannual to multidecadal anomalies in basin-wide precipitation and temperature is also evident with correlations between the two datasets of r = 0.96 and r = 0.98 respectively, with p < 0.001 in both cases.

We also assessed the influence of precipitation and temperature on streamflow using the nCLIMDIV dataset, as was done using PRISM (Fig. S11). Compared with PRISM, the nCLIMDIV data highlight a slightly greater influence of temperature on the low flows of the Dustbowl drought, in particular the driest years of that drought. The nCLIMDIV data also highlight the importance of precipitation in driving the lowest flows of the Turn-of-the-Century Drought, while demonstrating the sustained shift to negative temperature forcing of flow since 1984 and the importance of temperature in sustaining the Turn-of-the-Century drought through the second half of the 2000s, when UMRB precipitation was actually above its long-term trend line (Figs. S7, S10, & S11).

Finally, period comparisons of runoff efficiency based on PRISM were repeated using the nCLIMDIV dataset and are shown in Figure S12. In all cases, the distributions of temperature and precipitation are significantly different across time with later years of the 1984-2010 period and the Turn-of-the-Century Drought displaying higher temperatures and/or lower runoff efficiency than earlier periods ($p < 0.029$, all cases).

S4. Assessing the role of temperature, precipitation and evapotranspiration in driving trends in UMRB streamflow using a monthly water balance model

A simplified model for the generation of streamflow (Q) in the UMRB can be expressed as Q=P-ET, where P reflects precipitation and ET reflects evapotranspiration. Using PRISM precipitation anomalies and natural flow anomalies, we described decreasing UMRB runoff efficiency (RE) in recent decades (Fig. 5 a-c) where RE is expressed as Q-P. Given this simplified framework, decreasing RE implies increasing ET relative to P and establishes those increases in ET as a likely contributor to recent droughts in the basin. However, increases in ET implied by decreasing RE cannot be independently verified by analyses based on streamflow and precipitation data alone. In order to provide a process-based assessment of ET over the UMRB we used a well-verified monthly water balance model (MWBM) (14) to estimate ET and Q from UMRB precipitation and temperature data for comparison with those results based only on PRISM precipitation and natural streamflow data.

All MWBM records considered here are derived from aggregated precipitation and temperature data for the 36 HU 8 level watersheds listed in table S2, averaged first by sub-basin and water-year, and then by water-year for the entire basin. This reflects the same averaging approach used to generate the basin-wide estimate of natural streamflow from the 17 gages depicted in Figure 1. Figure S13 shows the model input data (precipitation and temperature) along with model output (ET and runoff). For comparison with MWBM runoff, runoff modeled by multiple linear regression (MLR) using water-year temperature, precipitation, and snow water equivalent (SWE) estimates are shown along with the basin-wide natural flow estimate. Both MWBM and MLR estimates were generated from raw PRISM data as well as U700 adjusted PRISM precipitation data (see supplemental text S3 for details

on U700 adjustment). For this reason, the comparisons shown in figure S13 are limited to the timeframe from 1950-2010, over which estimates of U700 wind speeds are available.

Since we expected increases in ET to be associated with decreases in RE as described above, we also directly examined the relationship between the two variables over time (Fig. 5, d and e; Fig S14). Here again, both raw PRISM and U700 adjusted PRISM precipitation data were used. To make this analysis comparable in time to that shown in Fig 5, a-c of the main text, we made two different comparisons of ET and RE in the URMB over time which are depicted in Figure S14. The first compares water-year values during the Dust Bowl drought to those during the Turn-of-the-Century Drought (Fig 5, d; Fig S14, a). The second compares the water-year values from 1906-1983 to those from 1984-2010 (Fig 5, e and Fig S14, b). Here it should be noted that because the analysis extends prior to 1950, U700 adjustment of PRISM precipitation was uncertain from 1906-1949. For these years, we used the level of adjustment used in 1950 as a constant level applied to the beginning of the record, which implies no further trend in U700 wind speeds back in time prior to 1950.

Several important observations in climate, modeled runoff, and natural streamflow trends are notable in Figure S13. First, a negative trend in natural streamflow (panels f,l) is evident over the 1950-2010 period, and runoff modeled by MLR and the MWBM also exhibit negative trends regardless of whether raw or U700-adjusted PRISM precipitation data are used (panels d,e,j,k). Over the same period, precipitation exhibits a slight positive trend (raw PRISM) (panel a), or no trend (U700 adjusted PRISM) (panel g), although all basin-wide trends in precipitation are insignificant at the 95% confidence level. Temperature and ET, on the other hand, exhibit significant positive trends over the period (panels b,c,h,i). Because runoff within the MWBM is roughly estimated as Q=P-ET (see McCabe and Wolock, 2011 for details), negative trends in runoff associated with a positive trend, or lack of trend in precipitation implies increased ET over time as estimated by the model.

Increasing ET in the UMRB is shown to be negatively associated with RE (Fig 5, d and e), and this is true for the U700 adjusted PRISM dataset as well (Fig. S14). Given the relationships discussed above, it follows that the lower runoff efficiencies observed in natural streamflow since 1984, and especially during the Turn-of-the-Century Drought, should be associated with increased ET relative to earlier periods. Kernel density estimates for the distributions of ET and RE based on MWBM output (Fig 5 d and e margins; Fig. S14 margins) reaffirm the tendency towards increased ET and decreased RE over the 20th and early 21st century implied by the estimates of RE based on precipitation and natural flow data alone (Figure 5, a-c). When assessed at the level of the individual HU 8 watersheds of the UMRB (Table S1), modeled runoff efficiency was significantly lower and modeled ET was significantly higher during the Turn-of-the-Century Drought relative to the Dust Bowl ($p<0.026$ all cases), and since 1984 relative to the 1906-1983 period ($p<0.001$ all cases).

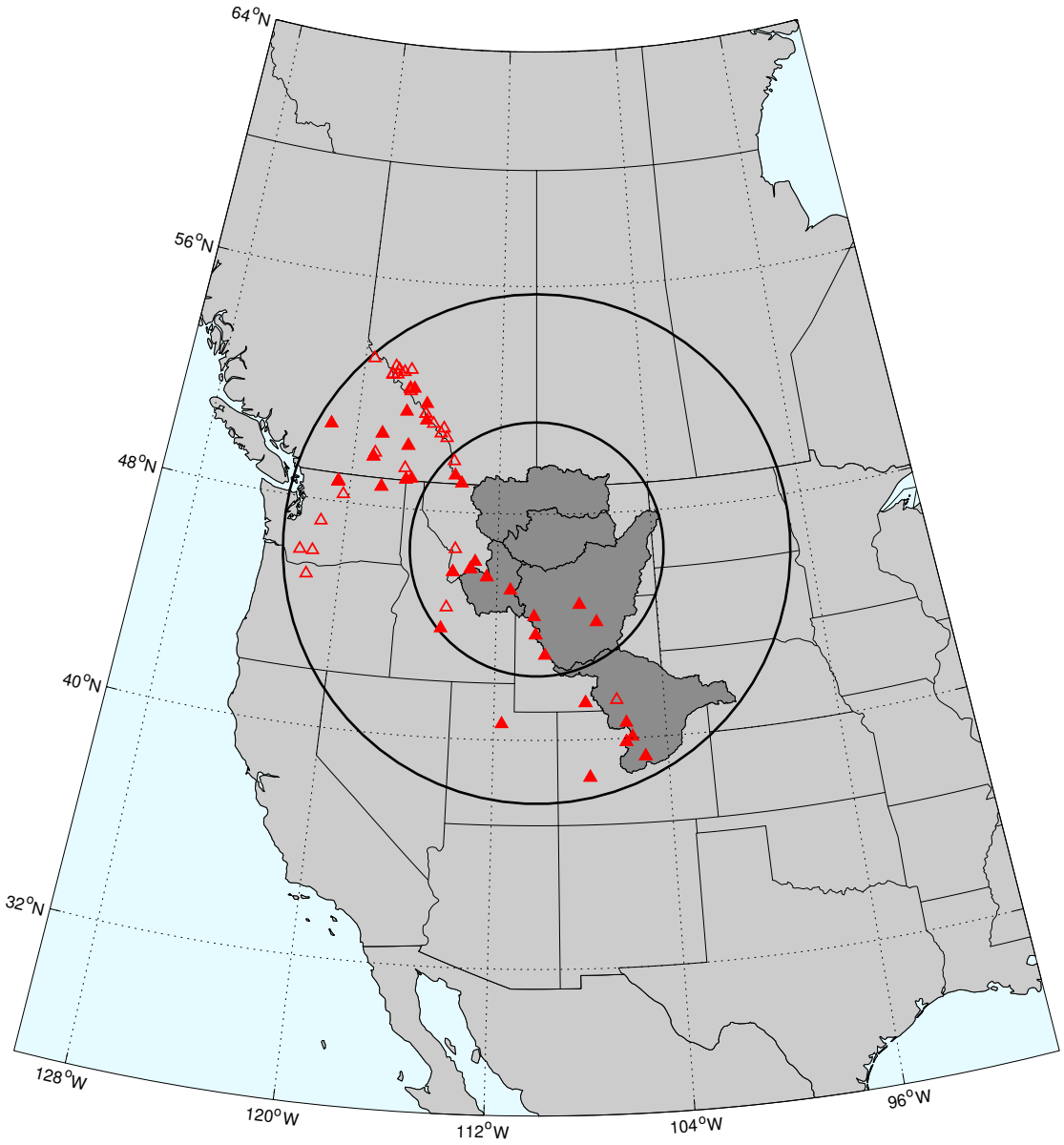

Fig. S1. Temperature reconstruction tree-ring records. The locations of high-elevation, ring-width and maximum latewood density (MXD) chronologies that are within the 1000 km UMRB search radius (outer circle), and are significantly ($p < 0.05$) positively correlated with runoff-season temperature. 500 km radius (inner circle) shown for reference. Filled symbols indicate chronologies that passed record screening and entered into the temperature reconstruction.

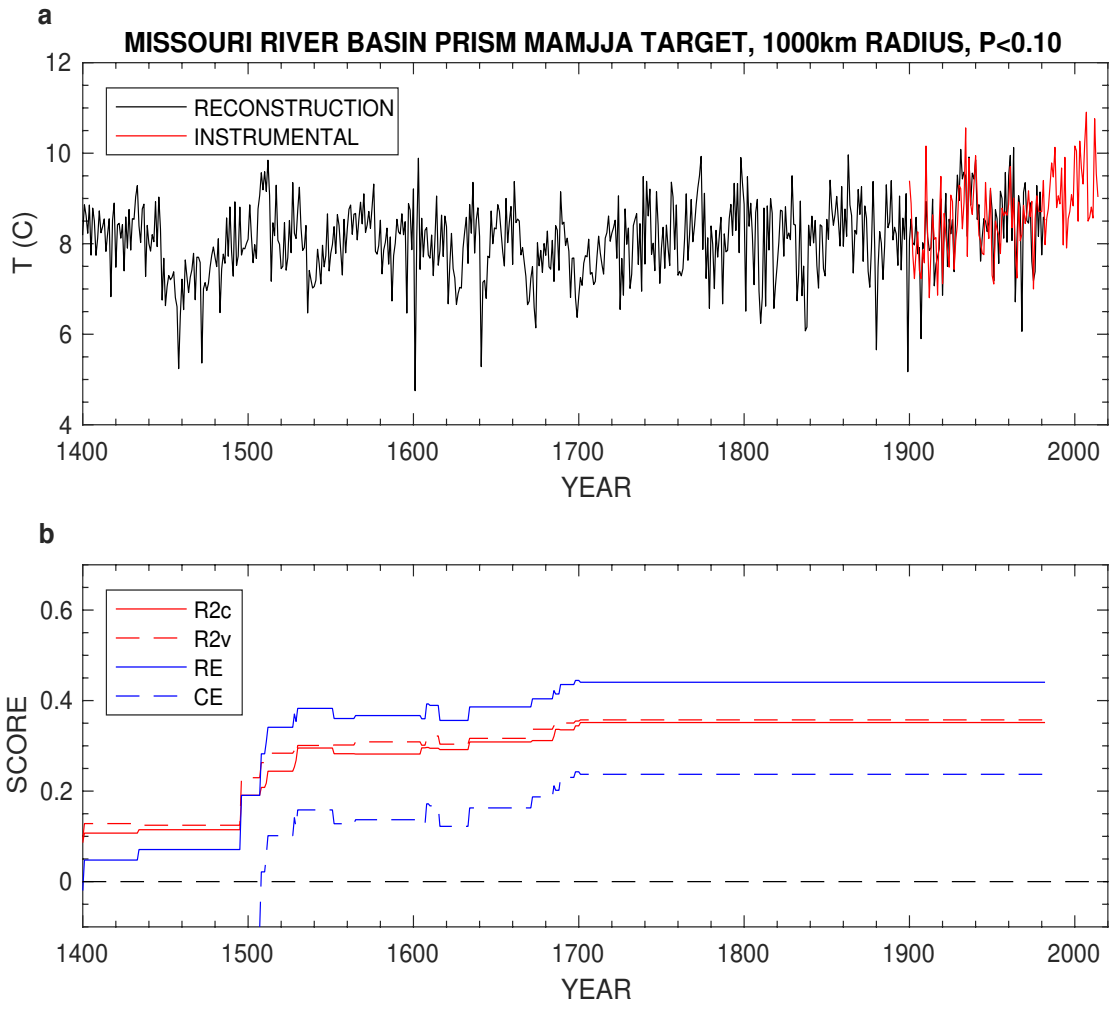

Fig. S2. UMRB runoff-season temperature reconstruction. Cross-validation statistics for the UMRB runoff-season (Mar-Aug) temperature reconstruction showing **(a)** the reconstruction (black) and target temperature record (red), and **(b)** validation R^2 (R2v), calibration R^2 (R2c), reduction of error statistic (RE), and coefficient of efficiency (CE).

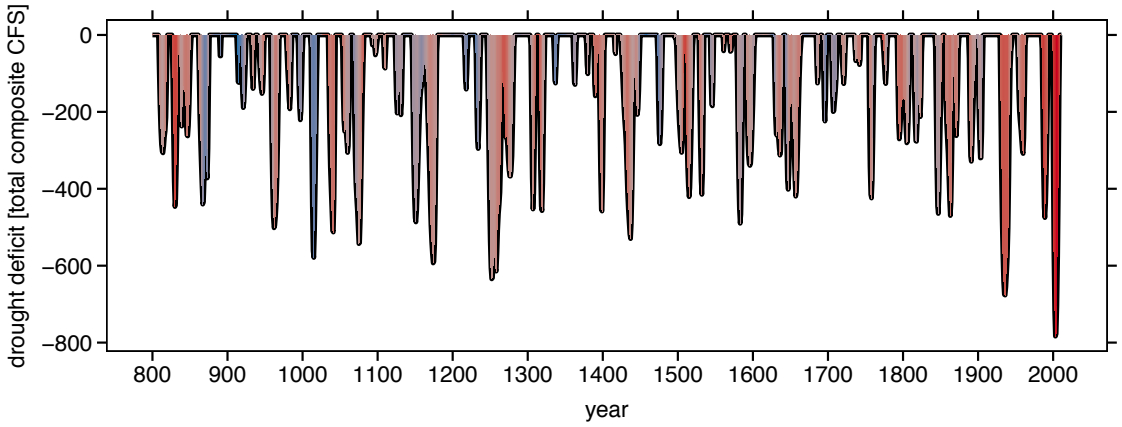

Fig. S3. UMRB un-standardized droughts. The basin-wide drought deficit reconstruction generated by averaging un-standardized streamflow values for the 17 gages in the UMRB composite record.

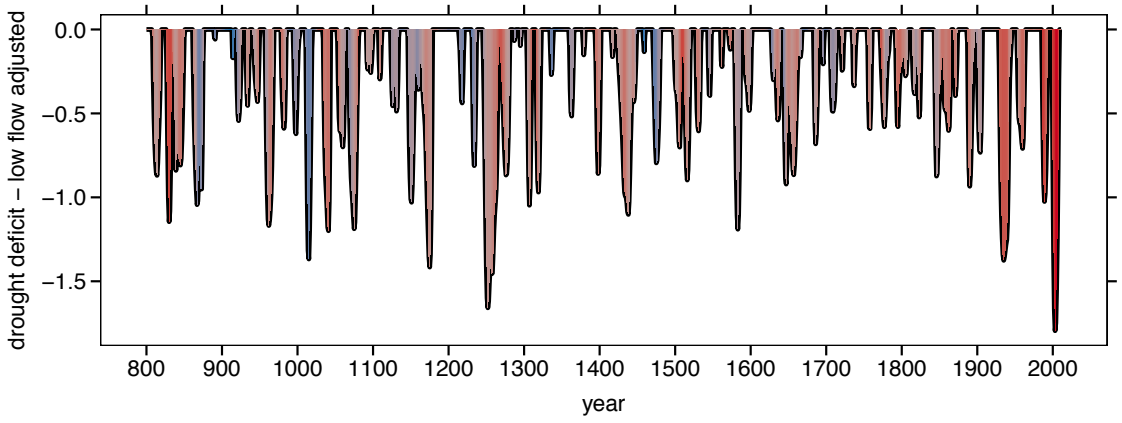

Fig. S4. UMRB droughts from low-flow aligned observed and reconstructed records. The basin-wide drought deficit reconstruction generated from the UMRB composite record in which reconstructed and observed flows are aligned based on their respective means over only the driest years of the calibration period.

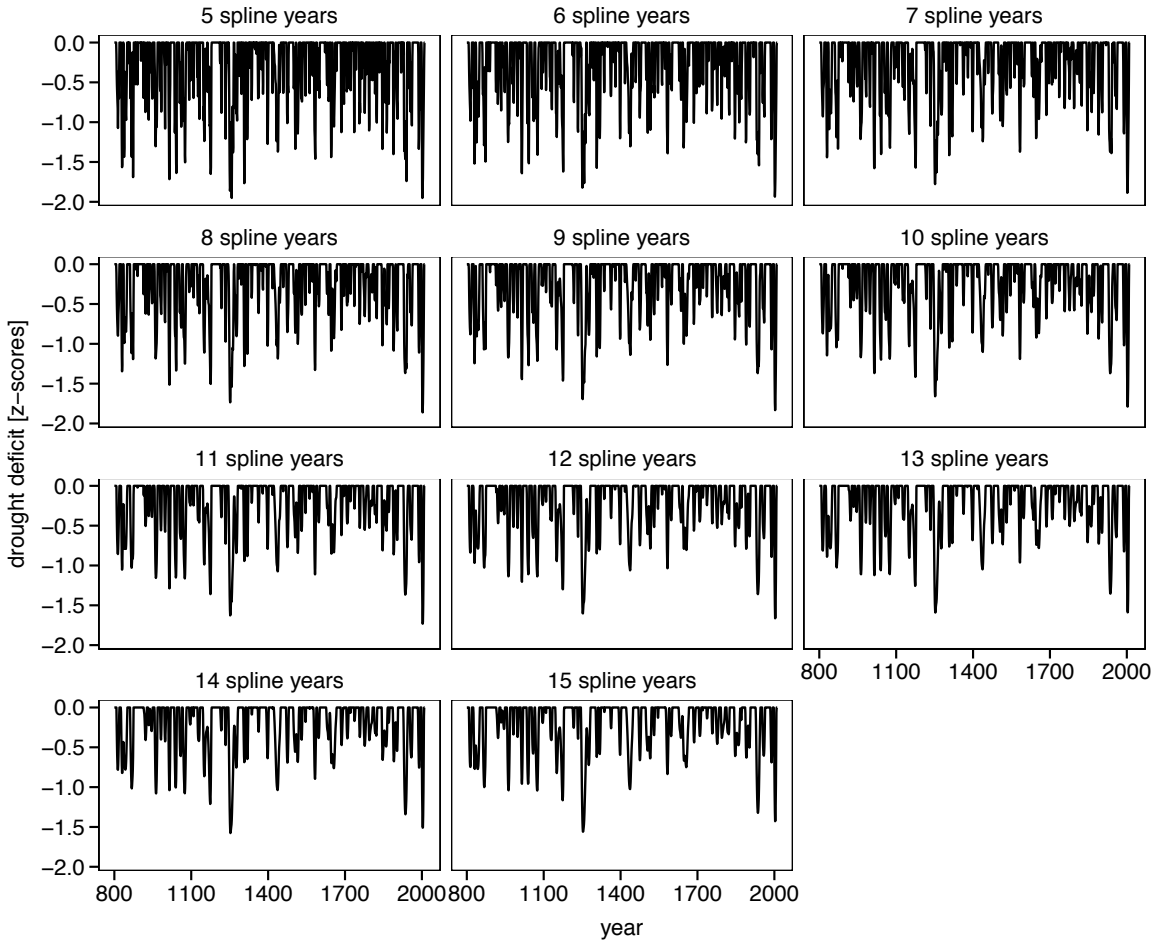

Fig. S5. Five to Fifteen-year UMRB drought splines. The basin-wide drought deficit reconstructions generated from splines lengths ranging from 5 to 15 years.

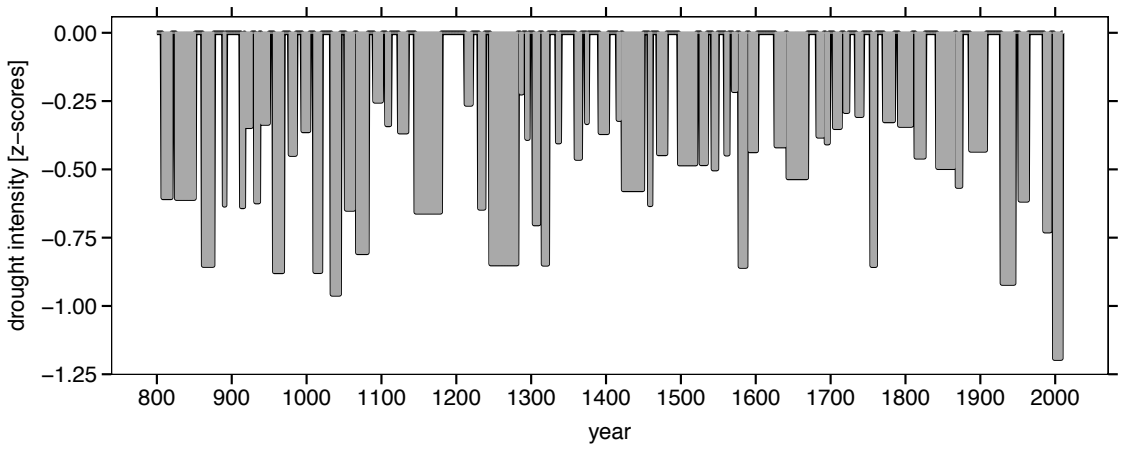

Fig. S6. UMRB drought intensity. The basin-wide decadal drought intensity record generated from droughts defined by the 10-year spline of streamflow and calculated as the cumulative deficit of the drought divided by the duration of the drought in units of z-scores.

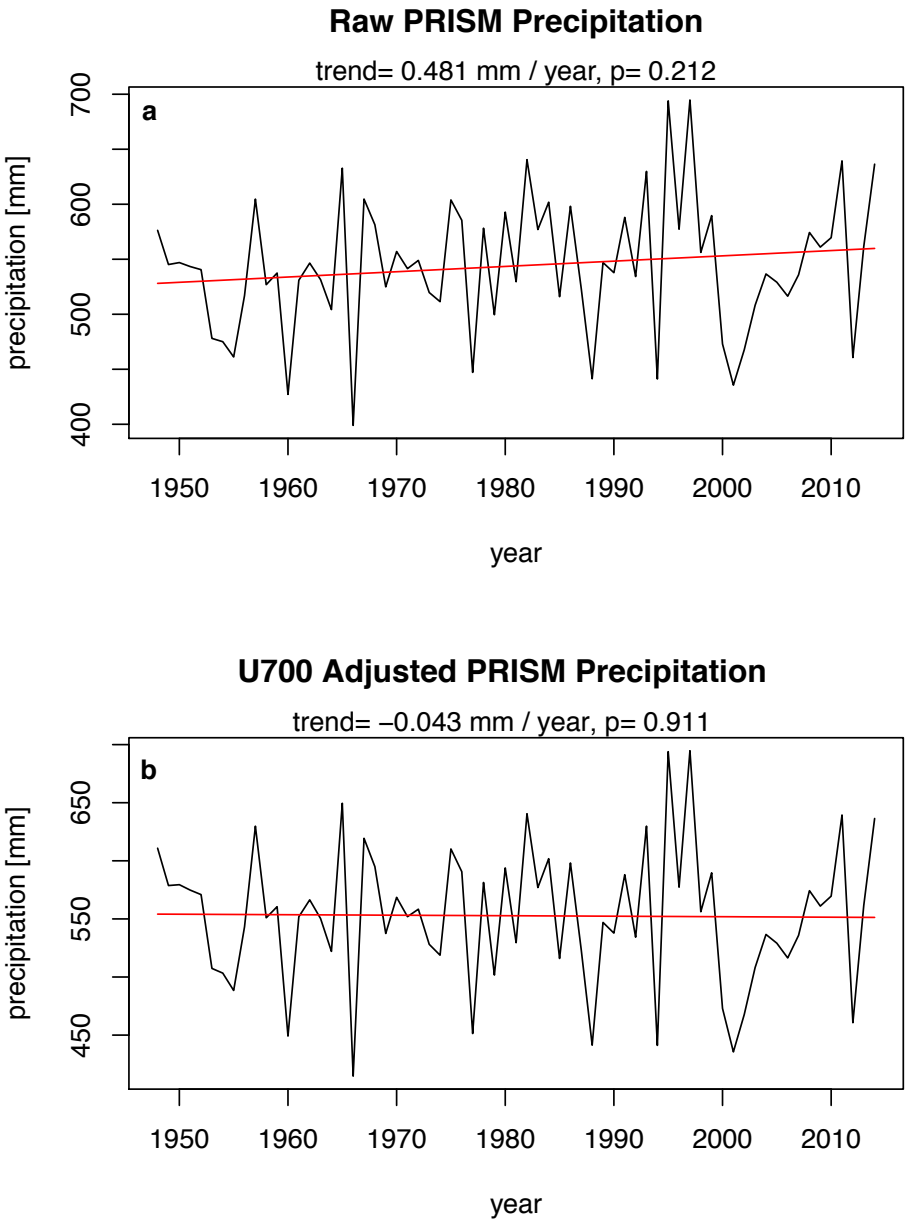

Fig. S7. Raw and U700-adjusted precipitation data. (**a**) Average total annual precipitation for PRISM grids within the UMRB derived from raw PRISM data, and (**b**) PRISM data adjusted for a hypothetical decrease in orographic enhancement of precipitation during the later 20th century.

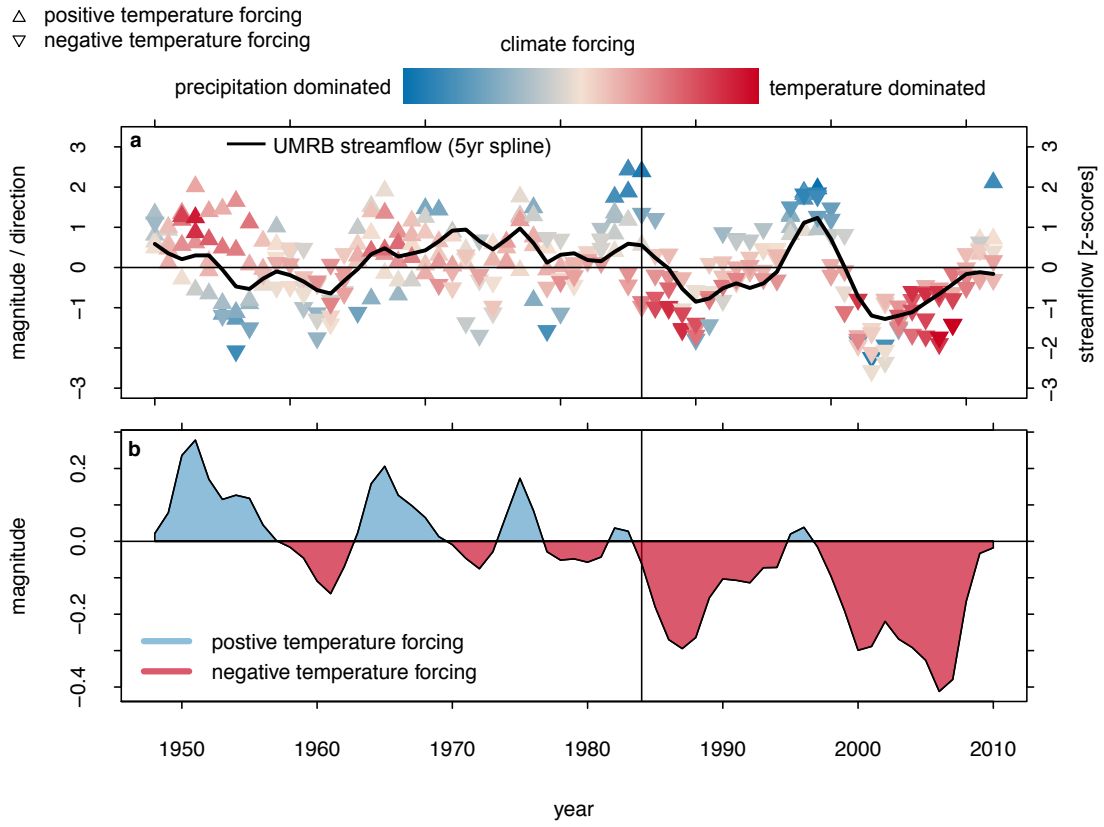

Fig. S8. Climate forcing of streamflow derived from U700-adjusted PRISM precipitation data. (a) The relative forcing of precipitation and temperature (arrows) on basin-wide mean annual streamflow (black line) in the UMRB. Colored arrows show the individual relative forcing estimates for each sub-region of the UMRB for each year. Colors denote which climate variable was more dominant in the combined forcing of streamflow relative to the long-term average influence of each variable. The y-axis shows the relative magnitude and direction of that combined forcing. The direction of the arrows shows the direction of forcing of the temperature component of that combined forcing (supporting or suppressing streamflow). All data shown are derived from the 5-year cubic smoothing splines of streamflow and climate data. **(b)** The relative forcing of temperature on streamflow determined as the temperature anomaly times its multiple-regression coefficient for predicting streamflow along with precipitation. The black line denotes the mean temperature forcing of temperature on streamflow for all sub-regions of the UMRB.

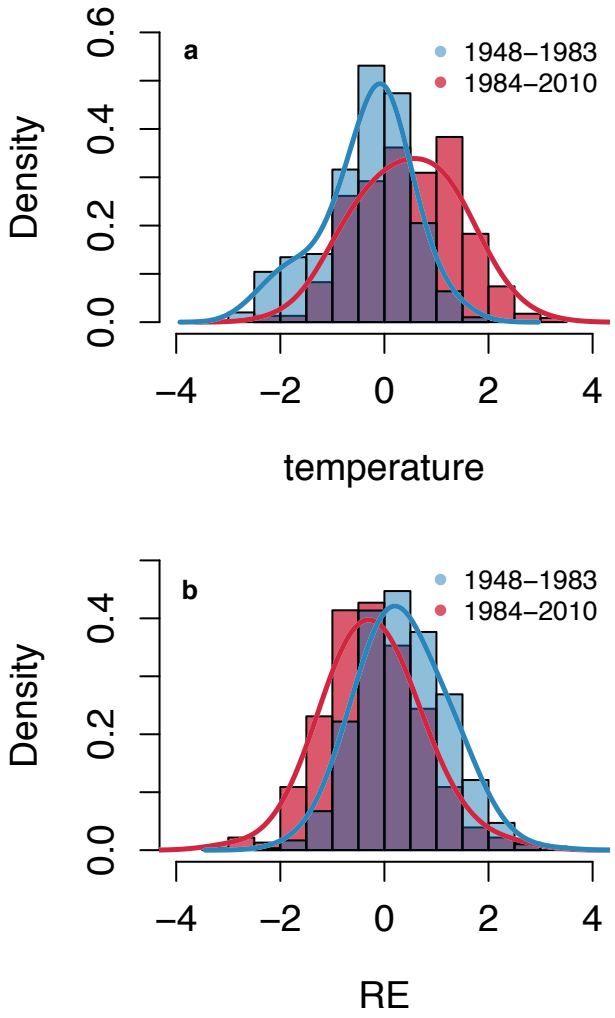

Fig. S9. UMRB basin-wide temperature and runoff efficiency derived from U700-adjusted PRISM precipitation data. (a) Distributions of temperature and **(b)** runoff efficiency from 1900-1983 (blue) and 1984-2010 (red). Lines show the kernel density estimates of the distributions.

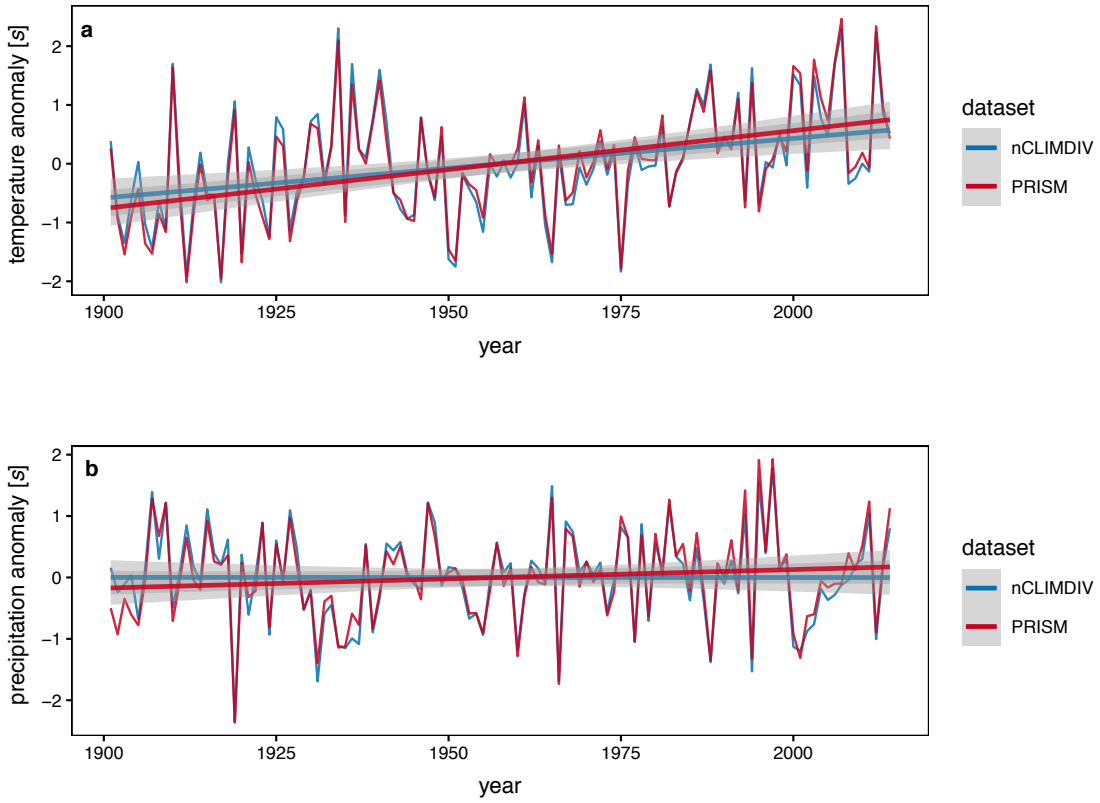

Fig. S10. UMRB basin-wide climate records. Comparison of (**a**) basin-wide average March-August temperature and (**b**) total annual precipitation anomalies in the UMRB derived from the PRISM and nCLIMDIV datasets. Least-squares regression trend lines are shown bracketed by their 95% confidence intervals (gray shading).

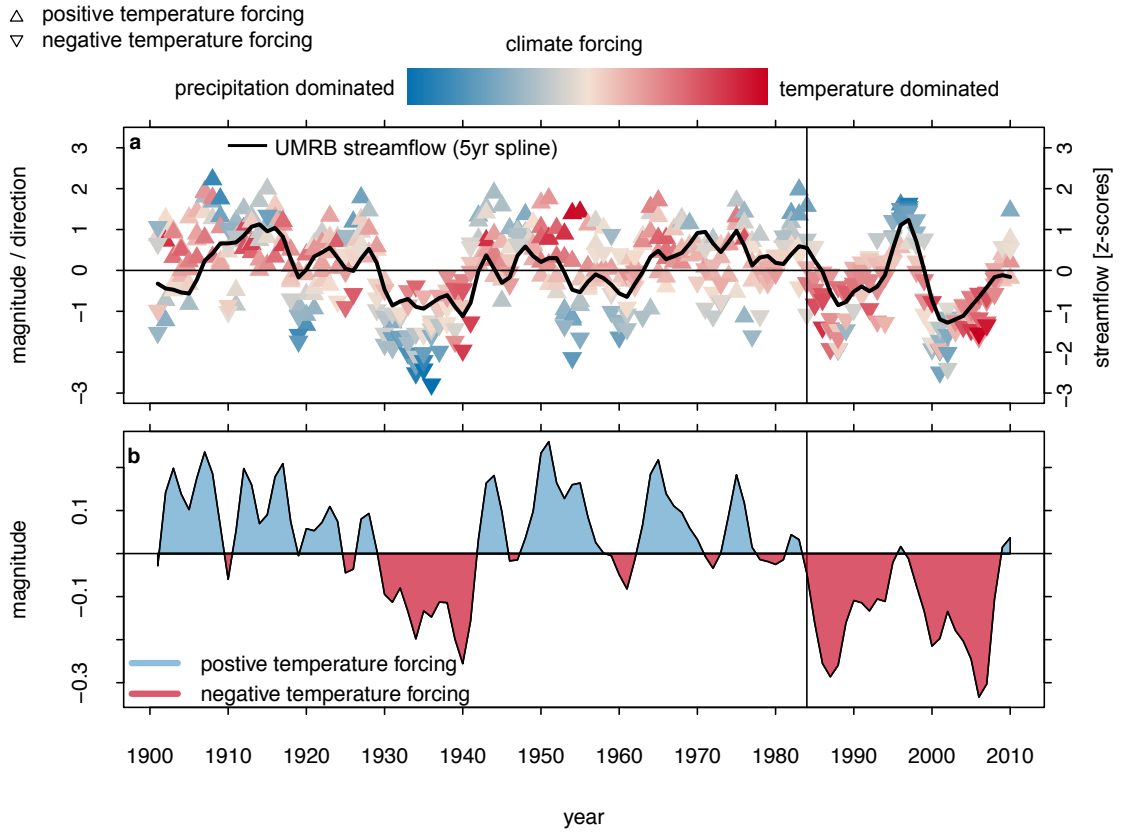

Fig. S11. Climate forcing of streamflow derived from the nCLIMDIV dataset. (a) The relative forcing of precipitation and temperature (arrows) on basin-wide mean annual streamflow (black line) in the UMRB. Colored arrows show the individual relative forcing estimates for each sub-region of the UMRB for each year. Colors denote which climate variable was more dominant in the combined forcing of streamflow relative to the long-term average influence of each variable. The y-axis shows the relative magnitude and direction of that combined forcing. The direction of the arrows shows the direction of forcing of the temperature component of that combined forcing (supporting or suppressing streamflow). All data shown are derived from the 5-year cubic smoothing splines of streamflow and climate data. **(b)** The relative forcing of temperature on streamflow determined as the temperature anomaly times its multiple-regression coefficient for predicting streamflow along with precipitation. The black line denotes the mean temperature forcing of temperature on streamflow for all sub-regions of the UMRB.

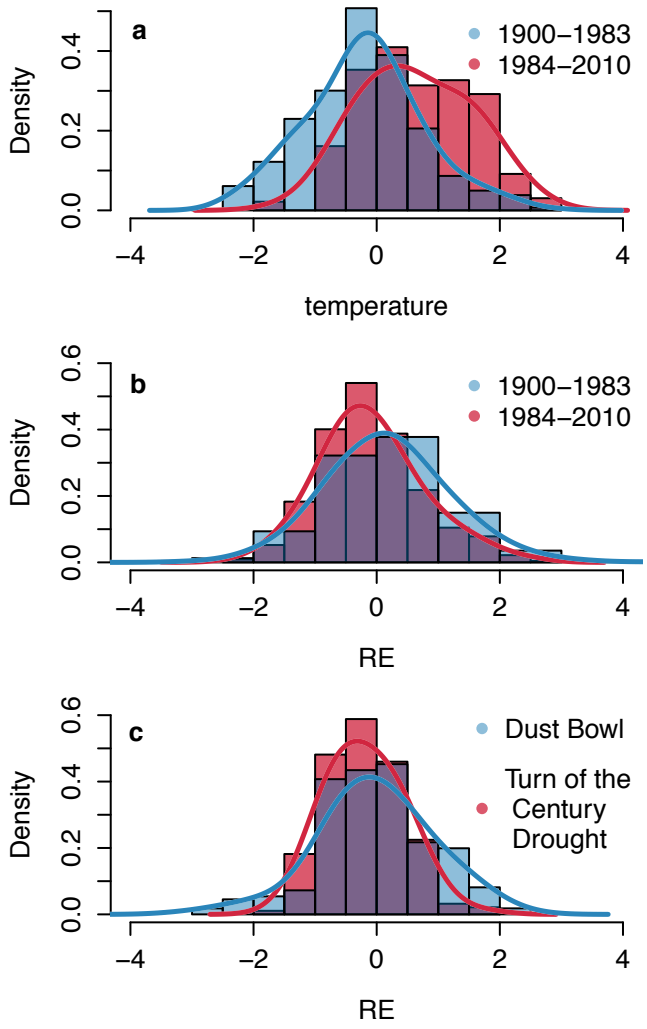

Fig. S12. UMRB basin-wide temperature and runoff efficiency derived from the nCLIMDIV dataset. Distributions of **(a)** temperature and **(b)** runoff efficiency from 1900-1983 (blue) and 1984-2010 (red). Panel **c** shows the distributions of runoff efficiency during the years of the Dust Bowl drought (blue) and Turn-of-the-Century Drought (red). Lines show the kernel density estimates of the distributions.

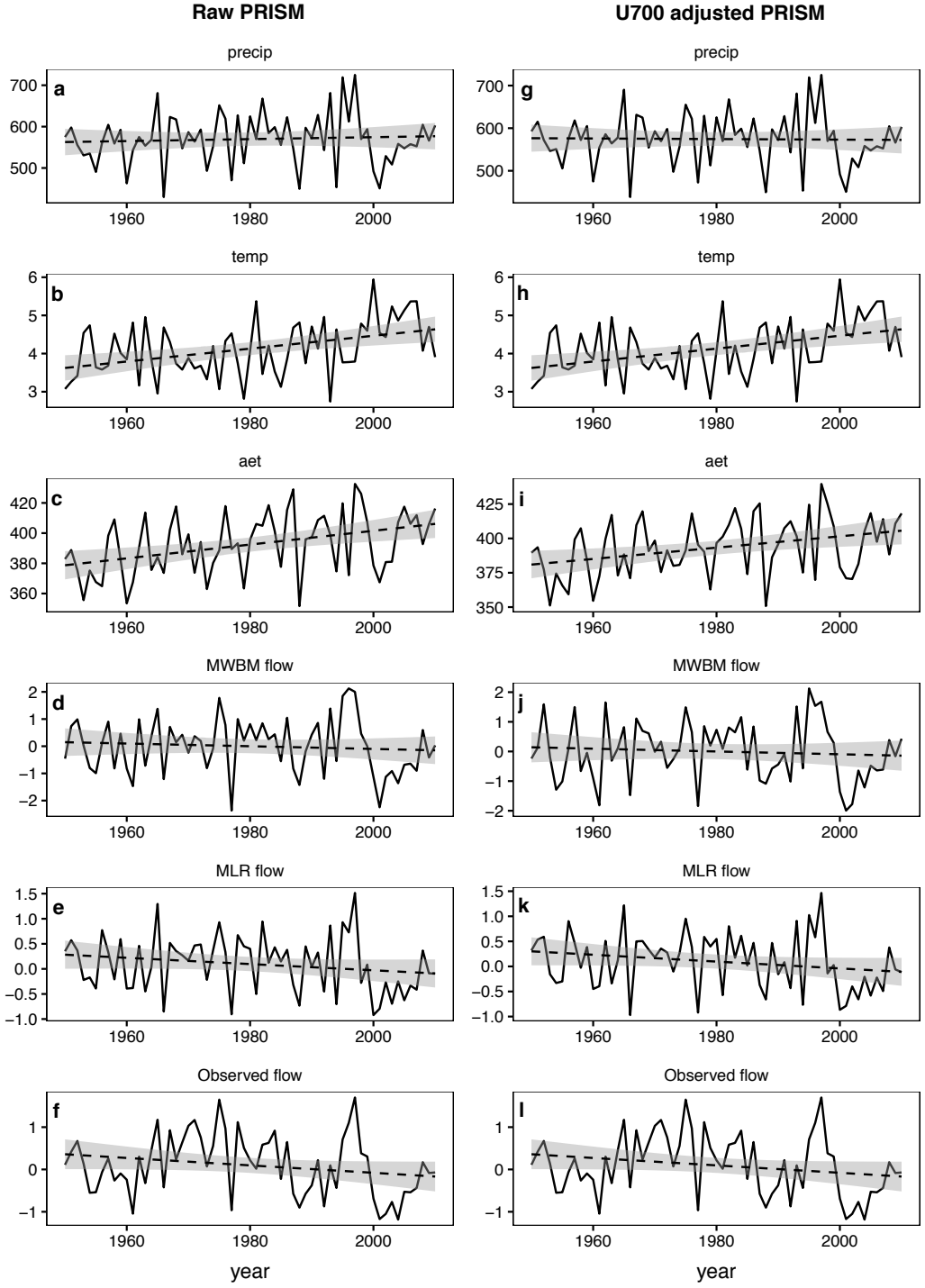

Fig. S13. UMRB basin-wide climate and streamflow records. (**a,g**) Water-year total precipitation [mm], (**b,h**) average temperature [°C], (**c,i**) evapotranspiration [mm], (**d,j**) monthly water balance model runoff [*z*], (**e,f**) linear regression modeled runoff [*z*], and (**f,l**) naturalized streamflow [*z*]. All records except naturalized flows are derived from precipitation and temperature data for the 36 HU 8 level watersheds listed in table S2, averaged by sub-basin and water-year, then by water-year. Data records in the left column use raw PRISM data while records in the right column use U700 adjusted PRISM precipitation data

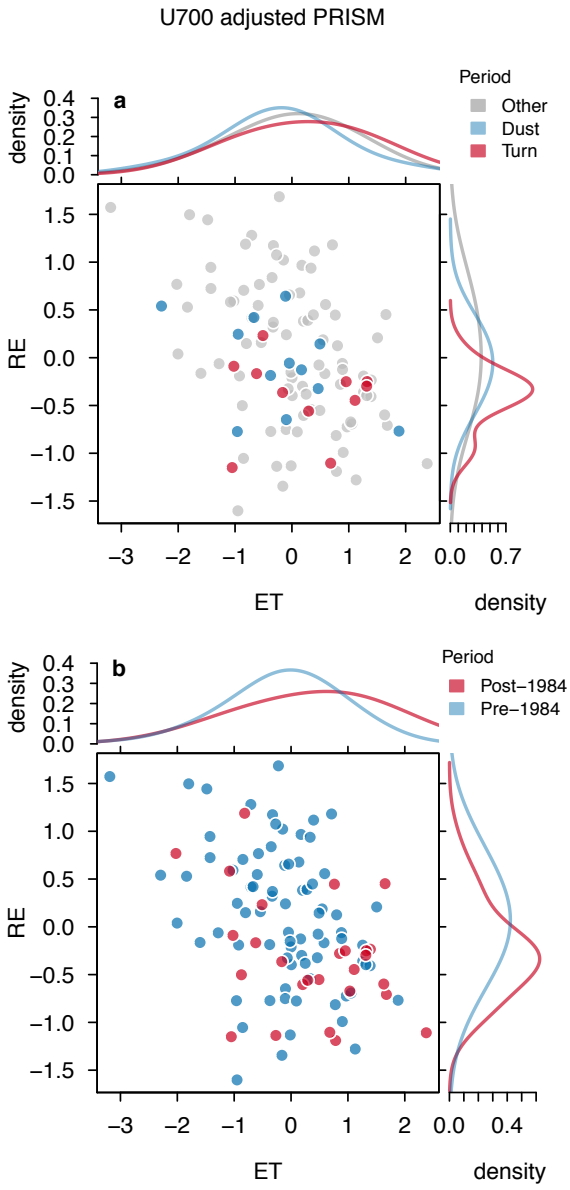

Fig. S14. U700-adjusted, UMRB basin-wide evapotranspiration (ET) versus runoff efficiency (RE). All values shown are derived from U700-adjusted PRISM precipitation and raw PRISM temperature data for the 36 HU 8 level watersheds listed in table S2, averaged by sub-basin and water-year, then by water-year (supplemental text S4). Statistics for the group comparisons ((**a**) Dust Bowl (Dust) vs. Turn-of-the-Century (Turn) Drought; and, (**b**) pre- vs. post-1984) are based on non-aggregated values for each watershed and each water-year within the groups being compared and are reported in supplementary table S1.

comparison	data	variable	effect size (sd)	t	p	n
Post'84-Pre'84	Raw PRISM	ET	0.34	9.14	<0.001	3780
Post'84-Pre'84	Raw PRISM	RE	-0.08	-5.82	<0.001	3780
Post'84-Pre'84	U700-adjusted PRISM	ET	0.27	7.06	<0.001	3780
Post'84-Pre'84	U700-adjusted PRISM	RE	-0.12	-6.58	<0.001	3780
Turn-Dust	Raw PRISM	ET	0.29	4.20	<0.001	828
Turn-Dust	Raw PRISM	RE	-0.05	-2.23	0.026	828
Turn-Dust	U700-adjusted PRISM	ET	0.17	2.34	0.019	828
Turn-Dust	U700-adjusted PRISM	RE	-0.06	-2.31	0.021	828

ET=evapotranspiration
RE=runoff efficiency

Table S1. Statistics for the group comparisons shown in Fig. S14 ((a&c) Dust Bowl (Dust) vs. Turn-of-the-Century (Turn) Drought; and, (b&d) pre- vs. post-1984) are based on non-aggregated values for each watershed and each water-year within the groups being compared.

Cluster	HUC
Missouri Headwaters	10020001
Missouri Headwaters	10020007
Missouri Headwaters	10020008
Missouri Headwaters	10020003
Missouri Headwaters	10020002
Missouri Headwaters	10020005
Missouri Mainstem	10030101
Missouri Mainstem	10030103
Missouri Mainstem	10030105
Missouri Mainstem	10040103
Missouri Mainstem	10040201
Yellowstone	10070001
Yellowstone	10080012
Yellowstone	10080013
Yellowstone	10080009
Yellowstone	10070006
Yellowstone	10070005
Yellowstone	10070002
Yellowstone	10080010
Yellowstone	10080001
Yellowstone	10080002
Yellowstone	10080003
Yellowstone	10080008
Northern Tributaries	10030104
Northern Tributaries	10030201
Northern Tributaries	17010207
Platte	10190001
Platte	10190002
Platte	14010002
Platte	14010001
Platte	10180001
Platte	10190007
Platte	10190006
Platte	10190005
Platte	10190004
Platte	10180002

Table S2. HUC 8 watersheds over which climate data were aggregated to serve as average sub-basin climate records.

SI References

1. Yang D, et al. (1998) Accuracy of NWS 8″ standard nonrecording precipitation gauge: Results and application of WMO intercomparison. *J Atmos Ocean Technol* 15(1):54–68.
2. Clark MP, Slater AG (2006) Probabilistic quantitative precipitation estimation in complex terrrain. *J Hydrometeorol* 7(1):3–22.
3. Groisman PY, Easterling DR (1994) Variability and trends of total precipitation and snowfall over the United States and Canada. *J Clim* 7(1):184–205.
4. Oyler JW, Dobrowski SZ, Ballantyne AP, Klene AE, Running SW (2015) Artificial amplification of warming trends across the mountains of the western United States. *Geophys Res Lett* 42:153–161.
5. McAfee SA, McCabe GJ, Gray ST, Pederson GT (2018) Changing station coverage impacts temperature trends in the Upper Colorado River basin. *Int J Climatol* 39(3):1517–1538.
6. Newman AJ, Clark MP, Longman RJ, Giambelluca TW (2019) Methodological intercomparisons of station-based gridded meteorological products: Utility, limitations, and paths forward. *J Hydrometeorol* 20(3):531–547.
7. Schaefer G, Paetzold R (2000) SNOTEL (SNOwpack TELemetry) and SCAN (soil climate analysis network). *Automated Weather Stations for Applications in Agriculture and Water Resources Management: Current Use and Future Perspectives*, p 30.
8. Vose RS, et al. (2014) Improved historical temperature and precipitation time series for U.S. climate divisions. *J Appl Meteorol Climatol* 53(5):1232–1251.
9. Luce CH, Abatzoglou J, Holden ZA (2013) The Missing Mountain Water: Slower Westerlies Decrease Orographic Enhancement in the Pacific Northwest USA. 342(6164):1360–1364.
10. Doggett M, Daly C, Smith J (2014) Challenges in Mapping Daily Temperature and Precipitation across the Conterminous United States. *AMS100* (Denver, CO).
11. Wise EK, Woodhouse CA, McCabe GJ, Pederson GT, St-Jacques JM (2018) Hydroclimatology of the Missouri River Basin. *J Hydrometeorol* 19(1):161–182.
12. Deems J, Hamlet AF (2010) *Historical Meteorological Driving Data Set* Available at: http://warm.atmos.washington.edu/2860/report/.
13. Oyler JW, Ballantyne A, Jencso K, Sweet M, Running SW (2015) Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. *Int J Climatol* 35(9):2258–2279.
14. McCabe GJ, Wolock DM (2011) Century-scale variability in global annual runoff examined using a water balance model. *Int J Climatol* 31(12):1739–1748.