Supplementary Information for

Changes in Streamflow Statistical Structure across the United States due to Recent Climate Change

Abhinav Gupta[*1], Rosemary W. H. Carroll[2], Sean A. McKenna[2]

[1]Division of Hydrologic Sciences, Desert Research Institute, 755 E. Flamingo Rd., Las Vegas, NV, 89169, United States of America

[2]Division of Hydrologic Sciences, Desert Research Institute, 2215 Raggio Pkwy, Reno, NV 89512, United States of America

Contents of this file

Text S1. Parameter estimation of FARIMA models

Text S2. FARIMA model validation

Text S3. Trends in streamflow temperature relationship

Text S4. Trend in rainfall runoff relationship

Text S5. Evidence of robustness of changes in streamflow statistical structure (SSS)

Figure S1. Deseasonalization illustration. (a) Streamflow time-series, (b) seasonal component obtained using LOWESS and simple averaging, and (c) deseasonalized streamflow using seasonal component obtained by LOWESS.

Figure S2. *1st time window*. The autocorrelation of fitted residuals at each time-lag for the models created in first time-window. For each time-lag, the autocorrelation is shown for all the watersheds as boxplot. This figure shows that for all the watersheds except a few, the autocorrelation at every lag was between -0.10 and 0.10. In fact, for most watersheds it was between -0.05 and 0.05.

Figure S3. *2nd time window*. Same as Figure S2 but for time-window 2.

Figure S4. *3rd time window*. Same as Figure S2 but for time-window 3.

Figure S5. *4th time window*. Same as Figure S2 but for time-window 4.

Figure S6. *5th time window*. Same as Figure S2 but for time-window 5.

Figure S7. *6th time window*. Same as Figure S2 but for time-window 6.

Figure S8. *7th time window*. Same as Figure S2 but for time-window 7.

Figure S9. *8th time window*. Same as Figure S2 but for time-window 8.

Figure S10. *9th time window*. Same as Figure S2 but for time-window 9.

Figure S11. Snow-dominated watersheds. Trend in snow signatures (a) rising limb slope (δ_{snow}), (b) rising limb intercept (β_{snow}), and (c) time to peak, and (d) trend in snow water equivalent

(SWE). The '+' sign indicates that the trend in statistically significant at 5% significance level. The number of watersheds in subplots (a), (b), and (c) are different from that in subplot (d) because some of the snow dominated watersheds did not follow the linear relationship.

Figure S12. Rain-dominated watersheds. Trend in mean parameter values of rainfall-runoff models: (a) λ, (b) CN, (c) α/β, and (d) $\sqrt{\alpha/\beta^2}$. The '+' sign indicates that the trend is statistically significant at 5% significance level. The units in subplot (c) and (d) are minutes time-window^{-1}.

Figure S13. Boxplots of the contribution of less than 2-weeks timescale components (F_4) for each time-window. These plots correspond to the watersheds where the first and second significance concluded that F_4 values changed statistically significantly while Mann-Kendall concluded that changed is statistically insignificant. Gauge number with each subplot refers to the unique ID of the streamflow gauge in CAMELS dataset. The lower and upper whiskers represent $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$, respectively. Here Q_1 and Q_3 denote 25th and 75th percentiles, and IQR denotes inter-quartile range, that is, $Q_3 - Q_1$.

Figure S14. Boxplots of the contribution of less than 2-weeks timescale components (F_4) for each time-window. These plots correspond to the watersheds where all three statistical agreed that the F_4 values changed statistically significantly. Gauge number with each subplot refers to the unique ID of the streamflow gauge in CAMELS dataset. The lower and upper whiskers represent $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$, respectively. Here Q_1 and Q_3 denote 25th and 75th percentiles, and IQR denotes inter-quartile range, that is, $Q_3 - Q_1$.

Figure S15. Boxplots of the contribution of less than 2-weeks timescale components (F_4) for each time-window. These plots correspond to the watersheds where first and second significance tests disagreed that the F_4 values changed statistically significantly. Gauge number with each subplot refers to the unique ID of the streamflow gauge in CAMELS dataset. The lower and upper whiskers represent $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$, respectively. Here Q_1 and Q_3 denote 25th and 75th percentiles, and IQR denotes inter-quartile range, that is, $Q_3 - Q_1$.

S1. Parameter estimation of FARIMA models

The calibration of FARIMA model parameter involved several steps. The procedure followed in this paper was similar to that described in Montanari et al. (1997). First, the streamflow time-series were deseasonalized by subtracting the corresponding seasonal components. Seasonal components were computed using two methods: the LOWESS method (Cleveland, 1979) and simple averaging. In the simple averaging method, the seasonal component was computed as follows:

$$X_{\text{sea},k} = \frac{1}{n}\sum_{i=1}^{n} X_{k,n}, \tag{S1}$$

where $X_{k,n}$ is the streamflow value at k^{th} day of the nth year and $X_{\text{sea},k}$ is average of streamflow values on k^{th} day computed over n years. The time-series $X_{\text{sea},k}, k = 1,2,...,365$ represents the seasonal component. Figure S1 shows an example of deseasonalization. The LOWESS and simple averaging methods yield similar values (Figure 4b); therefore, the LOWESS seasonal component method was used for deseasonalization. The LOWESS algorithm had two parameters to be specified by the user: number of iterations and fraction of data used for regression at each time-

step. In this study, 2 iterations were used with fraction of data used at each time-step equal to 0.02. In a few watersheds, the seasonal component obtained by the LOWESS model was unrealistic and was very different from the seasonal component obtained by the simple averaging method. In these watersheds, the seasonal component obtained by simple averaging approach was used for deseasonalization.

The parameter estimation procedure was carried out in two steps. In the first step, number of AR and MA parameters, denoted by p and q, respectively, were determined. In the second step, the value of the AR parameters, MA parameters, and d were determined. In the first step, a preliminary estimate of the d value was obtained as the average of the two values obtained by two methods: R/S method and aggregate variance method (Montanari et al., 1997). For a given d value, fractionally differenced time-series Y_t can be obtained as (using Equation 1 of the main text)

$$Y_t = (1-B)^d X_t = \Phi_p(B)^{-1} \Psi_q(B)\epsilon_t, \tag{S2}$$

where $(1-B)^d$ can be written in the polynomial form as

$$(1-B)^d = \sum_{i=0}^{\infty} a_i B^i,$$
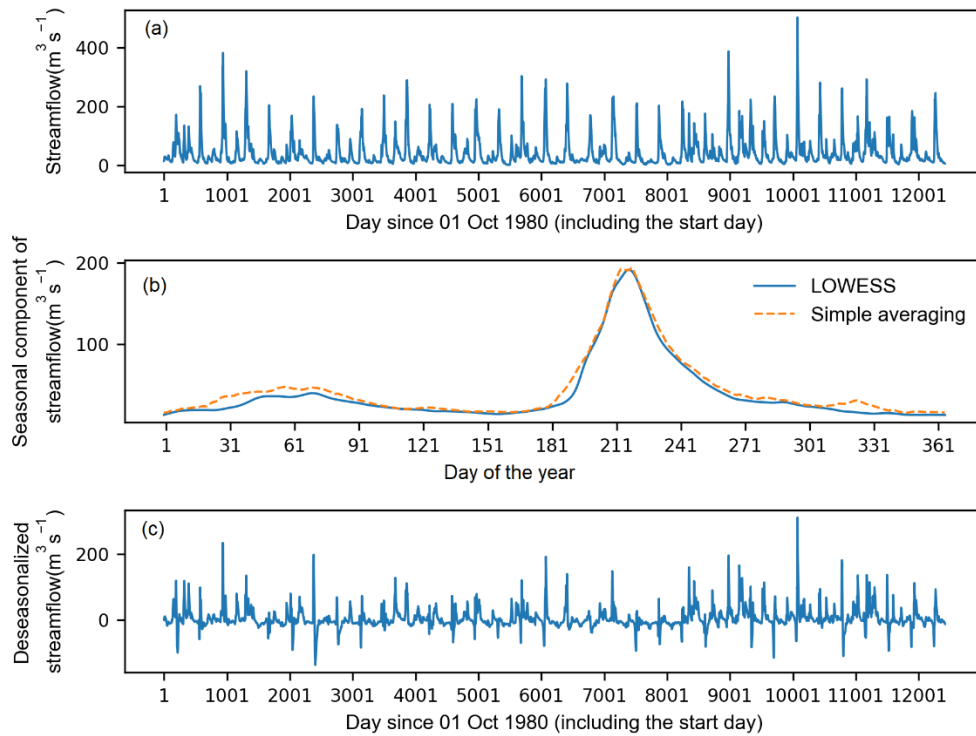$$a_i = \frac{\Gamma(-d+i)}{\Gamma(-d)\Gamma(i+1)}, \tag{S3}$$

Figure S1. Deseasonalization illustration. (a) Streamflow time-series, (b) seasonal component obtained using LOWESS and simple averaging, and (c) deseasonalized streamflow using seasonal component obtained by LOWESS.

where a_i is the coefficient of ith power of backward shift operator B, and it is a function of d and i as defined above. Using (S2),

$$Y_t = \sum_{i=0}^{\infty} a_i X_{t-i}. \tag{S4}$$

Equation (S2) tells us that Y_t is an ARMA process. Eq. (S4) tells us that Y_t can be computed from past and current values of X_t for a given value of d. In this study, Y_t was computed using the past 100 values and the current value of X_t. This is reasonable since the coefficients a_i's drop to zero very quickly and the contribution of the later terms to the summation is negligible. Once the time-series Y_t were obtained, several ARMA models were fitted to it using the python statsmodel package (Seabold & Perktold, 2010) with different values of p and q between 0 and 10. Thus, 121 sets of the parameters were obtained. The p and q corresponding to the set with minimum Aikaike Information Criterion (AIC, Akaike, 1973) were chosen as the optimal values.

In the second step, the values of the parameter were estimated using an iterative-d method. The values of p and q were fixed as obtained in the previous step. The d value was varied between 0 and 0.5 with a step of 0.01. Thus 51 sets of FARIMA parameters corresponding to each d value were obtained for a deseasonalized streamflow time-series and the set with minimum AIC was chosen as the optimum parameter set. Denote the parameter vector of the FARIMA model by $\boldsymbol{\theta}$. The confidence intervals (credible region in Bayesian language) over $\boldsymbol{\theta}$ were computed using the asymptotic Normality of posterior distribution (Berger, 1985) according to which the posterior distribution is approximated as Gaussian with maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$ as mean and covariance matrix $\boldsymbol{\Sigma}$ such that

$$\boldsymbol{\Sigma}^{-1} = -\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}, \tag{S5}$$

where $L(\boldsymbol{\theta})$ is the log-likelihood function of $\boldsymbol{\theta}$. Given the mean and covariance matrix of $\boldsymbol{\theta}$, the confidence intervals can be easily computed. In this study, instead of an exact log-likelihood, Whittle's approximate log-likelihood $L_{\text{W}}(\boldsymbol{\theta})$ was used for analytical convenience (Beran, 1994):

$$L_{\text{W}}(\boldsymbol{\theta}) = \sum_{i=1}^{m} h(\omega_i; \boldsymbol{\theta}) + \frac{I(\omega_i)}{h(\omega_i; \boldsymbol{\theta})},$$

$$m = \begin{cases} \dfrac{N}{2}, & N \text{ even} \\ \dfrac{N+1}{2}, & N \text{ odd.} \end{cases} \tag{S6}$$

where $h(\omega; \boldsymbol{\theta})$ is the same as in Equation (4) (in the main text) with dependency on parameter $\boldsymbol{\theta}$ shown explicitly, $I(\omega)$ is periodogram of deseasonalized streamflow (equivalent to observed PSD), m denotes the total number of discrete frequencies used in the summation (S6), and ω_i denotes the ith discrete frequency such that $\omega_i = i\frac{1}{N}$ day$^{-1}, i = 1, 2, \dots, m$.
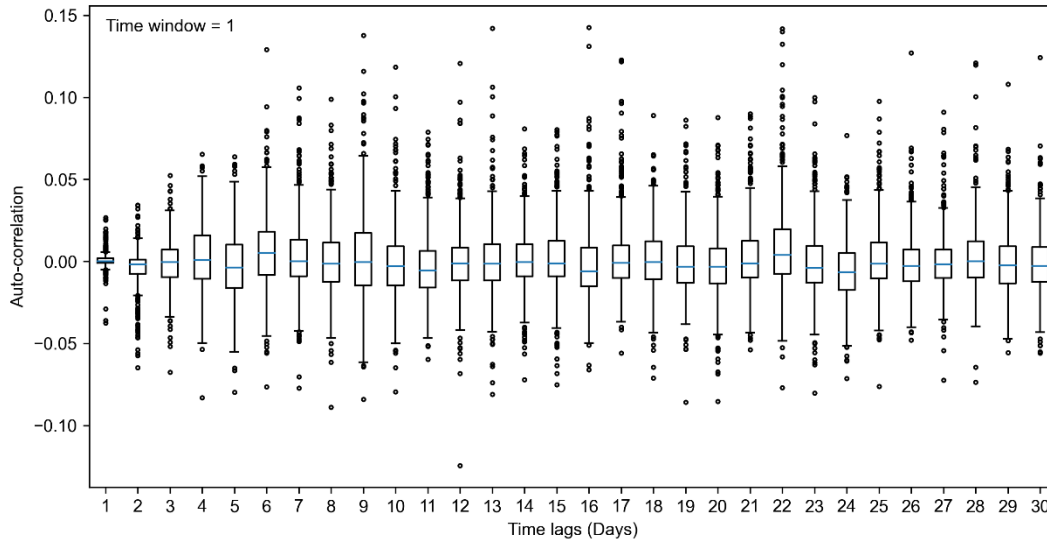
S2. FARIMA model validation

Figure S2. *1st time window*. The autocorrelation of fitted residuals at each time-lag for the models created in first time-window. For each time-lag, the autocorrelation is shown for all the watersheds as boxplot. This figure shows that for all the watersheds except a few, the autocorrelation at every lag was between -0.10 and 0.10. In fact, for most watersheds it was between -0.05 and 0.05.
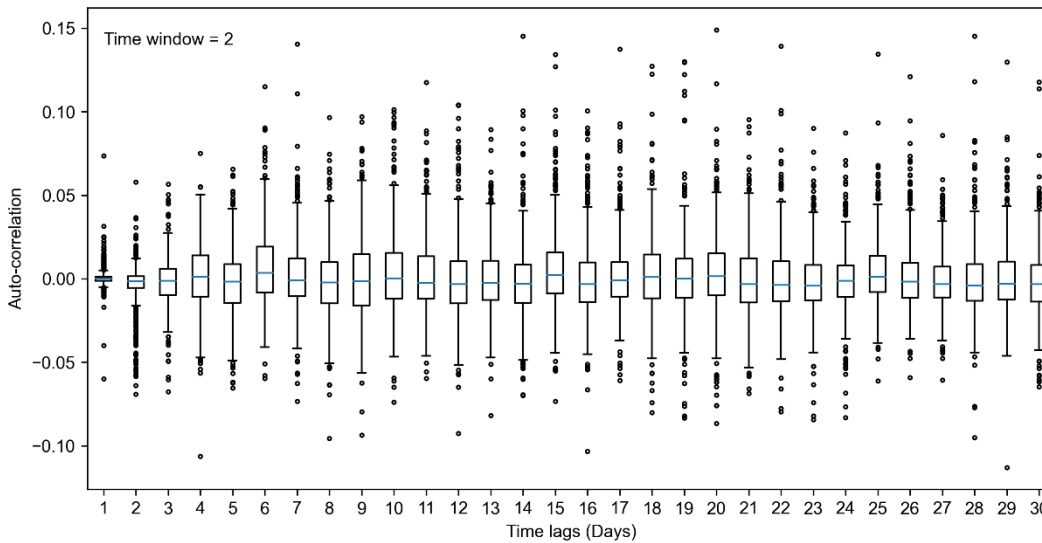

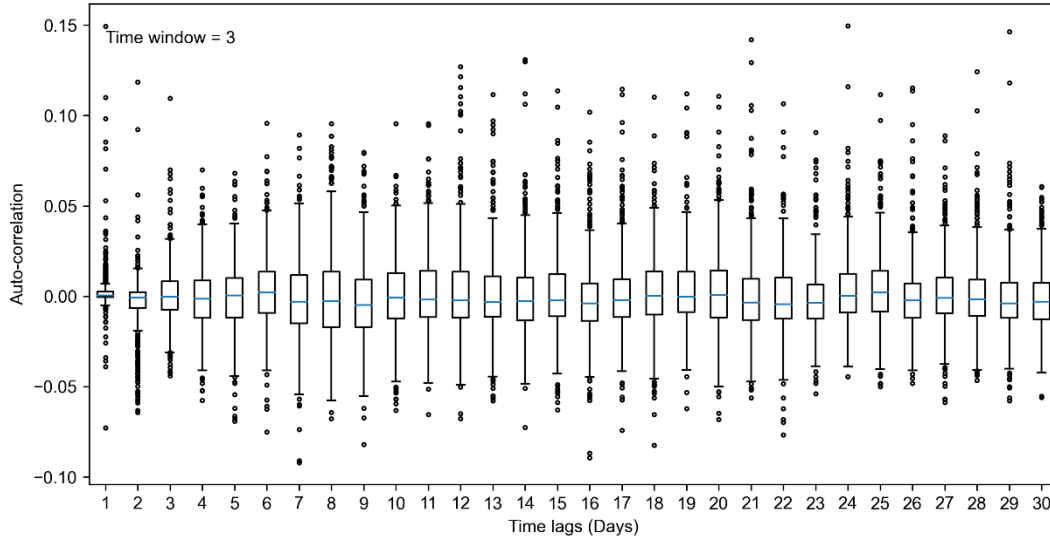Figure S3. *2nd time window*. Same as Figure S2 but for time-window 2.

Figure S4. *3rd time window*. Same as Figure S2 but for time-window 3.

Figure S5. *4th time window*. Same as Figure S2 but for time-window 4.
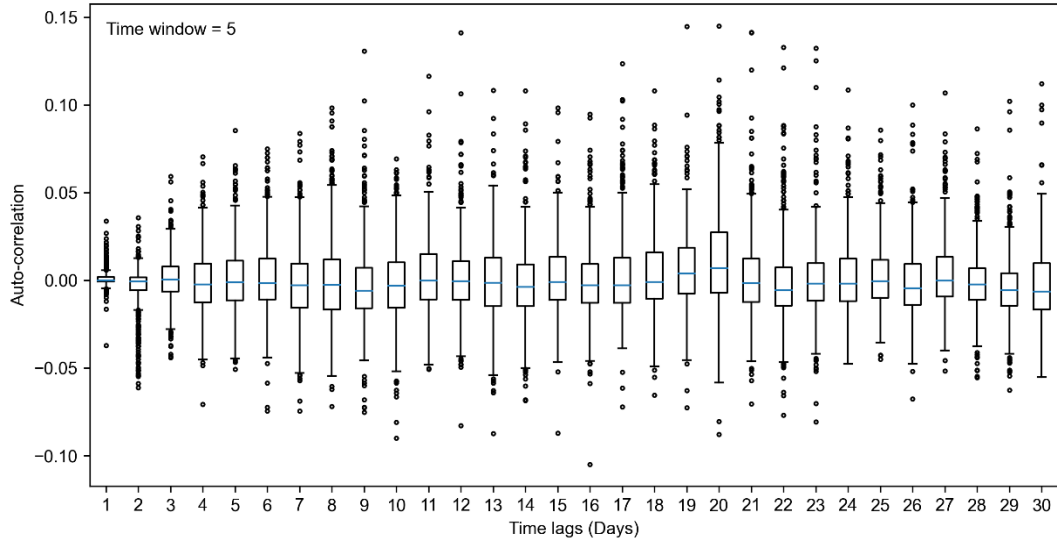
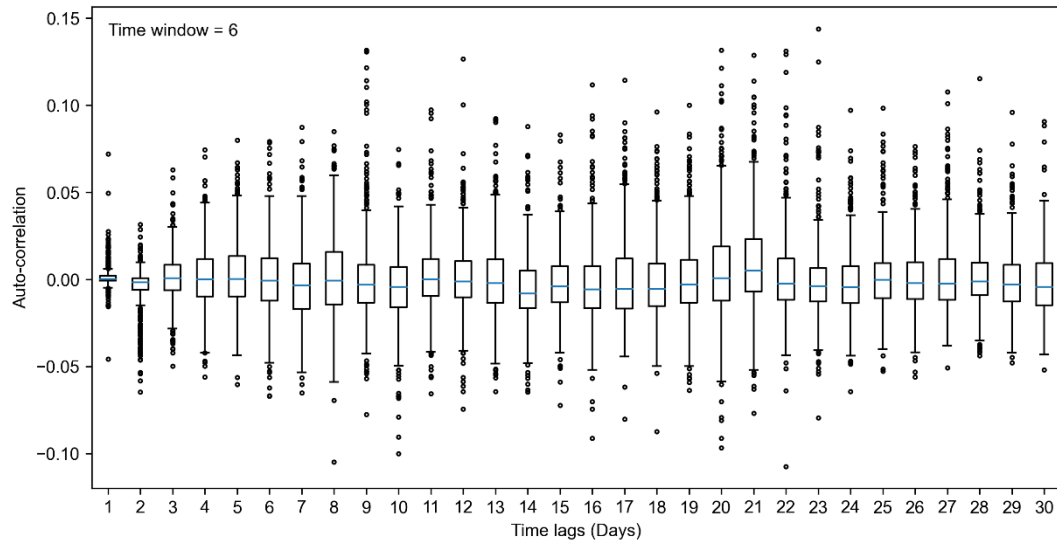Figure S6. *5th time window.* Same as Figure S2 but for time-window 5.

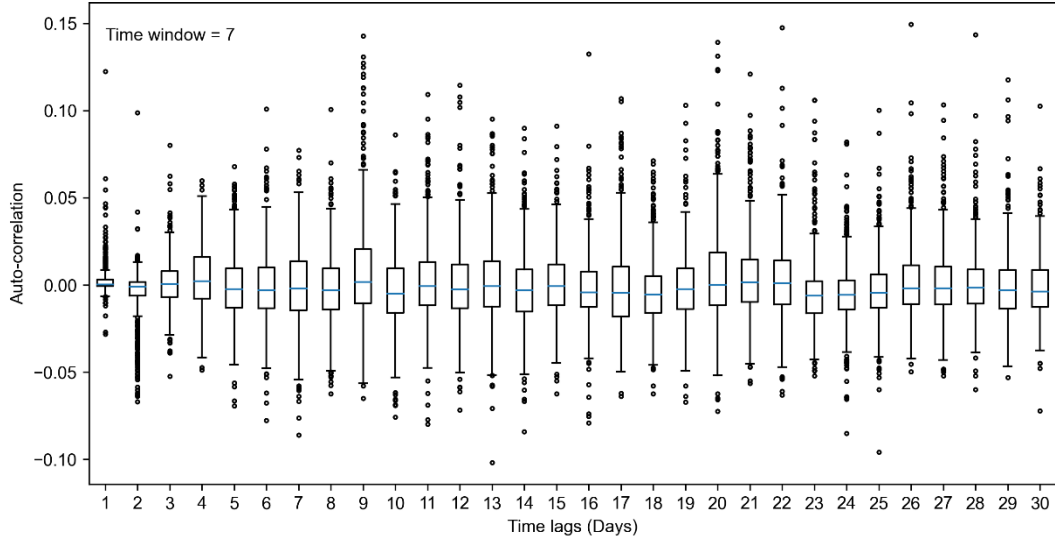Figure S7. *6th time window.* Same as Figure S2 but for time-window 6.

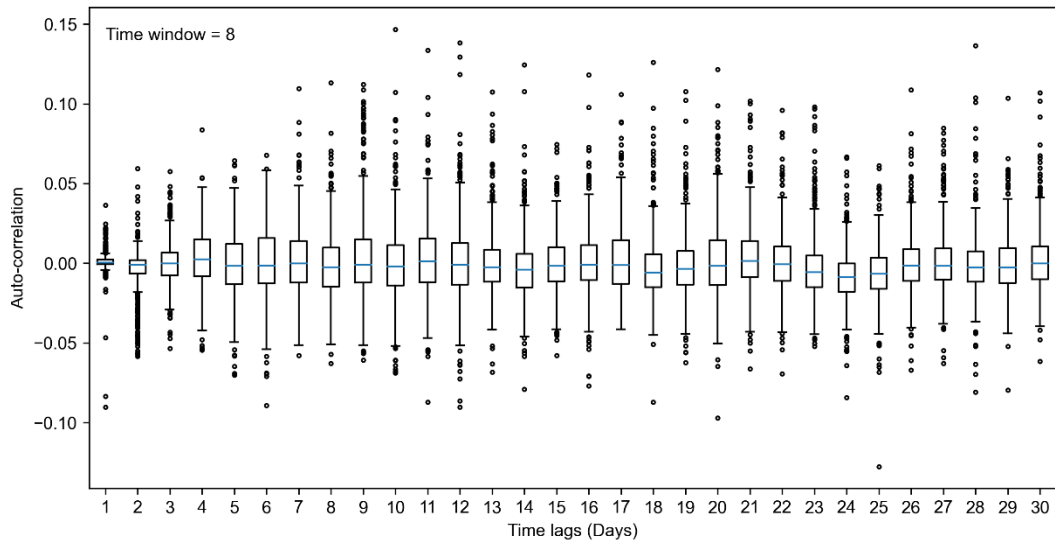Figure S8. *7th time window*. Same as Figure S2 but for time-window 7.

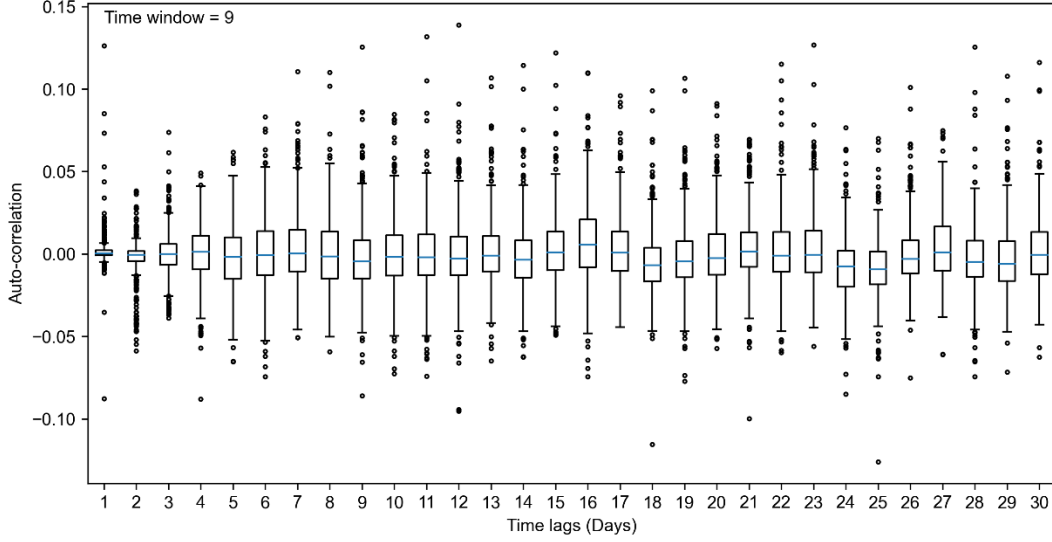Figure S9. *8th time window*. Same as Figure S2 but for time-window 8.

Figure S10. *9th time window*. Same as Figure S2 but for time-window 9.

S3. Trends in streamflow temperature relationship

In this study, snow signatures proposed by Horner et al. (2020) were used to identify the changes in temperature snow relationship. They defined streamflow, temperature, and SWE regimes as a 30-day moving average of their respective seasonal components. Let us denote streamflow, temperature, and SWE regimes by Q_{reg}, T_{reg}, and SWE_{reg}, respectively. Figure 8 (in main text) shows the relationship between temperature and streamflow regimes for a hypothetical snow-dominated watershed. The segment AB is the snowmelt period where both streamflow and temperature rise. Streamflow reaches its peak at point B. After point B, temperature continues to rise but streamflow decreases because of the lack of snow availability. During segment CD, temperature decreases without significant change in streamflow. During the segment DA, snow accumulates. The segments AB and CD capture the snowmelt dynamics. Horner et al. (2020) fitted linear relationships between temperature and streamflow regimes to model segments AB and CD and defined the slopes of these segments as snow signatures. In the study, we found that the linear relation was a good model for the segment AB in all the watersheds, but not for the segment BC in several watersheds (even though the line BC shown in Figure 8 (in main text) is linear). Therefore, we focused only on segment AB which we refer to as the rising limb of temperature-streamflow relationship. Let this relationship be modeled as

$$\hat{Q}_{\text{reg},i} = \delta_{\text{snow}} T_{\text{reg},i} + \beta_{\text{snow}}, \tag{S7}$$

where $T_{\text{reg},i}$ and $\hat{Q}_{\text{reg},i}$ denote the temperature and estimated streamflow regime value on i^{th} day of the water year during the first phase of snowmelt (limb AB), δ_{snow} and β_{snow} denote the slope and intercept of the relationship. We used both δ_{snow} and β_{snow} as the snow signatures.

The slope, δ_{snow}, is a measure of rate of increase of snowmelt per unit increase in temperature. The intercept β_{snow} is the streamflow when the mean temperature is zero and snowmelt has not started. An intuitive way of thinking about β_{snow} is as follows. For a given value of δ_{snow}, the

value of β_{snow} determines the point where line AB intersects with the x-axis ($Q_{reg} = 0$). By making Q_{reg} equal to 0 in Eq. (S7), one gets $T_{reg} = \beta_{snow}/\delta_{snow}$. Thus, given δ_{snow}, the intercept β_{snow} is the *measure* of threshold mean watershed temperature required to start the snowmelt. Keeping the δ_{snow} fixed, higher β_{snow} implies smaller values of threshold temperature and smaller values of β_{snow} implies larger values of threshold temperature. But note that β_{snow} is *not equal to* the threshold temperature required to start the snowmelt. Along with δ_{snow} and β_{snow}, time to peak – number of days since the start of the water year after which streamflow regime peaks – was also computed as a snow signature. We computed the snow signatures for the moving time windows of 10 years each as illustrated in Table 1 (in main text). Subsequently, trends in these signatures were computed over the time-windows. The trend values provide an estimate of change in snow signatures. In the context of this paper, trends in snow signature are related to the change in snowmelt dynamics.

Figure S11 shows the trend in the three snow signatures over snow-dominated CAMELS watersheds. Clustering of positive with positive and negative with negative trend watersheds is remarkable. Which indicates that in snow-dominated watersheds change in climatic statistics is responsible for hydrological changes and the differences in watershed properties has only secondary importance. Magnitude of trends in δ_{snow} and β_{snow} are larger in western watersheds than in eastern watersheds. The magnitude of β_{snow} decreased in most of the western watersheds. The trends in δ_{snow} are positive in most of the north-western watersheds and negative in south-western watersheds. Thus, in western USA mainly two kinds of trends are observed: (1) Increasing value of δ_{snow} and decreasing value of β_{snow}, and (2) decreasing value of δ_{snow} and decreasing value of β_{snow}. Increasing value of δ_{snow} combined with decreasing value of β_{snow}. indicates two possibilities:

(a) The threshold temperature to start snowmelt has increased. This, in turn, is likely due to a decrease in availability of snow at lower elevations. For snowmelt to begin at higher elevations, higher mean watershed temperatures are required. Another implication is that watershed is in a regime where $\partial Q/\partial T$ is higher which is plausible in higher temperatures regime. Increase in mean watershed temperature and lack of snow availability at lower elevations is a likely cause of increase in δ_{snow} and decrease in β_{snow}. We note that for this possibility to realize the increase in mean watershed temperature should be so large that mean temperature at higher elevations is also increased. This is supported by the fact that in north-western watersheds, increases in mean watershed temperature are very large, both daily minimum temperature and winter maximum (Figure 4 in the main text). Moreover, other researchers have noted using snow course observations and VIC model simulations that the effects of the increasing temperatures are more significant at low elevations than at high elevations in western watersheds (Knowles et al., 2006; Mote et al., 2005; Mote, 2006; Belmecheri et al., 2016; Berg & Hall, 2017). This is also supported by the Figure S10d which shows that SWE has decreased in the majority of the western watersheds. Also, note that the trend in β_{snow} is positive in a few Sierra Nevada watersheds located in California - in these watersheds SWE has indeed increased over the study period (Figure S11d).

(b) Another possibility is that the increase in δ_{snow} is so large that the threshold temperature to start snowmelt decreases even if β_{snow} has decreased. Decrease in temperature threshold is possible if more snow is available at lower elevations. However, this possibility seems unlikely to have realized because SWE has decreased over Western USA over the study

period (Figure S11) and snowline has been reported to move to higher elevations as discussed above.
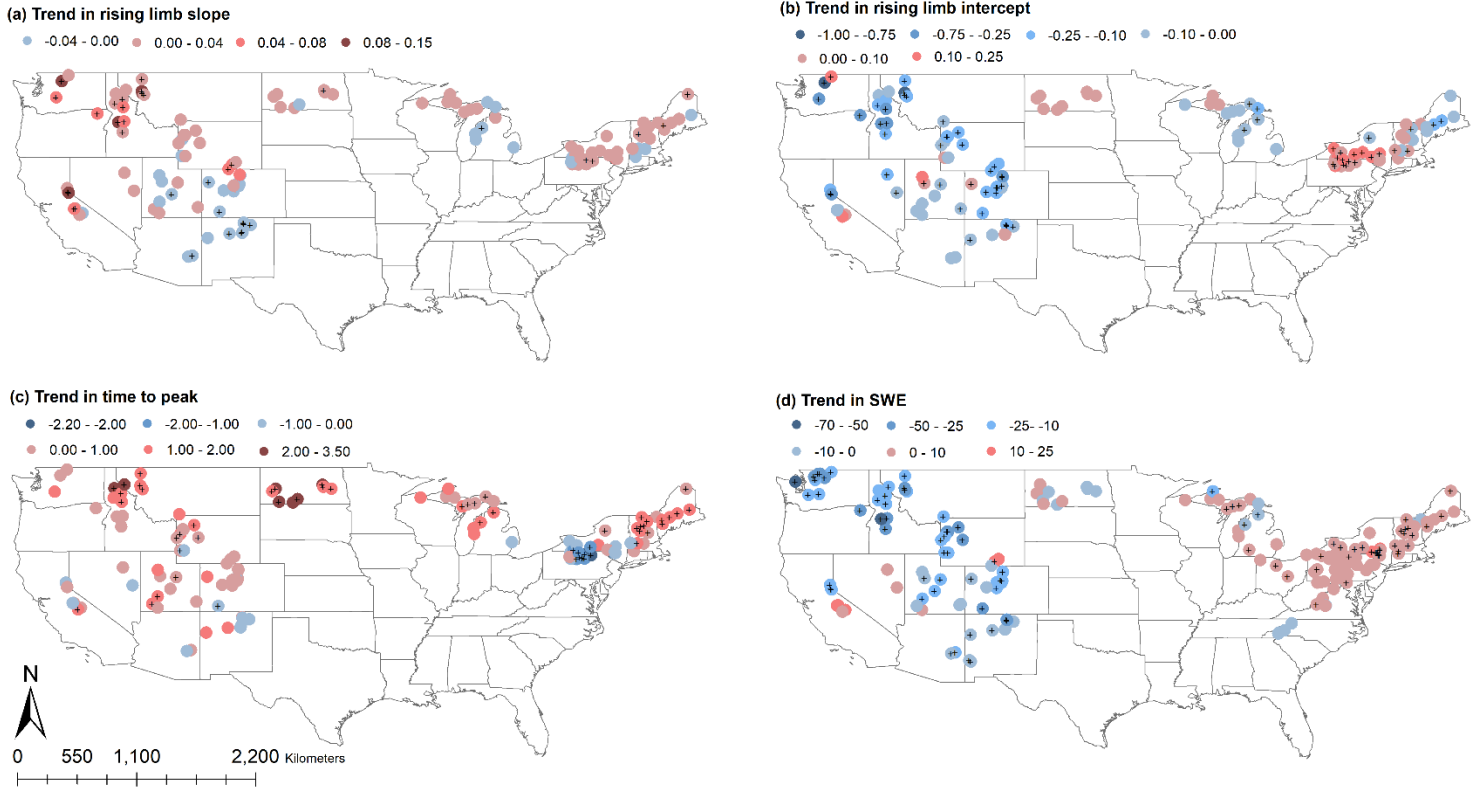

Figure S11. Snow-dominated watersheds. Trend in snow signatures (a) rising limb slope (δ_{snow}), (b) rising limb intercept (β_{snow}), and (c) time to peak, and (d) trend in snow water equivalent (SWE). The '+' sign indicates that the trend in statistically significant at 5% significance level. The number of watersheds in subplots (a), (b), and (c) are different from that in subplot (d) because some of the snow dominated watersheds did not follow the linear relationship.

Decreases in both δ_{snow} and β_{snow} indicate lack of snow availability at lower elevations and no change or decrease in temperatures at higher elevations. In summary, it seems likely that in the western USA (1) snow has moved to higher elevations, and (2) temperatures at higher elevations have increased in northern watersheds and changed negligibly in southern watersheds.

In eastern watersheds also two kinds of trends are observed: (1) Increasing δ_{snow} and increasing β_{snow}, and (2) Increasing δ_{snow} and decreasing β_{snow}. Increases in δ_{snow} and β_{snow} indicate that more snow is available at lower elevations and temperatures have increased. This is supported by the fact that SWE has increased over eastern snow dominated watersheds over the study period (Figure S11d). Increase in δ_{snow} and decrease in β_{snow} indicates that less snow is available at lower elevations and temperatures have increased significantly at higher elevations (as explained above). Since SWE has increased in these watersheds, this seems unlikely. Also note that in none of the eastern watersheds were both an increase in δ_{snow} and a decrease in β_{snow} statistically significant.

The only case not discussed so far is decreasing δ_{snow} and increasing β_{snow}. There exists only one watershed (in western USA) with a statistically significant negative trend in δ_{snow} and statistically

significant positive trend in β_{snow}. In the eastern region, no such watershed exists. This case indicates two possibilities:

(a) The temperature threshold to start snowmelt has decreased but $\partial Q / \partial T$ has decreased. This possibility can be realized either when more snow is available at lower elevations and temperatures have decreased significantly at lower elevations, or when less snow is available at lower elevations and the temperatures have increased significantly at higher elevations. The second possibility is unlikely since increasing temperature are expected to result in lesser snow at lower elevations. The first possibility is unlikely in western watersheds because snowline is known to have moved up in these watersheds.

(b) Temperature threshold to start snowmelt has increased and temperatures have decreased significantly in snow available regions. This possibility can be realized if the snow is only available at higher elevations compared to earlier time-period and temperature have decreased at higher elevations. The decrease in AMJ temperatures gives some evidence, albeit weak, toward this possibility.

Finally, Figure S11d shows that SWE has decreased (increased) in majority of the western (eastern) snow dominated watersheds. In summary, the discussion above indicates that snowmelt dynamics have changed significantly in many of the snow dominated watersheds. Some hypotheses along with preliminary evidence have also been proposed to explain the changes in snowmelt dynamics. In western watersheds, these changes are primary linked to increase in temperature and movement of the snowline to higher elevation.

S4. Trend in rainfall runoff relationship

The method of Lamb & Beven (1997) was used for hydrograph separation and identification of rainfall-runoff events. Baseflow corresponding to each rainfall-runoff event was identified using the method of Collischonn & Fan et al., (2013). The baseflow thus obtained was subtracted from the streamflow time-series to obtain streamflow according to excess rainfall which will be referred to as direct streamflow in what follows.

The event-based model had two components. First component processed rainfall time-series and yielded excess rainfall time-series and the second component routed excess rainfall to watershed outlet. The SCS-CN method was used to convert rainfall time-series to excess rainfall using the equation (Ponce & Hawkins, 1996). SCS-CN method has the advantage that it is parametrically simple with only two parameters, but this simplicity comes at the cost of the sacrificing physical realism. Also, it is argued that this method is applicable only to agricultural watersheds, but many authors have applied this method to watersheds with other land uses and reported good agreement between observed and simulated excess rainfall (e.g., Mishra & Singh, 1999; Geetha et al., 2007; Soulis & Valiantzas, 2012; Soulis & Valiantzas, 2013). Therefore, to keep the methodology manageable we used the SCS-CN method in this study. The SCS-CN model is follows:

$$P_e = \frac{(P - I_a)^2}{P - I_a + S}, \qquad t > 0,$$
$$I_a = \lambda S,$$
$$S = \frac{25400}{CN} - 254, \tag{S8}$$

where I_a denotes initial abstraction, S[in mm] denotes antecedent maximum retention capacity, the parameter λ denotes the fraction of S that is lost as initial abstraction, and the parameter CN determines maximum retention capacity and varied between $(0, 100)$, excluding the extreme values. Typically, the parameter λ is fixed at 0.20 and the parameter CN is determined based on land-use type, but in this study, these were treated as calibration parameter and their values varied from event to event.

The excess rainfall obtained using the SCS-CN method was routed to the watershed outlet using the D-duration unit hydrograph approach (Brutsaert, 2005). The D-duration unit hydrograph was obtained using instantaneous unit hydrograph (IUH). The instantaneous unit hydrograph was assumed to the form of a 2-parameter gamma distribution with parameter α and $\beta[\mathrm{T}^{-1}]$ (Botter et al., 2013):

$$f(t) = \frac{\beta^{\alpha} t^{\alpha-1} \mathrm{e}^{-\beta t}}{\Gamma(\alpha)}, \qquad t > 0, \tag{S9}$$

where t denotes time. The mean of the distribution is α/β, and the variance is α/β^2. The parameter λ, CN, α, and β were estimated for each rainfall-runoff event of a watershed so that the mean square error between simulated and observed direct streamflows was minimized. The optimization was carried out using the Dynamic Dimension Search (DDS) algorithm (Tolson & Shoemaker, 2007). A good fit could be obtained for most of rainfall-runoff events in all the watersheds. But there were a few rainfall-runoff events for which no set of the parameter could yield a good fit. We removed all the rainfall-runoff events with NSE of fit less than 0.75 from the further analyses. Once these parameters are obtained for each of the rainfall-runoff events, then the change in these parameters over time can be used as a measure of the change in the rainfall-runoff response of a watershed. One difficulty is that these parameters vary from event to event in a seemingly random fashion. Therefore, the change in probability distributions of these parameters had to be measured. This was achieved using the moving windows as illustrated in Table 1 (in the main text). All the events contained in a moving window were used to create a probability distribution of the four parameters. The change in probability distribution was measured by estimating the trend in several statistics of the probability distributions which includes mean, mean of 0-10 percentiles, mean of 10-30 percentiles, mean of 30-60 percentiles, mean of 60-90 percentiles, and mean of 90-100 percentiles. The important variables were recognized using the same method as in snow dominated watersheds.

Figure S12 shows the trends in mean values of the parameters of the rainfall-runoff model described above. Spatial clustering discussed above is observed in these plots also. Also, these watersheds have negligible anthropogenic influences. Thus, it appears that the change in these parameters is caused by change in climatic statistics. Trends in mean λ seem to be spatially random with very small correlation length scale. Trends in mean CN have a strong spatial structure: Majority of the watersheds in the east have negative trend and majority of the watersheds in the Pacific Northwest have positive trend. However, the largest magnitude of trends value is -2.4 time-window[-1] which translates to a decrease of 21.6; this is not a significant decrease. Decrease (increase) in mean CN indicates a decrease (increase) in equilibrium soil moisture in eastern (Pacific Northwestern) rain dominated watersheds. Decrease (increase) in equilibrium soil moisture is related to decrease (increase) in rainfall amount, and/or increase (decrease) in evapotranspiration. Thus, change in CN seem to be related to climate change, not to any physical characteristics of the watershed.
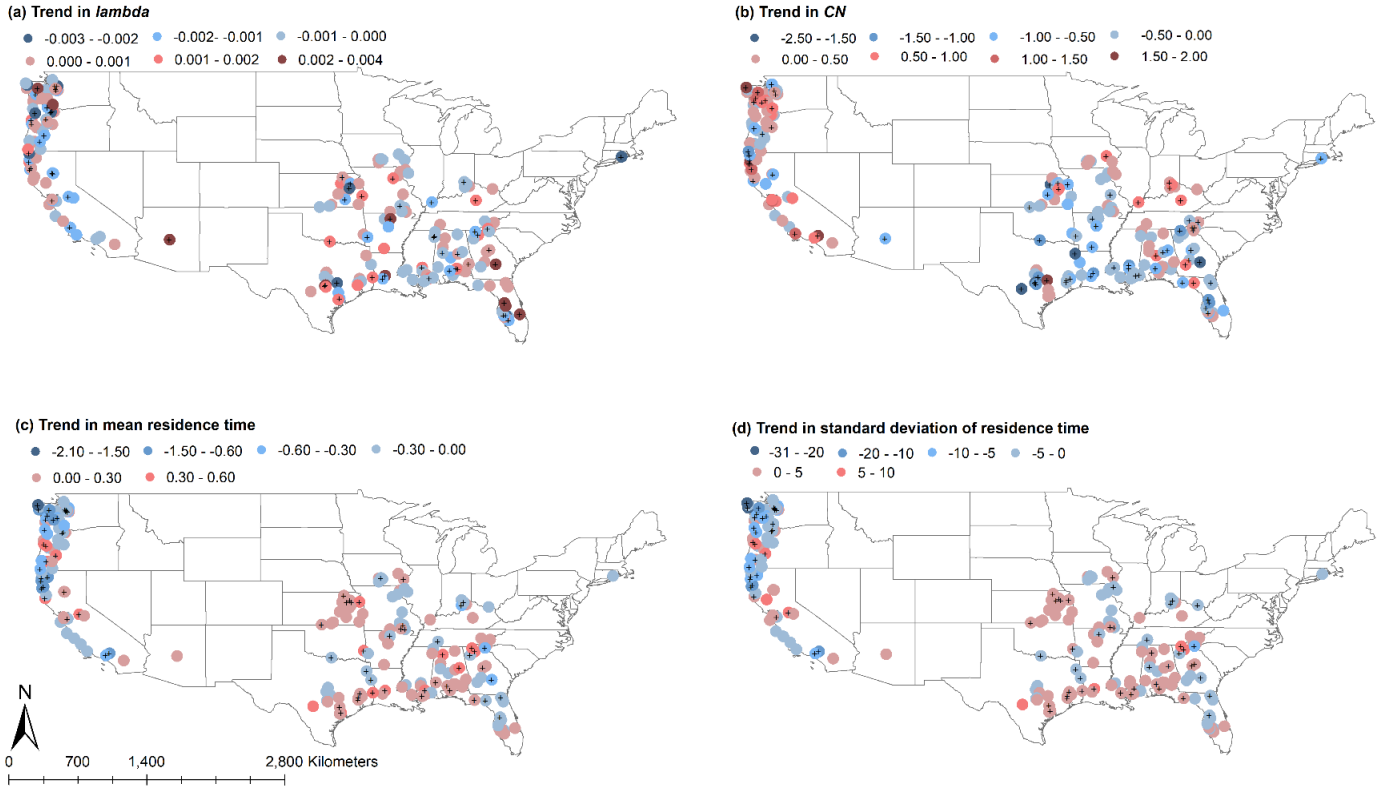
Figure S12. Rain-dominated watersheds. Trend in mean parameter values of rainfall-runoff models: (a) λ, (b) CN, (c) α/β, and (d) $\sqrt{\alpha/\beta^2}$. The '+' sign indicates that the trend is statistically significant at 5% significance level. The units in subplot (c) and (d) are minutes time-window^{-1}.

Trends in α/β and $\sqrt{\alpha/\beta^2}$ also show a spatial structure similar to that of CN: positive (negative) trends are more likely in East (West). Increase in both α/β and $\sqrt{\alpha/\beta^2}$ in a watershed implies that it takes more time to drain the watershed and the contribution of high frequency components to streamflow variations is decreasing. However, since the magnitude of change is small for both α/β and α/β^2, change in NPSD of streamflow is unlikely to be related to change in mean value of these parameters.

S5. Evidence of robustness of changes in streamflow statistical structure (SSS)

In this section, we discuss why the changes in the SSS reported in this study are robust. For the sake of brevity, only the changes in less than 1-month timescale components (F_4) are discussed. The same conclusions can be made for other components.

As explained in the main text, three tests were carried out to ascertain the statistical significance of the changes in SSS: (1) First significance test (see main text), (2) second significance test (see main text), and (3) Mann-Kendall test. The three tests agreed on the statistical significance of the changes for more than 70% of the watersheds. Figure S13 shows the boxplots of F_4 such that the changes in F_4 for the watersheds shown in were statistically significant according to the first and second significance tests, but statistically insignificant according to the Mann-Kendall test. The

Mann-Kendall test could not recognize these changes because the changes were abrupt. This illustrates that the changes in SSS reported in this study are robust.
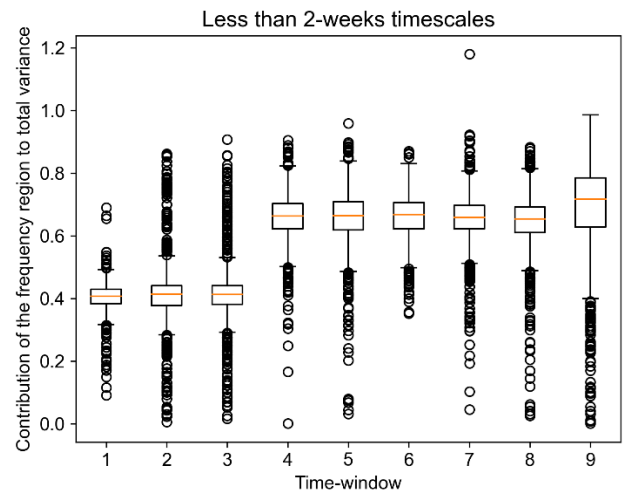
Figure S13 also illustrates that these changes are robust to the length of the time window used because the F_4 values in the last two or three time-windows (not just the last window) were significantly different from the F_4 values in the last two or three time-windows. This implies that even if we had estimated streamflow spectra using up to 16 years of daily streamflow data; the same changes in F_4 would have been obtained as obtained by using the 10-year time window.

The robustness to time-window length is also illustrated in Figure S14. Figure 14 shows the boxplots of F_4 values for each of the time-windows for a few of the watersheds for which all three tests (first significance test, second significance test, and Mann-Kendall test) agree that the F_4 values changed statistically significantly over the study period.

Finally, Figure S15 illustrates the effectiveness of the first and second significance tests. For the watersheds shown in this figure, the second significance concluded that F_4 changed significantly while the first significance concluded that F_4 did not change significantly. On visual inspection of the boxplots also, it can be concluded the changes indeed might be due to periodic fluctuations in streamflow rather than due to a systematic change.
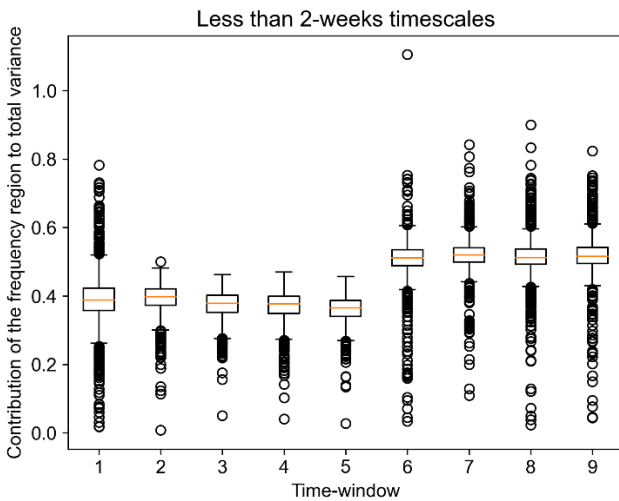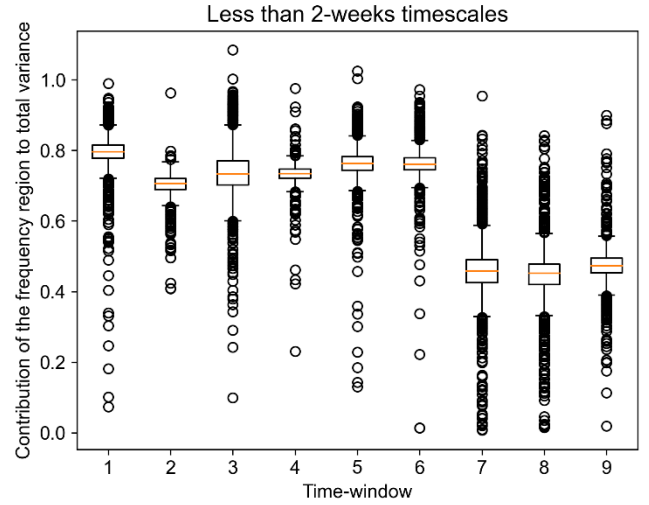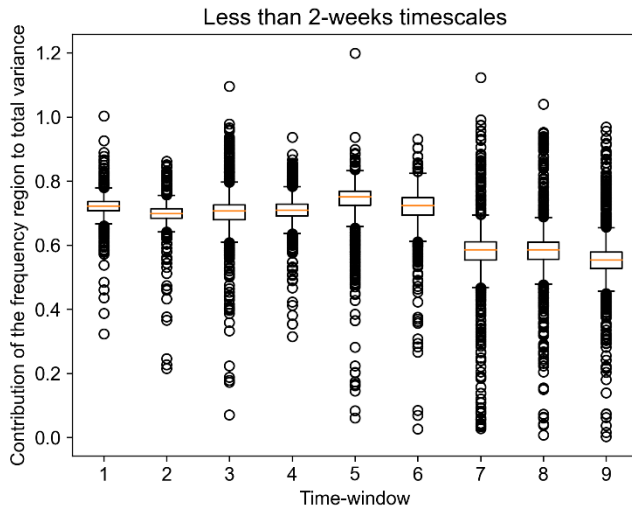

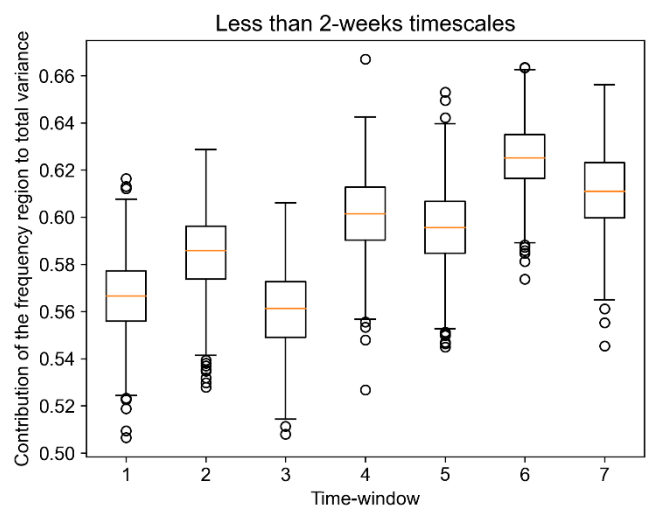Gauge number: 10259000   Gauge number: 08200000
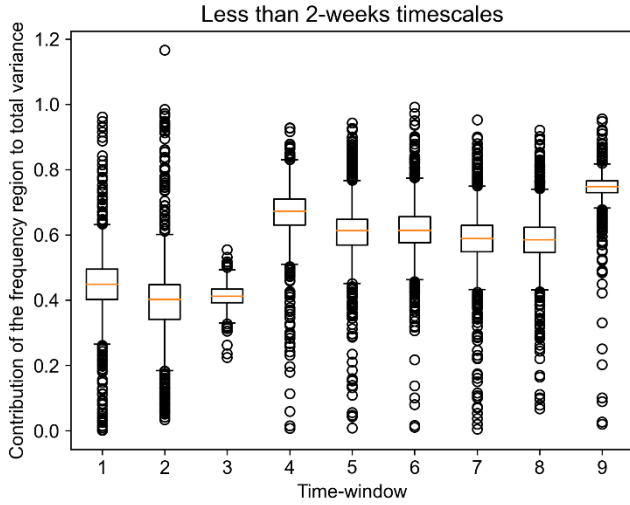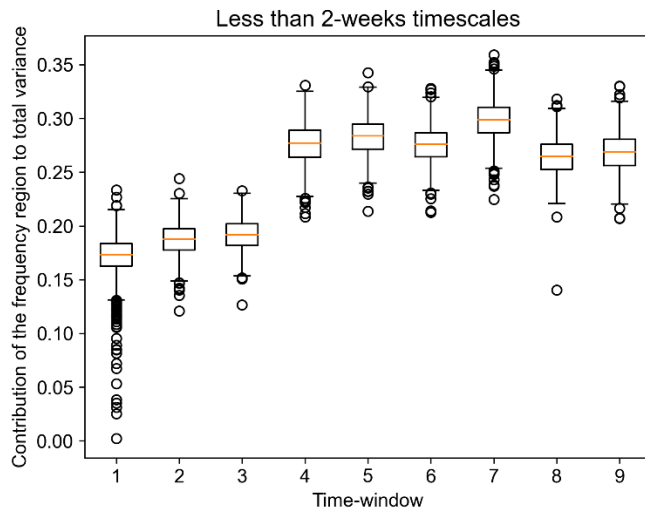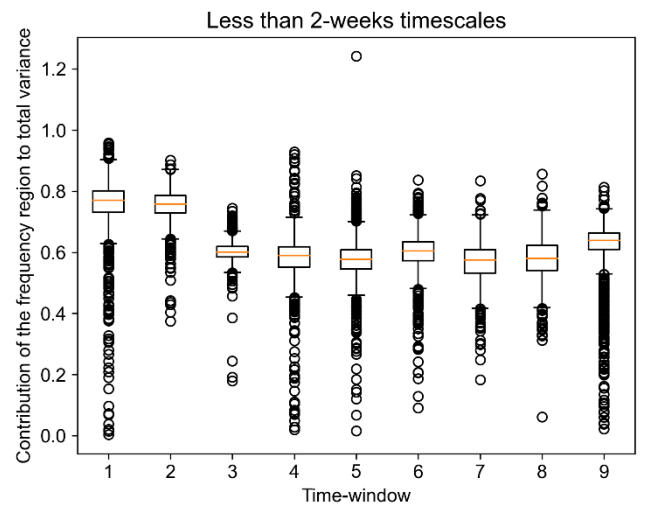
Gauge number: 08198500

Gauge number: 08086290

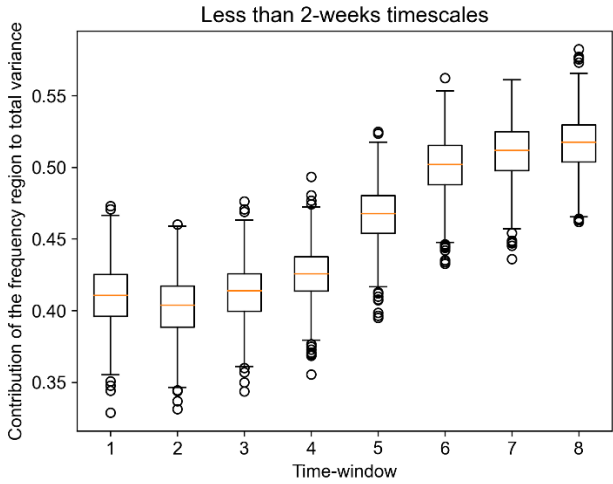Gauge number: 08086212

Gauge number: 03159540
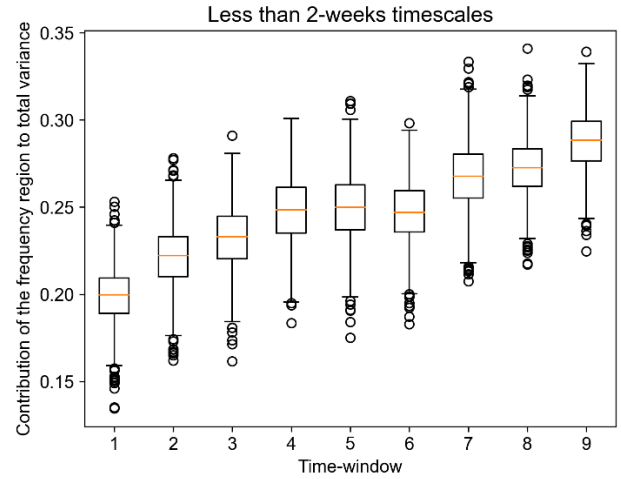

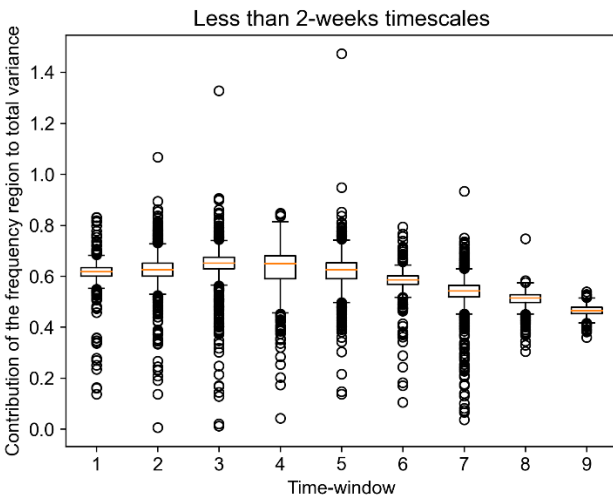Gauge number: 02092500

Gauge number: 02055100

Figure S13. Boxplots of the contribution of less than 2-weeks timescale components (F_4) for each time-window. These plots correspond to the watersheds where the first and second significance concluded that F_4 values changed statistically significantly while Mann-Kendall concluded that changed is statistically insignificant. Gauge number with each subplot refers to the unique ID of the streamflow gauge in CAMELS dataset. The lower and upper whiskers represent $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$, respectively. Here Q_1 and Q_3 denote 25th and 75th percentiles, and IQR denotes inter-quartile range, that is, $Q_3 - Q_1$.
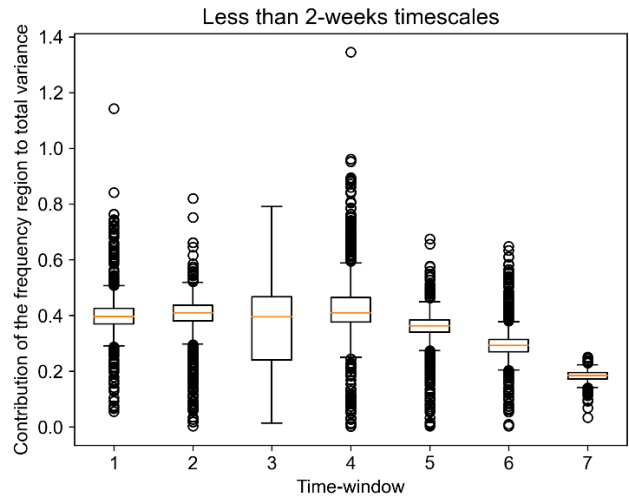

Gauge number: 03213700

Gauge number: 04105700

Gauge number: 08164600

Gauge number: 06479215
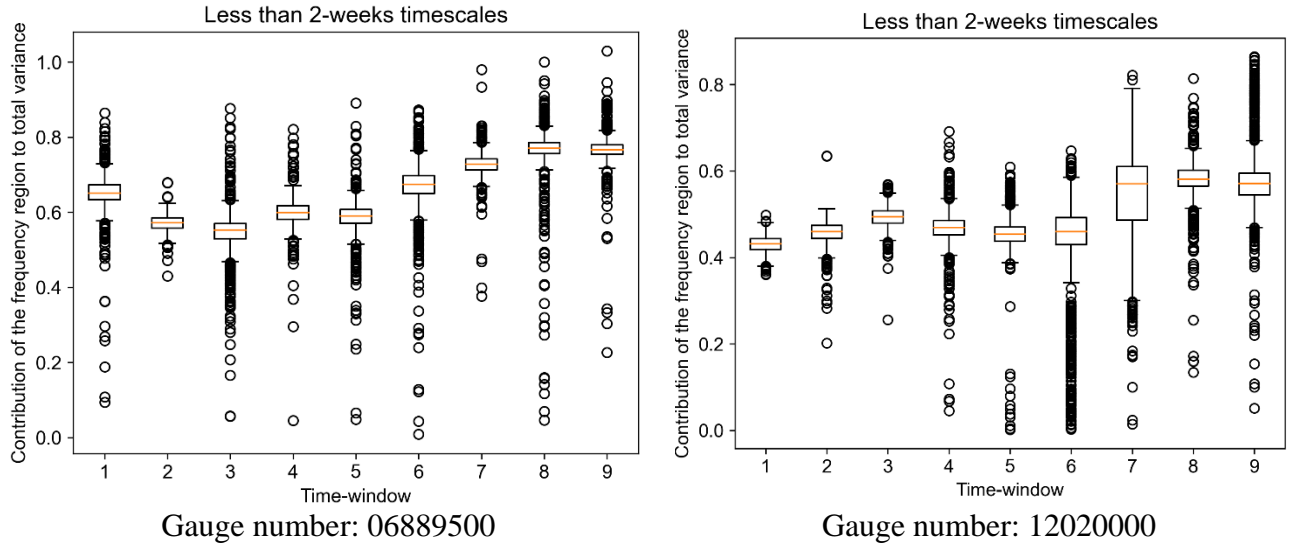
Gauge number: 06889500

Gauge number: 12020000

Figure S14. Boxplots of the contribution of less than 2-weeks timescale components (F_4) for each time-window. These plots correspond to the watersheds where all three statistical agreed that the F_4 values changed statistically significantly. Gauge number with each subplot refers to the unique ID of the streamflow gauge in CAMELS dataset. The lower and upper whiskers represent $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$, respectively. Here Q_1 and Q_3 denote 25th and 75th percentiles, and IQR denotes inter-quartile range, that is, $Q_3 - Q_1$.

Gauge number: 04161580
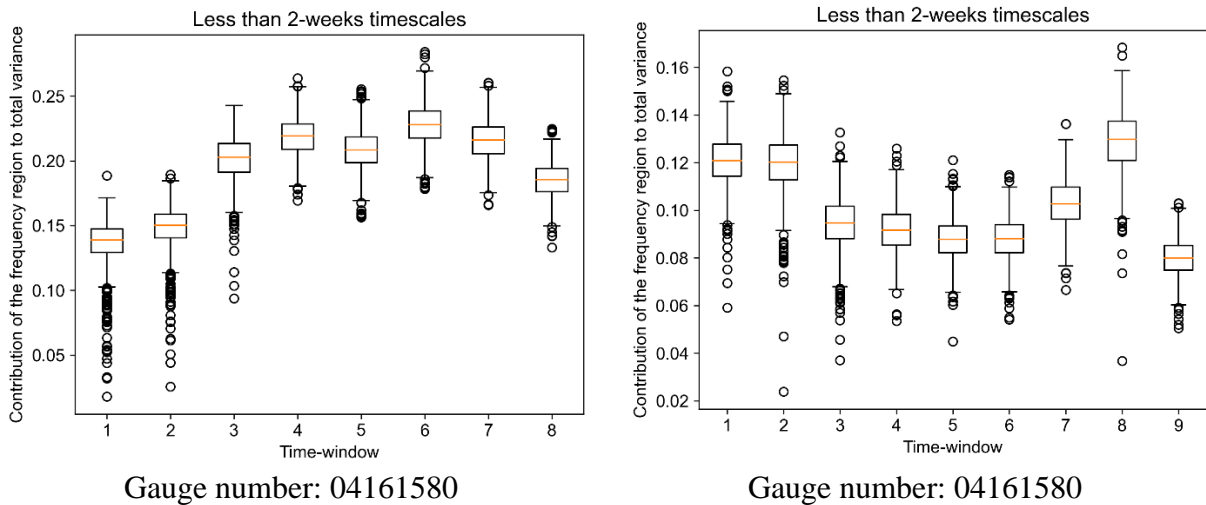
Gauge number: 04161580

Figure S15. Boxplots of the contribution of less than 2-weeks timescale components (F_4) for each time-window. These plots correspond to the watersheds where first and second significance tests disagreed that the F_4 values changed statistically significantly. Gauge number with each subplot refers to the unique ID of the streamflow gauge in CAMELS dataset. The lower and upper whiskers represent $Q_1 - 1.5 * IQR$ and $Q_3 + 1.5 * IQR$, respectively. Here Q_1 and Q_3 denote 25th and 75th percentiles, and IQR denotes inter-quartile range, that is, $Q_3 - Q_1$.

References:

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, in Petrov, B. N.; Csáki, F. (eds.), 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, pp. 267–281. Republished in Kotz, S.; Johnson, N. L., eds. (1992), Breakthroughs in Statistics, vol. I, Springer-Verlag, pp. 610–624.

Belmecheri, S., Babst, F., Wahl, E. R., Stahle, D. W., & Trouet, V. (2016). Multi-century evaluation of Sierra Nevada snowpack. *Nature Climate Change*, 6(1), 2-3.

Beran, J. (1994). *Statistics for long-memory processes.* Routledge.

Berg, N., & Hall, A. (2017). Anthropogenic warming impacts on California snowpack during drought. *Geophysical Research Letters*, 44(5), 2511-2518.

Brutsaert, W. (2005). *Hydrology: an introduction*. Cambridge University Press.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368), 829-836.

Collischonn, W., & Fan, F. M. (2013). Defining parameters for Eckhardt's digital baseflow filter. *Hydrological Processes*, 27(18), 2614-2622.

Geetha, K., Mishra, S. K., Eldho, T. I., Rastogi, A. K., & Pandey, R. P. (2007). Modifications to SCS-CN method for long-term hydrologic simulation. *Journal of Irrigation and Drainage Engineering*, 133(5), 475-486.

Horner, I., Branger, F., McMillan, H., Vannier, O., & Braud, I. (2020). Information content of snow hydrological signatures based on streamflow, precipitation and air temperature. *Hydrological Processes*, 34(12), 2763-2779.

Knowles, N., Dettinger, M. D., & Cayan, D. R. (2006). Trends in snowfall versus rainfall in the western United States. *Journal of Climate*, 19(18), 4545-4559.

Lamb, R., & Beven, K. (1997). Using interactive recession curve analysis to specify a general catchment storage model. *Hydrology and Earth System Sciences*, 1(1), 101-113.

Mishra, S. K., & Singh, V. P. (1999). Another look at SCS-CN method. *Journal of Hydrologic Engineering*, 4(3), 257-264.

Montanari, A., Rosso, R., & Taqqu, M. S. (1997). Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resources Research*, 33(5), 1035-1044.

Mote, P. W. (2006). Climate-driven variability and trends in mountain snowpack in western North America. *Journal of Climate*, 19(23), 6209-6220.

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). Declining mountain snowpack in western North America. *Bulletin of the American Meteorological Society*, 86(1), 39-50.

Ponce, V. M., & Hawkins, R. H. (1996). Runoff curve number: Has it reached maturity?. *Journal of Hydrologic Engineering*, 1(1), 11-19.

Seabold, S., & Perktold, J. (2010, June). Statsmodels: Econometric and statistical modeling with python. In Proceedings of the 9th Python in Science Conference (Vol. 57, p. 61).

Soulis, K. X., & Valiantzas, J. D. (2012). SCS-CN parameter determination using rainfall-runoff data in heterogeneous watersheds–the two-CN system approach. *Hydrology and Earth System Sciences*, 16(3), 1001-1015.

Soulis, K. X., & Valiantzas, J. D. (2013). Identification of the SCS-CN parameter spatial distribution using rainfall-runoff data in heterogeneous watersheds. *Water Resources Management*, 27(6), 1737-1749.

Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1).