

**Detection and Attribution of Climate Change Using Offline Model
Simulations with Applications to Runoff in the United States and
Streamflow in the Columbia River Basin**

**A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**Whitney Leeann Forbes
August 2018**

DEDICATION

I would like to dedicate this work to my beautiful girl Annabelle. May your tail always be wagging when I come home.

ACKNOWLEDGEMENTS

Thank you to all my family. You listened to all my complaining and offered me support when I needed it. Mom, Dave, Brandi, and Logan, you know how much I love you. Tony, you have definitely seen more than your fair share of tears and irritability these past 4 years, but I really appreciate the unwavering love you have shown me.

I would also like to thank my committee members and all the members of the Climate Change Science Institute at Oak Ridge National Laboratory. Dr. Jin, thank you for our frank relationship, helping to build my writing skills, and always promoting me to others. Dr. Mao, thank you for supporting my research in every way and always being understanding when I could not be present. Dr. Ricciuto and Dr. Shi, thank you for devoting so much of your time and climate modeling genius. Dr. Kao, thank you for introducing me to the fascinating subject of hydrology and for always being willing to discuss and brainstorm with me. Joe Kennedy and Anthony Walker, thank you for believing in me when I was losing hope.

ABSTRACT

Detection and attribution analysis of climate change is the processes of statistically detecting a change in a particular climate variable or variable affected by climate and then confidently attributing the change to effects from external forcings such as greenhouse gases, aerosols, and solar-volcanic. The variables studied here are annual and seasonal runoff in the contiguous United States and streamflow in the Columbia River Basin for the period 1950 – 2010 and 1950 – 2008, respectively. For forcings, the effects of climate change and variability, CO₂ [carbon dioxide] concentration, nitrogen deposition, and land use and land cover change are used in both studies. Monthly observations of runoff were provided by WaterWatch from the United States Geological Survey, and an ensemble of semi-factorial land surface model simulations were used to quantify the effects due to external forcings. The two limitations of the study conducted on runoff in the United States were: the inclusion of human regulation and irrigation withdrawals within the observations and not in the model simulations and a dry bias within the model simulations due to the precipitation driver. These limitations were overcome in the streamflow study for the Columbia River Basin due to the availability of a naturalized streamflow dataset and a new ensemble of semi-factorial land surface model simulations which were driven by less biased precipitation.

United States runoff had significant and insignificant increases in the east, north, and south, and a strong significant decrease in the west. These changes were detected in the effects of climate change and variability but could not be attributed due to the dry bias in the precipitation driver leading to underestimation in the model simulations. However, for the Columbia River Basin, the changes in annual total, center of timing of, and summer mean streamflow were attributed to climate change and variability. The most significant changes were the declines in the June – October months. On average, these months account for

approximately 49% of the annual total flow. More specifically, the greatest decline was 28% for June which comprised approximately 22% [22 percent] of the total annual flow.

TABLE OF CONTENTS

INTRODUCTION	1
CHAPTER I Contribution of Environmental Forcings to US Runoff Changes for the Period 1950 – 2010.....	6
Abstract.....	7
Introduction	8
Data and Methodology	12
Data and Data Processing.....	12
Trend and D&A Technique	16
Results	17
Trends	17
D&A Results	22
Discussion.....	25
Conclusions.....	29
References.....	32
Supplementary Material	44
CHAPTER II Streamflow in the Columbia River Basin: Quantifying changes over the period 1951 – 2008 and determining the drivers of those changes.....	68
Abstract.....	69
Introduction	70
Data.....	73
Observational Data	73
Climate Model Simulations	73
River Routing and Downstream Gauge Stations	75
Metrics and Methods of Analysis.....	79
Results	81
Monthly Climatologies.....	81
D&A Results	84
Discussion.....	88
Conclusions.....	90
References.....	92
Supplementary Material	98
CONCLUSION	109
VITA.....	113

LIST OF TABLES

Table 1. Ensembles of MsTMIP simulations used.....	15
Table 2. Meteorological Data Used by Each MsTMIP Model.	45
Table 3. R-Squared for Each Season Using WaterWatch and ALL. Results are for the model ensemble mean and the values shown in parenthesis correspond to the values for the individual models (CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6).	46
Table 4. Individual CRB Subbasins and Corresponding Downstream Stations.....	76
Table 5. Nash-Sutcliffe Efficiency between Monthly and Logarithmic Monthly Observations and ALL MME-Livneh for the Downstream Stations. VG=Very Good, G=Good, S=Satisfactory, US=Unsatisfactory	79

LIST OF FIGURES

Figure 1. Climate Model Schematic.	3
Figure 2. 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, CO ₂ , 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.	20
Figure 3. 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.	21
Figure 4. 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.	23
Figure 5. 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.	24
Figure 6. Comparison between WaterWatch stations and station from Dai's dataset (Dai et al., 2009). There are 59 stations in total.	44
Figure 7. Comparison between WaterWatch stations and station from Dai's dataset (Dai et al., 2009). There are 57 stations in total (59 minus the two high leverage points from figure S1(a)).	44
Figure 8. Spatial Patterns of Trends and Dominant Forcings using Pre-whitened 5-year Means from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings. (a-f) WaterWatch and MsTMIP MME pre-whitened 5-year mean runoff trends for 1950-2010, (mm/yr)/5years with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). The 'zyp' R package function based on Zhang (1999) was used to obtain the pre-whitened 5-year means and their corresponding Theil-Sen trends and Mann-Kendall Significance (Kendall, 1975; Mann, 1945; Sen, 1968; Theil,	

1950). (g) Dominant forcing when the trend values of CLMT, CO ₂ , 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.	48
Figure 9. SYNMAP Vegetation Type Trend using Theil-Sen (Sen, 1968; Theil, 1950). Trends are for the percent plant functional type of each cell for (a) trees and (b) crops.....	50
Figure 10. Dominant Season. The dominant season is the season that has the largest trend which has the same sign as the trend for ALL. White grid cells show locations where the sign of the trends for the seasons disagreed with the ALL forcing.....	51
Figure 11. Spatial Pattern of Trends for DJF Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME DJF 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, CO ₂ , 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.	52
Figure 12. Spatial Pattern of Trends for MAM Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME MAM 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, CO ₂ , 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.	53
Figure 13. Spatial Pattern of Trends for JJA Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME JJA 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, CO ₂ , 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.	54
Figure 14. Spatial Pattern of Trends for SON Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME SON 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,	

CO ₂ , 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.	55
Figure 15. Root Mean Square Difference (RMSD) Values between WaterWatch and ALL Annual Totals. The RMSD values were normalized by the RMSD value corresponding to approximately the 99 th percentile of values. The 99 th percentile was chosen rather than the maximum due to extreme values.	56
Figure 16. DJF RMSD. This figure is the same as figure S15 but instead of annual totals, the DJF season was used to determine the RMSD values....	56
Figure 17. MAM RMSD. This figure is the same as figure S15 but instead of annual totals, the MAM season was used to determine the RMSD values..	57
Figure 18. JJA RMSD. This figure is the same as figure S15 but instead of annual totals, the JJA season was used to determine the RMSD values. ...	57
Figure 19. SON RMSD. This figure is the same as figure S15 but instead of annual totals, the SON season was used to determine the RMSD values. .	58
Figure 20. Regional Trends of Annual Total Values for the M_{ST}MIP MME (black bars) and Each Individual Model. Each black bar denotes the beginning of a region. The regions are ordered as follows: east, north, south, west, and US as a whole.	59
Figure 21. Spatial Distributions of D&A Scaling Factors for DJF. 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.....	60
Figure 22. Same as figure S21 but for MAM.	60
Figure 23. Same as figure S21 but for JJA.	61
Figure 24. Same as figure S21 but for SON.	61
Figure 25. Autocorrelation Plot of the Residuals from the Regression using Northern Annual Totals. Blue lines represent the bounds for white noise.	62
Figure 26. Autocorrelation Plot of the Residuals from the Regression using Southern MAM.	62
Figure 27. Autocorrelation Plot of the Residuals from the Regression using Northern JJA.	63
Figure 28. Autocorrelation Plot of the Residuals from the Regression using US JJA.	63
Figure 29. Autocorrelation Plot of the Residuals from the Regression using Northern SON.	64
Figure 30. Scaling Factor Estimates (black asterisks) and Corresponding 95% Confidence Intervals for Annual Totals (a) and Seasonal Means (b-e) Using the MME of CLM4, CLM4VIC, and TEM6. 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.	65
Figure 31. Scaling factor estimates (asterisk) and corresponding 95% confidence intervals for 5-year means of the 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. Red asterisk in the bottom left corner for a region denotes where the residuals were autocorrelated. ..	66
Figure 32. CRU-NCEP vs PRISM Annual Precipitation for Each Region and the US CONUS Over the Period 1950 – 2010. Included within each figure is the 1950 – 2010 time series, mean value for the time period, Theil-Sen trend estimate, and Mann-Kendall Significance (Kendall, 1975; Mann, 1945; Sen, 1968; Theil, 1950). A significant trend is denoted by red dots on the trend line.	67
Figure 33. Locations of the Flowlines, Subbasins, and Farthest Downstream Stations Used with Topography.....	78
Figure 34. Monthly Climatologies at The Dalles for the Observations, ALL, and CLMT (Top) and CO2, NDEP, and LULCC (Bottom). Box plots (left y-axis) and trends (right y-axis) for the observations (black), ALL (blue), and CLMT (red) are shown in the top figure. The bottom figure shows CO2 (green), NDEP (orange), and LULCC (magenta). For the box plots, the horizontal lines of the box represent the lower quartile, median, and upper quartile. Lines extending from the box represent 1.5 times the interquartile range or the bottom and top 25% of the distribution. Data points outside 1.5 times the interquartile range are outliers and are shown by a “+” symbol. Trend values and corresponding 90% confidence intervals were estimated using Theil-Sen with significance determined by Mann-Kendall (Burkey, 2006).	82
Figure 35. Correlation-Based Detection Analysis for Annual Totals, Annual Maximums, Center of Timing, and Summer Means for ALL, CLMT, CO2, NDEP, and LULCC. Pearson correlation coefficients between the observations and ALL (blue), CLMT (red), CO2 (green), NDEP (orange), and LULCC (magenta) for each subbasin. Light gray lines denote the 2.5 th to 97.5 th percentiles of the correlations between the observations and the piControl simulations.	85
Figure 36. Scaling Factor Estimates and Corresponding 95% Confidence Intervals for Annual Totals, Annual Maximums, Center of Timing, and Summer Means using ALL and the Linear Combination of CLMT, CO2, NDEP, and LULCC. Scaling factors for CLMT, CO2, NDEP, and LULCC for each subbasin are shown on the left y-axis in red, green, orange, and magenta, respectively. Scaling factors for ALL (blue) are shown on the right y-axis in the same subbasin ordering. Light gray lines denote the values 0 and/or 1.	87
Figure 37. Monthly Climatologies at Lower Snake for the Observations, ALL, and CLMT (Top) and CO2, NDEP, and LULCC (Bottom). Box plots	

(left y-axis) and trends (right y-axis) for the observations (black), ALL (blue), and CLMT (red) are in the top figure. The bottom figure shows CO2 (green), NDEP (orange), and LULCC (magenta). For the box plots, the horizontal lines of the box represent the lower quartile, median, and upper quartile. Lines extending from the box represent 1.5 times the IQR. Data points outside 1.5 times the interquartile range are outliers and are shown by a “+” symbol. Trend values and corresponding 90% confidence intervals were estimated using Theil-Sen with significance determined by Mann-Kendall and denoted by asterisks.98

Figure 38. Same as figure 37 but for Middle Snake.	99
Figure 39. Same as figure 37 but for Upper Snake.	100
Figure 40. Same as figure 37 but for Mid Columbia.	101
Figure 41. Same as figure 37 but for Yakima.	102
Figure 42. Same as figure 37 but for Spokane.	103
Figure 43. Same as figure 37 but for Pend Oreille.	104
Figure 44. Same as figure 37 but for Kootenay.	105
Figure 45. Same as figure 37 but for Upper Columbia.	106
Figure 46. Regional Trends for Annual Totals, Annual Maximums, Center of Timing, and Seasonal Means. Trend values for the observations (black), ALL (blue), CLMT (red), CO2 (green), NDEP (orange), and LULCC (magenta). Trend values and corresponding 90% confidence intervals were estimated using Theil-Sen (Burkey, 2006). Significant trends are denoted by asterisks and were determined by Mann-Kendall.107	
Figure 47. Scaling Factor Estimates and Corresponding 95% Confidence Intervals for Annual Totals, Annual Maximums, Center of Timing, and Summer Means using the MME of ALL and the Linear Combination of the MMEs of CLMT, CO2, NDEP, and LULCC using the Simulations Produced Using CRUNCEP5, GSWP3, and Princeton. Scaling factors for CLMT, CO2, NDEP, and LULCC for each subbasin are shown on the left y-axis in red, green, orange, and magenta, respectively. Scaling factors for ALL (blue) are shown on the right y-axis in the same subbasin ordering. Light gray lines denote the values 0 and/or 1.108	

INTRODUCTION

As stated by the IPCC, in this paper, detection and attribution (D&A) take on the following definitions:

- Detection of change is defined as the process of demonstrating that climate or a system affected by climate has changed in some defined statistical sense without providing a reason for that change. An identified change is detected in observations if its likelihood of occurrence by chance due to internal variability alone is determined to be small.
- Attribution is defined as the process of evaluating the relative contributions of multiple casual factors to a change or event with an assignment of statistical confidence.

D&A analyses are usually completed using a form of optimal fingerprinting (OF) via hypothesis testing using ordinary least squares (OLS) or total least squares (TLS) regression. In this method, response patterns (i.e. fingerprints) from observational data is compared to that from model simulated factorial data via regression. For example, we may want to determine if decadal historical global mean near-surface air temperature changed due to the climate system's responses to anthropogenic forcings (ANT), such as greenhouse gases and aerosols, and natural forcings (NAT), such as solar irradiance and volcanic aerosols. In general, this takes the form $y = X\beta + \varepsilon_{IV}$ where the terms y , X , β , and ε_{IV} respectively represent the observational data, model simulated factorial data (e.g. ANT and NAT), scaling factors (regression coefficients), and natural internal variability. For TLS, rather than assuming the model simulated response patterns are perfectly known, noise due to the sampling uncertain from using a

finite number of realizations is also accounted for by using $x_i = x_i - \varepsilon_{i,su}$, where i denotes the forcings (e.g. ANT and NAT). While OLS and TLS are used the most, Error In Variables (EIV) is another well-known approach. EIV includes natural internal variability and sampling uncertainty while also considering noise from model error. The null hypothesis is that $\beta = 0$ implying that variations within the observations can be explained by the natural internal variability alone. Detection is then claimed if $\beta \neq 0$. Furthermore, attribution is claimed if the pattern in the model simulations (e.g. ANT or NAT) is found to be consistent with the pattern in the observations, $\beta = 1$. These D&A methodologies are generally applied using simulations from coupled global climate models (GCMs), which allow for feedbacks between the atmosphere, land, and ocean. Coupled GCMs are used because they usually include a piControl simulation. This simulation uses preindustrial values from 1850 in order to simulate a variable (e.g. near-surface air temperature) without the effects of anthropogenic influences. The piControl run is very important because it allows the researcher to estimate the natural internal variability of the variable, which can be essential in identifying the pattern fingerprints. Offline climate models use prescribed historical environmental and meteorological driver data to force a model for the atmosphere, land, or ocean. See figure 1 below for a simple schematic representation of the difference between coupled and offline models. When using offline (uncoupled) models, piControl runs do not exist because they are driven by historical observational environmental drivers and therefore, you cannot estimate the natural internal variability this way.

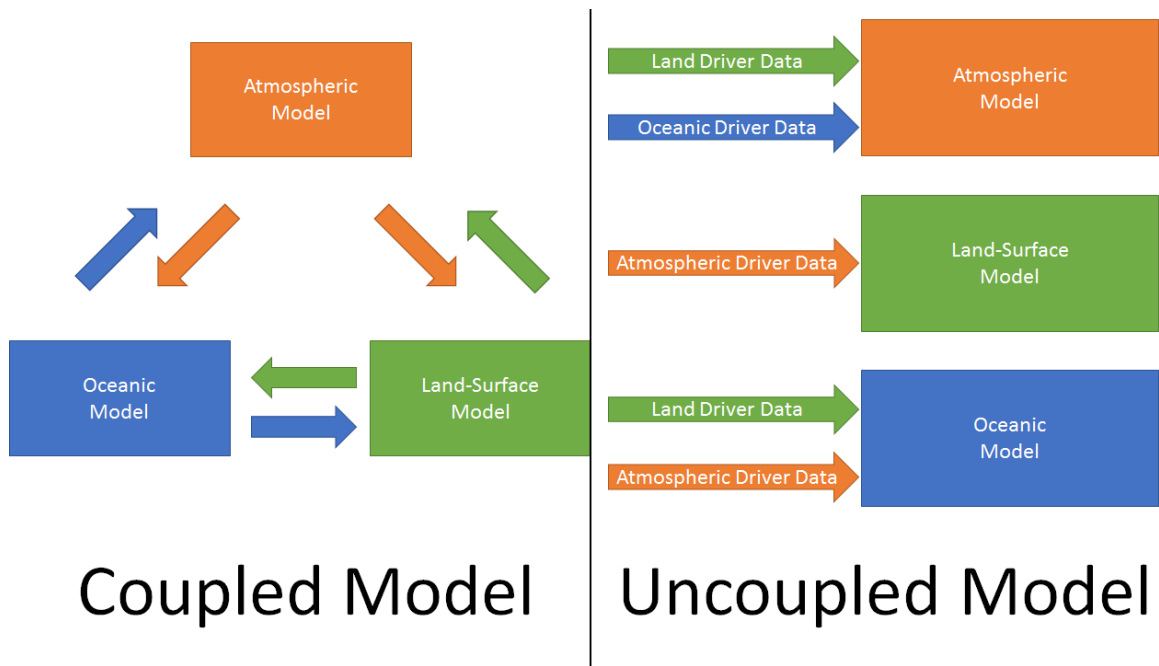


Figure 1. Climate Model Schematic.

Offline models have been used for comparison in previous studies, but due to not being able to apply the current D&A methodologies' use of the preindustrial control simulations for estimating the natural internal variability, the signal-to-noise ratio may be too low to be able to detect a distinguishable pattern fingerprint. Then the first question that comes to mind is - if offline models are ill fitted to using the current methodology, why would you want to use them? Running coupled GCMs is extremely cumbersome computationally. A global offline model takes at least 10 times less computational time than a global coupled model (from months to days) while also requiring approximately half the computational resources. Since offline models run much more quickly, they allow for many more realizations to be produced. Take historical monthly near-surface air temperature for example, most of the CMIP5 GCMs only have 3 realizations,

and that number is cut in half for 4 realizations, whereas the C20C+ offline models usually have 15, 50, or 100. So, if you cannot use this methodology, what do you do? Some studies have used offline models for the comparison to the observations while using coupled model preindustrial control simulations for estimating the natural internal variability, but this method still relies on the computationally cumbersome coupled models. Providing an answer to this question without needing to rely on coupled climate models is the goal of my dissertation.

For applications, we want to determine if runoff and streamflow are changing in the contiguous United States and the Columbia River Basin, respectively. And if they are changing, what is driving the change. For runoff in the contiguous US, we are considering annual, seasonal, and regional changes for the time period 1950 – 2010. Regional changes are considered because of the heterogeneity in the US annual and seasonal trends. Trends in the eastern, northern, and southern US are positive while the western US has a negative trend. We are testing to see if trends are being driven by responses due to changing climate, carbon dioxide concentration, nitrogen deposition, and/or land use and land cover change. Also, of the forcings driving the trends, which are contributing the most. Observational monthly runoff data for the contiguous US was provided by WaterWatch from the United States Geological Survey (USGS). As for the offline land-surface model simulations, we used 6 models from the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP). MsTMIP is a terrestrial biospheric modeling project in which all the models use the same set of prescribed environmental and meteorological drivers, boundary conditions (e.g. soil type), and implementation protocol (e.g. spin-up procedures,

ensemble generation, and factorial model experiments) but each modeling group can have their own representation of physical processes (e.g. algorithms for plant growth, hydrology, and nutrient cycling).

In order to identify challenges and solutions for local communities, the changes and underlying drivers need to be understood at individual basin areas with higher resolutions, spatially and temporally. Given that the Columbia River Basin is a relatively smaller region (compared to the US), we want to do a more comprehensive study by determining if streamflow is changing and what is its relationship with precipitation, near-surface air temperature, and possibly evapotranspiration. Observational data which did not contain the first-order effects of human regulation and irrigation, provided by the Bonneville Power Administration (BPA), were used. This is important because climate models do not include human regulation of water flow. While this has been shown not to affect large spatial scale studies of hydrology over annual or longer means, it can influence regional studies when the region contains a considerable amount of regulation. For the model simulations, we studied the same forcings as were used for US runoff study but with new simulations. The new simulations were needed due to the precipitation driver from the MsTMIP models being too dry in the western US. This dryness affects both the magnitude and variability of the simulated runoff. The new simulations are also going to allow us the opportunity to test the sensitivity of the model output to the precipitation driver chosen.

CHAPTER I

CONTRIBUTION OF ENVIRONMENTAL FORCINGS TO US

RUNOFF CHANGES FOR THE PERIOD 1950 – 2010

A version of this chapter was originally published by Whitney L. Forbes, Jiafu Mao, Mingzhou Jin, Shih-Chieh Kao, Wenting Fu, Xiaoying Shi, Daniel M. Ricciuto, Peter E. Thornton, Aurélien Ribes, Yutao Wang, Shilong Piao, Tianbao Zhao, Christopher R. Schwalm, Forrest M. Hoffman, Joshua B. Fisher, Akihiko Ito, Ben Poulter, Yuanyuan Fang, Hanqin Tian, Atul K. Jain, and Daniel J. Hayes. “Contribution of environmental forcings to US runoff changes for the period 1950 – 2010.” *Environmental Research Letters* Volume 13 (2018): 054023.

Jiafu Mao and I were the leaders of this work. I did the data analysis and initial writing. Jiafu provided advising and discussion. Also advising the work was Mingzhou Jin, Shih-Chieh Kao, Xiaoying Shi, Daniel M. Ricciuto, Peter E. Thornton, and Aurélien Ribes. Shih-Chieh Kao provided the USGS WaterWatch runoff data and Wenting Fu provided the area-weighted climate model simulations. Yutao Wang, Shilong Piao, Tianbao Zhao, Christopher R. Schwalm, Forrest M. Hoffman, Joshua B. Fisher, Akihiko Ito, Ben Poulter, Yuanyuan Fang, Hanqin Tian, Atul K. Jain, and Daniel J. Hayes are the respective contacts for each of the MsTMIP models used. After I wrote the initial draft, it was revised by the core team (i.e. Jiafu Mao, Mingzhou Jin, Shih-Chieh Kao, Xiaoying Shi, Daniel M. Ricciuto, and myself). Once we were satisfied, it was submitted to ORNL’s internal system where it was reviewed by Peter E. Thornton and Scott L. Painter. Then it was submitted to *Environmental Research Letters* where it was peer reviewed.

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 CO₂ 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.

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-Lanko 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-Lanko et al., 2014; Anderegg et al., 2012). Inversely, plants affect runoff through canopy interception, evaporation, and transpiration (Gerten et al., 2003; 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 CO₂ 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., 2003; 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. (2003) 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 CO₂ concentration (Barnett et al., 2008; Pierce et al., 2008; Hidalgo et al., 2009). The limitation of Groisman et al. (2003), 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, CO₂ 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.

Data and Methodology

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 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 et al. (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 S6-7).

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 CO₂ 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 CO₂ 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 fourth and the third simulations to achieve the effect of atmospheric CO₂ 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 CO₂ 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 CO₂ 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 ($0.5^\circ \times 0.5^\circ$), 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 S2. 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 for each region and season was able to reproduce the variability within the observations relatively well with the minimum

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

and maximum R-squared value being 0.739 and 0.921, respectively. All of the R-squared values are reported in table S3.

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

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

where y are the WaterWatch observations, β_i is the scaling factor for forcing i , x_i are the model 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.

Results

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 2(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 2(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 S8). The dominant forcing plot also shows the CLMT forcing agrees with the ALL forcing over a large area within the US [figure 2(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 CO₂ induced stomatal closure. There is also a large region in figure 2(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 S9(b)] (Jung et al., 2006). The supplementary material includes figures for the season which dominates the trend of the annual total values (figure S10) and the spatial patterns of trends and dominant forcings for each season (figures S11 – 14).

The division of the CONUS HUC2 basins (R01 – R18) into 4 regions (north, east, south, and west) is shown in the background of figure 3(a). For observational annual total runoff for the CONUS over 1950–2010, there is an estimated positive trend of approximately 0.2 mm/yr² (insignificant for $\alpha = 0.05$).

The estimated trend for the eastern region is twice that at approximately 0.4 mm/yr² (insignificant for $\alpha = 0.05$). For the western region, however, the observational trend is -0.9 mm/yr² (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 3(a), CO₂ is the only other positive forcing. This explains why there is a large area dominated by CO₂ in figure 2(h). Overall, the seasonal trends in figure 3(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 figure S15-19. Possible causes for the discrepancy between the estimated trends are in the Discussion. Annual total regional trend values for individual models are shown in the supplementary material (figure S20).

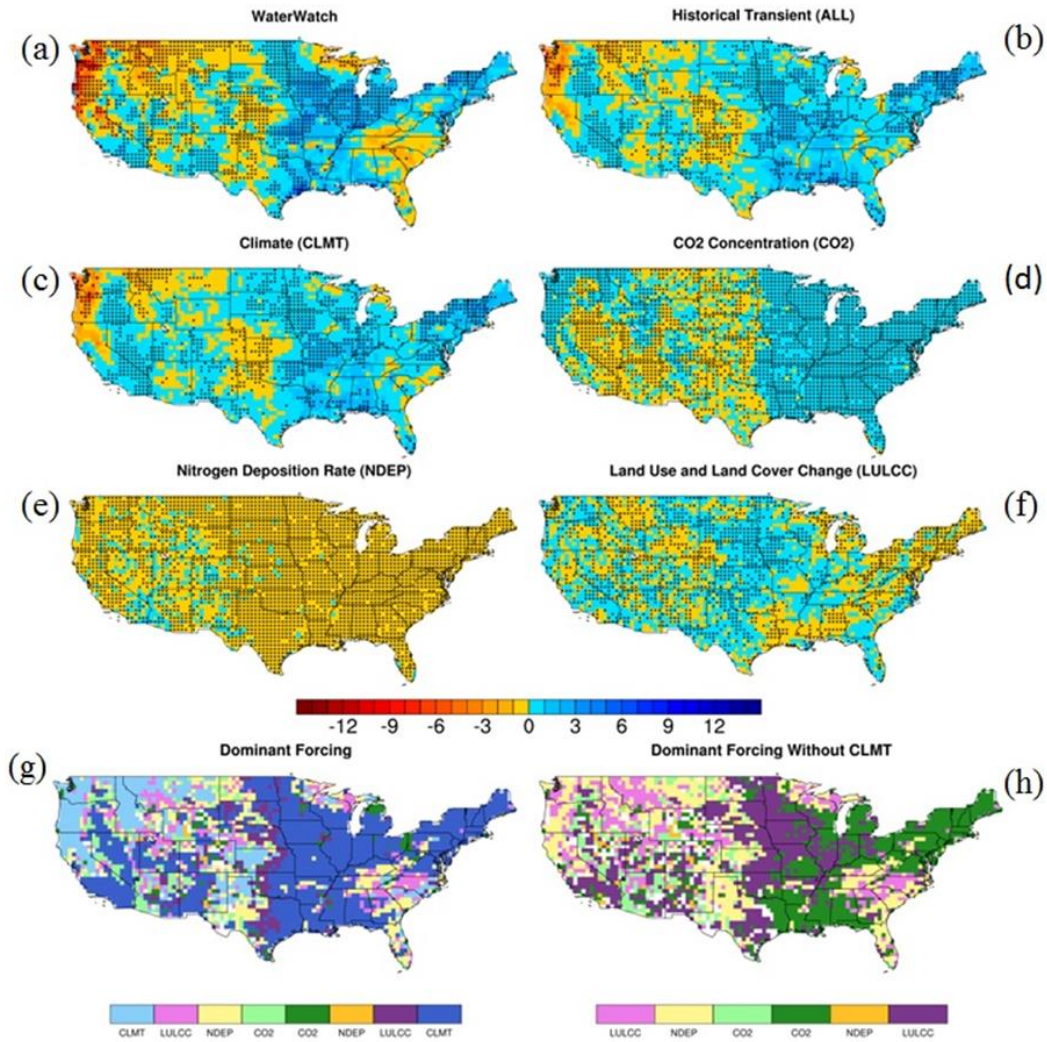


Figure 2. Spatial Pattern of Trends for Annual Total Runoff from WaterWatch and MSTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a) WaterWatch and (b-f) MSTMIP MME annual total runoff trends for 1950–2010, mm/yr² with dots representing grid cells with significant trends ($\alpha = 0.05$, 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.

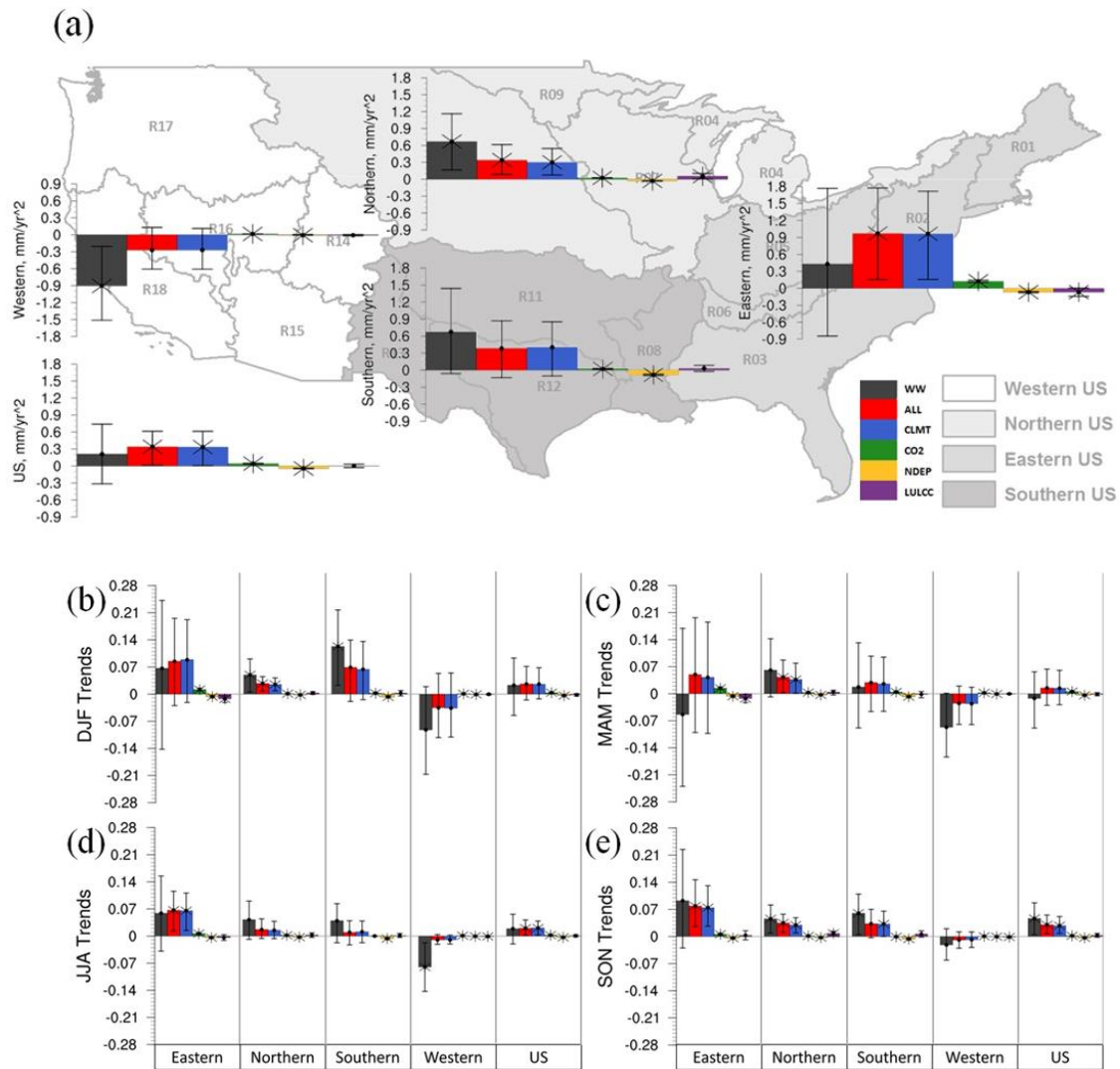


Figure 3. Regional Trends for Annual Totals (a) and Seasonal Means (b-e).

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.

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 2(g), there are large regions in the east and west where CLMT is the dominant forcing. These same regions can be seen in figure 4(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 [figure S21-24]. Scaling factor values are shown in figure 5(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 [figure S25-29]. 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

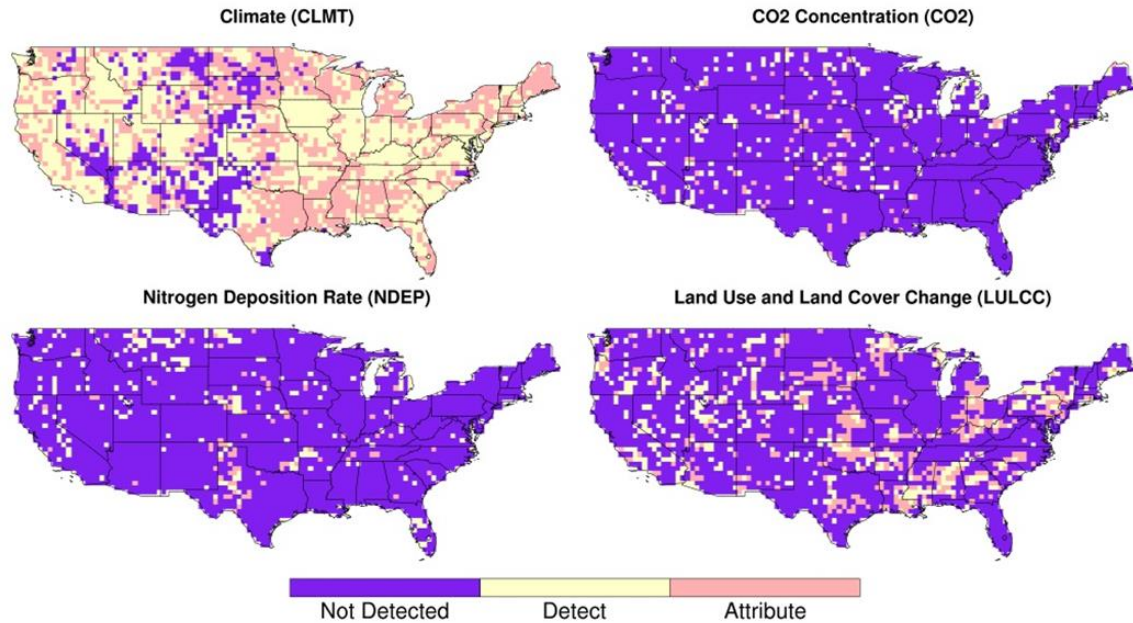


Figure 4. Spatial Distributions of D&A Scaling Factors. The D&A methodology was applied to each grid cell. Not detected (purple) denotes a scaling factor whose corresponding 95% confidence interval was less than zero or included zero. If the 95% confidence interval was greater than zero but did not include one, the forcing was detected (yellow). A positive confidence interval was labeled as attributed (pink) if it included one.

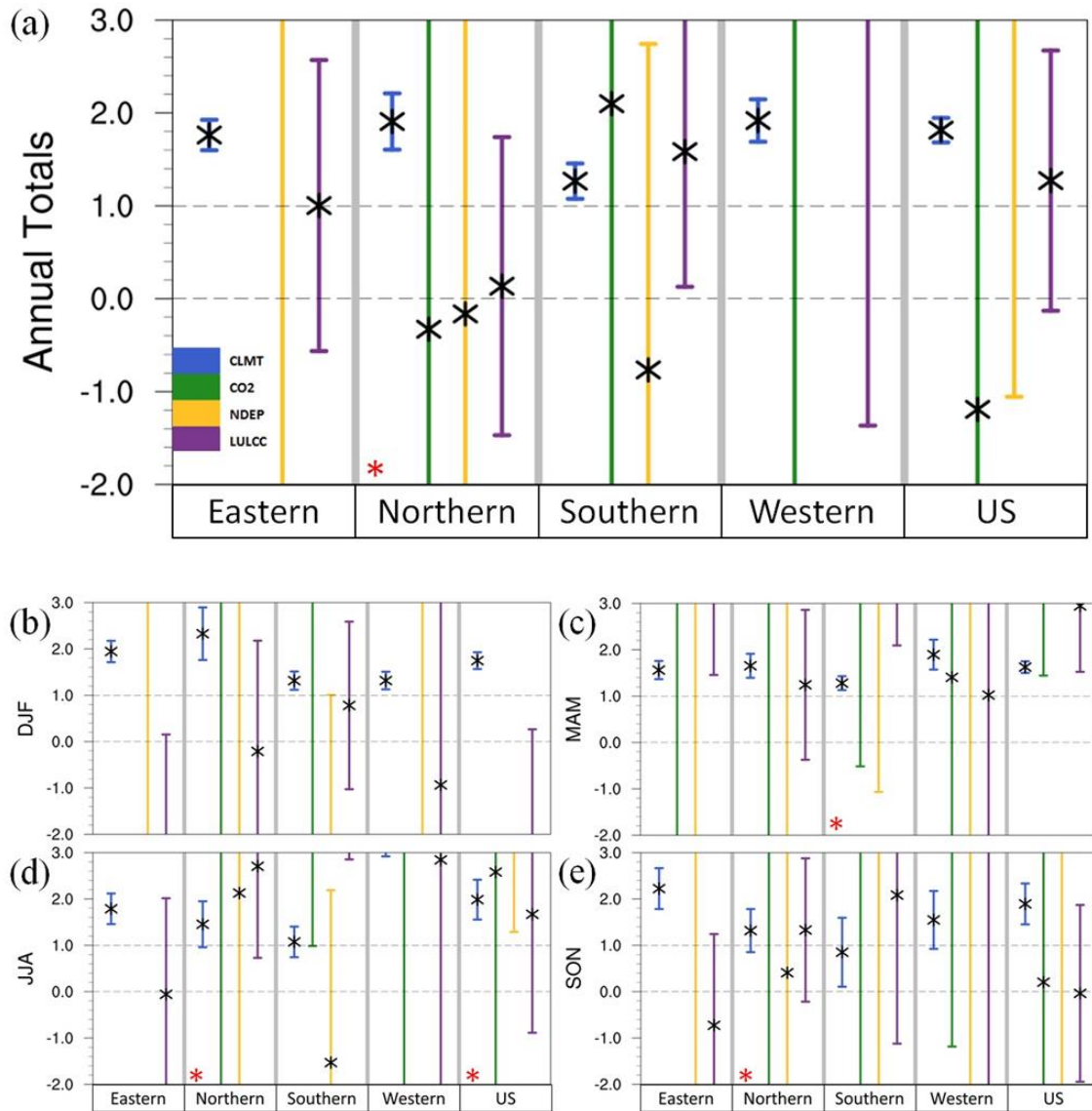


Figure 5. Scaling Factor Estimates (black asterisks) and Corresponding 95% Confidence Intervals for Annual Totals (a) and Seasonal Means (b-e). Dashed lines denote the values 0 and 1. Thick gray lines separate the results into different regions. A red asterisk in the bottom left corner for a region denotes where the residuals were autocorrelated.

large values (e.g. some confidence intervals are entirely outside the range of values considered), raising questions about physical realism. Results using only the three models with simulations for all of the forcings (CLM4, CLM4VIC, and TEM6) are shown in figure S30. Using only these three models leads to the same overall conclusions.

Discussion

Our results for the annual US runoff observations in figure 3(a) show the same spatio-temporal pattern as Groisman et al. (2003), 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. (2003) 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 CO₂ 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 CO₂ 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 CO₂ 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 detect the effects of climate, but rather than being overestimated, our results consistently show 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 CO₂ 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., 2003; 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 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 2 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 S15 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 β s 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 5-year means [figure S31], 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 S32 (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; Part 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.

Conclusions

Annual runoff observations for the period 1950–2010 had heterogeneous 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 led 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 project funded through the Terrestrial Ecosystem Science Program, and partially supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computing Scientific Focus Area project funded through the Regional and Global Climate 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. Oak Ridge National Laboratory is managed by UT-BATTELLE for DOE under contract 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. California Institute of

Technology. Government sponsorship acknowledged. JBF was supported in part by NASA SUSMAP, INCA, IDS, and CARBON programs. Atul K Jain was supported in part by DOE, NASA and NSF.

References

Alkama R, Ribes A, Decharme B and Marchand L 2013 Detection of global runoff changes: results from observations and CMIP5 experiments. *Hydrology and Earth System Sciences* **17** 2967–79 doi:10.5194/hess-17-2967-2013

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. *Proceedings of the National Academy of Sciences of the United States of America* **109**(1) 233–7 doi:10.1073/pnas.1107891109

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 doi:10.1002/2016JD025285

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 doi:10.1038/nature04141

Barnett T P *et al.* 2008 Human-induced changes in the hydrology of the Western United States. *Science* **319**(5866) 1080–3 doi:10.1126/science.1152538

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(8)** 05014004 doi:10.1061/(ASCE)HE.1943-5584.0000993

Betts R A *et al.* 2007 Projected increase in continental runoff due to plant responses to increasing carbon dioxide. *Nature* **448** 1037–U5
doi:10.1038/nature06045

Bindoff N L, Stott P A, AchutaRao K M, Allen M R, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov I I, Overland J, Perlwitz J, Sebbari R and Zhang X 2013 *Detection and Attribution of Climate Change: from Global to Regional. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)] (Cambridge, United Kingdom and New York, NY, USA: 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 doi:10.1016/0022-1694(82)90117-2

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. As.* **47(5)** 916–32 doi:10.1111/j.1752-1688.2011.00578.x

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, Washington. USA.
<http://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. Amer. Water Resour. Assoc.* **45** 279–90 doi:10.1111/j.1752-1688.2008.00277.x

Dai A, Qian T, Trenberth K E and Milliman J D 2009 Changes in Continental Freshwater Discharge from 1948 to 2004. *J. Climate* **22** 2773-92
doi:10.1175/2008JCLI2592.1

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. Climatol.* **28** 2031–64 doi:10.1002/joc.1688

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
doi:10.1002/2014GL059576

Diffenbaugh N S, Swain D L and Touma D 2015 Anthropogenic warming has increased drought risk in California. *Proc. Natl. Acad. Sci.* **112(13)** 3931–6 doi:10.1073/pnas.1422385112

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 doi:10.1073/pnas.1712381114

Fekete B M, Vörösmarty C J, Roads J O and Willmott C J 2004 Uncertainties in precipitation and their impacts on runoff estimates. *J. Climate* **17** 294-304 doi:10.1175/1520-0442(2004)017%3C0294:UIPATI%3E2.0.CO;2

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 doi:10.1038/nature04504

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. *Nature Geoscience* **7** 796-800 doi:10.1038/ngeo2263

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 doi:10.1016/j.jhydrol.2003.09.029

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 doi:10.1175/1525-7541(2004)005<0064:CCOTHC>2.0.CO;2

Hidalgo H G et al. 2009 Detection and attribution of streamflow timing changes to climate change in the Western United States. *J. Climate* **22** 3838-55 doi:10.1175/2009JCLI2470.1

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. *Geoscientific Model Development* **6** 2121-33 doi:10.5194/gmd-6-2121-2013

Jung M, Henkel K, Herold M and Churkina G 2006 Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sensing of Environment* **101(4)** 534–53 doi:10.1016/j.rse.2006.01.020

Kendall M G 1975 Rank correlation methods. Griffin, London

Kirshen P, Knee K and Ruth M 2008 Climate change and coastal flooding in Metro Boston: impacts and adaptation strategies. *Climatic Change* **90(4)** 453–73 doi:10.1007/s10584-008-9398-9

Koster R D et al. 2017 Hydroclimatic variability and predictability: a survey of recent research. *Hydrol. Earth Syst. Sci.* **21** 3777–98 doi:10.5194/hess-21-3777-2017

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 doi:10.5194/hess-12-1111-2008

Lettenmaier D P, Wood E F and Wallis J R 1993 Hydro-climatological trends in the continental United States, 1948 – 88. *J. Climate* **7** 586–607 doi:10.1175/1520-0442(1994)007<0586%3AHCTITC>2.0.CO%3B2

Lokupitiya E et al. 2016 Carbon and energy fluxes in cropland ecosystems: a model-data comparison. *Biogeochemistry* **129** 53–76 doi:10.1007/s10533-016-0219-3

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. *Environmental Research Letters* **10** 094008 doi:10.1088/1748-9326/10/9/094008

Mao J et al. 2016 Human-induced greening of the northern extratropical land surface. *Nature Climate Change* **6** 959–63 doi:10.1038/nclimate3056

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. *Global Planet. Change* **143** 100-117 doi:10.1016/j.gloplacha.2016.06.003

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 doi:10.1038/372666a0

Nicholls R J, Hanson S, Herweijer C, Patmore N, Hallegatte S, Corfee-Morlot J, Chateau J and Muir-Wood R 2008 *Ranking Port Cities with High Exposure and Vulnerability to Climate Extremes* OECD Environment Working Papers, No. 1 (Paris, France: Organisation for Economic Co-operation and Development (OECD) Publishing) pp 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(1)** 67–84 doi:10.5194/hess-18-67-2014

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.* **468-469** 139–50 doi:10.1016/j.jhydrol.2012.08.028

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 CO₂ on global river runoff trends, *Proc. Natl. Acad. Sci. U. S. A.*, 104(39), 15242–15247.

Pierce D W *et al.* 2008 Attribution of declining Western U.S. snowpack to human effects. *J. Climate* **21** 6425–44 doi:10.1175/2008JCLI2405.1

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 doi:10.5194/hess-19-2037-2015

Radcliffe D E and Mukundan R 2017 PRISM vs. CFSR precipitation data effects on calibration and validation of SWAT models. *Journal of the American Water Resources Association* **53(1)** 89–100 doi:10.1111/1752-1688.12484

Richardson G R A 2010 *Adapting to Climate Change: An Introduction for Canadian Municipalities* (Ottawa, ON, Canada: Natural Resources Canada) pp 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 [Barros V R, Field C B, Dokken D J, Mastrandrea M D, Mach K J, Bilir T E, Chatterjee M, Ebi K L, Estrada Y O, Genova R C, Girma B, Kissel E S, Levy A N, MacCracken S, Mastrandrea P R and White L L (eds.)] (Cambridge, United Kingdom and New York, NY, USA: 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**(24) 9320–25 doi:10.1073/pnas.1200311109

Scanlon B R, Duncan I and Reedy R C 2013 Drought and the water-energy nexus in Texas. *Environ. Res. Lett.* **8** 045033 doi:10.1088/1748-9326/8/4/045033

Schwalm C R 2015 How well do terrestrial biosphere models simulate coarse-scale runoff in the contiguous United States?. *Ecological Modelling* **303** 87–96 doi:10.1016/j.ecolmodel.2015.02.006

Shi X, Mao J, Thornton P E, Hoffman F M and Post W M 2011 The impact of climate, CO₂, nitrogen deposition and land use change on simulated contemporary global river flow. *Geophys. Res. Lett.* **38** L08704 doi:10.1029/2011GL046773

Sen P K 1968 Estimates of the regression coefficient based on Kendall's Tau. *J. Am. Stat. Assoc.* **63** 1379–89 doi:10.1080/01621459.1968.10480934

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 doi:10.1088/1748-9326/8/8/024012

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 doi:10.1002/2013WR014988

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. *Journal of the American Water Resources Association (JAWRA)* **53(2)** 258–79 doi:10.1111/1752-1688.12456

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 Mahowald N M 2007 Influence of carbon-nitrogen cycle coupling on land model response to CO₂ fertilization and climate variability. *Global Biogeochem. Cycles* **21** GB4018 doi:10.1029/2006GB002868

Wei Y *et al.* 2014 The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project: Part 2 - Environmental Driver Data. *Geoscientific Model Development* **7** 2875–93 doi:10.5194/gmd-7-2875-2014

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 U.S.A. *Climatic Change* **105(3-4)** 635–45 doi:10.1007/s10584-011-0024-x

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 doi:10.1029/2009WR008839

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. *Climatic Change* **109(Suppl 1)** S445–63 doi:10.1007/s10584-011-0329-9

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. Climate* **13** 1936–1950 doi:10.1175/1520-0442(2000)013<1936:VOMPIT>2.0.CO;2

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 doi:10.1080/07055900.2000.9649654

Zhu Z *et al.* 2016 Greening of the Earth and its drivers. *Nature Climate Change* **6**
791–5 doi:10.1038/nclimate3004

Supplementary Material

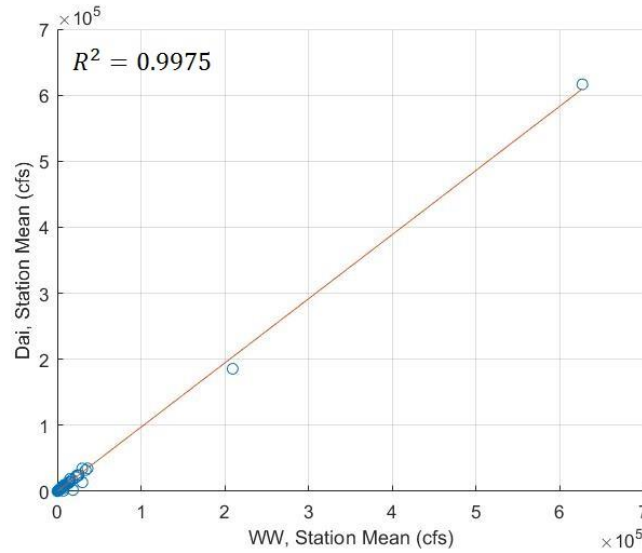


Figure 6. Comparison between WaterWatch stations and station from Dai's dataset (Dai et al., 2009). There are 59 stations in total.

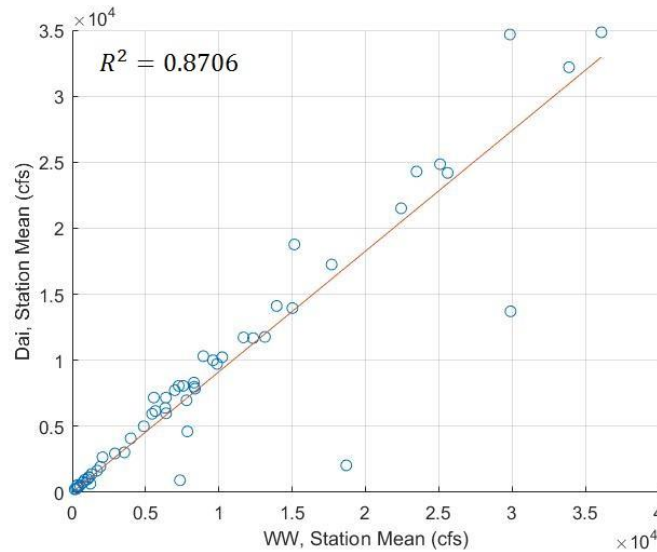


Figure 7. Comparison between WaterWatch stations and station from Dai's dataset (Dai et al., 2009). There are 57 stations in total (59 minus the two high leverage points from figure S1(a)).

Table 2. Meteorological Data Used by Each MsTMIP Model.

MsTMIP Model	Meteorological Variables Used
CLM4/CLM4VIC	Surface Air Temperature, Precipitation, Incoming Longwave Radiation, Incoming Shortwave Radiation, Specific Humidity, Surface Pressure, and Wind Speed
ISAM	Surface Air Temperature, Precipitation, Incoming Longwave Radiation, Incoming Solar Radiation, Specific Humidity, Surface Pressure, and Wind Speed
LPJ-wsl	Surface Air Temperature, Precipitation, Incoming Longwave Radiation, and Incoming Shortwave Radiation
VISIT	Surface Air Temperature, Precipitation, Incoming Longwave Radiation, Incoming Shortwave Radiation, Specific Humidity, and Wind Speed
TEM6	Surface Air Temperature, Precipitation, and Incoming Shortwave Radiation

Table 3. R-Squared for Each Season Using WaterWatch and ALL. Results are for the model ensemble mean and the values shown in parenthesis correspond to the values for the individual models (CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6).

	Eastern	Northern	Southern	Western	US
Annual Totals	0.891 (0.72, 0.576, 0.889, 0.86, 0.65, 0.905)	0.917 (0.759, 0.683, 0.841, 0.832, 0.829, 0.784)	0.904 (0.736, 0.586, 0.925, 0.799, 0.671, 0.858)	0.899 (0.885, 0.832, 0.906, 0.797, 0.865, 0.877)	0.921 (0.751, 0.704, 0.852, 0.875, 0.718, 0.922)
DJF	0.848 (0.619, 0.648, 0.914, 0.653, 0.638, 0.781)	0.8 (0.656, 0.471, 0.843, 0.389, 0.649, 0.466)	0.913 (0.762, 0.693, 0.92, 0.806, 0.674, 0.852)	0.896 (0.892, 0.878, 0.914, 0.748, 0.903, 0.822)	0.87 (0.728, 0.746, 0.901, 0.684, 0.683, 0.78)

Table 3. Continued.

	Eastern	Northern	Southern	Western	US
MAM	0.878 (0.732, 0.513, 0.916, 0.781, 0.741, 0.892)	0.878 (0.717, 0.685, 0.822, 0.78, 0.671, 0.779)	0.91 (0.671, 0.62, 0.914, 0.858, 0.674, 0.9)	0.887 (0.826, 0.831, 0.875, 0.823, 0.797, 0.82)	0.916 (0.76, 0.665, 0.899, 0.841, 0.812, 0.913)
JJA	0.859 (0.764, 0.622, 0.87, 0.724, 0.731, 0.76)	0.9 (0.611, 0.67, 0.84, 0.703, 0.783, 0.576)	0.754 (0.552, 0.452, 0.786, 0.612, 0.638, 0.442)	0.739 (0.613, 0.444, 0.35, 0.738, 0.272, 0.637)	0.789 (0.551, 0.555, 0.644, 0.644, 0.701, 0.713)
SON	0.891 (0.788, 0.706, 0.885, 0.774, 0.554, 0.837)	0.897 (0.761, 0.768, 0.857, 0.75, 0.782, 0.807)	0.752 (0.711, 0.496, 0.769, 0.673, 0.708, 0.644)	0.863 (0.853, 0.673, 0.819, 0.743, 0.759, 0.812)	0.891 (0.802, 0.67, 0.844, 0.765, 0.542, 0.827)

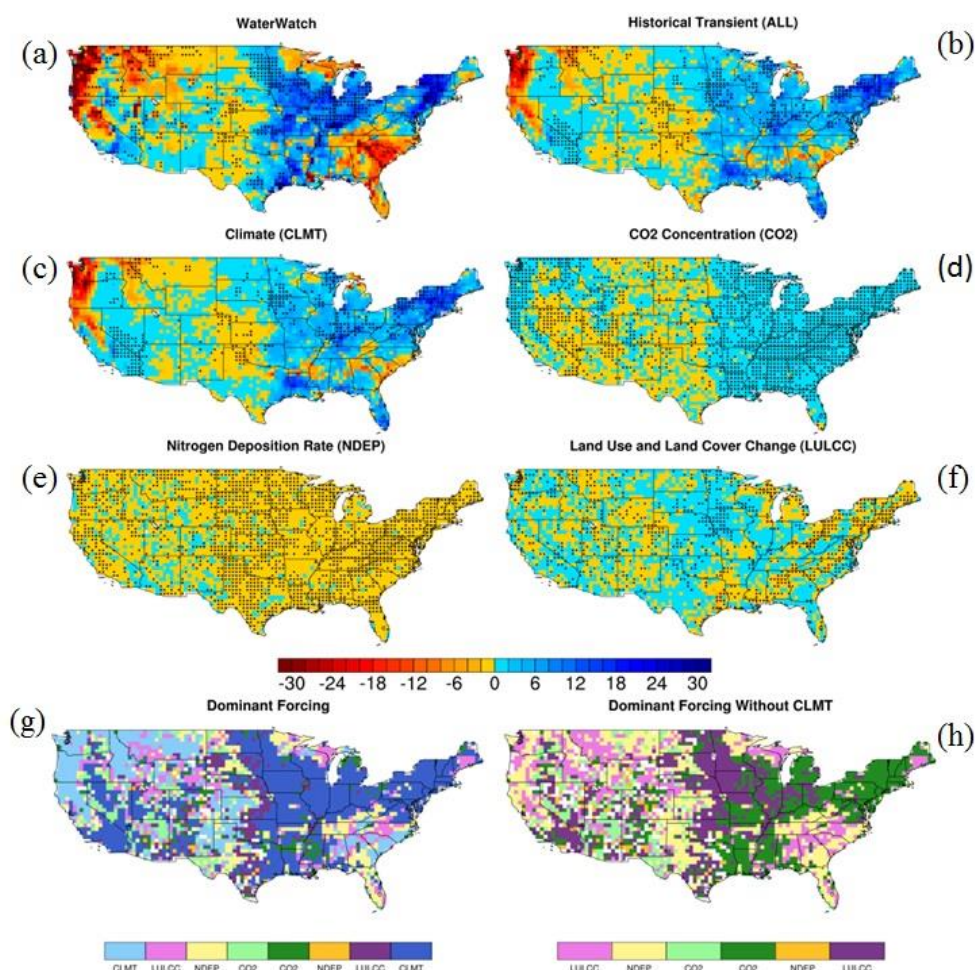


Figure 8. Spatial Patterns of Trends and Dominant Forcings using Pre-whitened 5-year Means from WaterWatch and MSTMIP Multi-Model Ensemble Mean (MME) Forcings. (a-f) WaterWatch and MSTMIP MME pre-whitened 5-year mean runoff trends for 1950-2010, (mm/yr)/5years with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). The 'zyp' R package function based on Zhang (1999) was used to obtain the pre-whitened 5-year means and their corresponding Theil-Sen trends and Mann-Kendall Significance (Kendall, 1975; Mann, 1945; Sen, 1968; Theil, 1950). (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.

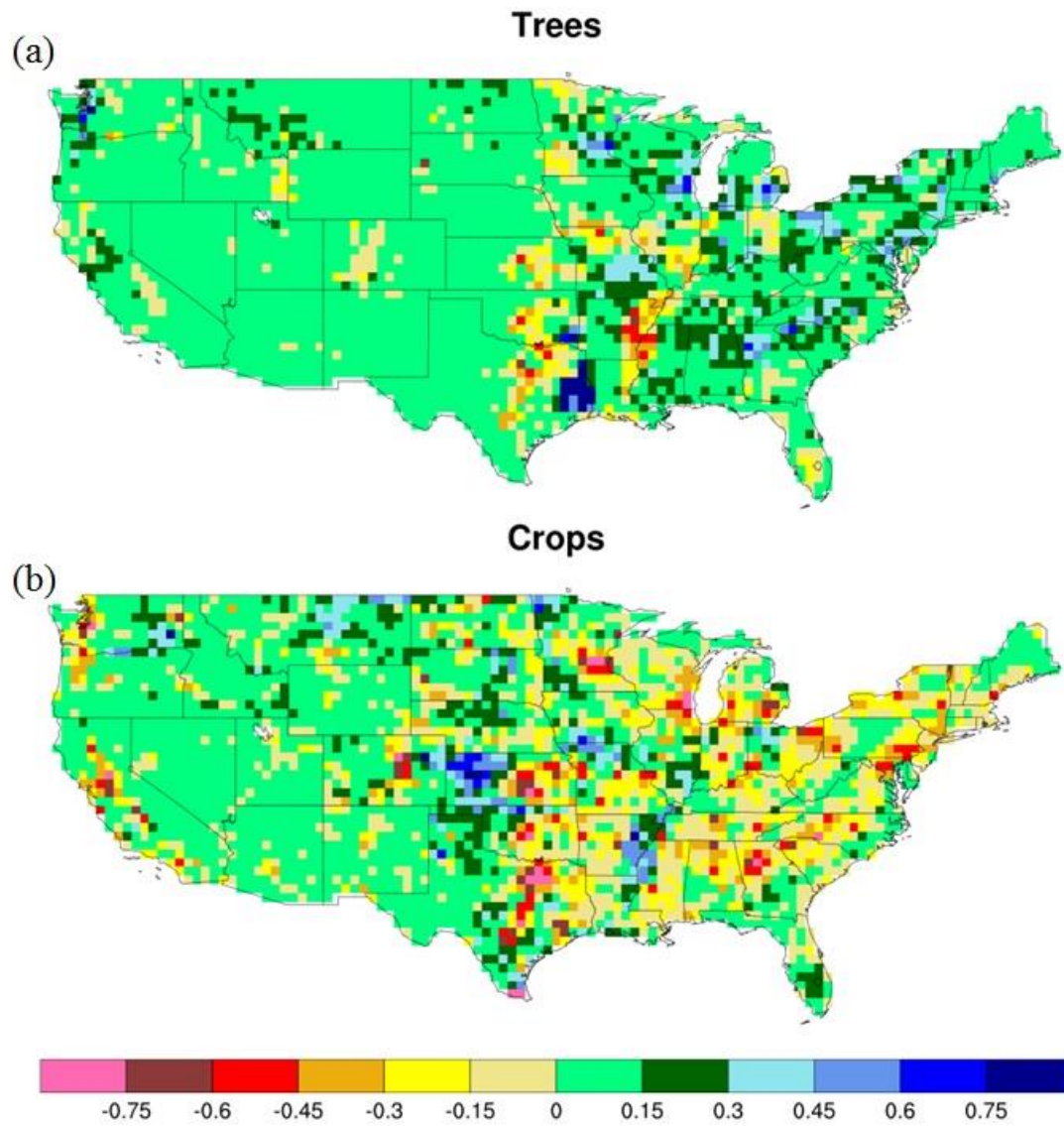


Figure 9. SYNMAP Vegetation Type Trend using Theil-Sen (Sen, 1968; Theil, 1950). Trends are for the percent plant functional type of each cell for (a) trees and (b) crops.

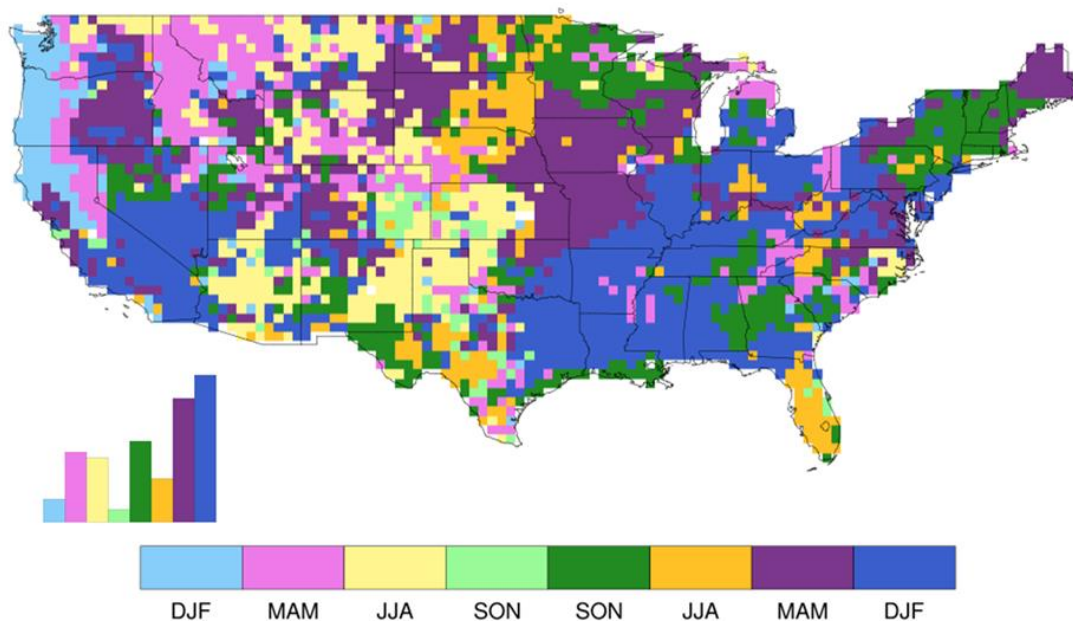


Figure 10. Dominant Season. The dominant season is the season that has the largest trend which has the same sign as the trend for ALL. White grid cells show locations where the sign of the trends for the seasons disagreed with the ALL forcing.

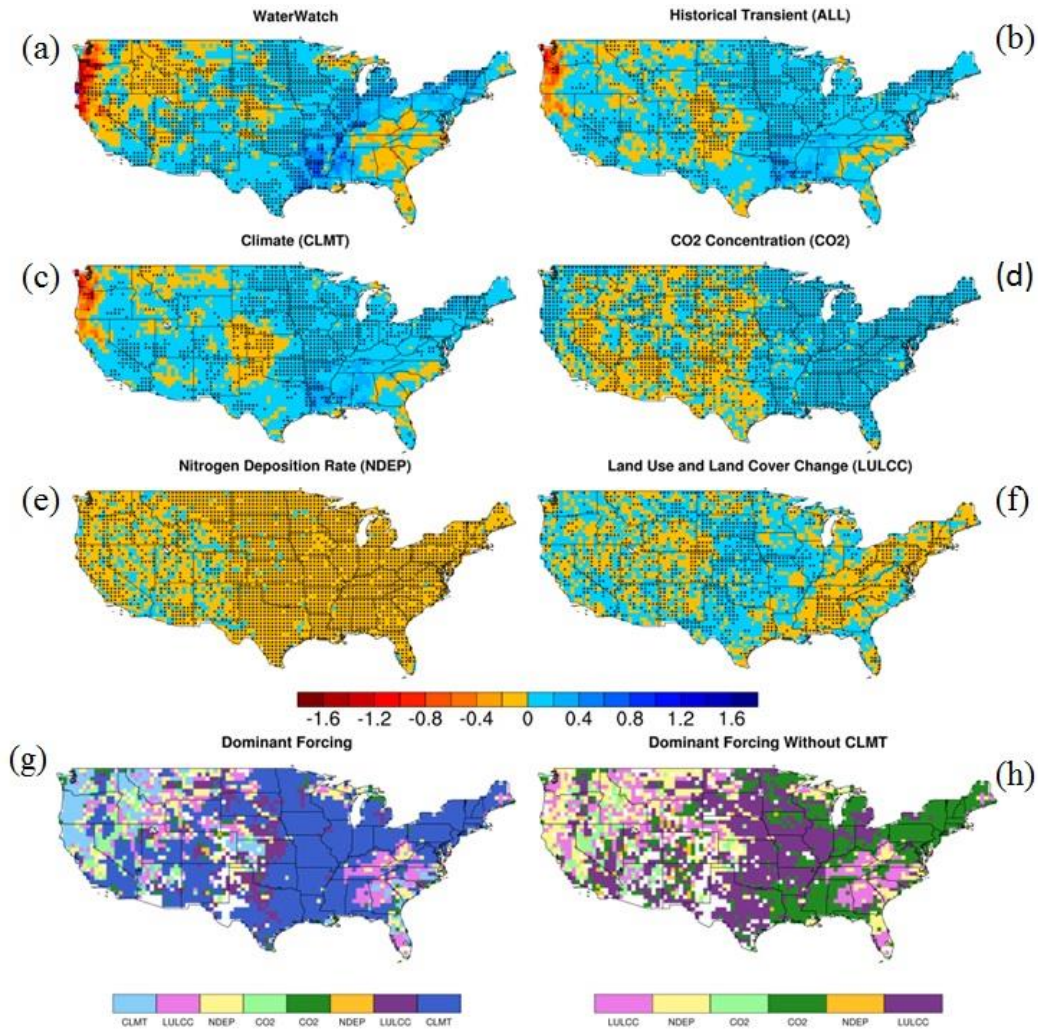


Figure 11. Spatial Pattern of Trends for DJF Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME DJF runoff trends for 1950-2010, mm/yr² with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). (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.

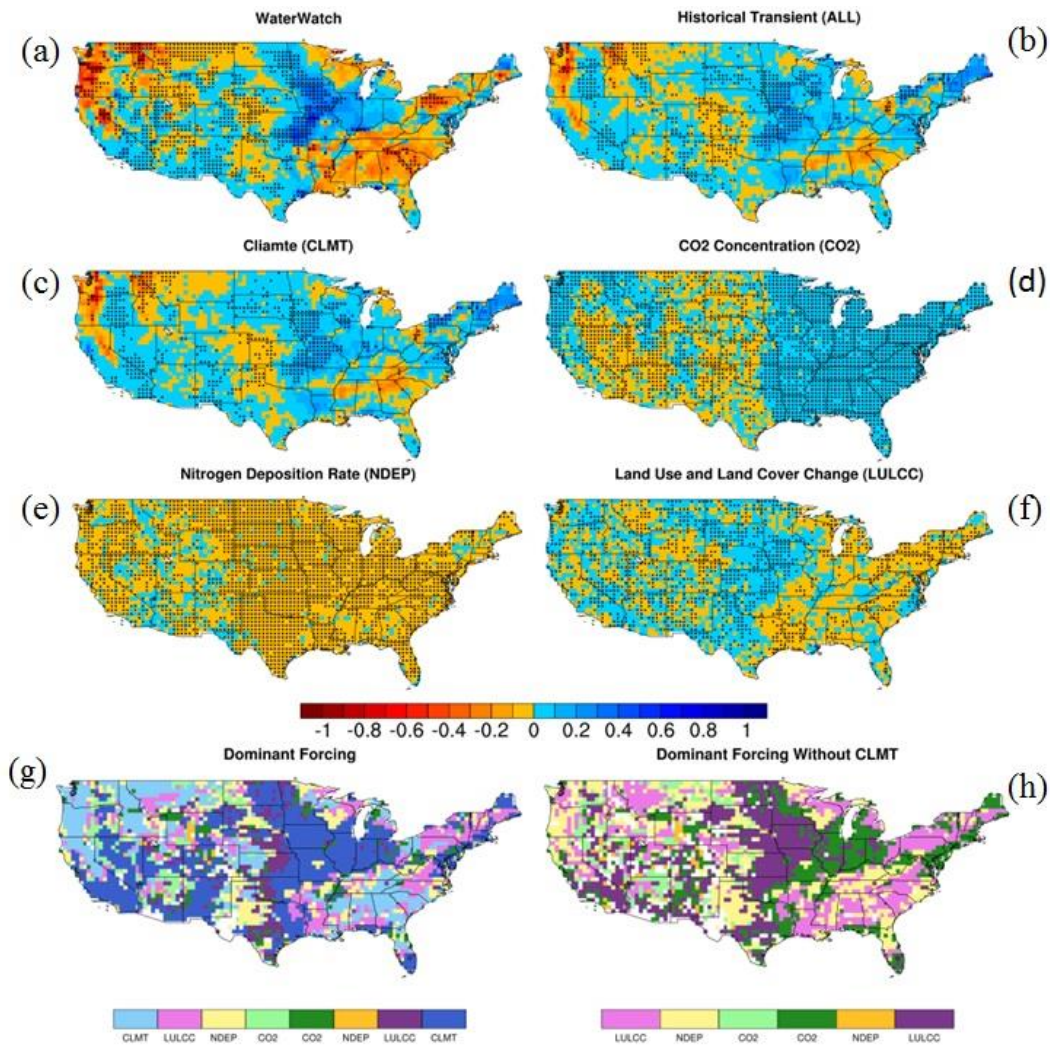


Figure 12. Spatial Pattern of Trends for MAM Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME MAM runoff trends for 1950-2010, mm/yr² with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). (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.

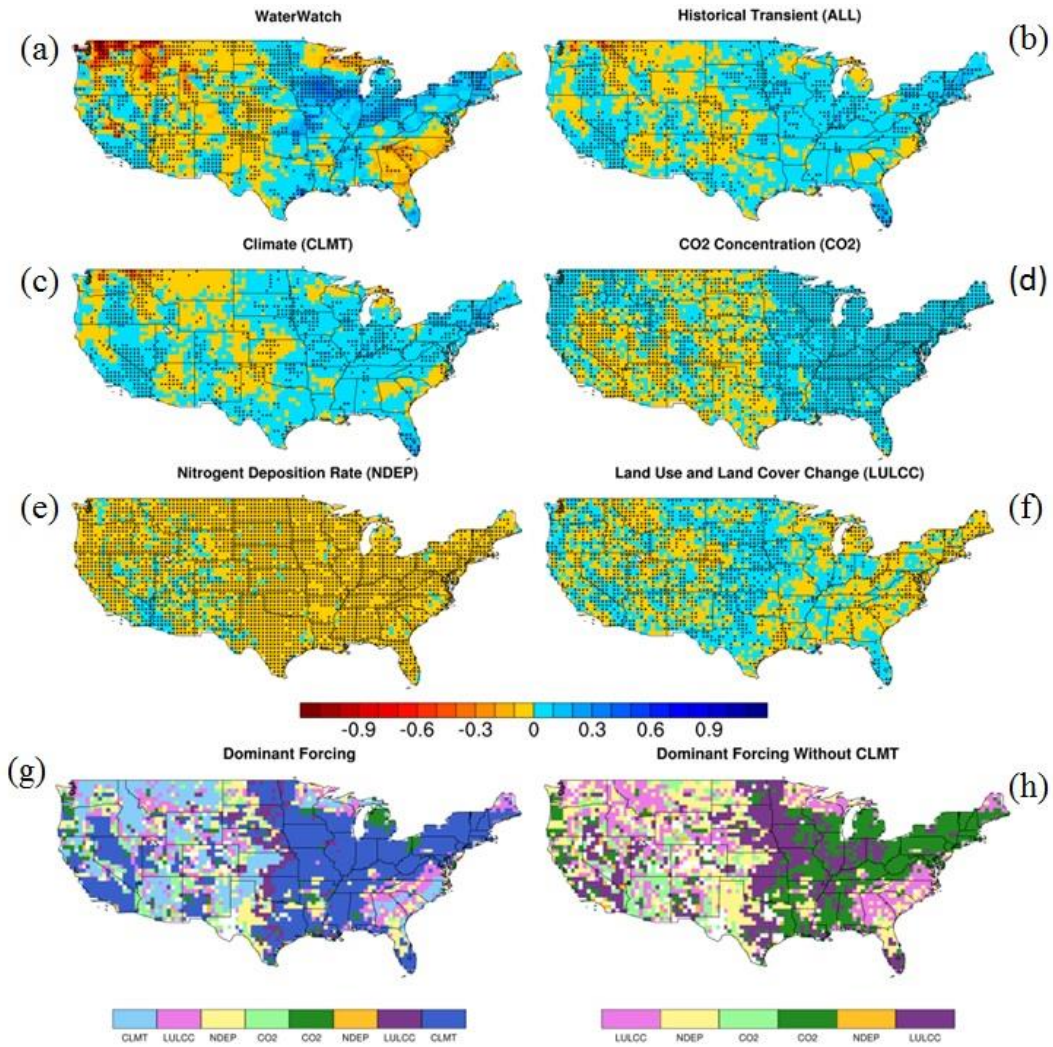


Figure 13. Spatial Pattern of Trends for JJA Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME JJA runoff trends for 1950-2010, mm/yr² with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). (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.

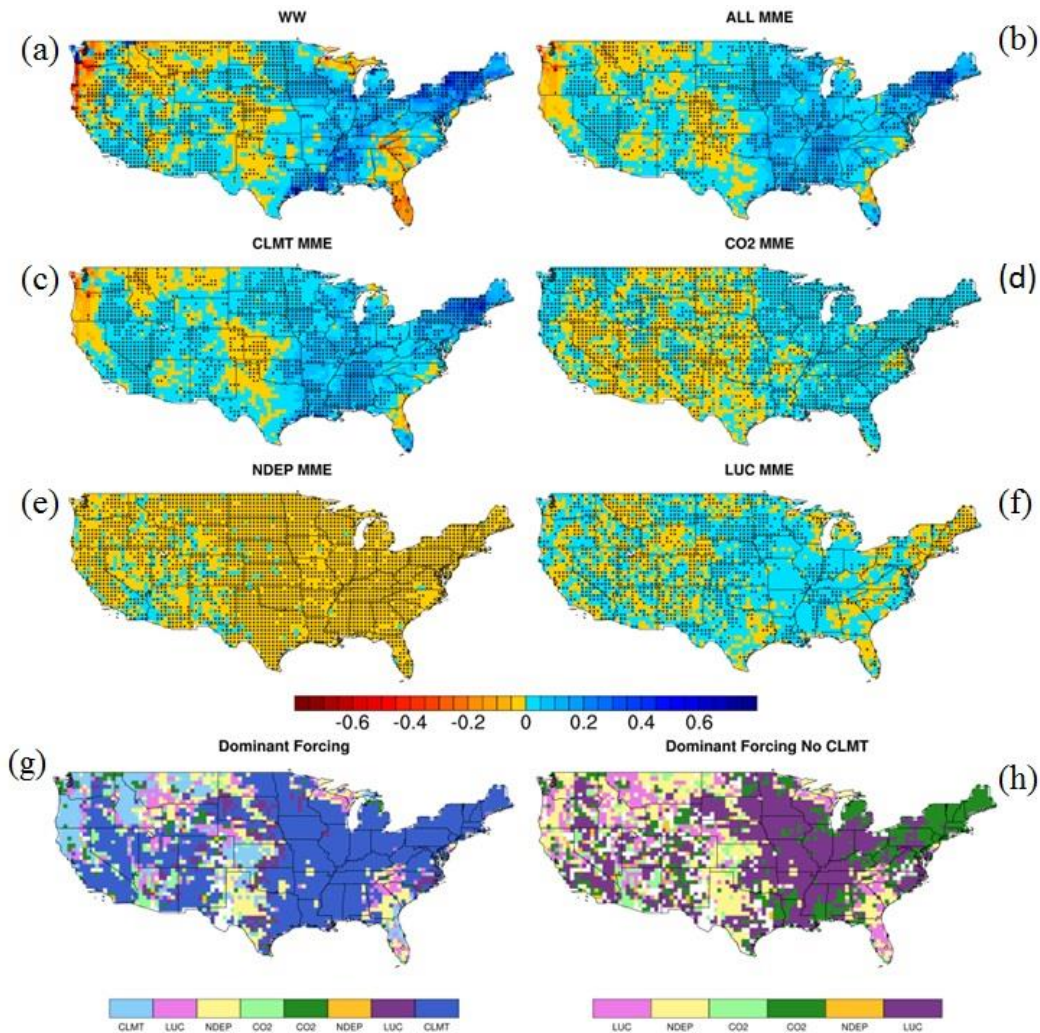


Figure 14. Spatial Pattern of Trends for SON Runoff from WaterWatch and MsTMIP Multi-Model Ensemble Mean (MME) Forcings and Dominant Forcings. (a-f) WaterWatch and MsTMIP MME SON runoff trends for 1950-2010, mm/yr² with dots representing grid cells with significant trends ($\alpha = 0.05$, Mann-Kendall). (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.

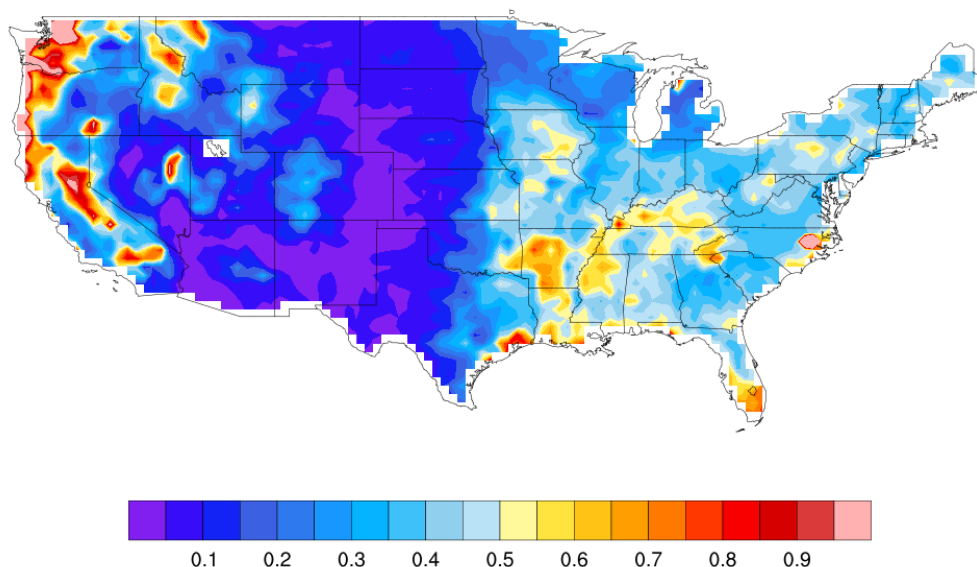


Figure 15. Root Mean Square Difference (RMSD) Values between WaterWatch and ALL Annual Totals. The RMSD values were normalized by the RMSD value corresponding to approximately the 99th percentile of values. The 99th percentile was chosen rather than the maximum due to extreme values.

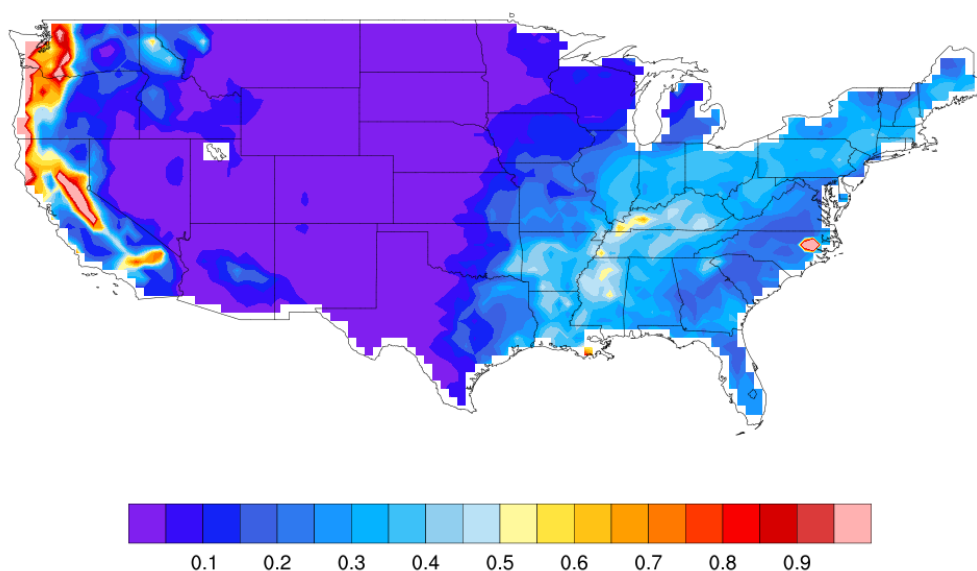


Figure 16. DJF RMSD. This figure is the same as figure S15 but instead of annual totals, the DJF season was used to determine the RMSD values.

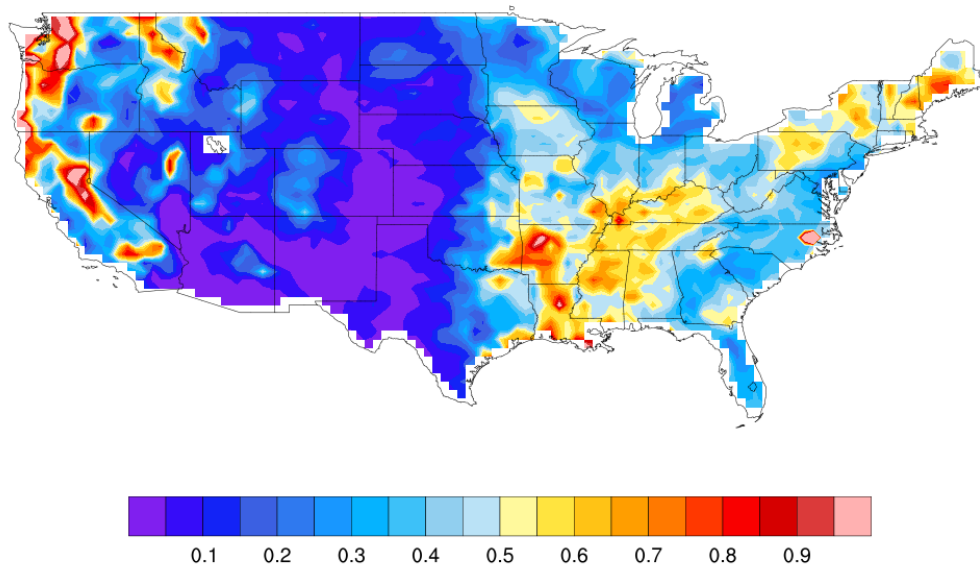


Figure 17. MAM RMSD. This figure is the same as figure S15 but instead of annual totals, the MAM season was used to determine the RMSD values.

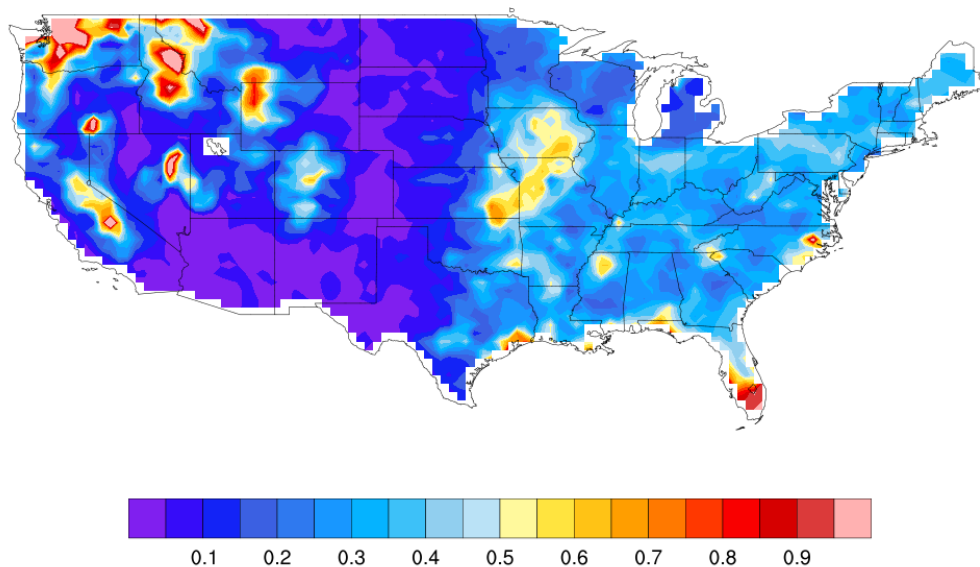


Figure 18. JJA RMSD. This figure is the same as figure S15 but instead of annual totals, the JJA season was used to determine the RMSD values.

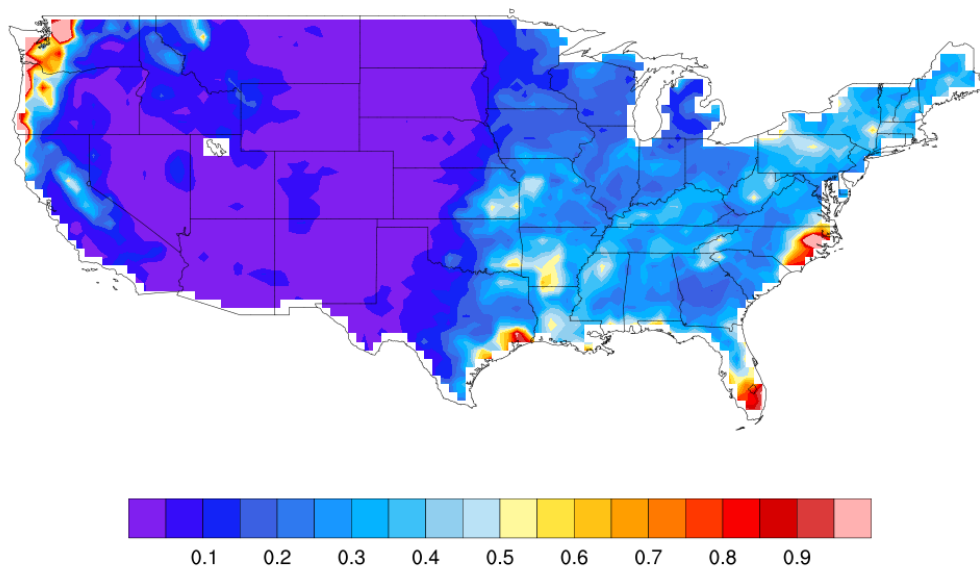


Figure 19. SON RMSD. This figure is the same as figure S15 but instead of annual totals, the SON season was used to determine the RMSD values.

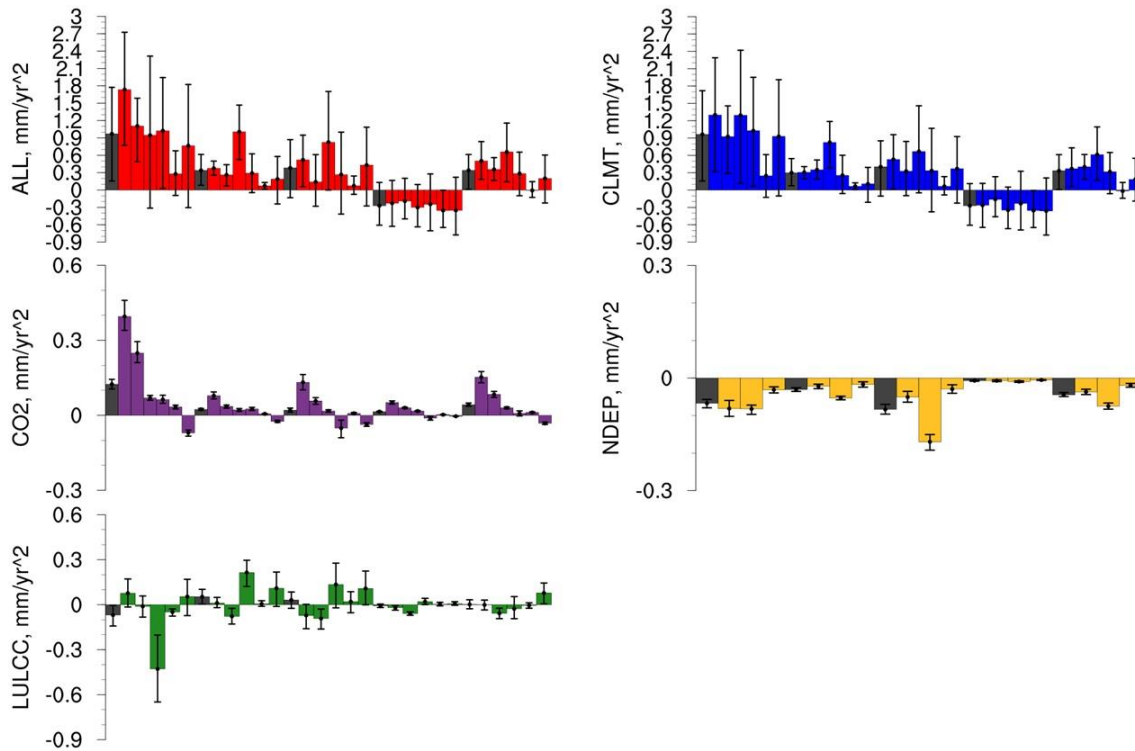


Figure 20. Regional Trends of Annual Total Values for the MstMIP MME (black bars) and Each Individual Model. Each black bar denotes the beginning of a region. The regions are ordered as follows: east, north, south, west, and US as a whole. The orders of the individual models are:

ALL	CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6
CLMT	CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6
CO2	CLM4, CLM4VIC, ISAM, LPJ-wsl, VISIT, TEM6
NDEP	CLM4, CLM4VIC, TEM6
LULCC	CLM4, CLM4VIC, ISAM, LPJ-wsl, TEM6

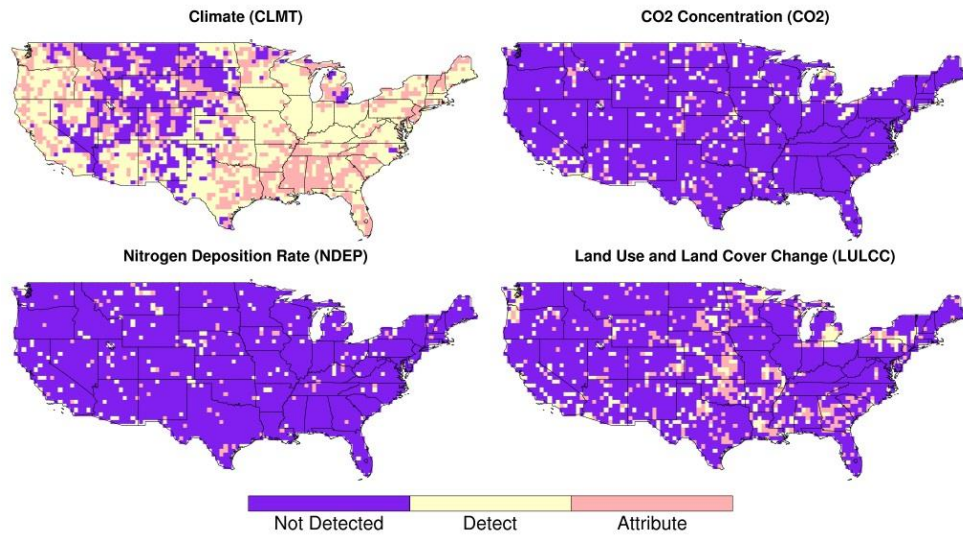


Figure 21. Spatial Distributions of D&A Scaling Factors for DJF. The D&A methodology was applied to each grid cell. Not detected (purple) denotes a scaling factor whose corresponding 95% confidence interval was less than zero or included zero. If the 95% confidence interval was greater than zero but did not include one, the forcing was detected (yellow). A positive confidence interval was labeled as attributed (pink) if it included one.

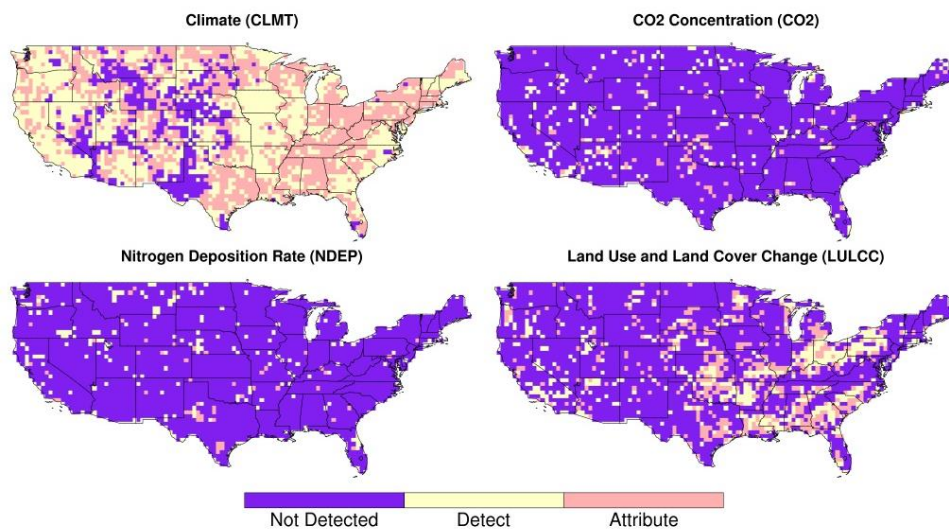


Figure 22. Same as figure S21 but for MAM.

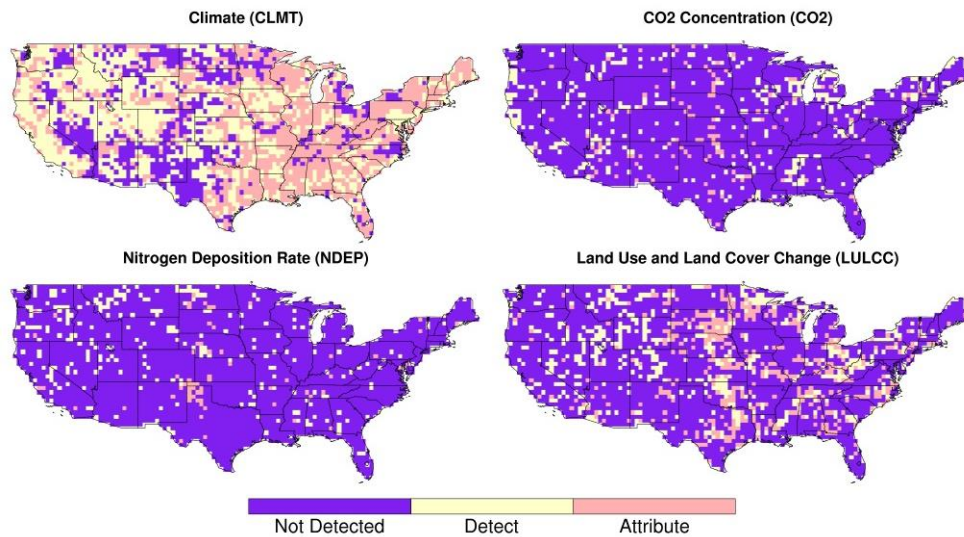


Figure 23. Same as figure S21 but for JJA.

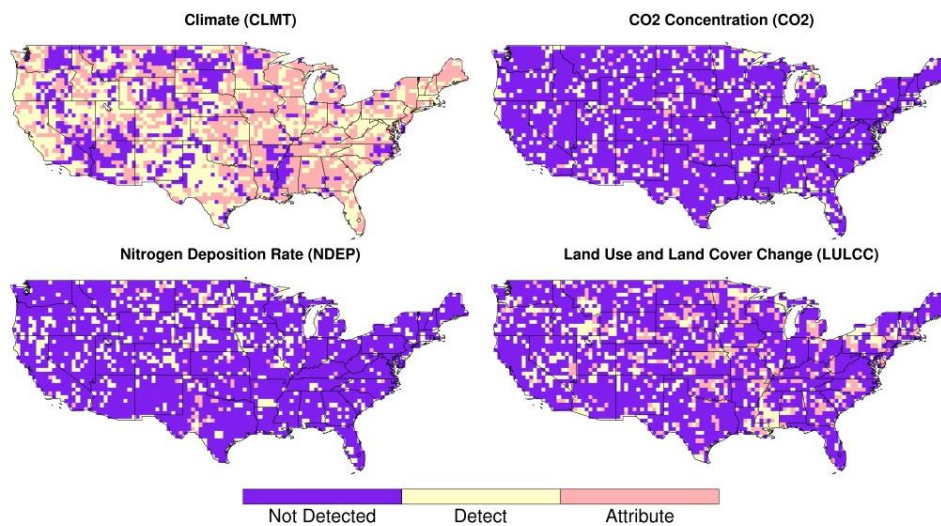


Figure 24. Same as figure S21 but for SON.

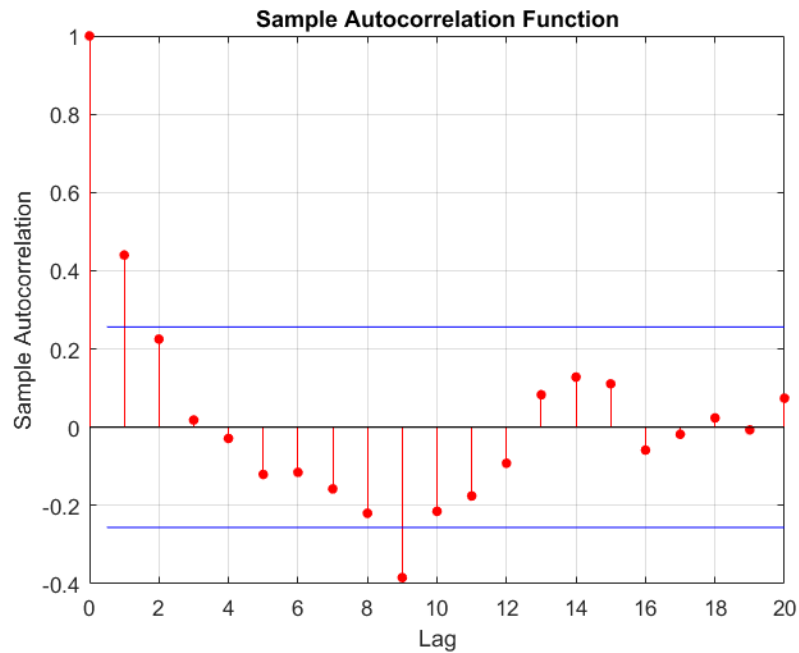


Figure 25. Autocorrelation Plot of the Residuals from the Regression using Northern Annual Totals. Blue lines represent the bounds for white noise.

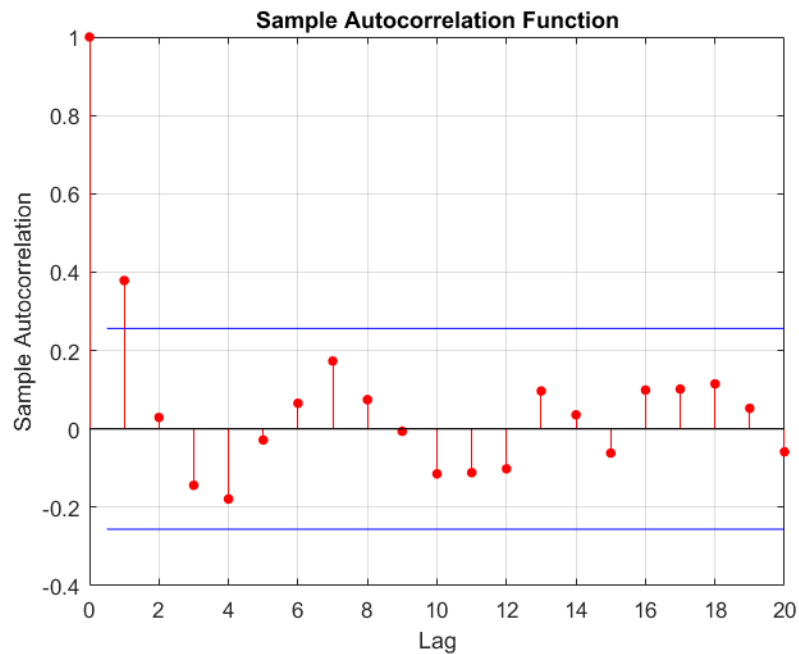


Figure 26. Autocorrelation Plot of the Residuals from the Regression using Southern MAM.

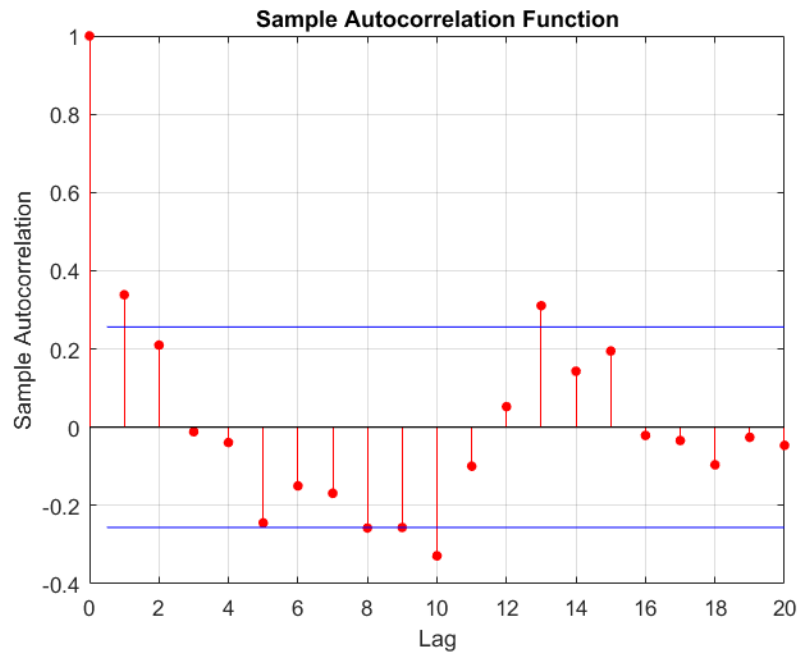


Figure 27. Autocorrelation Plot of the Residuals from the Regression using Northern JJA.

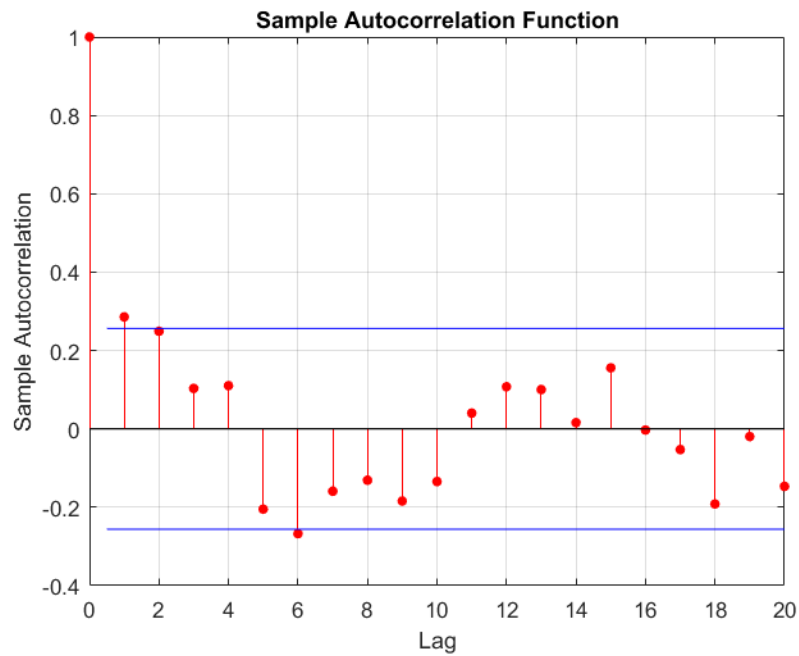


Figure 28. Autocorrelation Plot of the Residuals from the Regression using US JJA.

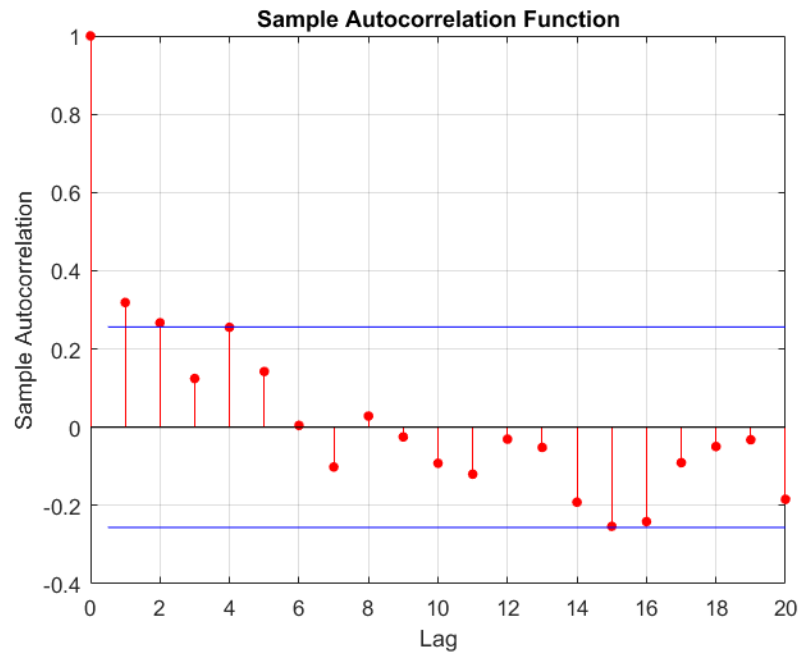


Figure 29. Autocorrelation Plot of the Residuals from the Regression using Northern SON.

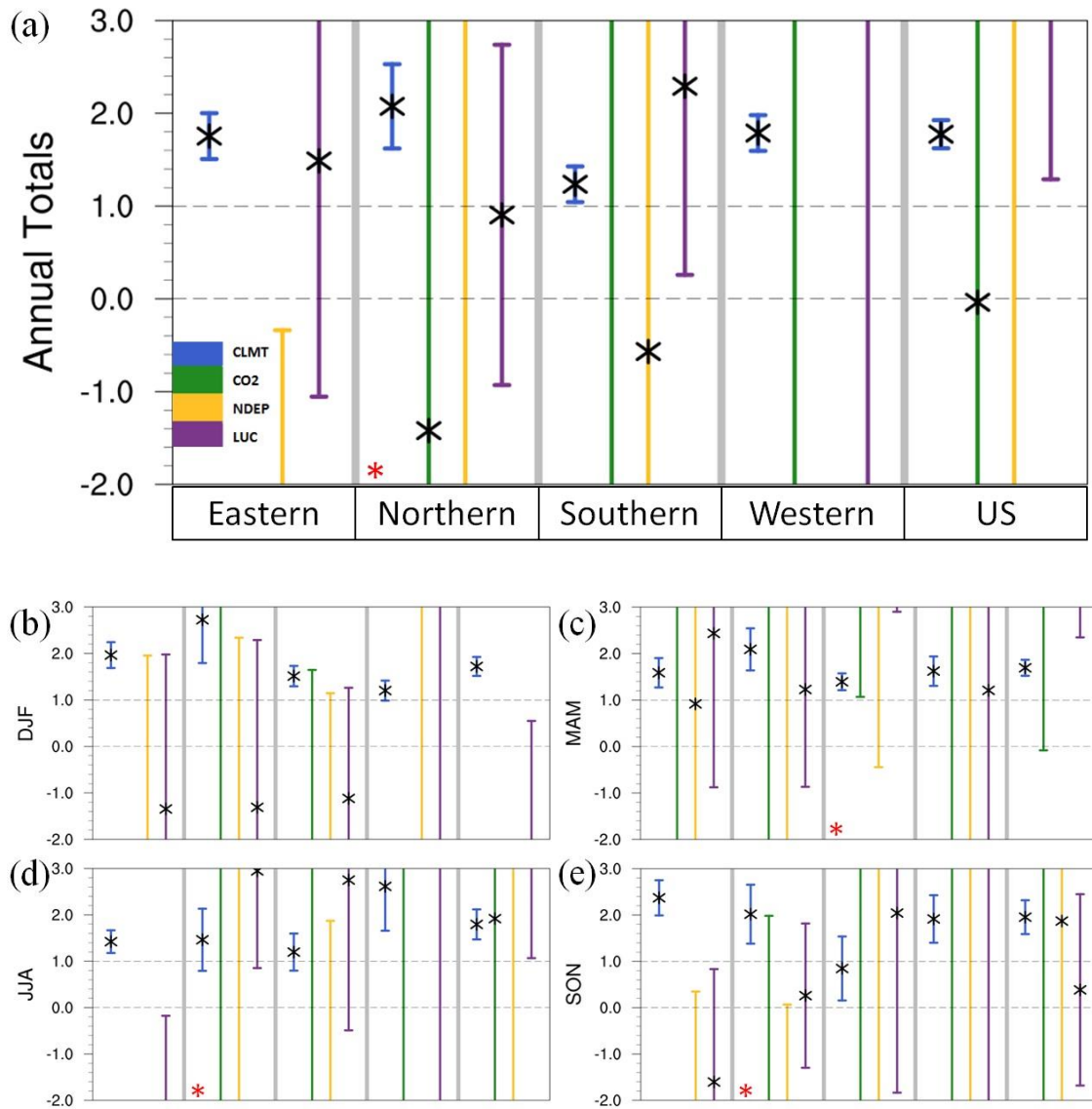


Figure 30. Scaling Factor Estimates (black asterisks) and Corresponding 95% Confidence Intervals for Annual Totals (a) and Seasonal Means (b-e) Using the MME of CLM4, CLM4VIC, and TEM6. Dashed lines denote the values 0 and 1. Thick gray lines separate the results into different regions. A red asterisk in the bottom left corner for a region denotes where the residuals were autocorrelated.

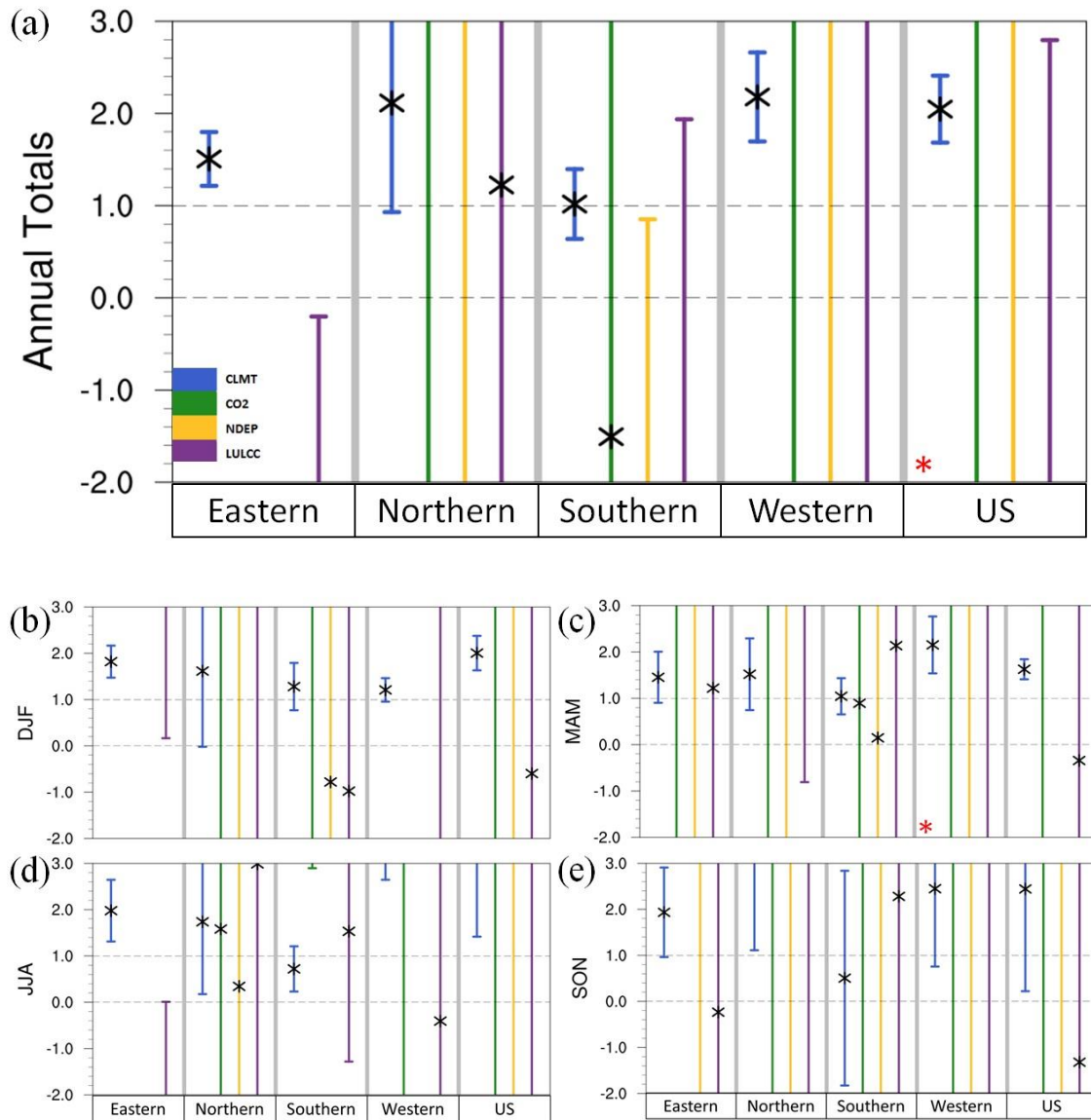


Figure 31. Scaling factor estimates (asterisk) and corresponding 95% confidence intervals for 5-year means of the annual totals (a) and seasonal means (b-e). Dashed lines denote the values 0 and 1. Thick gray lines separate the results into different regions. Red asterisk in the bottom left corner for a region denotes where the residuals were autocorrelated.

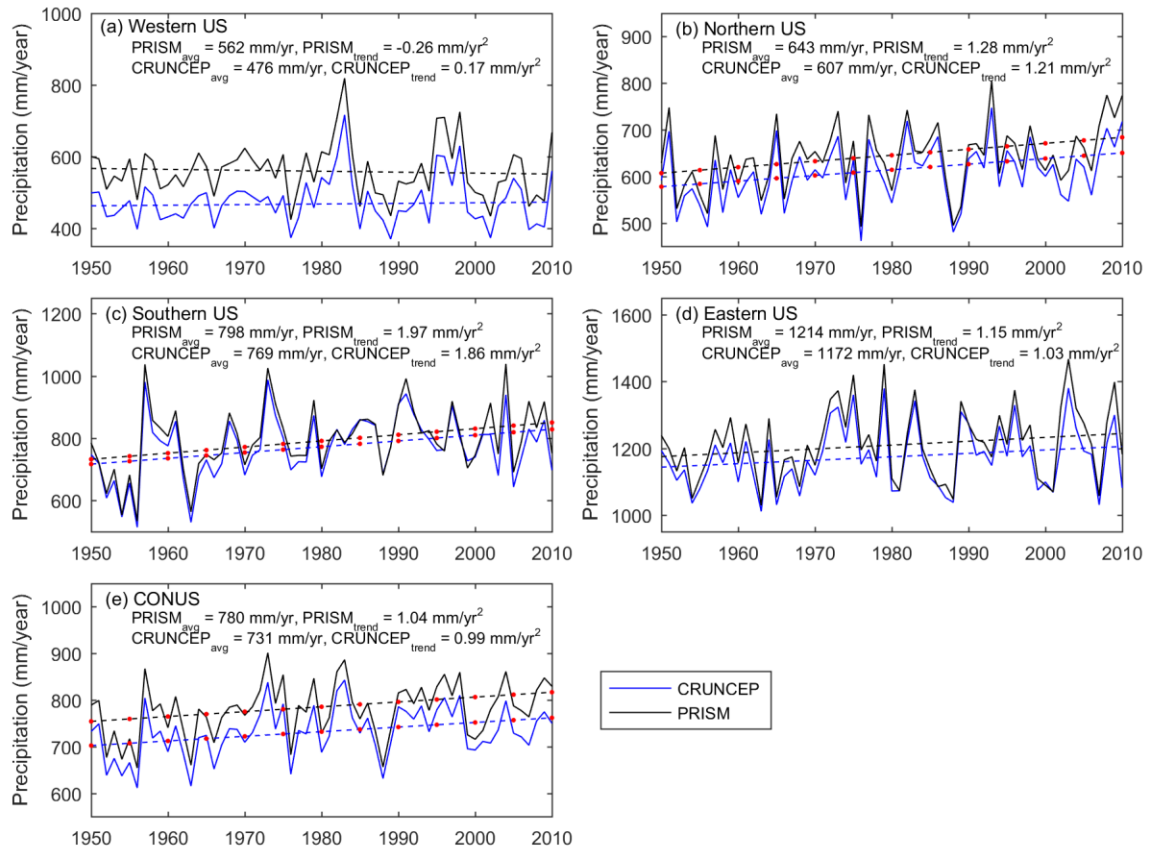


Figure 32. CRU-NCEP vs PRISM Annual Precipitation for Each Region and the US CONUS Over the Period 1950 – 2010. Included within each figure is the 1950 – 2010 time series, mean value for the time period, Theil-Sen trend estimate, and Mann-Kendall Significance (Kendall, 1975; Mann, 1945; Sen, 1968; Theil, 1950). A significant trend is denoted by red dots on the trend line.

CHAPTER II

**STREAMFLOW IN THE COLUMBIA RIVER BASIN: QUANTIFYING
CHANGES OVER THE PERIOD 1951 – 2008 AND DETERMINING
THE DRIVERS OF THOSE CHANGES**

This work has not been submitted to a journal yet, but it will be. Tentatively, the authors will be: Whitney L. Forbes, Jiafu Mao, Shih-Chieh Kao, Daniel M. Riccuito, Xiaoying Shi, Mingzhou Jin, Ahmad A. Tavakoly, and Peter E. Thornton. Shih-Chieh Kao and Daniel M. Riccuito prepped the driver data, Daniel M. Riccuito and Xiaoying Shi produced the climate simulations, and Shih-Chieh Kao ran the river routing model. Flow lines for the Canadian portion of the Columbia River Basin were provided by Ahmad A. Tavakoly. I analyzed the observational data and routed climate model simulations. Jiafu Mao and Mingzhou Jin were responsible for advising my work and providing thoughtful discussion and ideas.

Abstract

Trend and detection and attribution analyses were performed using naturalized streamflow observations and routed land surface model simulations for the Columbia River Basin covering the water years 1951 – 2008. The Columbia River Basin was separated into 10 subbasins with the flow from the Lower Columbia subbasin at The Dalles representing the flow for the entire basin. All the subbasins had significant ($\alpha = 0.10$) declines in the amount of annual total streamflow except Middle and Upper Snake and Upper Columbia. These declines in annual total flow were led by significant declines in the monthly flow for June – October. Declines in June – October also directly led to significant declines in the peak flow and July-August-September summer means. Trends for center of timing were not as consistent across all the subbasins, but a significant shift towards earlier center of timing was found at the Lower and Upper Columbia

and Kootenay subbasins. The RAPID routed semi-factorial E3SM land surface model simulations were driven by three different sets of meteorological drivers (i.e. CRUNCEP5, GSWP3, and Princeton) with the temperature and precipitation driver from Livneh. The mean of the three sets of simulations provided the historical changes in streamflow due to the effects of climate change (CLMT), CO₂ concentration (CO₂), nitrogen deposition (NDEP), and land use and land cover change (LULCC). Excluding the Snake River subbasins, LULCC had the same pattern of declines in monthly flow, but the period was shifted to May – September. The June – October pattern of significant trends was also found in NDEP; however, the trends showed significant increases in flow. While there were significant trends in CO₂, NDEP, and LULCC, the detection and attribution analysis showed that the change in annual total, center of timing of, and summer mean streamflow could only be attributed to CLMT. This is due to the signals in CO₂, NDEP, and LULCC being weak in comparison to the signal in CLMT and the natural internal variability found in streamflow. CLMT could only be detected in the annual maximum flow due to overestimation in the model simulations.

Introduction

Discovering and tracking trends in the hydrological cycle for the western United States has been an important area of research. In a series of papers, Barnett et al. (2008), Bonfils et al. (2008), Pierce et al. (2008), and Hidalgo et al. (2009) conducted detection and attribution (D&A) analyses on western US hydrology (i.e. river flow center of timing, winter air temperature, and snowpack), temperature (i.e. minimum and maximum daily temperatures, frost days, and

degree-days above 0°C), snowpack, and river flow center of timing, respectively. All 4 studies found that the trends in the respective variables over the period 1950 – 1999 were at least partially human-induced. Human-induced changes have also been discovered in summer streamflow in British Columbia (Najafi et al., 2017a). More recently, studies have been focusing on the change in snowpack and precipitation type. In addition to streamflow changes, Najafi et al. (2017b) also attributed declines in British Columbia spring snowpack to anthropogenic forcings. As for the western US, Mote et al. (2018) found that 33% of snow monitoring sites with observations for 1955 – 2016 had significant declines in snowpack. Berghuijs et al. (2014) showed that the fraction of annual precipitation falling as snow rather than rain has a significant influence on annual streamflow. More specifically, they determined that watersheds with a historically higher percentage of precipitation falling as snow, opposed to rain, had a higher long-term and inter-annual mean streamflow. This proportion of precipitation type is more closely linked to temperature variations than variations in precipitation amount in the western US (Safeeq et al.; 2016). For the future, Fyfe et al. (2017) projected that the snowpack will decline up to another 60% in the next 30 years. Safeeq et al. (2016) determined that the variability in the fraction of precipitation falling as snow for 1960 – 2003 resembles that projected when using the 2040-warming scenario of +1.8°C. The projected climate changes leading to decreases in snowpack also lead to decreases in mountain system groundwater recharge (Meixner et al., 2016). When assessing the sensitivity of western US flow regimes to climatic changes, Zhou et al. (2018) determined that in comparison to naturalized streamflow, the flow regime of regulated streamflows would change sooner, but the absolute change would be smaller. In 2005, Barnett et al. had

already predicted that by 2050, the Columbia River system would not be able to sustain both spring and autumn water releases for hydropower generation and releases for spring and summer salmon runs.

While detected changes have already been established and attributed for multiple hydrological variables in the western United States and British Columbia, in this study, the focus is the Columbia River Basin (CRB) and its subbasins. This is possible due to the availability of a naturalized daily streamflow dataset. We also took advantage of the capabilities of the recently developed river routing model, Routing Application for Parallel computation of Discharge (RAPID) (David et al., 2011a, b, 2013; Snow et al., 2016; Tavakoly et al., 2016). The same naturalized observational streamflow dataset was used in Hidalgo et al. (2009), but the temporal period was monthly and ended in 1999. Since then, the dataset has been extended to daily timescales through 2008. It is also important to determine if the trends found in previous studies are continuing. Assessing the changes in annual flow is needed, but the more important factor is seasonal changes. In order to investigate the seasonal flow, we looked at the center of timing, summer mean, and peak flow. Land surface model simulations of runoff routed by RAPID were used to quantify changes in the streamflow metrics due to changing climate and climate variability versus changes in evapotranspiration due to changes in carbon dioxide concentration, nitrogen deposition, and land use and land cover change. By also considering the subbasins within the CRB, spatial variability within these changes were also quantified.

Data

Observational Data

The observed streamflows used are from the Bonneville Power Administration (BPA). BPA operates 31 federal hydroelectric dams within the CRB. Using historical observations from the United States Geological Survey (USGS) National Water Information System (NWIS) and their own knowledge of the reservoir regulations imposed and depletions from irrigation, BPA estimated what the historical flows would have been without the first order effects of regulation and irrigation. These data are not true naturalized flows because they still contain the secondary effects of regulation and irrigation (e.g. lake attenuation, return flow lag, and ground water delays); however, they provide a good proxy for the naturalized flow. The dataset contains daily data for the time period July 1, 1928 – September 30, 2008 for 197 stations covering both the United States and Canadian portions of the CRB. The analysis reported here covered the period October 1, 1950 – September 30, 2008 (i.e. the water years 1951 – 2008). More information on the methodology for how this dataset was formed can be found in Climate and Hydrology Datasets for RMJOC Long-Term Planning Studies: Second Edition (RMJOC-II) Part I: Hydroclimate Projections and Analyses (Pytlak et al., 2018).

Climate Model Simulations

Semi-factorial land surface model simulations of runoff were produced using 6 sets of meteorological drivers to force the 0.5-degree resolution Energy Exascale

Earth System Model (E3SM) Land Model (ELM). The driver datasets considered were CRUNCEP5, GSWP3, and Princeton. Three more datasets were produced by replacing the precipitation and temperature variables from the respective driver datasets with the precipitation and temperature data from Livneh (CRUNCEP5-Livneh, GSWP3-Livneh, and Princeton-Livneh). 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 CO₂ 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 CO₂ 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 fourth and the third simulations to achieve the effect of atmospheric CO₂ concentration (named CO₂). To isolate the effect of nitrogen deposition, we use the difference of ALL and the fourth simulation (named NDEP). Since the radiative and physiological effects of CO₂ concentration on climate change cannot be separated by using offline ELM simulations and are included in the transient climate drivers, CO₂, NDEP, and LULCC thus represent the direct effects of CO₂ physiology, nitrogen deposition, and land use and land cover change, respectively (Gedney et al., 2014; Mao et al., 2016; Zhu et al., 2016).

All available daily CMIP5 piControl simulations of runoff with at least 58 years of data were also used in order to characterize the natural internal

variability of the system. Each simulation was separated into independent segments of 58 years resulting in 18, 3, 14, and 8 segments from CanESM2, CSIRO-Mk3-6-0, MIROC5, and NorESM1-1M, respectively. Each of the models was spatially interpolated to a 0.5-degree resolution to be comparable to the ELM factorial simulations.

River Routing and Downstream Gauge Stations

Each of the model simulations of runoff was routed using the RAPID river routing model. NHDPlus flowlines were used for the United States region, and Canadian flowlines were provided by Dr. Tavakoly at the United States Army Engineer Research and Development Center (Follum et al., 2016; NHDPlus, Horizon Systems Corporation, 2007). Using the flowlines and USGS latitude/longitude information from each observational station (179 total), a flowline segment was identified to represent RAPID's location for each gauge station.

The farthest downstream station was used to represent 10 subbasins within the CRB. The specific subbasins and representative stations analyzed along with their location (latitude/longitude), estimated drainage area from USGS and NHDPlus, elevation range, and corresponding USGS gauge station number (where applicable) are listed in table 4. Note that not all of the subbasins are independent of one another. For example, Mid-Columbia includes all the flow from the respective upstream watersheds. The independent subbasins are labeled with an asterisk in table 4. Even though the Bonneville station (BON) has the largest drainage area, we chose The Dalles to represent the Lower Columbia subbasin, and thus, the cumulative flow for the entire CRB. The spatial locations

Table 4. Individual CRB Subbasins and Corresponding Downstream Stations.

WATERSHED	STATION (ABBREV.)	USGS STATION NUMBER	LATITUDE, LONGITUDE	BPA BASIN AREA (KM²)	NHDPLUS BASIN AREA (KM²)
LOWER COLUMBIA	The Dalles (TDA)	14105700	45.61, -121.17	613830	572016
	Bonneville (BON)	14128870	45.63, -121.95	621341	579002
LOWER SNAKE	Ice Harbor (IHR)	13353000	46.25, -118.88	281015	247722
MIDDLE SNAKE	Hells Canyon (HCD)	13290450	45.25, -116.70	189846	159340
UPPER SNAKE*	King Hill, ID (SKHI)	13154500	43.00, -115.20	92722	66439
MID COLUMBIA	Priest Rapids (PRD)	12472800	46.63, -119.86	248640	241538
YAKIMA*	Kiona, WA (KLOW)	12510500	46.25, -119.48	14543	13198
SPOKANE*	Little Falls (LFL)	12433500	47.83, -117.94	16420	15071
PEND OREILLE*	Waneta (WAT)	----	49.00, -117.61	66822	----
KOOTENAY*	Brilliant (BRI)	----	49.32, -117.62	49987	----
UPPER COLUMBIA*	Murphy Creek (MUC)	----	49.18, -117.72	88098	88099

of the downstream stations (including BON) are shown in figure 33 along with their relation to the major flow lines and topography within the CRB.

The multi-model ensemble mean ALL forcing data from the routed ELM simulations driven by CRUNCEP5-Livneh, GSWP3-Livneh, and Princeton-Livneh (ALL MME-Livneh) for each downstream station was compared to the observations using the Nash-Sutcliffe Efficiency Coefficient (NSE). The logarithmic version of NSE was also calculated. In comparison, the logarithmic version is less sensitive to differences in peak or extreme values (Krause et al., 2005). The ensemble mean routed flow performed well overall. Using the classification from Moriasi et al. (2007) for NSE values using monthly data, NSE values >0.75 were considered very good, $0.65 < \text{NSE} \leq 0.75$ were good, $0.5 < \text{NSE} \leq 0.65$ were satisfactory, and ≤ 0.5 were unsatisfactory. For the non-logarithmic version, Lower and Mid Columbia were good; Kootenay and Upper Columbia were satisfactory; and Lower, Middle, and Upper Snake, Yakima, Spokane, and Pend Oreille were unsatisfactory. However, all the subbasins except Upper Snake were reproduced well with very good, good, or satisfactory logarithmic NSE values. The numerical NSE values for the downstream stations are listed in table 5.

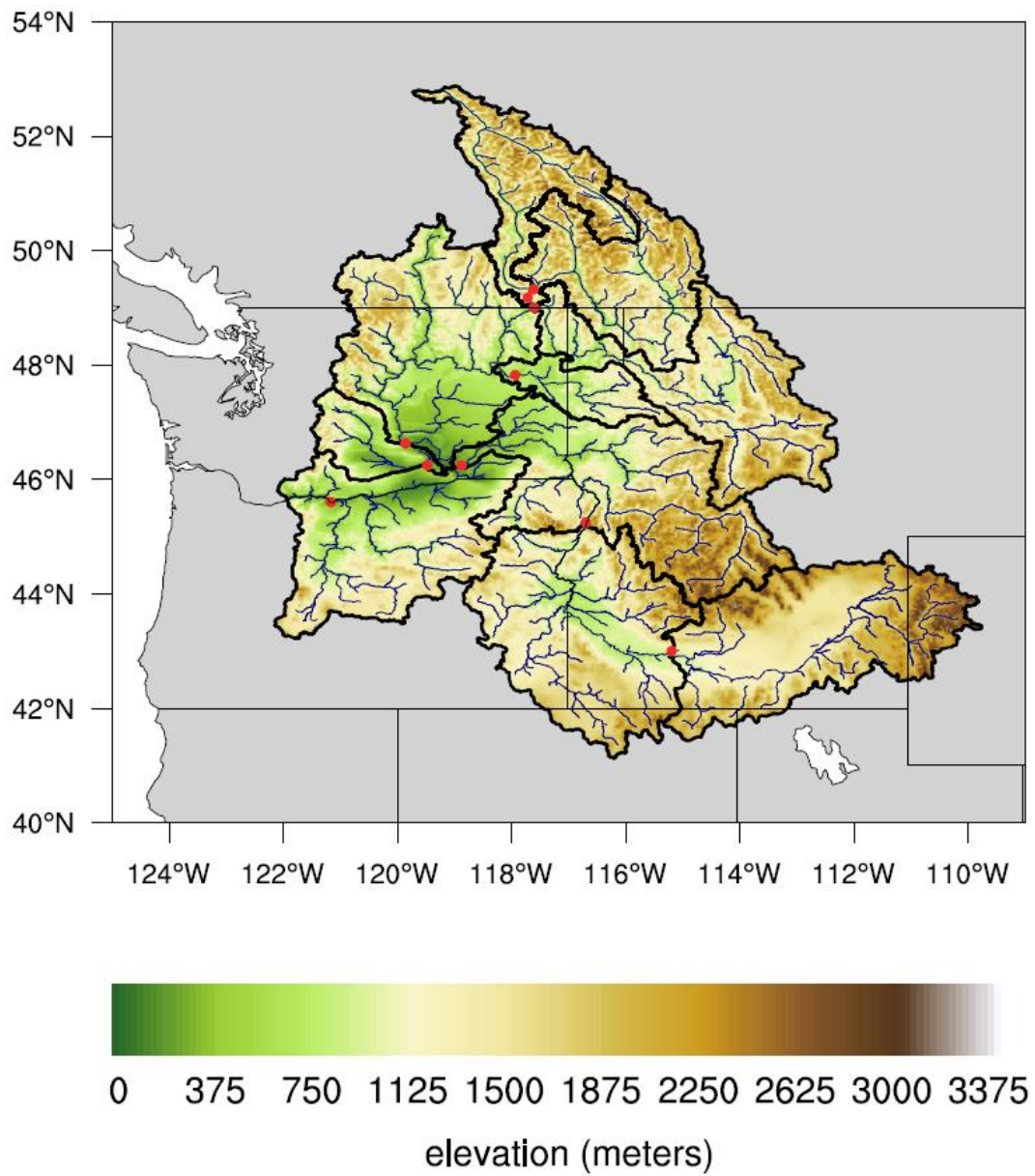


Figure 33. Locations of the Flowlines, Subbasins, and Farthest Downstream Stations Used with Topography.

Table 5. Nash-Sutcliffe Efficiency between Monthly and Logarithmic Monthly Observations and ALL MME-Livneh for the Downstream Stations.

VG=Very Good, G=Good, S=Satisfactory, US=Unsatisfactory

	TDA	IHR	HCD	SKHI	PRD
Monthly	0.6742, G	0.4684, US	0.4751, US	-0.0242, US	0.6733, G
Log Monthly	0.7995, VG	0.6800, G	0.5893, S	0.0028, US	0.7771, VG

	KIOW	LFL	WAT	BRI	MUC
Monthly	-0.5746, US	0.4371, US	0.0985, US	0.5037, S	0.5048, S
Log Monthly	0.5230, S	0.5887, S	0.5886, S	0.6870, G	0.6481, G

Metrics and Methods of Analysis

The following metrics were used to assess the annual and seasonal components of flow at each station: annual total flow, annual maximum flow, center of timing, and summer means. For center of timing, we used the center of mass as defined in Stewart et al. (2004) for month data. We used this definition rather than the day in which half of the annual total flow is exceeded due to the model's ability to more accurately reproduce the monthly streamflow in comparison to the daily

flow (Regonda et al., 2005; Maurer et al., 2007; Rauscher et al., 2008; Burn, 2008; Hidalgo et al., 2009; Wenger et al., 2010). Summer was defined as July – September to ensure summer started after the peak of the snowmelt flood each year.

Analyses were performed on the model mean of the three simulations produced using Livneh precipitation and temperature (i.e. CRUNCEP5-Livneh, GSWP3-Livneh, Princeton-Livneh). Analyses performed using the model mean of the simulations produced using CRUNCEP5, GSWP3, and Princeton are shown in the Supplementary Material. To assess the trends in the monthly climatologies and 4 metrics defined above, Theil-Sen was used to estimate the slope and Mann-Kendall to determine the significance using $\alpha = 0.10$ (Kendall, 1975; Mann, 1945; Sen, 1968; Theil, 1950; Burkey, 2006). As in Gudmundsson et al. (2017), we chose to use a correlation analysis as a first stage detection analysis. Pearson correlation coefficients between the observations and each model forcing (i.e. ALL, CLMT, CO2, NDEP, LULCC) were calculated, and to be detected, the correlation between the observations and a forcing had to be greater than the 97.5th percentile of the correlations between the observations and piControl segments. To form the distribution of correlations between the observations and piControl segments, 10,000 samples (with replacement) of size 3 were taken from the piControl segments and each sample was averaged. Samples of size 3 were taken to match the ensemble size of the ELM forcings.

The more in-depth, non-optimized version of D&A used in Forbes et al. (2018) was then implemented using the linear model

$$y = \beta_{CLMT}x_{CLMT} + \beta_{CO2}x_{CO2} + \beta_{NDEP}x_{NDEP} + \beta_{LULCC}x_{LULCC} + \varepsilon$$

where y denotes the observations, β_i is the scaling factor for forcing i , x_i is the model response to forcing i , and ε are the residuals. Unlike Forbes et al. (2018), uncertainty in the scaling factor estimates were calculated using the 43 piControl segments. Optimization of the signal-to-noise ratio was not used due to the limited number of available piControl segments. Using half of the segments for optimizing the observations and CLMT resulted in similar scaling factor estimates but led to wider uncertainty ranges due to using so few segments in the uncertainty calculation (not shown).

Results

Monthly Climatologies

The distribution of the monthly climatologies at The Dalles are shown in figure 34. For analyzing the monthly climatologies, the peak month was defined as the month with the greatest median and the period of peak flow began (ended) in the month that had a median greater than the previous (respectively, next) month's upper quartile. The period of peak flow for The Dalles occurred April – August with the peak flow occurring in June. The period of peak flow for all the subbasins started in March or April and ended in July or August with the peak flows occurring in either May or June. The subbasins with peaks ending in August were

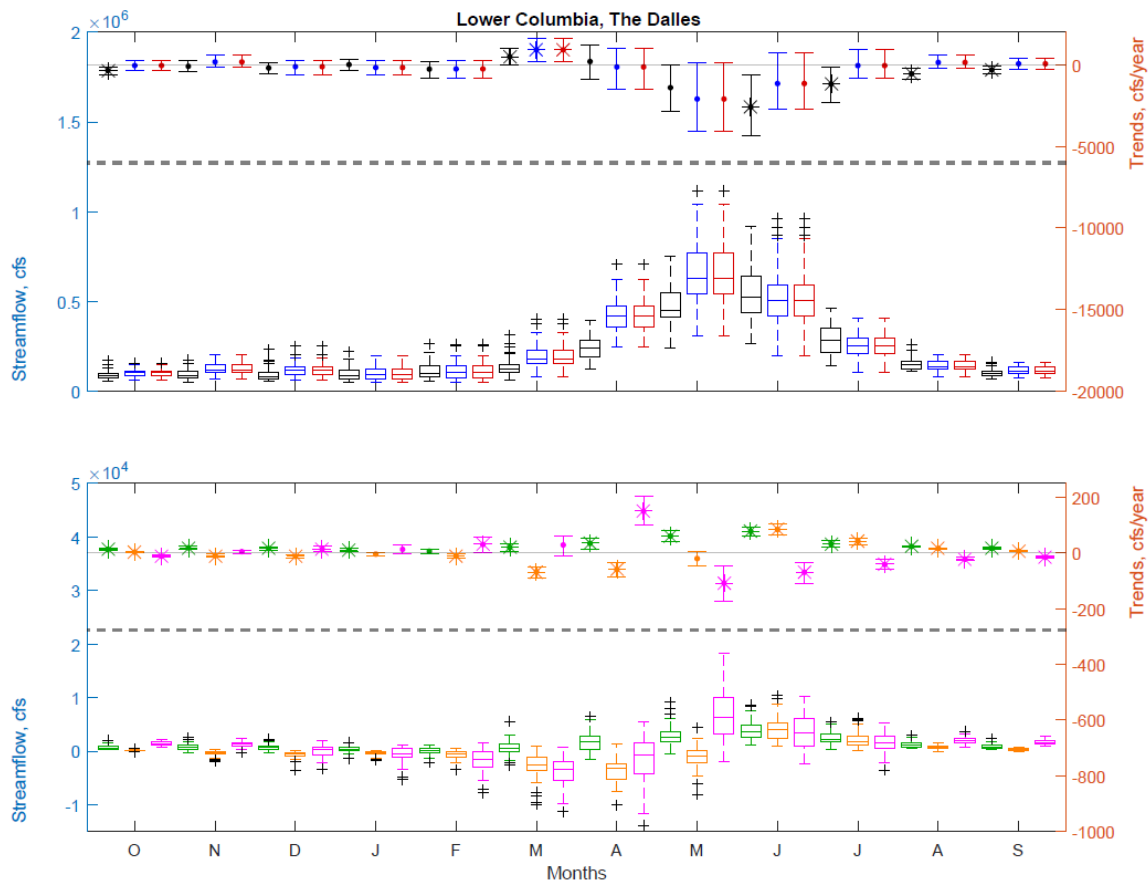


Figure 34. Monthly Climatologies at The Dalles for the Observations, ALL, and CLMT (Top) and CO₂, NDEP, and LULCC (Bottom). Box plots (left y-axis) and trends (right y-axis) for the observations (black), ALL (blue), and CLMT (red) are shown in the top figure. The bottom figure shows CO₂ (green), NDEP (orange), and LULCC (magenta). For the box plots, the horizontal lines of the box represent the lower quartile, median, and upper quartile. Lines extending from the box represent 1.5 times the interquartile range or the bottom and top 25% of the distribution. Data points outside 1.5 times the interquartile range are outliers and are shown by a “+” symbol. Trend values and corresponding 90% confidence intervals were estimated using Theil-Sen with significance determined by Mann-Kendall (Burkey, 2006).

Lower Columbia (The Dalles), Mid Columbia, Spokane, Pend Oreille, Kootenay, and Upper Columbia. Spokane, Pend Oreille, Kootenay, and Upper Columbia are the most northern independent subbasins, so the long period of the peak flow from the northern subbasins continued into the downstream subbasins (i.e. Lower and Mid Columbia). Even though June had the greatest median flow at The Dalles, it had the greatest significant decreasing trend for the period and was the start of significant decreases for the entire June – October season. This decreasing occurred generally consistently across all the subbasins (figures S37 – S45). These significant decreases led to significant decreases in the summer means. The only significant increase at The Dalles was in March, and while all the subbasins had an estimated increasing trend for March, only the trends for Kootenay and Upper Columbia were significant (figures S44 – S45). Following the estimated increases in March, April shows a mix of significant and insignificant positive and negative trends across the subbasins. However, all the subbasins but Upper Snake had a significant or insignificant decrease in May. Given the estimated positive trend in March and estimated negative trend in May/June, we can hypothesize that there has been a shift in the center of timing.

Monthly climatologies for CO₂, NDEP, and LULCC at The Dalles are shown in the bottom half of figure 34. Unsurprisingly, CO₂ increased in almost every month, and this is true for all subbasins. Excluding the subbasins covering the Snake River, the most interesting results come from the trends in NDEP and LULCC. Just as the observations showed significant decreasing trends for June – October, LULCC shows the same pattern but also includes May. NDEP has a pattern of significant trends for June – October as well, but rather than decreasing, NDEP had increases. In general, if a subbasin had a significant trend

for a particular month in both LULCC and NDEP, they had opposing signs. For most of the subbasins, the strongest positive trend for LULCC was in April and was followed by the strongest negative trend in May or June. However, for NDEP, the strongest negative trend was in March or April and was followed by the strongest positive trend in June. There are a couple other results which are interesting. The first being that specifically for Yakima, LULCC had significant positive trends for November – March. The other is significant negative (positive) trends for LULCC (NDEP) in December and January for Kootenay and Upper Columbia. Figures like figure 34 for the other subbasins are provided in the Supplementary Material (figure S37 – S45). Figure S46 shows the trend values at each subbasin for the annual totals, annual maximums, center of timing, and summer means.

D&A Results

The results for the correlation-based detection analysis are shown in figure 35. As expected, the correlation between the observations and ALL is greater than the 2.5th – 97.5th percentile of the correlations between the observations and piControl simulations for all 4 metrics and all 10 subbasins. Correspondingly, CLMT is detected for all metrics at all subbasins with correlation values approximately equal to those between the observations and ALL. The more informational results come from CO₂, NDEP, and LULCC. For the annual totals, LULCC was detected for Lower Columbia (The Dalles) and upstream throughout the Snake River subbasins. This holds for the annual maximums too, but also includes the Mid Columbia (above Yakima) and Pend Oreille subbasins. The Snake River and Pend Oreille subbasins are the most westerns subbasins and

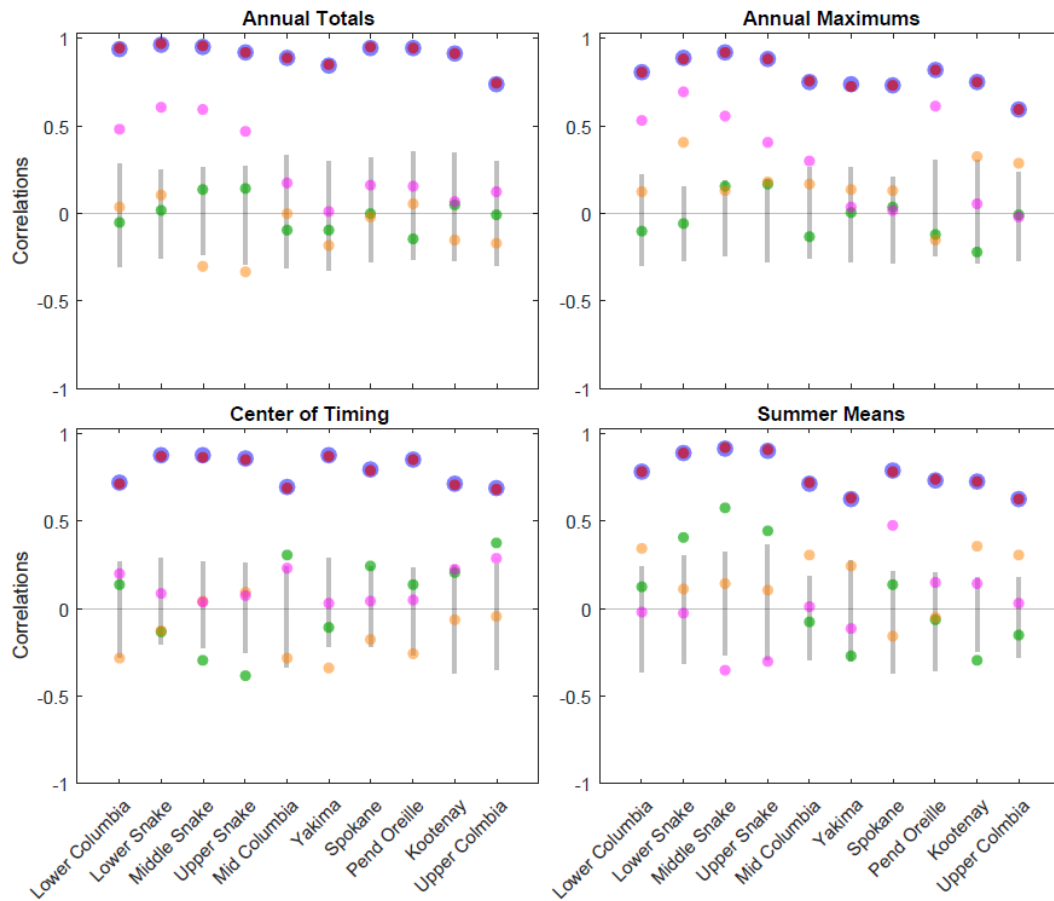


Figure 35. Correlation-Based Detection Analysis for Annual Totals, Annual Maximums, Center of Timing, and Summer Means for ALL, CLMT, CO₂, NDEP, and LULCC. Pearson correlation coefficients between the observations and ALL (blue), CLMT (red), CO₂ (green), NDEP (orange), and LULCC (magenta) for each subbasin. Light gray lines denote the 2.5th to 97.5th percentiles of the correlations between the observations and the piControl simulations.

share borders, so it is possible they are being influenced by the same driver of streamflow changes due to land use and land cover changes. NDEP is also detected for the annual maximums at the Lower Snake, Kootenay, and Upper Columbia subbasins. Other than CLMT, the only detected forcings for center of timing are CO₂ at Mid Columbia and CO₂ and LULCC at Upper Columbia. For the summer means, CO₂ was detected for all three Snake River subbasins, NDEP was detected for all three Columbia River subbasins, and LULCC was detected for Spokane.

D&A results using the linear regression methodology are shown in figure 36. For annual totals, annual maximums, center of timing, and summer means, 5, 0, 9, and 4 of the 10 subbasins show consistency between ALL and the observations, respectively. Results for the multifactor D&A are very similar to the results reported in Forbes et al. (2018) with changes being detected in CLMT, but for the other forcings, CO₂, NDEP, and LULCC, the scaling factors are too wide to be physically plausible. However, rather than only detecting the changes of streamflow in CLMT, the changes can be attributed to CLMT for many cases. More specifically, the changes in annual total streamflow can be attributed to CLMT for all subbasins except Upper Snake and Pend Oreille. Only detection of changes in CLMT were achieved for annual maximums due to overestimation in the model simulations. This overestimation is visible in the monthly climatology figures (figure 34, S37 – S45). Upper Columbia is the only subbasin for which the change in center of timing cannot be attributed to CLMT. For the summer means, only the changes can be attributed to CLMT for the Lower Columbia (The Dalles), Lower and Middle Snake, and Pend Oreille subbasins. One or more months from the July-August-September summer means were underestimated

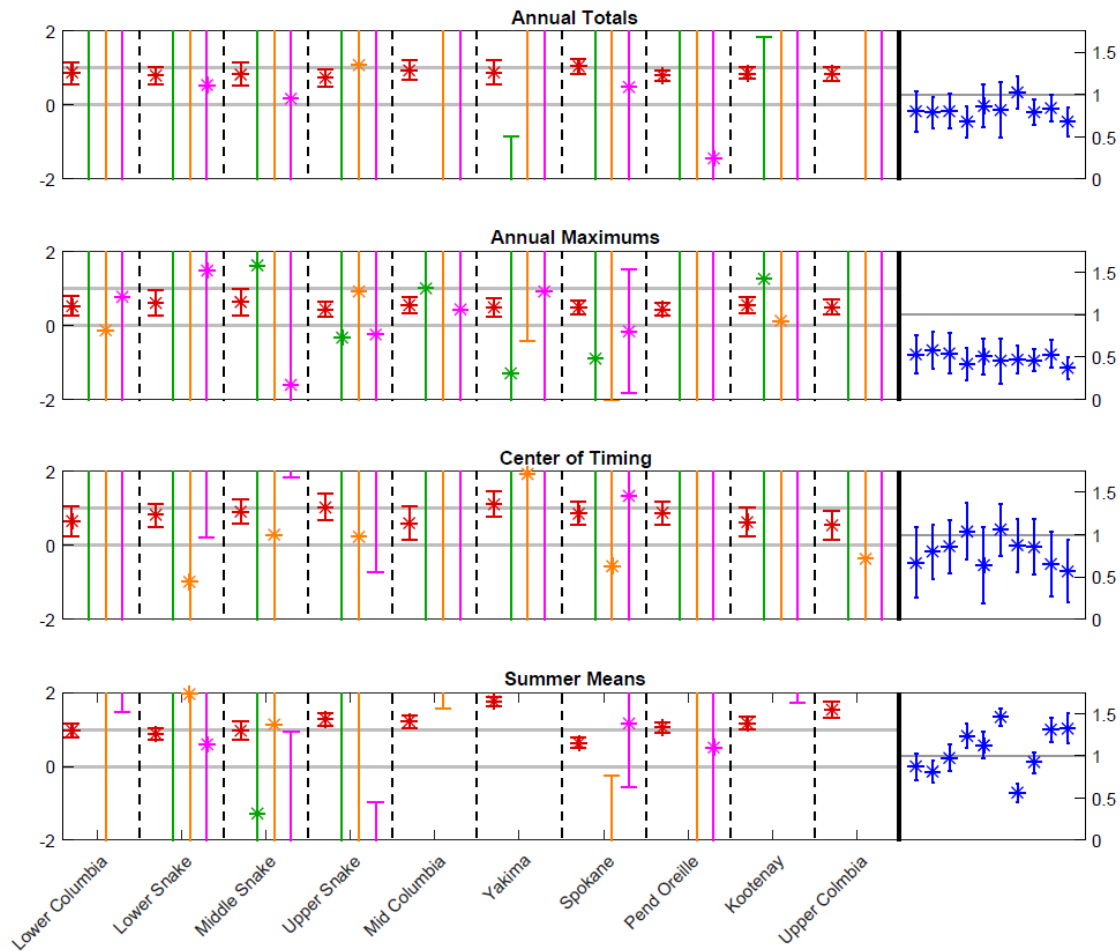


Figure 36. Scaling Factor Estimates and Corresponding 95% Confidence Intervals for Annual Totals, Annual Maximums, Center of Timing, and Summer Means using ALL and the Linear Combination of CLMT, CO2, NDEP, and LULCC. Scaling factors for CLMT, CO2, NDEP, and LULCC for each subbasin are shown on the left y-axis in red, green, orange, and magenta, respectively. Scaling factors for ALL (blue) are shown on the right y-axis in the same subbasin ordering. Light gray lines denote the values 0 and/or 1.

(overestimated) for the Upper Snake, Mid and Upper Columbia, and Kootenay (Yakima and Spokane) subbasins. In addition to the under or over estimation, the estimated trends in July, August, and September months for ALL were the opposite direction of the observations for the Upper Snake, Mid and Upper Columbia, and Yakima subbasins.

Discussion

Apart from the monthly climatologies showing a shift to an earlier center of timing and decreased flow June – October in the observations, NDEP and LULCC showed interesting changes in their distributions of flow. Excluding the Snake River subbasins, the decreasing trends within the observations for June – September were also found in LULCC. Unlike the observations, a shift in flow from May to April is more prominent in LULCC with 7 of the 10 subbasins having significant increasing in April and significant decreasing in May. For NDEP, the pattern of significance in the observations and LULCC in June – September is present, but rather than having significant decreasing, NDEP increased. Increases in streamflow due to nitrogen deposition could imply that the nitrogen saturation limit for the region has been met, and rather than increased vegetation growth due to nitrogen fertilization, the nitrogen level is restricting growth. The effects of land use and land cover change on streamflow caused a positive trend in the amount of flow for Yakima during November – March. The Yakima subbasin has a large percentage of agricultural land, but the amount of harvested cropland in Yakima County was relatively constant for 1925 – 2007 (Drennan 2013). However, the amount of pasture farmland in Yakima County has changed. There was an increase from approximately 800,000 acres in 1950 – 1955 to

approximately 1,500,000 in 1960 and then 1,700,000 acres in 1965. The acreage then decreased to approximately 1,350,000 by 1980 and remained somewhat steady through 2007. Within the changes of acreage were drastic changes in types of livestock. While the number of sheep and lambs decreased from 90,000 to 10,000 and hogs and pigs decreased from 17,000 to 500, cattle and calves increased from 90,000 to 220,000. There were also large fluctuations in the number of chickens with counts from 200,000 to 520,000. Large increases in fall and winter grazing livestock are important because they affect the amount of vegetative ground cover. Another interesting result is the significant trends in LULCC for the Kootenay and Yakima subbasins. LULCC is of particular interest in these subbasins because a large portion of the land is national parks.

The differences between the observations and the model simulated flow shown in the monthly climatologies is also present in the D&A results. The largest disagreement was found between the annual maximum flow. In some cases, the median maximum flow was shifted one month earlier than for the observations, but more importantly, the model simulated flow consistently overestimated the peak value. Thus, the scaling factors between the observations and ALL were less than one. For some of the subbasins, the July, August, and/or September flow was over- or under-estimated and the trend was in the opposing direction. There were also differences between the results found in the correlation-based D&A analysis for CO₂, NDEP, and LULCC in comparison to the regression-based methodology. This is due to the scale of the signals. Even though the pattern of change in CO₂, NDEP, or LULCC may correlate more closely with the observations than the observations correlated with the piControl runs, the strength of the signal is much smaller than the signal

found in the observations, ALL, and CLMT. An example of this difference in strength can be seen in the differences between the scales of the y-axes in figure 34.

The main limitation in the previous paper (Forbes et al., 2018) was the bias in the CRUNCEP precipitation driver used in the MSTMIP land surface model simulations. That limitation was overcome in this study by correcting the driver using the temperature and precipitation from Livneh. The regression-based D&A analysis was repeated using the MME of the simulations produced without the Livneh correction and are shown in figure S47. In comparison to the Livneh results (figure 36), the only metric for which ALL was more consistent with the observations is the annual maximums. For this paper, the main limitation was the limited availability of daily CMIP5 piControl simulations of runoff which were at least 58 years in length. In total, only 43 independent segments were available. So rather than using half of the segments for prewhitening the observations, ALL, and CLMT and the other half for estimating the uncertainty in the scaling factors, all 43 segments were used in estimating the scaling factor uncertainty. Analysis performed using the prewhitening led to approximately the same scaling factor estimates for ALL and CLMT, but the limited number of segments available for calculating the uncertainty of these estimates led to wider confidence intervals (not shown).

Conclusions

On average, the annual total streamflow for the Columbia River Basin decreased by approximately 15% between 1951 – 2008. Of that 15%, roughly 77% was during the June – October months with 40% solely in June (peak flow) and 31%

in July – September (summer mean). More specifically, flow in June has declined by 28% on average. The fact that these declines are in 5 consecutive months during the year is particularly worrisome. On average, these 5 months provided 49% of the annual total flow with June providing 22% itself. While the Columbia River Basin does have reservoirs, they cannot be used to hold and distribute water during the second half of the year as their primary purpose is to prevent storm surge flooding. Other than supplying municipal water sources, the lack of water during the second half of the year effects the amount of water for irrigation, summer salmon runs, and fall power generation. Even though on average, Upper Columbia saw an increase of 41% and 31% in March and April respectively, these months only accounted for approximately 9% of the total annual flow. Alternatively, June accounts for 25% and declined by 17%. The D&A analysis shows that these changes in annual total, center of timing of, and summer mean streamflow can be attributed to changing climate and variability. Some of the patterns in streamflow changes due to CO₂, nitrogen deposition, and land use and land cover change were detected using the correlation-based analysis, but the signals were not strong enough to be detected using the regression-based analysis.

References

Barnett T P *et al.* 2008 Human-induced changes in the hydrology of the Western United States. *Science* **319(5866)** 1080–3 doi:10.1126/science.1152538

Berghuijs W R, Woods R A and Hrachowitz M 2014 A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change* **4** 583–6 doi:10.1038/nclimate2246

Bonfils C *et al.* 2008 Detection and attribution of temperature changes in the mountainous western United States. *J. Climate* **21** 6404–24 doi:10.1175/2008JCLI2397.1

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, Washington. USA.
<http://www.mathworks.com/matlabcentral/fileexchange/authors/23983>

Burn D H 2008 Climatic influences on streamflow timing in the headwaters of the Mackenzie River Basin. *J. Hydrol.* **352** 225–38 doi:10.1016/j.hyrol.2008.01.019

David C H, Habets F, Maidment D R and Yang Z -L 2011a RAPID applied to the SIM-France model. *Hydrological Processes* **25(22)** 3412–25 doi:10.1002/hyp.8070

David C H, Maidment D R, Niu G -Y, Yang Z -L, Habets F and Eijkhout V 2011b River network routing on the NHDPlus dataset. *Journal of Hydrometeorology* **12(5)** 913–34 doi:10.1175/2011JHM1345.1

David C H, Yang Z -L and Hong S 2013 Regional-scale river flow modeling using off-the-shelf runoff products, thousands of mapped rivers and hundreds of stream flow gauges. *Environmental Modelling & Software* **42** 116–32 doi:10.1016/j.envsoft.2012.12.011

Drennan M 2013 Agricultural History of Yakima County.
<http://www.yakimacounty.us/DocumentCenter/View/2375/Agricultural-History-of-Yakima-County-PDF>

Feng X, Porporato A and Rodriguez-Iturbe I 2013 Changes in rainfall seasonality in the tropics. *Nature Climate Change* **3** 811–5 doi:10.1038/nclimate1907

Follum M L, Tavakoly A A, Niemann J D and Snow A D 2016 AutoRAPID: a model for prompt streamflow estimation and flood inundation mapping over regional to continental extents. *Journal of the American Water Resources Association* **53(2)** 280–99 doi:10.1111/1752-1688.12476

Forbes W L *et al.* 2018 Contribution of environmental forcings to US runoff changes for the period 1950 – 2010. *Environ. Res. Lett.* **13** 054023 doi:10.1088/1748-9326/aabb41

Fyfe J C *et al.* 2017 Large near-term projected snowpack loss over the western United States. *Nat. Commun.* **8** 14996 doi:10.1038/ncomms14966

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. *Nature Geoscience* **7** 796-800 doi:10.1038/ngeo2263

Gudmundsson L, Seneviratne S I and Zhang X 2017 Anthropogenic climate change detected in European renewable freshwater resources. *Nature Climate Change* **7** 813-6 doi:10.1038/NCLIMATE3416

Hidalgo H G *et al.* 2009 Detection and attribution of streamflow timing changes to climate change in the Western United States. *J. Climate* **22** 3838-55 doi:10.1175/2009JCLI2470.1

Horizon Systems Corporation 2007 National Hydrography Dataset Plus: Documentation. http://www.horizon-systems.com/NHDPlus/NHDPlusV2_documentation.php

Krause P, Boyle D P and Base F 2005 Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* **5** 89–97 doi:10.5194/adgeo-5-89-2005

Mao J *et al.* 2016 Human-induced greening of the northern extratropical land surface. *Nature Climate Change* **6** 959–63 doi:10.1038/nclimate3056

Maurer E P, Stewart I T, Bonfils C, Duffy P B and Cayan D R 2007 Detection, attribution, and sensitivity of trends toward earlier streamflow in the Sierra Nevada. *J. Geophys. Res.* **112** D11118 doi:10.1029/2006JD008088

Meixner T *et al.* 2016 Implications of projected climate change for groundwater recharge in the western United States. *J. Hydrol.* **534** 124–38 doi:10.1016/j.jhyrol.2015.12.027

Moriasi D N, Arnold J G, Van Liew M W, Bingner R L, Harmel R D and Veith T L 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* **50(3)** 885–900 doi:10.13031/2013.23153

Mote P W, Li S, Lettenmaier D P, Xiao M and Engel R 2018 Dramatic declines in snowpack in the western US. *npj Climate and Atmospheric Science* **1** doi:10.1038/s41612-018-0012-1

Najafi M R, Zweirs F and Gillet N 2017a Attribution of the observed spring snowpack decline in British Columbia to anthropogenic climate change. *J. Climate* **30(11)** 4113–30 doi:10.1175/JCLI-D-16-0189.1

Najafi M R, Zweirs F W and Gillet N P 2017b Attribution of observed streamflow changes in key British Columbia drainage basins. *Geophys. Res. Lett.* **44** 11012–20 doi:10.1002/2017GL075016

Pierce D W *et al.* 2008 Attribution of declining Western U.S. snowpack to human effects. *J. Climate* **21** 6425–44 doi:10.1175/2008JCLI2405.1

Pytlak E *et al.* 2018 Climate and Hydrology Datasets for RMJOC Long-Term Planning Studies: Second Edition (RMJOC-II) Part I: Hydroclimate Projections and Analyses. <https://www.bpa.gov/p/Generation/Hydro/hydro/cc/RMJOC-II-Report-Part-I.pdf>

Rauscher S A, Pal J S, Diffenbaugh N S and Benedetti M M 2008 Future changes in snowmelt-driven runoff timing over the western US. *Geophys. Res. Lett.* **35** L16703 doi:10.1029/2008GL034424

Regonda S K, Rajagopalan B, Clark M and Pitlick J 2005 Seasonal cycle shifts in hydroclimatology over the western United States. *J. Climate* **18** 372–84 doi:10.1175/JCLI-3272.1

Safeeq M, Shukla S, Arismendi I, Grant G E, Lewis S L and Nolin A 2016 Influence of winter season climate variability on snow-precipitation ratio in the western United States. *Int. J. Climatol.* **36** 3175–90 doi:10.1002/joc.4545

Snow A D, Christensen S D, Swain N R, Nelson E J, Ames D P, Jones N L, Deng D, Noman N S, David C H, Pappenberger F and Zoster E 2016 A high-resolution national-scale hydrologic forecast system from a global ensemble land surface model. *Journal of the American Water Resources Association* **52(4)** 950–64 doi:10.1111/1752-1688.12434

Stewart I T, Cayan D R and Dettinger M D 2004 Changes toward earlier streamflow timing across western North America. *J. Clim* **18** 1136–55 doi:10.1175/JCLI3321.1

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. *Journal of the American Water Resources Association* doi:10.1111/1752-1688.12456

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 Resources Research* **46** W09513 doi:10.1029/2009WR008839

Zhou T, Voisin N, Leng G, Huang M and Kraucunas I 2018 Sensitivity of regulated flow regimes to climate change in the western United States. *J. Hydrometeor.* **19** 499–515 doi:10.1175/JHM-D-17-0095.1

Zhu Z *et al.* 2016 Greening of the Earth and its drivers. *Nature Climate Change* **6** 791–5 doi:10.1038/nclimate3004

Supplementary Material

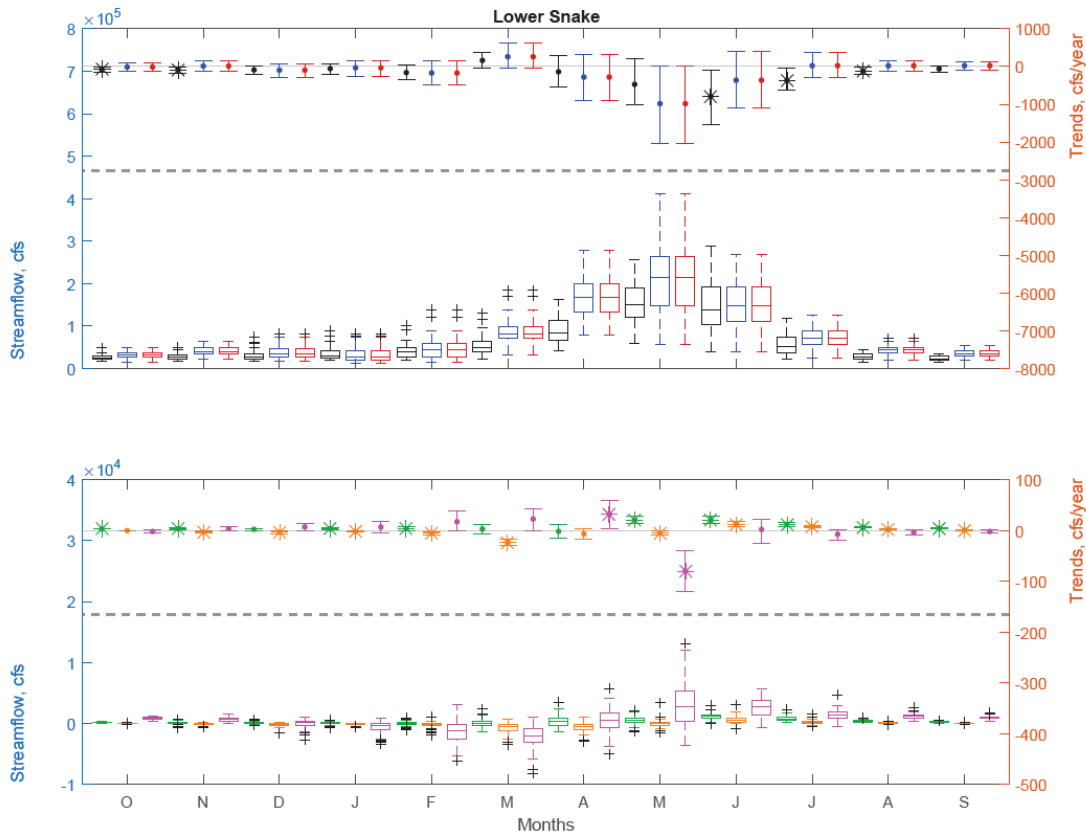


Figure 37. Monthly Climatologies at Lower Snake for the Observations, ALL, and CLMT (Top) and CO₂, NDEP, and LULCC (Bottom). Box plots (left y-axis) and trends (right y-axis) for the observations (black), ALL (blue), and CLMT (red) are in the top figure. The bottom figure shows CO₂ (green), NDEP (orange), and LULCC (magenta). For the box plots, the horizontal lines of the box represent the lower quartile, median, and upper quartile. Lines extending from the box represent 1.5 times the IQR. Data points outside 1.5 times the interquartile range are outliers and are shown by a “+” symbol. Trend values and corresponding 90% confidence intervals were estimated using Theil-Sen with significance determined by Mann-Kendall and denoted by asterisks.

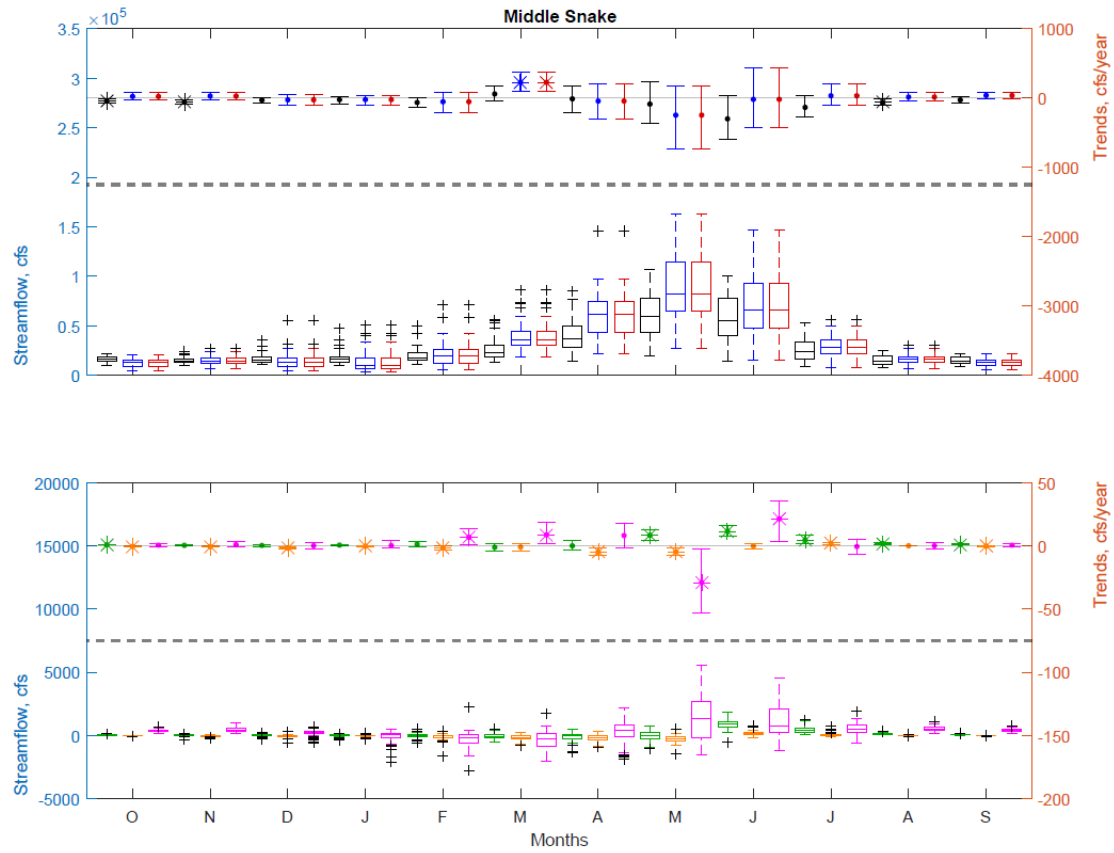


Figure 38. Same as figure 37 but for Middle Snake.

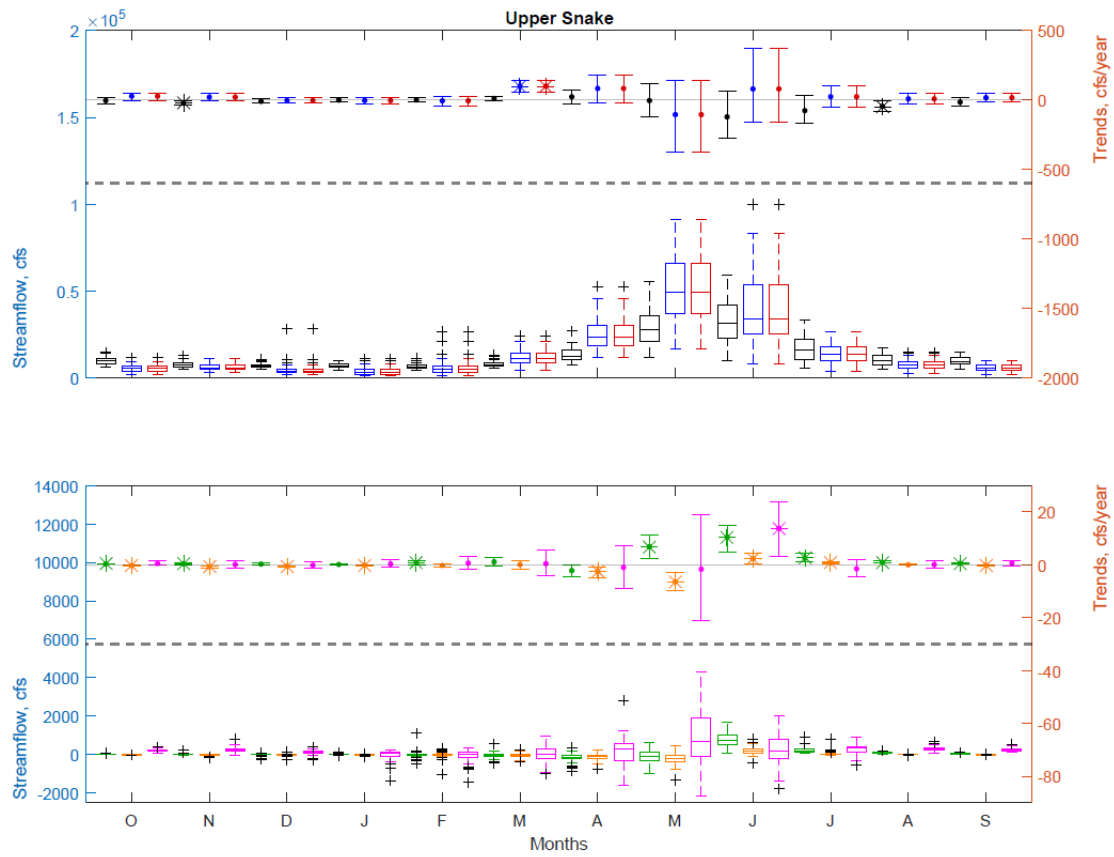


Figure 39. Same as figure 37 but for Upper Snake.

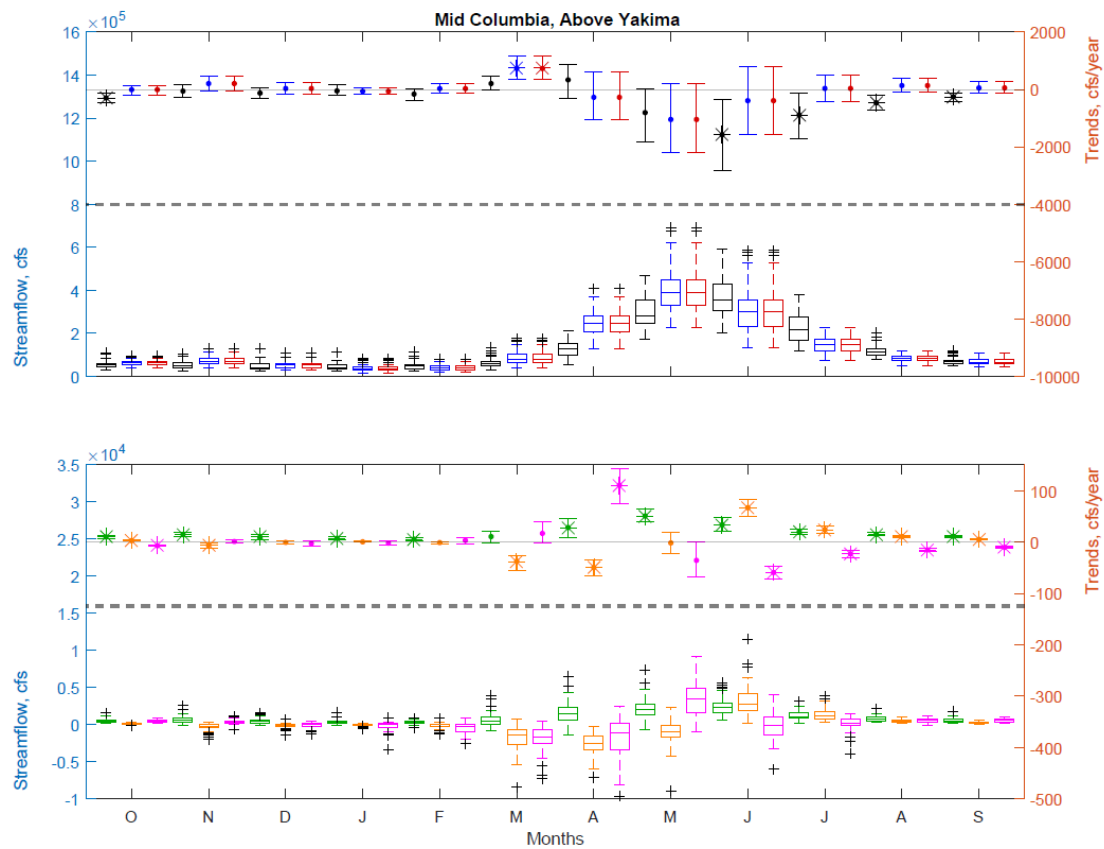


Figure 40. Same as figure 37 but for Mid Columbia.

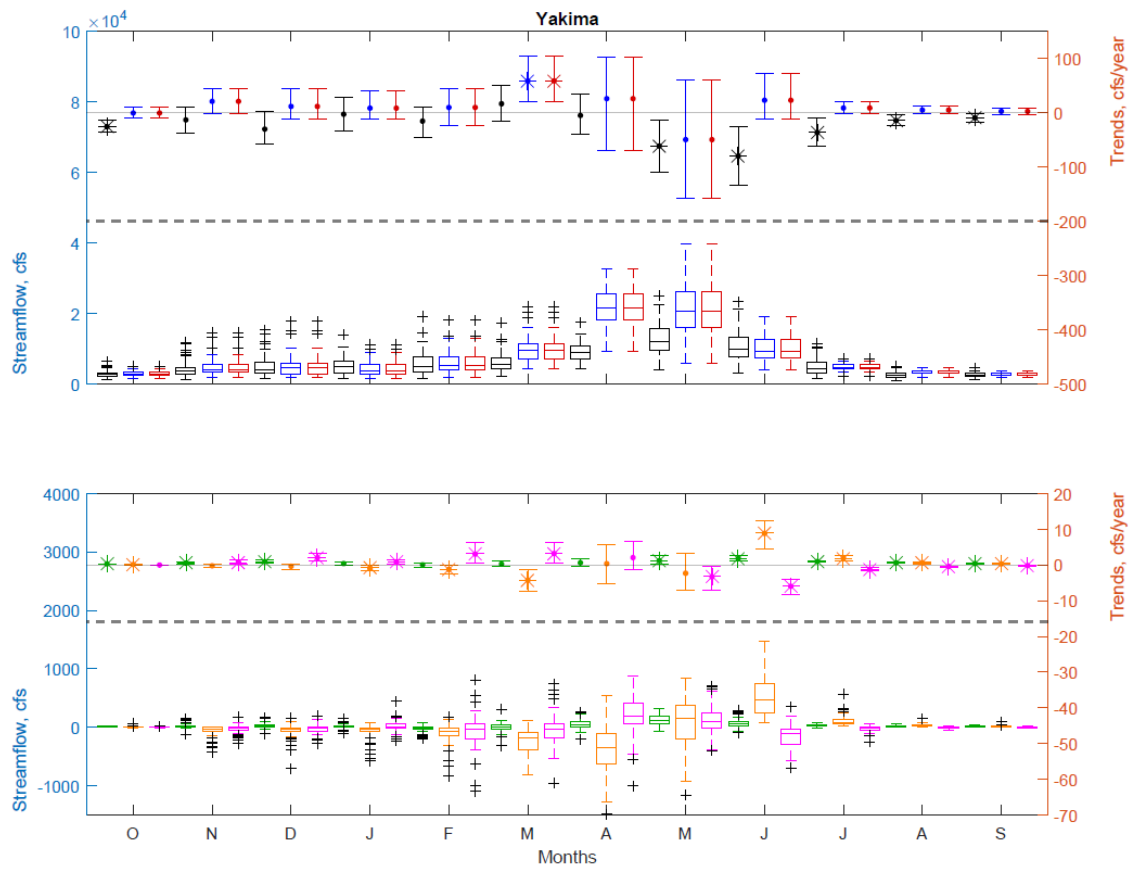


Figure 41. Same as figure 37 but for Yakima.

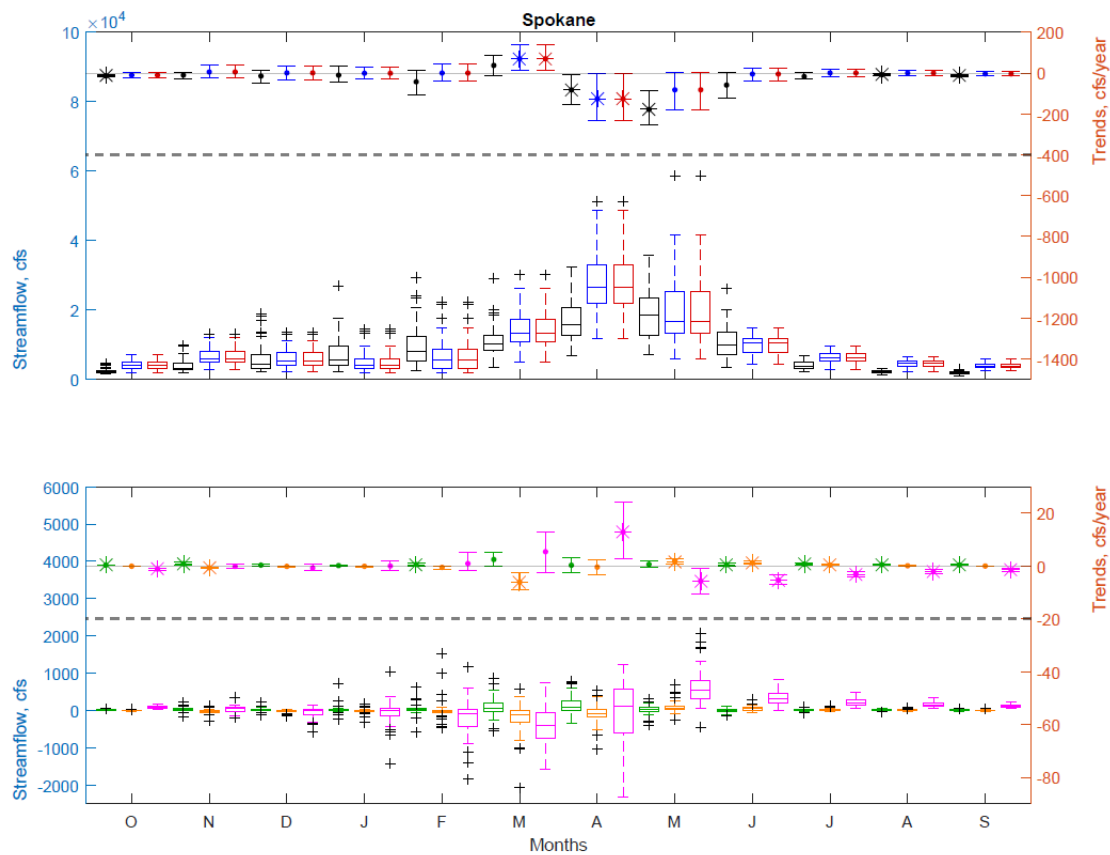


Figure 42. Same as figure 37 but for Spokane.

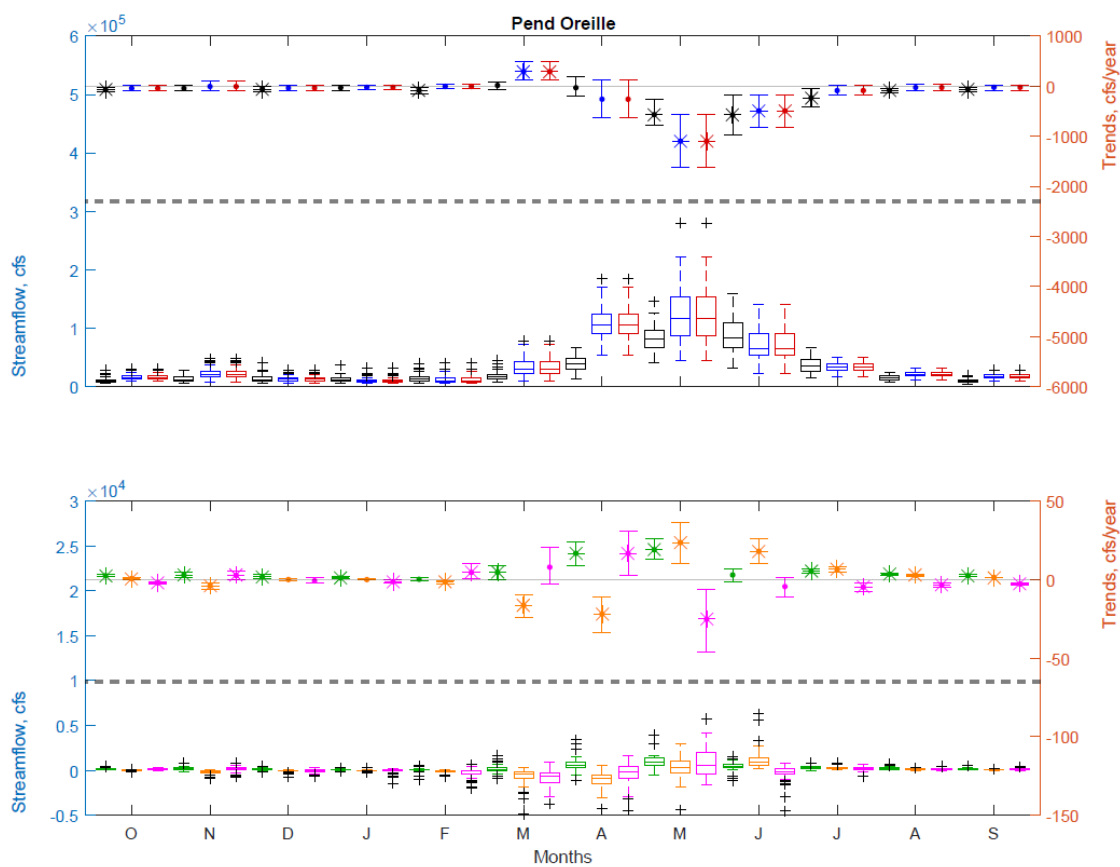


Figure 43. Same as figure 37 but for Pend Oreille.

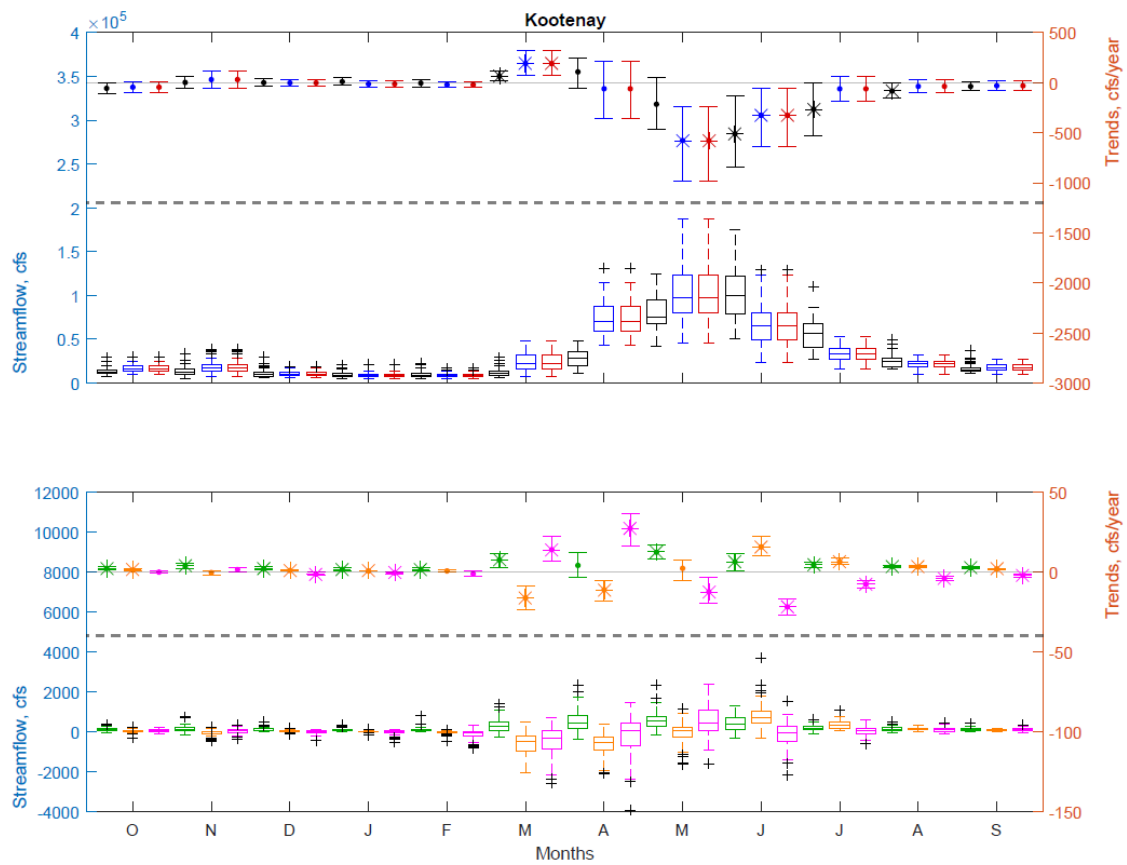


Figure 44. Same as figure 37 but for Kootenay.

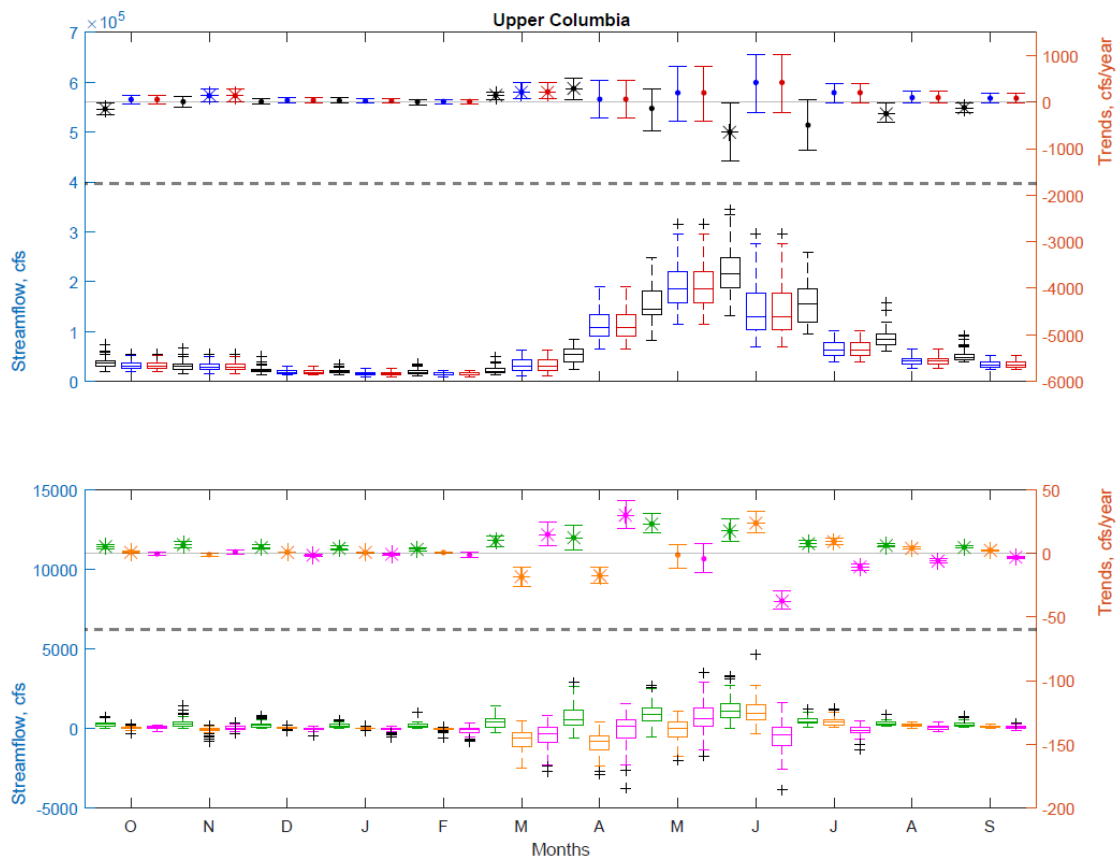


Figure 45. Same as figure 37 but for Upper Columbia.

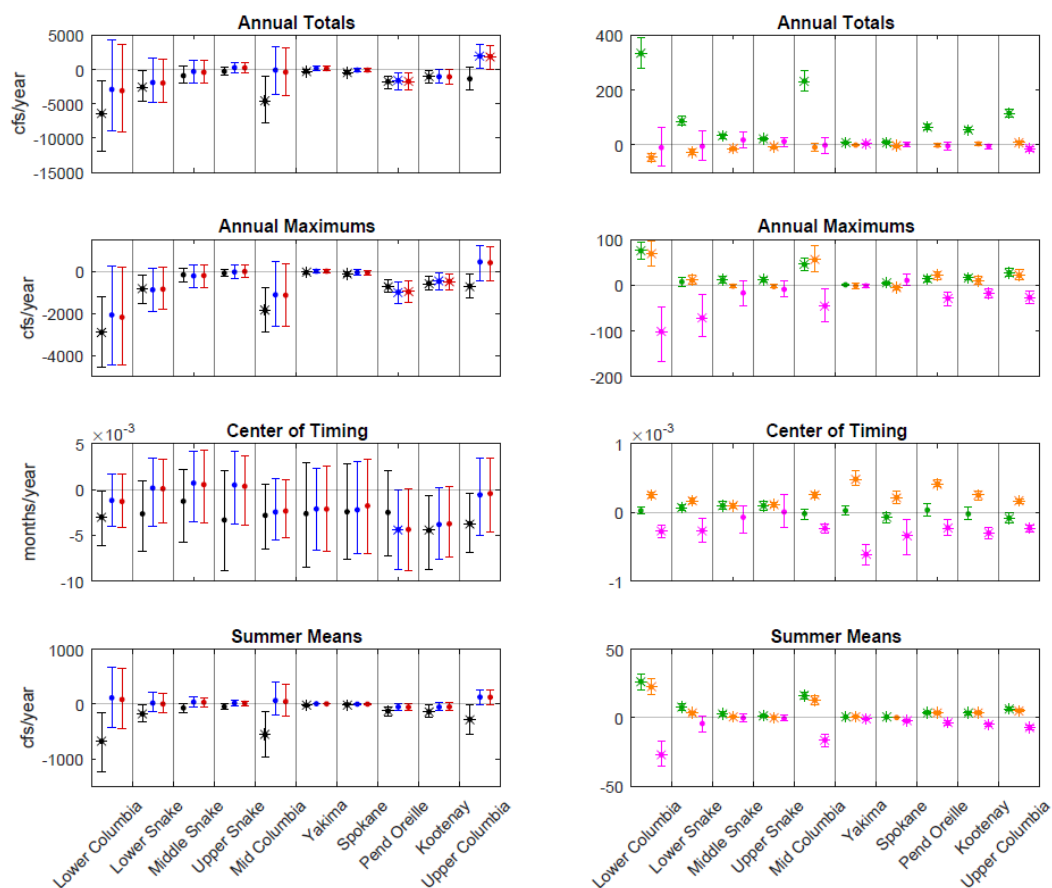


Figure 46. Regional Trends for Annual Totals, Annual Maximums, Center of Timing, and Seasonal Means. Trend values for the observations (black), ALL (blue), CLMT (red), CO2 (green), NDEP (orange), and LULCC (magenta). Trend values and corresponding 90% confidence intervals were estimated using Theil-Sen (Burkey, 2006). Significant trends are denoted by asterisks and were determined by Mann-Kendall.

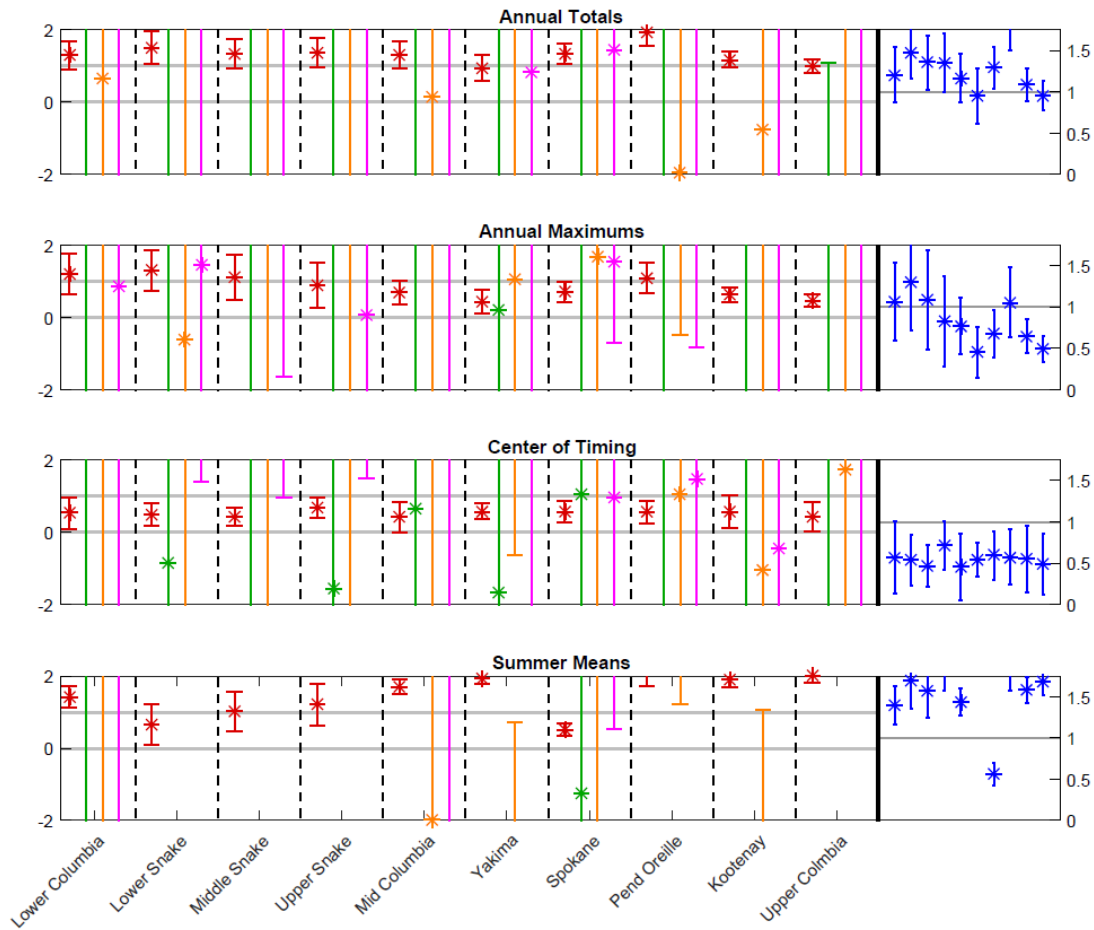


Figure 47. Scaling Factor Estimates and Corresponding 95% Confidence Intervals for Annual Totals, Annual Maximums, Center of Timing, and Summer Means using the MME of ALL and the Linear Combination of the MMEs of CLMT, CO2, NDEP, and LULCC using the Simulations Produced Using CRUNCEP5, GSWP3, and Princeton. Scaling factors for CLMT, CO2, NDEP, and LULCC for each subbasin are shown on the left y-axis in red, green, orange, and magenta, respectively. Scaling factors for ALL (blue) are shown on the right y-axis in the same subbasin ordering. Light gray lines denote the values 0 and/or 1.

CONCLUSION

Detecting changes in climate and systems effected by climate and attributing the causes of the change are important for our own well-being. While the idea of determining if patterns of change within external forcing are consistent with the change in observations seems simple, the noise due to contamination from natural internal variability, measurement error, model error, and sampling uncertainty can make isolating the patterns difficult. Natural internal variability is estimated using preindustrial control simulations, measurement error is generally negligible when using more recent observations, model error can be at least partially accounted for by using the ensemble mean of multiple perturbations from multiple models, and as always, sampling uncertainty can be minimized by collecting a bigger sample. Given that there is only one Earth, we cannot take multiple samples of the observations, but multiple samples of the external forcings can be taken by using more model simulations. This seems simple enough, but in reality, coupled general circulation models take months to run. The more viable option is uncoupled models which take approximately 10% the computational time and 50% the computational resources required by coupled models. However, using uncoupled models brings up a new difficulty: they do not offer the preindustrial control simulations used for estimating the natural internal variability. Generally, detection and attribution analysis is completed using a linear model of the form $y = \beta X + \varepsilon$ where y denotes the pattern of change in the observations, β are the scaling factors (i.e. regression coefficients), X is the collection of patterns of change due to external forcings, and ε is the natural internal variability. In this model, it is assumed that the patterns of change in both

the observations and external forcings have been prewhitened using the preindustrial control simulations in order to at least partially remove the noise due to natural internal variability, and thus, boosting the signal-to-noise ratio. In Chapters 1 and 2, uncoupled land surface model simulations were used for completing detection and attribution analyses on runoff in the United States during 1950 – 2010 and streamflow in the Columbia River Basin during 1951 – 2008. Without the availability of preindustrial control runs in the study of runoff in the United States, a slightly different linear model was used. While the equation used looked the same symbolically, the symbols represented slightly different things. y still represented the pattern of change in the observations, but it was the original, noise contaminated pattern. The same was true for X when considering the pattern of change in ALL and CLMT. Due to the semi-factorial experimental design used by the land surface models and simulation differencing, CO₂, NDEP, and LULCC did not include this contamination. Given that the patterns of change were not decontaminated, ε represented the residuals due to modeling error rather than natural internal variability, and thus, had to be checked for autocorrelation to ensure the assumptions required by Ordinary Least Squares regression were met. A similar method was used in the study of streamflow in the Columbia River Basin. However, for this study, preindustrial control simulations were used to estimate the uncertainty range of each scaling factor. Analysis was also produced (but not shown) using half of the preindustrial control segments to prewhiten the patterns of change and half to estimate the uncertainty range. This analysis led to similar values for the estimated scaling factors, but the uncertainty range was wide due to the small sample available for estimating the uncertainty. Using the method described above, the changes in US runoff were detected in

CLMT. The changes could not be attributed to CLMT due to the underestimated signal in the simulations. This underestimation was possibly due to biases in the CRUNCEP precipitation driver used in the simulations. To overcome this limitation, new simulations were produced for the CRB region. Instead of solely using CRUNCEP5 meteorological drivers, two groups of three simulations were produced using CRUNCEP5, GSWP3, Princeton, and Livneh. In the first set of three, the original drivers from CRUNCEP5, GSWP3, and Princeton were each used to produce a simulation. The second set of three was the exact same except the precipitation and temperature drivers were replaced by those from Livneh. The ALL simulation from the MME of the three Livneh simulations were more consistent with the observations for the annual totals, center of timing, and summer means, and the non-Livneh MME was more consistent for the annual maximums. The changes in all 4 metrics (i.e. annual totals, annual maximums, center of timing, and summer means) were attributed to CLMT. In both the US and CRB studies, the signals from CO₂, NDEP, and LULCC were not strong enough to be differentiated from the natural internal variability using the regression-based methodology. However, using a correlation-based detection method, the changes within the CRB streamflow could be detected in CO₂, NDEP, and LULCC for some cases.

In order to overcome the limitations that arise when using uncoupled models in detection and attribution, new methodologies or simulations must be established. A new methodology could consist of using particle filtering in place of the prewhitening using the preindustrial control simulations. As for new simulations, it is possible that adaptive methods could be used in order to recreate the coupled model preindustrial control runs using uncoupled models.

For example, maybe coupled model preindustrial control simulations of the meteorological drivers needed to run uncoupled land surface models could be used to create a pseudo preindustrial control simulation for land surface models.

VITA

Whitney L. Forbes, the daughter of David and Sherry Hunt, was born in Johnson City, TN. She is the second of two daughters: Brandi. She attended Happy Valley Elementary School in Carter County, TN then moved to T. A. Dugger Middle School and Elizabethton High School in Elizabethton, TN. After graduation, she attended East Tennessee State University where she obtained a Bachelor of Science degree in Mathematics with a minor in secondary education in 2012 and a Master of Science degree in Mathematics in 2014. She accepted a graduate research assistantship at The University of Tennessee, Knoxville, in the Industrial Engineering Doctoral Program. Her dissertation research was completed in conjunction with the Climate Change Science Institute at the Oak Ridge National Laboratory. Whitney graduated with a Doctor of Philosophy degree in Industrial Engineering in July 2018. She is planning on continuing her research career at the Oak Ridge National Laboratory.