Contribution of environmental forcings to US runoff changes for the period 1950–2010

Environmental Research Letters

LETTER

Contribution of environmental forcings to US runoff changes for the period 1950–2010

Whitney L Forbes[1,2], Jiafu Mao[2,20], Mingzhou Jin[1], Shih-Chieh Kao[2], Wenting Fu[3], Xiaoying Shi[2], Daniel M Riccuito[2], Peter E Thornton[2], Aurélien Ribes[4], Yutao Wang[5], Shilong Piao[6,7,8], Tianbao Zhao[9], Christopher R Schwalm[10,11], Forrest M Hoffman[12], Joshua B Fisher[13], Akihiko Ito[14], Ben Poulter[15], Yuanyuan Fang[16], Hanqin Tian[17], Atul K Jain[18] and Daniel J Hayes[19]

[1] Department of Industrial and Systems Engineering, University of Tennessee, Knoxville, TN, United States of America
[2] Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, United States of America
[3] Jackson School of Geosciences, the University of Texas, Austin, TX, United States of America
[4] Centre National de Recherches Météorologiques, Météo-France/CNRS, 42 Avenue Gaspard Coriolis, Toulouse 31057, France
[5] Fudan Tyndall Center, Department of Environmental Science and Engineering, Fudan University, Shanghai 200438, People's Republic of China
[6] Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100871, People's Republic of China
[7] Key Laboratory of Alpine Ecology and Biodiversity, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, People's Republic of China
[8] CAS Center for Excellence in Tibetan Plateau Earth Science, Beijing 100085, People's Republic of China
[9] Key Laboratory of Regional Climate-Environment Research for East Asia, Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences (CAS), Beijing, People's Republic of China
[10] Woods Hole Research Center, Falmouth, MA 02540, United States of America
[11] Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86011, United States of America
[12] Computer Science and Mathematics Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States of America
[13] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, United States of America
[14] Center for Global Environmental Research, National Institute for Environmental Studies (NIES), Onogawa 16-2, Tsukuba, Ibaraki 305–8506, Japan
[15] NASA Goddard Space Flight Center, Biospheric Sciences Lab., Greenbelt, MD 20771, United States of America
[16] Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, United States of America
[17] International Center for Climate and Global Change Research and School of Forestry and Wildlife Science, Auburn University, Auburn, AL, United States of America
[18] Department of Atmospheric Sciences, University of Illinois, Urbana, IL 61801, United States of America
[19] School of Forest Resources, University of Maine, Orono, ME 04459, United States of America
[20] These authors contributed equally to this work.

E-mail: maoj@ornl.gov **and** jin@utk.edu

Keywords: US runoff, detection and attribution, MsTMIP

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
30 January 2018

REVISED
28 March 2018

ACCEPTED FOR PUBLICATION
3 April 2018

PUBLISHED
4 May 2018

Abstract

Runoff in the United States is changing, and this study finds that the measured change is dependent on the geographic region and varies seasonally. Specifically, observed annual total runoff had an insignificant increasing trend in the US between 1950 and 2010, but this insignificance was due to regional heterogeneity with both significant and insignificant increases in the eastern, northern, and southern US, and a greater significant decrease in the western US. Trends for seasonal mean runoff also differed across regions. By region, the season with the largest observed trend was autumn for the east (positive), spring for the north (positive), winter for the south (positive), winter for the west (negative), and autumn for the US as a whole (positive). Based on the detection and attribution analysis using gridded WaterWatch runoff observations along with semi-factorial land surface model simulations from the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP), we found that while the roles of CO_2 concentration, nitrogen deposition, and land use and land cover were inconsistent regionally and seasonally, the effect of climatic variations was detected for all regions and seasons, and the change in runoff could be attributed to climate change in summer

and autumn in the south and in autumn in the west. We also found that the climate-only and historical transient simulations consistently underestimated the runoff trends, possibly due to precipitation bias in the MsTMIP driver or within the models themselves.

1. Introduction

Water is one of the most essential resources for the terrestrial biosphere as well as for human society; thus, it is important to detect and understand the potential drivers of changes in the hydrological cycle (Lettenmaier *et al* 1993, Barnett *et al* 2008, Gedney *et al* 2014, Koster *et al* 2017). Other than supplying drinking water, many facets of society and various ecosystems rely on freshwater resources and therefore are impacted by hydrological changes. Among these are irrigation, building and infrastructure planning, power generation, recreation, and plant/animal life cycles. Changes in the hydrological cycle can affect soil moisture that is crucial for agricultural activities. For regions experiencing drying, farmers need to switch to more drought-resistant crops or increase groundwater usage that may eventually lead to imbalance between groundwater withdrawal and recharge (Scanlon *et al* 2012). For regions getting wetter, more extreme events combined with land use and land cover change (LULCC) may lead to elevated flood likelihood and expanded flood plains (Collins 2008, Singh *et al* 2014). For example, without adaptive actions being taken, Metropolitan Boston is estimated to incur $26 billion in total losses due to climate change driven river flooding by 2100 (Romero-Lankao *et al* 2014, Kirshen *et al* 2008, Nicholls *et al* 2008, Richardson 2010, Weiss *et al* 2011). Using the Special Report on Emissions Scenarios (SRES) A2 emissions scenarios, Westerling *et al* (2011) estimated that by 2085 there will be significant increases in wildfire occurrence and burned area in California, due to effects on evapotranspiration through increased temperatures and reduced precipitation. According to Barnett *et al* (2005), by 2050 with projected climate change, the Columbia River system will not be able to sustain both water releases for summer and autumn hydroelectric power and spring and summer releases for salmon runs unless there is a 10%–20% reduction of hydropower generation. Compared to 2010, a drought in 2011 led to a 30% decrease in monitored reservoir storage for power plant cooling in Texas (Scanlon *et al* 2013). At the most basic level, changes in a region's hydrological cycle affect its natural ecosystems. This includes species of freshwater fish which need specific river flow conditions for breeding (Wenger *et al* 2010).

A decrease in water availability can also lead to plant mortality (Romero-Lankao *et al* 2014, Anderegg *et al* 2012). Inversely, plants affect runoff through canopy interception, evaporation, and transpiration (Gerten *et al* 2004, Betts *et al* 2007, Piao *et al* 2007, Mao *et al* 2015), and rooting strategy (Nepstad *et al* 1994, Fan *et al* 2017). Through physiological effects, increasing CO_2 may lead to reduced stomatal conductance or increased photosynthesis with either positive or negative impacts on plant transpiration (Betts *et al* 2007, Shi *et al* 2013, Mao *et al* 2015). Increasing nitrogen deposition can lead to more nitrogen fertilization causing increased vegetation growth and altered hydrologic dynamics in regions where nitrogen is limiting (Thornton *et al* 2007). LULCC directly affect the potential for evapotranspiration (Shi *et al* 2011). For example, deforestation leads to decreased evapotranspiration which then leads to increased runoff, whereas decreased runoff is possible after reforestation (Gerten *et al* 2004, Bosch and Hewlett 1982, Piao *et al* 2007).

Previous studies have found that key hydrological variables (e.g. precipitation, streamflow, and snowpack) are changing in the US. Over the entire contiguous US (CONUS) for the period 1950–2000, Groisman *et al* (2004) found increases in precipitation, temperature, and streamflow. Taking a more regional focus, they found an increase in precipitation and streamflow in the eastern US with an increase in dryness in the west. Petersen *et al* (2012) determined that the spatial variability in runoff seasonality in the eastern US depends on covariation between moisture and energy cycles, whereas the west shows a negative correlation leading to dependence on basin aridity and the seasonality of precipitation. Focusing on the western US, detection and attribution (D&A) studies attribute declining snowpack and streamflow timing changes to human effects, especially the human-induced elevation of CO_2 concentration (Barnett *et al* 2008, Pierce *et al* 2008, Hidalgo *et al* 2009). The limitation of Groisman *et al* (2004), Petersen *et al* (2012), and Alkama *et al* (2013), however, was that causality of changes in runoff could not be addressed due to solely using observational data or focusing on the detection issue. The work presented in this paper takes a step forward by addressing the causality using a gridded observational dataset and an ensemble of offline land surface models (LSMs) driven by the same observed environmental conditions in order to perform more robust D&A analysis.

We focus on the change of runoff in the US since it provides a 'spatial and temporal integrator of changes in the water cycle' (Gedney *et al* 2014). We examine if runoff is changing (detection), and also seek to understand how and why changes might occur (attribution). River runoff can be thought of as the difference between long-term precipitation and evapotranspiration without the effects of storage changes (Gedney *et al* 2014).

Thus, any mechanisms that affect precipitation or evapotranspiration affect river runoff. Model simulations are used to estimate the responses to individual external forcings (Gedney *et al* 2014). For D&A analysis, the corresponding response patterns are then used to estimate the amplitude of the change induced by each forcing in the observations. The environmental forcings considered here are climate change, CO_2 concentration, nitrogen deposition, and LULCC. A goal of this study is to determine if climate alone is driving changes in US runoff or if other major anthropogenic factors also have a significant impact for certain regions or seasons.

2. Data and methodology

2.1. Data and data processing

We used observed 1950–2010 monthly runoff from the US Geological Survey (USGS) WaterWatch runoff dataset (Brakebill *et al* 2011) to investigate the historical trends of US runoff. The period of study ends in 2010 due to the temporal coverage of the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) model simulations. Derived from the comprehensive USGS National Water Information System gauge observations, WaterWatch runoff is the assimilated time series of flow per unit of area calculated for each 8 digit hydrologic unit (HUC8) in the CONUS. For each HUC8, multiple National Water Information System (NWIS) gauge stations located within the HUC8 or downstream were used to estimate the runoff generated locally at each HUC8, with gauge weighting factors determined by joint contributing drainage areas (both gauge-to-HUC8 and HUC8-to-gauge). This approach effectively assimilates streamflow observations from multiple gauge stations as a consistent areal HUC8 runoff measurement with a unit similar to that for precipitation (depth/time). WaterWatch runoff has been used and discussed in several recent hydroclimate studies, including Beigi and Tsai (2014), Oubeidillah *et al* (2014), Schwalm (2015), and Naz *et al* (2016). Note that since WaterWatch does not explicitly exclude gauges that were under flow regulation, the runoff estimates in HUC8s with significant historical human impairments could be biased. To verify WaterWatch's applicability for this study, we compared its values with another commonly used data set (Dai *et al* 2009). When aggregating WaterWatch runoff to the same watersheds used by Dai *et al* (2009), a good agreement between both data sets was found (figure S1).

Simulated runoff from all-factor and single-factor simulations from the North American Carbon Program MsTMIP (Huntzinger *et al* 2013) was compared to WaterWatch runoff. For the all-factor simulation, all environmental drivers were allowed to vary throughout the fully transient simulation (named ALL). In the climate-only simulation, the climatic factors (e.g. temperature, precipitation, and shortwave radiation) are transient while CO_2 concentration, nitrogen deposition, and land use and land cover are held constant at their preindustrial values (named CLMT). The third simulation uses transient climate and land use and land cover, while the fourth simulation allows transient climate, land use and land cover change and CO_2 concentration. We use the difference between the third simulation and CLMT to isolate the effect of land use and land cover change (named LULCC), and use the difference between the forth and the third simulations to achieve the effect of atmospheric CO_2 concentration (named CO2). To isolate the effect of nitrogen deposition, we use the difference of ALL and the fourth simulation (named NDEP). In this paper, the term 'environmental forcings' is used because the radiative and physiological effects of CO_2 concentration on climate change cannot be separated by using offline LSM simulations and are included in the transient climate drivers (Gedney *et al* 2014, Mao *et al* 2016, Zhu *et al* 2016). CO2, NDEP, and LULCC thus represent the direct effects of CO_2 physiology, nitrogen deposition, and land use and land cover change, respectively. More details of the experimental design used within the MsTMIP modeling framework can be found in Huntzinger *et al* (2013) and Mao *et al* (2015). All MsTMIP models use the same spatial resolution ($0.5° \times 0.5°$), are forced with CRU-NCEP reanalysis meteorology, and use the same anthropogenic forcings (Wei *et al* 2014). The specific meteorological variables used by each model are listed in supplementary table S1 available at stacks.iop.org/ERL/13/054023/mmedia. Land use information was provided, but each modeling group customized the processing of this information to fit its unique definition of plant functional types. Ensemble sizes and specific MsTMIP models employed are listed in table 1. Analysis was completed using the multi-model ensemble means (MME), but some of the results for individual models are included in the supplementary material.

The HUC8-based WaterWatch runoff was remapped to the 0.5° MsTMIP grid for direct comparison. For each grid cell, the overlapping HUC8s and their overlapped areas were first identified using geographic information system (GIS). The overlapped areas were then used as weighting factors to average monthly runoff time series. The metrics used for the detection and attribution are annual and seasonal runoff (winter–December to February, DJF; spring–March to May, MAM; summer–June to August, JJA; autumn–September to November, SON). These metrics were examined at three different spatial resolutions: (1) individual grid cells, (2) US CONUS, and (3) 4 US regions (north, east, south, and west) used by Naz *et al* (2016) based on grouped 2 digit USGS hydrologic units (HUC2). The variability within the regional values for WaterWatch and ALL were compared in order to test the usability of the MsTMIP model ensemble. The model ensemble mean for ALL

Table 1. Ensembles of MsTMIP simulations used.

Experiment	Forcing	Ensemble size	MsTMIP models used
ALL	Historical Transient	6	CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6
CLMT	Climate Change	6	CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6
CO2	CO_2 Concentration	6	CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6
NDEP	Nitrogen Deposition Rate	3	CLM4, CLM4VIC, TEM6
LULCC	Land Use and Land Cover Change	5	CLM4, CLM4VIC, ISAM, LPJ-wsl, TEM6

for each region and season was able to reproduce the variability within the observations relatively well with the minimum and maximum R-squared value being 0.739 and 0.921, respectively. All of the R-squared values are reported in table S2.

2.2. Trend and D&A technique

Spatial patterns of the trend were estimated using the Theil-Sen estimator, and significance at the $\alpha = 0.05$ level was determined using Mann-Kendall's nonparametric test for a monotonic trend (Kendall 1975, Mann 1945, Sen 1968, Theil 1950). For the spatial trends, the dominant forcing for each grid cell was found. In this case, dominant forcing refers to the forcing trend that has the same sign as the trend from the ALL forcing and largest magnitude.

To investigate the contribution of various forcings to the observed trends we adapted the standard D&A methodology to the study of land surface only (instead of the entire climate system, classically). Within this framework, atmospheric boundary conditions are treated as one single forcing called 'climate'. This forcing contains both natural internal climate variability and climate change. The regression equation used here is of the form

$$y = \beta_{\mathrm{CLMT}} x_{\mathrm{CLMT}} + \beta_{\mathrm{CO2}} x_{\mathrm{CO2}} + \beta_{\mathrm{NDEP}} x_{\mathrm{NDEP}} + \beta_{\mathrm{LULCC}} x_{\mathrm{LULCC}} + \varepsilon, \quad (1)$$

where y are the WaterWatch observations, β_i is the scaling factor for forcing i, x_i are the model ensemble mean response to forcing i, and ε are the residuals. Data were centered by their means before performing the regression analysis. In D&A analysis, a forcing with a positive scaling factor and corresponding confidence interval which does not encompass zero is detected, meaning that the response to the considered forcing is significantly found in the observations. If a forcing is detected, it can be attributed if the scaling factor confidence interval includes one, meaning that the response found in the observations is consistent with the simulated one (Bindoff *et al* 2013).

In addition to adapting D&A to solely investigate the land surface, another difference comes from the treatment of internal variability in observations (y). The MsTMIP models, and uncoupled LSMs in general, do not offer preindustrial control simulations which D&A methodology relies heavily upon for estimating natural internal variability. This work uses the same method as in Gedney *et al* (2014). They used ordinary least squares (OLS) regression for estimating the scaling factors and then checked that the residuals were independent (not autocorrelated). The residuals ε therefore represent a model error rather than any kind of internal variability (at least in current LSMs, there is no internal variability in the land surface given atmospheric forcings). Residuals were defined to be significantly autocorrelated if the lag-one sample autocorrelation was outside the bounds of white noise or if multiple lags were outside the bounds.

3. Results

3.1. Trends

The simulated sign of runoff trends from ALL represents the WaterWatch spatial pattern of the trends relatively well, although the magnitudes are not as strong (figures 1(*a*)−(*b*)). Other than small differences in magnitude, the trends from ALL and CLMT forcings have a similar spatial pattern including the magnitude. It thus can be hypothesized that the CLMT forcing may be the leading driver of the observational runoff pattern. However, it is difficult to make any hypotheses about the CO2, NDEP, and LULCC forcings from the spatial trend plots (figures 1(*d*)−(*f*)). Most of the trends for CO2 and NDEP are significant, but this is due to the low variability found in CO2 and NDEP times series, where the variability related to the CLMT forcing is removed. It should be noted that these trends have smaller magnitudes (e.g. in comparison to the observations, ALL, and CLMT), but the variability is even smaller, which makes the signal-to-noise ratio relatively large (leading to statistical significance). The majority of the trend values for CO2 and NDEP remain significant even when using 5 year means which were pre-whitened using the method from Zhang *et al* (2000) (figure S2). The dominant forcing plot also shows the CLMT forcing agrees with the ALL forcing over a large area within the US (figure 1(*g*)). When the CLMT forcing is not considered, CO2 shows a prominence in the eastern region of the US, where the densely vegetated area dominates. Given the large amount of vegetation, this increased runoff could possibly be due to CO_2 induced stomatal closure. There is also a large region in figure 1(*h*) where LULCC determines the increasing trends of runoff. This region is mostly isomorphic to the region with increased historical cropland area shown in the Synergetic Land Cover Product (SYNMAP) vegetation type figure (figure S3(*b*)) (Jung *et al* 2006). The supplementary material includes figures for the season which dominates the trend of the annual total values (figure S4) and the spatial patterns of trends and dominant forcings for each season (figures S5–8).

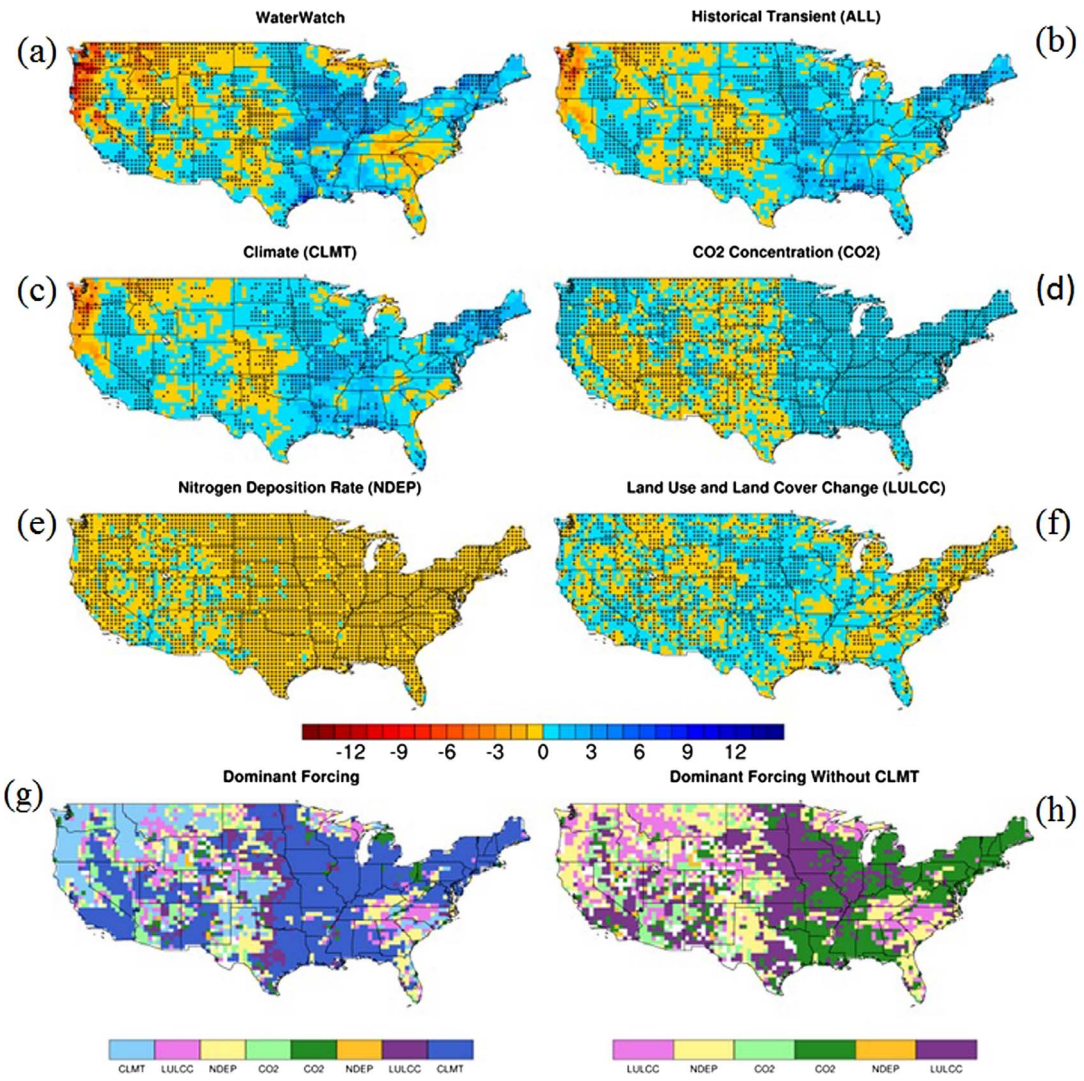

Figure 1. Spatial pattern of trends for annual total runoff from WaterWatch and MsTMIP multi-model ensemble mean (MME) forcings and dominant forcings. (*a*) WaterWatch and (*b*)–(*f*) MsTMIP MME annual total runoff trends for 1950–2010, mm yr^{-2} with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). (*g*) Dominant forcing when the trend values of CLMT, CO2, NDEP, and LULCC are compared to ALL. Lighter (darker) colors represent negative (positive) trends whereas white grid cells show spaces where the sign of the trends for the forcings disagreed with the ALL forcing. (*h*) Same as (*g*), but CLMT is not included.

The division of the CONUS HUC2 basins (R01–R18) into 4 regions (north, east, south, and west) is shown in the background of figure 2(*a*). For observational annual total runoff for the CONUS over 1950–2010, there is an estimated positive trend of approximately 0.2 mm yr^{-2} (insignificant for $\alpha = 0.05$). The estimated trend for the eastern region is twice that at approximately 0.4 mm yr^{-2} (insignificant for $\alpha = 0.05$). For the western region, however, the observational trend is −0.9 mm yr^{-2} (significant for $\alpha = 0.05$). Just as in the spatial trend plots, the ALL and CLMT forcings have the same sign and relatively close magnitude in every region. In each region, there is at least one forcing which has a sign opposite of the observations (figure 3). If we exclude the CLMT forcing for the eastern region in figure 2(*a*), CO2 is the only other positive forcing. This explains why there is a large area dominated by CO2 in figure 1(*h*). Overall, the seasonal trends in figures 2(*b*)–(*e*) have the same sign as the annual total trends. However, the observed negative MAM trend in the eastern region is not reproduced by the models. Combined with the large negative trend in the western region, this causes the observational trend for the US to be negative for MAM. The relative discrepancies between the annual total trend in the observations and the MsTMIP ALL MME forcing mostly come from the MAM and JJA seasons. The areas of the western region covering R17 and R18 include areas of large disagreement for all seasons. This is not surprising due to the regions sensitivity to precipitation and considerable amount of human water regulation (i.e. dams and irrigation). The largest difference for the eastern and southern regions is during MAM whereas it is JJA for the northern region. This can be seen in the normalized root mean squared difference (RMSD) values shown in figures S9(*a*)–(*e*). Possible causes for the discrepancy between the estimated trends are in the Discussion. Annual total regional

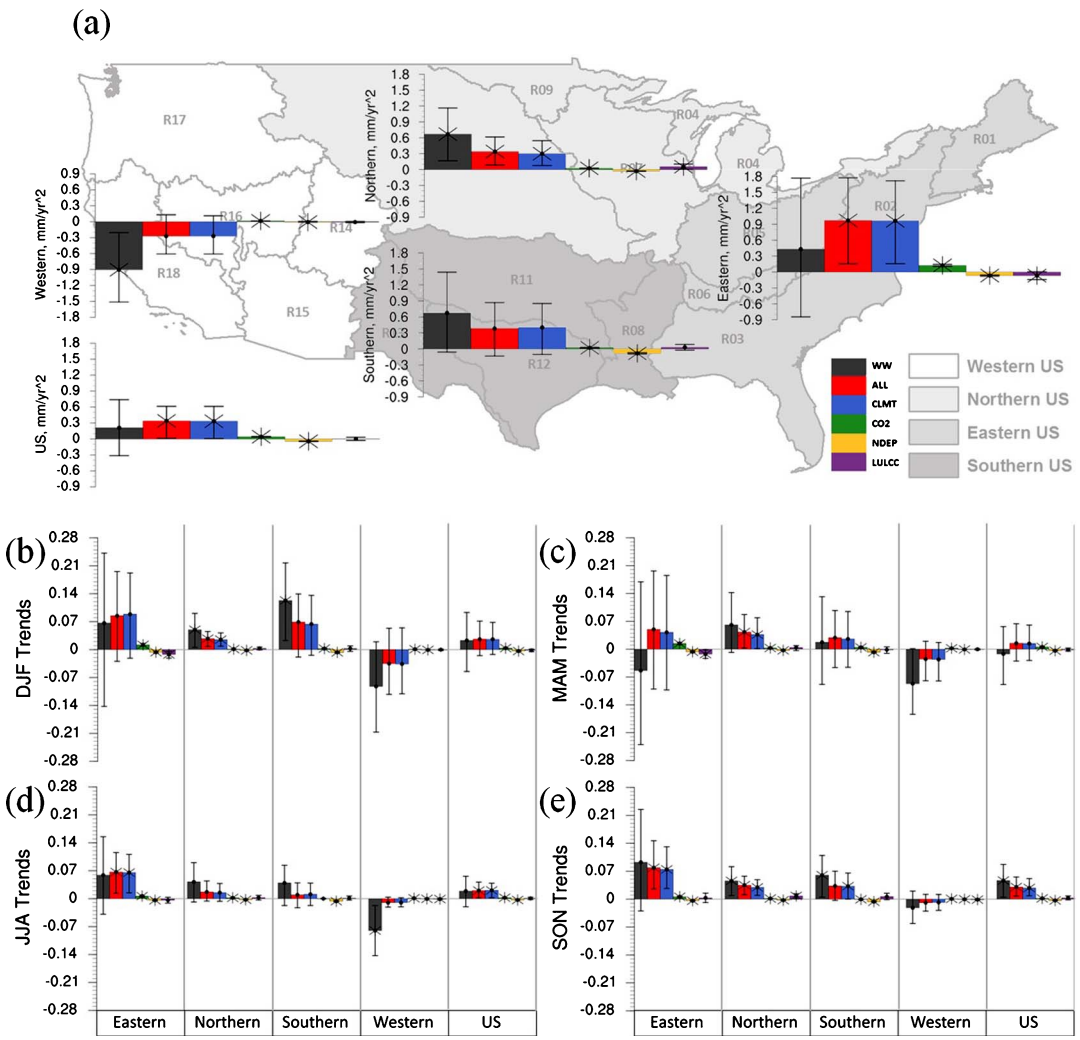

Figure 2. Regional trends for annual totals (*a*) and seasonal means (*b*)−(*e*). Trend values and corresponding 95% confidence intervals were estimated using Theil-Sen (Burkey 2006). Significant trends are denoted by asterisks. Significance is also denoted by error bars representing the 95% confidence interval for the Theil-Sen trend estimate (Burkey 2006). (*b*)−(*e*) Seasonal trends, (mm/month)/year, are grouped by region: eastern, northern, southern, western, and the US.

trend values for individual models are shown in the supplementary material (figure S10).

3.2. D&A results

D&A analysis was completed for each individual grid cell and each region for each season and the annual totals. Results from individual grid cells can be used to form hypotheses about the regional results. In figure 1(*g*), there are large regions in the east and west where CLMT is the dominant forcing. These same regions can be seen in figure 3(*a*). While CO2, NDEP, and LULCC do not show any areas of detectable and attributable cells large enough to make hypotheses about, LULCC does show multiple groupings of localized detection and attribution. It should also be noted that in some of the areas where CLMT is not detectable, CO2, NDEP, and/or LULCC can be detected and/or attributed, but this does not occur at a rate greater than that expected by chance (i.e. 5% of cases). The seasonal results from the D&A analysis using individual grid cells is in the supplementary material (figures S11(*a*)−(*d*)).

Scaling factor values are shown in figures 4(*a*)−(*e*) for all seasons/regions, but due to residuals failing the autocorrelation test (independence), results for northern annual totals, JJA, and SON; southern MAM; and US JJA are inconclusive. The autocorrelation plots are shown in the supplementary material (figures S12(*a*)−(*e*)). CLMT can be detected for all cases. The scaling factors consistently being greater than one implies that the multi-model mean underestimated the response to the CLMT forcing. This underestimation will be discussed more in section 4. Only in a few cases (southern JJA and SON and western SON) scaling factors are consistent with unity and we can attribute part of the observed changes to CLMT. Results for the other forcings are not quite as cohesive. This is due to the signal-to-noise ratio being low for CO2, NDEP, and LULCC. In a limited number of cases, forcings other than CLMT are detected, but then the estimated scaling factors take very large values (e.g. some confidence intervals are entirely outside the range of values considered), raising questions about physical realism.

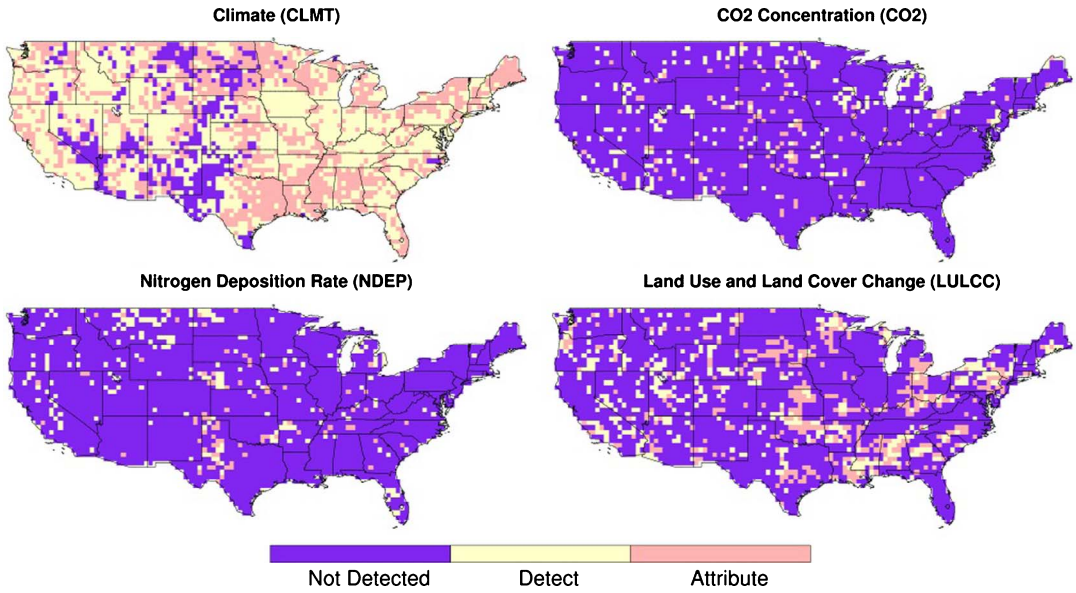

Figure 3. Spatial distributions of D&A scaling factors. The D&A methodology was applied to each grid cell. Not detected (purple) denotes a scaling factor whose corresponding 95% confidence interval was less than zero or included zero. If the 95% confidence interval was greater than zero but did not include one, the forcing was detected (yellow). A positive confidence interval was labeled as attributed (pink) if it included one.

Results using only the three models with simulations for all of the forcings (CLM4, CLM4VIC, and TEM6) are shown in figure S13. Using only these three models leads to the same overall conclusions.

4. Discussion

Our results for the annual US runoff observations in figure 2(*a*) show the same spatio-temporal pattern as Groisman *et al* (2004), positive in the east, north, south, of the US, whereas it is negative in the west. The west is already suffering from dry conditions which have led to numerous forest fires and water shortages (Dennison *et al* 2014, Diffenbaugh *et al* 2015). Continued drying will have more ecological effects along with effects to the western hydropower system. Groisman *et al* (2004) also found an increase in heavy and very heavy precipitation in the east. A general increase in wetting combined with an increase in heavy and very heavy precipitation will likely lead to more frequent flooding in the east.

After comparing observations with LSM simulated streamflow, Dai *et al* (2009) determined yearly streamflow for the world's largest rivers was more heavily impacted by climatic conditions than other environmental influences. While precipitation was not studied independently in this study, a detectable change in runoff due to climate has also been found. Contrary to our results, using a single LSM, Gedney *et al* (2006) found a direct CO_2 effect on continental (i.e. Africa, Asia, Europe, North America, South America) river runoff, but Dai *et al* (2009) determined that the results were model and data dependent. In 2014, Gedney and coauthors published another study focusing on the Northern Hemisphere where they addressed the concerns from the previous study. In the updated study, they again found significant effects from stomatal closure due to elevated CO_2 concentration. They were also able to detect solar dimming effects caused by aerosols. However, just as the results within this paper, for the basins they studied within the US (i.e. Mississippi, Hudson, and Neches basins), they were not able to detect CO_2 physiological effects and the scaling factor estimates had wide confidence intervals. Instead, climate was detected but overestimated. Further, land use effects were detected for the Neches basin in Gedney *et al* (2014). We also detected the effects of climate, but rather than being overestimated, our results consistently showed the MsTMIP climate simulations underestimated the trend and amplitude of runoff. In comparison to the Neches basin, we were also able to detect LULCC for annual totals in the southern region. Krakauer and Fung (2008) found that the effects of increasing temperature and CO_2 induced stomatal closure oppose each other in the CONUS and therefore cancel each other out. This provides a plausible explanation for the overall weak signal found in the CO2 forcing.

Improving from previous studies which only used observational data, one LSM, or focused on detection (Groisman *et al* 2004, Petersen *et al* 2012, and Alkama *et al* 2013, Gedney *et al* 2014), we used single-factor LSM simulations to conduct detailed D&A analysis in order to address the causality of changes in US runoff. We quantified the changes in runoff due to CLMT, CO2, NDEP, and LULCC using simulations from multiple LSMs by applying an adapted version of the classical regression-based methodology for D&A. In comparison to previous studies which must first

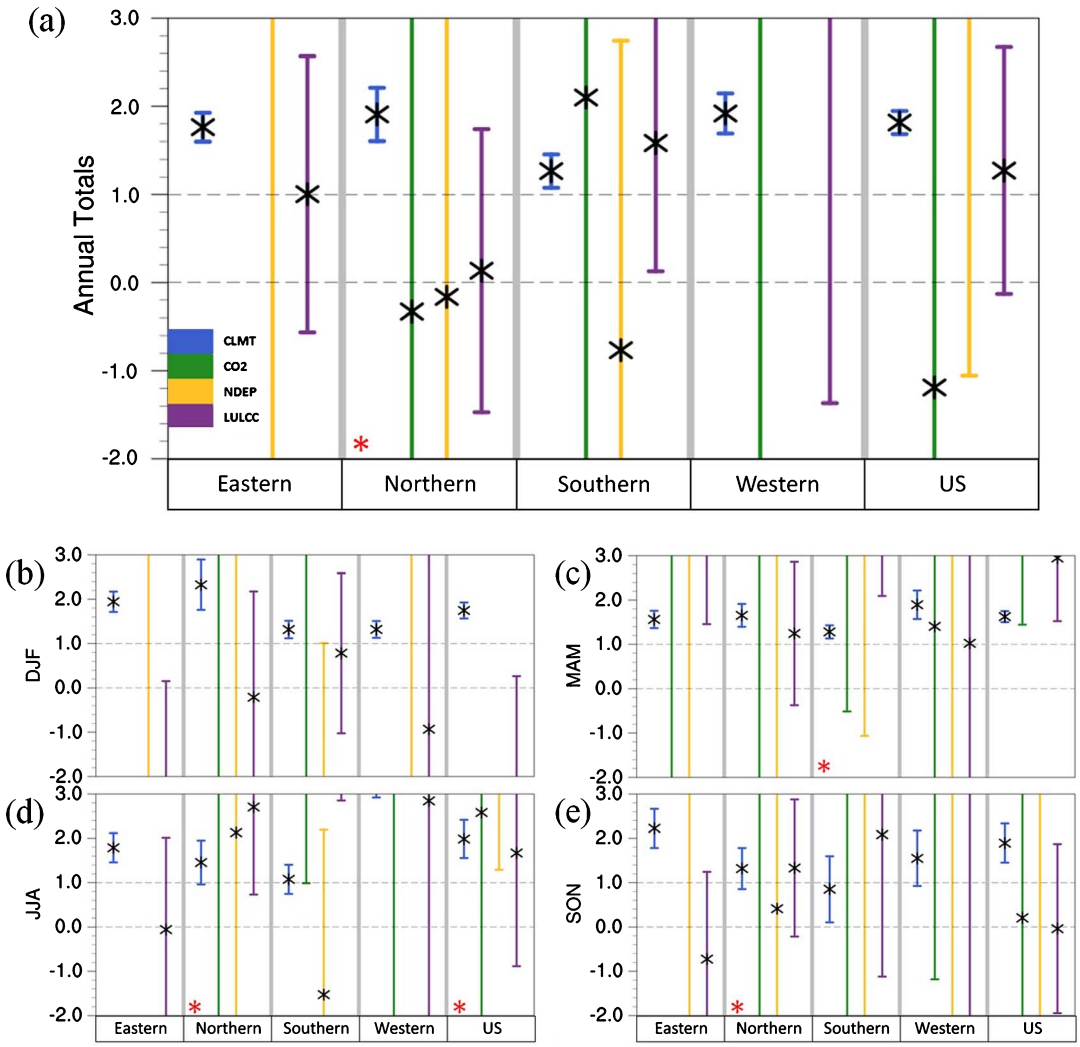

Figure 4. Scaling factor estimates (black asterisks) and corresponding 95% confidence intervals for annual totals (*a*) and seasonal means (*b*)−(*e*). Dashed lines denote the values 0 and 1. Thick gray lines separate the results into different regions. A red asterisk in the bottom left corner for a region denotes where the residuals were autocorrelated.

route the gridded model simulated flow in order to be comparable with station-based observations, we used the gridded WaterWatch observational dataset. This provided a more direct comparison to the gridded LSM simulations. The gridded observations also gave us the capability to study a broader extent of the US spatially in comparison to station-based studies which are linked to a subset of individual watersheds. The combination of these three attributes (i.e. gridded WaterWatch, LSM simulations, and D&A) provided us with the ability to conduct a more comprehensive study of runoff changes and their drivers for the CONUS.

Results from this study are mostly limited by two factors: the precipitation driver data used by the MsTMIP LSMs and human regulation within the WaterWatch observations. For the US spatial distribution, the largest RMSD values between WaterWatch and the MsTMIP MME ALL forcing annual totals in figure S9(*a*) are the areas in which Fekete *et al* (2004) found runoff simulations from water balance models to be the most sensitive to uncertainties in precipitation driver data. If the observations are regressed against the ALL ensemble mean, the average scaling factor (and min/max 95% confidence interval) for each region over all of the temporal metrics is $\beta_{east} = 1.53\ (1.26, 1.77)$, $\beta_{north} = 1.72\ (1.33, 2.22)$, $\beta_{south} = 1.39\ (1.04, 1.86)$, $\beta_{west} = 2.21\ (1.22, 4.89)$, and $\beta_{US} = 1.58\ (1.36, 1.98)$. This underestimation, indicated by the fact that βs are greater than 1 for all regions, is partially derived from the CRU-NCEP precipitation driver data being seemingly too dry. While the pattern of observational runoff is attributed to climate in more regions for more temporal metrics when using five year means (figures S14(*a*)−(*e*)), the response to this forcing overall is still underestimated. Part of this dryness is also shown by comparing the annual total time series and trends for 1950–2010 precipitation between CRU-NCEP and Parameter-elevation Regressions on Independent Slopes Model (PRISM) in

figure S15 (Daly *et al* 2008). PRISM was chosen due to its wide use in a variety of hydrologic studies as a baseline precipitation product for model evaluation and verification (e.g. Ashfaq *et al* 2016, Prat and Nelson 2015, Oubeidillah *et al* 2014, Widmann and Bretherton 2000). It is a gridded precipitation product which combines surface observations with a digital elevation model to account for the orographic enhancement of precipitation. Since PRISM does not incorporate assimilated information from numerical weather forecasting model or meteorologic reanalysis, it can usually result in better hydrologic modeling performance during calibration and validation (e.g. Radcliffe and Mukundan 2017). The inclusion of water management within WaterWatch is a limitation of this study given that it is not included in the MsTMIP models. However, Tavakoly *et al* (2016) found that even without the influence of water management, modeled river discharge at the continental scale was reasonably well reproduced. Individual region and season values are still vulnerable to biases due to the inclusion of water management within WaterWatch though. Tavakoly *et al* (2016) also showed that for the Mississippi River Basin, modeled flow was overestimated when not considering dams, lakes, and reservoirs. Modeled flow can also be overestimated in areas with significant amounts of human-managed land (e.g. cropland) by underestimating evapotranspiration due to overestimating sensible heat flux and underestimating latent heat flux and net ecosystem exchange when crop-specific parameterization is limited (Lokupitiya *et al* 2016). This implies that biases due to MsTMIP not including human management should lead to overestimation in ALL. Given that we found ALL to be underestimated, it would be more underestimated if human management was included. Thus, giving more support that the CRU-NCEP precipitation being too dry is driving the underestimation found in the MsTMIP ALL ensemble mean.

5. Conclusions

Annual runoff observations for the period 1950–2010 had heterogenous patterns of change regionally in the US. The eastern two-thirds of the US (USGS HUC2 R01–R13) has seen significant and insignificant increases in annual runoff while the western one-third (USGS HUC2 R14–R18) had a greater significant decrease. This heterogeneity lead to an insignificant increase for the US as a whole. Seasonally, autumn runoff significantly increased for the northern and southern regions and the US as a whole. Northern and southern runoff also significantly increased for the winter season. For the west, there was a significant decrease in summer runoff. The LSM simulations showed that the CLMT trend and time series were approximately equal to that of the ALL forcing. This consistency hypothesized a strong relationship between runoff and climate change, especially the precipitation variation. More formally, using D&A analysis, changes in observational runoff were detected in CLMT for all of the seasons and regions studied. ALL and CLMT were also both consistently underestimated, possibly due to uncertainties in the CRU-NCEP precipitation driver used by MsTMIP, leading to the changes in the observations only being detected in CLMT rather than detected in and attributed to CLMT in most cases. While the changes in observational runoff could be detected in and attributed to CO2, NDEP, and LULCC for certain cases, results were not consistent enough regionally and seasonally to draw any major conclusions.

The western US is at the greatest risk for water scarcity. Water availability in the region has already decreased, and shows signs of continued decreasing. Given that the northwestern US is a semi-arid region, it is very sensitive to uncertainties in precipitation. It is also the region that showed the largest disagreement between WaterWatch and the MsTMIP ALL forcing. For future work we plan to perform a comparison using a river basin which has a naturalized streamflow dataset. New higher resolution simulations using multiple pairings of environmental driver datasets will be used to test sensitivity to the precipitation driver. The most appropriate simulations will then be used to complete a D&A study for that river basin.

Acknowledgments

This work is supported by the Terrestrial Ecosystem Science Scientific Focus Area (TES SFA) project funded through the Terrestrial Ecosystem Science Program, partially supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computing Scientific Focus Area (RUBISCO SFA) project funded through the Regional and Global Climate Modeling Program, and partially supported by the Energy Exascale Earth System Model (E3SM) project funded through the Earth System Modeling Program, in the Climate and Environmental Sciences Division (CESD) of the Biological and Environmental Research (BER) Program in the US Department of Energy Office of Science. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC05-00OR22725. JBF contributed to this from the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. JBF was supported in part by NASA SUSMAP, INCA, IDS, and CARBON programs. Atul K Jain is funded by the US National Science Foundation (NSF-AGS-12-43071).

ORCID iDs

Jiafu Mao https://orcid.org/0000-0002-2050-7373
Shih-Chieh Kao https://orcid.org/0000-0002-3207-5328
Joshua B Fisher https://orcid.org/0000-0003-4734-9085
Akihiko Ito https://orcid.org/0000-0001-5265-0791

References

Alkama R, Ribes A, Decharme B and Marchand L 2013 Detection of global runoff changes: results from observations and CMIP5 experiments *Hydrol. Earth Syst. Sci.* **17** 2967–79

Anderegg W R L, Berry J A, Smith D D, Sperry J S, Anderegg L D L and Field C B 2012 The roles of hydraulic and carbon stress in a widespread climate-induced forest die-off *Proc. Natl Acad. Sci. USA* **109** 233–7

Ashfaq M, Rastogi D, Mei R, Kao S-C, Gangrade S, Naz B S and Touma D 2016 High-resolution ensemble projections of near-term regional climate over the continental United States *J. Geophys. Res.-Atmos.* **121** 9943–63

Barnett T P, Adam J C and Lettenmaier D P 2005 Potential impacts of a warming climate on water availability in snow-dominated regions *Nature* **438** 303–9

Barnett T P *et al* 2008 Human-induced changes in the hydrology of the Western United States *Science* **319** 1080–3

Beigi E and Tsai F T-C 2014 GIS-based water budget framework for high-resolution groundwater recharge estimation of large-scale humid regions *J. Hydrol. Eng.* **19** 05014004

Betts R A *et al* 2007 Projected increase in continental runoff due to plant responses to increasing carbon dioxide *Nature* **448** 1037–U5

Bindoff N L *et al* 2013 Detection and attribution of climate change: from global to regional *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker, D Qin, G K Plattner, M Tignor, S K Allen, J Boschung, A Nauels, Y Xia, V Bex and P M Midgley (Cambridge: Cambridge University Press) pp 867–952

Bosch J M and Hewlett J D 1982 A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration *J. Hydrol.* **55** 3–23

Brakebill J W, Wolock D M and Terziotti S E 2011 Digital hydrologic networks supporting applications related to spatially referenced regression modeling *J. Am. Water Resour. Assoc.* **47** 916–32

Burkey J 2006 A non-parametric monotonic trend test computing Mann-Kendall Tau, Tau-b, and Sens Slope written in Mathworks-MATLAB implemented using matrix rotations *King County, Department of Natural Resources and Parks, Science and Technical Services Section* (Seattle, WA) (www.mathworks.com/matlabcentral/fileexchange/authors/23983)

Collins M J 2008 Evidence for changing flood risk in New England since the late 20th century *J. Am. Water Resour. Assoc.* **45** 279–90

Dai A, Qian T, Trenberth K E and Milliman J D 2009 Changes in continental freshwater discharge from 1948–2004 *J. Clim.* **22** 2773–92

Daly C, Halbleib M, Smith J I, Gibson W P, Doggett M K, Taylor G H, Curtis J and Pasteris P P 2008 Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States *Int. J. Clim.* **28** 2031–64

Dennison P E, Brewer S C, Arnold J D and Moritz M A 2014 Large wildfire trends in the western United States 1984–2011 *Geophys. Res. Lett.* **41** 2928–33

Diffenbaugh N S, Swain D L and Touma D 2015 Anthropogenic warming has increased drought risk in California *Proc. Natl Acad. Sci.* **112** 3931–6

Fan Y, Miguez-Macho G, Jobbágy E G, Jackson R B and Otero-Casal C 2017 Hydrologic regulation of plant rooting depth *Proc. Natl Acad. Sci.* **114** 10572–7

Fekete B M, Vörösmarty C J, Roads J O and Willmott C J 2004 Uncertainties in precipitation and their impacts on runoff estimates *J. Clim.* **17** 294–304

Gedney N, Cox P M, Betts R A, Boucher O, Huntingford C and Stott P A 2006 Detection of a direct carbon dioxide effect in continental river runoff records *Nature* **439** 835–8

Gedney N, Huntingford C, Weedon G P, Bellouin N, Boucher O and Cox P M 2014 Detection of solar dimming and brightening effects on Northern Hemisphere river flow *Nat. Geosci.* **7** 796–800

Gerten D, Schaphoff S, Haberlandt U, Lucht W and Sitch S 2004 Terrestrial vegetation and water balance–hydrological evaluation of a dynamic global vegetation model *J. Hydrol.* **286** 249–70

Groisman P Y, Knight R W, Karl T R, Easterling D R, Sun B and Lawrimore J H 2004 Contemporary changes of the hydrological cycle over the contiguous United States: trends derived from *in situ* observations *J. Hydrometeor.* **5** 64–85

Hidalgo H G *et al* 2009 Detection and attribution of streamflow timing changes to climate change in the Western United States *J. Clim.* **22** 3838–55

Huntzinger D N *et al* 2013 The North American Carbon Program (NACP) Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP): Part I—overview and experimental design *Geosci. Model Dev.* **6** 2121–33

Jung M, Henkel K, Herold M and Churkina G 2006 Exploiting synergies of global land cover products for carbon cycle modeling *Remote Sens. Environ.* **101** 534–53

Kendall M G 1975 *Rank Correlation Methods* (London: Griffin)

Kirshen P, Knee K and Ruth M 2008 Climate change and coastal flooding in Metro Boston: impacts and adaptation strategies *Clim. Change* **90** 453–73

Koster R D *et al* 2017 Hydroclimatic variability and predictability: a survey of recent research *Hydrol. Earth Syst. Sci.* **21** 3777–98

Krakauer N Y and Fung I 2008 Mapping and attribution of change in streamflow in the coterminous United States *Hydrol. Earth Syst. Sci.* **12** 1111–20

Lettenmaier D P, Wood E F and Wallis J R 1993 Hydro-climatological trends in the continental United States 1948–88 *J. Clim.* **7** 586–607

Lokupitiya E *et al* 2016 Carbon and energy fluxes in cropland ecosystems: a model-data comparison *Biogeochemistry* **129** 53–76

Mann H B 1945 Nonparametric tests against trend *Econometrica* **13** 245–59

Mao J *et al* 2015 Disentangling climatic and anthropogenic controls on global terrestrial evapotranspiration trends *Environ. Res. Lett.* **10** 094008

Mao J *et al* 2016 Human-induced greening of the northern extratropical land surface *Nat. Clim. Change* **6** 959–63

Naz B S, Kao S-C, Ashfaq M, Rastogi D, Mei R and Bowling L C 2016 Regional hydrologic response to climate change in the conterminous United States using high-resolution hydroclimate simulations *Glob. Planet. Change* **143** 100–17

Nepstad D C, de Carvalho C R, Davidson E A, Jipp P H, Lefebvre P A, Negreiros G H, da Silva E D, Stone T A, Trumbore S E and Vieira S 1994 The role of deep roots in the hydrological and carbon cycles of the Amazonian forests and pastures *Nature* **372** 666–9

Nicholls R J, Hanson S, Herweijer C, Patmore N, Hallegatte S, Corfee-Morlot J, Château J and Muir-Wood R 2008 *Ranking Port Cities with High Exposure and Vulnerability to Climate Extremes OECD Environment Working papers, No. 1* (Paris: Organisation for Economic Co-operation and Development Publishing) p 62

Oubeidillah A A, Kao S-C, Ashfaq M, Naz B S and Tootle G 2014 A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US *Hydrol. Earth Syst. Sci.* **18** 67–84

Petersen T, Devineni N and Sankarasubramanian A 2012 Seasonality of monthly runoff over the continental United States: causality and relations to mean annual and mean monthly distributions of moisture and energy *J. Hydrol.* **468** 139–50

Piao S, Friedlingstein P, Ciais P, deNoblet-Ducoudré N, Labat D and Zaehle S 2007 Changes in climate and land use have a larger direct impact than rising CO_2 on global river runoff trends *Proc. Natl Acad. Sci. USA* **104** 15242–7

Pierce D W *et al* 2008 Attribution of declining Western US snowpack to human effects *J. Clim.* **21** 6425–44

Prat O P and Nelson B R 2015 Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge data sets at daily to annual scales (2002–2012) *Hydrol. Earth Syst. Sci.* **19** 2037–56

Radcliffe D E and Mukundan R 2017 PRISM vs. CFSR precipitation data effects on calibration and validation of SWAT models *J. Am. Water Resour. Assoc.* **53** 89–100

Richardson G R A 2010 *Adapting to Climate Change: An Introduction for Canadian Municipalities* (Ottawa: Natural Resources Canada) p 40

Romero-Lankao P, Smith J B, Davidson D J, Diffenbaugh N S, Kinney P L, Kirshen P, Kovacs P and Villers Ruiz L 2014 *North America. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed V R Barros *et al* (Cambridge: Cambridge University Press) pp 1439–98

Scanlon B R, Faunt C C, Longuevergne L, Reedy R C, Alley W M, McGuire V L and McMahon P B 2012 Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley *Proc. Natl Acad. Sci.* **109** 9320–25

Scanlon B R, Duncan I and Reedy R C 2013 Drought and the water-energy nexus in Texas *Environ. Res. Lett.* **8** 045033

Schwalm C R 2015 How well do terrestrial biosphere models simulate coarse-scale runoff in the contiguous United States? *Ecol. Model.* **303** 87–96

Shi X, Mao J, Thornton P E, Hoffman F M and Post W M 2011 The impact of climate, CO_2, nitrogen deposition and land use change on simulated contemporary global river flow *Geophys. Res. Lett.* **38** L08704

Sen P K 1968 Estimates of the regression coefficient based on Kendall's Tau *J. Am. Stat. Assoc.* **63** 1379–89

Shi X, Mao J, Thornton P E and Huang M 2013 Spatiotemporal patterns of evapotranspiration in response to multiple environmental factors simulated by the community land model *Environ. Res. Lett.* **8** 024012

Singh R, Wagener T, Crane R, Mann M E and Ning L 2014 A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: application to a watershed in Pennsylvania, USA *Water Resour. Res.* **50** 3409–27

Tavakoly A A, Snow A D, David C H, Follum M L, Maidment D R and Yang Z L 2016 Continental-scale river flow modeling of the Mississippi River Basin using high-resolution NHDPlus dataset *J. Am. Water Resour. Assoc.* **53** 258–79

Theil H 1950 A rank-invariant method of linear and polynomial regression analysis, 3; confidence regions for the parameters of polynomial regression equations *Proceedings KNAW*

Thornton P E, Lamarque J-F, Rosenbloom N A and MahowaldN M 2007 Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability *Glob. Biogeochem. Cycles* **21** GB4018

Wei Y *et al* 2014 The North American carbon program multi-scale synthesis and terrestrial model intercomparison project: Part 2—environmental driver data *Geosci. Model Dev.* **7** 2875–93

Weiss J L, Overpeck J T and Strauss B 2011 Implications of recent sea level rise science for low-elevation areas in coastal cities of the conterminous USA *Clim. Change* **105** 635–45

Wenger S J, Luce C H, Hamlet A F, Isaak D J and Neville H M 2010 Macroscale hydrologic modeling of ecologically relevant flow metrics *Water Resour. Res.* **46** W09513

Westerling A L, Bryant B P, Preisler H K, Holmes T P, Hidalgo H G, Das T and Shrestha S R 2011 Climate change and growth scenarios for California wildfire *Clim. Change* **109** S445–63

Widmann M and Bretherton C S 2000 Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States *J. Clim.* **13** 1936–50

Zhang X, Vincent L A, Hogg W and Niitsoo A 2000 Temperature and precipitation trends in Canada during the 20th century *Atmos.-Ocean* **38** 395–423

Zhu Z *et al* 2016 Greening of the Earth and its drivers *Nat. Clim. Change* **6** 791–5