The Fate of Western Headwaters: Climate Controls on Base-flow Decline

Caelum Mroczek

Abraham E Springer

Benjamin Lucas

2025-12-07

Abstract

Headwater streams comprise nearly 88% of the western U.S. river network and supply most of the region’s surface water, making them especially sensitive to warming, snowpack loss, and drought. As surface-water inputs decline, groundwater increasingly sustains streamflow, elevating the need to understand long-term trends in base flow and their climatic drivers. We analyzed 75 years of streamflow records (1950–2024) from 115 headwater basins to quantify historical changes, climate controls, and future trajectories of base flow. Using statistical analyses and cluster-specific Long Short-Term Memory models trained on downscaled climate data, we assessed both historical behavior and projections under SSP2-4.5 and SSP5-8.5. Historical results show widespread base-flow declines-most pronounced in early summer—driven by warming, reduced snowmelt, and declining antecedent moisture. Antecedent moisture emerged as the dominant positive driver, while snow and temperature exerted regime-dependent effects. Future projections indicate continued declines of 45–65% by late century, with earlier seasonal peaks and reduced summer flows. Snowmelt-dominated and arid basins experience the largest relative reductions, whereas mixed-regime systems contribute the greatest volumetric losses. These changes pose significant risks to municipal water supply, ecosystems, and wildfire resilience across the region. Collectively, our results highlight the vulnerability of groundwater-supported streamflow to climate change and demonstrate the value of integrated statistical and machine-learning approaches for regional hydrologic assessment.

#### Plain Language Summary

Headwater catchments supply much of the surface water in the western United States and are highly sensitive to warming and drying. Using 75 years of streamflow records from 115 headwater basins, we examined how groundwater-fed flow, known as base flow, has changed over time and how it is likely to change in the future. We combined statistical analyses with cluster-specific machine-learning models trained on downscaled climate data to identify the key climate controls and to project future streamflow conditions. Historical records show widespread declines in base flow, especially in early summer, driven by warmer temperatures, reduced snowmelt, and drier antecedent conditions. Antecedent moisture emerged as the strongest positive driver of base flow, while snow and temperature had region-specific effects that shaped the magnitude of change across hydroclimatic regimes. Future projections for SSP2-4.5 and SSP5-8.5 indicate continued declines of 45–65 percent by late century, along with earlier runoff and lower summer flows. Snow-dependent mountain basins and hot, arid basins show the largest relative reductions, while large mixed systems account for the greatest total volume of loss. These findings demonstrate that climate-driven reductions in snowpack and subsurface moisture will significantly diminish groundwater contributions to western headwater streams, with major implications for water supply, ecosystems, and wildfire resilience. Integrating groundwater–surface water dynamics into climate adaptation and water management strategies will be essential for maintaining the reliability and ecological function of these vulnerable systems.

#### Keywords

base flow; groundwater–surface water interactions; headwater streams; climate change

#### Key Points

* Widespread base-flow declines across western U.S. headwaters since 1950, driven by warming, snowpack loss, and reduced antecedent moisture.
* Antecedent moisture is the dominant positive control on base flow, while snow and temperature exert strong, region-specific effects across hydroclimatic regimes.
* Future projections show 45–65% annual declines and earlier seasonal peaks, emphasizing the vulnerability of groundwater-supported streamflow to climate change.

## Introduction

Increasing water demand, land-use change, and climate-driven shifts in precipitation are intensifying water scarcity across the western United States (Taylor et al. 2013; Diffenbaugh, Swain, and Touma 2015). Headwater catchments are the primary sources of streamflow, sustaining downstream ecosystems, agriculture, and communities. Globally, headwater systems make up approximately 80% of river networks (Golden et al. 2025) and contribute disproportionately to total streamflow, particularly in mountainous and semi-arid regions (Viviroli, Weingartner, and Messerli 2003). These catchments are critical zones of groundwater–surface water interaction, where recharge, storage, and discharge processes regulate the timing and persistence of streamflow (Thomas C Winter et al. 1998).

Base flow is the sustained portion of streamflow that is derived from groundwater discharge or other delayed sources (U. S. Geological Survey 2018). It provides a surface-water view of the groundwater system, maintaining perennial reaches, buffering flow during drought, supporting aquatic habitats, and sustaining downstream water supply. Base flow reflects the integrated response of the watershed to precipitation, snow accumulation and melt, evapotranspiration, and subsurface storage (Smakhtin 2001). As such, base-flow magnitude and timing provide key indicators of groundwater availability and watershed resilience under changing climate conditions (Tague and Grant 2009; Diffenbaugh, Swain, and Touma 2015).

Because base flow supplies a major share of annual discharge in many western headwaters, even moderate percentage changes can translate into substantial volumetric losses. These reductions have outsized consequences for downstream water users because late-season discharge is disproportionately groundwater-derived in arid and semi-arid systems (Saedi et al. 2022). As a result, long-term volumetric declines in base flow may directly affect water availability during the periods when ecological and human demands are highest.

Groundwater recharge, the source of base flow, is highly sensitive to shifts in temperature, snowpack, and precipitation. Warming across the western United States has already reduced snow accumulation, shifted precipitation regimes from snow to rain, and increased evaporative demand (Barnett, Adam, and Lettenmaier 2005; Mote et al. 2018). These changes alter groundwater recharge, advance seasonal streamflow peaks, and diminish late-season flows. However, the magnitude and timing of these effects differ across hydroclimatic regimes. Quantifying regime-specific climatic controls is essential for anticipating water availability under continued warming.

Previous studies have advanced understanding of base-flow variability (Mroczek et al. 2025; Beck et al. 2013; Santhi et al. 2008) and its relationship to climate drivers (Ayers et al. 2022; Tan, Liu, and Tan 2020) in different settings. Extending these analyses to incorporate climate projections is critical for anticipating how groundwater-supported flows, especially in headwater systems, will respond to future climate change. As noted by Golden et al. (2025), a key opportunity in headwater research is to move beyond isolated studies toward an integrative understanding of headwater flow regimes across space and time. This integrated approach is limited by the scarcity of adequately gauged headwater sites and by large-scale analyses that generalize hydrologic behavior across diverse climatic and physiographic settings.

Addressing these challenges requires an approach that captures both the historical evolution and the future trajectory of base flow. While headwater systems respond on the basin scale, the overarching effects of climate change on hydrologic systems is best understood on a regional scale (Gorelick and Zheng 2015). To integrate streamflow–climate relationships across diverse settings, we grouped study sites into distinct hydroclimatic clusters that share similar climatic and physiographic controls on base flow. This clustering links local headwater processes to broader regional patterns, allowing comparison of regime-specific responses to climate forcing. Within this framework, we integrate long-term streamflow observations with climate-driven modeling to examine how base flow has changed, what drives those changes, and how it is likely to evolve under future conditions.

Our approach forms a regional modelling framework to quantify base-flow changes and identify the climatic drivers shaping long-term trends. The objectives of this study are to:

1. Quantify seasonal and long-term changes in historic base flow across western U.S. headwater catchments (1950–2024);
2. Identify dominant climate drivers of base-flow variability (1980-2014); and
3. Project future base-flow trajectories using Long Short-term Memory (LSTM) models under both SSP2-4.5 (“Business as usual”) and SSP5-8.5 (“Fossil-fueled Development”) scenarios.

## Data & Methods

Statistical analyses, including Mann–Kendall trend test, Thiel-Sen slope, and mixed-effects modeling, provide the historical understanding of long-term patterns and are used to quantify climate sensitivity. However, these approaches cannot fully represent nonlinear and time-dependent systems, which are inherent in groundwater–surface water systems (Sivakumar 2009). To shore these gaps, we use cluster-specific Long Short-Term Memory (LSTM) neural networks, which excel at capturing time series dependencies and have been shown to capture complex climate–hydrology interactions (Kratzert, Klotz, et al. 2019). Our cluster-based framework groups basins by shared hydroclimatic characteristics, allowing the LSTM models to resolve regime-specific dynamics and improve regional interpretability.

### Data

Daily mean discharge data were obtained for 115 U.S. Geological Survey (USGS) streamgages located across the eleven western states [Figure 1](#fig-cluster-map). Streamflow records were downloaded from the USGS National Water Information System for the period of water years 1950–2024 (Geological Survey (U.S.) 2025). Sites were selected to have greater than 90% daily data completeness, with no gaps exceeding three consecutive years, and to have catchments located entirely within U.S. borders to avoid data restrictions. This period of record provides sufficient length to detect long-term trends and hydrologic regime shifts and encompasses three phases of the Pacific Decadal Oscillation (Newman et al. 2016), reducing the likelihood that results are biased by decadal climate variability. Sites with a dam on the main stem, as reported in the National Inventory of Dams (US Army Corps of Engineers, n.d.), were excluded to reduce the influence of direct flow regulation. Drainage areas ranged from 4 to 21,167 km², with a mean of 1,524 km² and a median of 558 km²

|  |
| --- |
| Figure 1: Study region and analysis timeline. (a) Study basins across the western United States grouped into four clusters representing distinct hydroclimatic regimes as defined in Section. (b) Study timeline showing observed streamflow (1950–2024) and historical (1980–2013) and projected (2014–2099) climate periods under SSP2-4.5 and SSP5-8.5 scenarios. |

Headwaters have been operational defined as streams with Strahler stream order 1 or 2 (Golden et al. 2025; Freeman, Pringle, and Jackson 2007; Imberger et al. 2023). It’s been found that large-scale stream network datasets may underestimate the extent of localized headwater reaches (Brinkerhoff 2024), indicating that these delineated headwater streams likely represent a minimum bound on actual headwater extent. To overcome this deficiency, in this study, streamgages located on streams of order 1 to 3 were included to capture the cumulative processes of headwater systems. By focusing on headwater systems, we overcome issues of scale that arise from the coarse resolution of national and global models and from the limited availability of catchments with adequate monitoring for both streamflow and climate variables.

Basin boundaries were delineated using the Hydro Network-Linked Data Index from the National Hydrography Dataset (Geological Survey (U.S.) 2025). Across the western United States, headwater systems, as defined above, constitute 88.15% of total river length. To assess the representativeness of study sites, their distribution was compared to the full headwater network across the western United States using Köppen climate classifications and the U.S. Forest Service Watershed Condition Framework. All seven major (two-letter) Köppen climate classes present in the region are represented in the study dataset (Peel, Finlayson, and McMahon 2007). Approximately 90 percent of study basins contained U.S. Forest Service lands, consistent with findings that most western water originates in forested areas (Brown, Hobbins, and Ramirez, n.d.). Across the region, 49.1 percent of forested headwater catchments are considered at risk, defined as Functioning at Risk or Impaired Function, the study basins represent this well with 48.5 percent of forested catchment areas falling into these categories.

Climate data were obtained from the Western United States Dynamically Downscaled Dataset (WUS-D3) (Rahimi et al. 2024), a dynamically downscaled product specifically developed for climate applications in the western United States. WUS-D3 contains simulations from multiple GCMs, including the Community Earth System Model v2 (CESM2) global climate model, downscaled with the Weather Research and Forecasting (WRF) model to a spatial resolution of 9 km (Danabasoglu et al. 2020). Climate forcings from CESM2 were selected because it is the only GCM within the WUS-D3 archive that provides simulations for both the SSP2-4.5 and SSP5-8.5 scenarios, allowing direct comparison of both moderate- and high-emissions futures within the modeling framework. The dataset includes a historical period from 1980 to 2013 and extends through 2099 for future projections. Variables used in this study included daily precipitation, daily mean temperature, daily minimum temperature, and daily maximum temperature. Snow precipitation was calculated by summing daily precipitation on days when maximum temperature was below 0 °C. Antecedent moisture was calculated as the cumulative precipitation over the previous three months, following the approach of Ayers et al. (2022) , and serves as a proxy for short-term water storage within the basin. Monthly, area-weighted climate summaries were calculated for each basin to align with the temporal resolution of the base-flow dataset and support climate–base flow relationship analyses.

### Base-flow Separation

Directly estimating base flow from streamflow records presents unique challenges because it cannot be measured directly at the gauge and must be inferred from the total hydrograph (Eckhardt 2008). Numerous approaches have been developed for separating base flow from total streamflow, including tracer studies (Gonzales et al. 2009), graphical interpolation methods (Institute of Hydrology 1980; Sloto and Crouse 1996) , and digital filtering techniques (Arnold et al. 1995; Eckhardt 2005; Nathan and McMahon 1990). The suitability of these approaches depends on factors such as spatial scale, record length, and study objectives. While the choice of method and parameterization can introduce subjectivity, prior work has shown that digital filters provide reliable and repeatable estimates when applied consistently within a study domain (Chapman 1999; Eckhardt 2005; Institute of Hydrology 1980; Ayers et al. 2022).

In this study we used the Eckhardt (2005) digital filter to estimate base flow from daily streamflow records using this equation:

where is the filtered base-flow response at time step , is the observed streamflow at time step , is the base-flow response at the previous time step, is the recession constant, and is the maximum possible base-flow index for the catchment. The recession constant was estimated for each site through hydrograph recession analysis. was determined for each site using the backwards filter method proposed by Collischonn and Fan (2013), which allows to be estimated from without requiring site-specific hydrogeologic field data. The Eckhardt filter was selected because it has shown strong performance in diverse hydrologic settings across the contiguous United States and has been recommended as a preferred base-flow separation method in large-sample studies (Xie et al. 2020).

### Hydroclimate Clustering

To classify study basins into groups with similar hydroclimatic regimes, we applied k-means clustering to basin-averaged climate and hydrologic variables. Clustering was conducted on long-term records of base flow magnitude and variability, precipitation, snow-derived precipitation, mean temperature, antecedent moisture, and static physiographic attributes including drainage area, relief, and elevation. To compare across watersheds with different areas, precipitation, and discharge, we normalized variables (z-scores) prior to clustering to ensure equal weighting. The optimal number of clusters was selected based on a combination of the elbow method and interpretability of known hydrologic regimes in the western United States. K-means was chosen because it is an efficient method for partitioning basins into internally cohesive groups which minimize within-group variance and enhances the detection of coherent regional trends in streamflow across diverse hydroclimatic settings (Dethier et al. 2020). This approach is well suited for large-sample hydrology applications where hydroclimatic gradients are continuous rather than categorical (Ikotun et al. 2023). The resulting clusters were used to stratify subsequent statistical and machine learning analyses, allowing for the identification of climate–base flow relationships and projected changes within distinct hydrologic response regimes.

### Statistical Models

We used the Mann-Kendall (MK) trend test to determine the presence of trends in the base-flow data at a monthly time-step. The MK test is a nonparametric test that detects monotonic trends in non-normally distributed data. The MK test is widely used in hydrologic studies (Murray, Ayers, and Brookfield 2023; Ayers et al. 2022; Woodhouse and Udall 2022; Chen and Teegavarapu 2021), and was used here to establish statistically significant trends ( < 0.05). Autocorrelation was present in many of the streamflow records used here, as such we employed the modified MK test proposed by Hamed and Ramachandra Rao (1998) . The trend magnitude was estimated using the Thiel-Sen slope, a non-parametric technique widely used in hydrologic studies (Rice et al. 2015; Tillman et al. 2022; Murray, Ayers, and Brookfield 2023).

We used a linear mixed-effects modeling framework to quantify relationships between monthly base flow and climate variables. Fixed effects included precipitation, snow, temperature, and antecedent moisture; which capture broad regional-scale climate drivers. A random intercept for each catchment accounts for site-level variability in base flow that is not explained by the fixed predictors. All models were fit using log-transformed monthly base flow as the response variable to reduce skewness, stabilize variance, and improve model performance. To address strong multicollinearity among the temperature variables (mean, minimum, and maximum daily temperature), we retained only mean temperature in the mixed-effects model to ensure interpretability of parameter estimates. This constraint was applied only in the mixed-effects framework; the LSTM modeling framework retained all three temperature variables, as deep learning approaches do not require independence among predictors and are capable of capturing nonlinear relationships and interactions (Razavi 2021).

### LSTM Neural Network Modeling

#### Model Framework

Long Short-Term Memory (LSTM) networks are a form of recurrent neural network designed to learn long-range temporal dependencies in sequential data [@hochreiter\_long\_1997], making them well suited for modeling hydrologic systems with lagged and nonlinear responses [@kratzert\_neuralhydrology\_2019]. This model is designed to predict monthly stream base flow using both climate history and watershed characteristics. It combines two types of input data: (1) a sequence of monthly climate variables over the past 24-months, and (2) static basin attributes like latitude, elevation, and area.

To model monthly base-flow dynamics, we trained cluster-specific LSTM models using overlapping 24-month climate sequences (e.g., months 1–24, 2–25, etc.), each used to predict base flow in the subsequent month. This sliding-window approach increases the number of training samples and enables the network to learn multi-seasonal temporal dependencies. Static basin attributes were appended to the data after the LSTM output so that the model could incorporate physiographic controls alongside climate inputs.

Each model consisted of a single LSTM layer with 256 units, followed by two dense layers with 128 and 64 units, respectively, with 40% dropout applied to the dense layers to reduce overfitting [@JMLR:v15:srivastava14a]. The final output is the predicted log-transformed base flow for the next month. Models were trained using the Adam optimizer (learning rate 0.001) with mean absolute error (MAE) as the loss function. Training continued for up to 50 epochs, with early stopping based on validation loss.

To reduce sensitivity to initialization, five models per cluster were trained using different random seeds. Seed-specific models were evaluated individually and used to generate both best-seed predictions (lowest validation MAE) and seed-ensemble predictions, with the ensemble computed as the median across all seed models. Ensemble predictions were used for all scenario analyses, as they were more stable across clusters.

|  |
| --- |
| Figure 2: Structure of the Long Short-term Memory (LSTM) model used to predict monthly base flow. (a) Conceptual layout of the LSTM-based recurrent neural network. (b) Model configuration showing 24-month climate inputs and static basin attributes processed through LSTM, dense, and output layers to predict next-month base flow. |

#### Model Training and Testing

Model training and hyperparameter selection were conducted separately for each hydroclimatic cluster. We tested a range of LSTM configurations varying the number of hidden units, dropout rates, learning rates, and batch sizes. The final hyperparameters were chosen based on minimizing cross-validated mean absolute error (MAE) during the 2006–2013 period (Table S1). Each model used a 24-month input sequence of climate predictors (precipitation, temperature, antecedent moisture indices, snow fraction, and seasonal harmonics) along with static basin characteristics (elevation, relief, drainage area, latitude, longitude). Inputs were standardized using the mean and standard deviation from the training period (1980–2005).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Predictors used in the LSTM models, grouped by general data type. Variables include raw climate inputs, lagged terms, derived precipitation metrics, static basin attributes, and seasonal harmonics.   | **Data Type** | **Predictor** | **Description** | | --- | --- | --- | | **Temperature** | t2 | Mean monthly near-surface (2 m) air temperature (°C) | |  | t2max | Monthly maximum temperature (°C) | |  | t2min | Monthly minimum temperature (°C) | |  | t2\_lag1 | Mean temperature, 1-month lag | |  | t2\_lag3 | Mean temperature, 3-month lag | | **Precipitation** | prec | Monthly precipitation (mm) | |  | prec\_lag1 | Precipitation, 1-month lag | |  | prec\_lag2 | Precipitation, 2-month lag | |  | prec\_lag3 | Precipitation, 3-month lag | |  | prec\_lag6 | Precipitation, 6-month lag | |  | rolling\_prec3 | 3-month rolling precipitation mean (mm) | |  | ppt\_change | Monthly precipitation change (current minus prior month) | | **Snow / Moisture** | snow | Precipitation falling when t2max < 0 °C (mm) | |  | prec\_mois | Antecedent moisture (3-month precipitation sum) | |  | prec\_mois\_lag1 | Antecedent moisture, 1-month lag | | **Seasonality** | month\_sin | Cyclical encoding of month (sine transform) | |  | month\_cos | Cyclical encoding of month (cosine transform) | | **Basin Attributes** | Area\_km | Watershed area (km²) | |  | Elev\_mean\_m | Mean basin elevation (m) | |  | Elev\_min\_m | Minimum basin elevation (m) | |  | Elev\_max\_m | Maximum basin elevation (m) | |  | Relief\_m | Basin relief (elev. max – elev. min, m) | |

For model evaluation, we validated each cluster-specific LSTM model against climate forcing from the WUS-D3 SSP2-4.5 and SSP5-8.5 datasets and observed base flow (2014–2024). Models were trained on data from 1980–2005 and cross-validated on 2006–2013, then applied to the scenario data to assess predictive skill outside the training period. To reduce sensitivity to initialization, we trained five models per cluster with different random seeds and combined them into an ensemble using the median prediction across seeds. Model skill was quantified using MAE, root mean square error (RMSE), and the Nash–Sutcliffe efficiency (NSE). The ensembled predictions showed improved performance over single-model predictions across clusters and scenarios.

## Results

### Hydroclimate Clustering

The k-means clustering identified four distinct hydroclimatic groups among the 115 study basins, reflecting differences in elevation, relief, basin size, climate regime, and base-flow response characteristics. These clusters capture gradients in snow influence, seasonal precipitation patterns, and base-flow variability across the western United States. Two clusters (1 and 3) are primarily high-elevation, snow-influenced headwaters but differ in their sensitivity to temperature and snowpack persistence. Cluster 2 contains the largest, mixed-regime basins with relatively stable base flow and Cluster 4 represents low-elevation, arid to monsoon-influenced basins with the lowest base flow and highest interannual variability. Descriptive statistics for each cluster, including key climate, physiographic, and base-flow metrics, are provided in [Table 2](#tbl-cluster-summary), and the spatial distribution of clusters across the study area is shown in [Figure 1](#fig-cluster-map).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Summary of physical, climatic, and base-flow characteristics for each hydroclimatic cluster. Values represent mean conditions for basins within each cluster; base-flow statistics are derived from the full 1950–2024 record. Base flow CV indicates the coefficient of variation ()).   | Cluster | n | Area (km²) | Elev. Mean (m) | Relief (m) | Prec. Mean (mm) | Temp. Mean (°C) | Monthly Base Flow Mean (cfs) | Base Flow CV | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 1 | 10 | 1193.3 | 2106 | 1930 | 107.48 | 5.51 | 4765 | 1.37 | | 2 | 17 | 4452.2 | 1266 | 2825 | 200.78 | 5.74 | 51840 | 0.74 | | 3 | 50 | 1124.8 | 2330 | 1772 | 85.69 | 3.97 | 4357 | 1.30 | | 4 | 38 | 865.1 | 735 | 1263 | 141.28 | 11.06 | 8071 | 1.77 | |

Cluster 1 (Snowmelt-Dominated Mountain Catchments) occupies steep, high-relief mountain terrain where winter precipitation is primarily stored as snowpack and released during a concentrated late-spring melt period. These basins are located across the study area in high-elevation mountains. Base flow peaks in June, approximately six months after the precipitation maximum, and inter-annual variability is high due to differences in snow accumulation and melt timing.

Cluster 2 (Mixed-Regime Large Catchments) represents moderate-elevation, high-relief basins with the largest drainage areas in the dataset. Most of these basins are found along the Cascade Range in the north west of the study area along CA, OR, and WA, including basins with perennial snow and glaciers (Pelto 2008). Base flow peaks earlier, in May, about five months after the precipitation maximum. Snowmelt remains important, but basin size and integrated flow paths buffer short-term variability, resulting in lower inter-annual variability during peak months.

Cluster 3 (Snowy, Responsive Headwaters) also peaks in June but differs from Cluster 1 in showing a stronger positive association between temperature and base flow, consistent with more immediate snowmelt responses during warm periods. These basins are found throughout the Intermountain West where they may experience earlier onset of melt or mid-winter melt events, leading to sustained base flow through early summer and moderate-to-high interannual variability in late spring.

Cluster 4 (Arid/Monsoon or Ephemeral Basins) occurs at the lowest elevations with minimal snow influence. Base flow peaks in March, only about three months after the precipitation maximum, reflecting rapid winter-to-spring runoff and limited storage. These basins are mostly located along the Pacific coast. Flows recede quickly, and summer base flow is minimal due to high evapotranspiration and low infiltration from monsoon rains.

Across all clusters, precipitation peaks in December, contributing roughly 15–18% of the annual total, and remains elevated through February ([Figure 3](#fig-bf-precip-contrib)). Snowmelt-dominated clusters (1 and 3) exhibit the largest variability in late-spring base flow, whereas Cluster 4 shows the strongest temperature-driven suppression of base flow in the dry season. These distinct seasonal signatures and sensitivities highlight the potential for divergent base-flow responses to future climate change across hydroclimatic regimes.

|  |
| --- |
| Figure 3: Monthly distribution of base flow and precipitation as a percentage of the annual total, stratified by hydroclimatic cluster. Base-flow boxes are colored by cluster (Cluster 1: green, Cluster 2: red, Cluster 3: yellow, Cluster 4: blue), precipitation is shown in gray. |

### Historical Base-flow Analysis

#### Climate-Base Flow Relationship

When the mixed-effects model was analyzed across all sites [Table 3](#tbl-mixed-effect_all-site), it revealed that antecedent moisture was the most consistently positive and influential driver of base flow across the study region. This supports the role of cumulative recharge and basin memory in sustaining flows, particularly during dry-season months. In contrast, snow precipitation had a strong negative contemporaneous effect in all clusters, consistent with winter accumulation storing water in the snowpack and delaying its release to streams. Mean temperature showed an overall negative association with base flow, suggesting that warmer conditions correspond to seasonal drying, likely through enhanced evapotranspiration or reduced soil moisture. Monthly precipitation had a generally weak to negative effect, with the strongest suppression in arid or monsoon-influenced basins.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: Mixed-effects model results for climate and static predictors of monthly base flow across all study basins. Estimates represent fixed effects on log-transformed base flow, with positive coefficients indicating a positive association and negative coefficients indicating a negative association. Significance codes: \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05; ns = not significant.   | Predictor | Estimate | Std. Error | t-value | Significance | | --- | --- | --- | --- | --- | | Precipitation | -0.085 | 0.009 | -8.99 | \*\*\* | | Snow Precipitation | -0.316 | 0.007 | -42.51 | \*\*\* | | Mean Temperature | -0.079 | 0.009 | -9.03 | \*\*\* | | Antecedent Moisture | 0.477 | 0.009 | 52.50 | \*\*\* | | Mean Elevation | -0.420 | 0.206 | -2.03 | \* | | Relief | 1.059 | 0.225 | 4.70 | \*\*\* | | Area | -0.012 | 0.223 | -0.06 | ns | |

When the model was evaluated by hydroclimatic cluster ([Table 4](#tbl-mixed-effect_cluster)), key differences in climate–base flow relationships emerged. Antecedent moisture remained strongly positive in all groups, with the largest effect in Cluster 4, indicating its particular importance in sustaining flows in water-limited environments. Snow suppression of base flow was strongest in Cluster 1, reflecting the storage-dominated snowmelt regime of steep mountain catchments. Temperature effects varied markedly: Cluster 3, comprising snow-responsive headwaters, was the only group with a positive temperature effect, consistent with warming-induced snowmelt boosting base flow during the transitional period between accumulation and melt. In contrast, Cluster 4 exhibited a strong negative temperature effect, suggesting enhanced evaporative losses or soil moisture depletion during warm periods. Terrain metrics were also important in certain regimes; elevation and relief had the strongest positive effects in Cluster 1, indicating that steep, high-relief basins may enhance routing of melt-water into groundwater and streams. Basin area had limited influence in most clusters but was strongly negative in Cluster 4, potentially reflecting inefficient runoff generation or storage losses in large, low-relief arid basins.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4: Mixed-effects model estimates for climate and static predictors of monthly base flow, stratified by hydroclimatic cluster. Estimates represent fixed-effects on log-transformed base flow; positive values indicate a positive association and negative values indicate a negative association.   | Variable | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | | --- | --- | --- | --- | --- | | Precipitation | -0.075 | -0.139 | -0.015 | -0.227 | | Snow Precipitation | -0.369 | -0.124 | -0.286 | -0.125 | | Mean Temperature | -0.064 | -0.104 | 0.325 | -0.568 | | Antecedent Moisture | 0.417 | 0.354 | 0.412 | 0.548 | | Mean Elevation | 0.474 | -0.402 | 0.031 | -0.268 | | Relief | 1.153 | 0.059 | 0.657 | -0.168 | | Area | 0.219 | -0.172 | 0.325 | -0.480 | |

These results highlight that while antecedent moisture is a dominant control on base flow in all regimes, the influence of snow, temperature, and terrain varies systematically across hydroclimatic clusters. These distinctions underscore the value of classification-informed modeling for understanding and projecting base-flow responses to climate variability and change.

#### Base-Flow Trends and Seasonality

Across the four hydroclimatic clusters, long-term, historical base-flow trends (1950-2024) varied in both magnitude and seasonal timing ([Table 5](#tbl-bf-trend_cluster); [Figure 4](#fig-hist_trends_monthly)). The most widespread declines occurred in early summer, especially in June and July, consistent with reductions in snowmelt-driven base flow. Historical rates in [Table 5](#tbl-bf-trend_cluster) are continuous Theil-Sen slopes (trend per decade) rather than baseline-referenced anomalies, and are thus not expressed in the same units as projected annual change (Section 3.3). Modest increases were observed in winter and early spring months, particularly January and March in Cluster 1 and March–April in Cluster 3. Cluster 1 showed a mix of increases in colder months and decreases in summer, resulting in a small overall positive annual trend. Cluster 2 exhibited consistent and large negative trends across nearly all months, with the strongest declines in February, June, and December. Clusters 3 and 4 showed modest negative trends overall, with variability in both direction and timing depending on the month. These patterns highlight seasonal shifts in base flow that differ across hydroclimatic regimes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 5: Median monthly base-flow trends by hydroclimatic cluster for 1950–2024, shown as percent change **per decade** from Theil–Sen slope estimates. Positive values indicate increasing base flow; negative values indicate decreasing base flow. “Pct. Sig. Up” and “Pct. Sig. Down” show the percentage of sites with statistically significant (p < 0.05) positive or negative trends.   | Cluster | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | Median Trend | Pct. Sig. Up | Pct. Sig. Down | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 1 | 1.22 | 0.38 | 5.60 | 3.92 | 0.50 | -1.31 | -1.71 | 0.08 | -3.22 | -0.64 | 3.14 | 1.10 | 0.44 | 20.69 | 16.38 | | 2 | -0.39 | -3.75 | -0.24 | -1.28 | -1.63 | -3.54 | -3.71 | -2.79 | -3.04 | -2.59 | -1.10 | -2.31 | -2.45 | 1.47 | 30.39 | | 3 | -0.61 | -0.95 | 1.12 | 1.12 | 0.04 | -2.33 | -3.10 | -3.26 | -2.25 | -0.92 | -0.77 | -1.81 | -0.94 | 8.60 | 17.37 | | 4 | -0.13 | -4.27 | 0.32 | 0.00 | -0.11 | -0.37 | -1.92 | -2.75 | -3.54 | -3.28 | -0.88 | -2.35 | -1.40 | 12.79 | 18.84 | |

|  |
| --- |
| Figure 4: Historic monthly base-flow trends across western U.S. headwater basins (1950-2024), colored by Theil–Sen slope and shaped by statistical significance. Upward- and downward-pointing triangles indicate significant increasing or decreasing trends (p < 0.05), respectively, while circles denote non-significant trends. Colors correspond to slope magnitude and direction, with warmer tones indicating negative trends and cooler tones indicating positive trends. |

### Projected Base-flow Analysis

#### Model Validation

|  |
| --- |
| Figure 5: Observed vs. predicted monthly log-transformed base flow for the ensemble LSTM models across four hydroclimatic clusters under two climate scenarios (SSP2-4.5 and SSP5-8.5). Each panel shows one cluster and scenario combination. The 1:1 dashed line indicates perfect prediction. Performance metrics (MAE, RMSE, NSE) are shown within each facet, demonstrating consistent model skill across clusters and scenarios, with strongest performance in Clusters 2 and 3 and higher variance in Cluster 4. |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 6: Ensemble LSTM evaluation metrics (MAE, RMSE, NSE) for each hydroclimatic cluster under SSP2-4.5 and SSP5-8.5 (2014–2024).   | Cluster | Scenario | MAE | RMSE | NSE | | --- | --- | --- | --- | --- | | 1 | SSP2-4.5 | 0.72 | 1.16 | 0.79 | |  | SSP5-8.5 | 0.69 | 1.10 | 0.81 | | 2 | SSP2-4.5 | 0.42 | 0.54 | 0.77 | |  | SSP5-8.5 | 0.42 | 0.54 | 0.77 | | 3 | SSP2-4.5 | 0.68 | 1.06 | 0.80 | |  | SSP5-8.5 | 0.65 | 1.00 | 0.82 | | 4 | SSP2-4.5 | 1.14 | 1.76 | 0.79 | |  | SSP5-8.5 | 1.13 | 1.73 | 0.80 | |

Model performance was evaluated using observed versus predicted monthly log-transformed base flow for each hydroclimatic cluster and future climate scenario ([Figure 5](#fig-obs_pred); [Table 6](#tbl-model_eval)). The ensemble LSTM models demonstrated strong predictive skill across all clusters and both climate scenarios during the validation period (2014–2024). Performance was highest for Clusters 2 and 3, with low MAE (0.42–0.68), low RMSE (<1.1), and high NSE (0.77–0.82), indicating accurate prediction of both magnitude and variability in monthly base flow. Cluster 1 showed moderate skill (MAE ≈ 0.7, NSE ≈ 0.8), while Cluster 4 had the largest errors (MAE ≈ 1.1–1.2, RMSE ≈ 1.7, NSE ≈ 0.8), reflecting the difficulty of modeling base flow in arid and monsoon-influenced basins. While Cluster 4 had higher errors, it retained good NSE due to stronger performance at higher flows, while low-flow values were more weakly predicted. Model skill was consistent between SSP2-4.5 and SSP5-8.5, indicating robustness across climate forcings.

In [Figure 5](#fig-obs_pred), the vertical alignment of points at very low observed values arises from two factors. First, a minimum threshold applied before log-transformation set all flows less than 0.01 cfs to the same log value, creating an artificial “fence” effect. Secondly, because the models were trained using MAE on log(base flow), errors at very low flows were weighted less by the loss function, reducing accuracy at low values, where observations are inherently noisy. Residual distributions are shown in Figures S1 and S2.

#### Magnitude and Scenario Dependence

Projections indicate persistent declines in base flow across western U.S. headwater systems throughout the 21st century, with the magnitude of reductions varying between emission scenarios ([Figure 6](#fig-annualized_anomaly) a). Relative to the 1980–2013 baseline, projected base flow under both SSP2-4.5 and SSP5-8.5 shows similar early-century (2025–2049) anomalies, followed by continued declines through the mid- (2050–2074) and late-century (2075–2099) periods. Trajectories remain comparable through mid-century in both scenarios, but, by late-century, SSP5-8.5 produces substantially larger reductions and greater inter-site variability ([Figure 6](#fig-annualized_anomaly) a).

Early-century base-flow anomalies are projected to range from –30 to –40 %, corresponding to basin-averaged losses of 3.0–3.4 million AF yr⁻¹. Mid-century reductions deepen to –40 to –50 % (3.6–4.1 million AF yr⁻¹), and by late-century, annual anomalies reach –45 to –65 %, or roughly 3.7–5.8 million AF yr⁻¹, relative to the 1980–2013 baseline ([Figure 6](#fig-annualized_anomaly) b). These patterns indicate that while early-century changes are similar across scenarios, the high-emission SSP5-8.5 pathway drives markedly greater declines in the late-century.

|  |
| --- |
| Figure 6: Projected magnitude of base-flow change across western U.S. headwaters under SSP2-4.5 and SSP5-8.5 scenarios. (a) Regional annual base-flow anomalies (percent difference from the 1980–2013 mean) with shaded 25–75 % ranges across basins. Both scenarios project sustained declines through the 21st century, with greater divergence after ~2050. Median losses are ~30–40 % by mid-century and ~45–65 % by late century. (b) Cluster-balanced mean annual volumetric change (million acre-feet yr⁻¹) summarized by period and scenario. Region-wide average losses equal approximately –3.7 million AF yr⁻¹ for SSP2-4.5 and –5.7 million AF yr⁻¹ for SSP5-8.5 by end-of-century, emphasizing the greater depletion expected under higher-emission conditions. |

#### Relative and Absolute Cluster Change

|  |
| --- |
| Figure 7: Monthly base-flow anomalies by hydroclimatic cluster, period, and emissions scenario. (a) Percent anomalies show the relative change in base-flow volume compared to the 1980–2013 baseline, summarized by cluster medians. Snow- and monsoon-influenced clusters (1 and 4) exhibit the largest relative declines, while Cluster 2 experiences smaller proportional reductions. (b) Corresponding volumetric anomalies (million acre-feet per month) highlight the absolute magnitude of change. Despite modest percent reductions, Cluster 2 contributes the greatest total volume loss across all future periods. |

Cluster-level projections reveal clear heterogeneity in base-flow responses between hydroclimatic clusters ([Figure 7](#fig-monthly-anomaly_period)). Snowy, responsive Cluster 3 and arid/monsoon-influenced Cluster 4 exhibit the largest relative declines, with median reductions of ~80–90 % by late-century under SSP5-8.5 ([Figure 7](#fig-monthly-anomaly_period) a). Cluster 1, snow-dominated mountainous catchments, showed less declines in the summer and fall. In contrast, mixed-regime Cluster 2 shows more moderate, but variable, relative declines of ~20–70% throughout the year, likely due to its larger size and integrated flow paths buffering against extreme variability. When considering absolute volumetric changes ([Figure 7](#fig-monthly-anomaly_period) b), Cluster 2 contributes the greatest monthly losses (≈0.5-3 million AF mo⁻¹) across all future periods, despite its smaller relative declines. This reflects its’ larger base-flow volumes and highlights the importance of considering both relative sensitivity (percent) and hydrologic significance (volume) when assessing climate impacts on streamflow.

#### Changes in Base-flow Timing

|  |
| --- |
| Figure 8: Monthly change in base-flow contribution by hydroclimatic cluster for end-of-century projections relative to the 1980–2013 baseline. Colored lines show SSP2-4.5 (blue) and SSP5-8.5 (red). Clusters 1–3 exhibit modest winter–spring increases and strong summer declines (up to ~5 %), while Cluster 4 shows earlier winter peaks and sustained reductions through summer. These shifts indicate earlier groundwater discharge and reduced summer base-flow persistence under warmer futures. |

Projected base-flow anomalies show clear shifts in seasonal timing ([Figure 8](#fig-monthly-contrib)). Relative to the 1980–2013 baseline, future projections indicate earlier contributions in spring and pronounced reductions in late-summer base flow, particularly under SSP5-8.5. The greatest reductions in base-flow contribution are seen in June across all clusters. Clusters 1, 2, and 3 exhibit modest winter–spring increases of approximately 1–5 %, followed by strong summer declines reaching −4 to −5 % of annual base flow. Cluster 4, the most water-limited regime, displays an earlier winter peak (~5 %) and sustained reductions through summer and fall. These patterns suggest earlier groundwater release and diminished late-season persistence, consistent with warming-driven snowpack loss, reduced recharge, and earlier runoff timing in snowmelt-dominated basins.

## Discussion

Our results demonstrate widespread declines in base flow across western U.S. headwaters, with the strongest reductions occurring in snow-dominated and mixed-regime basins and persistent, though smaller, decreases in arid systems. These patterns reflect long-term reductions in annual discharge and a shift toward earlier seasonal contributions, consistent with warming and snowpack loss (Barnett, Adam, and Lettenmaier 2005; Carroll et al. 2024). Across regimes, antecedent moisture and snow emerge as dominant climatic controls, aligning with nationwide analyses (Ayers et al. 2022). LSTM projections indicate that these trends will intensify under both SSP2-4.5 and SSP5-8.5, with annual base flow declining by 45–65% by late century. Together, historical and projected responses provide a coherent, regional synthesis of how climate forcing alters groundwater-supported flow, advancing recent calls for integrated headwater assessments (Golden et al. 2025; Beck et al. 2013).

Declines in late-summer and autumn base flow ([Figure 4](#fig-hist_trends_monthly)) are consistent with earlier melt, rain-on-snow processes, and widespread shifts in hydrologic timing documented across the West (Musselman et al. 2018; Stewart, Cayan, and Dettinger 2005). More than 90% of long-term snow records show declining snow water equivalent, with the largest losses in transitional regimes (Mote et al. 2018). These changes reduce the volume and duration of snow-derived recharge, particularly in basins where snowmelt has historically been the primary source of summer base flow (Tague and Grant 2009; Safeeq et al. 2013). Our findings echo this: clusters with high snow reliance showed the strongest declines, while basins with greater subsurface storage or higher antecedent moisture exhibited more muted reductions.

Antecedent moisture emerged as the most consistent positive predictor, underscoring the importance of multiseasonal storage in mediating climate impacts. Similar conclusions from national-scale studies (Ayers et al. 2021, 2022) highlight that storage can partially offset warming-driven drying. Conversely, basins with minimal snowmelt support show greater sensitivity to temperature and evapotranspiration. Rising ET is already linked to widespread groundwater depletion (Condon, Atchley, and Maxwell 2020), a pattern reflected in Cluster 4 where high temperatures and limited storage constrain year-round base flow.

These regime-dependent responses reinforce the role of basin characteristics in shaping climate sensitivity. Snow-dominated basins show the largest relative declines; mixed-regime systems reflect combined snow and rainfall controls; and arid basins exhibit weak climatic elasticity due to limited storage and infrequent recharge. This gradient is consistent with broader critical-zone theory indicating that subsurface permeability and storage capacity govern base-flow resilience (Wlostowski et al. 2021).

The clustering framework provides a meaningful structure for interpreting this heterogeneity. By grouping basins using hydrologic and climatic attributes (Olden, Kennard, and Pusey 2012), we reduce within-group variability and reveal distinct climate–base-flow relationships. This approach parallels hydro-region classifications (Dethier et al. 2020) and supports empirical evidence for hydrologic similarity across physiographically coherent regions (Wagener et al. 2007; Beck et al. 2020). It also facilitates improved modeling: cluster-specific LSTMs captured regime-dependent nonlinearities and lagged climate responses more effectively than a single regional model (Kratzert, Herrnegger, et al. 2019; Lees et al. 2022).

Future projections show substantial declines across all regimes, with the largest relative reductions in snow-dominated and monsoon-influenced basins and the greatest volumetric losses in large mixed-regime systems. SSP2-4.5 and SSP5-8.5 trajectories remain similar through mid-century but diverge after ~2060, with high-emission futures producing 10–20% larger declines ([Figure 6](#fig-annualized_anomaly); [Figure 7](#fig-monthly-anomaly_period)). These contrasts underscore the long-term implications of climate variability for groundwater-supported flow (Diffenbaugh, Swain, and Touma 2015).

Uncertainty remains due to nonstationarity, climate-model structure, and our reliance on CESM2 as the sole GCM available with both scenarios in the WUS-D3 dataset. While this constraint narrows the range of climate uncertainty represented (Milly et al. 2008), the direction of change is robust across clusters. Expanding to multi-model ensembles would enhance confidence in late-century projections.

These changes carry important implications for ecosystems and water management. Earlier discharge and diminished late-season base flow challenge reservoir operations, drought planning, and environmental-flow programs that rely on historical timing. More intermittent low flows may reduce habitat quality and limit ecological resilience (Hammond et al. 2021). Declining headwater contributions are also likely to increase supply variability for downstream communities that depend heavily on forested catchments (Brown, Hobbins, and Ramirez 2005). Stream restoration programs highlight the growing recognition of headwaters as essential water infrastructure, with benefits that extend beyond hydrology to include water quality, sediment regulation, wildfire mitigation, and nonmarket ecosystem values (Mueller, Soder, and Springer 2019; Soder et al. 2022).

Climate-adaptive strategies will need to prioritize protection of groundwater–surface water connectivity, including restoration of meadow and riparian systems, enhancement of shallow storage, and targeted recharge interventions. These approaches align with recent calls to elevate headwaters within conservation and policy frameworks (Golden et al. 2025; McDonnell et al. 2007; Creed et al. 2017), particularly as intermittent-flow segments expand under aridification (Datry, Larned, and Tockner 2014).

Finally, integrating statistical models with LSTMs provides complementary strengths: statistical approaches identify dominant climatic controls, while LSTMs represent nonlinear and lagged hydrologic responses (Ayers et al. 2022; Murray, Ayers, and Brookfield 2023; Shen et al. 2023; Nearing et al. 2021). This hybrid modeling enhances interpretability and predictive power, offering a foundation for more robust, climate-informed water management.

## Conclusions

This study examined historical and projected changes in groundwater-supported base flow across western U.S. headwaters using long-term observations, hydroclimatic clustering, and cluster-specific LSTM models. Across the region, base flow has declined since the mid-20th century, and our projections indicate these reductions will intensify through the end of the century under both SSP2-4.5 and SSP5-8.5. In addition to volumetric losses, base-flow contributions are expected to shift earlier in the year, with diminished late-summer flows that reduce the buffering role groundwater plays during dry periods.

The four hydroclimatic clusters reveal distinct pathways through which climate forcing alters base flow. Snow-dominated basins show the largest relative declines, driven by snowpack loss and earlier melt; mixed regimes exhibit moderate declines shaped by a combination of snow loss and rainfall-derived recharge; and arid basins remain constrained by high evaporative demand and limited storage. These contrasts demonstrate how physiographic and climatic setting shape hydrologic sensitivity and provide a coherent framework for interpreting spatial heterogeneity in climate–base-flow responses. Antecedent moisture emerged as the strongest positive driver of base flow across regimes, emphasizing the importance of multiseasonal storage in sustaining groundwater contributions under increasingly variable climate conditions.

Projections of 45–65% regional reductions in annual base flow suggest meaningful implications for both ecosystems and downstream water security. Earlier runoff timing and reduced late-season groundwater discharge are likely to exacerbate summer water scarcity, ecological stress, and wildfire risk—conditions already intensifying across much of the West. These shifts also challenge water-management systems that rely on historical flow timing, underscoring the need to update reservoir operations, drought planning, and environmental-flow frameworks to account for declining and more variable base flow.

By integrating interpretable statistical models with LSTM neural networks, this study demonstrates how hybrid approaches can capture both linear climatic sensitivities and the nonlinear, lagged processes that govern groundwater–surface water interactions. More broadly, the results highlight that headwater catchments—already recognized through watershed-investment programs as critical water infrastructure—are likely to experience substantial reductions in their ability to supply reliable late-season flows. Sustaining the water, ecological, and societal benefits provided by these systems will require management strategies informed by both historical trends and forward-looking projections. Recognizing base flow as a key indicator of watershed resilience offers a path toward more climate-adaptive planning for the millions of people and ecosystems that depend on western headwater streams.

## Acknowledgments

This work was supported by the Arizona Board of Regents Technology and Research Initiative Fund (TRIF) program. We thank the U.S. Geological Survey for providing streamflow data through the National Water Information System, and the developers of the WUS-D3 climate dataset for making their downscaled projections publicly available. We also appreciate the constructive comments from reviewers, which improved the clarity and strength of this work. Any errors or omissions are the responsibility of the authors.

## Data Availability

The streamflow data used in this study are publicly available from the U.S. Geological Survey National Water Information System (https://waterdata.usgs.gov/nwis). The WUS-D3 climate data can be accessed at https://doi.org/10.5194/gmd-17-2265-2024. The code and processed datasets used for analysis and modeling are available at request. ## References {.unnumbered}

Arnold, J. G., P. M. Allen, R. Muttiah, and G. Bernhardt. 1995. “Automated Base Flow Separation and Recession Analysis Techniques.” *Ground Water* 33 (6): 1010–18. <https://doi.org/10.1111/j.1745-6584.1995.tb00046.x>.

Ayers, Jessica R., Gabriele Villarini, Keith Schilling, and Christopher Jones. 2021. “Projected Changes in Monthly Baseflow Across the <Span Style="font-Variant:small-Caps;">u.s. Midwest</Span>.” *International Journal of Climatology* 41 (12): 5536–49. <https://doi.org/10.1002/joc.7140>.

Ayers, Jessica R., Gabriele Villarini, Keith Schilling, Christopher Jones, Andrea Brookfield, Samuel C. Zipper, and William H. Farmer. 2022. “The Role of Climate in Monthly Baseflow Changes Across the Continental United States.” *Journal of Hydrologic Engineering* 27 (5): 04022006. <https://doi.org/10.1061/(ASCE)HE.1943-5584.0002170>.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier. 2005. “Potential Impacts of a Warming Climate on Water Availability in Snow-Dominated Regions.” *Nature* 438 (7066): 303–9. <https://doi.org/10.1038/NATURE04141>.

Beck, Hylke E., Albert I. J. M. van Dijk, Diego G. Miralles, Richard A. M. de Jeu, L. A. Sampurno Bruijnzeel, Tim R. McVicar, and Jaap Schellekens. 2013. “Global Patterns in Base Flow Index and Recession Based on Streamflow Observations from 3394 Catchments: Global Patterns in Base Flow Characteristics.” *Water Resources Research* 49 (12): 7843–63. <https://doi.org/10.1002/2013WR013918>.

Beck, Hylke E., Ming Pan, Peirong Lin, Jan Seibert, Albert I. J. M. van Dijk, and Eric F. Wood. 2020. “Global Fully Distributed Parameter Regionalization Based on Observed Streamflow From 4,229 Headwater Catchments.” *Journal of Geophysical Research: Atmospheres* 125 (17): e2019JD031485. <https://doi.org/10.1029/2019JD031485>.

Brinkerhoff, Craig B. 2024. “The Importance of Source Data in River Network Connectivity Modeling: A Review.” *Limnology and Oceanography* 69 (12): 3033–60. <https://doi.org/10.1002/lno.12706>.

Brown, Thomas C, Michael T Hobbins, and Jorge A Ramirez. 2005. “The Source of Water Supply in the United States.”

———. n.d. “The Source of Water Supply in the United States.”

Carroll, Rosemary W. H., Richard G. Niswonger, Craig Ulrich, Charuleka Varadharajan, Erica R. Siirila-Woodburn, and Kenneth H. Williams. 2024. “Declining Groundwater Storage Expected to Amplify Mountain Streamflow Reductions in a Warmer World.” *Nature Water* 2 (5): 419–33. <https://doi.org/10.1038/s44221-024-00239-0>.

Chapman, Tom. 1999. “A Comparison of Algorithms for Stream Flow Recession and Baseflow Separation.” *Hydrological Processes* 13 (5): 701–14. <https://doi.org/10.1002/(SICI)1099-1085(19990415)13:5<701::AID-HYP774>3.0.CO;2-2>.

Chen, Hao, and Ramesh S. V. Teegavarapu. 2021. “Spatial and Temporal Variabilities in Baseflow Characteristics Across the Continental USA.” *Theoretical and Applied Climatology* 143 (3): 1615–29. <https://doi.org/10.1007/s00704-020-03481-0>.

Collischonn, Walter, and Fernando Mainardi Fan. 2013. “Defining Parameters for Eckhardt’s Digital Baseflow Filter.” *Hydrological Processes* 27 (18): 2614–22. <https://doi.org/10.1002/hyp.9391>.

Condon, Laura E., Adam L. Atchley, and Reed M. Maxwell. 2020. “Evapotranspiration Depletes Groundwater Under Warming over the Contiguous United States.” *Nature Communications* 11 (1): 873. <https://doi.org/10.1038/s41467-020-14688-0>.

Creed, Irena F., Charles R. Lane, Jacqueline N. Serran, Laurie C. Alexander, Nandita B. Basu, Aram J. K. Calhoun, Jay R. Christensen, et al. 2017. “Enhancing Protection for Vulnerable Waters.” *Nature Geoscience* 10 (11): 809–15. <https://doi.org/10.1038/ngeo3041>.

Danabasoglu, G., J.-F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, et al. 2020. “The Community Earth System Model Version 2 (CESM2).” *Journal of Advances in Modeling Earth Systems* 12 (2): e2019MS001916. <https://doi.org/10.1029/2019MS001916>.

Datry, Thibault, Scott T. Larned, and Klement Tockner. 2014. “Intermittent Rivers: A Challenge for Freshwater Ecology.” *BioScience* 64 (3): 229–35. <https://doi.org/10.1093/biosci/bit027>.

Dethier, Evan N., Shannon L. Sartain, Carl E. Renshaw, and Francis J. Magilligan. 2020. “Spatially Coherent Regional Changes in Seasonal Extreme Streamflow Events in the United States and Canada Since 1950.” *Science Advances* 6 (49): eaba5939. <https://doi.org/10.1126/sciadv.aba5939>.

Diffenbaugh, Noah S., Daniel L. Swain, and Danielle Touma. 2015. “Anthropogenic Warming Has Increased Drought Risk in California.” *Proceedings of the National Academy of Sciences* 112 (13): 3931–36. <https://doi.org/10.1073/pnas.1422385112>.

Eckhardt, K. 2005. “How to Construct Recursive Digital Filters for Baseflow Separation.” *Hydrological Processes* 19 (2): 507–15. <https://doi.org/10.1002/hyp.5675>.

———. 2008. “A Comparison of Baseflow Indices, Which Were Calculated with Seven Different Baseflow Separation Methods.” *Journal of Hydrology* 352 (1): 168–73. <https://doi.org/10.1016/j.jhydrol.2008.01.005>.

Freeman, Mary C., Catherine M. Pringle, and C. Rhett Jackson. 2007. “Hydrologic Connectivity and the Contribution of Stream Headwaters to Ecological Integrity at Regional Scales.” *JAWRA Journal of the American Water Resources Association* 43 (1): 5–14. <https://doi.org/10.1111/j.1752-1688.2007.00002.x>.

Geological Survey (U.S.). 2025. “National Hydrography Dataset.” <https://nhdplus.com/NHDPlus/>.

Golden, Heather E., Jay R. Christensen, Hilary K. McMillan, Christa A. Kelleher, Charles R. Lane, Admin Husic, Li Li, et al. 2025. “Advancing the Science of Headwater Streamflow for Global Water Protection.” *Nature Water* 3 (1): 16–26. <https://doi.org/10.1038/s44221-024-00351-1>.

Gonzales, A. L., J. Nonner, J. Heijkers, and S. Uhlenbrook. 2009. “Comparison of Different Base Flow Separation Methods in a Lowland Catchment.” *Hydrology and Earth System Sciences* 13 (11): 2055–68. <https://doi.org/10.5194/hess-13-2055-2009>.

Gorelick, Steven M., and Chunmiao Zheng. 2015. “Global Change and the Groundwater Management Challenge.” *Water Resources Research* 51 (5): 3031–51. <https://doi.org/10.1002/2014WR016825>.

Hamed, Khaled H., and A. Ramachandra Rao. 1998. “A Modified Mann-Kendall Trend Test for Autocorrelated Data.” *Journal of Hydrology* 204 (1): 182–96. <https://doi.org/10.1016/S0022-1694(97)00125-X>.

Hammond, John C., Margaret Zimmer, Margaret Shanafield, Kendra Kaiser, Sarah E. Godsey, Meryl C. Mims, Samuel C. Zipper, et al. 2021. “Spatial Patterns and Drivers of Nonperennial Flow Regimes in the Contiguous United States.” *Geophysical Research Letters* 48 (2): e2020GL090794. <https://doi.org/10.1029/2020GL090794>.

Ikotun, Abiodun M., Absalom E. Ezugwu, Laith Abualigah, Belal Abuhaija, and Jia Heming. 2023. “K-Means Clustering Algorithms: A Comprehensive Review, Variants Analysis, and Advances in the Era of Big Data.” *Information Sciences* 622 (April): 178–210. <https://doi.org/10.1016/j.ins.2022.11.139>.

Imberger, Moss, Belinda E. Hatt, Stephanie Brown, Matthew J. Burns, Ryan M. Burrows, and Christopher J. Walsh. 2023. “Headwater Streams in an Urbanizing World.” *Freshwater Science* 42 (3): 323–36. <https://doi.org/10.1086/726682>.

Institute of Hydrology. 1980. “Low Flow Studies Report 3 Catchment Characteristic Estimation Manual.”

Kratzert, Frederik, Mathew Herrnegger, Daniel Klotz, Sepp Hochreiter, and Günter Klambauer. 2019. “NeuralHydrology – Interpreting LSTMs in Hydrology.” In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, edited by Wojciech Samek, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller, 347–62. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-28954-6_19>.

Kratzert, Frederik, Daniel Klotz, Mathew Herrnegger, Alden K. Sampson, Sepp Hochreiter, and Grey S. Nearing. 2019. “Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning.” *Water Resources Research* 55 (12): 11344–54. <https://doi.org/10.1029/2019WR026065>.

Lees, Thomas, Steven Reece, Frederik Kratzert, Daniel Klotz, Martin Gauch, Jens De Bruijn, Reetik Kumar Sahu, Peter Greve, Louise Slater, and Simon J. Dadson. 2022. “Hydrological Concept Formation Inside Long Short-Term Memory (LSTM) Networks.” *Hydrology and Earth System Sciences* 26 (12): 3079–3101. <https://doi.org/10.5194/hess-26-3079-2022>.

McDonnell, J. J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, et al. 2007. “Moving Beyond Heterogeneity and Process Complexity: A New Vision for Watershed Hydrology.” *Water Resources Research* 43 (7). <https://doi.org/10.1029/2006WR005467>.

Milly, P. C. D., Julio Betancourt, Malin Falkenmark, Robert M. Hirsch, Zbigniew W. Kundzewicz, Dennis P. Lettenmaier, and Ronald J. Stouffer. 2008. “Stationarity Is Dead: Whither Water Management?” *Science* 319 (5863): 573–74. <https://doi.org/10.1126/science.1151915>.

Mote, Philip W., Sihan Li, Dennis P. Lettenmaier, Mu Xiao, and Ruth Engel. 2018. “Dramatic Declines in Snowpack in the Western US.” *Npj Climate and Atmospheric Science* 1 (1): 2. <https://doi.org/10.1038/s41612-018-0012-1>.

Mroczek, Caelum, Abraham E. Springer, Neha Gupta, Temuulen Sankey, and Benjamin Lucas. 2025. “Regional Base-Flow Index in Arid Landscapes Using Machine Learning and Instrumented Records.” *Journal of Hydrology: Regional Studies* 62: 102778. https://doi.org/<https://doi.org/10.1016/j.ejrh.2025.102778>.

Mueller, Julie M., Adrienne B. Soder, and Abraham E. Springer. 2019. “Valuing Attributes of Forest Restoration in a Semi-Arid Watershed.” *Landscape and Urban Planning* 184 (April): 78–87. <https://doi.org/10.1016/j.landurbplan.2018.12.012>.

Murray, J., J. Ayers, and A. Brookfield. 2023. “The Impact of Climate Change on Monthly Baseflow Trends Across Canada.” *Journal of Hydrology* 618 (March): 129254. <https://doi.org/10.1016/j.jhydrol.2023.129254>.

Musselman, Keith N., Flavio Lehner, Kyoko Ikeda, Martyn P. Clark, Andreas F. Prein, Changhai Liu, Mike Barlage, and Roy Rasmussen. 2018. “Projected Increases and Shifts in Rain-on-Snow Flood Risk over Western North America.” *Nature Climate Change* 8 (9): 808–12. <https://doi.org/10.1038/s41558-018-0236-4>.

Nathan, R. J., and T. A. McMahon. 1990. “Evaluation of Automated Techniques for Base Flow and Recession Analyses.” *Water Resources Research* 26 (7): 1465–73. <https://doi.org/10.1029/WR026i007p01465>.

Nearing, Grey S., Frederik Kratzert, Alden Keefe Sampson, Craig S. Pelissier, Daniel Klotz, Jonathan M. Frame, Cristina Prieto, and Hoshin V. Gupta. 2021. “What Role Does Hydrological Science Play in the Age of Machine Learning?” *Water Resources Research* 57 (3): e2020WR028091. <https://doi.org/10.1029/2020WR028091>.

Newman, Matthew, Michael A. Alexander, Toby R. Ault, Kim M. Cobb, Clara Deser, Emanuele Di Lorenzo, Nathan J. Mantua, et al. 2016. “The Pacific Decadal Oscillation, Revisited,” June. <https://doi.org/10.1175/JCLI-D-15-0508.1>.

Olden, Julian D., Mark J. Kennard, and Bradley J. Pusey. 2012. “A Framework for Hydrologic Classification with a Review of Methodologies and Applications in Ecohydrology.” *Ecohydrology* 5 (4): 503–18. <https://doi.org/10.1002/eco.251>.

Peel, M. C., B. L. Finlayson, and T. A. McMahon. 2007. “Updated World Map of the Köppen-Geiger Climate Classification.” *Hydrology and Earth System Sciences* 11 (5): 1633–44. <https://doi.org/10.5194/hess-11-1633-2007>.

Pelto, Mauri S. 2008. “Impact of Climate Change on North Cascade Alpine Glaciers, and Alpine Runoff.” *Northwest Science* 82 (1): 65–75. <https://doi.org/10.3955/0029-344X-82.1.65>.

Rahimi, Stefan, Lei Huang, Jesse Norris, Alex Hall, Naomi Goldenson, Will Krantz, Benjamin Bass, et al. 2024. “An Overview of the Western United States Dynamically Downscaled Dataset (WUS-D3).” *Geoscientific Model Development* 17 (6): 2265–86. <https://doi.org/10.5194/gmd-17-2265-2024>.

Razavi, Saman. 2021. “Deep Learning, Explained: Fundamentals, Explainability, and Bridgeability to Process-Based Modelling.” *Environmental Modelling & Software* 144 (October): 105159. <https://doi.org/10.1016/j.envsoft.2021.105159>.

Rice, Joshua S., Ryan E. Emanuel, James M. Vose, and Stacy A. C. Nelson. 2015. “Continental u.s. Streamflow Trends from 1940 to 2009 and Their Relationships with Watershed Spatial Characteristics.” *Water Resources Research* 51 (8): 6262–75. <https://doi.org/10.1002/2014WR016367>.

Saedi, Javad, Mohammad Reza Sharifi, Ali Saremi, and Hossein Babazadeh. 2022. “Assessing the Impact of Climate Change and Human Activity on Streamflow in a Semiarid Basin Using Precipitation and Baseflow Analysis.” *Scientific Reports* 12 (1): 9228. <https://doi.org/10.1038/s41598-022-13143-y>.

Safeeq, Mohammad, Gordon E. Grant, Sarah L. Lewis, and Christina. L. Tague. 2013. “Coupling Snowpack and Groundwater Dynamics to Interpret Historical Streamflow Trends in the Western United States.” *Hydrological Processes* 27 (5): 655–68. <https://doi.org/10.1002/hyp.9628>.

Santhi, C., P. M. Allen, R. S. Muttiah, J. G. Arnold, and P. Tuppad. 2008. “Regional Estimation of Base Flow for the Conterminous United States by Hydrologic Landscape Regions.” *Journal of Hydrology* 351 (1): 139–53. <https://doi.org/10.1016/j.jhydrol.2007.12.018>.

Shen, Chaopeng, Alison P. Appling, Pierre Gentine, Toshiyuki Bandai, Hoshin Gupta, Alexandre Tartakovsky, Marco Baity-Jesi, et al. 2023. “Differentiable Modelling to Unify Machine Learning and Physical Models for Geosciences.” *Nature Reviews Earth & Environment* 4 (8): 552–67. <https://doi.org/10.1038/s43017-023-00450-9>.

Sivakumar, Bellie. 2009. “Nonlinear Dynamics and Chaos in Hydrologic Systems: Latest Developments and a Look Forward.” *Stochastic Environmental Research and Risk Assessment* 23 (7): 1027–36. <https://doi.org/10.1007/s00477-008-0265-z>.

Sloto, Ronald, and Michele Crouse. 1996. “HYSEP: A Computer Program for Streamflow Hydrograph Separation and Analysis.” 96-4040. USGS. <https://doi.org/10.3133/wri964040>.

Smakhtin, V. U. 2001. “Low Flow Hydrology: A Review.” *Journal of Hydrology* 240 (3-4): 147–86. <https://doi.org/10.1016/S0022-1694(00)00340-1>.

Soder, Adrienne B., Julie M. Mueller, Abraham E. Springer, and Katelyn E. LaPine. 2022. “Geospatial Analysis of Nonmarket Values to Prioritize Forest Restoration.” *Land* 11 (9): 1387. <https://doi.org/10.3390/land11091387>.

Stewart, Iris T., Daniel R. Cayan, and Michael D. Dettinger. 2005. “Changes Toward Earlier Streamflow Timing Across Western North America,” April. <https://doi.org/10.1175/JCLI3321.1>.

Tague, Christina, and Gordon E. Grant. 2009. “Groundwater Dynamics Mediate Low-Flow Response to Global Warming in Snow-Dominated Alpine Regions: GROUNDWATER DYNAMICS AND LOW-FLOW RESPONSE.” *Water Resources Research* 45 (7). <https://doi.org/10.1029/2008WR007179>.

Tan, Xuejin, Bingjun Liu, and Xuezhi Tan. 2020. “Global Changes in Baseflow Under the Impacts of Changing Climate and Vegetation.” *Water Resources Research* 56 (9): e2020WR027349. <https://doi.org/10.1029/2020WR027349>.

Taylor, Richard G., Bridget Scanlon, Petra Döll, Matt Rodell, Rens Van Beek, Yoshihide Wada, Laurent Longuevergne, et al. 2013. “Ground Water and Climate Change.” *Nature Climate Change* 3 (4): 322–29. <https://doi.org/10.1038/nclimate1744>.

Thomas C Winter, Judson W Harvey, O Lehn Franke, and William M Alley. 1998. “Ground Water and Surface Water: A Single Resource.”

Tillman, Fred D, Natalie K. Day, Matthew P. Miller, Olivia L. Miller, Christine A. Rumsey, Daniel R. Wise, Patrick C. Longley, and Morgan C. McDonnell. 2022. “A Review of Current Capabilities and Science Gaps in Water Supply Data, Modeling, and Trends for Water Availability Assessments in the Upper Colorado River Basin.” *Water* 14 (23): 3813. <https://doi.org/10.3390/w14233813>.

U. S. Geological Survey. 2018. “Water Science Glossary.” <https://www.usgs.gov/water-science-school/science/water-science-glossary>.

US Army Corps of Engineers. n.d. “National Inventory of Dams.”

Viviroli, Daniel, Rolf Weingartner, and Bruno Messerli. 2003. “Assessing the Hydrological Significance of the World’s Mountains.” *Mountain Research and Development* 23 (1): 32–40. <https://doi.org/10.1659/0276-4741(2003)023[0032:ATHSOT]2.0.CO;2>.

Wagener, Thorsten, Murugesu Sivapalan, Peter Troch, and Ross Woods. 2007. “Catchment Classification and Hydrologic Similarity.” *Geography Compass* 1 (4): 901–31. <https://doi.org/10.1111/j.1749-8198.2007.00039.x>.

Wlostowski, Adam N., Noah Molotch, Suzanne P. Anderson, Susan L. Brantley, Jon Chorover, David Dralle, Praveen Kumar, et al. 2021. “Signatures of Hydrologic Function Across the Critical Zone Observatory Network.” *Water Resources Research* 57 (3): e2019WR026635. <https://doi.org/10.1029/2019WR026635>.

Woodhouse, Connie A., and Bradley Udall. 2022. “Upper Gila, Salt, and Verde Rivers: Arid Land Rivers in a Changing Climate.” *Earth Interactions* 26 (1): 1–14. <https://doi.org/10.1175/EI-D-21-0014.1>.

Xie, Jiaxin, Xiaomang Liu, Kaiwen Wang, Tiantian Yang, Kang Liang, and Changming Liu. 2020. “Evaluation of Typical Methods for Baseflow Separation in the Contiguous United States.” *Journal of Hydrology* 583 (April): 124628. <https://doi.org/10.1016/j.jhydrol.2020.124628>.