ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR031575

Variable Streamflow Response to Forest Disturbance in the Western US: A Large-Sample Hydrology Approach

Sara A. Goeking[1,2] and David G. Tarboton[3]

[1]Department of Watershed Sciences, Utah State University, Logan, UT, USA, [2]USDA Forest Service, Rocky Mountain Research Station, Forest Inventory & Analysis Program, Ogden, UT, USA, [3]Utah Water Research Laboratory, Utah State University, Logan, UT, USA

Key Points:

- Large-sample analyses found that while streamflow often increased following forest disturbance, it decreased in some watersheds
- The direction of streamflow response to forest disturbance (increase vs. decrease) is dependent on aridity
- Forest disturbance is more likely to occur in arid locations, which is also where disturbance tends to result in decreased streamflow

Correspondence to:

S. A. Goeking,
sara.goeking@usda.gov

Citation:

Goeking, S. A., & Tarboton, D. G. (2022). Variable streamflow response to forest disturbance in the western US: A large-sample hydrology approach. *Water Resources Research*, *58*, e2021WR031575. https://doi.org/10.1029/2021WR031575

Received 6 NOV 2021
Accepted 25 MAY 2022

Author Contributions:

Conceptualization: Sara A. Goeking, David G. Tarboton
Data curation: Sara A. Goeking, David G. Tarboton
Formal analysis: Sara A. Goeking, David G. Tarboton
Investigation: Sara A. Goeking, David G. Tarboton
Methodology: Sara A. Goeking, David G. Tarboton
Resources: Sara A. Goeking
Supervision: David G. Tarboton
Validation: Sara A. Goeking, David G. Tarboton
Visualization: Sara A. Goeking, David G. Tarboton
Writing – original draft: Sara A. Goeking, David G. Tarboton
Writing – review & editing: Sara A. Goeking, David G. Tarboton

Abstract

Forest cover and streamflow are generally expected to vary inversely because reduced forest cover typically leads to less transpiration and interception. However, recent studies in the western U.S. have found no change or even decreased streamflow following forest disturbance due to drought and insect epidemics. We investigated streamflow response to forest cover change using hydrologic, climatic, and forest data for 159 watersheds in the western U.S. from the CAMELS data set for the period 2000–2019. Forest change and disturbance were quantified in terms of net tree growth (total growth volume minus mortality volume) and mean annual mortality rates, respectively, from the U.S. Forest Service's Forest Inventory and Analysis database. Annual streamflow was analyzed using multiple methods: Mann-Kendall trend analysis, time trend analysis to quantify change not attributable to annual precipitation and temperature, and multiple regression to quantify contributions of climate, mortality, and aridity. Many watersheds exhibited decreased annual streamflow even as forest cover decreased. Time trend analysis identified decreased streamflow not attributable to precipitation and temperature changes in many disturbed watersheds, yet streamflow change was not consistently related to disturbance, suggesting drivers other than disturbance, precipitation, and temperature. Multiple regression analysis indicated that although change in streamflow is significantly related to tree mortality, the direction of this effect depends on aridity. Specifically, forest disturbances in wet, energy-limited watersheds (i.e., where annual potential evapotranspiration [PET] is less than annual precipitation) tended to increase streamflow, while post-disturbance streamflow more frequently decreased in dry water-limited watersheds (where the PET to precipitation ratio exceeds 2.35).

Plain Language Summary

Forest disturbance is typically expected to lead to increased runoff, and therefore more water available for aquatic ecosystems and people, because loss of forest vegetation results in less water being taken up and transpired by plants. We examined streamflow and forest change in 159 watersheds in the western U.S. to test this expectation. We found that not all disturbed watersheds experienced increased streamflow. Very dry watersheds were more likely to produce less runoff following forest disturbance and were also more likely to experience forest disturbance.

1. Introduction

Based on decades of research, forest cover and streamflow are generally expected to vary inversely (Andréassian, 2004; Bosch & Hewlett, 1982; Hibbert, 1967; Troendle, 1983). Such research is based on a combination of paired watershed experiments (e.g., Brown et al., 2005; Moore et al., 2020), post-hoc analysis of streamflow data in unpaired watersheds where streamflow can be modeled as a function of climatic observations (e.g., Biederman et al., 2015; Zhao et al., 2010), and simulation modeling that encompasses various levels of complexity (e.g., Bennett et al., 2018; Buma & Livneh, 2015; Sun et al., 2018). The mechanism behind the inverse relationship between forest cover and streamflow includes a combination of reduced evaporation of canopy-intercepted precipitation, and reduced canopy transpiration following forest cover loss (Adams et al., 2012; Hibbert, 1967; Pugh & Gordon, 2012). Conversely, forest recovery or afforestation are assumed to increase total transpiration and evaporative losses of canopy-intercepted precipitation, thus leading to decreased runoff (Andréassian, 2004; Hibbert, 1967).

Contrary to the hypothesis of an inverse relationship between forest cover and streamflow, observed streamflow changes following recent forest disturbances have been variable in magnitude and direction (Boisramé et al., 2017; Goeking & Tarboton, 2020; Ren et al., 2021; Slinski et al., 2016). Over the past two decades, widespread but low- to moderate-severity forest disturbance has occurred as a result of drought stress, insect epidemics, and disease epidemics, as well as altered wildfire regimes (Adams et al., 2012; Williams et al., 2013), thus providing opportunities to identify circumstances leading to decreased post-disturbance streamflow. Most exceptions to the inverse relationship between forest cover and streamflow occurred as post-disturbance decreases in streamflow, typically at low latitudes and south-facing aspects with high aridity, high incoming solar radiation (SRAD), and/or where tree canopies were replaced by rapid growth of dense grasses or shrubs (Bennett et al., 2018; Goeking & Tarboton, 2020; Guardiola-Claramonte et al., 2011; Morillas et al., 2017; Ren et al., 2021). Even in studies that found conforming streamflow increases following disturbance, the magnitude of streamflow increases was modulated by aridity (Saksa et al., 2019). Although such findings are anomalous in the larger context of decades of forest hydrology research, they highlight alternative hypotheses to the inverse relationship between forest cover and streamflow. One such alternative hypothesis is that although streamflow typically increases following forest disturbance, post-disturbance conditions that lead to increased evaporation (i.e., increased energy at snowpack or soil surface) or increased transpiration (i.e., replacement of sparse trees with dense shrubs) lead to a reduced streamflow response.

While numerous studies of runoff response to forest change have focused on site-specific treatments (e.g., harvest, planting) or severe disturbance (e.g., stand-replacing wildfire, clearcuts) in one or two small watersheds, fewer studies have examined lower severity disturbances across broader geographic areas or across more gradual timescales than episodic timber harvesting or wildfire (Andréassian, 2004; Hallema et al., 2017; Wine et al., 2018). Response to low to moderate severity forest disturbances (generally defined as <70% tree mortality) may fundamentally differ from severe, stand-replacing disturbances due to their different effects on energy balances affecting snowpack and soil moisture as well as different transpiration rates for pre-disturbance vs. post-disturbance vegetation (Adams et al., 2012; Pugh & Gordon, 2012; Reed et al., 2018). Recent tree die-off events spanning western North America have provided the opportunity to examine streamflow responses to disturbance that is less severe but more widespread than the forest changes considered in most previous forest hydrology studies (Adams et al., 2012; Hallema et al., 2017). Studies based on both observations (Biederman et al., 2014, 2015; Guardiola-Claramonte et al., 2011) and simulations (Bennett et al., 2018; Ren et al., 2021) have found unexpected post-disturbance decreases in streamflow. Streamflow response to disturbance at broader scales may not reflect hypotheses developed from study of small watersheds that are commonly the focus of paired watershed experiments (Andréassian, 2004), which underscores the value of broad-scale evaluation of hypotheses that were developed at fine scales.

A challenge in testing such hypotheses is the need to balance breadth with depth, that is, gathering fine-scale observations from individual watersheds vs. coarser observations from many watersheds (Gupta et al., 2014). Large-sample hydrology can complement fine-scale studies of individual small watersheds by identifying broad-scale patterns in streamflow response to forest disturbance. Fine-scale studies have produced useful information about the response of streamflow (e.g., Biederman et al., 2015; Guardiola-Claramonte et al., 2011), snowpack (e.g., Broxton et al., 2016; Moeser et al., 2020), and individual ecohydrological processes to forest change (e.g., Biederman et al., 2014; Reed et al., 2018). In contrast, large-sample hydrology can evaluate hypotheses across many watersheds to identify circumstances that conform to or deviate from hypothesized relationships (Addor et al., 2019; Gupta et al., 2014; Newman et al., 2015). Another challenge is accounting for the effects of climate variability in streamflow assessments, such that the effects of vegetation change on streamflow are not confounded with climate effects. To address this challenge, quantitative models of streamflow response to vegetation change often include precipitation and temperature as explanatory variables (Zhao et al., 2010).

In this study, we used a large sample of catchments to test hypotheses about the direction of runoff response following forest disturbance in semi-arid catchments. Observations consisted of streamflow, vegetation, and climate data, which allowed us to account for streamflow changes related to variability in precipitation and temperature and thus disentangle climate from vegetation effects. Based on previous studies finding exceptions to the inverse relationship between forest cover and streamflow, we developed two alternative hypotheses. First, post-disturbance runoff in catchments conforms with the commonly held paradigm that runoff increases with tree mortality or reductions in net growth. Second, an alternative hypothesis is that in watersheds with higher aridity and incoming SRAD, runoff is more likely to decrease or not change than in watersheds with lower aridity and SRAD. A corollary of this hypothesis is that a threshold of aridity index exists above which disturbance results in a decrease in runoff. Our results find this threshold to be an aridity index of 2.35.

Table 1
The Four Questions Addressed in This Study, the Analytical Framework Used to Address Each Question, and the Variables Included in the Analysis

Question	Analytical framework	Variables analyzed
1) To what extent and where is there a consistent trend in annual Q, Q/P, P, PET, and T, regardless of forest change effects?	Mann-Kendall trend tests (univariate)	Annual Q, Q/P, P, PET, and T
2) To what extent and where do trends in runoff ratio and forest density demonstrate an inverse relationship?	Trend in Q/P vs. net tree growth	Trend (Kendall's Tau) in annual Q/P; net tree growth
3) To what extent has streamflow changed in watersheds with substantial forest disturbance?	Time trend analysis (comparison of observed vs. predicted Q)	Annual Q, P, and T; disturbance (disturbed/not disturbed)
4) How well does the severity of forest disturbance, and the interaction of disturbance severity with aridity, predict change in streamflow?	Multiple regression	Annual Q, P, T; tree mortality; aridity (PET/P)

Note. Q, streamflow; P, precipitation; PET, potential evapotranspiration; T, temperature.

2. Data and Methods

We combined data from the CAMELS large-sample hydrology data set (CAMELS; Addor et al., 2017) and the U.S. Forest Service's Forest Inventory and Analysis (FIA) forest monitoring data set (Bechtold & Patterson, 2005) to answer four questions (Table 1). The ability of each question's analytical framework to disentangle climatic from forest disturbance effects on streamflow successively increases from the first to the fourth question. For analyses that do not explicitly permit such disentangling, we interpret the results in the context of factors that were not included in the analysis.

2.1. Data Sources

2.1.1. Streamflow and Climate Data

Watersheds were selected from the CAMELS data set, which was compiled for watersheds that have little or no known land-use change and whose streamflow is relatively unimpacted by storage or diversions (Addor et al., 2017). However, watersheds in the CAMELS data set have been subject to disturbance from wildfire and other causes of tree mortality that have been quantified by FIA. From the entire CAMELS data set, we first constrained our analysis to watersheds in the western U.S. for which we could obtain estimates of forest characteristics from the FIA data set. Then we removed watersheds where runoff ratio was calculated as larger than 1.0 (runoff greater than precipitation) in any 1 yr, which indicates an impossible water budget and where data is presumed to be in error. Precipitation and streamflow data within the CAMELS data set were derived from Daymet climate data and USGS streamflow gages, respectively (Addor et al., 2017), and these separate data sources do not impose constraints of water budget closure. While we recognize that some catchments may have runoff ratios greater than 1.0, for example, in volcanic or karst landscapes, and that runoff ratios near but less than 1.0 may be similarly implausible, we had no means of quantifying realistically vs. unrealistically high runoff ratios. These constraints yielded 159 watersheds, out of 211 candidate watersheds as 52 (25%) had runoff ratio greater than 1.0. The fact that 25% of watersheds had runoff ratios greater than 1.0 is indicative of the uncertainty and difficulty in compiling quality controlled data over large samples, even for curated data sets such as CAMELS. The watersheds selected had a wide range of physical and land cover characteristics (Table 2), runoff

Table 2
Characteristics of 159 Watersheds Used in This Study

	Area (km^2)	Mean slope (m/km)	Mean elevation (m)	Runoff ratio	P (mm/yr)	PET (mm/yr)	Fraction forested
Median	238	92.8	1,613	0.419	822	1,084	0.76
Mean	649	92.0	1,650	0.409	1,062	1,088	0.64
Standard deviation	1,454	35.3	882	0.241	674	206	0.34

Note. Values are summarized from CAMELS attributes (Addor et al., 2017).

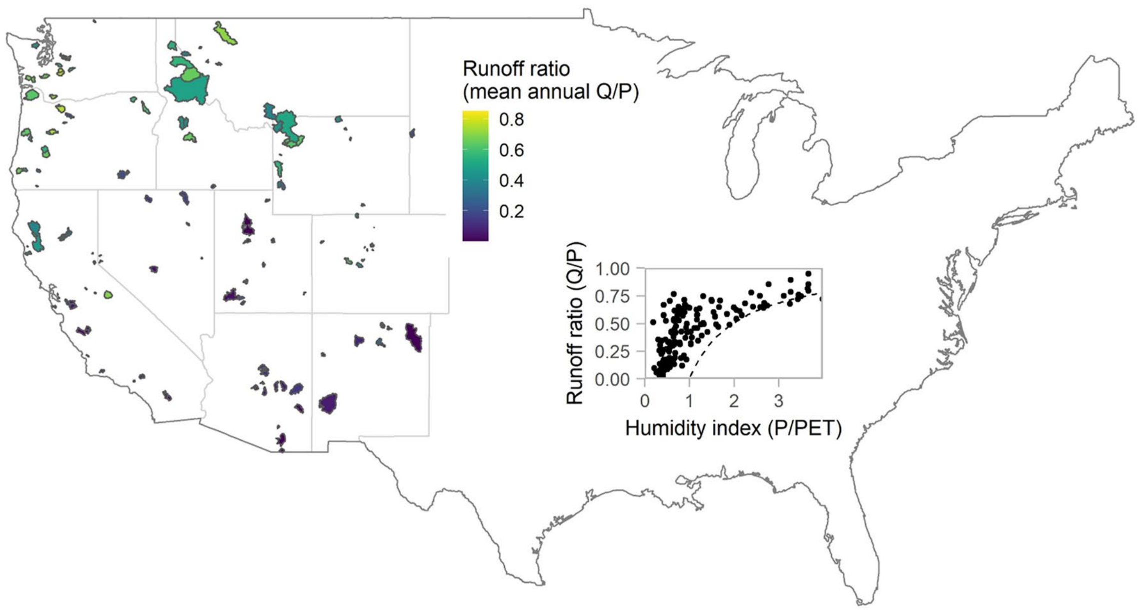

Figure 1. Watersheds from the CAMELS database used in our analyses ($n = 159$). Inset plot shows watersheds in nondimensional space based on long-term CAMELS attributes; the dashed curve represents energy limitation on streamflow, expressed as $Q = P - \text{PET}$ framed in terms of the dimensionless axes as $Q/P = 1 - 1/(P/\text{PET})$, where Q = annual streamflow, P = annual precipitation, and PET = annual potential evapotranspiration.

ratios, and humidity indices (Figure 1), giving the study a broad degree of generality. Given the criteria for inclusion in the CAMELS data set (Addor et al., 2017), we assumed that stream gauges for each watershed quantify actual runoff, and that withdrawals, transfers, and changes in storage are negligible.

The CAMELS data set includes daily time series of climatic variables and streamflow as well as time-averaged catchment characteristics. We used temporally averaged variables representing basin characteristics such as mean incoming SRAD, and aridity, defined as the ratio of mean annual potential evapotranspiration (PET) to mean annual precipitation, all from the CAMELS data set (Addor et al., 2017). We summed CAMELS daily streamflow and precipitation values to get total annual water year streamflow and precipitation. Annual mean temperature was calculated by first averaging CAMELS minimum and maximum daily temperature to get daily mean temperature and then averaging the daily mean temperature. Additionally, we estimated annual PET by first using the Hamon method (Hamon, 1963; Lu et al., 2005) to estimate daily PET based on precipitation, temperature, and day length from the CAMELS data set, and then aggregating daily values to annual PET.

Because the CAMELS data set extends only through water year 2014, while available forest data extend through 2019, we used USGS streamflow data and Daymet gridded climate data for water years 2015–2019 to extend the record of our analysis through water year 2019. USGS streamflow data were obtained through the R package *DataRetrieval* (Hirsch & De Cicco, 2015). Daymet gridded precipitation, minimum temperature, and maximum temperature values were downloaded using the R package *daymetr* (Hufkens et al., 2018) and extracted as area-weighted averages within each CAMELS catchment boundary, following the methods used to construct the CAMELS time series (Newman et al., 2015). That extraction process yielded time series analogous to the time series within the CAMELS data set. We then aggregated daily values to annual values in the same manner as described above for the CAMELS time series. We cross checked our extended data set by ensuring that we could replicate water year 2014 in the CAMELS data, finding that the only differences were due to numerical rounding.

2.1.2. Forest and Disturbance Data

Data on forest conditions and disturbances were obtained from the U.S. Forest Service's FIA program. The FIA program established plot locations using probabilistic sampling to obtain a representative sample with mean spacing of 5 km across all forest types and owner groups (Bechtold & Patterson, 2005). In the western U.S., 10% of plots are measured each year and each plot is therefore measured once every 10 yr. Each year's subsample of plots is spatially distributed such that the sample of forest conditions is both spatially and temporally balanced.

This sampling design was developed to produce unbiased estimates of forest attributes that represent discrete areas such as watersheds (Bechtold & Patterson, 2005).

Data collected from FIA plots include detailed tree measurements that permit calculation of plot-level volume of both live and dead trees, volume of net tree growth, volume of trees that recently died (i.e., “mortality trees”), and many other variables (USDA, 2010). Each plot is associated with an expansion factor that facilitates estimation of forest characteristics and their associated sampling errors for discrete areas, based on data from multiple plots over the same sampling period (Bechtold & Patterson, 2005; Burrill et al., 2018). FIA estimates are updated annually based on a 10 yr moving window such that the estimate in any 1 yr is based on data collected during the previous 10 yr (e.g., an estimate with a nominal date of 2019 is based on data collected during 2010–2019). FIA implemented this nationally consistent, probabilistic sample in 2000, although the onset of data collection varied among states, with Wyoming being the last state to fully implement this design in 2011.

We characterized forest disturbance using FIA's estimates of net tree growth and tree mortality and their associated standard errors, for the period 2010–2019, from the publicly accessible EVALIDator tool (USDA, 2020). Each estimate was constrained to a watershed represented by an 8-digit Hydrologic Unit Code (HUC8) that contains a CAMELS catchment. Although ideally, we would have produced FIA estimates at the scale of CAMELS watersheds, these smaller watersheds contained small sample sizes of FIA plots and thus were associated with high uncertainty at the CAMELS scale. The forested portions of most HUC8 catchments exist at relatively high elevations that tend to be less impacted by water transfers and human activities (i.e., nonforest land uses), which is also where CAMELS watersheds occur (Addor et al., 2017). To test whether forest conditions in CAMELS vs. HUC8 watersheds were similar, we computed the percentage of area at each scale that experienced forest change between 2001 and 2019 as determined from the National Land Cover Database change product (Homer et al., 2020). We found that the distributions of forest change at the two scales were not significantly different based on $p = 0.51$ from the Kolmogorov-Smirnov test for equal distributions. This result supports the use of FIA data at the HUC8 scale as representative of CAMELS watersheds.

Mean annual net growth and mortality rates are expressed as volume per year (Burrill et al., 2018) rather than numbers of trees because under normal conditions with no disturbance, small trees typically die at higher rates than larger or older trees due to self-thinning that occurs naturally as forest stands develop over time (Reineke, 1933; Yoda et al., 1963). Net growth is defined as volumetric growth of all live trees minus the total volume of trees that died in the previous 10 yr (i.e., mortality volume). Values of net growth greater than zero indicate that tree growth has outpaced mortality, while negative net growth is indicative of mortality that occurred faster than growth of live trees. To assess the severity of forest disturbance, we estimated each watershed's mean annual mortality rate and standardized that rate by the total of live volume plus mortality volume. Note that watersheds with high mean annual mortality can also have positive net growth if post-disturbance recovery and live tree growth occurs more rapidly than mortality. A strength of using net growth and mortality estimates is that it permits assessment of quantitative relationships between forest conditions and hydrologic variables, as opposed to being limited by categorical mapping of disturbance or rules-of-thumb such as having >20% of area affected (Goeking & Tarboton, 2020).

2.2. Methods

We used multiple analytical methods to address our objectives. First, we used trend analysis to identify monotonic trends in individual water budget components and drivers. Second, we qualitatively related trends in runoff ratio to forest change across gradients of latitude and aridity. Third, we used time trend analysis (Zhao et al., 2010) to quantify the magnitude of streamflow change that cannot be attributed to precipitation and temperature drivers, and then correlated the magnitude of unattributed streamflow change with forest disturbance, latitude, SRAD, and aridity. Fourth, we evaluated the relative importance of several factors—including temperature, precipitation, and the interaction of forest disturbance and aridity—for predicting change in streamflow across decades using a multiple regression model.

2.2.1. Trends in Water Budget Components and Drivers

Our first question was whether runoff ratio has changed over time, that is, whether there is any monotonic trend, regardless of climate or forest disturbance effects. We answered this question using the nonparametric Mann-Kendall trend test, which determines whether the central tendency of a variable changes solely as a function of time (Helsel et al., 2020). We tested for trends in annual runoff ratio (Q/P) as well as water budget components and drivers, including annual streamflow (Q), annual total precipitation (P), annual mean temperature (T), and annual PET. Each variable was tested independently of vegetation effects. Each test evaluated two time periods: first, the period 2000–2019, which was the basis for our subsequent analyses of streamflow response to forest disturbance, and second, 1980–2019, for the purpose of determining whether any other long-term trends exist that extend prior to the period covered in our analysis.

Watersheds with significant trends in Q, P, Q/P, T, and PET were identified based on two-sided p-values associated with Kendall's tau (Helsel et al., 2020) evaluated with the *MannKendall* function in the *Kendall* package (McLeod, 2011) for R statistical analysis software (R Core Team, 2020). Two-sided p-values <0.1, which correspond to one-side p-values <0.05, were considered statistically significant.

2.2.2. Runoff Ratio and Forest Density Change

Our second question was whether there is general support for the hypothesis that forest cover is inversely related to annual runoff, across a large sample of watersheds spanning a range of aridity, incoming SRAD, and latitude. Under this hypothesis, we expected that most watersheds that experienced forest cover loss (i.e., disturbance) exhibited increases in runoff ratio, and that watersheds that experienced forest cover gain (i.e., increased tree density in the absence of disturbance) exhibited decreases in runoff ratio. An alternative hypothesis, based on recent observations of decreased streamflow following forest disturbance as summarized by Goeking and Tarboton (2020), is that post-disturbance runoff sometimes decreases in more arid, low-latitude watersheds with higher incoming SRAD.

To characterize watersheds as disturbed vs. undisturbed and as having increased vs. decreased runoff ratio, we determined whether net growth and trend in runoff ratio (Q/P) were each positive or negative for each watershed. Watersheds were characterized as having increased vs. decreased runoff ratio on the basis of Kendall's tau, which allows dimensionless comparison of trends in runoff ratio across watersheds whose runoff ratios may vary widely (Helsel et al., 2020), again using R package *Kendall* (McLeod, 2011).

Net tree growth estimates for 2010–2019 encompass a temporal averaging period beginning in 2000 for plots measured in 2010, and in 2009 for plots measured in 2019, because growth is calculated from individual tree growth representing the 10 yr prior to plot measurement (USDA, 2010). Therefore, we conducted trend analysis for the period 2000–2019, which encompasses the averaging period for FIA plot measurements.

We categorized watersheds into two groups: those that met the expectation that the change in runoff ratio is inversely related to forest cover change (conforming watersheds), and those that did not meet this expectation (nonconforming watersheds). Conforming watersheds included watersheds where tree volume increased (i.e., positive tree growth) and Q/P decreased, as well as those where tree volume decreased (i.e., negative tree growth) and Q/P increased. Similarly, nonconforming watersheds consisted of those where both tree volume and Q/P increased and where both tree volume and Q/P decreased. This categorization resulted in four combinations of change in tree volume and trend in Q/P.

We assessed differences in aridity, SRAD, and latitude among the four categories of conforming and nonconforming watersheds. Aridity was compared among watersheds in the context of evaporative index and aridity index, as defined by Budyko (Budyko & Miller, 1974), to assess whether nonconforming watersheds (i.e., those with forest disturbance and decreased streamflow) were more likely to occur in water-limited watersheds than in energy-limited ones. Evaporative index represents the proportion of precipitation that evaporates, on a mean annual basis, and is equal to the quantity $1 - Q/P$. Aridity index is the ratio of mean annual PET to mean annual P. Long-term values of mean annual Q, mean annual P, aridity, and incoming SRAD for each watershed were obtained from the CAMELS data set (Addor et al., 2017). We also tested for significant differences in latitude, aridity, and SRAD among conforming vs. nonconforming watersheds using the nonparametric Kruskal-Wallis test for multiple comparisons, which was conducted using the function *Kruskal* in R package *agricolae* (de Mendiburu, 2020).

2.2.3. Expected Streamflow Change in Watersheds With and Without Forest Disturbance

To address the question of whether streamflow has changed as a result of forest disturbance over discrete time periods, we used time trend analysis, which is an analytical framework used to quantify streamflow change resulting from vegetation change (Zhao et al., 2010). The premise of time trend analysis is that expected streamflow

10.1029/2021WR031575

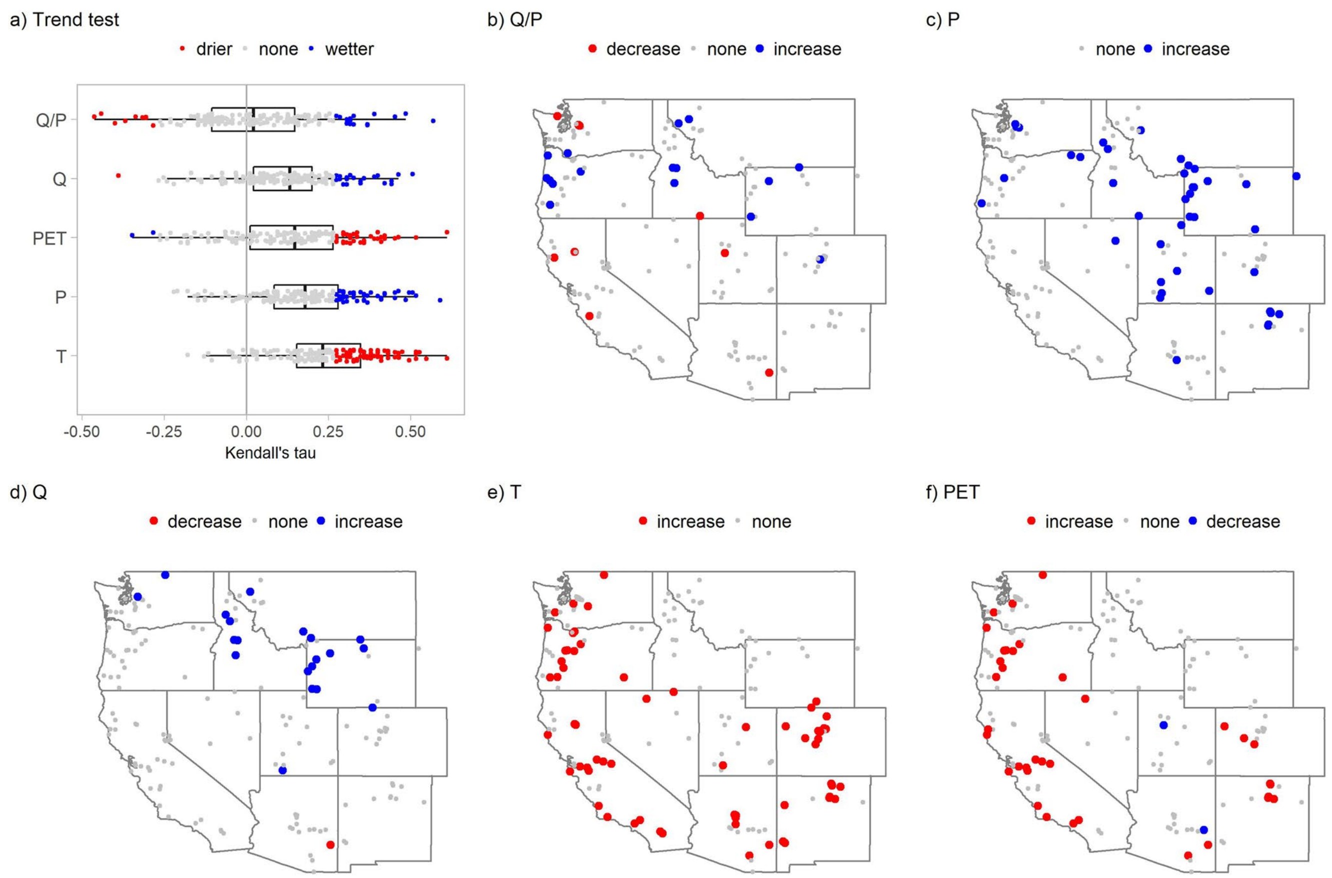

Figure 2. Significant trends in annual water budget components and drivers over the period 2000–2019, based on the Mann-Kendall trend test ($p < 0.1$). Q = streamflow; P = precipitation; T = temperature; and PET, potential evapotranspiration.

can be predicted from a small number of predictor variables for a calibration period, and then applied to a later time period to compare predicted to observed runoff for that time period. Computationally, a linear regression model is calibrated on an initial time period, applied to a second time period, and the residuals (i.e., the difference between the observed and predicted values in the second time period) are assumed to be due to factors not included in the model. Although previous applications of time trend analysis have used a linear regression model, we initially attempted to conduct this analysis using a machine learning model structure, specifically random forests (Breiman, 2001), but found that random forests performed similarly to linear regression but presented the disadvantage of not producing easily interpretable coefficients.

For the purposes of time trend analysis, we split our period of record into two time periods: 2000–2009 and 2010–2019. We calibrated and validated the linear regression model for time trend analysis using data from water years 2000–2009. Odd-numbered years were used for calibration, and even-numbered years for validation. Preliminary analysis indicated that our data set met the assumptions required for linear regression (Helsel et al., 2020). Given that temperature exhibited a significant positive trend at many watersheds (Figure 2) and was a significant predictor, we included it in our model. Thus, the regression model took the form:

$$Q_1 = a_1 * P_1 + b_1 * T_1 + c_1 + e \quad (1)$$

In Equation 1, Q = annual streamflow; P = annual precipitation; T = annual mean temperature; subscripts represent values from the calibration/validation period (time 1, or 2000–2009); a, b, and c are coefficients; and e represents model residuals. We also tested whether the model improved when we included the interaction of T and P as a product term, and seasonal rather than annual T and P; neither of these options improved model fit, so we proceeded with the simpler Equation 1. The regression held a and b the same across all watersheds, for two reasons. First, the processes that relate P and T to streamflow should be consistent across all watersheds, and second, allowing these coefficients to vary would effectively create a separate model for each watershed, which would result in many watersheds being omitted due to years with missing data during the calibration period. The intercept, c, was allowed to vary among watersheds to capture watershed specific differences with respect to factors that were not included in this linear model. The application of this model to the evaluation period (time 2) uses time 1 coefficients and time 2 observations of annual precipitation and temperature to predict annual streamflow over time period 2 (2010–2019):

$$Q'_2 = a_1 * P_2 + b_1 * T_2 + c_1 \tag{2}$$

The difference between observed ($\overline{Q_2}$) and predicted ($\overline{Q'_2}$) mean annual streamflow during the evaluation period is represented as the quantity:

$$\overline{Q_{\text{obs-exp}}} = \overline{Q_2} - \overline{Q'_2} \tag{3}$$

where $\overline{Q_{\text{obs-exp}}}$ represents the magnitude of streamflow change that cannot be attributed to precipitation and temperature and thus is typically interpreted to be due to vegetation change (Zhao et al., 2010).

One objective of time trend analysis was to determine how runoff responds to disturbance. As in our other analyses, we hypothesized that runoff is likely to increase in disturbed watersheds, although a secondary hypothesis was that runoff response depends not only on magnitude of disturbance but also on aridity and/or incoming SRAD. To answer the question of whether streamflow has increased or decreased in disturbed watersheds, we interpreted significant change in streamflow, from our time trend analysis results (i.e., deviation in observed Q from predicted Q) in the context of disturbance. Significant change in annual streamflow was identified using a one-sample t test (Biederman et al., 2015), wherein the null hypothesis was that there has been no change in streamflow due to factors other than precipitation and temperature ($Q_{\text{obs-exp}} = 0$). P-values less than 0.05 were identified as significant deviations in streamflow. Disturbed watersheds were defined as those where tree mortality exceeded 10% of initial live tree volume.

2.2.4. Streamflow Change as a Function of Disturbance Severity and Climate

We used multiple regression to address two objectives: (a) to evaluate the relative importance of several factors for predicting change in streamflow (ΔQ), which allowed isolation of the relative contributions of climate vs. disturbance to ΔQ, and (b) to determine whether the interaction of forest disturbance severity with aridity or SRAD affects runoff response to forest disturbance. A regression model was developed to predict ΔQ across two discrete time periods, 2000–2009 vs. 2010–2019.

To enable disentangling the confounding effects of climate vs. vegetation changes, we initially considered a large set of predictor variables encompassing time varying climatic variables (e.g., change in mean annual precipitation) as well as time-invariant climate descriptors (e.g., long-term mean incoming SRAD) that are specific to each watershed. The initial set of potential predictors included baseline Q and baseline P for 2000–2009 ($\overline{Q_1}$ and $\overline{P_1}$, respectively), mean watershed aridity and SRAD, tree mortality during 2010–2019, and change in temperature, precipitation, and PET between the two time periods. To meet the assumption of noncollinearity among predictors, we then reduced the number of predictors by evaluating pairwise correlations among all predictors and removing predictors with correlation coefficients with absolute values of 0.6 or greater, where the predictor with the lower correlation with ΔQ was removed. In this manner, PET, SRAD, and aridity were removed due to their respective correlations with temperature and $\overline{P_1}$; SRAD and aridity were represented in the model in interaction terms with tree mortality. Due to multicollinearity between the interactions of mortality with SRAD and aridity, we removed the interaction of mortality with SRAD as it was a less useful predictor than the interaction of mortality with aridity. Thus, the final regression model took the form:

$$\Delta Q = b_0 + b_1 \overline{P_1} + b_2\, \Delta P + b_3\, \Delta T + b_4\, \text{mortality} + b_5\, \text{mortality} * \text{aridity} \tag{4}$$

where $\overline{P_1}$ represents mean annual precipitation for 2000–2009; ΔP and ΔT were differences in mean annual precipitation (mm) and mean annual temperature (°C) between 2000–2009 and 2010–2019; and b_x refer to coefficients. As before, we tested whether model fit improved with the inclusion of a product term representing interactions between ΔP and ΔT, and also using differences in seasonal rather than annual P and T to consider the effects of precipitation phase and snowpack, and the model did not improve so we implemented Equation 4 using annual observations of P and T. For this analysis, mortality was standardized by total volume of trees in the watershed, that is, as the volume of trees that died during the study period relative to initial live tree volume, thus having possible values of 0–1 (USDA, 2020). The last term, mortality * aridity, represents the interaction of tree mortality with aridity, which was included to test the hypothesis that streamflow response to forest change is influenced by aridity. We used the p-value associated with the coefficient of each predictor variable in Equation 4 to assess its significance as a predictor of ΔQ. We then compared standardized regression coefficients for each variable to determine the relative importance of climatic factors, forest disturbance, and interaction of forest disturbance with aridity for predicting ΔQ.

Based on the predominant hypothesis that runoff increases following forest disturbance, we expected that tree mortality would have a positive coefficient in the regression model, that is, that larger levels of tree mortality would lead to positive ΔQ. Our alternative hypothesis—that disturbance may decrease runoff at high aridity or SRAD—led to the expectation that the coefficient for the interaction of tree mortality with aridity or SRAD would be negative, even as the coefficient for tree mortality alone was positive. To interpret the ability of each predictor variable to explain additional variability in ΔQ, we examined partial regression plots for each predictor (Moya-Laraño & Corcobado, 2008). Partial regression plots, also known as added variable plots, isolate the explanatory capability of a single variable relative to that of all other variables (Moya-Laraño & Corcobado, 2008). Although pairwise scatterplots between a predictor and ΔQ would be appropriate for simple (single-variable) regression, in the context of multiple regression, such plots ignore the effects of other variables in the model and can thus be misleading representations of the contribution of each variable to explaining variability in the response variable (Moya-Laraño & Corcobado, 2008). Partial regression plots were developed to address this concern using the R package *car* (Fox & Weisberg, 2019). To visualize the interactive effect of disturbance severity and aridity on streamflow change, we also examined marginal effects of the interaction between mortality and aridity using R package *sjPlot* (Lüdecke, 2021).

To interpret our regression model in the context of climatic warming, we used the regression model (Equation 4) to evaluate the sensitivity of streamflow changes to tree mortality and aridity, both with and without 1°C of warming. We compared our results to those of previous studies that projected decreases in streamflow with climate warming across the western U.S. (McCabe et al., 2017; Udall & Overpeck, 2017).

3. Results

3.1. Trends in Water Budget Components and Drivers

Most watersheds (>60%) did not experience significant monotonic trends in any water budget components or drivers during 2000–2019 (Figure 2). P increased significantly between 2000 and 2019 in 26% of watersheds, driving some increasing trends in Q (13%) and Q/P (10%). P and Q decreased in <1% of watersheds, and Q/P decreased significantly 6% of watersheds. T and PET increased significantly in 40% and 23% watersheds, respectively, and both decreased in ≤1% of watersheds (Figure 2), which is consistent with general climate warming. Significant changes in Q/P, P, Q, T, and PET were widespread with no clear geographic patterns (Figures 2a–2f).

When we repeated the Mann-Kendall trend test for the entire period of record (1980–2019), results were very different than for 2000–2019. More watersheds experienced significant decreases in P, Q/P, and Q (7%, 24%, and 17%, respectively), and only 8% of watersheds exhibited significant increases in Q and Q/P. This pattern coincides with significant increases in T (84%) and PET (81%), both of which decreased in <1% of watersheds. Thus, while an appreciable percentage of watersheds show evidence for long-term (1980–2019) increases in T and PET, only a small percentage show evidence for changes in Q and Q/P.

3.2. Runoff Ratio and Forest Change

This analysis sought to test the hypothesis that forest cover is inversely related to runoff, and comparison of trends in runoff ratio (Q/P) to net tree growth demonstrated only moderate support for this hypothesis. Slightly less than half of all watersheds (43%) met the expectation that Q/P is inversely related to change in forest density (Figure 3, upper left and lower right quadrants, with 24 and 44 watersheds, respectively), and the remaining watersheds (57%) did not conform to this expectation (Figure 3, lower left and upper right quadrants). However, a small

10.1029/2021WR031575

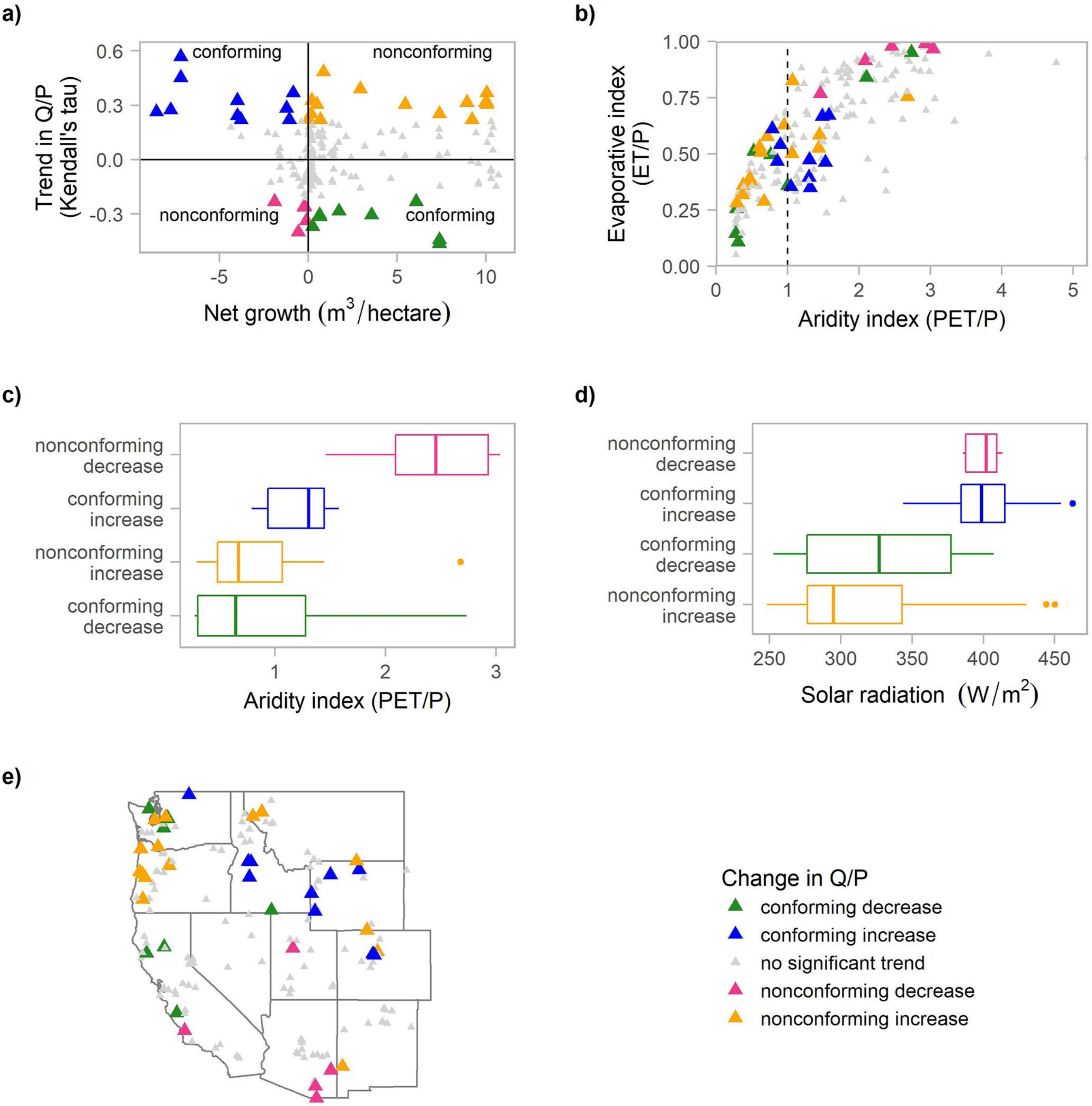

Figure 3. (a) Relationship between trend in Q/P (measured as Kendall's tau) and net growth of trees for 2000–2019. Positive values of Kendall's tau indicate a monotonic increase in Q/P. Colors for watersheds with significant trend over time are assigned based on quadrants, where upper left and lower right quadrants conform to expected Q/P response to forest changes, and lower left and upper right exhibit runoff ratio trends do not conform to expectations. (b) Position of watersheds in the Budyko framework of evaporative index ($1 - Q/P$) vs. aridity index (PET/P). (c and d) Aridity and incoming solar radiation, with watersheds grouped into the quadrants in (a). Boxes represent interquartile ranges; horizontal bars within boxes represent medians. Boxes were not statistically significantly different, based on Kruskal-Wallis test ($\alpha = 0.1$). (d) Geographic distribution of watersheds, with colors as assigned in (a). Q, streamflow; P, precipitation; ET, evapotranspiration; PET, potential evapotranspiration.

proportion of watersheds exhibited statistically significant trends in Q/P, as we found in the previous section. Note that in Figure 3a, watersheds in both left quadrants experienced negative net tree growth, that is, mortality exceed growth by surviving or newly established trees, which indicates disturbance and decrease in volumetric forest density. To quantify the degree to which estimated net growth might reflect random sample variability

10.1029/2021WR031575

or noise, which is higher in smaller watersheds due to smaller sample sizes, we examined the standard errors associated with the estimated net growth in each watershed as produced by the EVALIDator tool. For >75% of watersheds, net growth differed from 0 by more than one standard error. Thus, we inferred that most watersheds have sufficient sample size to reliably indicate positive vs. negative net growth.

Trends in Q/P that contradict the expectation that Q/P is inversely related to change in forest density occurred in two situations. First, Q/P decreased in watersheds with negative net tree growth, that is, greater mortality than live tree growth (Figure 3a, lower left quadrant). This response was observed mainly in water-limited catchments where PET/$P > 1$ and at lower latitudes in the southwestern U.S. (Figures 3b–3e, magenta symbols). Second, Q/P increased while net tree growth was positive (Figure 3a, upper right quadrant). This response was generally observed in energy-limited or moderately water-limited (PET/$P < 2$) watersheds at higher latitudes of the Pacific Northwest and northern Rocky Mountains (Figures 3b–3e).

Given recent research questioning the inverse relationship between forest cover and runoff (Goeking & Tarboton, 2020), an alternative hypothesis is that runoff ratio is more likely to decrease following forest disturbance in watersheds with high aridity and at lower latitude. However, we found that forest disturbance itself was more widespread and severe within water-limited watersheds, as evidenced by the preponderance of magenta and blue symbols where PET/$P > 1$ (Figures 3b and 3c) and where incoming SRAD is relatively high (Figure 3d). Results of the Kruskal-Wallis test showed no significant differences in aridity or SRAD among disturbed watersheds with increased vs. decreased runoff ratio, nor were there significant differences among relatively undisturbed watersheds with increased vs. decreased runoff ratio (Figures 3c and 3d). However, these results do not account for an increasing trend in P over 2000–2019 (see previous section). The following two analyses do account for this effect and thus allow better separation of forest disturbance vs. climate effects on streamflow.

3.3. Streamflow Change as a Function of Precipitation and Temperature Versus Other Drivers

Time trend analysis and subsequent t tests for significant deviations in streamflow indicated that observed streamflow changed significantly in 44 (28% of) watersheds in 2010–2019 relative to 2000–2009 (Figure 4) due to factors other than precipitation and temperature. Of these watersheds, streamflow decreased and increased by statistically significant magnitudes in 30 and 14 watersheds, respectively (Table 3). Validation of the linear model (Equation 1) had adjusted $r^2 = 0.98$. As expected, both precipitation and temperature were significant predictors ($p < 0.01$ for both variables).

Only 26 watersheds experienced both disturbance and significant change in streamflow, as determined by time trend analysis, and streamflow decreased in 20 of these watersheds (Table 3). This finding contradicts the hypothesis that streamflow increases following disturbance. The geographic distribution of significant decreases in streamflow in disturbed watersheds (Figure 4) partially supports our secondary hypothesis that streamflow response to disturbance is influence by factors such as incoming SRAD, aridity, or latitude. Additionally, 18 undisturbed watersheds had significant changes in streamflow (10 decreases and eight increases; Figure 4). These results imply that deviations in observed vs. expected streamflow, as predicted from a linear model based on precipitation and temperature, cannot be attributed to vegetation change alone, which has commonly been an interpretation of time trend analysis (Biederman et al., 2015; Zhao et al., 2010). However, unlike the univariate trends shown in Figures 2 and 3, time trend analysis accounts for changes in P and T over time and evaluates Q relative to those changes.

We considered the possibility that our choice of disturbance threshold could affect our results and therefore evaluated the direction of streamflow response given different disturbance thresholds. Among all watersheds, 67 met our initial disturbance criterion of >10% tree mortality during 2010–2019. Different thresholds (5%, 15%, and 20%) did not lead to different conclusions about the proportion of disturbed watersheds that experience decreased vs. increased streamflow. For all thresholds of disturbance, a slight majority (>54%) of disturbed watersheds exhibited decreased streamflow, based on observed streamflow compared to that predicted by the time trend analysis model.

10.1029/2021WR031575

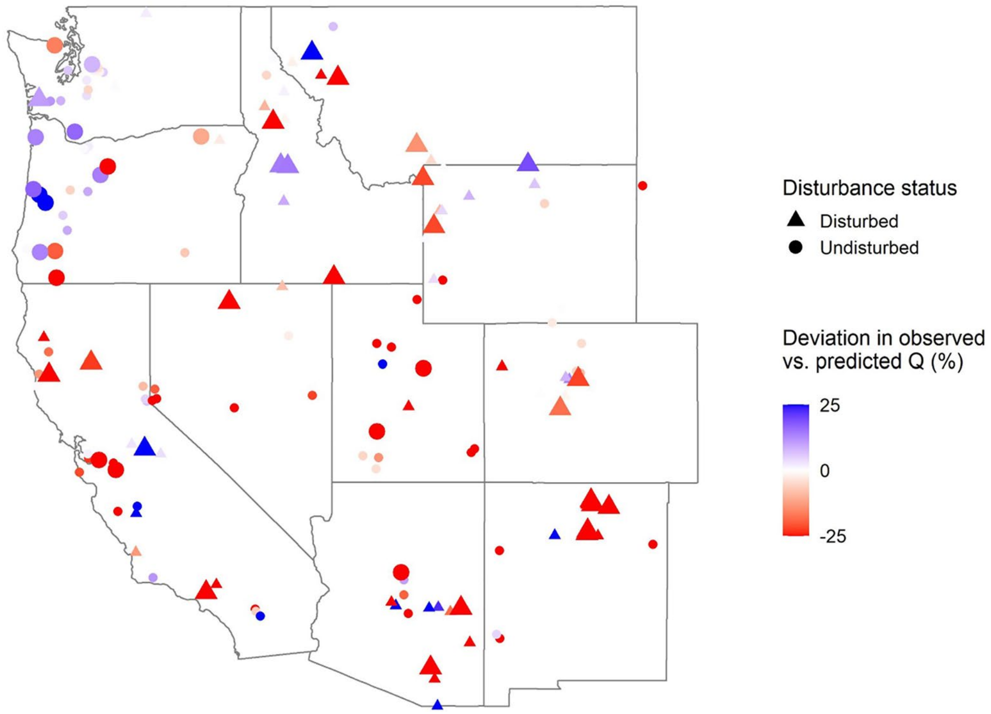

Figure 4. Percent deviation in observed mean annual streamflow (Q) for 2010–2019, relative to Q predicted by time trend analysis (calibrated for 2000–2009). Watersheds with statistically significant deviation in Q (large symbols) were identified using on a one-sample t test ($p < 0.05$); small symbols represent watersheds with no significant deviation in Q ($p \geq 0.05$). Disturbed watersheds (triangles) are those where tree mortality exceeded 10% of initial live tree volume.

3.4. Streamflow Change as a Function of Climate and Disturbance

All coefficients in the multiple regression model for ΔQ (Equation 4) were statistically significant ($p < 0.05$; Table 4) with adjusted model $r^2 = 0.70$ ($p < 0.01$). The average change in runoff (ΔQ) across all 159 watersheds during the time period considered in this analysis was positive (63 mm/yr), consistent with an increase in P (mean ΔP was 91 mm/yr). Standardized regression coefficients indicate the direction and relative impact of each predictor on ΔQ (Figure 5a) and indicate that $\overline{P_1}$ had the largest impact on ΔQ, which may be due to a positive association of $\overline{P_1}$ and ΔP between 2000–2009 and 2010–2019 in watersheds that were already relatively wet. Specifically, the wettest watersheds—those in the Pacific Northwest—had the largest $\overline{P_1}$ and ΔP, and also had low tree mortality. $\overline{P_1}$, ΔP, and mortality all had positive coefficients and thus positive effects on ΔQ, while ΔT and the interaction of mortality with aridity had negative coefficients (Table 4; Figure 5a). Partial regression plots (Figures 5b–5f) illustrate the ability of each predictor variable to explain variability in ΔQ that is not specifically accounted for by other predictors. Note that partial regression plots are not scatterplots of pairwise variables but instead represent the effect on model residuals of adding an additional model term to an existing model. The slopes of the lines in the partial regression plots (Figures 5b–5f) are equal to the regression coefficients and are all significantly different than zero (Table 4), which indicates that each predictor provides useful information in predicting ΔQ. Examination of model diagnostics verified that residuals were normally distributed and independent of predictor values. Figure 5 shows that some observations exert high leverage for some predictors.

Table 3
Results of Time Trend Analysis, Which Predicts Mean Annual Streamflow From Observed Precipitation and Temperature and Then Compares Observed to Predicted Streamflow for a Future Time Period

	Runoff lower than expected (decreased Q)		Runoff higher than expected (increased Q)	
	Any change	Significant change	Any change	Significant change
Disturbed ($n = 67$)	42	20	25	6
Not disturbed ($n = 92$)	56	10	36	8
Total	98	30	61	14

Note. Disturbed watersheds are defined as those where tree mortality exceeded 10% of initial live tree volume. Significant change in annual streamflow was identified as $p < 0.05$ from a one-sample t test.

Table 4
Regression Coefficients, Standard Errors, t-Statistics, and Associated p-Values for Multiple Linear Regression of ΔQ Between 2000–2009 and 2010–2019

Variable	Units	Coefficient	Standard error	*t*-statistic	*P*-value
Intercept	mm/yr	−29.20	10.20	−2.860	0.005
$\overline{P_1}$	mm/yr	0.087	0.008	11.473	<0.001
ΔP	mm/yr	0.107	0.047	2.279	0.024
ΔT	°C	−27.85	6.895	−4.038	<0.001
Mortality	Proportion	250.3	67.91	3.685	<0.001
Mortality * aridity	Proportion	−108.4	43.59	−2.488	0.014

One purpose of this regression analysis was to test the hypothesis that runoff increases following tree mortality, and as an alternative hypothesis, that the sign (positive or negative) of runoff response to disturbance is affected by aridity. Our results provide partial support for both hypotheses. As expected, the coefficient for tree mortality was positive (Table 4; Figure 5a); the statistical significance of this positive coefficient supports the first hypothesis that runoff increases with decreased forest cover. However, the significant and negative coefficient for the interaction of mortality and aridity also supports our alternative hypothesis that mortality does not result in increased runoff in all cases. In particular, runoff response to disturbance may be negative in very arid watersheds. Figure 6a illustrates ΔQ as a function of mortality and aridity based on observations (i.e., not modeled values), demonstrating two important results. First, relatively wet watersheds (aridity < 1.5) generally had positive ΔQ, and ΔQ was larger for watersheds with more tree mortality. Second, very dry watersheds (aridity > 2.5) generally experienced negative ΔQ, and higher mortality was associated with larger decreases in Q. In interpreting these results, it is important to note that overall ΔP was positive, which is expected to contribute to positive ΔQ; thus, the dashed line representing ΔP in Figure 6a provides a more neutral axis of reference than $\Delta Q = 0$.

Figure 6b illustrates predictions and 90% prediction intervals for ΔQ as a function of tree mortality for aridity at its observed 5th percentile, median, and 95th percentile, assuming that all other variables are held constant at their mean observed values. The value of aridity at which tree mortality was predicted to have a negative effect on Q was 2.35. Thus, for watersheds with PET/$P \geq 2.35$, ΔQ decreased with tree mortality. Thus, in these very water-limited watersheds there is an inverse relationship between ΔQ and tree mortality. Note that 95% of watersheds experienced levels of tree mortality less than 33%, so predictions above this level of mortality are beyond the range of most data and therefore uncertain.

As shown in Equation 4, the regression model accounted for changes in precipitation and temperature. The modeled relationship between mortality, aridity, and ΔQ (Figure 6b) demonstrates the same variable response to disturbance as that shown by observations (Figure 6a), illustrating that the response of ΔQ to disturbance and the interaction of disturbance with aridity is not explained by precipitation and temperature changes alone. Thus, decreased streamflow in response to increased temperature or decreased precipitation may be modulated (in wet watersheds) or exacerbated (in dry watersheds) by disturbance.

To assess the overall sensitivity of our modeled ΔQ to potential warming, we summarized ΔQ for several values of mortality and aridity, with and without 1°C of warming (Table 5) and with no change in precipitation. Specifically, Equation 4 was applied with $\Delta P = 0$ and $\Delta T = 0$ or 1. The model predicted a mean decrease in streamflow of 5.6% for 1°C of warming. Regression-based estimates for ΔQ at various levels of tree mortality and aridity generally suggest that streamflow is expected to increase at increasing levels of disturbance for watersheds at low to moderate values of aridity, while the opposite is true in very arid watersheds, specifically with PET/$P > 2.35$, as manifested in the rightmost column of Table 5. Left to right in Table 5, the model indicates greater percentage increases in streamflow following disturbance in more humid watersheds, trending down to a decrease in streamflow for the most arid watersheds. For 1°C of warming, the 5.6% decrease in streamflow is superimposed on these trends.

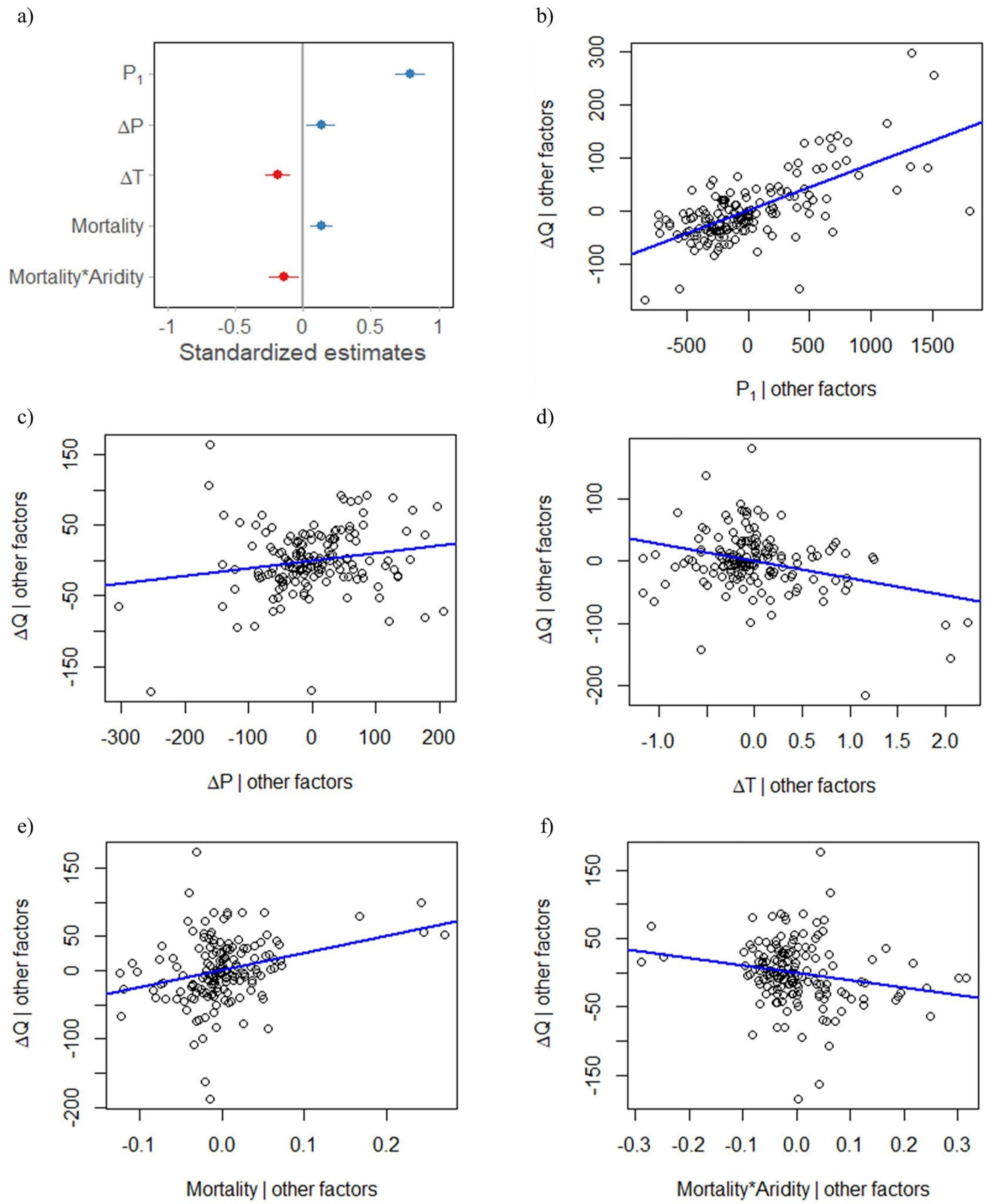

Figure 5. Effect of each variable on change in annual streamflow (ΔQ), in mm/yr, from 2000–2009 to 2010–2019: (a) unitless standardized coefficient estimates, which indicate the magnitude of change in ΔQ, in standard deviations, for a change equal to one standard deviation of each predictor variable. $\overline{P_1}$ = mean annual P for 2000–2009, ΔP, change in precipitation, and ΔT, change in temperature. (b–f) Partial regression plots for each predictor variable. Each plot depicts the relationship between the named predictor and ΔQ while accounting for the explanatory capability of all other predictors. Values along the x axis of each plot represent the residuals of a model omitting the named variable, values along the y axis represent the residuals of a model of the named predictor as a function of all other predictors, and the slope of the line is equal to the multiple regression coefficient for the named variable.

4. Discussion

We found variable runoff response to forest disturbance using multiple analysis methods: Mann-Kendall trend analysis, time trend analysis of predicted vs. observed streamflow based on observed precipitation and temperature, and multiple regression using both climatic and disturbance variables. Collectively, our results confirm, via

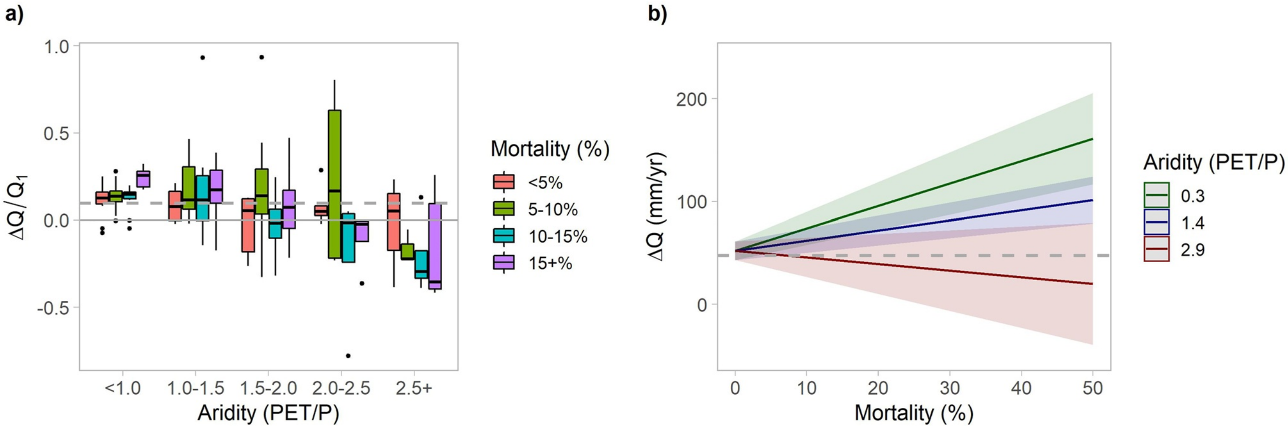

Figure 6. Interacting effect of tree mortality and aridity on ΔQ (2000–2009 vs. 2010–2019). (a) Boxplots of ΔQ (as a proportion of Q_1) based on observed values from 159 watersheds. (b) Marginal effects of mortality and aridity, based on the multiple regression model (i.e., values of ΔQ for different values of mortality and aridity when values of other predictors are held constant); values of aridity represent the 5th percentile (0.3), median (1.4), and 95% percentile (2.9) of watersheds examined in this study. In both plots, horizontal dashed lines represent ΔP times P_1/Q_1, (relative to Q_1 for 6a), which illustrates the expected ΔQ based solely on ΔP.

systematic broad-scale analysis, that the generally held hypothesis that forest cover and streamflow are inversely related is not universal in semi-arid western watersheds. Examination of the relationship between Mann-Kendall trend in *Q*/*P* vs. net tree growth allowed us to identify two scenarios that do not conform to this relationship (Figure 3). First, statistically significant decreases in *Q*/*P* occurred during a period of forest cover loss in a small number of watersheds (four) that occur in areas of high aridity (PET/*P*) and high incoming SRAD. Second, 10 watersheds exhibited statistically significant increases in *Q*/*P* during a period of forest cover growth. Time trend analysis indicated that among watersheds with significant changes in streamflow, 77% (20 of 26) of disturbed watersheds, and only 56% (10 of 18) undisturbed watersheds, experienced decreased streamflow. Thus, significantly decreased streamflow was more prevalent in disturbed than undisturbed watersheds, counter to commonly held expectations. Increased streamflow in 44% (8 of 18) of undisturbed watersheds coincided with higher precipitation overall in 2010–2019 compared to 2000–2009. Multiple regression analysis showed that mortality explains some variability in ΔQ that is not explained by climatic drivers, and that the direction of streamflow response to mortality (i.e., increase vs. decrease) is affected by aridity.

Among our analysis methods, only the multiple regression quantitatively assessed change in streamflow as a function of both climatic and disturbance variables in a way that allowed isolating and quantifying climate and disturbance effects. Therefore, the finding that disturbance severity (i.e., magnitude of tree mortality) is a significant predictor with a positive coefficient supports the overarching hypothesis that streamflow increases as a result of disturbance, and that disturbance effects on streamflow are separable from climate effects. However, the interaction of mortality and aridity had a negative coefficient, which signifies a decrease in streamflow as a result of disturbance in very arid watersheds. Observational data (Figure 6a) as well as our multiple regression

Table 5
Predicted Change in Mean Annual Streamflow (Expressed as a Percentage of Q_1, or Initial Mean Q) for Different Levels of Tree Mortality and Aridity, With and Without a 1°C Temperature Increase and Assuming No Change in Precipitation

		Aridity (PET/*P*)				
	Tree mortality	0.30 (5th percentile)	0.77 (25th percentile)	1.44 (median)	2.08 (75% quantile)	2.93 (95th percentile)
No warming	0%	0.0%	0.0%	0.0%	0.0%	0.0%
	10%	4.4%	3.4%	1.9%	0.5%	−1.3%
	25%	11.0%	8.5%	4.8%	1.3%	−3.4%
1°C warming	0%	−5.6%	−5.6%	−5.6%	−5.6%	−5.6%
	10%	−1.2%	−2.3%	−3.7%	−5.1%	−7.0%
	25%	5.4%	2.8%	−0.9%	−4.4%	−9.1%

10.1029/2021WR031575

results (Figure 6b) provide quantitative evidence that disturbances at high aridity are more likely to result in decreased streamflow than those at lower aridity. These findings are consistent with a recent modeling study (Ren et al., 2021), which concluded that of runoff responds variably to forest disturbance caused by mountain pine beetle, that the response depends on both mortality level and aridity, and that drier years tend toward decreased post-disturbance streamflow. In that study, the inflection from increased to decreased runoff occurred between aridity values of 2.0 and 3.0, or in wetter areas with mortality levels less than 40%, and decreased runoff was explained by either increased canopy evapotranspiration or increased ground transpiration following disturbance (Ren et al., 2021).

Independent of forest cover changes, we observed decreased streamflow associated with increased T and PET. Our multiple regression model predicted a mean decrease in streamflow of 5.6% for 1°C of warming, which is consistent with the 6% reduction per °C that is predicted for the entire Colorado River Basin (Udall & Overpeck, 2017) and 6%–7% reductions per degree that are predicted for the Upper Colorado River Basin (McCabe et al., 2017; Udall & Overpeck, 2017). Our study period, 2000–2019, coincides with the onset of above-average temperatures in the Colorado River Basin that began in 2000 and contributed to below-average streamflow (Udall & Overpeck, 2017). Although this trend has been previously documented in western U.S. watersheds (Brunner et al., 2020; Udall & Overpeck, 2017), the time trend and multiple regression analyses presented here disentangle climate from vegetation effects and offer a refined understanding of the role of forest change effects on streamflow in these trends.

Increasing T and PET are driving not only decreases in streamflow in many western watersheds (Brunner et al., 2020; Udall & Overpeck, 2017) but also increases in tree mortality (Williams et al., 2013). Our analysis of trend in Q/P relative to net tree growth, and our regression model of ΔQ as a function of tree mortality, show relatively high forest disturbance in watersheds with high aridity and SRAD (Figures 3c and 3d). Higher T and PET may affect streamflow both directly, via increased evaporative demand, and indirectly via vegetation-mediated effects such as replacement of trees with vegetation that may actually have higher total evapotranspiration (Bennett et al., 2018; Guardiola-Claramonte et al., 2011; Morillas et al., 2017). Additionally, increases in T and PET that result in increased soil evaporation can increase vegetation moisture stress and susceptibility to disturbance such as wildfire (Groisman et al., 2004).

Possible mechanisms for nonconforming decreases in runoff in watersheds with decreased forest cover (i.e., lower left quadrant in Figure 3a) may be a combination of increased transpiration by surviving or newly established vegetation, as well as increased SRAD reaching snowpack and soil surfaces, either of which may increase total evapotranspiration. The first mechanism, net increase in evapotranspiration due to increased total transpiration, has been observed following insect outbreaks with rapid growth of surviving trees (Biederman et al., 2014), simulated tree die-off that resulted in increased herbaceous transpiration (Guardiola-Claramonte et al., 2011), and replacement of trees with dense shrubs (Bennett et al., 2018); all three of these studies were conducted in semiarid to arid watersheds. Further, short-term streamflow response may contradict longer-term response as young trees grow rapidly during forest recovery (Perry & Jones, 2017) in a phenomenon known as the Kuczera effect (Kuczera, 1987), and the use of net growth as a disturbance metric can quantify the extent to which post-disturbance regrowth may produce this effect. The second mechanism, increased SRAD as a result of canopy loss, could result in earlier snowpack ablation (Lundquist et al., 2013) driven by increased sublimation (Biederman et al., 2014) and increased evapotranspiration from soil and non-canopy vegetation (Morillas et al., 2017; Reed et al., 2018). Changes to post-disturbance energy budgets have been observed following multiple disturbance types and severities (Cooper et al., 2017; Maness et al., 2013). Just as net increases in evapotranspiration can occur following forest disturbance and lead to decreased streamflow, the converse is that net decreases in evapotranspiration can occur during periods of forest cover growth and thus lead to increased streamflow (i.e., upper right quadrant in Figure 3a). Independently of forest disturbance or growth, an additional contributing factor to decreased runoff may be a long-term decline in deep soil moisture due to recent droughts (Iroumé et al., 2021; Peterson et al., 2021; Williams et al., 2020).

Another potential confounding effect is the type of winter precipitation (rain vs. snow). In this study, we accounted for precipitation and temperature at annual and not seasonal time scales; neither the regression model used for time trend analysis nor the multiple regression model for ΔQ improved appreciably when seasonal rather than annual timescales were tested. Previous work has observed both streamflow increases (Hammond & Kampf, 2020) and decreases (Berghuijs et al., 2014) in response to winter precipitation phase (snow to rain) shifts. Warmer temperatures have been observed to result in decreased streamflow in watersheds with high snow fraction, that is, >0.15, although the causal mechanism for this observation is unknown (Berghuijs et al., 2014). In contrast, Hammond and Kampf (2020) observed both increased and decreased streamflow following shifts from snow to mixed rain and snow. Streamflow response to snow-to-rain transitions appear to be more strongly associated with the seasonal timing, particularly relative to the seasonal timing of maximum annual evapotranspiration, than the type of precipitation (de Lavenne & Andréassian, 2018; Knighton et al., 2020; Robles et al., 2021). In our study, increasing trends in Q/P and simultaneous increases in tree growth occurred in a wide variety of environments (Figure 3e), including the temperate Pacific Northwest, where snow fraction may be less than 0.15, as well as high-elevation forested watersheds across the western U.S. where winter precipitation phase change may translate to more rain-on-snow events that produce rapid winter runoff. Because seasonal snowpack represents storage of water that becomes available for transpiration by plants during the growing season, seasonal asynchrony between water availability and the growing season may dampen any relationship between forest cover changes and streamflow response (Knighton et al., 2020).

Results of our time trend analysis demonstrate that streamflow has deviated from predictions based on precipitation and temperature at many watersheds across the western U.S., regardless of forest disturbance (Table 3). An assumption of time trend analysis is that any change not predicted by factors included in the model, typically precipitation and temperature, is due to factors not included in the model, typically vegetation (i.e., land cover) change or land use change (Zhao et al., 2010). However, time trend analysis provides observational but not causal links of change in streamflow to factors such as vegetation change. Incongruities between the subset of watersheds that were disturbed and those with significant streamflow change (Table 3) call into question the underlying premise of time trend analysis that deviations of observed from predicted streamflow are due to vegetation change alone (Zhao et al., 2010). In our exploration of whether changes in streamflow were correlated with changes in T and PET over longer time periods, we found that although T and PET increased in most watersheds, increases in T and PET were not strongly correlated with changes in streamflow or runoff ratio. Given that Mann-Kendall trend tests detected significant increases in T and PET for 1980–2019 that were not detectible during the period covered by our time trend analysis (2000–2019), it is possible that model coefficients for T over multiple decades may not remain constant as temperature increases beyond the range of observed T during 2000–2009. In other words, the assumptions inherent in time trend analysis may not hold in a nonstationary climate as changes may go beyond ranges for which the model was calibrated. Other possible explanations for significant changes in streamflow include shifts in winter precipitation phase (from snow to rain), the timing of seasonal precipitation, longer-term increases in T and PET that are occurring beyond the timeframe considered in this analysis, seasonal T and precipitation extremes that are not reflected in annual mean values, and/or forest disturbance below the threshold considered in our analysis.

A caveat of this study is that we characterized disturbance across entire watersheds, when in reality, disturbance is typically patchy and may include a combination of stand-replacing and nonstand-replacing disturbances. For example, less severe disturbance may be uniformly distributed throughout a watershed whereas more intense disturbances that may affect only small portions of a watershed, where both scenarios would lead to comparable watershed-scale metrics of forest cover loss or tree mortality. Previous studies illustrated that forest structure affects snowpack (Broxton et al., 2016; Moeser et al., 2020), so this distinction may be important for determining disturbance effects on runoff. The ability to project future changes in streamflow due to both changing climate and forest disturbance will likely improve with enhanced spatial representation of forest characteristics.

Several challenges exist in combining observational data sets from different disciplines and using different temporal and spatial sampling frames, and here we describe some of those challenges and potential future solutions. First, the analyses conducted in this study required using forest inventory data collected across multiple years rather than an annual time step. It is not currently possible to produce estimates of the FIA attributes used in this analysis at an annual time step at the scale of individual watersheds, and this constraint undoubtedly dampens observed hydrologic response to acute, episodic disturbances such as severe wildfire. Ongoing work in the area of statistical small area estimation (Coulston et al., 2021; Hou et al., 2021) demonstrates promising capabilities for characterizing forest attributes at finer spatial and temporal scales. Combining FIA-based estimates with other data sets, for example, the Monitoring Trends in Burn Severity data set that delineates large wildfires by severity class (Eidenshink et al., 2007), could illuminate how specific disturbances may have unique or compounding effects on streamflow and snowpack. Application of such techniques to future investigations will require identification of appropriate lag effects and legacy effects (e.g., response to recovery from severe disturbance vs. persistent response to the initial severe disturbance).

Second, most CAMELS watersheds are smaller than the encompassing HUC8 watersheds that we used to summarize forest data, although we found that forest change metrics from the National Land Cover Database (Homer et al., 2020) were statistically similar at the two scales. Compatibility of these data sets could be improved by combining ground observations from forest monitoring plots with remote sensing and other ancillary data, for example, via the small area estimation techniques described above. Ongoing extension of the period of record and improved precision in estimates for individual watersheds will enhance our ability to relate forest characteristics and dynamics to changes in hydrologic processes and flux magnitudes. In particular, improved precision of future monitoring may help quantify important relationships among modulating factors such as aridity and incoming SRAD.

Correlation is not causation, and therefore we cannot be sure that any observed changes in streamflow are due to forest disturbance or the lack thereof. Our results, which are based on observations across many watersheds, underscore the need for process-based modeling to understand where, why, and to what degree unexpected streamflow responses may occur as a result of the combined effects of forest change and climate change. Although there may indeed be forest disturbance effects on streamflow, hydrologic responses may be modulated, offset, or intensified by factors such as aridity and incoming SRAD and by changes in forcing such as increasing temperature.

5. Conclusions

We used a large-sample hydrology approach to combine hydrologic, climatic, and forest data within 159 watersheds in the western U.S. to assess evidence for the hypothesis that forest cover loss leads to increased streamflow. This study expanded on previous studies that have linked streamflow to climatic drivers by also considering quantitative forest disturbance information, which allowed us to disentangle climate effects from forest disturbance effects on streamflow. Multiple analysis methods—including simple trend analysis, time trend analysis accounting for climate variables, and multiple regression—demonstrated that streamflow in some disturbed watersheds was lower than expected based on climatic drivers (i.e., P and T) alone. Results of both observations and multiple regression modeling showed that streamflow response to disturbance was modulated by aridity. Although disturbed watersheds exhibited increased streamflow at low to intermediate aridity, which is consistent with the hypothesis that reduced forest cover produces increased water yield, we found that disturbance in very arid watersheds (aridity > 2.35) was associated with reduced streamflow. Disturbance was also more prevalent in watersheds with high SRAD and high aridity, the very watersheds that are more likely to be vulnerable to decreased streamflow following disturbance. These results suggest that very arid watersheds may be more susceptible to both increased forest disturbance and decreased streamflow in the future.

Acknowledgments
The authors thank Jianning Ren, Ge Sun, and three anonymous reviewers for thoughtful feedback on previous versions of this paper; the National Center for Atmospheric Research (NCAR) for making the CAMELS data set publicly available; the USDA Forest Service for making FIA data publicly available; and to the FIA field crews who collected the forest data used in this study. This work was supported in part by the U.S. Department of Agriculture, Forest Service, and by the Utah Water Research Laboratory at Utah State University. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. This article was prepared in part by employees of the USDA Forest Service as part of official duties and is therefore in the public domain in the U.S.

Data Availability Statement

In an effort to make this study reproducible, the data and computational scripts used to produce the study results have been made publicly available in HydroShare (Goeking & Tarboton, 2022).

References

Adams, H. D., Luce, C. H., Breshears, D. D., Allen, C. D., Weiler, M., Hale, V. C., et al. (2012). Ecohydrological consequences of drought—and infestation—triggered tree die-off: Insights and hypotheses. *Ecohydrology*, *5*, 145–159. https://doi.org/10.1002/eco.233

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., & Mendoza, P. A. (2019). Large-sample hydrology: Recent progress, guidelines for new data sets and grand challenges. *Hydrological Sciences Journal*, *65*, 712–725. https://doi.org/10.1080/02626667.2019.1683182

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*, 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Andréassian, V. (2004). Waters and forests: From historical controversy to scientific debate. *Journal of Hydrology*, *291*, 1–27. https://doi.org/10.1016/j.jhydrol.2003.12.015

Bechtold, W. A., & Patterson, P. L. (2005). *The enhanced Forest Inventory and Analysis Program: National sampling design and estimation procedures*. U.S. Dept. of Agriculture, Forest Service, Southern Research Station GTR-80, 85.

Bennett, K. E., Bohn, T. J., Solander, K., McDowell, N. G., Xu, C., Vivoni, E., & Middleton, R. S. (2018). Climate-driven disturbances in the San Juan River sub-basin of the Colorado River. *Hydrology and Earth System Sciences*, *22*, 709–725. https://doi.org/10.5194/hess-22-709-2018

Berghuijs, W. R., Woods, R. A., & Hrachowitz, M. (2014). A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change*, *4*, 583–586. https://doi.org/10.1038/nclimate2246

Biederman, J. A., Harpold, A. A., Gochis, D. J., Ewers, B. E., Reed, D. E., Papuga, S. A., & Brooks, P. D. (2014). Increased evaporation following widespread tree mortality limits streamflow response. *Water Resources Research*, *50*, 5395–5409. https://doi.org/10.1002/2013WR014994

Biederman, J. A., Somor, A. J., Harpold, A. A., Gutmann, E. D., Breshears, D. D., Troch, P. A., et al. (2015). Recent tree die-off has little effect on streamflow in contrast to expected increases from historical studies. *Water Resources Research*, *51*, 9775–9789. https://doi.org/10.1002/2015WR017401

Boisramé, G., Thompson, S., Collins, B., & Stephens, S. (2017). Managed wildfire effects on forest resilience and water in the Sierra Nevada. *Ecosystems*, *20*, 717–732. https://doi.org/10.1007/s10021-016-0048-1

Bosch, J. M., & Hewlett, J. D. (1982). A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. *Journal of Hydrology*, *55*, 3–23. https://doi.org/10.1016/0022-1694(82)90117-2

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/a:1010933404324

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *Journal of Hydrology*, *310*, 28–61. https://doi.org/10.1016/j.jhydrol.2004.12.010

Broxton, P. D., Dawson, N., & Zeng, X. (2016). Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. *Earth and Space Science*, *3*, 246–256. https://doi.org/10.1002/2016EA000174

Brunner, M. I., Melsen, L. A., Newman, A. J., Wood, A. W., & Clark, M. P. (2020). Future streamflow regime changes in the United States: Assessment using functional classification. *Hydrology and Earth System Sciences*, *24*, 3951–3966. https://doi.org/10.5194/hess-24-3951-2020

Budyko, M. I., & Miller, D. H. (1974). *Climate and life*. New York: Academic Press.

Buma, B., & Livneh, B. (2015). Potential effects of forest disturbances and management on water resources in a warmer climate. *Forest Science*, *61*, 895–903. https://doi.org/10.5849/forsci.14-164

Burrill, E. A., Wilson, A. M., Turner, J. A., Pugh, S. A., Menlove, J., Christiansen, G., et al. (2018). *The Forest Inventory and Analysis Database: Database Description and User Guide for Phase 2 (version 8.0) [WWW Document]*. US Dep. Agric. For. Serv. North. Res. Stn. Retrieved from https://www.fia.fs.fed.us/library/database-documentation/current/ver80/FIADB%20User%20Guide%20P28-0.pdf

Cooper, L. A., Ballantyne, A. P., Holden, Z. A., & Landguth, E. L. (2017). Disturbance impacts on land surface temperature and gross primary productivity in the western United States. *Journal of Geophysical Research: Biogeosciences*, *122*, 930–946. https://doi.org/10.1002/2016JG003622

Coulston, J. W., Green, P. C., Radtke, P. J., Prisley, S. P., Brooks, E. B., Thomas, V. A., et al. (2021). Enhancing the precision of broad-scale forestland removals estimates with small area estimation techniques. *Forestry: An International Journal of Forest Research*, *94*, 427–441. https://doi.org/10.1093/forestry/cpaa045

de Lavenne, A., & Andréassian, V. (2018). Impact of climate seasonality on catchment yield: A parameterization for commonly-used water balance formulas. *Journal of Hydrology*, *558*, 266–274. https://doi.org/10.1016/j.jhydrol.2018.01.009

de Mendiburu, F. (2020). *agricolae: Statistical procedures for agricultural research, R package version 1.5*. https://CRAN.R-project.org/package=agricol

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007). A project for monitoring trends in burn severity. *Fire Ecology*, *3*, 3–21. https://doi.org/10.4996/fireecology.0301003

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Thousand Oaks, CA: Sage.

Goeking, S. A., & Tarboton, D. G. (2020). Forests and water yield: A synthesis of disturbance effects on streamflow and snowpack in western coniferous forests. *Journal of Forestry*, *118*, 172–192. https://doi.org/10.1093/jofore/fvz069

Goeking, S. A., & Tarboton, D. G. (2022). Data for variable streamflow response to forest disturbance in the western U.S.: A large-sample hydrology approach. *HydroShare*. https://doi.org/10.4211/hs.2a674715887a4604ad951d87bdb3c847

Groisman, P. Y., Knight, R. W., Karl, T. R., Easterling, D. R., Sun, B., & Lawrimore, J. H. (2004). Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. *Journal of Hydrometeorology*, *5*, 64–85. https://doi.org/10.1175/1525-7541(2004)005<0064:CCOTHC>2.0.CO;2

Guardiola-Claramonte, M., Troch, P. A., Breshears, D. D., Huxman, T. E., Switanek, M. B., Durcik, M., & Cobb, N. S. (2011). Decreased streamflow in semi-arid basins following drought-induced tree die-off: A counter-intuitive and indirect climate impact on hydrology. *Journal of Hydrology*, *406*, 225–233. https://doi.org/10.1016/j.jhydrol.2011.06.017

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: A need to balance depth with breadth. *Hydrology and Earth System Sciences*, *18*, 463–477. https://doi.org/10.5194/hess-18-463-2014

Hallema, D. W., Sun, G., Bladon, K. D., Norman, S. P., Caldwell, P. V., Liu, Y., & McNulty, S. G. (2017). Regional patterns of postwildfire streamflow response in the western United States: The importance of scale-specific connectivity. *Hydrological Processes*, *31*, 2582–2598. https://doi.org/10.1002/hyp.11208

Hammond, J. C., & Kampf, S. K. (2020). Subannual streamflow responses to rainfall and snowmelt inputs in snow-dominated watersheds of the western United States. *Water Resources Research*, *56*. https://doi.org/10.1029/2019WR026132

Hamon, W. R. (1963). Estimating potential evapotranspiration. *Transactions of the American Society of Civil Engineers*, *128*, 324–338. https://doi.org/10.1061/TACEAT.0008673

Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E. J. (2020). *Statistical methods in water resources (USGS Numbered Series No. 4-A3), Techniques and Methods*. U.S. Geological Survey.

Hibbert, A. R. (1967). Forest Treatment Effects on Water Yield. *International Symposium on Hydrol*, 527–543.

Hirsch, R. M., & De Cicco, L. A. (2015). *User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (USGS Numbered Series No. 4-A10), Techniques and Methods*. U.S. Geological Survey.

Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., et al. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 184–199. https://doi.org/10.1016/j.isprsjprs.2020.02.019

Hou, Z., Domke, G. M., Russell, M. B., Coulston, J. W., Nelson, M. D., Xu, Q., & McRoberts, R. E. (2021). Updating annual state- and county-level forest inventory estimates with data assimilation and FIA data. *Forest Ecology and Management*, *483*, 118777. https://doi.org/10.1016/j.foreco.2020.118777

Hufkens, K., Basler, D., Milliman, T., Melaas, E. K., & Richardson, A. D. (2018). An integrated phenology modeling framework in r. *Methods in Ecology and Evolution*, *9*, 1276–1285. https://doi.org/10.1111/2041-210X.12970

Iroumé, A., Jones, J., & Bathurst, J. C. (2021). Forest operations, tree species composition and decline in rainfall explain runoff changes in the Nacimiento experimental catchments, south central Chile. *Hydrological Processes*, *35*, e14257. https://doi.org/10.1002/hyp.14257

Knighton, J., Vijay, V., & Palmer, M. (2020). Alignment of tree phenology and climate seasonality influences the runoff response to forest cover loss. *Environmental Research Letters*, *15*, 104051. https://doi.org/10.1088/1748-9326/abaad9

Kuczera, G. (1987). Prediction of water yield reductions following a bushfire in ash-mixed species eucalypt forest. *Journal of Hydrology*, *94*, 215–236. https://doi.org/10.1016/0022-1694(87)90054-0

Lu, J., Sun, G., McNulty, S. G., & Amatya, D. M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. *Journal of the American Water Resources Association*, *41*, 621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science, R package version 2.8.9*. Retrieved from https://CRAN.R-project.org/package=sjPlot

Lundquist, J. D., Dickerson-Lange, S. E., Lutz, J. A., & Cristea, N. C. (2013). Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling. *Water Resources Research*, *49*, 6356–6370. https://doi.org/10.1002/wrcr.20504

Maness, H., Kushner, P. J., & Fung, I. (2013). Summertime climate response to mountain pine beetle disturbance in British Columbia. *Nature Geoscience*, *6*, 65–70. https://doi.org/10.1038/ngeo1642

McCabe, G. J., Wolock, D. M., Pederson, G. T., Woodhouse, C. A., & McAfee, S. (2017). Evidence that recent warming is reducing Upper Colorado River flows. *Earth Interactions*, *21*, 1–14. https://doi.org/10.1175/EI-D-17-0007.1

McLeod, A. I. (2011). *Kendall: Kendall rank correlation and Mann-Kendall trend test, R package v2.2*. https://CRAN.R-project.org/package=Kendall

Moeser, C. D., Broxton, P. D., Harpold, A., & Robertson, A. (2020). Estimating the effects of forest structure changes from wildfire on snow water resources under varying meteorological conditions. *Water Resources Research*, *56*, e2020WR027071. https://doi.org/10.1029/2020WR027071

Moore, R. D., Gronsdahl, S., & McCleary, R. (2020). Effects of forest harvesting on warm-season low flows in the Pacific Northwest: A review. *Confluence: Journal of Watershed Science and Management*, *4*, 1–29. https://doi.org/10.22230/jwsm.2020v4n1a35

Morillas, L., Pangle, R. E., Maurer, G. E., Pockman, W. T., McDowell, N., Huang, C.-W., et al. (2017). Tree mortality decreases water availability and ecosystem resilience to drought in Piñon-Juniper woodlands in the southwestern U.S. *Journal of Geophysical Research: Biogeosciences*, *122*, 3343–3361. https://doi.org/10.1002/2017JG004095

Moya-Laraño, J., & Corcobado, G. (2008). Plotting partial correlation and regression in ecological studies. *Web Ecology*, *8*, 35–46. https://doi.org/10.5194/we-8-35-2008

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., et al. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, *19*, 209–223. https://doi.org/10.5194/hess-19-209-2015

Perry, T. D., & Jones, J. A. (2017). Summer streamflow deficits from regenerating Douglas-fir forest in the Pacific Northwest, USA. *Ecohydrology*, e1790. https://doi.org/10.1002/eco.1790

Peterson, T. J., Saft, M., Peel, M. C., & John, A. (2021). Watersheds may not recover from drought. *Science*, *372*, 745–749. https://doi.org/10.1126/science.abd5085

Pugh, E., & Gordon, E. (2012). A conceptual model of water yield effects from beetle-induced tree death in snow-dominated lodgepole pine forests. *Hydrological Processes*, *27*, 2048–2060. https://doi.org/10.1002/hyp

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reed, D. E., Ewers, B. E., Pendall, E., Frank, J., & Kelly, R. (2018). Bark beetle-induced tree mortality alters stand energy budgets due to water budget changes. *Theoretical and Applied Climatology*, *131*, 153–165. https://doi.org/10.1007/s00704-016-1965-9

Reineke, L. H. (1933). Perfection a stand-density index for even-aged forest. *Journal of Agricultural Research*, *46*, 627–638.

Ren, J., Adam, J. C., Hicke, J. A., Hanan, E. J., Tague, C. L., Liu, M., et al. (2021). How does water yield respond to mountain pine beetle infestation in a semiarid forest? *Hydrology and Earth System Sciences*, *25*, 4681–4699. https://doi.org/10.5194/hess-25-4681-2021

Robles, M. D., Hammond, J. C., Kampf, S. K., Biederman, J. A., & Demaria, E. M. C. (2021). Winter inputs buffer streamflow sensitivity to snowpack losses in the salt river watershed in the Lower Colorado River Basin. *Water*, *13*, 3. https://doi.org/10.3390/w13010003

Saksa, P. C., Bales, R. C., Tague, C. L., Battles, J. J., Tobin, B. W., & Conklin, M. H. (2019). Fuels treatment and wildfire effects on runoff from Sierra Nevada mixed-conifer forests. *Ecohydrology*, *13*(3), e2151. https://doi.org/10.1002/eco.2151

Slinski, K. M., Hogue, T. S., Porter, A. T., & McCray, J. E. (2016). Recent bark beetle outbreaks have little impact on streamflow in the western United States. *Environmental Research Letters*, *11*, 074010. https://doi.org/10.1088/1748-9326/11/7/074010

Sun, N., Wigmosta, M., Zhou, T., Lundquist, J., Dickerson-Lange, S., & Cristea, N. (2018). Evaluating the functionality and streamflow impacts of explicitly modeling forest-snow interactions and canopy gaps in a distributed hydrologic model. *Hydrological Processes*, *32*, 2128–2140. https://doi.org/10.1002/hyp.13150

Troendle, C. A. (1983). The potential for water yield augmentation from forest management in the Rocky Mountain region. *Water Resources Bulletin*, *19*, 359–373. https://doi.org/10.1111/j.1752-1688.1983.tb04593.x

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

USDA. (2010). *Forest Inventory & Analysis national core field guide, version 5.0*. U.S. Department of Agriculture, Forest Service, Northern Research Station.

USDA. (2020). *Forest Service, Forest Inventory EVALIDator web-application Version 1.8.0.01*. St. Paul, MN: U.S. Department of Agriculture, Forest Service, Northern Research Station.

Williams, A. P., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., et al. (2013). Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change*, *3*, 292–297. https://doi.org/10.1038/nclimate1693

Williams, A. P., Cook, E. R., Smerdon, J. E., Cook, B. I., Abatzoglou, J. T., Bolles, K., et al. (2020). Large contribution from anthropogenic warming to an emerging North American megadrought. *Science*, *368*, 314–318. https://doi.org/10.1126/science.aaz9600

Wine, M. L., Cadol, D., & Makhnin, O. (2018). In ecoregions across western USA streamflow increases during post-wildfire recovery. *Environmental Research Letters*, *13*, 014010. https://doi.org/10.1088/1748-9326/aa9c5a

Yoda, K., Kira, T., Ogana, H., & Hozumi, K. (1963). Self-thinning in overcrowded pure stands under cultivated and natural conditions. *Journal of Biological Osaka City University*, *14*, 107–129.

Zhao, F., Zhang, L., Xu, Z., & Scott, D. F. (2010). Evaluation of methods for estimating the effects of vegetation change and climate variability on streamflow. *Water Resources Research*, *46*(3). https://doi.org/10.1029/2009WR007702