Earth's Future

Supporting Information for

Peak Runoff Timing is Linked to Global Warming Trajectories

Donghui Xu[1*], Valeriy Y. Ivanov[1*], Xiuyuan Li[2], Tara J. Troy[3]

[1]Univeristy of Michigan-Ann Arbor

[2]Lehigh University

[3]Department of Civil Engineering, University of Victoria, Victoria, BC Canada

Corresponding authors: Donghui Xu (donghui.xu@pnnl.gov),

Valeriy Y. Ivanov (ivanov@umich.edu)

† Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, tel.: 734-730-8326.

Contents of this file

SM. 1 Historical runoff dataset

The absolute majority of watersheds with monitored streamflow are much smaller than the grid cell size of typical GCM resolution, therefore, it is not appropriate to directly correct biases of GCM runoff estimates with streamflow observations. Long-term estimates of daily runoff (surface water yield per unit area) provided by *Livneh et al.* (2013) is used in this study as true "observations" within the Bayesian framework of multi-model downscaling. This daily runoff is obtained as the output of the Variable Infiltration Capacity (VIC) model (*Liang et al.*, 1994) forced with precipitation and temperature, at the spatial resolution of $1/16^{\circ} \times 1/16^{\circ}$.

To verify the consistency of the *Livneh et al.* (2013) runoff data set, we selected 5,217 daily streamflow gauges from the United States Geological Survey (USGS) data record. We selected records with at least 15 years of data between 1961 and 1990. The accuracy of the average annual runoff magnitude and annual maximum runoff date of the *Livneh et al.* (2013) data set is verified in Figure S1. Specifically, the discharge time series of the USGS gauge data were converted to annual runoff by dividing the series by the respective watershed area (*Fekete et al.*, 2002). Runoff of all grid cells in *Livneh et al.* (2013) that are located inside a watershed corresponding to a given USGS gauged outlet were aggregated. The annual peak runoff dates from both datasets were subsequently identified and averaged for the selected time period (e.g., 1961-1990) with the method of directional statistics (*Berens*, 2009).

SM.2 Robustness of GCM estimates

Although the Bayesian framework is used to increase the confidence of multi-model projections, the projection uncertainty due to model-specific biases can still be significant, particularly at a

local scale (*Xu et al.*, 2018). To investigate the consistency of the GCMs used, a robustness metric (ψ) introduced in *Knutti and Sedláček* (2012) is used in this work:

$$\psi = 1 - \frac{A_1}{A_2},$$

where A_1 is defined as the integral of the squared area between two cumulative density functions (CDFs), one constructed from the combined individual model projections and the other one from the BWA projection for a future period. Variable A_2 is the integral of the squared area between two CDFs, both constructed from the BWA of model outputs for future and past periods. The A_1 and A_2 represent 'model noise' (equal to zero if individual model projections perfectly overlap; attains a large value if they do not) and 'model signal', respectively. Therefore, the robustness metric can be interpreted as a measure of model relative agreement on the change in future projections. In the original application of (*Knutti and Sedláček*, 2012), the Gaussian distribution is assumed for CDFs and their parameters are estimated from the sample mean and variance. In this work, we use the von Mise distribution (*Abramowitz*, 1974) because of the circular nature of the downscaled DOY variables. The two parameters of the distribution, preferred direction (μ) and concentration (κ), are estimated from the samples with the circular statistics tool as in (*Berens*, 2009). Specifically, the parameters μ and $1/\kappa$ are analogous to the mean and variance in the normal distribution.

SM.3 Example of linear variable conversion

For example, for a given grid cell location, the observed DOY of annual peak runoff is 5 (i.e., $X_{i0} = 5$), and two climate models can produce outputs of annual peak runoff on days 10 and 365

(i.e., $X_{i1} = 10$ $X_{i2} = 365$). Although both of such outputs correspond to the dates equally distant from the observation, the model with the value of 365 would be regarded as having a poor skill in the original BWA due to the large absolute difference of DOY values placed on a linear scale. In Eq (10) and (11), directional statistics (*Berens*, 2009) is used to calculate the difference between $X_{i1}(X_{i2})$ and X_{i0} , which leads to $\tilde{X}_{i1} = 5$ and $\tilde{X}_{i2} = -5$. These two models will therefore be assigned equal weights for evaluating the multi-model ensemble after the conversion because of the equal absolute deviation. Therefore, $\tilde{X}_{ij}$ and $\tilde{Y}_{ij}$ are used as the variable of interest instead of X_{ij} and Y_{ij} in the modified BWA approach.

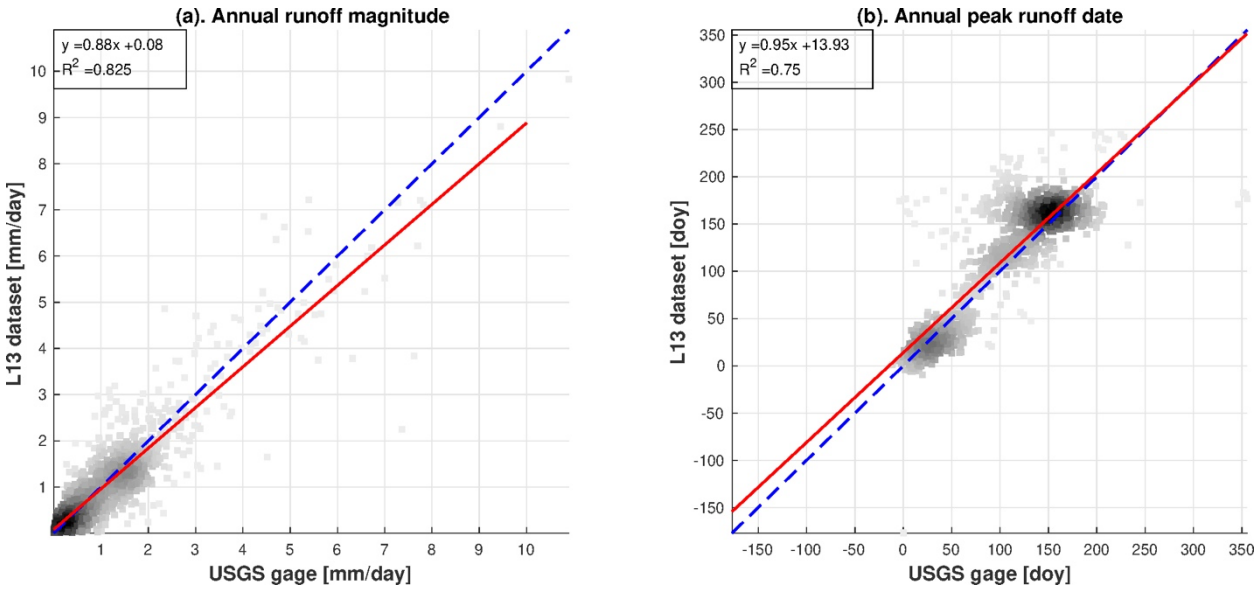

Figure S1. Validation of runoff data of *Livneh et al.* (2013). 2,240 USGS streamflow gauges for which the concentration metric of *Bloschl et al.* (2017) is larger than 0.6 are used to validate (a) the mean annual runoff magnitude, and (b) the mean annual peak runoff date over 1961 – 1990. The X-axis represents the data based on observations for USGS gauges and the Y-axis represents the data product of *Livneh et al.* (2013). The blue dashed line is the 1:1 reference line and the solid red line is the linear regression line. The shading level depth is positively associated with the point density.

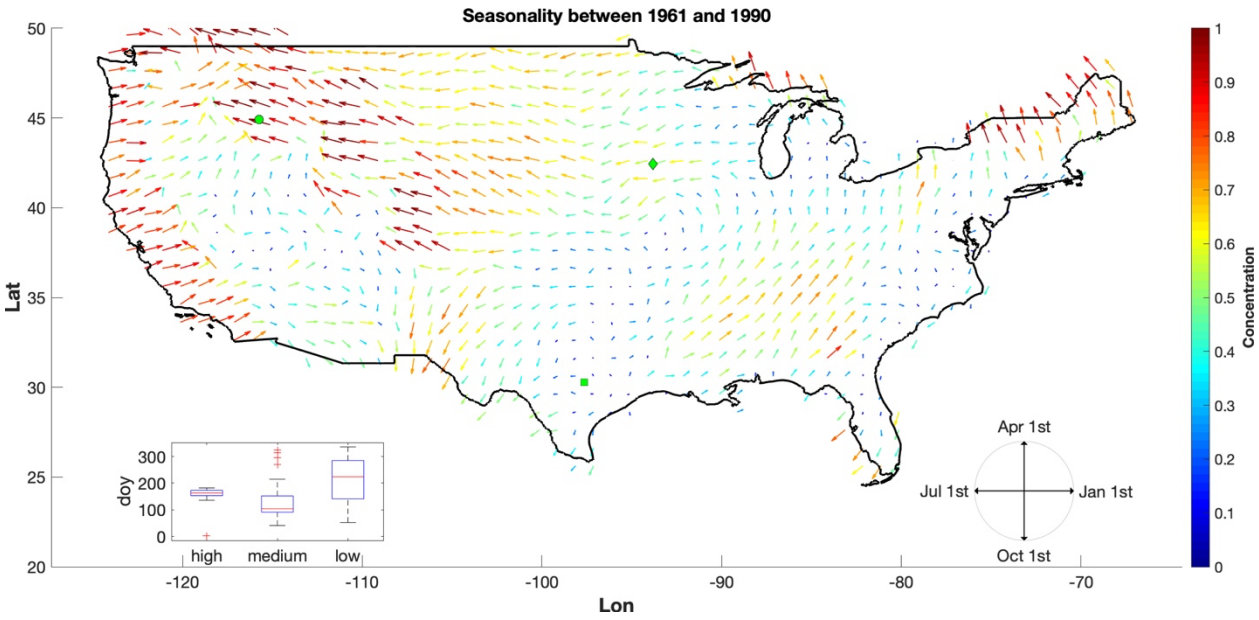

Figure S2. Peak annual runoff timing seasonality between 1961 and 1990 computed using daily runoff product of Livneh et al. (2013). The arrow direction represents the mean occurrence date computed over the 30-year period, and the length and color represent the concentration metric (ψ) of Bloschl et al. (2017). The concentration metric ψ varies between 0 and 1, where 1 means that peak annual runoff occurs on the same date, and 0 implies peak runoff date can occur on any day within a year. Three exemplary locations are selected to visualize the distribution of peak annual runoff dates in the plot insert. The green circle, diamond, and square correspond to high ($\psi = 0.9$), medium ($\psi = 0.6$) and low ($\psi = 0.1$) values of the concentration metric.

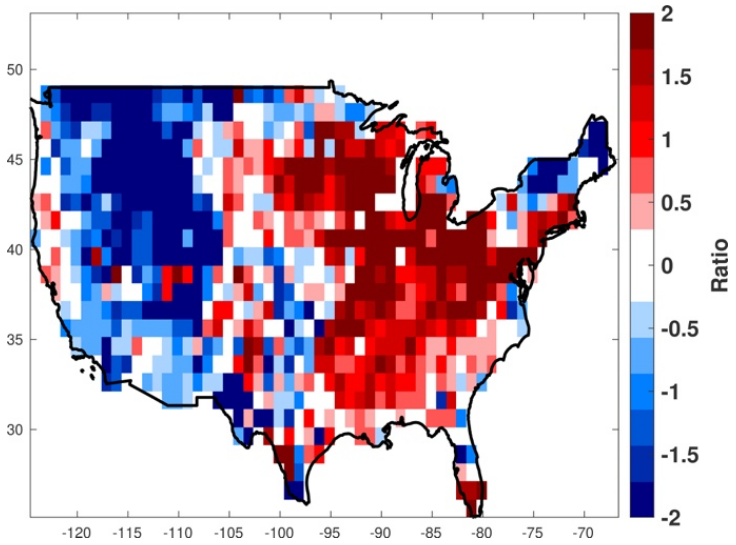

Figure S3. Significance of changes in timing of the peaks compared to historical variability. The value of each grid cell shows the ratio between the peak runoff timing change (from the Bayesian posterior distribution corresponding to maximum likelihood) and the standard deviation of peak runoff timing in the historical period. The projection for the end-of-entry, RCP 8.5 scenario is used here.

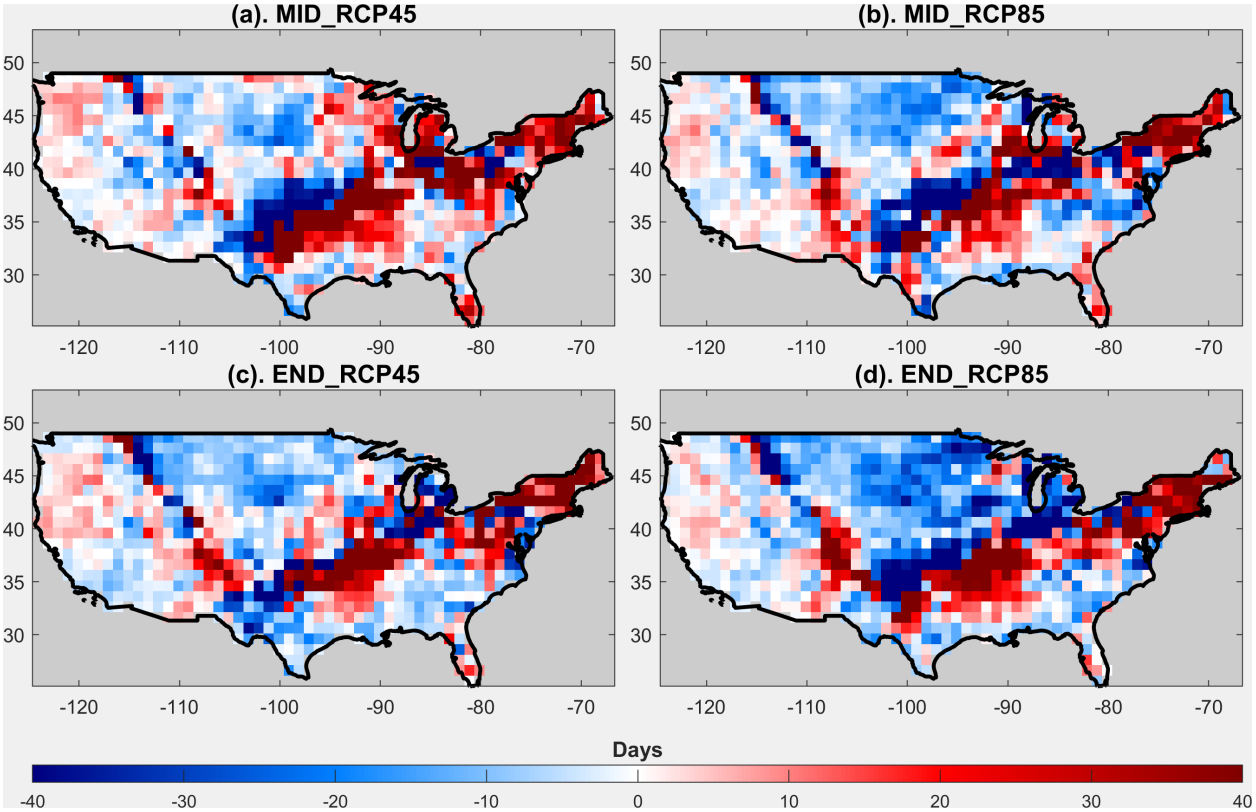

Figure S4. Change of annual peak precipitation time between the control (CTL) and future periods (FUT). "MID" (subplots (a) and (b)) represents the date difference with respect to 2041-2070 and 1961-1990 periods, and "END" (subplots (c) and (d)) represents the difference with respect to 2071-2100 and 1961-1990 periods. The differences are calculated from the simple average mean of selected GCMs in terms of annual peak precipitation time (e.g., future – control).

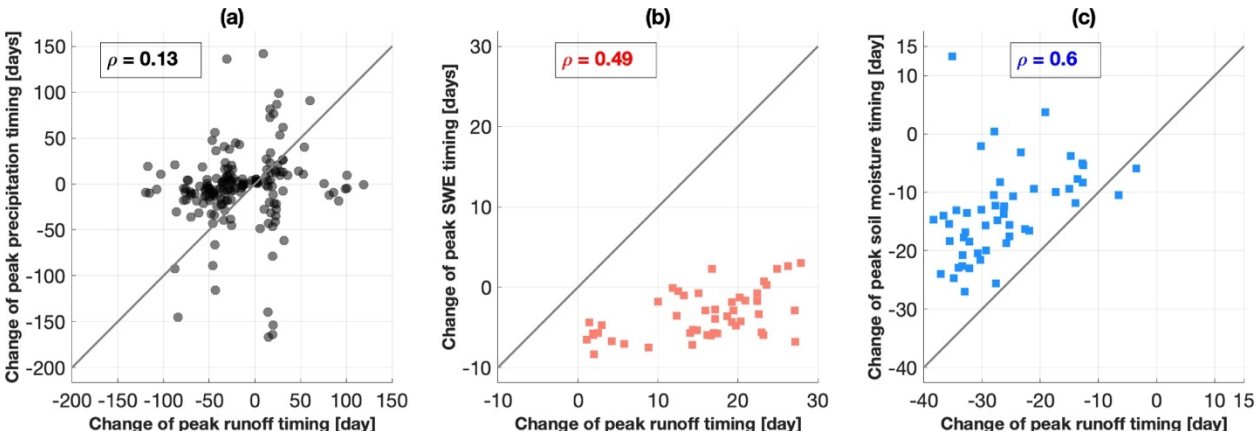

Figure S5. Regressions between the peak annual runoff timing change and attribution variables. (a). Peak annual precipitation timing, (b) change of the date of maximum snow water equivalent, and (c) the shift of spring soil wetness. In all subplots, only grid cells with the change robustness metric larger than 0.6 for peak annual runoff timing are presented. In subplot (a), both positive and negative values of peak runoff timing are taken. In subplot (b), only the grid cells with positive changes are used (i.e., the red cells with inner green circles in Fig. 1d). In subplot (c), only the grid cells with negative changes are presented (i.e., the blue cells with inner green circles in Fig. 1d). The grey solid line represents the 1:1 reference line, and ρ is the correlation coefficient.

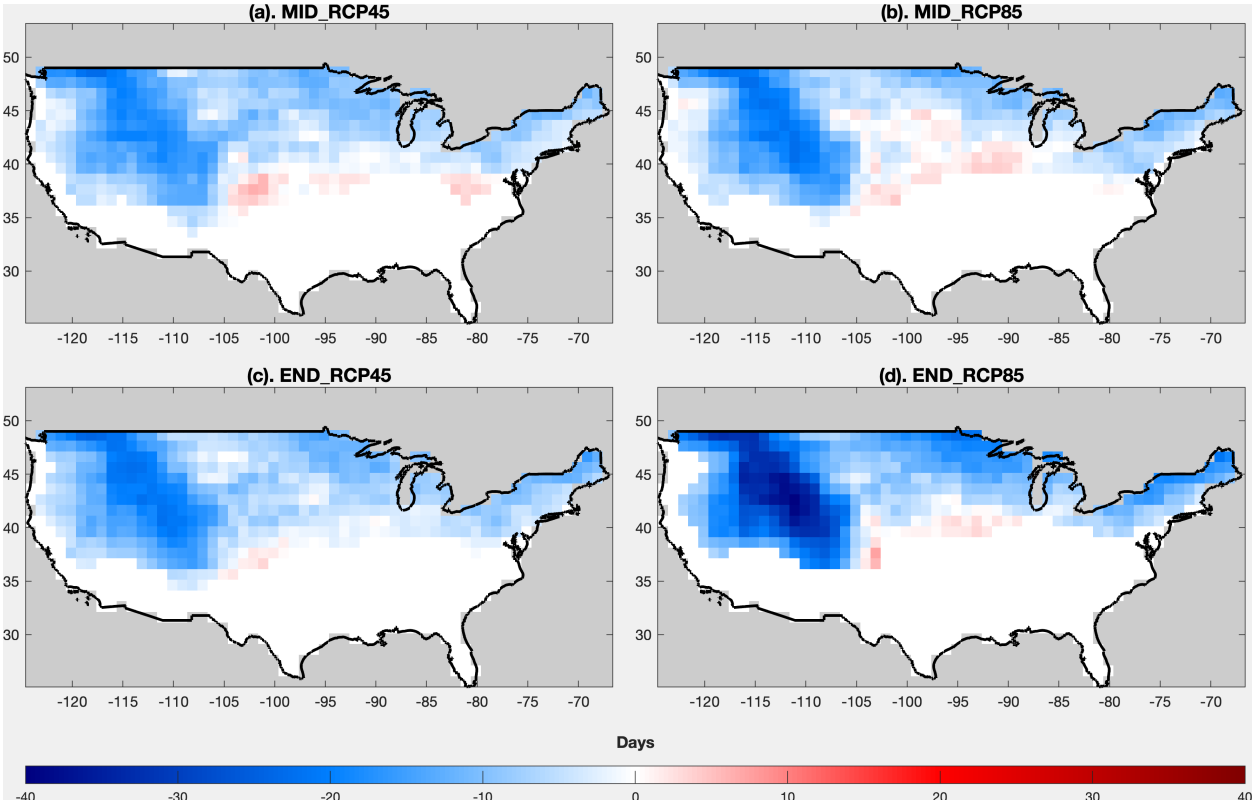

Figure S6. Change of the date of annual maximum snow water equivalent between control (CTL) and future periods (FUT). The differences are calculated from the simple average mean of annual timing of maximum snow water equivalent (i.e., future - control). The white pixels represent the maximum snow water equivalent less than 15 [kg/m^2] (i.e.,15 [mm] as liquid water depth) in the future period.

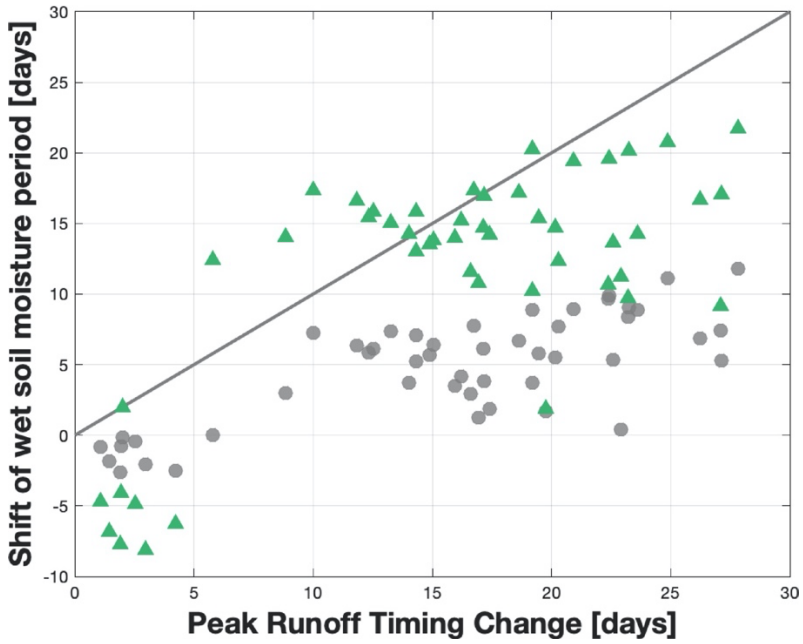

Figure S7. Shift of wet soil moisture period at 25% and 75% frequency of the CDF of dates of springtime extreme soil wetness. The regression between the peak annual runoff timing change and the shift of CDF of wet soil moisture periods during springtime (February to May) at 25% (grey dots) or 75% (green triangles) for the end-century and RCP 8.5 scenario. Only pixels with a delay shift of peak annual runoff timing with high robustness metric are presented. The correlation coefficients are 0.77 and 0.71, respectively. The grey line represents the 1:1 reference line.

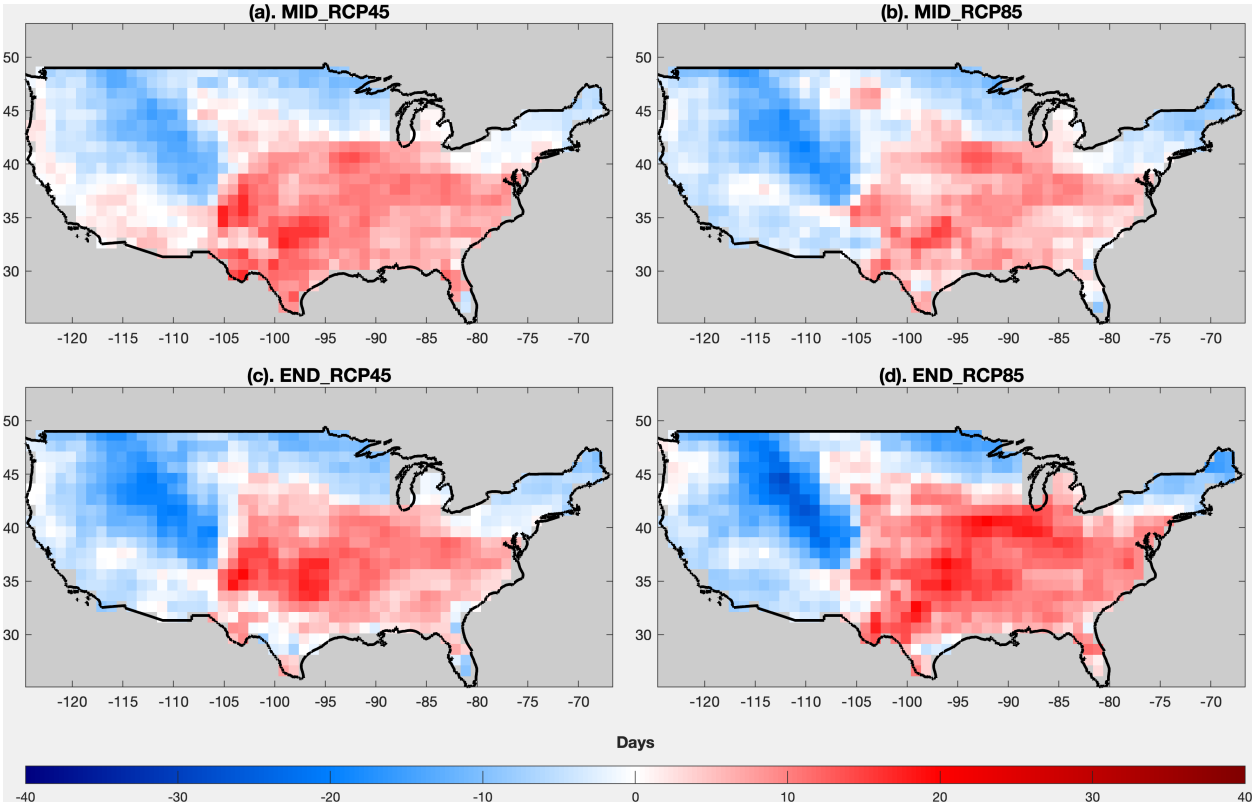

Figure S8. The shift of centroid date of extreme soil saturation during the springtime period between the control (CTL) and future periods (FUT). The period includes dates between February and May on which soil moisture is larger than 95% of the saturated soil moisture content. CDFs of soil saturation dates are constructed using all of the selected GCMs in each grid cell. The values shown in the subplots are the differences between the CDFs for the control and future periods at 50% percentile.

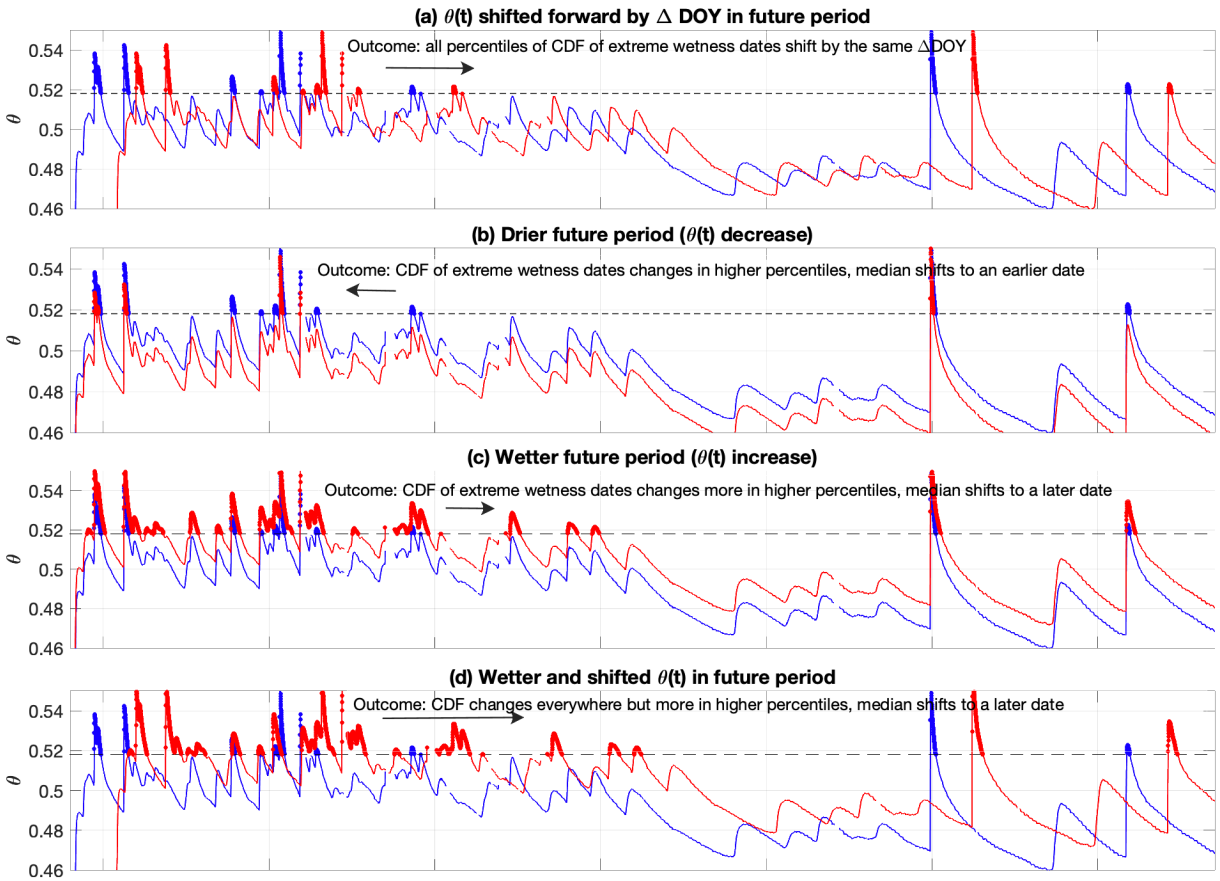

Figure S9. Conceptual examples of soil moisture time series with a shift of wetness centroid between control and future periods. The blue solid line and the red solid line represent (synthetic/hypothetical) soil moisture time series for the control and future period, respectively. (a) An example in which the entire soil moisture time series is shifted forward in future. (b) Soil is drier in the future, which results in the shift of centroid of the CDF to an earlier date. (c) Soil is wetter in the future, and the centroid of the CDF shifts to a later date. (d) A combination of (a) and (c) leading to a more significant shift of the centroid of the CDF to a later date.

References

Abramowitz, M. (1974), *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc.

Berens, P. (2009), CircStat: A MATLAB Toolbox for Circular Statistics, *2009*, *31*(10), 21.

Bloschl, G., et al. (2017), Changing climate shifts timing of European floods, *Science*, *357*(6351), 588-590.

Fekete, B. M., C. J. Vörösmarty, and W. Grabs (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochemical Cycles*, *16*(3), 15-11-15-10.

Knutti, R., and J. Sedláček (2012), Robustness and uncertainties in the new CMIP5 climate model projections, *Nature Climate Change*, *3*(4), 369-373.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415-14428.

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier (2013), A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions, *Journal of Climate*, *26*(23), 9384-9392.

Xu, D., V. Y. Ivanov, J. Kim, and S. Fatichi (2018), On the use of observations in assessment of multi-model climate ensemble, *Stochastic Environmental Research and Risk Assessment*.