

Spring 5-2018

# SEASONAL IMPACTS OF CLIMATE CHANGE ON FUTURE PEAK RIVER DISCHARGE IN THE U.S. NORTHEAST

Christina Wu

Follow this and additional works at: [https://scholarworks.umass.edu/cee\\_ewre](https://scholarworks.umass.edu/cee_ewre)



Part of the [Environmental Engineering Commons](#)

---

Wu, Christina, "SEASONAL IMPACTS OF CLIMATE CHANGE ON FUTURE PEAK RIVER DISCHARGE IN THE U.S. NORTHEAST" (2018). *Environmental & Water Resources Engineering Masters Projects*. 91.  
<https://doi.org/10.7275/fge1-ah65>

This Article is brought to you for free and open access by the Civil and Environmental Engineering at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Environmental & Water Resources Engineering Masters Projects by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

**SEASONAL IMPACTS OF CLIMATE CHANGE ON FUTURE PEAK RIVER  
DISCHARGE IN THE U.S. NORTHEAST**

A Master's Project Presented

by

CHRISTINA Y. WU

Submitted to the Graduate School of the  
University of Massachusetts Amherst in partial fulfillment  
of the requirements for the degree of

MASTER OF SCIENCE  
CIVIL ENGINEERING

May 2018

Environmental & Water Resources Engineering

**SEASONAL IMPACTS OF CLIMATE CHANGE ON FUTURE PEAK RIVER  
DISCHARGE IN THE U.S. NORTHEAST**

A Masters Project Presented

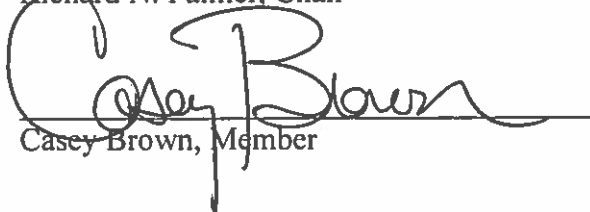
by

CHRISTINA Y. WU


Approved as to style and content by:



Richard N. Palmer, Chair



Casey Brown, Member



Richard N. Palmer, Department Head  
Civil & Environmental Engineering

## ABSTRACT

### SEASONAL IMPACTS OF CLIMATE CHANGE ON FUTURE PEAK RIVER DISCHARGE; IN SUPPORT OF THE U.S. NORTHEAST REGIONAL HYDROLOGIC MODELING

MAY 2018

CHRISTINA Y. WU, B.S., UNIVERSITY OF MASSACHUSETTS AMHERST

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. Richard N. Palmer

This research uses projected changes in future precipitation to calculate the changes in the magnitude, frequency, and timing of streamflow, particularly peak annual flows, in the U.S. Northeast through 20 representative watersheds. Temperature and precipitation data on a 3-hourly time step from five climate projection from the North American Regional Climate Change Assessment Program (NARCCAP) are downscaled and bias-corrected using the Spatial Downscaling and Bias Correction (SDBC) method. These data are used to force a regional hydrological model (WRF-Hydro) to estimate daily future streamflow. The changes in magnitude at various return intervals of peak flow events are determined through the comparison between peak annual flow values during the historical period (1968-1999) and during the future period (2038-2070). The frequencies of high daily streamflow in each month are evaluated using a peak-over-threshold (POT) analysis of both high precipitation days and high streamflow days to understand the correlation (if any) between the two in this particular region. The results indicate an overall average increase of 10%, 15%, and 18 % in the 2-year, 50-year, and 100-year return interval magnitude of peak floods in the U.S. Northeast region, respectively. The POT analysis reveals increases in the number of extreme precipitation days during the winter months (DJF) which is expected to result in higher peaks in streamflow. This correlation is less apparent during the summer (JJA) months, suggesting a significant role of soil moisture and snowpack. The degree of climate change impacts vary by season, leading to differing flow regimes. Shifts in the seasonality of future peak flow events are observed in the results and further explain the changes in flood magnitudes and frequencies. They suggest similar trends in the inundation processes that directly influence soil moisture; consequently exacerbating flood and drought events that require new adaptation and mitigation strategies in the region.

# TABLE OF CONTENTS

	Page
ABSTRACT .....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
1 Introduction .....	1
1.1 Infrastructure Vulnerability .....	1
1.2 Project Objectives.....	2
1.3 Literature Review .....	3
1.3.1 Global and Regional Climate Models.....	5
1.3.2 Modeling Climate Change Impacts on Flooding Events .....	7
2 Study Area & data.....	10
2.1 Study Region.....	10
2.2 Climate Data .....	11
2.3 River Discharge & Channel Data.....	13
3 Methods & Experimental Design .....	13
3.1 Spatial Downscaling & Bias Correction of NARCCAP Climate Data .....	13
3.2 Estimating River Discharge .....	15
3.2.1 WRF-Hydro .....	15
3.2.2 Calibration of WRF-Hydro .....	16
3.2.3 Watershed Selection.....	17
3.2.4 Routing Scheme.....	19
3.3 Watershed-Based Analysis of Results.....	20
3.3.1 Model Performance.....	20
3.3.1.1 Spatial Downscaling & Bias Correction Skill on NARCCAP Data .....	20
3.3.1.2 WRF- Hydro Simulations of Peak Annual Flows .....	21
3.3.2 Investigating Hydrologic Responses to Heavy Precipitation Events.....	<b>Error! Bookmark not defined.</b>
3.3.2.1 Peaks over Threshold (POT).....	22
3.3.2.2 Climate Indices.....	23
4 Results .....	23
4.1 Modeling Performance & Projections .....	24
4.1.1 Skill of SDBC Method.....	24
4.1.2 Peak Annual Flows .....	28
4.1.2.1 Accuracy of Simulated Historical Peak Annual Flows .....	28

4.1.2.2	Future Peak Annual Flows .....	30
4.2	Hydrologic Responses to Climate Change .....	32
4.2.1	Projected Changes in Monthly Precipitation.....	32
4.2.2	Seasonality of Peak Annual Flows .....	34
4.2.3	Seasonal Peaks over Threshold .....	37
4.2.4	ClimDEX Indices.....	39
5	Discussion & Conclusions.....	41
6	Recommendations & Future Work .....	45
	References .....	47
	WRF-HYDRO CALIBRTION RESULTS.....	50
	MUSKINGUM ROUTING SCHEME – PYTHON CODE	<b>Error! Bookmark not defined.</b>
	SUPPLEMENTARY RESULTS FIGURES .....	53

## LIST OF TABLES

Table	Page
Table 4-1: Selected Watershed Characteristics.....	19
Table 5-1. GEV Goodness of Fit Metrics.....	30
Table 5-2. Regional Average Percent Change in Peak Annual Flows.....	31

## LIST OF FIGURES

Figure	Page
Figure 2-1 Map of Modeled Region.....	11
Figure 3-1 Modular Calling Structure of WRF-Hydro (from [ <i>Gochis et al.</i> , 2015]).....	16
Figure 3-2 Locations of Selected Watersheds .....	18
Figure 4-1 SDBC Results of NARCCAP – Change in Mean Monthly Temperature (°C) (Average of 5 NARCCAP SDBC Models) .....	25
Figure 4-2 SDBC Results of NARCCAP – Change in Total Monthly Precipitation (%) (Average of 5 NARCCAP SDBC Models) .....	25
Figure 4-3 SDBC Method Mean Monthly Error in Mean Daily Temperature.....	26
Figure 4-4 Mean Percent Bias in Daily Precipitation Accumulation from the SDBC Method during the Calibration and Validation Periods. ....	27
Figure 4-5 Percent Change in Peak Annual Flow .....	32
Figure 4-6 Monthly Annual Total Precipitation .....	34
Figure 4-7: CDF of Julian Day of Annual Maximum Flow and Precipitation .....	35
Figure 4-8 Seasonal Precipitation and Streamflow POT (50 <sup>th</sup> Percentile Threshold) .....	38
Figure 4-9 Winter Precipitation and Streamflow POT.....	39
Figure 4-10: Max 5-Day Precipitation .....	41



# 1 INTRODUCTION

The continued release of greenhouse gases and their impacts on increases in temperature continue to threaten both the built and natural environment. An understanding of the hydrologic cycle and its interactions with the built and natural environment is crucial to estimate risk and vulnerabilities. Hydrologic extremes, such as floods, have amounted to millions of dollars in damages to infrastructure including culvert washouts and bridge overtopping, as well as the displacement of communities and the restriction of access to emergency resources. In the past, flood risk was determined based on historical peak streamflow data at a particular location, and fitted to a distribution to extrapolate return intervals of interest. However, practitioners often encounter a paucity of historical data at locations of interest and the duration of that data may be limited. Streamflow data are fundamentally important because they provide the best estimate of the true natural system. Addressing future hydrologic risk in a changing climate becomes more challenging, in part due to the question of stationarity, among other modeling uncertainties.

## 1.1 Infrastructure Vulnerability

Climate science researchers forecast increases in temperature and extreme hydrologic events in the Northeast U.S. and throughout much of the globe [Melillo *et al.*, 2014a]. Increased levels of flooding, as demonstrated by weather events like Hurricane Sandy [Kaufman *et al.*, 2012] suggest the need for more effective mitigation strategies and increased investments in our infrastructure. In the U.S. Northeast, where temperatures and precipitation events are expected to increase in both magnitude and frequency, infrastructure such as highways, bridges, and culverts - particularly those built

in river flood plains - are becoming more vulnerable, as their ability to pass extreme flows diminishes [*Karmalkar et al.*, 2017].

Bridges and culverts are designed based on a defined storm frequency (perhaps the 50 or 100-year storm). Unfortunately, there is often incomplete or no data available for specific locations and “rules of thumb” or simple nomographs are used to estimate the design flow. Designs based on past experience may prove inadequate, especially since the magnitudes of peak flows have shown a positive trend in the Northeastern U.S. as a result of the projected increasing intensity of precipitation events [*Demaria et al.*, 2016]. The failure of bridges and culverts due to high flows often impacts transportation services. Failures limit access to critical services like hospitals and safe zones during emergencies. In addition to transportation considerations, improperly sized bridges and culverts can also create barriers for fish passage and other aquatic wildlife.

## **1.2 Project Objectives**

This research estimates changes in the peak annual streamflow (PAF), changes in seasonality, and peaks-over-thresholds (POT) associated with extreme meteorological events in the Northeast U.S. by forcing climate model projections through a regional hydrological model and extracting results for a selected number of watersheds. PAF is defined as the maximum annual discharge on a particular stream for each year within a time period; these annual values are fitted to a distribution from which probabilities (also known as return intervals) are extracted. Seasonality refers to the day or month of occurrence of a particular event, i.e. how the timing of extreme precipitation and streamflow events is changing. POT is defined here as the number of days above a particular magnitude provided by a historical percentile threshold. This information can

support decision makers and stakeholders in addressing and identifying vulnerable infrastructure within the region. For the U.S. Northeast, a physically-based hydrologic model is developed. The hydrology model is calibrated and verified for a set of selected watersheds. Simulated temperature and precipitation data from five different regional climate models are used to force the hydrology model to obtain five different realizations of future streamflow conditions. The return period and value of annual peak flows are calculated using a generalized extreme values distribution (GEV) statistical model. The results are assessed for changes from a historical period (1968-1999) to a future period (2038-2070) in various metrics including 1) peak annual flow, 2) precipitation and streamflow seasonality (i.e. timing), 3) seasonal precipitation and streamflow peaks-over-threshold (POT), and 4) climate indices such as monthly precipitation accumulation and monthly maximum 5-day precipitation. This adds depth to a traditional return period analysis in an attempt to develop wide-ranging and useful understanding of flood for stakeholders and decision makers in the region.

### **1.3 Literature Review**

In the Northeast U.S. region, climate change projections have identified challenges to environmental, social, and economic systems due to the increases in extreme events [Horton and Yohe, 2014]. Recent analysis of the latest climate model simulations from CMIP5 (Coupled Model Intercomparison Project Phase 5) indicate that the U.S. Northeast is projected to experience the fastest warming; reaching an increase of 3°C when the global average increase reaches 2°C [Karmalkar *et al.*, 2017]. In the first half of the 20<sup>th</sup> century, the region has experienced an average increase of 0.14°F per decade, followed by an increase to a rate of 0.5°F per decade in the latter half [Dawson,

2013]. Based on the Clausius-Clapeyron relationship, increases in temperature have the potential to increase precipitation [Wasko and Sharma, 2017]. An analysis of precipitation trends in the region show a significant increase in the quantity of precipitation during very heavy events, coupled with an increasing warming trend [Horton and Yohe, 2014]. This agrees with other analyses that have shown that increases in the number of extreme precipitation events in this region [Parr and Wang, 2014; Ivancic and Shaw, 2015]. The projected accelerated warming of the U.S. Northeast compared to other U.S. regions is reinforced by the U.S. Northeast experiencing the greatest increase of extreme precipitation in the past half-century [Parr and Wang, 2014].

Increases in the magnitude and frequency of extreme precipitation events have encouraged researchers to translate changes in the atmosphere to changes on the Earth's surface, measured by soil moisture and river discharge. A trend analysis conducted on historical mean values in the Connecticut River Basin indicates, in addition to increasing precipitation, increases in discharge, runoff ratios, and soil moisture [Parr and Wang, 2014]. However, a global assessment of the relationship between extreme precipitation and streamflow report a lack of evidence in a strong correlation between the two [Wasko and Sharma, 2017]. Similarly, an analysis of 390 watersheds across the U.S. further supports insignificant trends in high streamflow despite increasing heavy precipitation events in the U.S. Northeast [Ivancic and Shaw, 2015]. The disagreement over the relationship between extreme precipitation and peak streamflow (almost analogous to the differences in the rate of warming that the U.S. Northeast is expected to experience in comparison to other regions) may be explained by the very different trends and hydrologic regimes of this particular region.

Furthermore, it is essential that asset and risk management include assessments of possible hydrologic futures that the built environment may experience, especially in the wake of aging and inadequate infrastructure [Palmer *et al.*, 2013]. Many state agencies and municipalities have begun to form adaptation plans that incorporate climate impact studies in long-term planning efforts, where “the key is to link adaptation strategies with capital improvement cycles and adjustment of plans to incorporate emerging climate projections” [Melillo *et al.*, 2014b].

The impacts of climate change on society and its resources are modeled on various temporal and spatial scales. For example, future climates are estimated using atmosphere-ocean general circulation models (GCMs) that are driven with projections of greenhouse gas emissions to produce realizations of changes in meteorological metrics such as temperature and precipitation across the globe and over time. Likewise, impacts on water resources and flood infrastructure can be estimated with hydrological models that utilize the outputs of GCMs to produce estimates of streamflow, runoff, and soil moisture. However, one challenge in producing accurate and serviceable estimates lies in both model selection and their appropriate spatial and temporal scales [Gutmann *et al.*, 2014].

### **1.3.1 Global and Regional Climate Models**

Impact assessments at the regional or watershed scale require finer spatial scales than that of those provided by GCMs or other coarse-scale climate models. On a gridded spatial scale, often referred to as distributed, hydrological models can be categorized into lumped parameter models, or physically-based models [DeVantier and Feldman, 1993]. A distributed model uses a regularly spaced grid in which each grid is treated as an

independent entity that is resolved within the model. This often creates the necessity of dynamical and statistical downscaling models when the gridded inputs are on a larger scale than the hydrological model.

Statistical downscaling takes a more direct approach by deriving empirical relationships between the large-scale and fine-scale variables, thereby bypassing the need to solve for mass, energy, and momentum transfers in physically based processes [*Jang and Kavvas, 2015*]. There are numerous procedures used to reproduce a particular variable in a simulated meteorological dataset. A simple statistical downscaling technique is the delta-change method that superimposes trends found in the climate model to transform a historical time series; therefore the original variability remains intact [*Wood et al., 2004; Kay et al., 2009*]. The disadvantage of this method is the very strong assumption of stationarity, where the transformed time series is heavily dependent on the order and variability of events in the original historical time series, which may have a significant impact on modeled hydrologic processes [*Arnell, 2003*]. This finding echoes the general criticisms of statistical downscaling in that relationships between the large-scale and fine-scale product are derived using observed and simulated historical data that are then applied to the future period [*Tryhorn and Degaetano, 2011*].

Due to the computational demands of dynamical downscaling, statistical downscaling is the more widely used approach [*Jang and Kavvas, 2015*]. Qiao et al. 2014 compared two sets of outputs from a Variable Infiltration Capacity (VIC) hydrological model; one using an RCM dataset provided by the North American Regional Climate Change Assessment (NARCCAP), and the other using a CMIP3 GCM dataset statistically downscaled using the Bias-Corrected and Spatially Downscaled

(BCSD) method. Their results found that the RCM climate data yielded greater accuracy in the VIC outputs due to the dataset's ability to capture more instances of meso-scale driven convective rainfall. Tryhorn et al. 2011 conducted a similar study in which the HadCM3 GCM dataset downscaled using the statistical BCSD method was compared to the dynamically downscaled NARCCAP dataset (HadRM3) over the U.S. Northeast. Their study focused on the technique's ability to reproduce extreme precipitation. They found that the dynamically downscaled (HadRM3) dataset overestimated mean and extreme precipitation, while the statistically downscaled GCM dataset (HadCM3) only produced larger errors with higher return intervals [Tryhorn and Degaetano, 2011]. These two studies' respective preference of downscaling method is influenced by their ultimate objective; where Qiao et al. 2014's inclination towards dynamic downscaling aims to produce estimates of streamflow, while the preferred statistical downscaling in Tryhorn et al. 2011 aims to produce estimates of extreme precipitation. Therefore, the selection of a downscaling method for impact assessments is dependent on location, and in this case, the ability to produce accurate streamflow estimates in the U.S. Northeast.

### **1.3.2 Modeling Climate Change Impacts on Flooding Events**

Despite the vast array of hydrological models and relevant parameters, studies have shown that a single hydrological model, when driven with various climate datasets that utilize different downscaling or bias correction techniques, can produce differing outcomes [Hwang and Graham, 2014; Qiao et al., 2014]. For example, Hwang et al. 2014 investigated three different statistical downscaling methods on multiple GCMs for their skill in estimating streamflow through a hydrological model in west-central Florida, 1) a modified BCSD method (BCSD\_daily) in which daily values were used instead of

the traditional use of monthly values, 2) a spatial-disaggregation and bias-correction (SDBC) method in which the steps in the modified BCSD method are reversed, and 3) a bias-correction and stochastic analog method (BCSA). Their results noted an overestimation of evapotranspiration due to the spatially smoothed light precipitation events produced by the BCSD\_daily method, leading to underestimated mean streamflow. The reverse method, or SDBC, also suffers from highly spatially correlated precipitation events, but is able to capture the observed temporal standard deviation of daily precipitation; leading to an overestimation of high streamflow during the wet season, but accurate estimates of daily streamflow [Hwang and Graham, 2014]. This illustrates the significance of accurate temporal representations of climate variables and indices in modeling extreme hydrological events, where the smoothed temporal distributions of climate variables influence estimates of high streamflow.

Conducting a more standardized analysis, Werner et al. 2016 evaluated seven downscaling methods using a collection of climate indices (ClimDEX), and compared the 3-day and 7-day peak streamflow using VIC in the Peace River Basin in BC, Canada. To compare with the previously mentioned studies, the BCCI downscaling method (analogous to the SDBC method) passed a greater number of Pearson's correlation and KS tests overall for both ClimDEX indices and hydrological extremes. A combination of the BCCI method with a bias-corrected constructed analogs (BCCA) method (termed BCCAQ) was able to surpass all other downscaling methods tested for reproducing extreme hydrological events by removing overly spatially correlated meteorological events [Werner and Cannon, 2016]. This Werner et al. 2016 study also identifies select ClimDEX indices, including the number of consecutive dry days, total precipitation, and



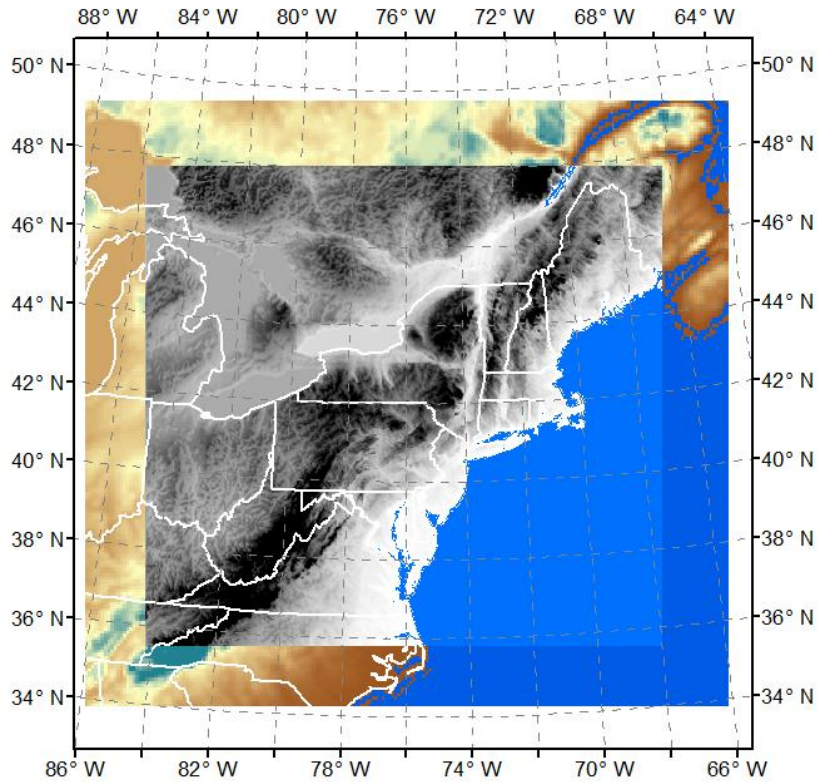
maximum 5-day precipitation, that are characteristic of the downscaling methods that produced good representations of the hydrologic extremes. These select ClimDEX indices allude to the significance of soil moisture in estimating peak streamflow.

According to the Intergovernmental Panel on Climate Change (IPCC), studies have been inconsistent their findings of the direction of change in the magnitude and frequency of floods across the globe [*Mallakpour and Villarini, 2015*]. For example, Mallakpour et al. 2015 investigated the trends in annual peak discharge in 774 watersheds in the central U.S. region, and found no statistically significant trends in most of the watersheds. The 20% of watersheds that did show an increasing trend in flood magnitude are located in urban areas, where imperious surfaces do not provide any dampening of runoff, are most likely due increases in built-up and impervious areas. Despite the lack of increasing flood magnitude, the frequency of flood events has increased [*Mallakpour and Villarini, 2015*]. Ivancic et al. 2015 examined the relationship between extreme precipitation and peak streamflow in 290 watersheds across the U.S.; finding that extreme precipitation has a greater correlation with extreme discharge during wet soil conditions than during dry soil conditions. The literature suggests that previous desires to bias correct climate simulations of extreme precipitation events to match extreme climate observations may have misguided their relationship to extreme flood events, as soil moisture is being recognized in having perhaps a greater influence [*Small et al., 2006; Qiao et al., 2014; Ivancic and Shaw, 2015*].

## 2 STUDY AREA & DATA

### 2.1 Study Region

The U.S. Northeast is the region of interest in this thesis covering 15 states (Figure 2-1). Although mostly dominated by forest, this region is home to approximately 41% of the national population, creating dense urban areas among swaths of agricultural and ecologically dense environments [Horton *et al.*, 2014]. This region receives approximately 40 inches per year of precipitation, characterizing it with warm and humid summers and cold and wet winters [Horton *et al.*, 2014]. With dense regions of both urban and rural areas, runoff attenuation varies at either end of the urban/rural spectrum; mostly attributed to impervious surfaces and land types within watersheds. As described in the following sections, the simulated gridded climate data is spatially downscaled and bias corrected to be forced through the hydrological model over this entire region. The resulting runoff outputs are input through a Muskingum channel routing scheme written in Python in the selected watersheds. Data analysis is performed in each watershed to evaluate the performance of the climate downscaling method, the hydrologic model, and projected changes in flood magnitude, seasonality, and frequency.



**Figure 2-1 Map of Modeled Region**

## 2.2 Climate Data

The climate datasets used as gridded observed time-series are provided by Livneh *et al.* 2013, and by the North American Land Data Assimilation System (NLDAS). Five gridded climate model projections are provided by the North American Regional Climate Change Assessment Program (NARCCAP); the variables within the datasets include forecasted temperature and precipitation and are used to force the hydrologic model.

The simulated NARCCAP data are the results of atmosphere-ocean general circulation models (AOGCMs) forced with the SRES A2 emissions scenario that are then forced through a set of regional climate models (RCMs). This produces gridded precipitation and temperature data at a 50 kilometer (km) spatial resolution [NARCCAP, 2007]. Additional calibration of the RCMs to improve accuracy in representing climate

variables over North America involves an initial run with NCEP Reanalysis II data [NARCCAP, 2007]. The five NARCCAP datasets used in this research are: CGCM3\_CRCM, CCSM\_CRCM, GFDL\_HRM3, CGCM3\_RCM3, and GFDL\_RCM3; where the first term represents the GCM model used and the second represents the RCM applied. Although five climate projection datasets are used, these are combinations of three GCMs and three RCMs; in essence, these datasets provide 5 different realizations from 3 climate projections. These data span a historical (1968-1999) and future (2038-2070) time period. The use of these five simulated datasets as forcing for the hydrological model requires a finer spatial resolution than the provided 50 km resolution; therefore, the Statistical Downscaling and Bias Correction (SDBC) method is applied to downscale the gridded climate projections to a  $1/16^\circ$  scale (i.e. the spatial scale of the gridded observations).

The gridded observational dataset provided by Livneh *et al.* 2013 is used to guide the SDBC method and enables the simulated NARCCAP datasets to be resampled to the spatial resolution of such dataset. This observational dataset includes gridded estimates of temperature, precipitation, wind, and radiation. These were derived using algorithms that take advantage of daily temperature and precipitation observations from 20,000 NOAA Cooperative Observer (COOP) stations and disaggregated to 3-hourly time steps [Livneh *et al.*, 2013]. This historical observed dataset spans from 1915 – 2011 and is at a  $1/16^\circ$  spatial resolution on a 3-hourly time step. These data are hosted at the University of Washington, and is available for download online (<ftp://ftp.hydro.washington.edu/pub/blivneh/CONUS/>).

The assimilated NLDAS dataset also contains sub-daily units of the modeling parameters required for the hydrologic model, including radiation, surface pressure, humidity, temperature, and precipitation [Mitchell, 2004]. This dataset provides climate data for preliminary hydrologic model runs to ‘ramp-up’ and calibrate the model outputs to observed USGS data. More information on the NLDAS dataset is available in Mitchell et al. 2004.

### **2.3 River Discharge & Channel Data**

Observations of daily streamflow are obtained from U.S. Geological Survey (USGS) gages at each watershed outlet. These observed streamflow time series include the historical period of 1968-1999. These data are used to compare to initial hydrological model simulations of NLDAS data, as well as the GEV distribution of estimated streamflow from the five NARCCAP driven simulations.

The river routing method relies on the National Hydrography Dataset Plus (NHDPlus) dataset to provide a network of flowlines within each selected watershed. The NHDPlus network is used to geo-locate the stream-lines for which simulations of runoff are routed.

## **3 METHODS & EXPERIMENTAL DESIGN**

### **3.1 Spatial Downscaling & Bias Correction of NARCCAP Climate Data**

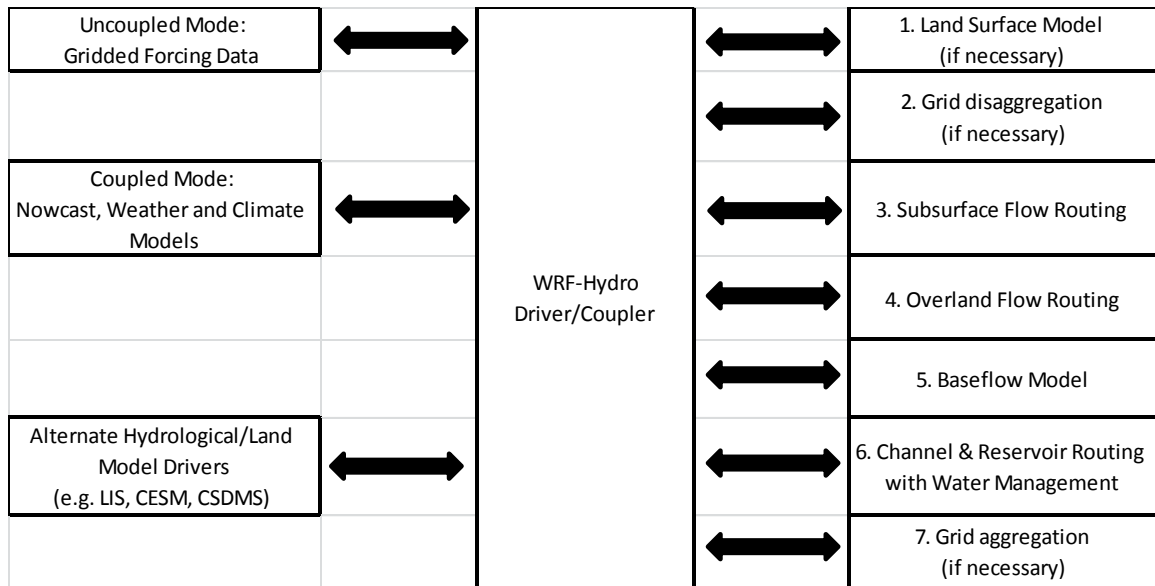
The spatial downscaling and bias correction (SDBC) procedure is performed on the five climate datasets provided by the North American Regional Climate Change Assessment Program (NARCCAP) following a similar procedure to Hwang et al. 2014. This procedure consists of three basic steps; 1) resample the simulated NARCCAP

datasets to the spatial resolution of the observed dataset, 2) bias correct the data via quantile mapping, and 3) disaggregate the results from daily to sub-daily. The initial temporal resolution of both datasets are 3-hourly; however, the Livneh et al. 2013 gridded observational precipitation dataset are provided as daily values equally divided into 3-hour time series. The first step, resampling the simulated datasets to the desired spatial resolution, employs the inverse-distance-weighting (IDW) of the four nearest NARCCAP simulation points to each gridded observation point for each 3-hour time-step. Once the NARCCAP simulation points and observed points represent the same location, a direct comparison of monthly CDFs within the calibration time period (1968-1988) can be made; this process is the bias correction step and is performed on a daily time-step. The factors used in correcting the quantiles within the temperature and precipitation CDFs are derived based on total daily precipitation accumulation and mean daily temperature during the calibration period (1968-1988). The results are both evaluated in the validation period (1989-1999) and applied to the projection period (2038-2070). Lastly, the bias-corrected simulated daily time-series are disaggregated into a 3-hourly time step. Disaggregation factors are derived from the 3-hourly signal within the spatially downscaled and bias-corrected NARCCAP simulated time-series. The factors used in the daily bias correction process (multiplication for precipitation and addition for temperature) throughout the daily simulated time-series are again used on the 3-hourly data. This preserves the sub-daily signal provided by the NARCCAP datasets. The results from the five simulated NARCCAP datasets, after the SDBC method are applied, are used to force the hydrologic model.

## 3.2 Estimating River Discharge

### 3.2.1 WRF-Hydro

The National Center for Atmospheric Research (NCAR) developed a Weather Research and Forecasting hydrological model extension package (WRF-Hydro) to improve representations of terrestrial hydrologic processes. This model is designed to provide a link between various earth system models (i.e. atmospheric and terrestrial models); however, this research uses it as a stand alone land surface model (LSM). Included in WRF-Hydro is the Noah multi-parameterization (Noah-MP) LSM; providing a 1-dimensional vertical land surface parameterization using inputs of short and longwave radiation, specific humidity, air temperature, surface pressure, near surface wind, and precipitation rate, to calculate vertical energy fluxes in the form of heat and radiation, moisture, and soil states. The modular structure of WRF-Hydro allows consideration of the physics of surface overland flow and saturated subsurface flow (Figure 3-1). WRF-Hydro utilizes the subsurface lateral flow module to add any exfiltration from fully saturated grid cells to Noah-MP outputs of infiltration excess, allowing for the overland flow routing module to calculate surface runoff. Because the baseflow module within WRF-Hydro is conceptual rather than physical, this module is more used as calibration parameters to the Noah-MP LSM. Although the WRF-Hydro framework provides a channel routing module, this research provides a separate routing scheme that operates as a post processor to the surface runoff output. Additional information on the WRF-Hydro model is provided in the *NCAR WRF-Hydro Technical Description and User's Guide* [Gochis et al., 2015].



**Figure 3-1 Modular Calling Structure of WRF-Hydro (from [Gochis et al., 2015])**

### 3.2.2 Calibration of WRF-Hydro

WRF-Hydro, compared to other hydrology models, is computationally intensive; therefore, the number of runs to calibrate the model is limited. The Statistical Parameter Optimization Tool in Python (SPOTPY), developed by Houska et al. 2015, is used to calibrate the parameters within the baseflow module of WRF-Hydro. Senatore et al. 2015 identified the two most relevant parameters for estimating an hourly hydrograph; the infiltration factor (REFKDT) and a coefficient governing deep drainage that adjusts such in the deepest soil layer (SLOPE). An additional parameter, the saturated soil hydraulic conductivity (DKSAT), as suggested by the Noah LSM User's Guide for calibration, is also employed [Mitchell et al., 2005].

The Monte Carlo Markov Chain algorithm in SPOTPY is used to calibrate the identified calibration parameters. Previous calibration efforts have optimized parameters according to their Nash-Sutcliffe Efficiency (NSE); however, this research also employs measures of Kling-Cupta Efficiency (KGE) in an effort to minimize bias [Gupta et al.,

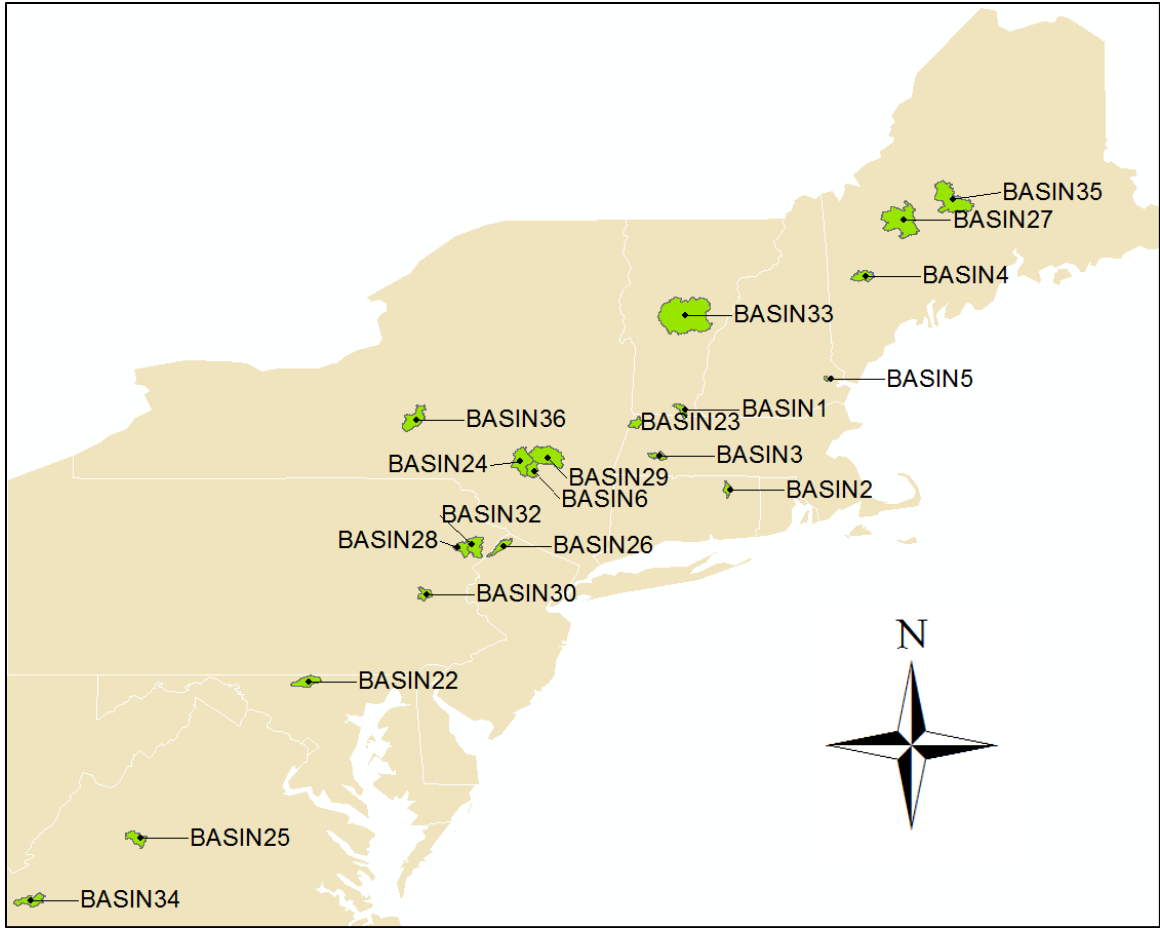


2009]. This algorithm has been widely used in hydrological modeling as a technique to find the global optimum of selected parameters [Houska *et al.*, 2015]. The equations used for NSE and KGE are provided in Appendix B.

### **3.2.3 Watershed Selection**

To analyze the WRF-Hydro gridded output from the U.S Northeast region, 20 watersheds are selected to illustrate the spatial distribution of changes in peak annual flow (PAF) events in the region, where PAF is defined as the maximum streamflow in a given year. Observational climate data provided by the North American Land Data Assimilation System (NLDAS) are used as forcings to the WRF-Hydro model to allow calibration and validation measures. Thereby using these preliminary results to select the 20 watersheds based on their respective NSE and KGE values when compared to the observed USGS unimpeded streamflow gages' historical record. The watersheds are chosen based on the following metrics: NSE or KGE values above 0.5, or an absolute percentage bias below 10% for the 36 years of daily flow values [D. N. Moriasi *et al.*, 2007].

Out of the ~15 Northeast states modeled, watersheds are selected from 10 states; 1 in CT, 3 in MA, 1 in MD, 3 in ME, 1 in NH, 1 in NJ, 4 in NY, 3 in PA, 2 in VA, and 1 in VT (Figure 3-2, Table 3-1). Their respective drainage areas range from 31.30 km<sup>2</sup> to 1790.24 km<sup>2</sup>.



**Figure 3-2 Locations of Selected Watersheds**

**Table 3-1: Selected Watershed Characteristics**

Basin ID	Watershed	USGS ID	State	Drainage Area, km <sup>2</sup>	Latitude	Longitude	ELEVATION, METERS				
							MIN	MAX	RANGE	MEAN	STD
5	Oyster River near Durham, NH	01073000	NH	31.3	43.148696	-70.96506	21	117	96	59	20
2	Mount Hope River near Warrenville, CT	01121000	CT	70.25	41.843709	-72.16897	75	383	308	186	50
1	Green River near Colrain, MA	01170100	MA	106.99	42.703417	-72.67065	130	736	607	411	114
23	Green River at Williamstown, MA	01333000	MA	112.16	42.708969	-73.19677	180	1061	881	472	171
30	Jordan Creek near Schnecksville, PA	01451800	PA	135.84	40.661762	-75.62685	103	478	374	202	35
26	Flat Brook near Flatbrookville, NJ	01440000	NJ	167.72	41.106667	-74.95222	96	499	403	272	84
6	Esopus Creek at Allaben, NY	01362200	NY	169.19	42.117034	-74.38015	305	1135	831	672	165
28	Brodhead Creek near Analomink, PA	01440400	PA	175.21	41.084815	-75.21463	176	666	489	422	107
4	Little Androscoggin River near South Paris, ME	01057000	ME	190.92	44.303992	-70.53968	116	733	616	281	93
25	Tye River near Lovingson, VA	02027000	VA	240.77	37.715419	-78.98169	173	1239	1066	549	267
3	West Branch Westfield River at Huntington, MA	01181000	MA	243.5	42.237312	-72.89565	110	608	498	382	96
22	Big Pipe Creek at Bruceville, MD	01639500	MD	267.18	39.612361	-77.23744	103	336	233	196	42
34	S F Roanoke River near Shawsville, VA	02053800	VA	280.72	37.140132	-80.26643	401	1196	795	702	141
32	Bush Kill at Shoemakers, PA	01439500	PA	305.88	41.088151	-75.03768	123	621	498	386	72
36	Ostelic River at Cincinnatus, NY	01510000	NY	382.99	42.541181	-75.89964	311	652	340	484	67
24	East Brook Delaware River at Margaretville, NY	01413500	NY	424.1	42.14481	-74.65349	389	1180	791	664	140
29	Schoharie Creek at Prattsville, NY	01350000	NY	612.51	42.319528	-74.43654	344	1233	889	652	148
35	Piscataquis River near Dover-Foxcroft, ME	01031500	ME	769.05	45.175008	-69.3147	109	798	689	303	107
27	Carrabassett River near North Anson, ME	01047000	ME	909.1	44.8692	-69.9551	94	1291	1197	376	213
33	White River at West Hartford, VT	01144000	VT	1790.24	43.714236	-72.41815	113	1150	1037	447	155

### 3.2.4 Routing Scheme

A prominent factor in determining adequate and robust engineering designs for bridges and culverts is the magnitude of peak flows. Estimating peak flows involves the estimation of travel time and attenuation of the flood waves; commonly termed this flood routing. For most applications, flows are considered unsteady in open-channels and can be solved for using the Saint-Venant Equations developed by Barre de Saint-Venant in 1848. This technique accounts for continuity and momentum for one-dimensional flow [Heatherman, 2012]. The momentum equation accounts for gravity, pressure variation, and friction due to the channel walls. A full solution that employs all terms of the Saint-Venant equations is considered a ‘hydraulic’ solution, whereas simplifications are called ‘hydrologic’ solutions, in which one or more terms in the momentum equations are omitted. Because a hydraulic solution is very computationally intensive, most

approaches are simplifications of the Saint-Venant equations in which solutions are empirical approximations [Heatherman, 2012].

To minimize overall computational time in this research, the gridded outputs of surface runoff from the WRF-Hydro model are routed through channel networks within the 20 selected watersheds, rather than through the entire U.S. Northeast domain. This is achieved by independently employing the Muskingum method in each watershed by using a Python module written to approximate the Saint-Venant equations. The Muskingum method relies on two coefficients that are used to represent the travel time and attenuation of flood waves,  $K$ , and  $X$ , respectively. This routing approximation maintains constant calibration parameters that do not vary with flow. This method is employed in this research through a code written in Python and is provided for reference in Appendix B.

### **3.3 Watershed-Based Analysis of Results**

#### **3.3.1 Model Performance**

##### **3.3.1.1 Spatial Downscaling & Bias Correction Skill on NARCCAP Data**

The ability of the SDBC method to correct the monthly distributions of daily simulated NARCCAP precipitation and temperature data is measured using a simple error metric, along with percent bias in precipitation. Although the final product of this method provided 3-hourly data to WRF-Hydro, surface runoff outputs are on a daily time step, therefore, the climate downscaling and bias correction is also evaluated using daily values. To minimize computation time, the time series at each gridded point within a watershed is aggregated to create a single precipitation and temperature time series on

which analyses are performed. Error values during the calibration (1968-1988) period are compared to those of the validation (1989-1999) period to evaluate accuracy in the future projection (2038-2070) period.

### **3.3.1.2 WRF- Hydro Simulations of Peak Annual Flows**

The peak annual flows (PAF) in the historical (1968-1999) period and the future (2038-2070) period are extracted after routing the WRF-Hydro runoff outputs. For each selected watershed, the maximum river discharge is obtained from each year from USGS observed streamflow and the six estimated (simulated) river discharges (1 NDLAS dataset, 5 NARCCAP datasets). These PAF are fit to a generalized extreme values (GEV) distribution using RStudio. Once performed, this distribution allows estimation of the upper, lower, and median confidence intervals of various annual return intervals through statistical bootstrapping. As a final calibration step, a scaling factor is determined from the comparison between the GEV distributions of observed USGS peak flows and peak flows simulated using NLDAS data. This provides calibration of the peak streamflow outputs from the Muskingum routing scheme. This scaling factor is applied to the GEV distribution of peak flows simulated from the 5 NARCCAP datasets for both the historical and future time periods, thereby assuming some stationarity within the model. Model performance is evaluated using goodness-of-fit metrics provided by an R package called “HydroGOF” that compares the GEV distribution of simulated historical flows to the observed distribution found in the USGS data. Projected changes in PAF are also calculated within each NARCCAP dataset between the historical and future time periods, in each watershed.

### **3.3.2 Hydrologic Responses to Climate Change**

#### **3.3.2.1 Peaks over Threshold (POT)**

Because a peak flow analysis uses only a single value each year, an alternative analysis is conducted using a peaks-over-threshold (POT) approach [Mallakpour and Villarini, 2015]. For both streamflow and precipitation, thresholds are defined as the 50<sup>th</sup>, 95<sup>th</sup>, 97<sup>th</sup>, and 99<sup>th</sup> quantiles of monthly CDFs, partitioned into historical or future time period. For example, all daily January values of precipitation in the historical period (31 days x 30 years) are used to create a CDF, from which values of selected thresholds (quantiles) are extracted. These quantile values (i.e. predetermined historical magnitudes) serve as thresholds, where the number of days exceeding each threshold is summed for each month in each year within the period.

#### **3.3.2.2 Seasonality of Extreme Events**

Investigating the changes in seasonality of peak annual flows (PAF) and peak annual precipitation (PAP) provides information on the time of year of such events and the relationship between the two. The day of the year (i.e. Julian day) is extracted from PAF and PAP events during the historical and future time periods. These values are used to create 4 CDFs for each watershed; 1) Julian days of PAF in the historical period, 2) Julian days of PAP in the historical period, 3) Julian days of PAF in the future period, and 4) Julian days of PAP in the future period. Changes in the relationship between the timing of PAF and PAP suggests changes in runoff attenuation, perhaps indicating growing urban areas or changes in hydrologic regimes such as decreasing a snowpack.

### 3.3.2.3 Climate Indices

Supplementing the POT and seasonality analysis, and furthering the investigation of the significance of precipitation, select indices from the standard ClimDEX indices are quantified in the five downscaled NARCCAP products. ClimDEX indices were developed by the Expert Team on Climate Change Detection and Indices (ETCCDI) [Zhang, n.d.]; this software is hosted and maintained by the Climate Research Division of Environment Canada. The selection of ClimDEX indices are based on their potential influence on soil moisture in terms of precipitation magnitude and intensity.

## 4 RESULTS

The hydrologic model generates an ensemble of five simulated time series of streamflow at the outlet of the 20 selected watersheds; from which projected changes in various metrics are calculated. The preceding steps, including the climate data downscaling approach and calibration of the hydrology model, are validated in terms of their skill in reproducing events in the observed time series provided by a gridded observational dataset (Livneh et al. 2014) and USGS streamflow measurements. These results are presented beginning with model skill during the historical period, followed by a summary of the projected changes during the future period. The projected changes in climate and peak annual flows (PAF) during the future period are further investigated through a PAF seasonality analysis, a seasonal peaks-over-threshold (POT) analysis, and a ClimDEX indices analysis.

The projected changes in PAF indicate a wide range of increases and decreases in the 20 selected watersheds that prove to require additional analysis when exploring their underlying causes. The seasonality analysis of PAF indicates a shift in the month of

occurrence that then, coupled with projected changes in precipitation of that particular month; more clearly represents the influence of extreme precipitation on high streamflow events. The seasonal POT analysis provides a type of frequency assessment that reveals the varying levels of influence of extreme precipitation events, where a heavy precipitation event does not always create high streamflow and is rather dependent on season. Additional climate conditions such as precipitation magnitude and intensity are further examined using select ClimDEX indices to understand the changing influence of heavy precipitation on streamflow in certain seasons.

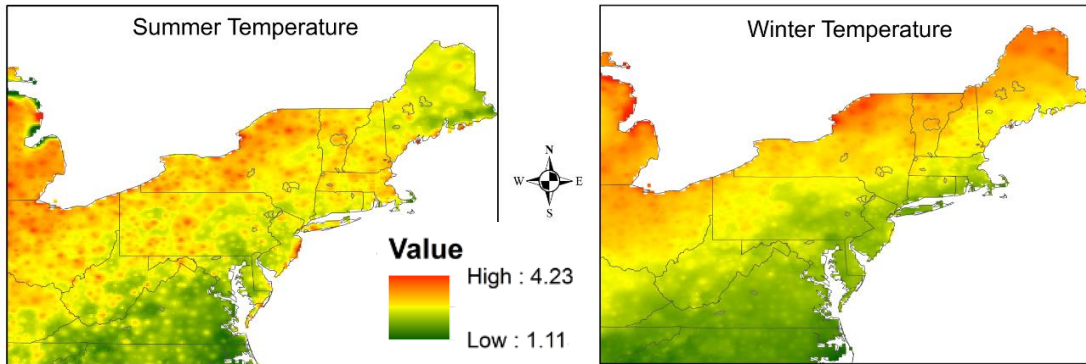
## **4.1 Modeling Performance & Projections**

### **4.1.1 Skill of SDBC Method**

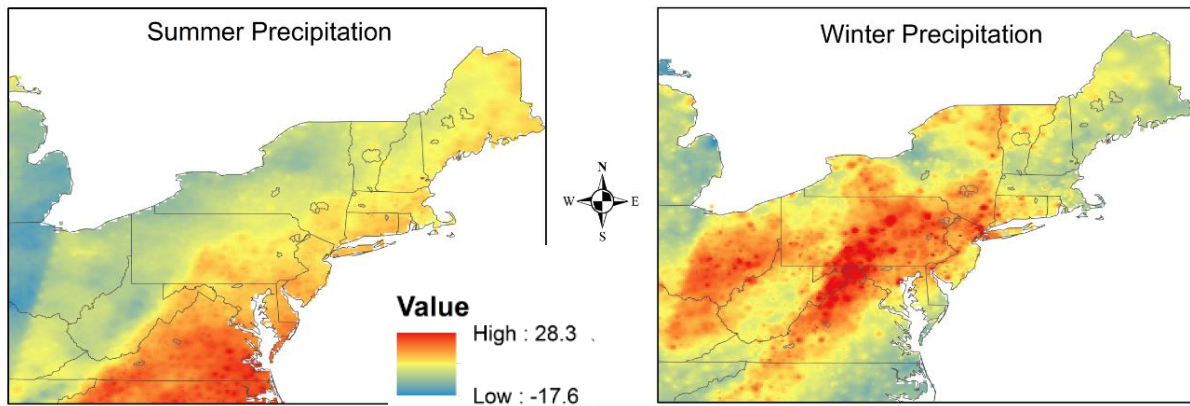
The downscaled NARCCAP climate data provides temperature and precipitation data on a 3-hourly time step to the hydrological model; however, results are analyzed on a daily time step. The observed time series used to ‘train’ the SDBC method spans the years 1968-1988, while the remaining years 1989-1999 are used as a validation period. Because the data are downscaled on a monthly basis, performance metrics are also evaluated at this scale. The SDBC method uses the calibration period (1968-1988) as ‘training’ data, so it is not surprising that the downscaled NARCCAP datasets performs reasonably well during this period. Skill is measured as error for daily mean temperature and percent bias for daily precipitation accumulation in each month and each watershed. For example, daily values in January over the historical period (31 days x 30 years) from the 5 NARCCAP datasets are analyzed for error and percent bias. The following figures characterize the results of the SDBC procedure that provide a finer spatial scale (1/16°)



outlook at changes in summer and winter temperature and precipitation (Figure 4-1 and Figure 4-2).



**Figure 4-1 SDBC Results of NARCCAP – Change in Mean Monthly Temperature (°C) (Average of 5 NARCCAP SDBC Models)**



**Figure 4-2 SDBC Results of NARCCAP – Change in Total Monthly Precipitation (%) (Average of 5 NARCCAP SDBC Models)**

The error analysis performed on mean daily temperature shows excellent skill in reproducing the monthly distribution of observed values during the calibration period (Figures 4-3, 4-4). Downscaling performance decreases during the validation period, with the largest errors in the cooler months of the region (Figure 4-3). The variation in error is much larger during the validation period compared to that of the calibration period. This may be due to multi-decadal temporal variations in the temperature signal

that the relatively short (20 year) calibration period is unable to capture; however, median values of monthly error are generally below 1°C, with the exception of December, when error values are greater than 1°C in almost all watersheds (Figure 4-3). The majority of months indicate a slight over-prediction, with the exception of January and February, this almost to be expected due to the model projected increases in temperature as a result of increased greenhouse gasses. Although over-prediction of mean daily temperature may result in over-prediction of evaporation within the hydrologic model, ensemble-mean errors are generally less than 1°C and merely represent a number of possible futures that the region may experience. The biases within each NARCCAP dataset will also be contained, as any reported changes in future values are relative to its historical period.

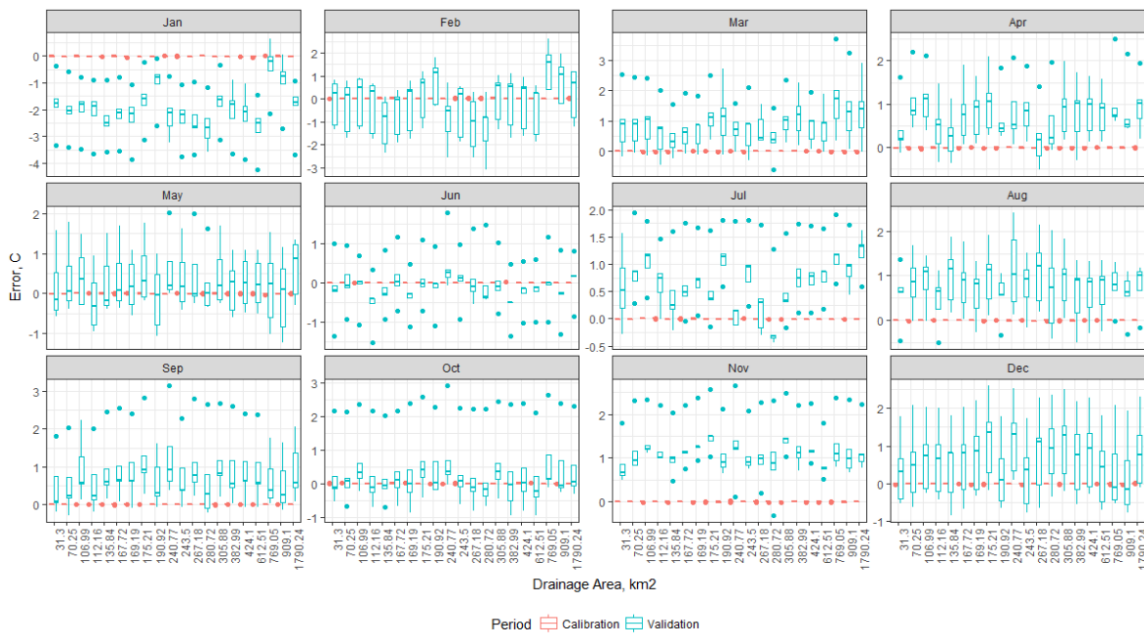
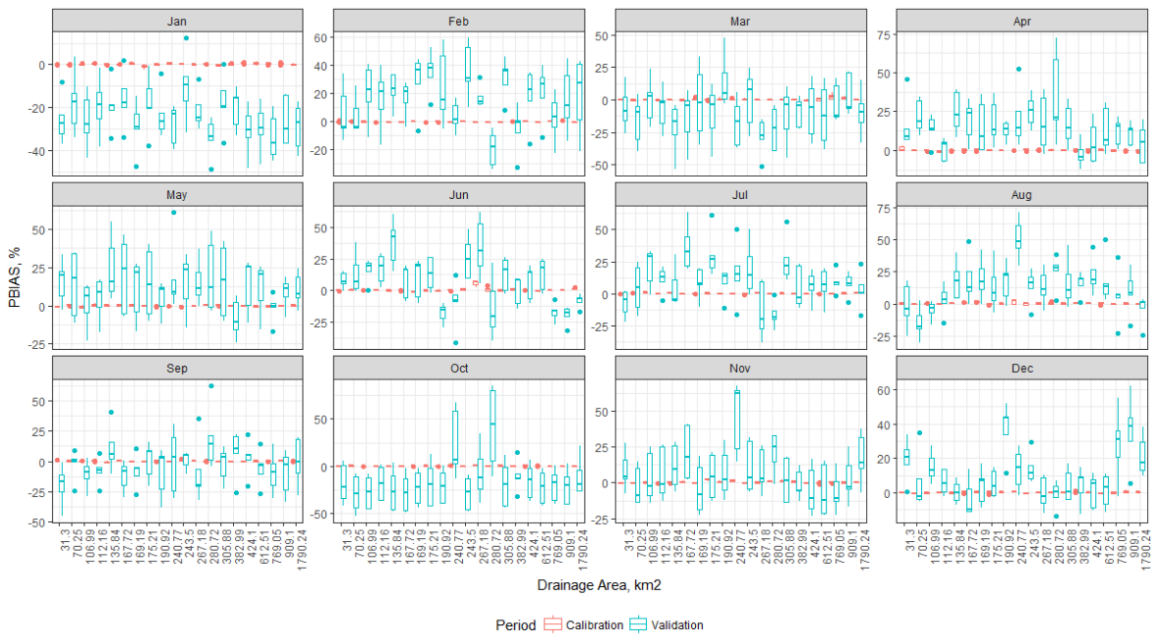


Figure 4-3 Mean Monthly Error in Mean Daily Temperature from the SDBC Method during the Calibration and Validation Periods.

The percent bias (PBIAS) analysis performed on daily precipitation accumulation indicates a slightly more varied skill than the mean daily temperature values in both the calibration and validation periods. During the calibration period, the variability in PBIAS

seems to be dependent on month, and watershed. For example, in March, the larger watersheds show a significantly larger range of error than other watersheds, however, in August, a mid-sized watershed shows the greatest variability (Figure 4-4). Interestingly, the range of error during the calibration period does not always propagate to the validation period. For example, the largest range of error in August (as previously mentioned) during the calibration period leads to the largest range of error during the validation period; however, the opposite happens in November, where the largest range of error during the validation period does not originate from a large range of error in the calibration period (Figure 4-4).



**Figure 4-4 Mean Percent Bias in Daily Precipitation Accumulation from the SDBC Method during the Calibration and Validation Periods.**

Errors during the calibration period are relatively small; the majority of watersheds have a precipitation error less than 1mm, in comparison to the errors in the validation period where values range from at least 1mm to a maximum of 8mm (Appendix C). The month of December shows a significant error in all watersheds in the

validation period despite the opposite behavior during the calibration period. While plotted in the order of increasing watershed size along the x-axis, visual inspection suggests little to no trend in the performance of the downscaling method. It is important to note that these errors, especially during the validation period, may be comparable to the projected changes in mean temperature and daily precipitation from the downscaled and bias corrected NARCCAP datasets. However, not all changes in the projected climate are accounted for in the error analysis. For example, the average change in total monthly winter precipitation indicate increases in the selected watersheds (Figure 4-2) while errors during the winter months indicate minimal error, underestimation, and overestimation in December, January, and February, respectively (Figure 4-4). Although performance varies from month to month during the validation period, when further aggregated to the seasonal or annual scale, errors are not as high, thus allowing for additional analysis on a larger time scale.

## **4.1.2 Peak Annual Flows**

### **4.1.2.1 Simulated Historical Peak Annual Flows**

Peak annual flows are of great interest to stakeholders because of their impacts on infrastructure; thus the ability to recreate historical values is important in validating values in the future period. It is important to ensure that the model is calibrated such that the physics within the model provide accurate estimates of peak annual flows. This process utilizes observed climate assimilations as inputs to the model, thereby producing estimates of streamflows. These are compared to USGS observed streamflow to estimate the accuracy of the physics within the hydrologic model. The hydrologic model results

from forcing NLDAS climate data are fitted to a GEV distribution of peak annual flows (PAF) and compared to a fitted GEV distribution of PAF from the USGS time-series. These watersheds were selected using preliminary runs of NLDAS data. As expected, the majority of the selected watersheds perform well in their goodness-of-fit metrics from their respective GEV distributions (Table 4-1).

Nash-Sutcliffe efficiency (NSE) is a measure of the relative magnitude of residual variance of simulated values compared to the variance of observed data, where a value of 0 indicates that simulated values are as accurate as the mean of observed data. All watersheds show a NSE value above 0, with the majority above 0.5 (Table 4-1). Kling-Gupta efficiency (KGE) was developed as an alternative to NSE to address its shortcomings when being optimized for hydrological modeling calibration; especially due to its tendency to underestimate runoff peaks [Gupta *et al.*, 2009]. Because this research is centered on predicting high flows, KGE is preferred due to its ability to simultaneously account for bias, correlation, and variability [Gupta *et al.*, 2009]. Almost all watersheds report a lower KGE value than their respective NSE values, suggesting that the variability in the simulated values is underestimated. Volumetric efficiency (VE) provides another goodness-of-fit metric that has physical significance by, in this case, representing the volumetric fraction of water captured by peak annual flows over the historical time period [Criss and Winston, 2008]. The watersheds in this research show high VE, further validating the hydrological model's performance at these locations, thereby providing some confidence in projected future values (Table 4-1).

One watershed (Schoharie Creek in Prattsville, NY) reported a low NSE value of 0.19; a value outside of the watershed NSE selection criteria. Further analysis of the

observed and simulated GEV distributions show that the simulated values are accurate in representing the magnitude of stream flows at the lower return intervals (or more frequently seen flows), followed by a growing under-prediction of flows at higher return intervals (Appendix A). Despite this, an R2 value of 0.99 for the GEV distributions of Schoharie Creek in Prattsville, NY shows that the shape of the observed distribution is well maintained in simulated values (Table 4-1). Interestingly, this watershed also reported a higher KGE value, contrary to all other watersheds. This may be because NSE has a high weighting of residuals, while KGE optimally weights bias, variability, and correlation; validating the previous visual inspection of the GEV distributions of this particular watershed. GEV distributions of simulated and observed peak annual flows (PAF) at all 20 watersheds are in Appendix A.

**Table 4-1 GEV Goodness of Fit Metrics**

BASIN ID	Location	KGE	NSE	R2	RMSE	PBIAS %	VE
22	BIG PIPE CREEK AT BRUCEVILLE, MD	0.66	0.82	0.99	31.09	-15.6	0.83
28	Brodhead Creek near Analomink, PA	0.74	0.81	0.99	23.06	-18.4	0.81
32	Bush Kill at Shoemakers, PA	0.9	0.99	1	15.33	0	0.95
27	Carrabassett River near North Anson, Maine	0.37	0.53	0.88	77.92	-17.5	0.63
24	EAST BR DELAWARE R AT MARGARETVILLE NY	0.76	0.9	1	11.21	-8.5	0.91
6	ESOPUS CREEK AT ALLABEN NY	0.7	0.87	0.99	13.37	10.1	0.87
26	FLAT BROOK NEAR FLATBROOKVILLE NJ	0.73	0.88	1	28.84	-12.4	0.87
23	GREEN RIVER AT WILLIAMSTOWN, MA	0.91	0.99	1	4.99	-1.6	0.96
1	GREEN RIVER NEAR COLRAIN, MA	0.67	0.82	1	3.83	-16.8	0.81
30	Jordan Creek near Schnecksville, PA	0.7	0.83	0.99	18.64	8	0.9
4	Little Androscoggin River near South Paris, Maine	0.67	0.76	1	7.88	14.8	0.85
2	MOUNT HOPE RIVER NEAR WARRENVILLE, CT.	0.99	1	1	0.45	0.9	0.99
36	OTSELIC RIVER AT CINCINNATUS NY	0.44	0.55	0.99	138	12.5	0.82
5	OYSTER RIVER NEAR DURHAM, NH	0.67	0.85	0.99	9.41	5.6	0.88
35	Piscataquis River near Dover-Foxcroft, Maine	0.53	0.66	0.96	72.41	8.3	0.86
34	S F ROANOKE RIVER NEAR SHAWSVILLE, VA	0.55	0.63	0.96	116.06	-18.9	0.81
29	SCHOHARIE CREEK AT PRATTSVILLE NY	0.4	0.19	0.99	35.8	28.4	0.72
25	TYE RIVER NEAR LOVINGSTON, VA	0.89	0.99	1	10.17	-2.9	0.95
3	WEST BRANCH WESTFIELD RIVER AT HUNTINGTON, MA	0.84	0.95	0.99	4.71	-4.9	0.93
33	WHITE RIVER AT WEST HARTFORD, VT	0.87	0.96	1	57.11	-9.2	0.91

#### 4.1.2.2 Future Peak Annual Flows

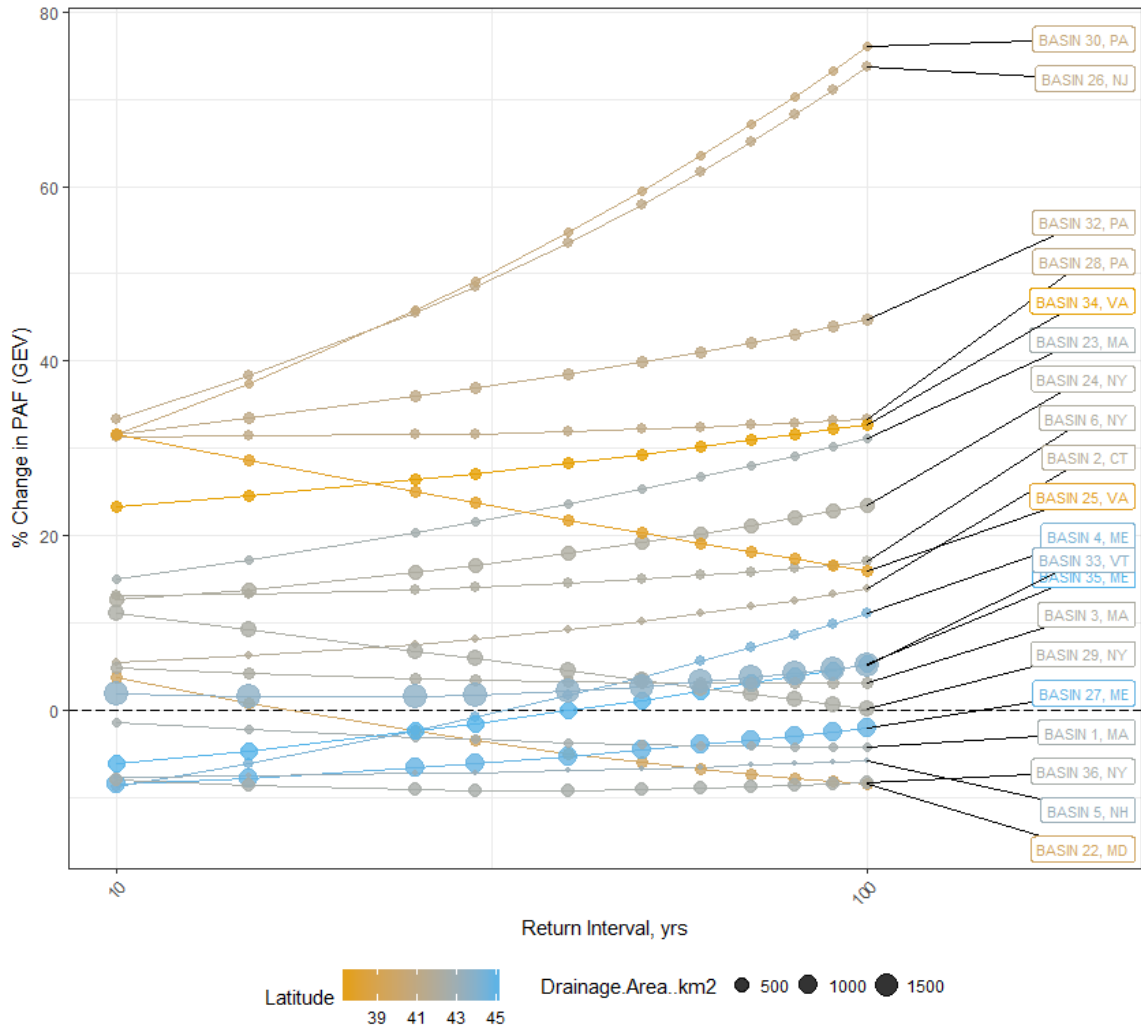
The simulated historical and future streamflow time-series generated by the hydrological modeling process are sampled for their PAF from their respective GEV

distribution to calculate percent changes in the future period at 15 different return intervals (Table 4-2). Percent changes are calculated for each of the 5 NARCCAP futures and averaged to produce an ensemble-mean. The region represented by the average percent change among the 20 watersheds, show increases at every return interval ranging from 10% to 18% (Table 4-2).

**Table 4-2 Regional Average Percent Change in Peak Annual Flows.**

<b>Return Interval, years</b>	2	3	5	7	10	15	25	30	40	50	60	70	80	90	100
<b>Regional Average % Change (of 20 Watersheds)</b>	10	10	10	10	10	11	12	13	14	15	15	16	17	17	18

These regional trends of increasing PAF are analogous to the increases in the total annual precipitation (Appendix C). This is indicative of the hydrologic model being able to translate changes in precipitation to changes in stream flows. However, the 20 watersheds within the region do not always agree in terms of the direction of change (Figure 4-5). For example, a few watersheds indicate decreases in the magnitude of PAF at all return intervals, while nearby watersheds indicate increases at all return intervals (Figure 4-5). When evaluated on an annual scale, the differences in changes in PAF within individual seasons may aggregate to an overall decrease in frequency; however, disaggregating to a seasonal analysis provides a slightly finer temporal scale to evaluate changes. This broaches the question of the seasonal characteristics of a precipitation event that cause extreme flood events.



**Figure 4-5 Percent Change in Peak Annual Flow**

## 4.2 Hydrologic Responses to Climate Change

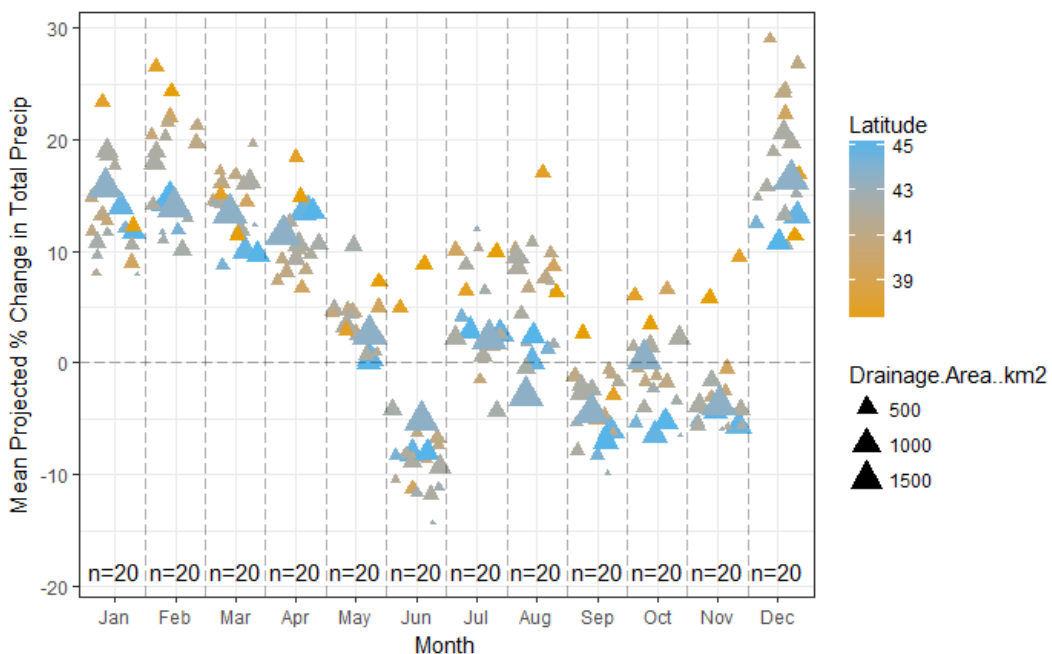
### 4.2.1 Projected Changes in Monthly Precipitation

Given the ranges of errors in the downscaling and bias correction of the NARCCAP simulated precipitation data (Figure 4-3& 4-4), daily values are aggregated to the monthly scale and averaged within the historical and future time series at each simulation point in each respective basin to eliminate some uncertainty in the direction of change. This aggregation from daily to monthly precipitation values provides a medium



to report projected climate impacts in place of the otherwise noisy daily signal or the over-smoothed annual signal.

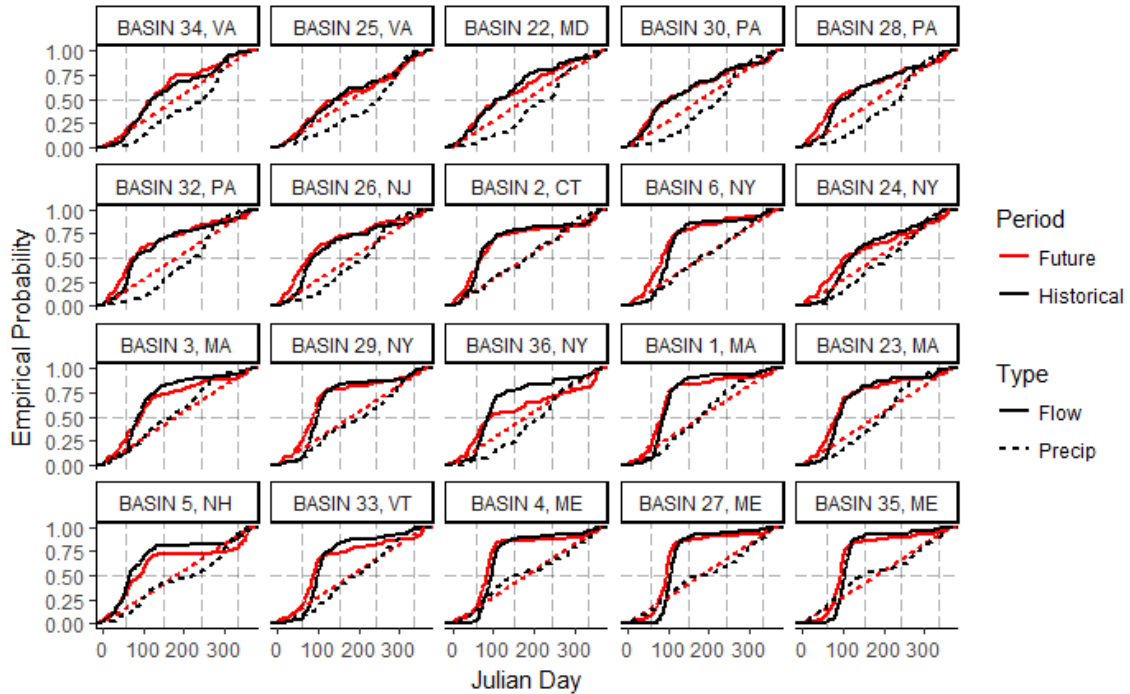
The total monthly precipitation per year in each time period are averaged, and used to calculate percent change in each of the 5 downscaled and bias corrected NARCCAP datasets. The winter (DJF) months in all the investigated watersheds show between 8-25% increases at the ensemble-mean (Figure 4-6). During the dryer seasons of fall (SON), watersheds at different latitudes show somewhat different climatology; where the smaller southern basins experience increases, while the larger northern basins experience decreases in total monthly precipitation (Figure 4-6). June is similar; however, the latter summer months show increases except for a few watersheds. The disagreement in the direction of change at the monthly scale, despite agreements at the annual scale and amongst the differing trends in PAF in each watershed, suggests a need to investigate the seasonality in the relationship between extreme precipitation and high streamflow events (Figure 4-6, Appendix C).



## **Figure 4-6 Monthly Annual Total Precipitation**

### **4.2.2 Seasonality of Peak Annual Flows**

As a preliminary investigation of the relationship between extreme precipitation and extreme flood events, the Julian day of PAF and peak annual precipitation (PAP) are calculated for each year of the historical and future periods. For example, the Julian day of each PAF of each year in the historical period is used to create a cumulative density plot (CDF); these are developed for both PAF and PAP in the historical and future time periods. This analysis indicates the seasonality in the relationship between maximum precipitation and maximum streamflow (Figure 4-7). Many of the southern watersheds indicate an increased probability of PAP occurring in the winter, summer, and spring. The majority of PAF events happen in the winter or spring months, despite the tendency for PAP events to occur relatively evenly over all seasons (Figure 4-7). This suggests that there are other significant drivers of extreme flows during the spring and summer other than large precipitation events.



**Figure 4-7: CDF of Julian Day of Annual Maximum Flow and Precipitation**

As indicated previously, model results indicate that the U.S. Northeast region as a whole will see increases in PAF at all return intervals; however, a number of watersheds show decreases at all return intervals (Table 4-2, Figure 4-5). At the 100-year return interval, the watersheds showing the greatest percentage of PAF increase (30-85% in BASIN 30, 26, 32, 28) can be explained by their shift towards an earlier PAF timing toward the winter in the future period; likely due to the shift in peak precipitation events (Figures 4-5 & 4-7). The next grouping watersheds showing increases in PAF (5-30% in BASIN 23, 24, 2, 3) also show a spread towards such events occurring earlier in the winter and less likely in the spring (Figures 4-5 & 4-7). Flows during the spring tend to be of greater magnitude, particularly in the more northern and snow-dominated watersheds. Because this particular group of watersheds indicates a strong concentration of PAF occurring in the spring during the historical period, a shift towards earlier flows may be indicative of earlier and less snowmelt. This theory, counterweighted with the

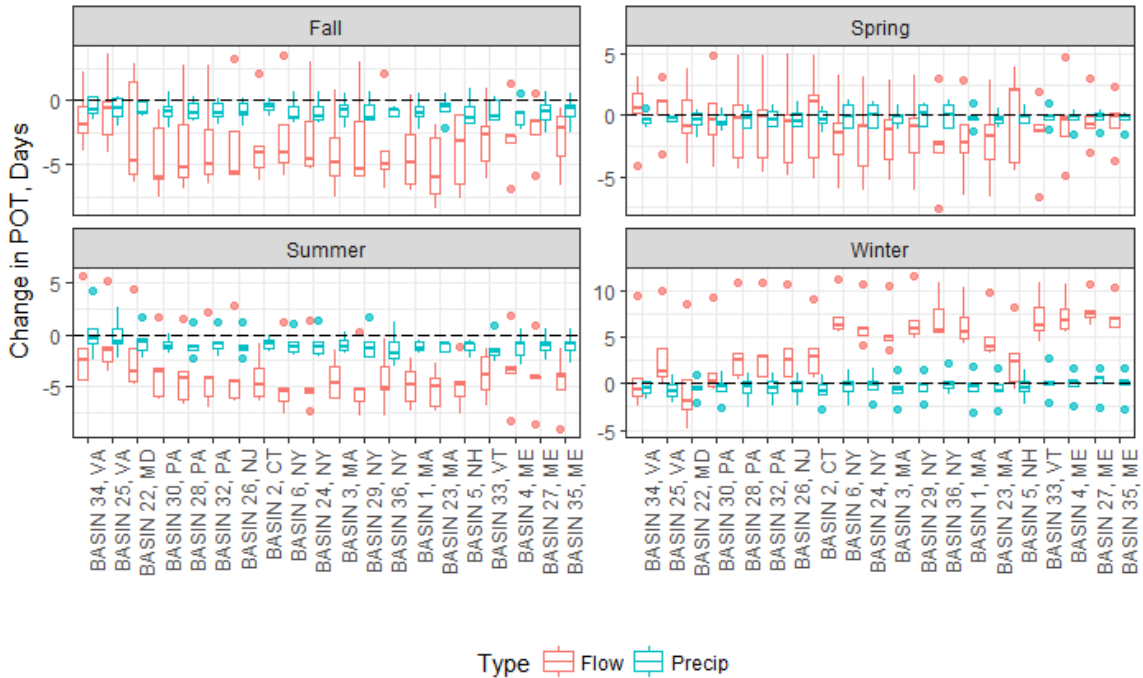
projected increases in winter precipitation, explains the significantly greater percent change in PAF in the PA and NJ watersheds, where there is little to no snowpack and a shift to more winter PAF. These watersheds indicate a more direct correlation between precipitation and streamflow, especially during the winter due to the lack of a snowpack buffer. The few watersheds indicate opposite trends in changes in PAF between its return intervals (i.e. -10% at the 10-year return interval and 10% at the 100-year return interval in BASIN 4) are snow-dominated watersheds that also show a greater number of earlier winter and spring flood events. Three watersheds (BASIN 25, 29, and 22) indicate a diminishing percent change in PAF from the 10-year to 100-year return interval (Figure 4-5). The southern-most watersheds (BASIN 25 and 22) indicate a more correlated relationship between extreme precipitation and peak streamflow during the future period. This suggests that the buffer between the two (e.g. snowpack or pervious surfaces) has diminished; therefore resulting in higher return period streamflow exhibiting less change than more frequent streamflow at lower return intervals.

The shifts in seasonality are directly related to the changes in climate, where increases in temperature in conjunction with increases in winter precipitation have led to an increased likelihood in peak annual events occurring earlier in the year. This effect in snow-dominated watersheds can promote a greater number of smaller snowmelts throughout the winter rather than a large concentrated snowmelt event in the spring. However, increases in winter precipitation may also lead to more precipitation falling as snow, resulting in greater magnitudes in winter or spring flood events. This seasonality analysis provides an improved sense of the changing processes that cause PAF events.

### 4.2.3 Seasonal Peaks over Threshold

To further this investigation, a count of peaks-over-threshold (POT) is performed on both the precipitation and stream flow time series to examine any changes in the frequency of peak annual events. The thresholds are defined as the 50<sup>th</sup>, 95<sup>th</sup>, 97<sup>th</sup>, and 99<sup>th</sup> percentile of historical daily values aggregated by month. For example, the 95<sup>th</sup> percentile flow threshold for January is extracted from a CDF of daily precipitation from all Januarys in the historical period (1968-1999). The numbers of daily values that exceed each threshold are counted in each month and in each year. The change in daily values greater than each threshold (or POT) is the difference in the average number of POT within the historical and future time periods, respectively.

Figure 4-8 contains boxplots of the change in the number of precipitation and streamflow POT from the historical to the future period within each NARCCAP dataset at the 50<sup>th</sup> percentile threshold. The differences between precipitation and streamflow changes in POT at the 50<sup>th</sup> percentile indicate that such events are more correlated in the spring months than in the fall, summer, and winter months. The higher thresholds (95<sup>th</sup>, 97<sup>th</sup>, and 99<sup>th</sup> percentiles) show a greater correlation between the two, suggesting that days of extreme precipitation may be contributing to days of extreme flows (Appendix C).

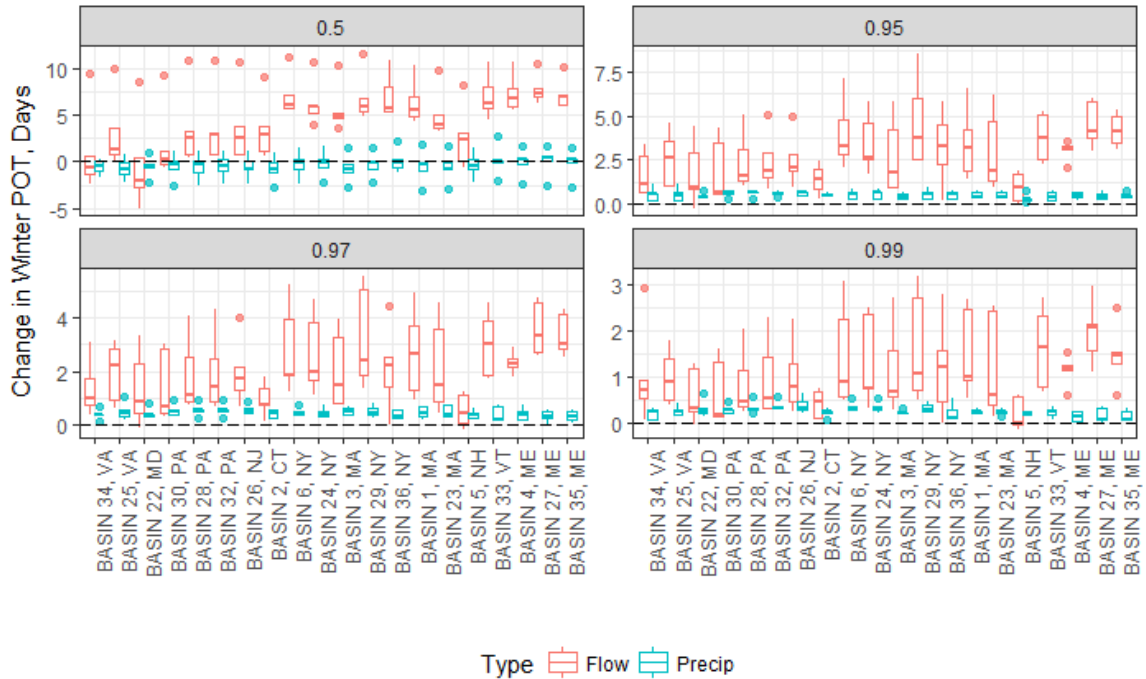


**Figure 4-8 Seasonal Precipitation and Streamflow POT (50<sup>th</sup> Percentile Threshold)**

The most apparent changes in the future period are decreases in streamflow POT in the fall, spring, and summer streamflow along with significant increases in the winter at the 50<sup>th</sup> percentile threshold. Despite the minimal change in the number of winter precipitation POT, future streamflow projections indicate an increase in winter POT at all thresholds (Figure 4-9) perhaps due to an increased in rain on snow events, along with earlier and rapid snowpack melt. When the watersheds are ordered from south to north, an increasing trend emerges in the changes in winter POT demonstrating the greater effect that increasing winter precipitation and temperatures have on these watersheds. Similar trends are not seen in the other seasons at the higher thresholds, suggesting that changes in streamflow in the spring, fall, and summer are mostly occurring at median values rather than in more extreme streamflow events.

This POT analysis, in conjunction with the analysis of total monthly precipitation, suggests that precipitation in each season provides a different level of influence on

streamflow POT. Because future streamflow POT indicate increases solely in the winter months despite the lack of the same trend in winter precipitation POT (Figure 4-9), and coupled with the projected increases in winter total monthly precipitation (Figure 4-6), suggests that extreme flood events in the future are not caused by increases in the magnitude of extreme precipitation.



**Figure 4-9 Winter Precipitation and Streamflow POT**

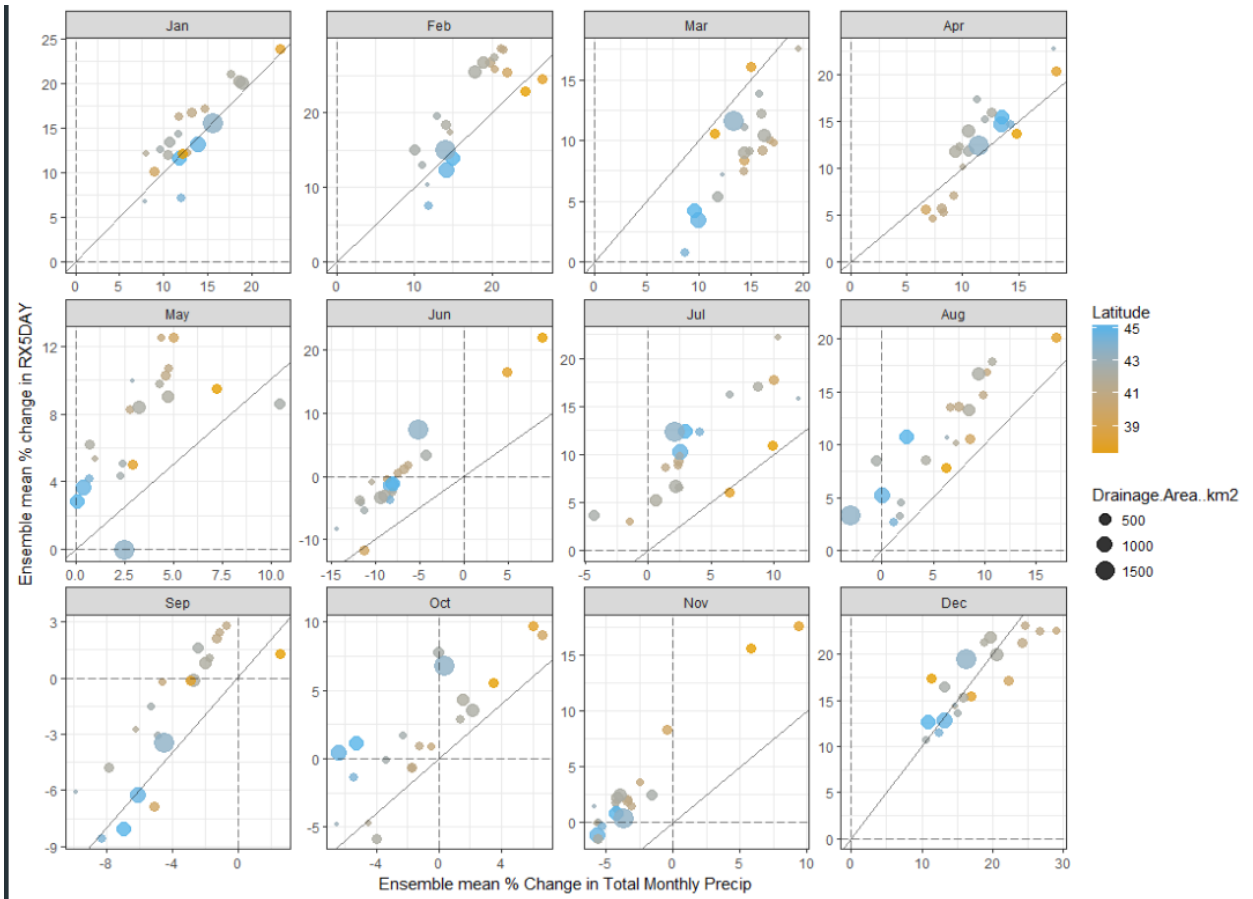
#### 4.2.4 ClimDEX Indices

The ClimDEX indices are selected for this analysis to explore climate trends. The previous analyses in this research indicate the lack of influence of extreme precipitation events on peak seasonal streamflow in most of the selected watersheds, raising the question of what seasonal conditions contribute to increases in high streamflow. The indices selected are total monthly precipitation and monthly maximum 5-day precipitation (RX5DAY).

Changes in total monthly precipitation are presented in previous sections; however, the addition of a RX5DAY analysis provides insight to the characteristics of changing precipitation events. The results suggest an increase in monthly precipitation accumulations in almost all seasons except the fall despite minimal change in precipitation POT as previously noted (Figure 4-6 & 4-8). RX5DAY indicates increases in all selected watersheds for the winter, spring, and the majority of the summer months, along with decreases in the majority of the fall months (Figure 4-10).

Comparisons of changes in RX5DAY and changes in monthly precipitation accumulation provide insight to how precipitation events are changing. For example, the winter months (DJF) indicate increases in both total monthly precipitation and RX5DAY, suggesting that additional precipitation accumulation will fall as a larger multi-day (i.e. 5-day) event. A similar trend is seen in the spring months (MAM); however, the additional precipitation falling in March indicates that the precipitation is being delivered over multiple events rather than one day events due to the change in total monthly accumulation being greater than the change in RX5DAY (Figure 4-10). The opposite is observed in May, where the change in RX5DAY is greater than the change in total monthly accumulation, suggesting an increase in the intensity of a multi-day (i.e. 5-day) event and perhaps decreases in magnitude of other precipitation events within the month (Figure 4-10). The dryer months of the summer and fall seasons indicate both increase and decreases in both metrics. The fall months are slightly more varied in their results; for example, October indicates increases in RX5DAY despite having both increases and decreases in total monthly precipitation in the 20 watersheds (Figure 4-10).





**Figure 4-10: Max 5-Day Precipitation**

## 5 DISCUSSION & CONCLUSIONS

This research provides stakeholders and decision makers in the U.S. Northeast region with estimates of changes in the magnitude and frequency of peak streamflow events, and changes in select precipitation indices. A future period (2038-2070) is compared to a historical period (1968-1999) in terms of various precipitation and streamflow metrics in 20 selected watersheds in the U.S. Northeast. The years included in the future and historical time period are selected based on the available years of the NARCCAP datasets. Five simulated NARCCAP climate datasets are used to model five different hydrological futures using WRF-Hydro in each watershed. As mentioned previously, these five datasets originate from three different GCMs which provides a

platform to evaluate the effect of different RCMs on a GCM; however, due to the large computation effort required, such analysis is limited in this research.

The climate data spatial downscaling and bias correction (SDBC) efforts are validated by quantifying error in mean daily temperature and percent bias (PBIAS) in daily precipitation accumulation from the observational dataset (Livneh et al. 2013). The errors in the statistically downscaled temperature and precipitation values indicate better performance during the calibration period than the validation period. The performance of downscaled daily precipitation when aggregated is similar amongst the 5 NARCCAP datasets; however, performance in a particular watershed may vary (Appendix C). Winter periods demonstrated the largest precipitation PBIAS during the validation period in comparison to the other seasons. The SDBC performance on daily mean temperature show similar patterns in that there is little variability in error amongst the 5 NARCCAP models in the region overall, excluding the winter months, where errors are more variable between the models (Appendix C).

To reduce computational time in analyzing results, the gridded points within each watershed are aggregated to produce a single set of results, rather than a set at each grid point. This watershed aggregation approach applied may lead to overestimates of extreme precipitation events due to the exaggeration of its spatial extent [*Hwang and Graham, 2014*]; though, this research does not find a consistent trend between precipitation PBIAS and watershed area. However, errors that are greater during the validation period may be indicative of nonstationarity, as concluded by Salvi et al. 2016 where visual inspection suggested nonstationarity in the U.S. Northeast in high population areas. Unfortunately, this research cannot confirm whether such errors are due

to nonstationarity or model biases due to the use of a relatively short calibration period (20 years) that is likely unable to capture multi-decadal variability.

Projected peak annual flows are validated using USGS streamflow observations at the outlet of each selected watershed to test the goodness-of-fit of the GEV distributions developed from the WRF-Hydro simulated peak annual flows driven by NLDAS reanalysis data. This process ensures that the physically-based hydrologic model performs satisfactory in representing physical processes in the twenty selected watersheds in terms of their respective NSE and KGE values. After forcing WRF-Hydro with the five NARCCAP datasets, model outputs of future streamflow indicate an increase in peak annual flows at all return intervals in the majority of the selected watersheds, excluding a number of northern watersheds that may be snow-dominated in which decreases are observed due the diminishing accumulation of snowpack.

Previous literature has indicated that antecedent soil moisture has a larger influence than extreme precipitation on extreme flood events based on their finding that a 99<sup>th</sup> percentile precipitation event leads to a 99<sup>th</sup> percentile discharge event 36% of the time, and suggests that increases in extreme precipitation may not always lead to increase in extreme discharges [Ivancic and Shaw, 2015]. These percentiles are defined for each watershed using their entire period of record, and employ a 5-day lag time for the watershed outlet to realize a heavy precipitation event. It is reported that the temporal distribution of soil moisture more closely matches that of streamflow, rather than that of precipitation [Ivancic and Shaw, 2015]. This result aligns with the POT analysis of precipitation and streamflow included in this research since future climate and streamflow indicate seasonal differences in the influence of extreme precipitation on

flood events (Figure 4-8). For example, spring is generally known as the wettest season, thus the frequency of streamflow events is closely correlated with precipitation events of the same threshold (Figure 4-8). The opposite is observed in the summer, generally known as a dry season, where decreases in summer precipitation leads to decreases in soil moisture, displays decreases in summer streamflow despite minimal change in the frequency of precipitation events (Figure 4-8). This result is further corroborated by a study on snow-dominated watersheds in New York State, where 60% of peak annual flows are found to be correlated with moderate rainfall and very wet soil conditions [Shaw and Riha, 2011]. This relationship between streamflow and soil moisture is stronger in the northern watersheds than in the southern watersheds.

The watersheds in this region also indicate the significance of a changing seasonality in projected changes in the magnitude of PAFs. A number of watersheds show an increased correlation in the future period between PAF and peak annual precipitation than in the historical period, especially in the southern-most watersheds (Figure 4-7). This may be indicative of the disappearance of snowpack storage or an increase in impervious surfaces within the watershed. Increases in impervious surfaces lessen the influence of soil moisture on streamflow, thus providing an explanation for the southern watersheds exhibiting a weaker relationship between the two.

The ClimDEX indices were able to identify that southern watersheds will experience a greater magnitude and more intense multi-day precipitation events than the northern watersheds. In general, the more northern watersheds more often exhibit decreases or lesser increases in both total monthly precipitation and RX5DAY than the southern-most watersheds. This reflects an increase in precipitation intensity in the

southern watersheds. This observation coincides with that observed in the changes in PAF, where the more southern watersheds show greater increases in the magnitude of streamflow than the northern watersheds. In terms of frequency, the winter streamflow POT analysis indicates greater increases in the northern watersheds than the southern watersheds; however, because the northern watersheds tend to have PAF occurring in the spring, this does not translate to increases in the magnitude of PAF at these watersheds. As noted by Ivancic and Shaw et al. 2015, increases winter streamflow POT may also be indicative of an increase in soil moisture; suggesting that precipitation in the subsequent seasons may lead to greater discharges despite a lack of increasing flood frequency in such seasons. Although greater discharges may not be characterized as a flood (or peak annual flow) in the winter, the changing seasonality of such events may bring about a new set of challenges for managing resources and mitigation strategies.

## **6 RECOMMENDATIONS & FUTURE WORK**

Although there may be uncertainties associated with the emissions scenario, the selected GCMs, the climate downscaling method, the hydrological model and its parameters, and natural internal climate variability, it is argued that the most uncertainty originates from the GCM structure [*Kay et al.*, 2009]. Quantifying uncertainty is essential in ensuring accurate model results, and provides valuable insight in decision-making and risk assessments. Although this research does not attempt an uncertainty analysis, the projected changes in precipitation and flood events can still be made use of through bottom up decision scaling [*Brown et al.*, 2012]. Once decision-makers and stakeholders identify vulnerable states of a system using relevant parameters to a

particular climate change problem or plan, such climate information or projected changes in extreme events can reveal and prioritize risk mitigation strategies.

Future work should include quantifying uncertainty in the above-mentioned sources to further the understanding of GCM and downscaling uncertainty in the Northeast U.S. region [Wilby and Harris, 2006]. Additionally, some regions have suggested that the natural variability of the hydrological system may still be larger than that brought by climate change [Wilby and Harris, 2006; Kay et al., 2009; Sjerps et al., 2017]. However, proper partitioning between variability and uncertainty will identify regions that are vulnerable to extreme events; whether it is from natural variability, or climate change. In addition to furthering analyses on the model results, it is important to translate these changes in climate and streamflow to actionable science appropriate for stakeholders and policy making. One approach to accomplishing this is incorporating these projected changes into a systems model, where future conditions can be tested against status quo to identify the most vulnerable states of the system to develop mitigation strategies.

## REFERENCES

- Arnell, N. W. (2003), Relative effects of multi-decadal climatic variability and changes in the mean and variability of climate due to global warming: Future streamflows in Britain, *J. Hydrol.*, 270(3–4), 195–213, doi:10.1016/S0022-1694(02)00288-3.
- Brown, C., Y. Ghile, M. Laverly, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, *Water Resour. Res.*, 48(9), 1–12, doi:10.1029/2011WR011212.
- Criss, R. E., and W. E. Winston (2008), Application of a mass-balance model to a Himalayan glacier, *Hydrol. Process.*, (22), 2723–2725, doi:10.1002/hyp.
- D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith (2007), Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Trans. ASABE*, 50(3), 885–900, doi:10.13031/2013.23153.
- Demaria, E. M. C., R. N. Palmer, and J. K. Roundy (2016), Regional climate change projections of streamflow characteristics in the Northeast and Midwest U.S., *J. Hydrol. Reg. Stud.*, 5, 309–323, doi:10.1016/j.ejrh.2015.11.007.
- DeVantier, B. A., and A. D. Feldman (1993), Review of GIS Applications in Hydrologic Modeling, *J. Water Resour. Plan. Manag.*, 119(2), 246–261.
- Gochis, D., W. Yu, and D. Yates (2015), *WRF- Hydro Technical Description and User's Guide Version 3*.
- Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003.
- Gutmann, D. A., T. Pruitt, M. Clark, L. Brekke, J. Arnold, D. Raff, and R. Rasmussen (2014), An Intercomparison of Statistical Downscaling Methods used for water resource assessments in the United States, , 7167–7186, doi:10.1002/2014WR015559.Received.
- Horton, R., G. Yohe, W. Easterling, R. Kates, M. Ruth, E. Sussman, A. Whelchel, D. Wolfe, and F. Lipschultz (2014), Chapter 16: Northeast, *Clim. Chang. Impacts United States Third Natl. Clim. Assess.*, (October), 1–24, doi:10.7930/J0SF2T3P.On.
- Houska, T., P. Kraft, A. Chamorro-Chavez, and L. Breuer (2015), SPOTting model parameters using a ready-made python package, *PLoS One*, 10(12), 1–22, doi:10.1371/journal.pone.0145180.
- Hwang, S., and W. D. Graham (2014), Assessment of Alternative Methods for Statistically Downscaling Daily GCM Precipitation Outputs to Simulate Regional Streamflow, *J. Am. Water Resour. Assoc.*, 50(4), 1010–1032, doi:10.1111/jawr.12154.
- Ivancic, T. J., and S. B. Shaw (2015), Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge, *Clim. Change*, 133(4), 681–693, doi:10.1007/s10584-015-1476-1.
- Jang, S., and M. Kavvas (2015), Downscaling Global Climate Simulations to Regional Scales: Statistical Downscaling versus Dynamical Downscaling, *J. Hydrol. Eng.*, 20(1), A4014006, doi:10.1061/(ASCE)HE.1943-5584.0000939.
- Karmalkar, A. V. et al. (2017), Consequences of Global Warming of 1.5 °C and 2 °C for

- Regional Temperature and Precipitation Changes in the Contiguous United States, *PLoS One*, 12(1), e0168697, doi:10.1371/journal.pone.0168697.
- Kaufman, S., C. Qing, N. Levenson, and M. Hanson (2012), Transportation During and After Hurricane Sandy, *Rudin Cent. Transp. NYU Wagner Grad. Sch. Public Serv.*, (November), 1–36.
- Kay, A. L., H. N. Davies, V. A. Bell, and R. G. Jones (2009), Comparison of uncertainty sources for climate change impacts: Flood frequency in England, *Clim. Change*, doi:10.1007/s10584-008-9471-4.
- Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier (2013), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions, *J. Clim.*, 26(23), 9384–9392, doi:10.1175/JCLI-D-12-00508.1.
- Mallakpour, I., and G. Villarini (2015), The changing nature of flooding across the central United States, *Nat. Clim. Chang.*, doi:10.1038/nclimate2516.
- Melillo, J. M., T. (T. C. . Richmond, and G. Yohe (2014a), *Climate Change Impacts in the United States: The Third National Climate Assessment*, U.S. Global Change Research Program.
- Melillo, J. M., T. (T. C. . Richmond, and 2014 Gary W. Yohe, Eds. (2014b), *Climate Change Impacts in the United States Climate Change Impacts in the United States*.
- Mitchell, K. et al. (2005), *Noah Land Surface Model (LSM) User's Guide*.
- Mitchell, K. E. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, 109(D7), D07S90, doi:10.1029/2003JD003823.
- NARCCAP (2007), No Title, *About NARCCAP*. Available from: <http://narccap.ucar.edu/about/index.html>
- Palmer, R. N., H. E. Cardwell, M. A. Lorie, and W. Werick (2013), Disciplined planning, structured participation, and collaborative modeling - applying shared vision planning to water resources, *J. Am. Water Resour. Assoc.*, doi:10.1111/jawr.12067.
- Parr, D., and G. Wang (2014), Hydrological changes in the U.S. Northeast using the Connecticut River Basin as a case study: Part 1. Modeling and analysis of the past, *Glob. Planet. Change*, 122, 208–222, doi:10.1016/j.gloplacha.2014.08.009.
- Qiao, L., Y. Hong, R. McPherson, M. Shafer, D. Gade, D. Williams, S. Chen, and D. Lilly (2014), Climate Change and Hydrological Response in the Trans-State Oologah Lake Watershed-Evaluating Dynamically Downscaled NARCCAP and Statistically Downscaled CMIP3 Simulations with VIC Model, *Water Resour. Manag.*, 28(10), 3291–3305, doi:10.1007/s11269-014-0678-z.
- Shaw, S. B., and S. J. Riha (2011), Assessing possible changes in flood frequency due to climate change in mid-sized watersheds in New York State, USA, *Hydrol. Process.*, 25(16), 2542–2550, doi:10.1002/hyp.8027.
- Sjerps, R. M. A., T. L. ter Laak, and G. J. J. G. Zwolsman (2017), Projected impact of climate change and chemical emissions on the water quality of the European rivers Rhine and Meuse: A drinking water perspective, *Sci. Total Environ.*, 601–602(June), 1682–1694, doi:10.1016/j.scitotenv.2017.05.250.
- Small, D., S. Islam, and R. M. Vogel (2006), Trends in precipitation and streamflow in the eastern U.S.: Paradox or perception?, *Geophys. Res. Lett.*, 33(3), 2–5,

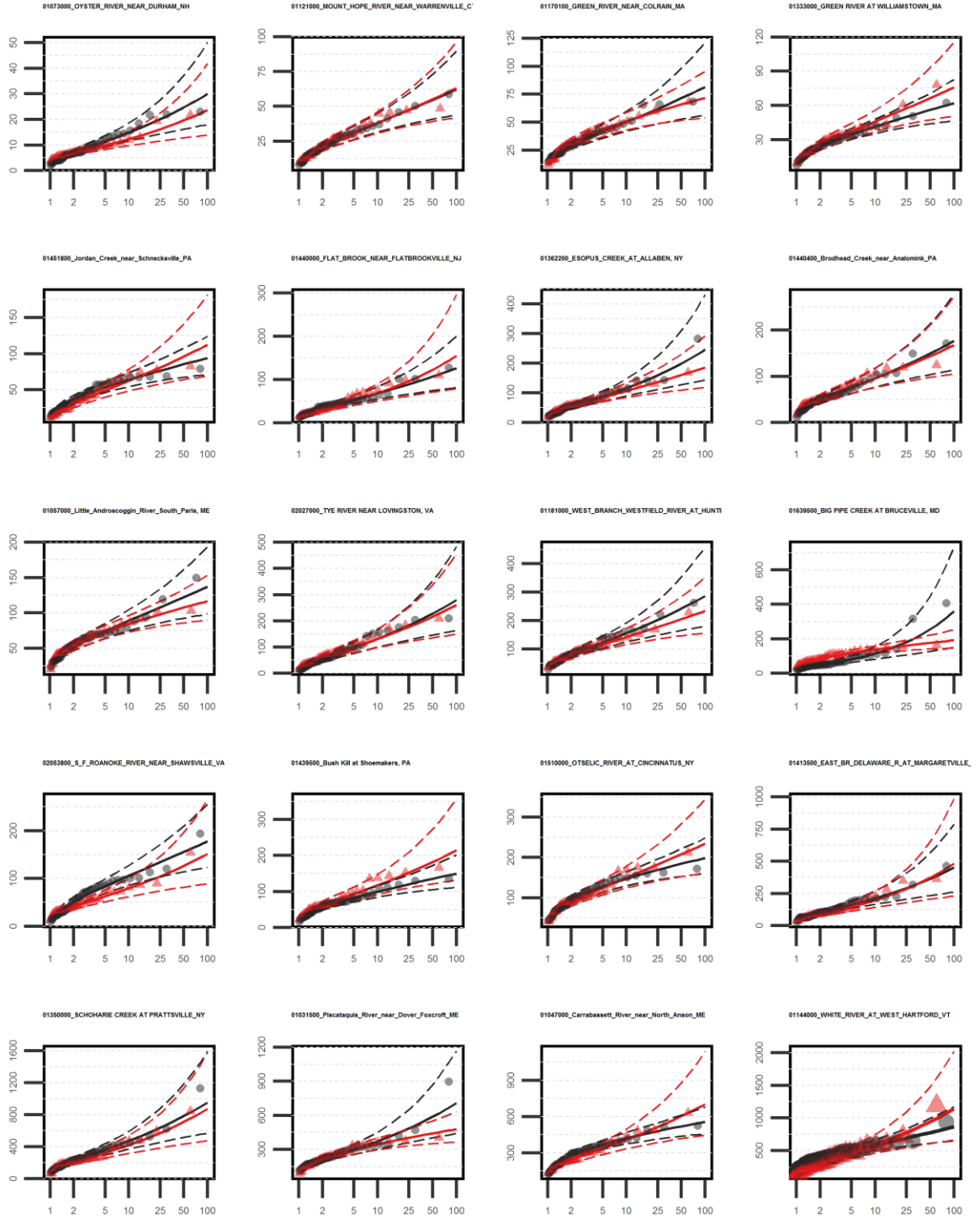


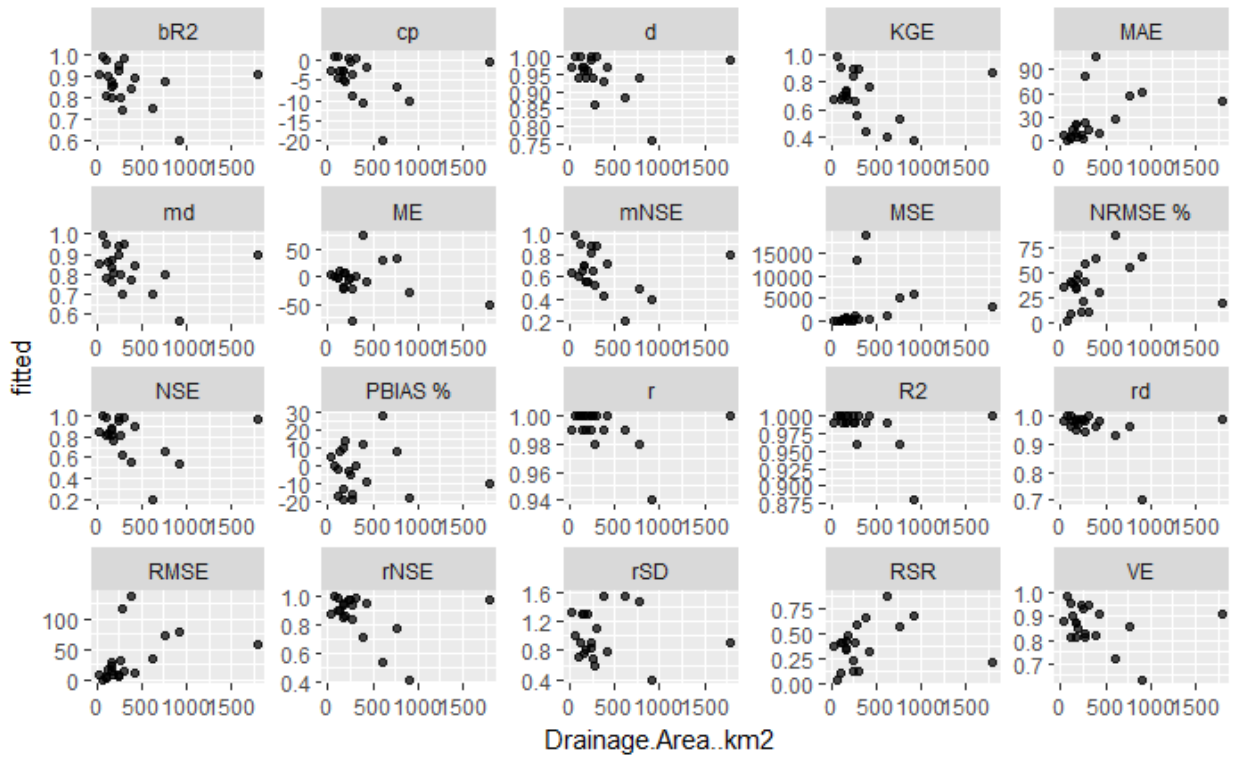
- doi:10.1029/2005GL024995.
- Tryhorn, L., and A. Degaetano (2011), A comparison of techniques for downscaling extreme precipitation over the Northeastern United States, *Int. J. Climatol.*, *31*(13), 1975–1989, doi:10.1002/joc.2208.
- Wasko, C., and A. Sharma (2017), Global assessment of flood and storm extremes with increased temperatures, *Sci. Rep.*, *7*(1), 7945, doi:10.1038/s41598-017-08481-1.
- Werner, A. T., and A. J. Cannon (2016), Hydrologic extremes - An intercomparison of multiple gridded statistical downscaling methods, *Hydrol. Earth Syst. Sci.*, *20*(4), 1483–1508, doi:10.5194/hess-20-1483-2016.
- Wilby, R. L., and I. Harris (2006), A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, *Water Resour. Res.*, *42*(2), 1–10, doi:10.1029/2005WR004065.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004), Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Clim. Change*, *62*(1–3), 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.
- Zhang, X. (n.d.), ETCCDI/CRD Climate Change Indices, *Clim. Res. Div. Environ. Canada*. Available from: <http://etccdi.pacificclimate.org/index.shtml> (Accessed 1 January 2018)

# APPENDIX A

## WRF-HYDRO CALIBRATION RESULTS

page 1 of 1





## APPENDIX B

### EQUATIONS AND CODES

#### Details

$$KGE = 1 - ED$$

$$ED = \sqrt{(s[1] * (r - 1))^2 + (s[2] * (vr - 1))^2 + (s[3] * (\beta - 1))^2}$$

$r$  = Pearson product-moment correlation coefficient

$$\beta = \mu_s / \mu_o$$

$$vr = \begin{cases} \alpha & , \text{method}="2009" \\ \gamma & , \text{method}="2012" \end{cases}$$

$$\alpha = \sigma_s / \sigma_o$$

$$\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s / \mu_s}{\sigma_o / \mu_o}$$

Kling-Gupta efficiencies range from -Inf to 1. Essentially, the closer to 1, the more accurate the model is.

<https://cran.r-project.org/web/packages/hydroGOF/hydroGOF.pdf>

#### Details

$$NSE = 1 - \frac{\sum_{i=1}^N (S_i - O_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}$$

The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance ("noise") compared to the measured data variance ("information") (Nash and Sutcliffe, 1970).

NSE indicates how well the plot of observed versus simulated data fits the 1:1 line.

Nash-Sutcliffe efficiencies range from -Inf to 1. Essentially, the closer to 1, the more accurate the model is.

- ) NSE = 1, corresponds to a perfect match of modelled to the observed data.
- ) NSE = 0, indicates that the model predictions are as accurate as the mean of the observed data,
- ) -Inf < NSE < 0, indicates that the observed mean is better predictor than the model.

<https://cran.r-project.org/web/packages/hydroGOF/hydroGOF.pdf>

# APPENDIX C

## SUPPLEMENTARY RESULTS FIGURES

