

**1 Headwaters of the Western US: Base-flow Analyses and
2 Projections**

3 Caelum Mroczek¹, Abraham E Springer¹, Benjamin Lucas²

4 ¹School of Earth and Sustainability, Northern Arizona University,

5 ²Department of Mathematics and Statistics, Northern Arizona University,

6 **Abstract**

7 Headwater streams make up nearly 88% of the stream network in the western
 8 United States and supply most of the region's surface water. These systems are
 9 highly sensitive to climate change, as warming, snowpack loss, and drought al-
 10 ter groundwater recharge and the timing of streamflow. As surface-water inputs
 11 decline, streams increasingly depend on groundwater to sustain base flow, under-
 12 scoring the need to understand long-term trends and climate controls on these
 13 flows. We analyzed 75 years of streamflow records (1950–2024) from 115 headwa-
 14 ter basins across the western United States to quantify historical patterns, climatic
 15 drivers, and projected changes in base flow. Statistical analyses (base-flow sepa-
 16 ration, Mann–Kendall, mixed-effects modeling) and cluster-specific Long Short-Term
 17 Memory (LSTM) models trained on dynamically downscaled climate data (WUS-D3,
 18 CESM2) were used to assess both historical and future conditions. Historical results
 19 show widespread base-flow declines, especially in early summer, linked to warm-
 20 ing, reduced snowmelt, and declining antecedent moisture. Antecedent moisture is
 21 the dominant positive driver across all regions, while snow and temperature exert
 22 regime-dependent negative effects. Future projections under SSP2-4.5 and SSP5-8.5
 23 indicate continued base-flow declines of 45–65% by late century, with earlier sea-
 24 sonal peaks and shorter summer flow duration. Snowmelt-dominated and arid basins
 25 show the largest relative declines, while mixed-regime basins contribute the greatest
 26 volumetric losses. These results highlight the vulnerability of groundwater-supported
 27 streamflow to warming and drying trends and demonstrate the value of combining
 28 statistical and machine-learning approaches to assess regional-scale climate sensitiv-
 29 ity.

30 **1 Introduction**

31 Increasing water demand, land-use change, and climate-driven shifts in precipita-
 32 tion are intensifying water scarcity across the western United States (Diffenbaugh
 33 et al., 2015; Taylor et al., 2013). Headwater catchments are the primary sources
 34 of streamflow, sustaining downstream ecosystems, agriculture, and communities.
 35 Globally, headwater systems make up approximately 80% of river networks (Golden
 36 et al., 2025) and contribute disproportionately to total streamflow, particularly in
 37 mountainous and semi-arid regions (CITE). These catchments are critical zones of
 38 groundwater–surface water interaction, where recharge, storage, and discharge pro-
 39 cesses regulate the timing and persistence of streamflow (Thomas C Winter et al.,
 40 1998).

41 Base flow is the sustained portion of streamflow that is derived from groundwater
 42 discharge or other delayed sources (U. S. Geological Survey, 2018). It provides a
 43 surface-water view of the groundwater system, maintaining perennial reaches, buffer-
 44 ing flow during drought, supporting aquatic habitats, and sustaining downstream
 45 water supply. Base flow reflects the integrated response of the watershed to precip-
 46 itation, snow accumulation and melt, evapotranspiration, and subsurface storage
 47 (Smakhtin, 2001). As such, base-flow magnitude and timing provide key indicators
 48 of groundwater availability and watershed resilience under changing climate condi-
 49 tions (Diffenbaugh et al., 2015; Tague & Grant, 2009).

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

58 Previous studies have advanced understanding of base-flow variability (Beck et
 59 al., 2013; Santhi et al., 2008) and its relationship to climate drivers (Ayers et al.,
 60 2022; Tan et al., 2020) in different settings. Extending these analyses to incorporate
 61 climate projections is critical for anticipating how groundwater-supported flows, es-
 62 pecially in headwater systems, will respond to future climate change. As noted by
 63 Golden et al. (2025), a key opportunity in headwater research is to move beyond iso-
 64 lated studies toward an integrative understanding of headwater flow regimes across
 65 space and time. This integrated approach is limited by the scarcity of adequately
 66 gauged headwater sites and by large-scale analyses that generalize hydrologic behav-
 67 ior across diverse climatic and physiographic settings.

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

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

89 Our approach forms a regional modelling framework to quantify base-flow changes
 90 and identify the climatic drivers shaping long-term trends. Specifically, this study
 91 aims to:

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

98 2 Data & Methods

99 2.1 Data

100 Daily mean discharge data were obtained for 115 U.S. Geological Survey (USGS)
 101 streamgages located across the eleven western states Figure 1. Streamflow records
 102 were downloaded from the USGS National Water Information System for the period
 103 of water years 1950–2024 (Geological Survey (U.S.), 2025). Sites were selected to
 104 have > 90% daily data completeness, with no gaps exceeding three consecutive years,
 105 and to have catchments located entirely within U.S. borders to avoid data restric-
 106 tions. This period of record provides sufficient length to detect long-term trends
 107 and hydrologic regime shifts and encompasses three phases of the Pacific Decadal
 108 Oscillation (Newman et al., 2016), reducing the likelihood that results are biased
 109 by decadal climate variability. Sites with a dam on the main stem, as reported in

110 the National Inventory of Dams (US Army Corps of Engineers, n.d.), were excluded
 111 to reduce the influence of direct flow regulation. Drainage areas ranged from 4 to
 112 21,167 km², with a mean of 1,524 km² and a median of 558 km²

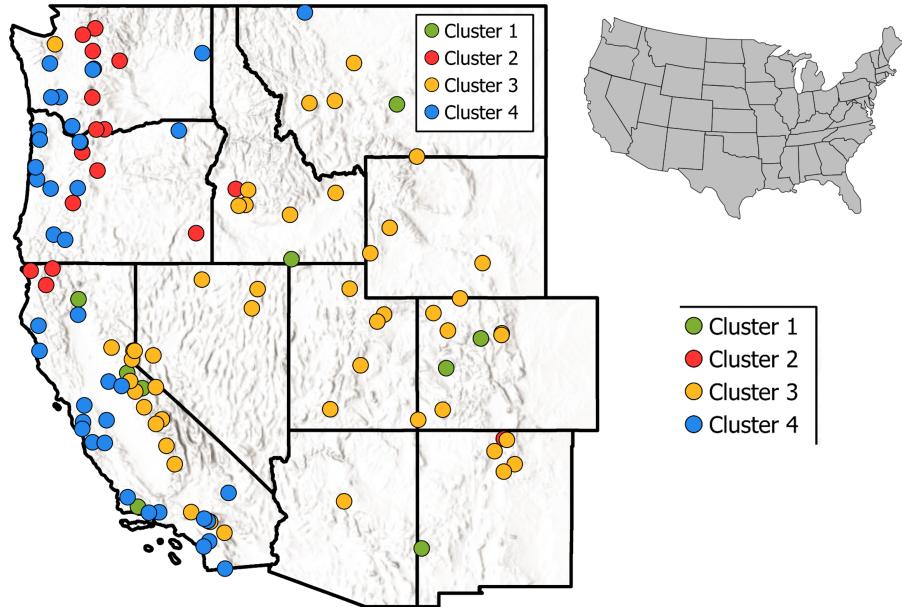
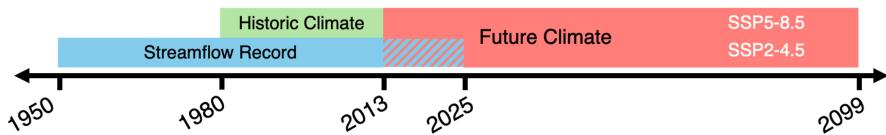
a**b**

Figure 1: MOCKUP > 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.

113 Headwaters have been operational defined as streams with Strahler stream order
 114 1 or 2 (Freeman et al., 2007; Golden et al., 2025; Imberger et al., 2023). It's been
 115 found that large-scale stream network datasets may underestimate the extent of
 116 localized headwater reaches (Brinkerhoff, 2024), indicating that these delineated
 117 headwater streams likely represent a minimum bound on actual headwater extent.
 118 To overcome this deficiency, in this study, streamgages located on streams of order
 119 1 to 3 were included to capture the cumulative processes of headwater systems.
 120 By focusing on headwater systems, we overcome issues of scale that arise from the
 121 coarse resolution of national and global models and from the limited availability of
 122 catchments with adequate monitoring for both streamflow and climate variables.

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

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

155 **2.2 Base-flow Separation**

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

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

$$b_t = \frac{(1 - BFI_{max})\alpha b_{t-1} + (1 + \alpha)BFI_{max}Q_t}{(1 - \alpha BFI_{max})}$$

170 where b_t is the filtered base-flow response at time step t , Q_t is the observed stream-
 171 flow at time step t , b_{t-1} is the base-flow response at the previous time step, α is
 172 the recession constant, and BFI_{max} is the maximum possible base-flow index for

173 the catchment. The recession constant α was estimated for each site through hydro-
 174 graph recession analysis. BFI_{max} was determined for each site using the backwards
 175 filter method proposed by Collischonn & Fan (2013), which allows BFI_{max} to be
 176 estimated from α without requiring site-specific hydrogeologic field data. The Eck-
 177 hardt filter was selected because it has shown strong performance in diverse hydro-
 178 logic settings across the contiguous United States and has been recommended as a
 179 preferred base-flow separation method in large-sample studies (Xie et al., 2020).

180 2.3 Hydroclimate Clustering

181 To classify study basins into groups with similar hydroclimatic regimes, we applied k-
 182 means clustering to basin-averaged climate and hydrologic variables. Clustering was
 183 conducted on long-term records of base flow magnitude and variability, precipitation,
 184 snow-derived precipitation, mean temperature, antecedent moisture, and static phys-
 185 iographic attributes including drainage area, relief, and elevation. To compare across
 186 watersheds with different areas, precipitation, and discharge, we normalized variables
 187 (z-scores) prior to clustering to ensure equal weighting. The optimal number of clus-
 188 ters was selected based on a combination of the elbow method and interpretability
 189 of known hydrologic regimes in the western United States. K-means was chosen be-
 190 cause it is an efficient method for partitioning basins into internally cohesive groups
 191 which minimize within-group variance and enhances the detection of coherent re-
 192 gional trends in streamflow across diverse hydroclimatic settings (Dethier et al.,
 193 2020). This approach is well suited for large-sample hydrology applications where
 194 hydroclimatic gradients are continuous rather than categorical (Ikotun et al., 2023).
 195 The resulting clusters were used to stratify subsequent statistical and machine learn-
 196 ing analyses, allowing for the identification of climate–base flow relationships and
 197 projected changes within distinct hydrologic response regimes.

198 2.4 Statistical Models

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

209 We used a linear mixed-effects modeling framework to quantify relationships be-
 210 tween monthly base flow and climate variables. Fixed effects included precipitation,
 211 snow, temperature, and antecedent moisture; which capture broad regional-scale
 212 climate drivers. A random intercept for each catchment accounts for site-level vari-
 213 ability in base flow that is not explained by the fixed predictors. All models were
 214 fit using log-transformed monthly base flow as the response variable to reduce skew-
 215 ness, stabilize variance, and improve model performance. To address strong multi-
 216 collinearity among the temperature variables (mean, minimum, and maximum daily
 217 temperature), we retained only mean temperature in the mixed-effects model to
 218 ensure interpretability of parameter estimates. This constraint was applied only in
 219 the mixed-effects framework; the LSTM modeling framework retained all three tem-
 220 perature variables, as deep learning approaches do not require independence among
 221 predictors and are capable of capturing nonlinear relationships and interactions
 222 (Razavi, 2021).

223 **2.5 LSTM Neural Network Modeling**

224 ***2.5.1 Model Framework***

225 This model is designed to predict monthly stream base flow using both climate
 226 history and watershed characteristics. It combines two types of input data: (1) a
 227 sequence of monthly climate variables over the past 24-months, and (2) static basin
 228 attributes like latitude, elevation, and area.

229 To capture monthly flow dynamics, the model was trained using overlapping 24-
 230 month climate input sequences (e.g., months 1–24, 2–25, 3–26, etc.), where each
 231 sequence was used to predict base flow in the subsequent month (Figure 2). This
 232 sliding-window approach increases the number of training samples and enables the
 233 model to learn temporal dependencies and incremental changes in base flow across
 234 consecutive months. The climate sequence is combined with the static features and
 235 passed through two fully connected layers (also called dense layers) (Hochreiter
 236 & Schmidhuber, 1997), with dropout included to reduce overfitting (Srivastava et
 237 al., 2014). The final output is a single number: the predicted log-transformed base
 238 flow for the subsequent month. The model is trained to minimize the difference
 239 between its predictions and observed log(base flow), using a loss function based on
 240 mean absolute error (MAE). This hybrid architecture allows the model to learn both
 241 temporal patterns and site-specific differences in hydrologic behavior.

242 ***2.5.2 Model Training and Testing***

243 Model training and hyperparameter selection were conducted separately for each hy-
 244 droclimatic cluster. We tested a range of LSTM configurations varying the number
 245 of hidden units, dropout rates, learning rates, and batch sizes. The final hyper-
 246 parameters were chosen based on minimizing cross-validated mean absolute error
 247 (MAE) during the 2006–2013 period (Table S1). Each model used a 24-month input
 248 sequence of climate predictors (precipitation, temperature, antecedent moisture in-
 249 dices, snow fraction, and seasonal harmonics) along with static basin characteristics
 250 (elevation, relief, drainage area, latitude, longitude). Inputs were standardized using
 251 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)

Data Type	Predictor	Description
Snow / Moisture	ppt_change	Monthly precipitation change (current minus prior month)
	snow	Precipitation falling when $t_{2\max} < 0$ °C (mm)
	prec_moist	Antecedent moisture (3-month precipitation sum)
	prec_moist_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 (km ²)
	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.

3 Results

3.1 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, and the spatial distribution of clusters across the study area is shown in Figure 1.

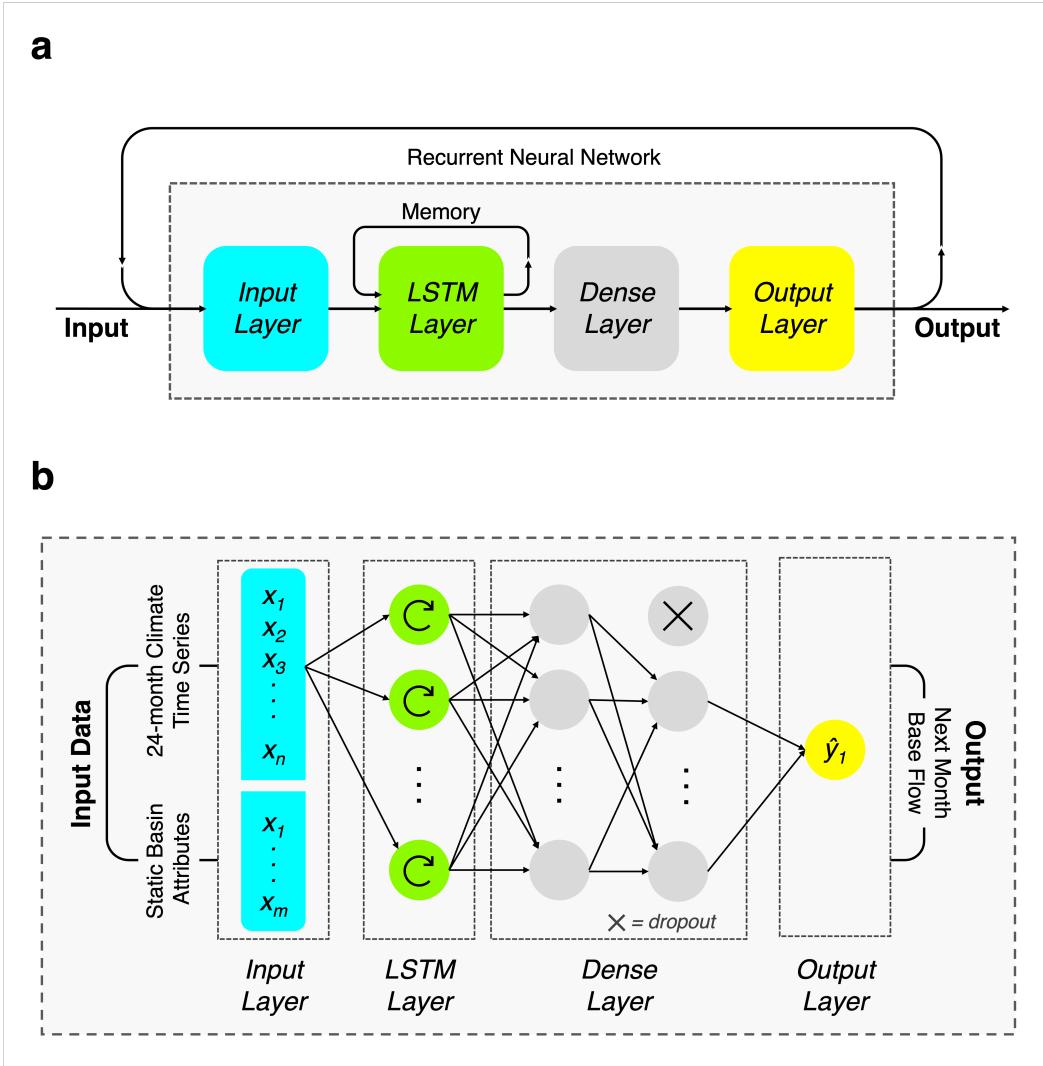


Figure 2: Structure of the Long Short-term Memory (LSTM) model used to predict monthly base flow. (a) Conceptual layout of the LSTM-based recurrent neural network. (b) Model configuration showing 24-month climate inputs and static basin attributes processed through LSTM, dense, and output layers to predict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

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 (σ/μ)).

Cluster	n	Area (km ²)	Elev. Mean (m)	Relief (m)	Prec. Mean (mm)	Temp. Mean (°C)	Flow Mean (cfs)	Base Flow CV
Monthly Base Flow								
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

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

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

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

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

301 Across all clusters, precipitation peaks in December, contributing roughly 15–18%
 302 of the annual total, and remains elevated through February (Figure 3). Snowmelt-
 303 dominated clusters (1 and 3) exhibit the largest variability in late-spring base flow,
 304 whereas Cluster 4 shows the strongest temperature-driven suppression of base flow
 305 in the dry season. These distinct seasonal signatures and sensitivities highlight the
 306 potential for divergent base-flow responses to future climate change across hydrocli-
 307 matic regimes.

3.2 Historical Base-flow Analysis

3.2.1 Climate-Base Flow Relationship

310 When the mixed-effects model was analyzed across all sites Table 3, it revealed
 311 that antecedent moisture was the most consistently positive and influential driver

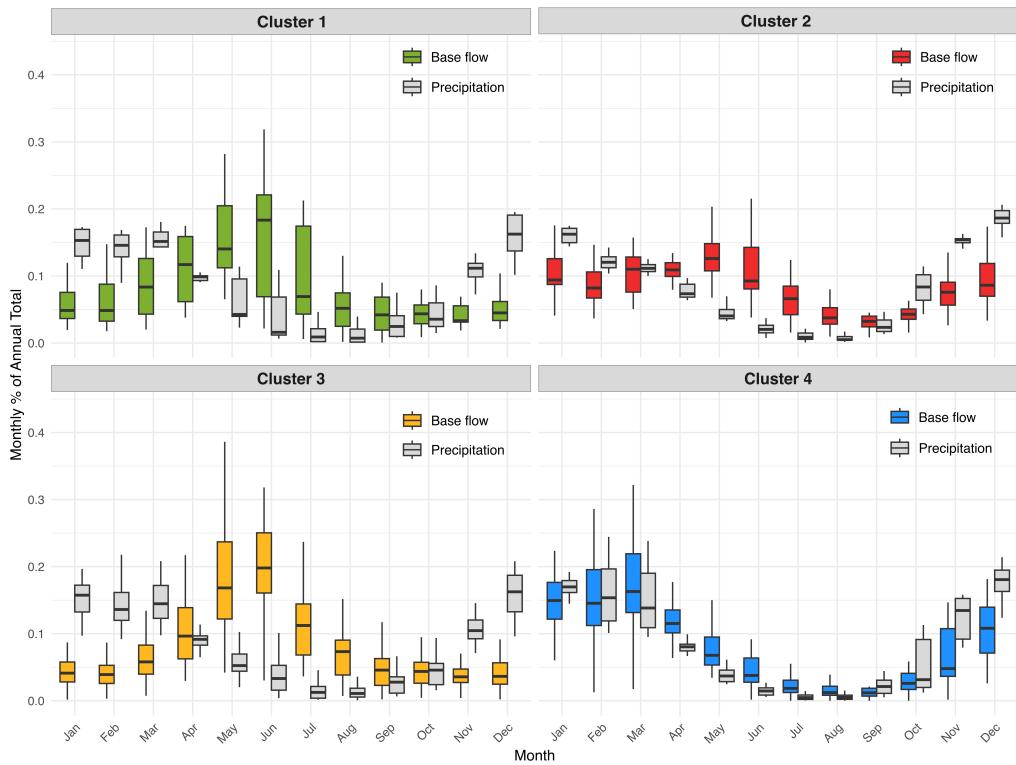


Figure 3: Monthly distribution of base flow and precipitation as a percentage of the annual total, stratified by hydroclimatic cluster. Base-flow boxes are colored by cluster (Cluster 1: green, Cluster 2: red, Cluster 3: yellow, Cluster 4: blue), precipitation is shown in gray.

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), 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

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
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.

3.2.2 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; Figure 4). 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 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	Trend	Pct. Media	Pct. Sig.	Pct. Sig.
														Up	Down	
1	1.22	0.38	5.60	3.92	0.50	-	-	0.08	-	-	3.14	1.10	0.44	20.69	16.38	
						1.31	1.71		3.22	0.64						
2	-	-	-	-	-	-	-	-	-	-	-	-	-	1.47	30.39	
	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			
3	-	-	1.12	1.12	0.04	-	-	-	-	-	-	-	-	8.60	17.37	
	0.61	0.95				2.33	3.10	3.26	2.25	0.92	0.77	1.81	0.94			
4	-	-	0.32	0.00	-	-	-	-	-	-	-	-	-	12.79	18.84	
	0.13	4.27			0.11	0.37	1.92	2.75	3.54	3.28	0.88	2.35	1.40			

3.3 Projected Base-flow Analysis

3.3.1 Model Validation

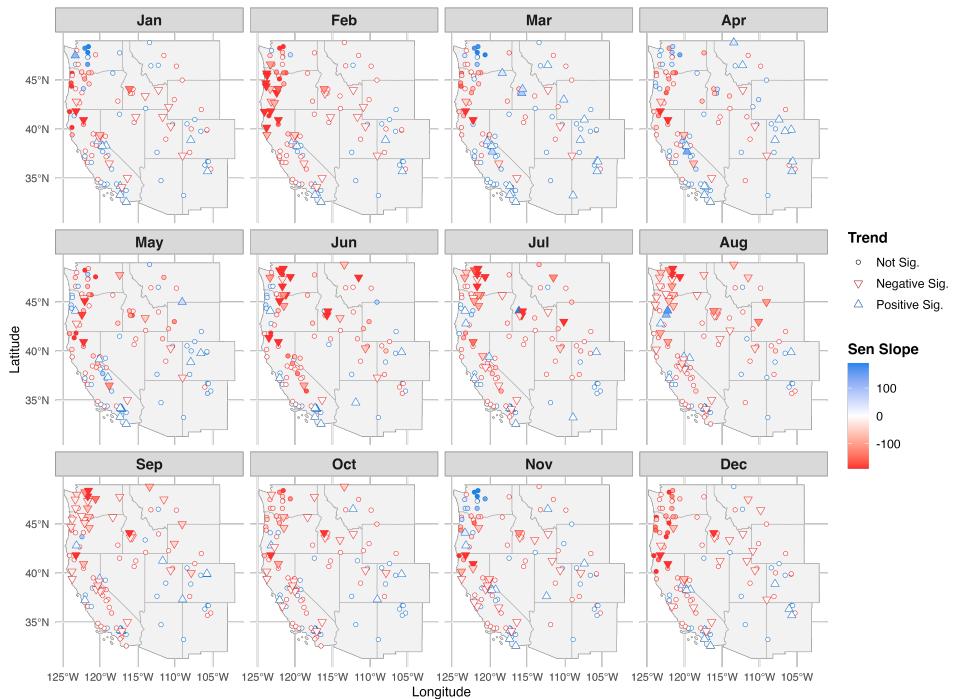


Figure 4: Historic monthly base-flow trends across western U.S. headwater basins (1950–2024), colored by Theil-Sen slope and shaped by statistical significance. Upward- and downward-pointing triangles indicate significant increasing or decreasing trends ($p < 0.05$), respectively, while circles denote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

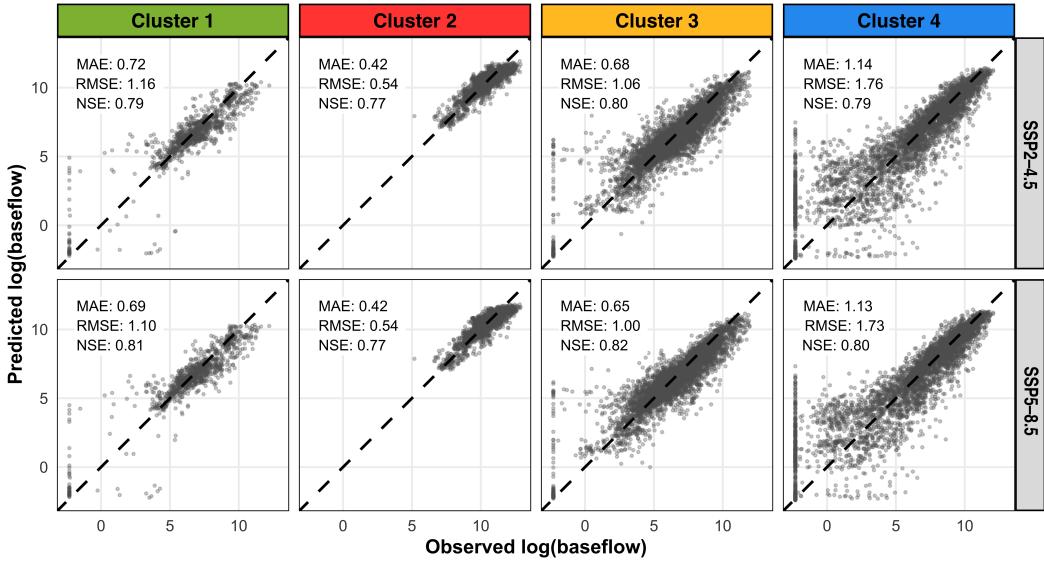


Figure 5: Observed vs. predicted monthly log-transformed base flow for the ensemble LSTM models across four hydroclimatic clusters under two climate scenarios (SSP2-4.5 and SSP5-8.5). Each panel shows one cluster and scenario combination. The 1:1 dashed line indicates perfect prediction. Performance 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; Table 6). 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, 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 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.

3.3.2 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 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 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 yr¹. Mid-century reductions deepen to –40 to –50 % (3.6–4.1 million AF yr¹), and by late-century, annual anomalies reach –45 to –65 %, or roughly 3.7–5.8 million AF yr¹, relative to the 1980–2013 baseline (Figure 6 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.

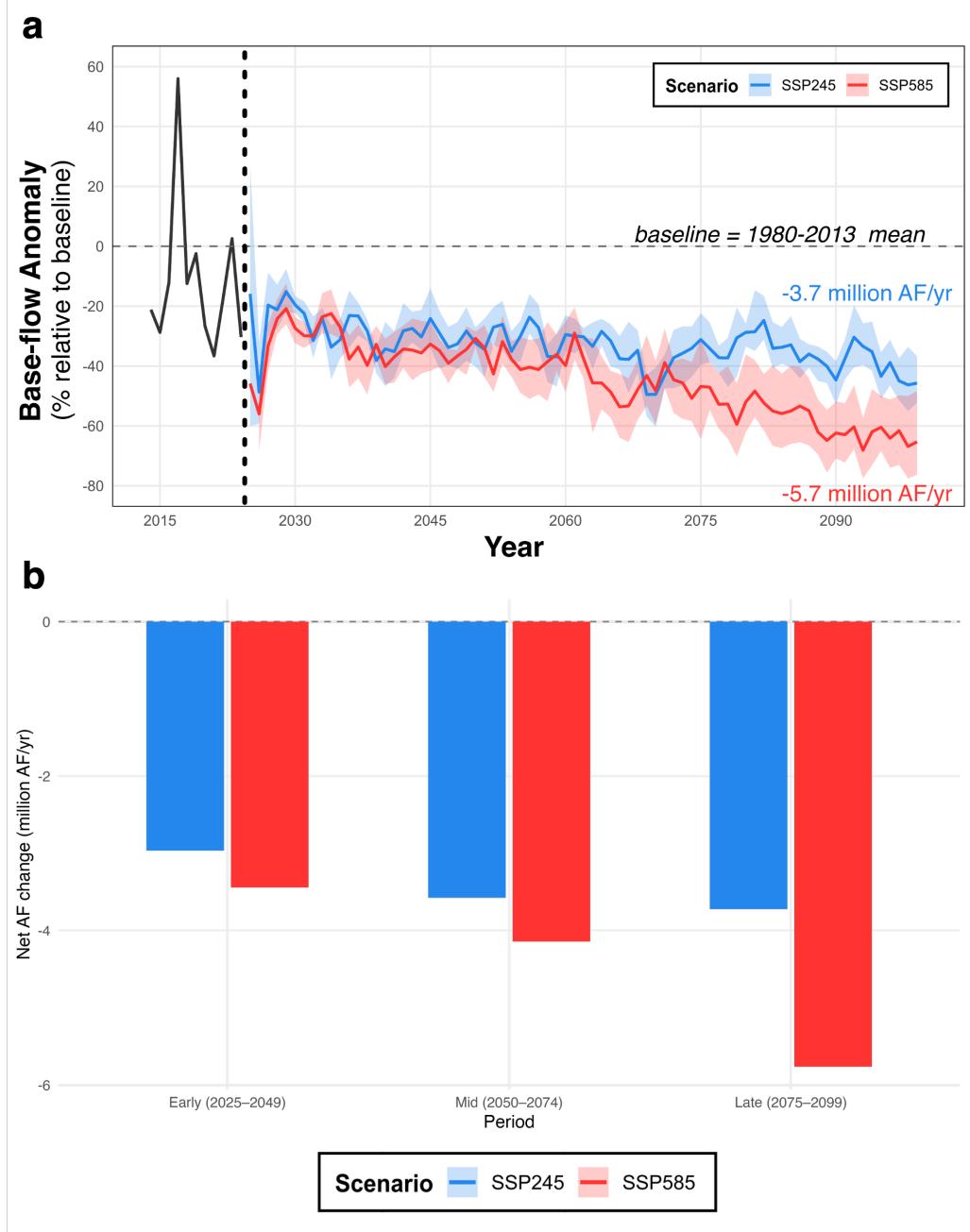


Figure 6: Projected magnitude of base-flow change across western U.S. headwaters under SSP2-4.5 and SSP5-8.5 scenarios. (a) Regional annual base-flow anomalies (percent difference from the 1980–2013 mean) with shaded 25–75 % ranges across basins. Both scenarios project sustained declines through the 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.

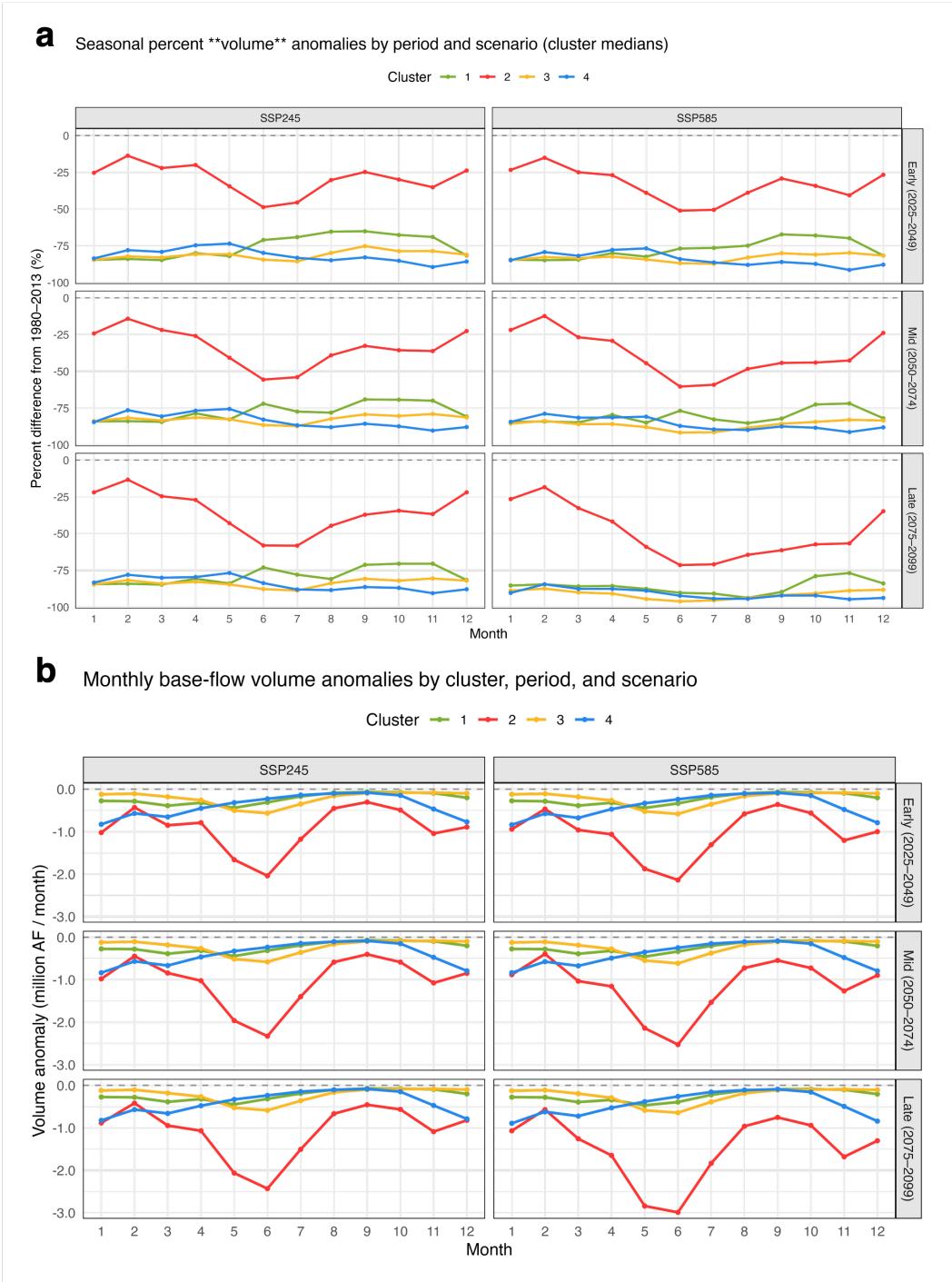


Figure 7: Monthly base-flow anomalies by hydroclimatic cluster, period, and emissions scenario. (a) Percent anomalies show the relative change in base-flow volume compared to the 1980–2013 baseline, summarized by cluster medians. Snow- and monsoon-influenced clusters (1 and 4) exhibit the largest relative 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.

3.3.3 Relative and Absolute Cluster Change

Cluster-level projections reveal clear heterogeneity in base-flow responses between hydroclimatic clusters (Figure 7). 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 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 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.

3.3.4 Changes in Base-flow Timing

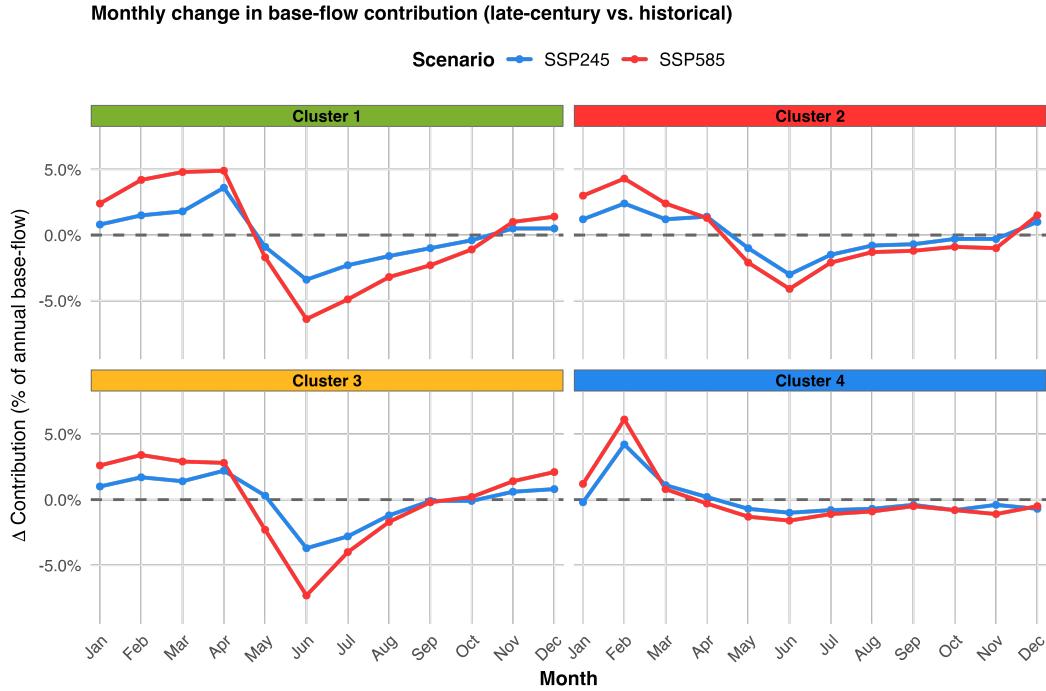


Figure 8: Monthly change in base-flow contribution by hydroclimatic cluster for end-of-century projections relative to the 1980–2013 baseline. Colored lines show SSP2-4.5 (blue) and SSP5-8.5 (red). Clusters 1–3 exhibit modest winter–spring increases and strong summer declines (up to ~5 %), while Cluster 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). 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.

4 Discussion

4.1 Overview of main findings

The results of this study reveal consistent regional patterns of declining base flow across western U.S. headwaters, with the strongest reductions occurring in snow- and mixed-regime basins and more moderate declines in arid and transitional systems. These trends reflect both a long-term downward trajectory in annual base flow and a shift toward earlier seasonal contributions, which is consistent with observed warming and snowpack loss throughout the region (CITE). Relative base-flow reductions in Cluster 2 align with previous findings from the Pacific Northwest, where declining low-flow quantiles have been attributed to temperature-driven decreases in snowpack and late-season recharge (Luce & Holden, 2009). Across clusters, base flow is most sensitive to antecedent moisture and snow contributions, echoing nationwide patterns identified by Ayers et al. (2022). Looking forward, LSTM projections indicate continued annual declines culminating in 45–65% reductions by late century under both SSP2-4.5 and SSP5-8.5 scenarios. Together, these findings provide a regionally consistent, climate-based synthesis of base-flow behavior across western U.S. headwaters, building on recent calls to integrate diverse headwater systems into coherent regional frameworks (Beck et al., 2013; Golden et al., 2025).

4.2 Historical base-flow changes and climate drivers

Observed historical declines in base flow across western U.S. headwaters reflect a combination of warming-driven snowpack loss, earlier snowmelt, and increasing evaporative demand. Our results show reductions in late-summer and early-autumn base flow (Figure 4), consistent with widespread evidence of hydrologic timing shifts over the past half century. These patterns are consistent with observed and projected trends toward warmer winters, rain-on-snow events, and earlier peak flows in snow-dominated basins (Musselman et al., 2018). Stewart et al. (2005) documented regional advances of spring snowmelt and streamflow by 1–4 weeks across western North America, driven primarily by rising winter and spring temperatures. Similarly, Mote et al. (2018) found that over 90% of long-term snow monitoring sites in the western U.S. show declining snow water equivalent, with the greatest losses occurring in transitional snow regimes typical of the PNW. These findings align with the patterns we observe in snow-dominated and mixed-regime clusters (Clusters 1–3), where reduced snowpack storage and earlier melt have translated into diminished groundwater recharge and shortened base-flow duration.

The link between reduced snow storage and declining summer flow has been widely recognized in mountain systems (Safeeq et al., 2013; Tague & Grant, 2009). In our analysis, the strongest base-flow declines occurred in clusters with large snowmelt contributions and steep topographic gradients (Table 5), regions where snow loss directly limits recharge. Tague & Grant (2009) found that in the Cascades range in Oregon, summer streamflow declines were four times greater than in fast-draining basins, emphasizing how subsurface storage capacity controls the sensitivity of base flow to warming. Our results echo this relationship: catchments with greater storage or higher antecedent moisture exhibit smaller declines, suggesting that storage buffers can offset some climatic drying.

Antecedent moisture emerged as the dominant positive predictor of base flow in our mixed-effects models, highlighting the importance of soil–groundwater storage in sustaining streamflow between storm events. Similar results were reported by Ayers et al. (2021) and Ayers et al. (2022), which identified antecedent moisture as the most consistent driver of base-flow variability across the conterminous U.S. These

470 findings emphasize that even in warming climates, multiseasonal storage integration
 471 can maintain low flows when recharge is insufficient. In some regions, this effect
 472 may counterbalance drying trends. For example, Douglas et al. (2000) suggests that
 473 increased rainfall, in place of snow, can enhance shallow groundwater storage, poten-
 474 tially raising low flows even as high flows decline. This mechanism may explain the
 475 modest increases we observed in winter base flow in some transitional basins.

476 Conversely, the combined effects of warming and rising evapotranspiration (ET) are
 477 likely accelerating groundwater depletion and constraining recharge, particularly
 478 in non-snowmelt-dominated systems. Condon et al. (2020) demonstrated that in-
 479 creased ET under warming leads to net groundwater loss across the contiguous U.S.,
 480 even in regions with stable precipitation. In our study, these dynamics are evident in
 481 Cluster 4, where elevated temperatures coincide with persistently low base-flow frac-
 482 tions. These responses emphasize the role of water balance partitioning to determine
 483 the direction, magnitude, and leading cause of base-flow change across regimes.

484 In all, these results illustrate the regime-dependent nature of base-flow responses
 485 to climate forcing. Snow-dominated systems show the largest relative declines, con-
 486 sistent with reduced snowpack and earlier melt (Mote et al., 2018; Stewart et al.,
 487 2005). Mixed regimes exhibit middling declines, reflecting both snow loss and mod-
 488 est recharge from rainfall. Non-snowmelt-dominated basins remain most limited by
 489 storage and precipitation frequency, with warming-driven ET likely outweighing the
 490 potential of rainfall gains. These findings align with the broader understanding of
 491 hydrologic partitioning in the critical zone where catchments with greater subsur-
 492 face permeability and snow-derived recharge maintain higher base-flow indices and
 493 reduced sensitivity to seasonal drying (Wlostowski et al., 2021). Our regional syn-
 494 thesis thus reinforces that the sensitivity of base flow to climate is mediated by both
 495 the magnitude and timing of water inputs and by the storage characteristics that
 496 regulate basin streamflow.

497 **4.3 Spatial heterogeneity and cluster-based insights**

498 We developed a set of base-flow-specific clusters following the principles outlined by
 499 Olden et al. (2012), classifying basins based on data-driven hydrologic and climate
 500 attributes tailored to groundwater-surface water interactions rather than pre-defined
 501 regions. Regional hydroclimatic clustering provides a powerful framework for in-
 502 terpreting base-flow variability across diverse headwater systems. This method
 503 minimizes within-group variability (including differences in elevation, snowmelt tim-
 504 ing, and precipitation regime) while maximizing between-group contrast. Similar
 505 to the hydro-region framework of Dethier et al. (2020), our cluster-based method
 506 allows for the identification of regionally consistent signals in base-flow behavior and
 507 climate response. This methodology supports the idea that hydrologic similarity can
 508 be identified empirically through shared climatic and physiographic attributes, as
 509 emphasized by Wagener et al. (2007) and Beck et al. (2020).

510 Differences among clusters reveal distinct climate–base-flow relationships driven by
 511 regional water balance controls. In snow-dominated systems (Cluster 1, 3), snow pre-
 512 cipitation and melt timing govern interannual variability, while in transitional and
 513 mixed regimes (Cluster 2), the balance between rainfall and antecedent moisture de-
 514 termines the persistence of base flow. Arid clusters (Cluster 4) are most sensitive to
 515 temperature-driven evapotranspiration and storage limitation, showing weak climatic
 516 elasticity and minimal recovery following droughts. These contrasting responses
 517 mirror the spatial structure of recharge and storage capacity across the western U.S.
 518 and show how hydrologic regimes operate along a spectrum from energy- to water-
 519 limited systems. Importantly, the clustering approach offers a scalable framework
 520 for regional hydrologic modeling, enabling data-driven models, such as LSTMs, to
 521 learn regime-specific dynamics more effectively [Kratzert et al. (2019); Lees et al.
 522 (2022)]. This framework strengthens the integrative understanding of headwater be-

523 behavior across space and time suggested by Golden et al. (2025) and provides a path
 524 towards multi-regime hydrologic synthesis.

525 **4.4 Future baseflow trajectories and implications for water resources**

526 Our projections indicate substantial, continued base-flow declines of roughly 45–65%
 527 by late century depending on SSP2-4.5 and SSP5-8.5 scenarios, along with earlier
 528 seasonal peaks and shorter summer flow duration. The greatest relative declines oc-
 529 cur in snow-dominated basins and monsoon-influenced basins (Clusters 3, 4), while
 530 mixed regimes experience the largest total losses (Cluster 2). These trends align
 531 with broader projections of warming, drying, and increased drought frequency across
 532 the western United States (Diffenbaugh et al., 2015).

533 Earlier base-flow timing and reduced late-season flow may have important conse-
 534 quences for aquatic ecosystems and downstream water supply reliability. For water
 535 management, these shifts suggest that existing reservoir operations and drought
 536 plans, which are often based on historical flow timing (CITE), may become less
 537 effective. Increasingly variable and intermittent flow regimes could make it more
 538 difficult to meet environmental flow needs and sustain base-flow-dependent systems
 539 (Hammond et al., 2021).

540 Direct comparisons between SSP2-4.5 and SSP5-8.5 emphasize how emissions path-
 541 ways amplify base-flow declines. Although anomalies under both scenarios overlap
 542 through mid-century, divergence becomes more evident after ~2060. By late-century,
 543 basins under SSP5-8.5 are projected to experience 10–20% larger declines relative to
 544 SSP2-4.5 (Figure 6), with the gap particularly pronounced in snowmelt-dominated
 545 regions (Figure 7). These contrasts underscore the long-term importance of emis-
 546 sions mitigation in reducing the magnitude of base-flow losses across western head-
 547 waters.

548 While some uncertainty remains due to model structure, climate forcing, and the
 549 nonstationary nature of hydrologic systems (Milly et al., 2008), the direction of
 550 change is consistent across scenarios and clusters. The integration of statistical trend
 551 analysis with cluster-specific LSTM models captures both linear sensitivities and
 552 nonlinear climate–hydrology interactions (Kratzert et al., 2019), providing a more
 553 complete picture of how groundwater-supported flow may evolve under future warm-
 554 ing. Continued efforts to couple machine learning with process-based understanding
 555 and to expand headwater monitoring networks will be essential for improving projec-
 556 tions and supporting climate-resilient water management.

557 **4.5 Methodological advances and integration of statistical + machine 558 learning approaches**

559 By combining statistical and machine learning approaches, this study bridges histori-
 560 cal interpretation with predictive modeling to better understand base-flow dynamics
 561 across diverse hydroclimatic regimes. Statistical models provide interpretable esti-
 562 mates of long-term change and quantify the strength of climatic drivers, aligning
 563 with similar approaches in previous studies linking base flow to precipitation, tem-
 564 perature, and snow metrics (Ayers et al., 2022; Murray et al., 2023). However, these
 565 linear models are limited in their ability to capture the nonlinear, lagged feedbacks
 566 that characterize groundwater–surface water interactions. Integrating LSTM neural
 567 networks addressed this gap, allowing the detection of complex, time-dependent re-
 568 lationships between climate inputs and base-flow responses that vary seasonally and
 569 across clusters.

570 Across hydroclimatic clusters, the LSTM models successfully captured both the
 571 magnitude and variability headwater base flow. In particular, the models accurately
 572 reproduced seasonal lags between climate forcings and streamflow response. Model
 573 performance was strongest in snow- and mixed-regime basins where long-term cli-
 574 mate signals are more coherent, while arid clusters exhibited higher uncertainty due

575 to more variation in low-flows and weaker climate–base-flow relationships. These
 576 factors highlight the importance of dataset length, cluster representativeness, and
 577 model transferability when applying deep learning to hydrologic prediction. Even so,
 578 the hybrid framework demonstrated how process-informed statistical analyses and
 579 data-driven models can complement one another with statistical methods identifying
 580 dominant drivers and LSTMs modeling their nonlinear integration. This approach
 581 attempts to answer recent calls for *process-based machine learning* (Nearing et al.,
 582 2021; Shen et al., 2023) and contributes to emerging efforts to combine empirical
 583 studies, physical understanding, and regional synthesis in hydrologic science (Golden
 584 et al., 2025; Kratzert et al., 2019).

585 **4.6 Implications and future directions**

586 Our projections rely on CESM2 from the WUS-D3 dataset, as it is the only GCM
 587 with full coverage for both SSP2-4.5 and SSP5-8.5 in this region-specific dataset.
 588 Using a single model ensures internal consistency between scenarios but narrows
 589 the range of climate uncertainty represented. Future work should incorporate multi-
 590 GCM ensembles to assess the robustness of projected declines and to quantify sce-
 591 nario and model spread.

592 Our results point to heightened vulnerability of groundwater-supported flow under
 593 continued aridification, with the greatest risks during late summer and early autumn
 594 when ecological needs and human demands converge. Protecting and managing
 595 headwaters, where groundwater–surface water exchange sustains downstream supply,
 596 will require strategies that prioritize aquatic refugia, riparian and meadow restora-
 597 tion, and recharge-enhancing practices. This is in line with recent calls to elevate
 598 headwaters in conservation and policy frameworks (Creed et al., 2017; Golden et
 599 al., 2025; McDonnell et al., 2007). Growing non-perennial segments and increased
 600 intermittent flow emphasize the need to integrate headwater systems into regional
 601 water-management planning, including environmental flow targets that explicitly
 602 account for declining base flow and earlier seasonal timing (Datry et al., 2014).

603 Future work should prioritize reducing uncertainty and improving the decision-
 604 making relevance of base-flow projections. Expanding monitoring in data-sparse
 605 regions, especially in non-perennial reaches and headwater catchments, will be criti-
 606 cal for capturing emerging hydrologic change. The hybrid modeling framework used
 607 here, which couples LSTM models (to capture nonlinear, lagged responses) with
 608 process-based or statistical models, can improve model transferability across regimes
 609 and can help bridge knowledge gaps in the near term. To better constrain possible
 610 climate futures, future analyses should draw on multi-model climate ensembles that
 611 capture a broader range of warming and precipitation trajectories. The cluster-based
 612 framework presented in this study provides a scalable way to generalize climate–
 613 streamflow relationships and to translate insights between site-level and regional
 614 scales. Together, these steps can strengthen the connection between scientific knowl-
 615 edge and water management, by putting forth a headwater-focused approach to
 616 climate resilience that recognizes base flow as a key indicator of groundwater-surface
 617 water connectivity, watershed health, and water-supply reliability.

618 **5 Conclusions**

619 This study analyzed base-flow trends and seasonality across western U.S. headwaters
 620 to understand both historical patterns and future changes under different climate
 621 scenarios. We found persistent regional declines in base-flow volume that are pro-
 622 jected to continue and intensify through the 21st century, regardless of emissions
 623 pathway. Along with these volumetric reductions, base-flow contributions are ex-
 624 pected to shift earlier in the year, with earlier spring peaks and lower summer flows.
 625 Using cluster-wise LSTM models, which captured both temporal and spatial het-
 626 erogeneity, we modeled distinct hydroclimatic responses that reflect the diversity of
 627 western headwater systems.

628 The four hydroclimatic clusters highlight meaningful contrasts in both historical
 629 base-flow behavior and climate sensitivity. Clusters 1 and 3 are both snowmelt-
 630 dominated but differ in how temperature affects flow. Cluster 1 shows pronounced
 631 snow suppression and a weakly negative temperature effect, reflecting delayed melt
 632 and storage-dominated runoff. In contrast, Cluster 3 exhibits a positive temperature
 633 effect, consistent with enhanced snowmelt contributions to base flow during trans-
 634 sitional warming periods. Cluster 2 exhibits the smallest projected declines, likely
 635 reflecting the buffering influence of perennial snow and glacier melt in the Cascade
 636 Range. In contrast, Cluster 4 is characterized by low and highly variable base flow,
 637 dominated by short recharge events and strong evaporative losses typical of arid
 638 and monsoon-driven systems. These distinctions highlight how physiographic and
 639 climatic setting shape base-flow resilience and provide a meaningful framework for
 640 interpreting spatial and cluster-wise hydrologic change.

641 Antecedent moisture was consistently the most positive driver of base flow, em-
 642 phasizing that cumulative recharge is key in sustaining groundwater-supported
 643 streamflow. Snow precipitation had strong negative effects, while temperature and
 644 precipitation effects varied by regime with positive effects in snow-responsive basins
 645 and negative effects in arid basins. Future projections indicate continued declines
 646 in base flow across all clusters, with regional reductions of between 45–65% by late
 647 century depending on climate scenario. Although early-century changes are similar
 648 across scenarios, stronger declines and greater variability emerge after mid-century
 649 under higher emissions.

650 Overall, our results show that base flow in western headwaters is highly sensitive to
 651 both temperature-driven hydrologic shifts and long-term wetness conditions. Earlier
 652 runoff timing and reduced summer flows will likely worsen late-season water scarcity,
 653 ecosystem stress, and wildfire risk. By combining statistical and machine-learning
 654 approaches, this work offers a framework for identifying climate-sensitive hydrologic
 655 regimes and improving regional water management under a warming climate.

656 6 Acknowledgments

657 7 Open research

658 References

- 659 Arnold, J. G., Allen, P. M., Muttiah, R., & Bernhardt, G. (1995). Automated base
 660 flow separation and recession analysis techniques. *Ground Water*, 33(6), 1010–
 661 1018. <https://doi.org/10.1111/j.1745-6584.1995.tb00046.x>
- 662 Ayers, J. R., Villarini, G., Schilling, K., & Jones, C. (2021). Projected changes
 663 in monthly baseflow across the u.s.
 664 Midwest. *International Journal of Climatology*, 41(12), 5536–5549.
 665 <https://doi.org/10.1002/joc.7140>
- 666 Ayers, J. R., Villarini, G., Schilling, K., Jones, C., Brookfield, A., Zipper, S. C.,
 667 & Farmer, W. H. (2022). The role of climate in monthly baseflow changes
 668 across the continental united states. *Journal of Hydrologic Engineering*, 27(5),
 669 04022006. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0002170](https://doi.org/10.1061/(ASCE)HE.1943-5584.0002170)
- 670 Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of
 671 a warming climate on water availability in snow-dominated regions. *Nature*,
 672 438(7066), 303–309. <https://doi.org/10.1038/NATURE04141>
- 673 Beck, H. E., Dijk, A. I. J. M. van, Miralles, D. G., Jeu, R. A. M. de, Sampurno Brui-
 674 jnzeel, L. A., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base
 675 flow index and recession based on streamflow observations from 3394 catchments:
 676 Global patterns in base flow characteristics. *Water Resources Research*, 49(12),
 677 7843–7863. <https://doi.org/10.1002/2013WR013918>
- 678 Beck, H. E., Pan, M., Lin, P., Seibert, J., Dijk, A. I. J. M. van, & Wood, E. F.
 679 (2020). Global Fully Distributed Parameter Regionalization Based on Observed
 680 Streamflow From 4,229 Headwater Catchments. *Journal of Geophysical Re-*

- 681 search: *Atmospheres*, 125(17), e2019JD031485. <https://doi.org/10.1029/2019JD031485>
- 682 Brinkerhoff, C. B. (2024). The importance of source data in river network connectivity modeling: A review. *Limnology and Oceanography*, 69(12), 3033–3060. <https://doi.org/10.1002/lno.12706>
- 683 Brown, T. C., Hobbins, M. T., & Ramirez, J. A. (n.d.). *The Source of Water Supply in the United States*.
- 684 Chapman, T. (1999). A comparison of algorithms for stream flow recession and baseflow separation. *Hydrological Processes*, 13(5), 701–714. [https://doi.org/10.1002/\(SICI\)1099-1085\(19990415\)13:5%3C701::AID-HYP774%3E3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-1085(19990415)13:5%3C701::AID-HYP774%3E3.0.CO;2-2)
- 685 Chen, H., & Teegavarapu, R. S. V. (2021). Spatial and temporal variabilities in baseflow characteristics across the continental USA. *Theoretical and Applied Climatology*, 143(3), 1615–1629. <https://doi.org/10.1007/s00704-020-03481-0>
- 686 Collischonn, W., & Fan, F. M. (2013). Defining parameters for eckhardt's digital baseflow filter. *Hydrological Processes*, 27(18), 2614–2622. <https://doi.org/10.1002/hyp.9391>
- 687 Condon, L. E., Atchley, A. L., & Maxwell, R. M. (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>
- 688 Creed, I. F., Lane, C. R., Serran, J. N., Alexander, L. C., Basu, N. B., Calhoun, A. J. K., et al. (2017). Enhancing protection for vulnerable waters. *Nature Geoscience*, 10(11), 809–815. <https://doi.org/10.1038/ngeo3041>
- 689 Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., 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>
- 690 Datry, T., Larned, S. T., & Tockner, K. (2014). Intermittent Rivers: A Challenge for Freshwater Ecology. *BioScience*, 64(3), 229–235. <https://doi.org/10.1093/biosci/bit027>
- 691 Dethier, E. N., Sartain, S. L., Renshaw, C. E., & Magilligan, F. J. (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>
- 692 Diffenbaugh, N. S., Swain, D. L., & Touma, D. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*, 112(13), 3931–3936. <https://doi.org/10.1073/pnas.1422385112>
- 693 Douglas, E. M., Vogel, R. M., & Kroll, C. N. (2000). Trends in floods and low flows in the united states: Impact of spatial correlation. *Journal of Hydrology*, 240(1), 90–105. [https://doi.org/10.1016/S0022-1694\(00\)00336-X](https://doi.org/10.1016/S0022-1694(00)00336-X)
- 694 Eckhardt, K. (2005). How to construct recursive digital filters for baseflow separation. *Hydrological Processes*, 19(2), 507–515. <https://doi.org/10.1002/hyp.5675>
- 695 Eckhardt, K. (2008). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*, 352(1), 168–173. <https://doi.org/10.1016/j.jhydrol.2008.01.005>
- 696 Freeman, M. C., Pringle, C. M., & Jackson, C. R. (2007). Hydrologic connectivity and the contribution of stream headwaters to ecological integrity at regional scales¹. *JAWRA Journal of the American Water Resources Association*, 43(1), 5–14. <https://doi.org/10.1111/j.1752-1688.2007.00002.x>
- 697 Geological Survey (U.S.). (2025). National hydrography dataset. Retrieved from <https://nhdplus.com/NHDPlus/>
- 698 Golden, H. E., Christensen, J. R., McMillan, H. K., Kelleher, C. A., Lane, C. R., Husic, A., 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., Nonner, J., Heijkers, J., & Uhlenbrook, S. (2009). Comparison of different base flow separation methods in a lowland catchment. *Hydrology and Earth System Sciences*, 13(11), 2055–2068. <https://doi.org/10.5194/hess-13-2055-2009>
- Gorelick, S. M., & Zheng, C. (2015). Global change and the groundwater management challenge. *Water Resources Research*, 51(5), 3031–3051. <https://doi.org/10.1002/2014WR016825>
- Hamed, K. H., & Ramachandra Rao, A. (1998). A modified mann-kendall trend test for autocorrelated data. *Journal of Hydrology*, 204(1), 182–196. [https://doi.org/10.1016/S0022-1694\(97\)00125-X](https://doi.org/10.1016/S0022-1694(97)00125-X)
- Hammond, J. C., Zimmer, M., Shanafield, M., Kaiser, K., Godsey, S. E., Mims, M. C., 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>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178–210. <https://doi.org/10.1016/j.ins.2022.11.139>
- Imberger, M., Hatt, B. E., Brown, S., Burns, M. J., Burrows, R. M., & Walsh, C. J. (2023). Headwater streams in an urbanizing world. *Freshwater Science*, 42(3), 323–336. <https://doi.org/10.1086/726682>
- Institute of Hydrology. (1980). *Low flow studies report 3 catchment characteristic estimation manual*.
- Kratzert, F., Herrnegger, M., Klotz, D., Hochreiter, S., & Klambauer, G. (2019). NeuralHydrology – interpreting LSTMs in hydrology. In W. Samek, G. Mon-tavon, A. Vedaldi, L. K. Hansen, & K.-R. Müller (Eds.), *Explainable AI: Interpreting, explaining and visualizing deep learning* (pp. 347–362). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-28954-6_19
- Lees, T., Reece, S., Kratzert, F., Klotz, D., Gauch, M., De Bruijn, J., et al. (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>
- Luce, C. H., & Holden, Z. A. (2009). Declining annual streamflow distributions in the Pacific Northwest United States, 1948–2006. *Geophysical Research Letters*, 36(16), 2009GL039407. <https://doi.org/10.1029/2009GL039407>
- McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., 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., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, 319(5863), 573–574. <https://doi.org/10.1126/science.1151915>
- Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. (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>
- Murray, J., Ayers, J., & Brookfield, A. (2023). The impact of climate change on monthly baseflow trends across canada. *Journal of Hydrology*, 618, 129254. <https://doi.org/10.1016/j.jhydrol.2023.129254>
- Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., et al. (2018). Projected increases and shifts in rain-on-snow flood risk over western north america. *Nature Climate Change*, 8(9), 808–812. <https://doi.org/10.1038/s41558-018-0236-4>

- 791 Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for
 792 base flow and recession analyses. *Water Resources Research*, 26(7), 1465–1473.
 793 <https://doi.org/10.1029/WR026i007p01465>
- 794 Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame,
 795 J. M., et al. (2021). What Role Does Hydrological Science Play in the Age
 796 of Machine Learning? *Water Resources Research*, 57(3), e2020WR028091.
 797 <https://doi.org/10.1029/2020WR028091>
- 798 Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Lorenzo, E.
 799 D., et al. (2016). The pacific decadal oscillation, revisited. <https://doi.org/10.1175/JCLI-D-15-0508.1>
- 800 Olden, J. D., Kennard, M. J., & Pusey, B. J. (2012). A framework for hydrologic
 801 classification with a review of methodologies and applications in ecohydrology.
 802 *Ecohydrology*, 5(4), 503–518. <https://doi.org/10.1002/eco.251>
- 803 Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the
 804 köppen-geiger climate classification. *Hydrology and Earth System Sciences*, 11(5),
 805 1633–1644. <https://doi.org/10.5194/hess-11-1633-2007>
- 806 Pelto, M. S. (2008). Impact of climate change on north cascade alpine glaciers, and
 807 alpine runoff. *Northwest Science*, 82(1), 65–75. <https://doi.org/10.3955/0029-344X-82.1.65>
- 808 Rahimi, S., Huang, L., Norris, J., Hall, A., Goldenson, N., Krantz, W., et al. (2024).
 809 An overview of the western united states dynamically downscaled dataset (WUS-
 810 D3). *Geoscientific Model Development*, 17(6), 2265–2286. <https://doi.org/10.5194/gmd-17-2265-2024>
- 811 Razavi, S. (2021). Deep learning, explained: Fundamentals, explainability, and
 812 bridgeability to process-based modelling. *Environmental Modelling & Software*,
 813 144, 105159. <https://doi.org/10.1016/j.envsoft.2021.105159>
- 814 Rice, J. S., Emanuel, R. E., Vose, J. M., & Nelson, S. A. C. (2015). Continental
 815 u.s. Streamflow trends from 1940 to 2009 and their relationships with wa-
 816 tershed spatial characteristics. *Water Resources Research*, 51(8), 6262–6275.
 817 <https://doi.org/10.1002/2014WR016367>
- 818 Safeeq, M., Grant, G. E., Lewis, S. L., & Tague, Christina. L. (2013). Coupling
 819 snowpack and groundwater dynamics to interpret historical streamflow trends
 820 in the western United States. *Hydrological Processes*, 27(5), 655–668. <https://doi.org/10.1002/hyp.9628>
- 821 Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., & Tuppad, P. (2008). Re-
 822 gional estimation of base flow for the conterminous united states by hydrologic
 823 landscape regions. *Journal of Hydrology*, 351(1), 139–153. <https://doi.org/10.1016/j.jhydrol.2007.12.018>
- 824 Shen, C., Appling, A. P., Gentine, P., Bandai, T., Gupta, H., Tartakovsky, A.,
 825 et al. (2023). Differentiable modelling to unify machine learning and physical
 826 models for geosciences. *Nature Reviews Earth & Environment*, 4(8), 552–567.
 827 <https://doi.org/10.1038/s43017-023-00450-9>
- 828 Sloto, R., & Crouse, M. (1996). *HYSEP: A computer program for streamflow hy-
 829 drograph separation and analysis* (No. 96-4040). USGS. <https://doi.org/10.3133/wri964040>
- 830 Smakhtin, V. U. (2001). Low flow hydrology: a review. *Journal of Hydrology*, 240(3-
 831 4), 147–186. [https://doi.org/10.1016/S0022-1694\(00\)00340-1](https://doi.org/10.1016/S0022-1694(00)00340-1)
- 832 Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R.
 833 (2014). Dropout: A simple way to prevent neural networks from overfitting.
 834 *Journal of Machine Learning Research*, 15(56), 1929–1958. Retrieved from
 835 <http://jmlr.org/papers/v15/srivastava14a.html>
- 836 Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward Earlier
 837 Streamflow Timing across Western North America. <https://doi.org/10.1175/JCLI3321.1>

- 845 Tague, C., & Grant, G. E. (2009). Groundwater dynamics mediate low-flow response
846 to global warming in snow-dominated alpine regions: GROUNDWATER DY-
847 NAMICS AND LOW-FLOW RESPONSE. *Water Resources Research*, 45(7).
848 <https://doi.org/10.1029/2008WR007179>
- 849 Tan, X., Liu, B., & Tan, X. (2020). Global changes in baseflow under the im-
850 pacts of changing climate and vegetation. *Water Resources Research*, 56(9),
851 e2020WR027349. <https://doi.org/10.1029/2020WR027349>
- 852 Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., et al.
853 (2013). Ground water and climate change. *Nature Climate Change*, 3(4), 322–
854 329. <https://doi.org/10.1038/nclimate1744>
- 855 Thomas C Winter, Judson W Harvey, O Lehn Franke, & William M Alley. (1998).
856 *Ground Water and Surface Water: A Single Resource*.
- 857 Tillman, F. D., Day, N. K., Miller, M. P., Miller, O. L., Rumsey, C. A., Wise, D. R.,
858 et al. (2022). A review of current capabilities and science gaps in water supply
859 data, modeling, and trends for water availability assessments in the upper col-
860 orado river basin. *Water*, 14(23), 3813. <https://doi.org/10.3390/w14233813>
- 861 U. S. Geological Survey. (2018). Water Science Glossary. Retrieved from <https://www.usgs.gov/water-science-school/science/water-science-glossary>
- 862 US Army Corps of Engineers. (n.d.). National inventory of dams.
- 863 Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment Classifica-
864 tion and Hydrologic Similarity. *Geography Compass*, 1(4), 901–931. <https://doi.org/10.1111/j.1749-8198.2007.00039.x>
- 865 Wlostowski, A. N., Molotch, N., Anderson, S. P., Brantley, S. L., Chorover, J.,
866 Dralle, D., et al. (2021). Signatures of Hydrologic Function Across the Critical
867 Zone Observatory Network. *Water Resources Research*, 57(3), e2019WR026635.
868 <https://doi.org/10.1029/2019WR026635>
- 869 Woodhouse, C. A., & Udall, B. (2022). Upper gila, salt, and verde rivers: Arid land
870 rivers in a changing climate. *Earth Interactions*, 26(1), 1–14. <https://doi.org/10.1175/EI-D-21-0014.1>
- 871 Xie, J., Liu, X., Wang, K., Yang, T., Liang, K., & Liu, C. (2020). Evaluation of typ-
872 ical methods for baseflow separation in the contiguous united states. *Journal of*
873 *Hydrology*, 583, 124628. <https://doi.org/10.1016/j.jhydrol.2020.124628>