Is Hot Drought a Risk in the US Mid-Atlantic? - a Potomac Basin Case Study

C. L. Schultz[1], A. Seck[1], and S. N. Ahmed[1]

[1]Interstate Commission on the Potomac River Basin, Section for Cooperative Water Supply Operations on the Potomac, 30 West Gude Drive, Suite 450, Rockville, MD 20850

C. L. Schultz (cschultz@icprb.org)

Key Points:

- We estimate changes in the cumulative probability distribution of annual Potomac River flows based on an ensemble of climate projections.
- Although long-term average flow is projected to increase, annual flow decreases in an extreme drought year in most of our scenarios.
- Our method can provide annual flow scaling factors which can be used to construct inputs for water supply planning models.

(The above elements should be on a title page)

Abstract

We propose a new nonparametric approach for assessing future changes in annual stream flows in extreme drought years based on an ensemble of climate projections. We apply the method to the Potomac River basin, investigating whether future flows in the river may be impacted by "hot drought", that is, increasing severity of hydrological drought caused by rising temperatures coupled with variability in precipitation. Long time series representative of annual climate in time periods of interest are constructed by pooling and concatenating shorter time series sampled from an ensemble of bias corrected and spatially downscaled climate projections, where the K-nearest neighbor method is used to select pool members. The pooled time series are of sufficient length to allow estimation of the probability distribution of a full range of future annual flows, including 1st percentile values, indicative of flow in an extreme drought year. An empirically derived climate response function for annual mean flow is used as this study's simple hydrologic model. The resulting set of cumulative probability distributions can be used to compute scaling factors for future annual Potomac River flow which demonstrate the disparate impacts of climate change on high flow, average flow, and low flow years. For most scenarios considered, results indicate that though long-term mean precipitation and river flow will increase modestly in future years, annual flows in an extreme drought year will decrease. This new approach can provide multi-model consensus inputs for water supply planning models to support decision-making regarding new infrastructure for climate resilience.

Plain Language Summary

The Potomac River, located in the Mid-Atlantic region of the United States, is the primary source of drinking water for the Washington, DC, metropolitan area. Climate change is expected to bring moderately wetter conditions, on average, to the Potomac basin, but year to year variations in rainfall combined with rising temperatures could result in years in which river flows are lower than ever experienced in the past, termed by some as "hot drought". We propose an approach to better understand future changes in river flows that distinguishes the disparate impacts of climate change on wet years, average years, and dry years. Teams of scientists around the world have built computer models to simulate future climate conditions, and because it's not possible to determine which of these models produces the best predictions, our approach incorporates results from many global climate models. We examine results for a number of scenarios which reflect uncertainty in future global carbon emissions and uncertainty in the physical response of watershed processes to rising temperatures. We find that for most of the future scenarios we consider, river flows will fall in extreme drought years while increasing in average and in wet years.

1 Introduction

Extreme events are key drivers in the development of water management strategies, and water supply planners need tools to help better understand the effects of climate change on future extreme drought (Ehsani et al., 2017; Watts et al., 2012; Zeff et al., 2016). Projected changes in long-term average precipitation and streamflow can provide a first look at a region's water availability under future climate, and global studies indicate that increases are likely to be experienced in some regions and decreases in others (P. C. D. Milly et al., 2005; Tang & Lettenmaier, 2012). But projected increases in meteorological and hydrological variability, at multiple time scales, will potentially lead to more severe and frequent flooding and also to increases in extreme drought (Fowler et al., 2003; Kay et al., 2021; Tebaldi et al., 2006). At the regional scale, where projections for long-term mean precipitation may vary widely or even indicate that modest increases are to be expected, studies show that hydrological droughts may become more severe due to natural variability in precipitation coupled with increasing temperatures (Hayhoe et al., 2007; McCabe et al., 2017; Xue & Ullrich, 2022), threatening to result in serious events which have been characterized as "hot drought" (Udall & Overpeck, 2017).

A look at projected changes over time of the cumulative probability distributions of climatologic and hydrologic variables can give a more complete picture of how future climate will impact water resources, with the lower tails of precipitation and flow distributions providing information on drought severity and the upper tails indicating changes in wet weather events. At the global scale, spatial patterns of changes in statistics representing low flow conditions have been found to differ from patterns of average values (Arnell & Gosling, 2013; Döll & Schmied, 2012; Hirabayashi et al., 2008). Hayhoe et al. (2007) examined the probability distribution of annual streamflows as part of their study on the impact of climate change in the northeast United States and found that the two tails of the distribution exhibited opposite trends, with flows increasing in medium to high flow years, represented by the 50th and 95th percentiles, and flows decreasing in low flow years, represented by the 25th and 5th percentiles.

The Washington, DC, metropolitan area (WMA) relies on the Potomac River as its primary water source, and there are indications that average conditions in the Potomac basin will become wetter. Basin-wide drought is relatively infrequent, with two severe droughts occurring in 1930 and 1966 and two moderate droughts in 1999 and 2002. Reservoir storage for the WMA system is sufficient to meet supply needs and environmental requirements during moderate or short-term drought, and planning decisions are to a large degree driven by the risk of future extreme drought (Ahmed et al., 2020). The Potomac River basin is in the United States' Mid-Atlantic region and situated within the larger Chesapeake Bay watershed. Projections of future precipitation for the Chesapeake Bay region have varied widely (Najjar et al., 2009; Pyke & Najjar, n.d.), with more recent projections indicating that long-term average precipitation will increase (Shenk et al., 2021) while there continues to be uncertainty regarding the sign of change in future stream flows (Hinson et al., 2022). Analyses of historical data indicate that the Potomac basin has been in a transition region with respect to changes in climate and hydrology over the past century. Trend analyses of observed precipitation in areas along the eastern coastal region of the United States generally show increases to the north of the basin and decreases to the south (Yang et al., 2015). Historical trends in streamflow metrics in the Chesapeake Bay watershed indicate that both annual mean flows and low flows are increasing in most areas north of the Potomac basin and decreasing in most areas south of the basin (Fleming et al., 2021; Rice & Hirsch, 2012) .

In a changing climate where the assumption of stationarity is no longer valid (P. C. D. Milly et al., 2008), new tools are needed to characterize trends in extreme quantile values, for example, values in the lower tails of the cumulative probability distributions of streamflow associated with extreme hydrologic drought. Trends in long-term climatological and hydrological statistics are typically investigated based on results computed from time periods that are several decades in length. But this approach is limited by sample sizes too small to compute, for example, the 0.01 quantile value (1st percentile value) of annual streamflow, whose magnitude has a probability of one percent of not being exceeded. This value of annual streamflow is indicative of an extreme drought event in the Potomac basin which in the past, based on the assumption of an unchanging climate, would have been associated with a 100-year return period and referred to as a "100-year drought". Parametric approaches are sometimes used to estimate extreme quantile values , but these rely on assumptions about the probability distributions which may or may not be valid.

To investigate future trends in extreme hydrologic drought, we propose a new nonparametric approach based on an ensemble of downscaled climate projections which provides multi-model consensus estimates of the cumulative probability distribution of future annual river flow. We begin with a multi-model ensemble of downscaled climate projections derived from General Circulation Models (GCMs). For a given time window of interest, we select multiple time series of annual temperature and precipitation from the ensemble and concatenate them to create long pooled climate time series representative of conditions in the window. The pooled time series are of sufficient length to allow investigation of trends in statistics indicative of the severity of future extreme drought. The K-nearest neighbor (K-NN) method is used in the selection process. K-NN is a nonparametric classification method used in many fields, including pattern recognition (Cover & Hart, 1967; Fix & Hodges, 1951) and machine learning (see review by Weinberger and Saul, 2009), with applications in climatology and hydrology including rainfall-runoff forecasting (Karlsson & Yakowitz, 1987), stochastic weather and climate generation models (Groves et al., 2008; Rajagopalan & Lall, 1999; Sharif & Burn, 2006; Yates et al., 2003), and generation of synthetic streamflow time series (Lall & Sharma, 1996; Prairie et al., 2006). K-NN has been used to produce long duration daily time series by resampling output from a single regional climate model (Leander & Buishand, 2007), but not by sampling time series from an ensemble of models, as is proposed in our study. Combining results from multiple global climate models and from multiple runs of a single global model is common in weather and seasonal climate forecasting, where it has been found that multi-model consensus forecasts can provide superior performance and skill over predictions derived from single models (Hagedorn et al., 2005; Krishnamurti et al., 1999), with similar conclusions reached in studies on streamflow forecasting (Baker et al., 2021; Block et al., 2009; Georgakakos et al., 2004; Regonda et al., 2006). It has been argued that this is because multiple model ensembles not only better reflect uncertainty in initial conditions (Toth & Kalnay, 1993), but also may improve forecasts by incorporating variations in model physics and numerics into consensus forecasts (Fritsch et al., 2000; Hagedorn et al., 2005). Multi-model ensembles have also been used to estimate future changes in long-term statistics related to drought (Rashid et al., 2020).

A climate response function (CRF) for the Potomac River watershed is used in this study as a simple hydrologic model to predict annual stream flow from projections of annual climate. The CRF was developed using multiple regression analyses of historic streamflow, temperature and precipitation (P. Milly et al., 2018; Revelle & Waggoner, 1983; Risbey & Entekhabi, 1996). CRFs have also been developed by perturbing the inputs of land surface or other hydrologic models to predict changes in long-term mean streamflow for a range of future climate scenarios (Nash & Gleick, 1991; Schaake, 1990; Vano et al., 2012). Use of a simple CRF allows the processing of a large number of climate projections, thus providing the computational efficiency to support a risk-based multi-model analysis (Brown et al., 2012). Sample sets of annual flow time series can thus be constructed by using the pooled climate time series as inputs to the CRF. Statistics for these sample sets are computed by successive time windows covering the simulation period to examine trends in the severity of extreme drought.

2 Study Area

The Potomac River is the primary source of drinking water for Washington, DC, and its adjacent suburbs in Maryland and Virginia, providing on average 78% of the water demand of the region's three major water suppliers (Ahmed et al., 2020), who participate in a cooperative, interstate system of drought planning and management (Hagen et al., 2005; Sheer, D.P. & Flynn, K., 1983). The Potomac basin is in the Mid-Atlantic region of the United States, covering parts of the states of Maryland, Pennsylvania, Virginia, and West Virginia, as well as the District of Columbia. Land use is relatively undeveloped, with 53% forest and 26% agriculture (Moltz et al., 2020). Flow in the freshwater portion of the Potomac River is measured at the USGS stream gage at Little Falls dam near Washington, DC, located just below the intakes of the metropolitan area water suppliers and a few kilometers above the head of tide in the Potomac estuary. With few major impoundments in the 29,940 km^2 drainage area above Little Falls, river flow is largely unregulated and highly variable (Cummins et al., 2010). Some degree of storage is provided by the underlying fractured bedrock aquifers, but baseflow recession rates are typically on the order of months so this storage can be rapidly depleted during periods of low precipitation (Schultz et al., 2014). Precipitation above Little Falls averages 1024 mm annually, with evapotranspiration averaging 65%. Precipitation is fairly uniform throughout the year, but river flow exhibits a pronounced seasonal variation due to higher evapotranspiration rates during the March through September growing season which reduce both groundwater recharge and runoff (Trainer & Watkins, 1975). Flow tends to be highest in the month of March, with a long-term mean of 671 m3/s, and lowest in September, with a long-term mean of 110 m3/s. The snowpack that may accumulate at higher elevations during the winter months slightly increases median river flows in March and April but does not persist long enough to have a significant impact on summertime flows (Cummins et al., 2010). Basin-wide drought is fairly infrequent in the Potomac. The most prolonged severe drought in the historic record began in the summer of 1930 and persisted through December of that year, with average annual precipitation in the watershed above Little Falls at just 54% of its long-term mean. The second most serious drought occurred in 1966, when daily river flow fell to its lowest recorded value. But unlike the drought of 1930, the drought of 1966 as well as two moderate droughts which occurred in 1999 and 2002 all ended in the late summer/early fall with the onset of weather events related to tropical storms.

3 Data Sources

Historical annual climate and streamflow time series from 1896 through 2017 are used to develop the CRF. The flow time series represents “natural” Potomac River flow at the US Geological Survey’s stream gage near the Washington, DC, Little Falls pump station (Station No. 01646500), located just downstream of the WMA water supply intakes, and in this paper, Potomac River flow will refer to natural flow. To estimate natural flows at Little Falls, the starting point was the USGS’s flow data for Potomac River (adjusted) near Washington, DC (Station No. 01646502), which is based on observed flows at Little Falls, with adjustments made to account for water supply diversions near the WMA. Then amounts equal to estimated upstream consumptive losses were added and estimated effects of two large reservoirs in the watershed were removed: Savage Reservoir in Alleghany County, Maryland, completed in 1952, and Jennings Randolph Reservoir, situated between Alleghany County, Maryland, and Mineral County, West Virginia, completed in 1982. Annual flows for the early years of the Little Falls time series, February 1895 through February 1930, were reconstructed using data from two upstream gages, the Potomac River near Point of Rocks, Maryland (Station No. 01638500) and the Monocacy River near Frederick, Maryland (Station No. 01642000).

For the historical climate time series, this study relied on the Precipitation-Elevation Regression on Independent Slopes (PRISM) model dataset from the PRISM Climate Group at Oregon State University (Daly et al., 2008). PRISM uses climate observations from a wide range of monitoring networks and a series of regression models to develop spatially explicit climate maps at a regional scale. Monthly time series of 4 x 4 km PRISM gridded data for air temperature and precipitation were downloaded for the time period, 1895-2017 (available at https://prism.oregonstate.edu/historical/ for 1895-1980 and https://prism.oregonstate.edu/recent/ for 1981-2017). Values were spatially averaged over the Potomac River drainage area above Little Falls to create time series of monthly basin-wide averages of temperature and precipitation. In the initial development of the CRF, results obtained using PRISM were compared with those obtained using a second dataset: the CLIMGRID dataset from NOAA's National Centers for Environmental Information (NCEI) (Vose et al., 2014). The CLIMGRID dataset is derived from NOAA's Global Historical Climatology Network (GHCN) and is recommended for calculations of regional climate trends. Gridded 5km GHCN-Daily Temperature and Precipitation Dataset (nClimGrid/CLIMGRID), version 1 (available at https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00332) were downloaded for the period 1895-2013. To obtain the Potomac basin CRF (see below), regression results were computed from both PRISM and nClimGrid data and compared. The best-performing regression model was obtained from the PRISM set of annual (calendar year) precipitation and temperature time series. For this model, the Nash-Sutcliffe efficiency (NSE) is 0.77, compared with 0.55 for the model derive from the nClimGrid dataset. Because of its stronger performance in the Potomac basin, the historical PRISM climate dataset was relied upon in this study.

The climate projections used in this study are derived from the Coupled Multi-model Intercomparison Project, Phase 5 (CMIP5), statistically downscaled using monthly bias-correction and spatial disaggregation (BCSD) (Reclamation, 2013), available from the "Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections" archive (available at https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html). The BCSD data are monthly time series for precipitation and air temperature extending from 1950 through 2099 for a grid of 1⁄8 degree by 1⁄8 degree, providing a spatial resolution of approximately 12 x 12 km. An ensemble of 231 BCSD projections, clipped and spatially averaged over the drainage area of the Potomac basin upstream of the USGS stream gage at Little Falls Pump station near Washington, DC (38.9375 degrees north and -77.1875 degrees west), were downloaded from the archive. The ensemble included runs from 36 global climate models for four representative concentration pathways: RCP 2.6 (53 members), RCP 4.5 (71 members), RCP 6.0 (37 members), and RCP 8.5 (70 members).

As a verification and filtering step, the nonparametric Kolmogorov-Smirnov test was applied to compare the empirical distribution function (ecdf) of observed and simulated climate in the Potomac basin for each BCSD ensemble member in the reference period, 1950-1979, and also for the subsequent 30-year period, 1980-2009, where the observed distributions were calculated from PRISM data. For 1950-1979, there were no significant differences in observed versus simulated distributions of annual precipitation or annual temperature for any of the 231 ensemble members at the $p = 0.05$ significance level. But for 1980-2009 there were significant differences in the case of 86 ensemble members. After discarding these, 145 members remained in the filtered BCSD ensemble: RCP 2.6 (33 members), RCP 4.5 (46 members), RCP 6.0 (25), and RCP 8.5 (41 members).

4 Methods

We present a new nonparametric approach to investigating future trends in extreme values of annual climatic and hydrologic variables by sampling and combining time series from an ensemble of annual climate projections. For a given time window several decades in length, longer "pooled" time series of annual climate are constructed by concatenating multiple ensemble member projections selected using the K-NN method. K-NN is used to sample multi-decadal climate sequences in a manner that treats sequences from different GCMs or different runs of a single GCM as different possible realizations of the climatic future and aims to capture the potential serial correlation of annual climate in a given time window. A CRF, developed from a regression analysis of historical data, then serves as a simple hydrologic model to obtain corresponding long pooled time series of annual streamflow. Each pool is representative of conditions in the shorter time window, but the length of each pooled time series is sufficient to allow estimates to be made of extreme quantile values, including values representative of annual flow in an extreme drought year.

4.1 Construction of pooled annual climate time series

Assume that an ensemble of projections of annual climate has N members that are each N_t years in length, beginning at year t_1 and ending at year t_{Nt}, where $(t_{Nt} - t_1 + 1) = N_t$, and assume this time period can be divided into W time windows of equal length, L, where L will typically be several decades. Then $N_t = L*W$ where L and W are positive integers. Let λ be an index which denotes one of the N ensemble members, that is, λ = 1, 2, ..., N, and $P^{\lambda}{}_{i}$ and $T^{\lambda}{}_{i}$ denote annual precipitation and temperature for ensemble member, λ, in the i^{th} year of the simulation period, t_i, where i = 1, 2, ..., N_t. Let $\hat{P}^{\lambda\, w}{}_{j}$ and $\hat{T}^{\lambda\, w}{}_{j}$ denote annual precipitation and temperature in the j^{th} year of the w^{th} time window, where j = 1, 2, …, L, and ω = 1, 2, ..., W. Then

$$\hat{P}^{\lambda\,\omega}{}_{j} = P^{\lambda}{}_{t0 + L*(\omega -1) + j - 1}$$
$$\hat{T}^{\lambda\,\omega}{}_{j} = T^{\lambda}{}_{t0 + L*(\omega -1) + j - 1}$$

Equation 1

Define $\hat{Y}^{\lambda\, w}{}_{j}$ to be the 2-dimensional vector of annual precipitation and temperature in the j^{th} year of the w^{th} time window from ensemble member, λ, that is,

$$\hat{Y}^{\lambda\,\omega}{}_{j} = (\hat{P}^{\lambda\,\omega}{}_{j} , \hat{T}^{\lambda\,\omega}{}_{j})$$

Equation 2

and define $\hat{Y}^{\lambda\, w}$ to be paired precipitation and temperature time series of length L from ensemble member, λ, which we view as an instance of climate conditions and annual variability representative of the w^{th} time window, that is,

$$\hat{Y}^{\lambda\,\omega} = (\hat{Y}^{\lambda\,\omega}{}_{1}, \hat{Y}^{\lambda\,\omega}{}_{2}, \ldots, \hat{Y}^{\lambda\,\omega}{}_{L})$$

Equation 3

Because of the nonstationarity of the climate time series, depending on the time series length, L, it may be preferable to detrend the time series in each time window of interest before constructing the $\hat{Y}^{\lambda_1\,\omega}$.

Long pooled climate time series are created by concatenating multiple time series from ensemble members, each representative of conditions in a shorter window of time. That is, to create a pooled time series representative of conditions in a given time window w, of length L, M ensemble members are concatenated as follows. First a member, λ_1, is randomly selected from the ensemble of annual climate projections, where λ_1 has an L-year climate time series, $\hat{Y}^{\lambda_1\,\omega}$, in time interval, w. Then a second ensemble member, λ_2, is selected using a weighted random sampling process from the K nearest neighbors of λ_1 and concatenated with λ_1. This process is continued until (M-1) additional time segments from the time window, w, have been selected and concatenated, forming the time series ($\hat{Y}^{\lambda_1\,\omega}, \hat{Y}^{\lambda_2\,\omega}, \ldots, \hat{Y}^{\lambda_M\,\omega}$), which is of length, M*L years, and where each ensemble member time segment, λ_i, has been selected from the nearest neighbors of the preceding segment, λ_{i-1}. In this manner, a sample set of N_p climate pools, $\hat{Z}^{\nu\,w}$, can be created for a given time window, w, where $\nu = 1, 2, \ldots, N_p$. Each climate pool, $\hat{Z}^{\nu\,w}$, is a time series L*M years in length consisting of a concatenation of M of the L-year time series, that is, M instances of climate conditions representative of the period, w, selected from the ensemble members. Thus

$$\hat{Z}^{\nu\,\omega} = (\hat{Y}^{\lambda_{\nu_1}\,\omega}, \hat{Y}^{\lambda_{\nu_2}\,\omega}, \ldots, \hat{Y}^{\lambda_{\nu_M}\,\omega})$$

Equation 4

To use the K-NN method in the selection process described above, we follow Lall and Sharma (1996), and identify nearest neighbors by first defining a feature vector and a successor vector for each ensemble member, where the feature vector provides a forecast of the successor vector. For a given L-year climate time series, $\hat{Y}^{\lambda\,w}$, for ensemble member, λ, and time window, w, the feature vector, $\mathbf{x}_{\lambda\,w}$, is chosen to be the annual climate in the last year of the time window, that is, $\hat{Y}^{\lambda\,w}{}_L$, and the successor vector is annual climate in the first year of the next time window, that is, $\hat{Y}^{\lambda\,w+1}{}_1$. Because the aim of the concatenation process is to follow a climate time series, $\hat{Y}^{\lambda\,w}$, with a succeeding time series, $\hat{Y}^{\lambda'\,w}$, in a manner that preserves temporal correlations, nearest neighbors, λ', are defined as ensemble members which have similar values of precipitation and temperature in the last year of the preceding time window, (w -1), that is, based on the feature vectors, $\mathbf{x}_{\lambda'\,w-1}$. In this way, if there are serial correlations present in the climate time series, the last year of climate conditions in the preceding time window will have some ability to forecast climate in the first year of time window, w, and a concatenation of the two climate time series should to some degree reflect any correlations expected between annual climate in the last year of $\hat{Y}^{\lambda\,\omega}$ and the first year of $\hat{Y}^{\lambda'\,\omega}$.

Thus, the feature vector, $\mathbf{x}_{\lambda\,w}$, for ensemble member, λ, and time window, w, is the two-dimensional vector of annual precipitation and temperature in the last, that is the L^{th} year of the time window, that is,

$$\mathbf{x}_{\lambda\,w} = \left(\frac{\hat{P}_L^{\lambda\,\omega}}{\sigma_{P\;\omega}}, \frac{\hat{T}_L^{\lambda\,\omega}}{\sigma_{T\;\omega}} \right)$$

Equation 5

standardized by dividing by standard deviations. Nearest neighbors, λ', are identified based on Euclidean distances between the feature vectors of the current and preceding time windows, $\|\boldsymbol{x}_{\lambda\,\omega} - \boldsymbol{x}_{\lambda'\,\omega-1}\|$, as discussed above. For a given pool, $\hat{Z}^{\nu\,w}$, the first time series, $\hat{Y}^{\lambda_{\nu_1}\,\omega}$, is selected from ensemble members via random sampling. The second series, $\hat{Y}^{\lambda_{\nu_2}\,\omega}$, is selected by random sampling from K nearest neighbors of $\hat{Y}^{\lambda_{\nu_1}\,\omega}$ using a weighted sampling algorithm. The third series, $\hat{Y}^{\lambda_{\nu_3}\,\omega}$ is selected from the K nearest neighbors of $\hat{Y}^{\lambda_{\nu_2}\,\omega}$. We use the common approach of setting K as the square root of the sample size, that is, of the number of ensemble members in the sample. Successive series used in the concatenation are selected using the same procedure. Sampling weights, W, are assigned by defining the kernel as,

$$W_k = \left(\frac{1}{k}\right) \Big/ \sum_{k=1}^{K} \frac{1}{k}$$

Equation 6

where k denotes the indices of an ordered set of K nearest neighbors, and where we note that the sum of the K weights equals 1.

4.2 Climate response function

An empirical climate response function (CRF) was developed, based on historical data, to serve as this study's simple hydrological model for predicting annual mean river flow from climate projections. The CRF was constructed using multiple regression analysis applied to time series of historic streamflow and climate, using a form of a regression equation similar to that of Milly et al. (2018), with annual mean precipitation and air temperature as predictor variables and including a lagged flow term to simulate interannual storage (P. Milly & Dunne, 2002). The resulting CRF for Potomac River flow successfully simulates the historic record and serves as this study's simple hydrologic model for annual mean flow. For a given simulation year i, Potomac River annual natural flow at Little Falls (Q_i) is predicted as a function of annual watershed average precipitation (P_i), annual watershed average air temperature (T_i), and previous year's mean flow. The regression equation differs from Milly et al. in that it includes a quadratic precipitation term to capture potential nonlinear effects (Risbey & Entekhabi, 1996):

$$\frac{(Q_i - \bar{Q})}{\bar{Q}} = \beta_1 \frac{(Q_{i-1} - \bar{Q})}{\bar{Q}} + \beta_2 (T_i - \bar{T}) + \beta_3 \frac{(P_i - \bar{P})}{\bar{P}} + \beta_4 \left(\frac{P_i - \bar{P}}{\bar{P}}\right)^2 + \varepsilon_i$$

Equation 7

where $\bar{Q}$, $\bar{T}$, and $\bar{P}$ are long term average values computed over the historical period used in the analysis, 1896 – 2017.

5 Results

We apply the approach described above to the Potomac River basin above Little Falls dam. To characterize projected changes in climate and river flow over the 150-year period of the BCSD data, 1950 to 2099, we divide the simulation period into five successive 30-year time windows. That is, according to the formalism presented above, we set W=5 and L=30, where N_t=150 = L*W. We define the study "baseline period" to be the first 30-year time interval, 1950-1979, and note that this period represents pre-climate change conditions in the Potomac basin reasonably well. The baseline period is important because in later sections, baseline results will be a starting point to evaluate projected changes in natural annual Potomac River flow. Observed mean precipitation during the baseline period, 992 mm, is essentially equal to mean precipitation over the historic record, 1896-1979, 991 mm, based on PRISM data. Similarly, observed mean temperatures for the baseline versus the longer historic period are similar: 11.04 versus 11.19 °C, a difference of 0.15 °C. Finally, mean observed natural Potomac River flow for the baseline period is 342 mm, which differs by only 1.5% from the mean flow over the historic record, which is 337 mm.

Trends in climate and flow statistics are investigated by RCP. For each RCP, five sample sets of pooled climate time series representative of the five successive 30-year time windows are constructed. Each pool is 180 years in length, formed by concatenation of six 30-year time series selected from the BCSD ensemble by the method described above (that is, M=6), and each pool set contains 100 pools, that is, $N_p = 100$. Results discussed below showed little sensitivity to values of M and N_p. Before concatenation, each of the 30-year temperature and precipitation time series is detrended using linear regression. To apply the K-NN nearest neighbor sampling method, the value of K is set at the square root of the number of members in the RCP sub-ensemble, following Lall and Sharma (1996). Then for each pooled climate time series, a corresponding time series of river flow can be computed by means of the CRF, Eq. 7. An example of a pooled time series is shown in Figure 1: a 180-year time series of annual temperature representing conditions in the period, 2080-2099, constructed by concatenating six 30-year time series for 2080-2099, all selected from the sub-ensemble of RCP 4.5 climate projections. This graph also shows values of components of the feature vectors used in the K-NN selection process (Equation 5), that is, the value of annual temperature in the last year of the time window of interest (in this case, ω=5) and of the preceding time window (ω=4) where nearest neighbors are identified based on Euclidean distances between the feature vectors of the current and preceding time windows.

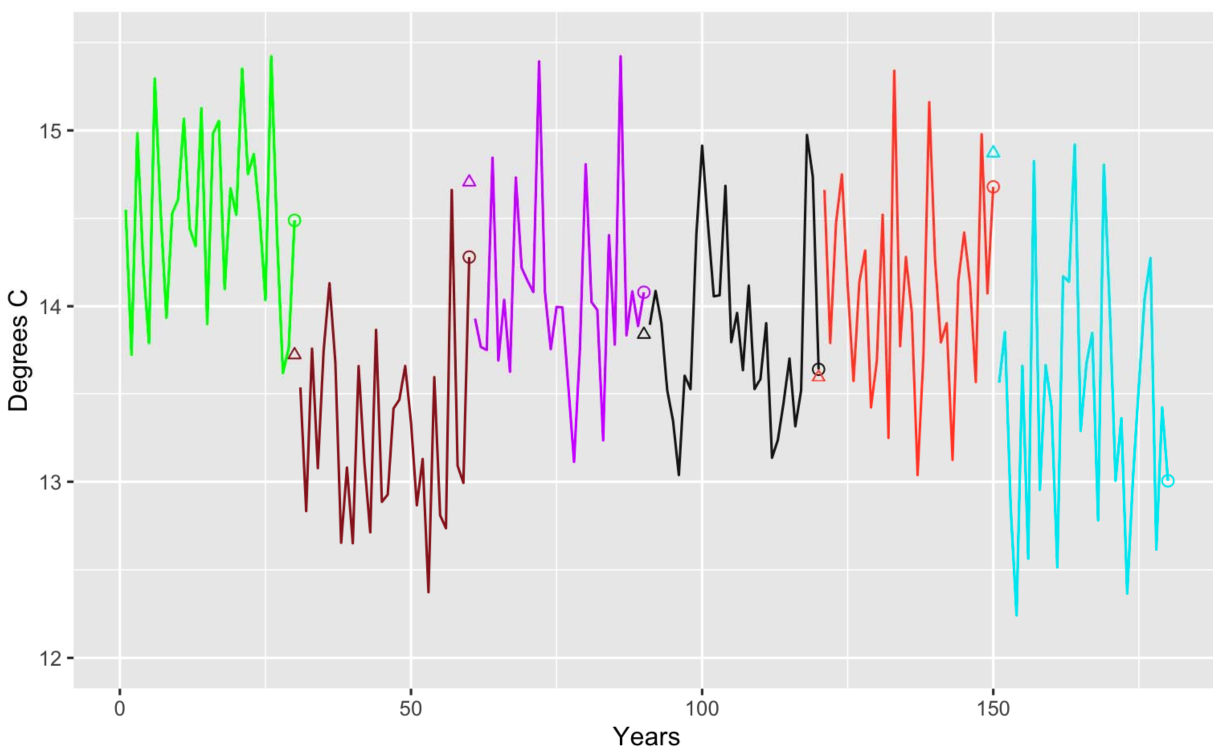

Figure 1: Example of a pool representative of annual temperature in the 30-year window, 2080-2099, where each segment was selected from the RCP 4.5 sub-ensemble. This graph also shows values of the feature vector used in the K-NN selection method - the last value of a 30-year segment (circles) and the last value of the previous window of the next segment (triangles).

5.1 Performance of the pooled climate time series

To evaluate the performance of the pooled time series, we compare long-term statistics of annual temperature and precipitation characterizing 30-year windows of the simulation period computed directly from the filtered ensemble members and computed from the pooled time series. Selected statistics are computed for each 30-year ensemble member time series (n=30) and for each 180-year pooled time series (n=180). The algorithm associated with the Weibull plotting position is used for computing quantiles (type=6 in the quantile function of the R scripting language) because it provides an unbiased estimator (Vogel & Fennessey, 1995), not affected by sample size. For a given 30-year window and a given RCP, this provides sample sets of statistics for 100 pools and for 25 to 46 ensemble members, depending on the size the RCP sub-ensemble. We compare statistics computed from the 30-year time series versus the longer pooled time series first by looking at boxplots of quantile values, showing medians and interquartile ranges, for all RCPs, and then by looking at sample set means of time series means, standard deviations, and Kendall's tau. Figures 2 and 3 show boxplot comparisons for precipitation and temperature of five different quantile values: 0.05, 0.25, 0.50, 0.75, and 0.95. Increases over the simulation period, 1950-2099, are evident for both precipitation and temperature for all displayed quantiles for all RCPs. From the data underlying Figures 2 and 3, the increases in median (0.50 quantile) precipitation from the baseline period, 1950-1979, to the last 30-year window, 2070-2099, range from 11% for RCP 2.6 to 16% for RCP 8.5. Increases in median temperature range from 2.0 degrees C for RCP 2.6 to 5.1 degrees C for RCP 8.5. In Figures 2 and 3, there is a good match between the boxplots in most cases, with an overlap in the range of uncertainty of the median values, as indicated by the boxplot "notches". Exceptions are the somewhat lower pool set values in some of the 0.95 quantile results for precipitation, representative of conditions in very wet years, and conversely, some of the pool set values for the lowest two temperature quantiles, 0.05 and 0.25, which tend to be lower, especially in the case of RCP 8.5. Finally, there is a noticeable difference in the interquartile ranges of the boxplots, with the smaller interquartile ranges of the pool set boxplots indicating that there is more similarity between the pools than between the original ensemble member time series. This is consistent with the fact that the RCP sub-ensembles sampled to create the pools are relatively small: ranging from just 25 members for RCP 6.0 to 46 members for RCP 4.5. This means that in a 100-pool sample set, where each pool consists of a concatenation of M=6 sub-ensemble members, there is repetition in the use of sub-ensemble members, leading to some similarity in the pools. It can be verified that for smaller M's the interquartile range of the pool sets increases and for larger M's it decreases.

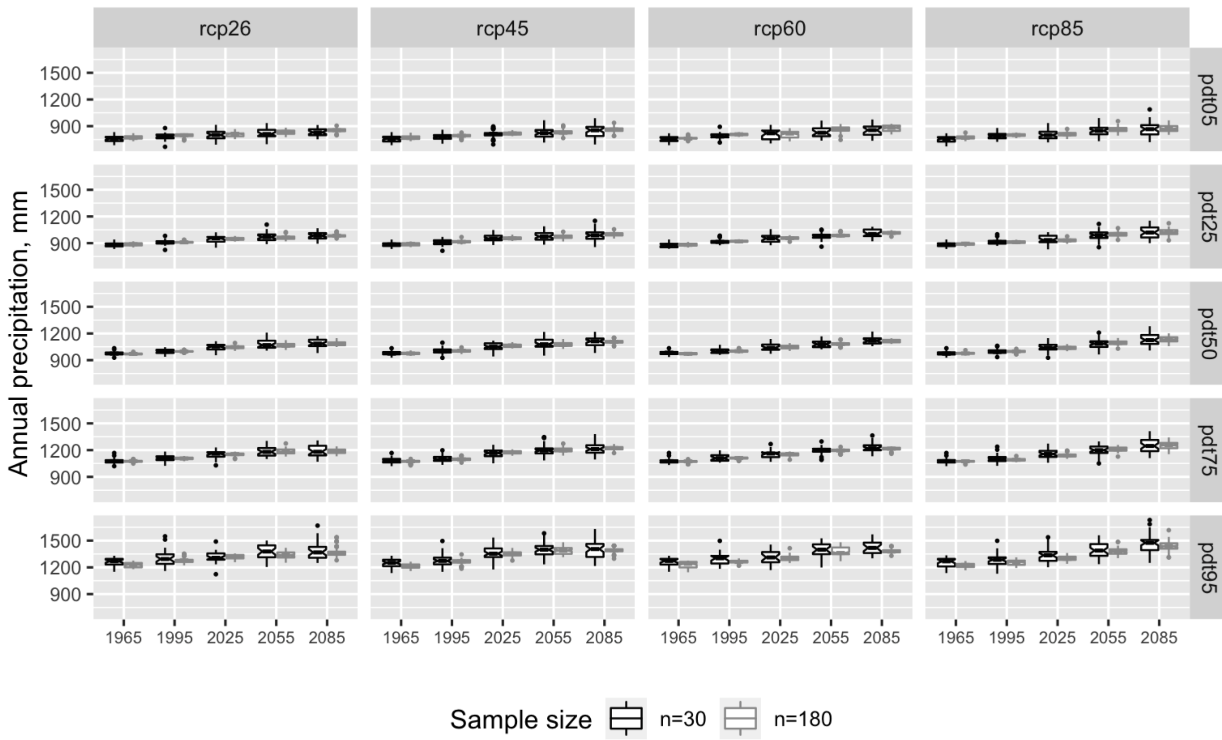

Figure 2: Boxplots of annual precipitation quantile values, 0.05, 0.25, 0.50, 0.75, and 0.95, for the sample sets of the ensemble member time series (n=30) and the pooled time series (n=180), by RCP, where the labels on the x-axis denote midpoints of the 30-year windows.

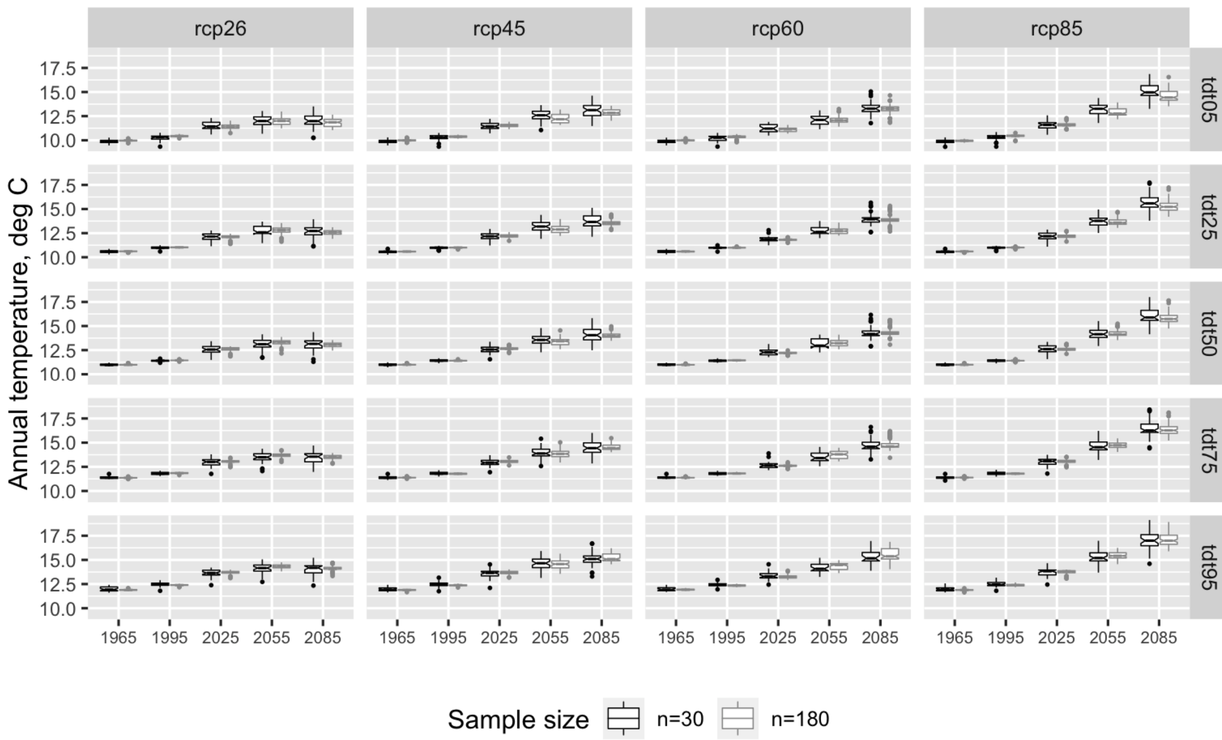

Figure 3: Boxplots of annual temperature quantile values, 0.05, 0.25, 0.50, 0.75, and 0.95, for the sample sets of the ensemble member time series (n=30) and the pooled time series (n=180), by RCP, where the labels on the x-axis denote midpoints of the 30-year windows.

Table 1 provides a more detailed look at trends in selected long-term statistics computed from the ensemble time series and the pooled time series. Results are shown for RCP 4.5, but they are illustrative of results for all four RCPs in terms of the match between the two sets of values. For each 30-year window, ensemble set means of ensemble member time series means (n=30) are very close to pool set means of the pooled time series means (n=180), with differences of less than a half a percent for precipitation and less than 0.1 degrees C for temperature. Ensemble set means of the standard deviations of annual precipitation and temperature ensemble member time series (n=30) grow over time in the case of precipitation and remain relatively constant in the case of temperature. However, pool set means of the pooled time series standard deviations (n=180) are up to 7% larger than the corresponding ensemble means of standard deviations in the case of precipitation and up to 0.2 degrees C in the case of temperature, with the largest discrepancies occurring in the 2070-2099 window. These results indicate that the variability of the pooled time series exceeds that of the original ensemble time series, especially in the later years. An explanation of this is provided by the remaining statistics in Table 1 - the standard deviations of the time series means. These also grow with time in the case of the ensemble statistics, and to a lesser degree for the pool statistics. For example, the standard deviation of long-term mean temperature in the 2070-2099 window is 0.6 degrees C. Thus, the long-term means of ensemble members differ more and more over the course of the simulation period. Because ensemble member time series are being sampled and combined to form the pools, some pools are concatenations of ensemble members with very different 30-year means, especially in the case of temperature, as is visible in the example of a pooled temperature time series shown in Figure 1. In short, in a given pool there may be a considerably wider range of temperatures than in a given ensemble window, due to the variability of the 30-year ensemble means superimposed on the annual variability. On the other hand, it should be noted that multi-decadal variability in climatic conditions in the Potomac basin is not necessarily unrealistic. Paleo reconstructions of Potomac River flow from tree ring chronologies indicate that prior to the instrumental record, considerable multidecadal variability was present (Maxwell et al., 2011; Torbenson & Stagge, 2021).

The K-NN sampling approach used in this study aims to preserve serial correlations that may be present in the original time series of projected climate as they are concatenated into longer pooled time series, as described above. Table 1 provides comparisons of a non-parametric measure of serial correlation, Kendall's tau, for both the sub-ensemble time series and the pooled time series. The value of Kendall's tau for observed climate over the historical period, 1896-2017, is -0.01 for annual precipitation, indicating no significant serial correlation, and 0.11 for annual temperature, indicating serial correlation at the 0.10 significance level. In the case of annual precipitation, the results in Table 1 are consistent with the historical results, with no serial correlation detected in either the n=30 ensemble time series or the n=180 pooled time series for both historical periods and future periods. In the case of annual temperature, neither the original ensemble time series nor the pooled time series indicate the presence of serial correlation for the two historical time windows, 1950-1979 and 1980-2009. However, Kendall tau values for the later three time windows do indicate significant serial correlation of annual temperature in many of the pools. A review of results for individual pools in the 2070-2099 window shows that this effect is related to the increase in the annual standard deviations of long-term temperature means, discussed above. Individual pools that are a concatenations of ensemble members with very different 30-year temperature means have both high standard deviations and high Kendall's tau for annual temperature whereas pools constructed from ensemble members with similar 30-year temperature means have low standard deviations and low Kendall's tau.

Table 1: Ensemble versus pool set means and standard deviations of ensemble member climate time series statistics (n=30) versus pooled climate time series statistics (n=180). Results shown are for RCP 4.5.

		1950-1979	1980-2009	2010-2039	2040-2069	2070-2099
Mean of long-term precipitation means (mm)	n=30	986	1011	1062	1093	1107
	n=180	983	1012	1069	1091	1114
Mean of annual precipitation standard deviations (mm)	n=30	136	139	152	160	152
	n=180	132	141	159	169	163
Standard deviation of long-term precipitation means (mm)	n=30	18	26	37	48	57
	n=180	7	10	14	21	17
Mean of Kendall's tau for annual precipitation time series	n=30	-0.07	-0.01	0.01	-0.07	-0.05
	n=180	-0.04	0.00	0.03	0.01	0.03
Mean of long-term temperature means (°C)	n=30	11.0	11.4	12.5	13.6	14.1
	n=180	11.0	11.4	12.6	13.4	14.0
Mean of annual temperature standard deviations (°C)	n=30	0.6	0.6	0.6	0.6	0.6
	n=180	0.6	0.6	0.6	0.7	0.7
Standard deviation of long-term temperature means (°C)	n=30	0.1	0.1	0.3	0.6	0.7
	n=180	0.0	0.0	0.2	0.4	0.3
Mean of Kendall's tau for	n=30	0.03	0.06	-0.01	-0.03	0.00

annual temperature time series	n=180	0.05	0.05	0.11	0.23	0.23

5.2 Climate response function

The coefficients of the CRF, Eq. 7, were computed using a multiple regression analysis of observed annual flow, precipitation, and temperature data extending from 1896 through 2017. Two different definitions of year were used to convert monthly values to annual values and compared in an effort to obtain the best CRF: calendar year (January 1 through December 31) and USGS water year (October 1 through the following September 30). The best-performing regression model was from the PRISM-calendar time series, resulting in a Nash-Sutcliffe efficiency (NSE) of 0.77, compared with 0.69 for the model derived from the water year datasets. The model coefficients are given in Table 2, along with their standard errors and p-values. The coefficient of determination is 0.76 and the standard error of the prediction is 0.16. A comparison of observed annual Potomac River flow and flow predicted by the regression model is shown in Figure 2.

Table 2: Annual flow regression model coefficients.

Coefficient	**β_1**	**β_2**	**β_3**	**β_4**
Value	0.14	-0.054	1.845	1.027
Standard Error	0.04	0.025	0.114	0.336
p-value	0.002	0.017	4.45E-32	0.003

In Table 2, the value of the coefficient, $ß_1$ is 0.14, representing the portion of annual flow provided by storage from the previous year, and is close to the median, 0.16, of values obtained by Milly et al. (2018) resulting from a similar regression analysis for 2673 basins around the world. The sensitivity of flow to precipitation can be compared with other studies if Eq. 7 is first differentiated and the values for $ß_3$ and $ß_4$ are substituted into the result, giving

$$\frac{\partial\left(\left(Q_i-\bar{Q}\right)/\bar{Q}\right)}{\partial\left(\left(P_i-\bar{P}\right)/\bar{P}\right)} = 1.845 + 1.027\left(\frac{P_i-\bar{P}}{\bar{P}}\right)$$

Equation 8

Thus Eq. 8, with coefficients from Table 2, captures the rising sensitivity of flow to precipitation with increasing precipitation (P. Milly et al., 2018; Revelle & Waggoner, 1983). For the range of values of $P_i/\bar{P}$ in the historic record used in this study, 0.54 to 1.55, the percent change in flow from a 1% change in precipitation from Eq. 8 ranges from 0.9% to 3.0%. For the interquartile range of values of $P_i/\bar{P}$, 0.92 to 1.08, the percent change in flow is 1.7% to 2.0%, which is very similar to results in other studies. Sankarasubramanian and Vogel (Sankarasubramanian & Vogel, 2003) used a nonparametric estimator to compute the ratio of change in flow to change in precipitation for 1337 basins in the United States and found that values generally ranged from 1.0 to 2.5. Finally, the value of the temperature coefficient in Table 2, $ß_2$ = -0.054, indicates that a 1° C increase in annual mean temperature decreases mean annual flow in the Potomac River by 5.4%.

The sensitivity of flow to change in temperature is a crucial factor in determining the impact of climate change on water availability, since projections of future precipitation tend to vary widely in the Potomac, as in many other regions, but future temperatures rise significantly in all projections. Temperature sensitivity affects the degree to which future increases in precipitation can counteract future increases in temperature, and it determines the severity of future hot droughts. But there is debate about the degree to which temperature sensitivity can be accurately estimated. Some studies have indicated that variability in flow is almost completely explained by variability in precipitation, and that the role of temperature is small (Gleick, 1986; Karl & Riebsame, 1989; McCabe & Wolock, 2011). Others point to the significant differences in estimates of temperature sensitivity in widely studied regions such as the upper Colorado River Basin (CRB), where estimates have ranged from -2 percent to -15 percent (P. C. Milly & Dunne, 2020). Recent work for the upper CRB supports the importance of temperature in determining river flow, indicating that the impact of rising temperature is now evident in the instrumental record (McCabe et al., 2017; Udall & Overpeck, 2017; Woodhouse et al., 2016). In the results section below, we explicitly take into account the uncertainty in the sensitivity of flow to changes in temperature by considering three different temperature sensitivity scenarios based on the range of values of the temperature coefficient, β_2, defined by ± one standard error, that is, -2.9% to – 7.9%.

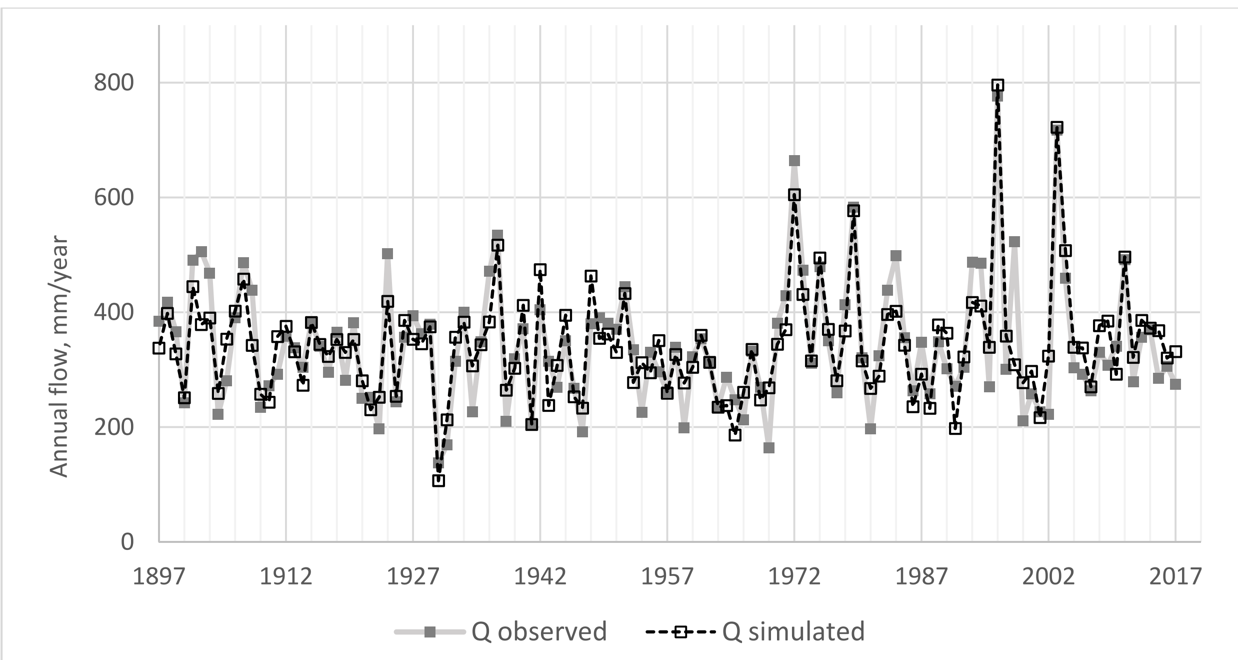

Figure 4: Observed annual Potomac River observed flows compared with flows simulated using the CRF and historical climate.

5.3 Performance of the pooled flow time series

Projected future flows in the Potomac basin are computed by using the climate time series discussed above as inputs to the CRF, Eq. 7, with the regression coefficients given in Table 2. Here we examine long-term statistics for flows in the 30-year windows, comparing results for flows computed directly from the filtered ensemble members (n=30) with those computed from the pooled time series (n=180). We expect that the variability of annual flows in the Potomac basin will increase because of the increasing variability of precipitation. The competing effects of rising precipitation and temperature, as reflected in Eq. 7, can be expected to further increase flow variability.

To investigate trends in projected annual flows, long-term statistics representing the 30-year windows, by RCP, are computed for each pool in the 100-member flow pool set and are compared with those computed for 30-year time intervals from the filtered sub-ensemble representing the RCP. Figure 5 compares box plots of annual flow percentile values obtained by these two methods, ranging from the 5^{th} to the 95^{th} percentile values, by RCP and by successive 30-year time interval. Each "30-year" boxplot characterizes percentile values computed for the sample of sub-ensemble members for each RCP (ranging from N=25 for RCP 6.0 to N=46 for RCP 4.5), each of length n=30. Each corresponding "180-year" boxplot characterizes percentile values computed from a sample of size N_p=100 of pooled time series, each of length n=180.

There is a reasonably good match between the pairs of boxplots in Figure 5, indicating that using the methods proposed in this study, probability distributions of annual flows characteristic of given 30-year time intervals can be constructed that are consistent with those obtained using the conventional method of relying on 30-year time series to estimate long-term statistics. The medians of the sets of annual flow quantile values tend to be slightly higher for the n=180 time series than the n=30 time series in the case of drought years (0.05 quantile), for all 30-year windows, and slightly lower in the case of high flow years (0.95 quantile). This is consistent with the fact that the n=180 time series tend to slightly over-estimate precipitation in very dry years and slightly under-estimate precipitation in very wet years (see Figure 3).

Table 3 provides a more detailed comparison of selected long-term statistics for annual flow computed using the two methods. Again, results are only shown for RCP 4.5, but are illustrative of those obtained for all four RCPs. Long-term means are very similar for all 30-year windows, differing by at most 2%. Sample set means of long-term standard deviations are somewhat higher for the n=180-year time series, by up to 10% for 2070-2099. Again, we attribute this to the fact that standard deviations of long-term flow means are quite significant for the n=30 sample set, reflecting the fact that the flow time series were computed from sets of climate time series which in many cases exhibit considerable multi-decadal variability, as discussed above. Finally, Table 3 gives sample set means of Kendall tau values, which range from 0.03 to 0.09 for the n=30 sample sets and from 0.05 to 0.12 for the n=180 sample sets. The Kendall tau for a historical time series of annual flows extending from 1896 through 2017 is 0.06 with a p-value of 0.29, indicating no serial correlation.

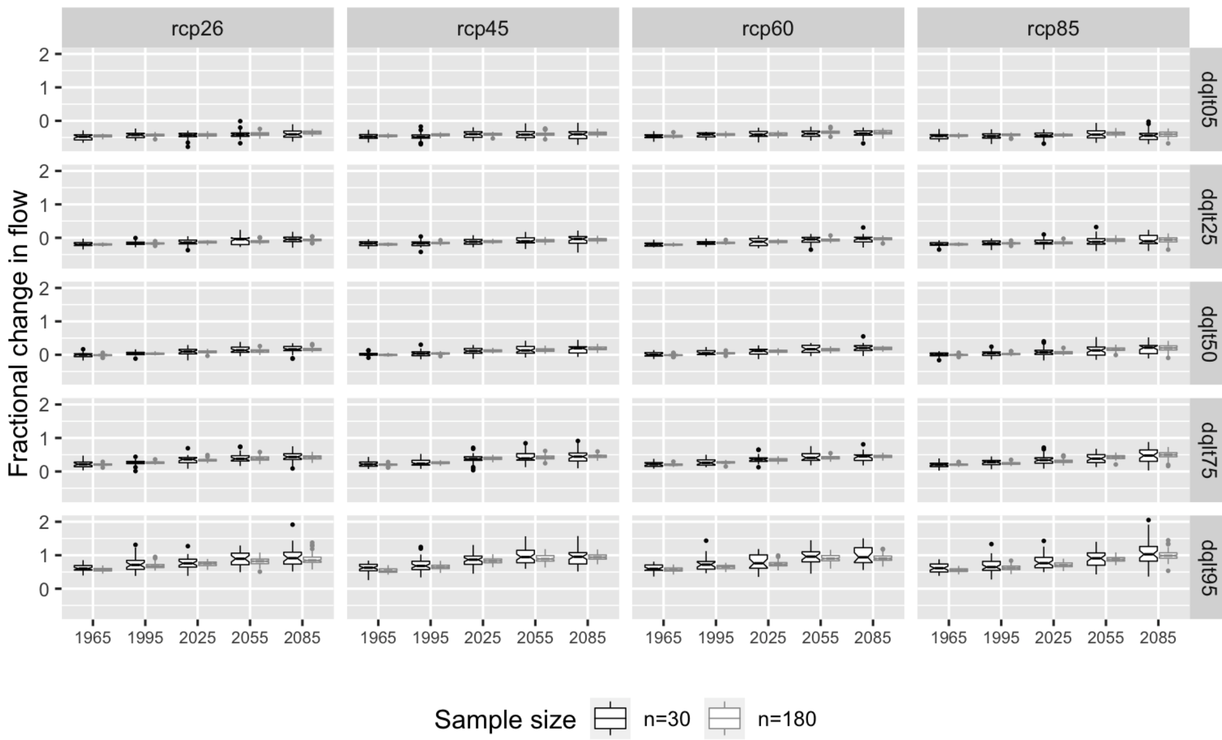

Figure 5: Boxplots of annual fractional flow change quantile values, 0.05, 0.25, 0.50, 0.75, and 0.95, for the sample sets of the ensemble member time series (n=30) and the pooled time series (n=180), by RCP, where the labels on the x-axis denote midpoints of the 30-year windows.

Table 3: Ensemble versus pool set means and standard deviations of ensemble member flow time series statistics (n=30) versus pooled flow time series statistics (n=180). Results shown are for RCP 4.5.

		1950-1979	1980-2009	2010-2039	2040-2069	2070-2099
Mean of long-term flow means (mm)	n=30	349	353	371	377	376
	n=180	343	356	375	378	383
Mean of annual flow standard deviations (mm)	n=30	100	109	119	129	122
	n=180	102	109	127	134	134
Standard deviation of long-term flow means (mm)	n=30	18	22	39	40	48
	n=180	7	8	14	16	19
Mean of Kendall's tau for annual flow time series	n=30	0.03	0.07	0.09	0.05	0.07
	n=180	0.05	0.09	0.12	0.10	0.10

5.4 Future trends in the probability distribution of annual flows

Studies from around the world indicate that a warming climate will lead to increasing severity of both wet weather and dry weather events. The methodology proposed above allows quantitative predictions to be made for future changes in river flows at the annual time scale, throughout the empirical cumulative probability distributions, including the extreme lower and upper tails. The annual time scale is relevant to water supply planning studies, since for many municipal systems droughts that stress water supply resources are those that persist one or more years. The annual time scale is less relevant for flood risk analyses, where increases in daily precipitation variability play a key role.

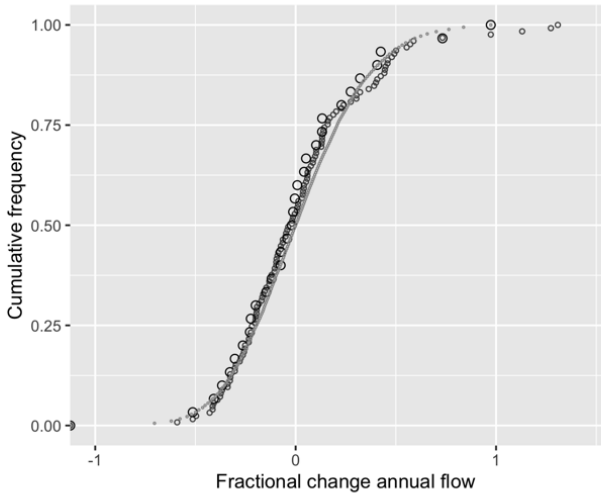

Figure 6: Comparison of observed and simulated annual flow ecdf's for this study's pre-climate change record, 1896 – 1979, and baseline period,1950 – 1979, with simulated flows from the RCP 4.5 sub-ensemble.

To examine trends in annual flow quantiles, we begin with the 100-member sample sets of pooled 180-year annual climate time series, which have been constructed for each of the five successive 30-year time windows covering the period, 1950-2099, by RCP. A flow time series can be computed from each pooled climate time series, providing a sample of annual flows of size n=180 from which quantile statistics can be calculated. Pool sample set means can then computed, providing a multi-model consensus ecdf of flow quantile values.

Figure 6 shows the ecdf of simulated flows, reported as fractional change from the historical mean flow, over the baseline period, 1950-1979, compared with observed values during the baseline period and the longer historic period, 1897-1979. The simulated flows are for the RCP 4.5 multi-model sample set, but results for other RCPs are very similar. This graph indicates that the ecdf of observed annual flows for the baseline period matches the distribution for the longer historic period quite well, supporting use of the 30-year baseline, 1950-1979, as a good approximation to pre-climate change flow conditions. In addition, Kolmogorov-Smirnov tests on the distributions indicate that they do not differ at the 5% significance level.

Figure 7 shows the change over time of ecdf's of annual flows, by RCP, where flows were computed assuming three different sensitivities of flow to temperature change. Results for "medium" temperature sensitivity used annual flows computed using Eq. 7 and the regression coefficients in Table 2, where the temperature coefficient is $ß_2 = -0.054$. Flows for the "low" and "high" temperature sensitivity results were computed in the same way except that the temperature coefficients used are the regression analysis value plus or minus one standard error, that is, $ß_2 = -0.054 \pm 0.025 = -0.029$ and -0.079, respectively.

The graphs in Figure 7 demonstrate the profound uncertainty in future river flows stemming from the physical response of watershed hydrologic processes to rising temperatures and uncertainty in future global carbon emissions. The four graphs in the top row of Figure 7, computed under the assumption that the sensitivity of river flows to a 1 deg C rise in temperature is low, -2.9%, all indicate that future flows will be higher even in extreme drought years and that climate change will not have an adverse impact on water supplies in the Potomac basin. The four graphs in the middle row of Figure 7, computed under the assumption that the response of flows to temperature is medium, -5.4%, indicate that future extreme droughts will likely be more severe than the historical drought of record even though flows will be higher in medium and high flow years. Examining in more detail the data used to create the four middle row graphs, the analysis predicts that annual Potomac River flows during an extreme drought year, that is, a year in which annual flow does not exceed its first quantile value, will change as follows: in the period, 2040-2069, by -14%, -13%, +4%, and -19% for RCPs 2.6, 4.5, 6.0, and 8.5, respectively, and by +1%, -11%, -11%, and -46% by the period, 2070-2099. Finally, the four graphs in the last row of Figure 7, computed under the assumption that the sensitivity of Potomac River flow to temperature is high, -7.9%, indicate that there will be a substantial decrease in Potomac River flows in future extreme drought years.

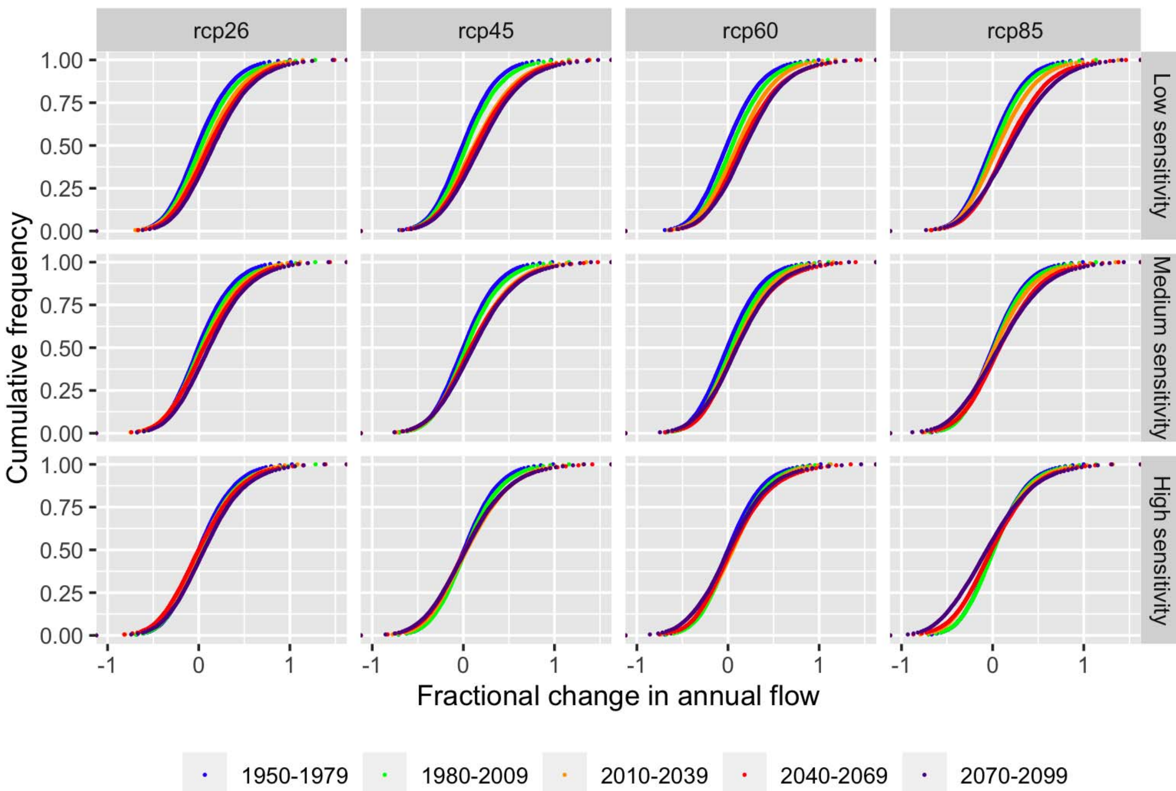

Figure 7: Changes in the ecdf for annual river flow for successive 30-year time windows beginning with the baseline period, 1950-1979, computed from pooled time series by RCP, for three flow-temperature sensitivities: Low (β_2 = -2.9%,), Medium (β_2 = -5.4%,), and High (β_2 = -7.9%,).

5.5 Developing inputs for water supply planning models

Projections of annual flow distributions are valuable for water supply planning studies because they provide quantitative estimates of the disparate impacts of climate change on low flow, medium flow, and high flow years. This is particularly important for regions like the Potomac River basin, where the possibility of future severe drought may be discounted in the face of evidence that the basin is getting wetter. But results from the current study indicate that though future Potomac River flows will increase in average years and in high flow years they may diminish in extreme drought years. Below we demonstrate how the methods described above, which integrate results from multi-model ensembles of climate projections, can be used to construct daily river flow time series. Such flow time series are often key inputs of water supply planning models.

Beginning with ecdf's characterizing annual flow in future time periods, quantile scaling can be applied, similar to its use with rainfall projections (Johnson & Sharma, 2011). This approach captures the changes in interannual variability present in the projections. Following Johnson and Sharma, scaling factors, F^{future}, for a given quantile, q, of annual flow are computed as the ratios of future and baseline quantile values, that is,

$$F_q^{future} = \frac{Q_q^{future}}{Q_q^{baseline}}$$

Equation 9

These annual flow scaling factors can then be used to project the impact of climate change on the historic record by multiplying by observed daily flows,

$$\left(Q_i^{future}\right)_{QS} = Q_{i\,q}^{obs} F_q^{future}$$

Equation 10

where $Q_{i\,q}^{obs}$ is the observed daily flow on Julian day, i, with q denoting its quantile in the historic record of annual flows. A similar method of nonparametric scaling was used by Nowak et al. (Nowak et al., 2010) in their method of multisite disaggregation of annual to daily streamflows. This approach is particularly well-suited to the need of Potomac basin water supply planning models because it preserves the relationship between daily flows at upstream and downstream locations, enabling use of flow forecasting techniques in simulations of the performance of the WMA water supply system under future climate change.

In Table 4, annual flow scaling factors for selected quantiles are shown for two time periods of interest to planners in the Potomac basin, 2040-2069 and 2070-2099, where change is computed from the "pre-climate change" baseline period, 1950-1979. Such scaling factors can be computed from ecdf's for any multi-model ensemble, or even from individual pooled time series, but the results shown below are from the RCP 4.5 subset of the BCSD ensemble. The uncertainty in the response of flow to change in temperature is significant, as illustrated in Figure 7, and the scaling factors in Table 4 were computed for three different temperature sensitivity scenarios determined by a range of plus or minus one standard error around $ß_2$, the temperature coefficient in the CRF.

Table 4: Projected scaling factors for flow from baseline period, 1950-1979 (pool sample set means of annual flow quantiles, RCP 4.5 only).

	Sensitivity of annual flow to annual temperature					
Annual flow quantile	Low (β_2 = -0.029)		Medium (β_2 = -0.054)		High (β_2 = -0.079)	
	2040-2069	2070-2099	2040-2069	2070-2099	2040-2069	2070-2099
0.01	1.07	1.12	0.87	0.89	0.66	0.64
0.05	1.10	1.13	0.97	0.97	0.84	0.81
0.10	1.11	1.15	1.00	1.01	0.90	0.87
0.25	1.13	1.17	1.05	1.06	0.96	0.96
0.50	1.15	1.20	1.08	1.10	1.01	1.02
0.75	1.18	1.21	1.13	1.14	1.07	1.07
0.90	1.21	1.24	1.16	1.17	1.11	1.11
0.95	1.24	1.26	1.19	1.20	1.14	1.14
0.99	1.30	1.29	1.26	1.24	1.22	1.19

Of particular interest to planners in the Potomac basin are the factors in Table 4 for the 0.01 quantile, representing projected change in annual flow in an extreme drought year. These are 0.87 and 0.89, respectively, for the medium temperature sensitivity scenarios. These results indicate that if extreme drought strikes the region in the period, 2040-2069, annual river flow will be 13 percent less than flows in a corresponding historical drought, and 11 percent less in the period, 2070-2099, based on the projections of the RCP 4.5 sub-ensemble. In contrast, annual flows in an extremely high flow year (0.99 quantile) are projected to rise by 26% and 24% in these same two future periods.

6 Summary and Conclusions

In the Potomac River basin, the major water supply source of the Washington, DC, metropolitan area, there is a growing expectation that climate change will bring increased precipitation. But because variability is also expected to increase, it's important that water supply planners investigate the risk that future extreme drought may be more severe than droughts in the past due to increased precipitation variability coupled with rising temperatures, termed by some as hot drought. Common assessment approaches that can capture this risk include detailed hydrologic modeling based on climate inputs from a subset of available GCMs projections (Arnell and Delaney, 2006; Manning et al., 2009; Matonse et al., 2013; O'Hara and Georgakakos, 2008; Paton et al., 2014) and vulnerability assessments exploring a wide range of possible future climates to aid planners in understanding what mitigation options need to be "on the table" (Brown et al., 2011; Steinschneider et al., 2015). But studies based on a limited number of climate projections or worst-case scenarios from vulnerability assessments may not be enough to compel decision makers to move forward in cases where the need for costly infrastructure projects is indicated.

Trends in climate and hydrology are typically investigated by comparing conditions in successive time windows, each several decades in length. Our proposed approach provides a method of constructing annual climate and flow time series representative of a given time window that are of sufficient length to compute statistics indicative of the severity of extreme drought, for example, a drought severity with a probability of exceedance of just 1%. Applying this approach to the Potomac River, pooled time series, each 180 years in length, were created to characterize climatic conditions in 30-year time windows, and these were then used as input into the CRF to compute corresponding annual flow time series. Each pair of pooled temperature-precipitation time series incorporates information from multiple climate models and could be used individually to investigate the range of conditions expected in a future climate. In this study, we added an averaging step: we created a large sample set of pools and took means of quantile and other statistics computed from individual pools. Comparisons of quantile values in the range, 0.05 to 0.95, computed from the pools and from the original ensemble member time series indicate that means of quantile values computed from the sample set of pools provide reasonably good estimates of those computed from shorter time series, lending confidence to this study's results for more extreme quantiles. We did note that standard deviations were considerably higher for the pooled temperature time series and to a lesser degree for the pooled flow time series, and also Kendall's tau values for temperature, and attributed these differences to the multi-decadel variability introduced by concatenating multiple 30-year temperature time series from multiple GCM's.

Our results for the Potomac basin indicate that future declines in river flow in extreme drought years may become more severe even as long-term mean flows increase, reflective of processes indicative of hot drought. Twelve future scenarios were considered to investigate the competing influences of rising precipitation and rising temperatures, based on four scenarios for future global emissions and three scenarios for the sensitivity of flow to temperature change. For all of the high and three out of four of the medium flow-temperature sensitivity scenarios, the 0.01 quantile value of annual flow decreases in both the 2040-2069 and the 2070-2099 planning periods. Our results for future trends in flow quantiles are consistent with those obtained by

Hayhoe et al. (2007) for the US northeast, showing that climate change will increase annual flows in the upper tail of the cumulative distribution function and decrease annual flows in the lower tail. But our proposed methodology provides quantitative information on changes in extreme drought, as represented by the 0.01 quantile, whereas standard empirical nonparametric approaches, because of sample size, are limited to changes in a more moderate drought.

The proposed approach can be useful for water supply planning studies, especially in the case of systems with significant reservoir storage where the annual time scale may be appropriate for investigation of future changes in water availability. It provides a method for constructing long time series of annual climate, incorporating projections from multiple GCMs, that are representative of shorter time intervals of interest. Applicability to a given watershed depends on whether a multiple regression analysis indicates a significant relationship between annual flow and annual temperature and precipitation, allowing the development of a CRF. If so, ecdf's for annual flow can be constructed, as described above, and scaling factors for annual flow, as a function of annual flow quantile, can easily be computed. These factors can be applied, via quantile scaling, to historic monthly or daily time series to create inputs to water supply planning models that reflect projected impacts of climate change. Alternatively, annual flow scaling factors could be applied to flow time series that have been synthetically generated based on historic data. In the current study, risk is explored via a scenario approach, by considering four representative pathways for future greenhouse gas concentrations and three scenarios for the sensitivity of flow to temperature change. For practical applications, planners could generate a single scenario for use in their planning models by first determining the level of risk appropriate for their study, that is, an RCP and an assumption of the sensitivity of flow to temperature change, or alternatively, generate results for a set of scenarios of their choice.

Acknowledgments

Support for this work was from the Section for Cooperative Water Supply Operations on the Potomac of the Interstate Commission on the Potomac River Basin, with funding provided by the CO-OP water suppliers: Fairfax Water, WSSC Water, and the Washington Aqueduct Division of the U.S. Army Corps of Engineers.

Open Research

Data Availability Statement

Historical climate datasets used in this study are: 4 x 4 km PRISM gridded monthly air temperature and precipitation data from the PRISM Climate Group at Oregon State University, available at https://prism.oregonstate.edu), and 5km GHCN-Daily gridded Temperature and Precipitation Dataset (nClimGrid/CLIMGRID), version 1, from NOAA's National Centers for Environmental Information (NCEI) available at https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00332). The climate projections used in this study, downscaled to the Potomac River basin, are derived from the Coupled Multi-model Intercomparison Project, Phase 5 (CMIP5), bias-correction and spatial disaggregation (BCSD) dataset available at https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html). The annual time series of historical natural Potomac River flow at Little Falls are available upon request from the corresponding author. Figures and tables were created with R version 4.3.0 (R Core Team, 2023) and RStudio version 2023.03.0 (Posit Team, 2023). Code and data for reproduction of results is available on Github at https://github.com/icprbcoop/cc_br.

References

Ahmed, S. N., Moltz, H. L. N., Schultz, C. L., & Seck, A. (2020). 2020 Washington metropolitan area water supply reliability study—Demand and resource availability forecast for the year 2050, ICPRB Report No. 20-3. ICPRB. https://www.potomacriver.org/wp-content/uploads/2020/12/2020-WMA-Water-Supply-study-FINAL-September-2020.pdf

Arnell, N. W., & Delaney, E. K. (2006). Adapting to climate change: Public water supply in England and Wales. Climatic Change, 78(2–4), 227–255. https://doi.org/10.1007/s10584-006-9067-9

Arnell, N. W., & Gosling, S. N. (2013). The impacts of climate change on river flow regimes at the global scale. Journal of Hydrology, 486, 351–364. https://doi.org/10.1016/j.jhydrol.2013.02.010

Baker, S. A., Rajagopalan, B., & Wood, A. W. (2021). Enhancing Ensemble Seasonal Streamflow Forecasts in the Upper Colorado River Basin Using Multi-Model Climate Forecasts. JAWRA Journal of the American Water Resources Association, 57(6), 906–922. https://doi.org/10.1111/1752-1688.12960

Block, P. J., Souza Filho, F. A., Sun, L., & Kwon, H. (2009). A streamflow forecasting framework using multiple climate and hydrological models 1. JAWRA Journal of the American Water Resources Association, 45(4), 828–843. https://doi.org/10.1111/j.1752-1688.2009.00327.x

Brown, C., Ghile, Y., Laverty, M., & Li, K. (2012). Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resources Research, 48(9). https://doi.org/10.1029/2011WR011212

Brown, C., Werick, W., Leger, W., & Fay, D. (2011). A Decision-Analytic approach to managing climate risks: Application to the upper great lakes 1. JAWRA Journal of the American Water Resources Association, 47(3), 524–534. https://doi.org/10.1111/j.1752-1688.2011.00552.x

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964

Cummins, J., Buchanan, C., Haywood, C., Moltz, H., Griggs, A., Jones, R. C., Kraus, R., Hitt, N., Bumgardner, R. V., & Branch, A. E. (2010). Potomac basin large river environmental flow needs. Interstate Commission on the Potomac River Basin. Publication Number ICPRB, 10–03. https://www.potomacriver.org/wp-content/uploads/2015/02/ICP10-3_Cummins..pdf

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology: A Journal of the Royal Meteorological Society, 28(15), 2031–2064. https://doi.org/10.1002/joc.1688

Döll, P., & Schmied, H. M. (2012). How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis. Environmental Research Letters, 7(1), 014037. https://doi.org/10.1088/1748-9326/7/1/014037

Ehsani, N., Vörösmarty, C. J., Fekete, B. M., & Stakhiv, E. Z. (2017). Reservoir operations under climate change: Storage capacity options to mitigate risk. Journal of Hydrology, 555, 435–446. https://doi.org/10.1016/j.jhydrol.2017.09.008

Fix, E., & Hodges, J. L. (1989). Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties. International Statistical Review / Revue Internationale de Statistique, 57(3), 238–247. https://doi.org/10.2307/1403797

Fleming, B. j., Archfield, S. a., Hirsch, R. m., Kiang, J. e., & Wolock, D. m. (2021). Spatial and Temporal Patterns of Low Streamflow and Precipitation Changes in the Chesapeake Bay Watershed. JAWRA Journal of the American Water Resources Association, 57(1), 96–108. https://doi.org/10.1111/1752-1688.12892

Fowler, H. J., Kilsby, C. G., & O'Connell, P. E. (2003). Modeling the impacts of climatic change and variability on the reliability, resilience, and vulnerability of a water resource system. Water Resources Research, 39(8). https://doi.org/10.1029/2002WR001778

Fritsch, J. M., Hilliker, J., Ross, J., & Vislocky, R. L. (2000). Model consensus. Weather and Forecasting, 15(5), 571–582. https://doi.org/10.1175/1520-0434(2000)015%3C0571:MC%3E2.0.CO;2

Fu, G., Charles, S. P., & Chiew, F. H. (2007). A two-parameter climate elasticity of streamflow index to assess climate change effects on annual streamflow. *Water Resources Research, 43*(11). https://doi.org/10.1029/2007WR005890

Georgakakos, K. P., Seo, D.-J., Gupta, H., Schaake, J., & Butts, M. B. (2004). Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. Journal of Hydrology, 298(1–4), 222–241. https://doi.org/10.1016/j.jhydrol.2004.03.037

Gleick, P. H. (1986). Methods for evaluating the regional hydrologic impacts of global climatic changes. Journal of Hydrology, 88(1–2), 97–116. https://doi.org/10.1016/0022-1694(86)90199-X

Groves, D. G., Yates, D., & Tebaldi, C. (2008). Developing and applying uncertain global climate change projections for regional water management planning. Water Resources Research, 44(12). https://doi.org/10.1029/2008WR006964

Hagedorn, R., Doblas-Reyes, F. J., & Palmer, T. N. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. Tellus A: Dynamic Meteorology and Oceanography, 57(3), 219–233. https://doi.org/10.1029/2008WR006964

Hagen, E. R., Holmes, K. J., Kiang, J. E., & Steiner, R. C. (2005). BENEFITS OF ITERATIVE WATER SUPPLY FORECASTING IN THE WASHINGTON, DC, METROPOLITAN AREA 1. JAWRA Journal of the American Water Resources Association, 41(6), 1417–1430. https://doi.org/10.1111/j.1752-1688.2005.tb03809.x

Hayhoe, K., Wake, C. P., Huntington, T. G., Luo, L., Schwartz, M. D., Sheffield, J., Wood, E., Anderson, B., Bradbury, J., DeGaetano, A., Troy, T. J., & Wolfe, D. (2007). Past and future changes in climate and hydrological indicators in the US Northeast. Climate Dynamics, 28(4), 381–407. https://doi.org/10.1007/s00382-006-0187-8

Hinson, K. E., Friedrichs, M. A. M., Najjar, R. G., Herrmann, M., Bian, Z., Bhatt, G., St-Laurent, P., Tian, H., & Shenk, G. (2022). Impacts and uncertainties of climate-induced changes in watershed inputs on estuarine hypoxia. EGUsphere, 1–46. https://doi.org/10.5194/egusphere-2022-1028

Hirabayashi, Y., Kanae, S., Emori, S., Oki, T., & Kimoto, M. (2008). Global projections of changing risks of floods and droughts in a changing climate. Hydrological Sciences Journal, 53(4), 754–772. https://doi.org/10.1623/hysj.53.4.754

Johnson, F., & Sharma, A. (2011). Accounting for interannual variability: A comparison of options for water resources climate change impact assessments. Water Resources Research, 47(4). https://doi.org/10.1029/2010WR009272

Karl, T. R., & Riebsame, W. E. (1989). The impact of decadal fluctuations in mean precipitation and temperature on runoff: A sensitivity study over the United States. Climatic Change, 15(3), 423–447. https://doi.org/10.1007/BF00240466

Karlsson, M., & Yakowitz, S. (1987). Nearest-neighbor methods for nonparametric rainfall-runoff forecasting. Water Resources Research, 23(7), 1300–1308. https://doi.org/10.1029/WR023i007p01300

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. Advances in Water Resources, 25(8–12), 1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8

Kay, A. L., Griffin, A., Rudd, A. C., Chapman, R. M., Bell, V. A., & Arnell, N. W. (2021). Climate change effects on indicators of high and low river flow across Great Britain. Advances in Water Resources, 151, 103909. https://doi.org/10.1016/j.advwatres.2021.103909

Krishnamurti, T., Kishtawal, C. M., LaRow, T. E., Bachiochi, D. R., Zhang, Z., Williford, C. E., Gadgil, S., & Surendran, S. (1999). Improved weather and seasonal climate forecasts from multimodel superensemble. Science, 285(5433), 1548–1550. https://doi.org/10.1126/science.285.5433.1548

Lall, U., & Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic time series. Water Resources Research, 32(3), 679–693. https://doi.org/10.1029/95WR02966

Leander, R., & Buishand, T. A. (2007). Resampling of regional climate model output for the simulation of extreme river flows. Journal of Hydrology, 332(3–4), 487–496. https://doi.org/10.1016/j.jhydrol.2006.08.006

Manning, L., Hall, J., Fowler, H., Kilsby, C., & Tebaldi, C. (2009). Using probabilistic climate change information from a multimodel ensemble for water resources assessment. Water Resources Research, 45(11). https://doi.org/10.1029/2007WR006674

Matonse, A. H., Pierson, D. C., Frei, A., Zion, M. S., Anandhi, A., Schneiderman, E., & Wright, B. (2013). Investigating the impact of climate change on New York City's primary water supply. Climatic Change, 116, 437–456. https://doi.org/10.1007/s10584-012-0515-4

Maxwell, R. S., Hessl, A. E., Cook, E. R., & Pederson, N. (2011). A multispecies tree ring reconstruction of Potomac River streamflow (950–2001). Water Resources Research, 47(5). https://doi.org/10.1029/2010WR010019

McCabe, G. J., & Wolock, D. M. (2011). Independent effects of temperature and precipitation on modeled runoff in the conterminous United States. Water Resources Research, 47(11). https://doi.org/10.1029/2011WR010630

McCabe, G. J., Wolock, D. M., Pederson, G. T., Woodhouse, C. A., & McAfee, S. (2017). Evidence that Recent Warming is Reducing Upper Colorado River Flows. Earth Interactions, 21(10), 1–14. https://doi.org/10.1175/EI-D-17-0007.1

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? Science, 319(5863), 573–574. https://doi.org/10.1126/science.1151915

Milly, P. C. D., Dunne, K. A., & Vecchia, A. V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. Nature, 438(7066), Article 7066. https://doi.org/10.1038/nature04312

Milly, P. C., & Dunne, K. A. (2020). Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation. Science, 367(6483), 1252–1255. https://doi.org/10.1126/science.aay9187

Milly, P., & Dunne, K. (2002). Macroscale water fluxes 2. Water and energy supply control of their interannual variability. Water Resources Research, 38(10), 24–1. https://doi.org/10.1029/2001WR000760

Milly, P., Kam, J., & Dunne, K. A. (2018). On the sensitivity of annual streamflow to air temperature. Water Resources Research, 54(4), 2624–2641. https://doi.org/10.1002/2017WR021970

Moltz, H. L., Wallace, C. W., Sharifi, E., & Bencala, K. (2020). Integrating sustainable water resource management and land use decision-making. Water, 12(8), 2282. https://doi.org/10.3390/w12082282

Najjar, R., Patterson, L., & Graham, S. (2009). Climate simulations of major estuarine watersheds in the Mid-Atlantic region of the US. Climatic Change, 95(1–2), 139–168. https://doi.org/10.1007/s10584-008-9521-y

Nash, L. L., & Gleick, P. H. (1991). Sensitivity of streamflow in the Colorado Basin to climatic changes. Journal of Hydrology, 125(3), 221–241. https://doi.org/10.1016/0022-1694(91)90030-L

Nowak, K., Hoerling, M., Rajagopalan, B., & Zagona, E. (2012). Colorado River basin hydroclimatic variability. *Journal of Climate, 25*(12), 4389-4403. https://doi.org/10.1175/JCLI-D-11-00406.1

Nowak, K., Prairie, J., Rajagopalan, B., & Lall, U. (2010). A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow. Water Resources Research, 46(8). https://doi.org/10.1029/2009WR008530

O'Hara, J. K., & Georgakakos, K. P. (2008). Quantifying the urban water supply impacts of climate change. Water Resources Management, 22, 1477–1497. https://doi.org/10.1007/s11269-008-9238-8

Paton, F. L., Dandy, G. C., & Maier, H. R. (2014). Integrated framework for assessing urban water supply security of systems with non-traditional sources under climate change. Environmental Modelling & Software, 60, 302–319. https://doi.org/10.1016/j.envsoft.2014.06.018

Posit team (2023). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. http://www.posit.co/

Prairie, J. R., Rajagopalan, B., Fulp, T. J., & Zagona, E. A. (2006). Modified K-NN model for stochastic streamflow simulation. Journal of Hydrologic Engineering, 11(4), 371–378. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:4(371)

Pyke, C. R., & Najjar, R. (n.d.). 2008. Climate Change and the Chesapeake Bay: State-of-the-Science Review and Recommendations. A Report from the Chesapeake Bay Program Science and Technical Advisory Committee (STAC), Annapolis, MD. 59 pp. https://www.chesapeake.org/stac/Pubs/climchangereport.pdf

R Core Team (2023). R: A Language and Environment for Statistical Computin. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rajagopalan, B., & Lall, U. (1999). A k-nearest-neighbor simulator for daily precipitation and other weather variables. Water Resources Research, 35(10), 3089–3101. https://doi.org/10.1029/1999WR900028

Rashid, M. M., Sharma, A., & Johnson, F. (2020). Multi-model drought predictions using temporally aggregated climate indicators. Journal of Hydrology, 581, 124419. https://doi.org/10.1016/j.jhydrol.2019.124419

Reclamation, S. (2013). Downscaled CMIP3 and CMIP5 climate and hydrology projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs.

Regonda, S. K., Rajagopalan, B., Clark, M., & Zagona, E. (2006). A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin. Water Resources Research, 42(9). https://doi.org/10.1029/2005WR004653

Revelle, R. R., & Waggoner, P. E. (1983). Effects of a carbon dioxide-induced climatic change on water supplies in the western United States. In *National Research Council. Changing climate: Report of the carbon dioxide assessment committee* (pp. 419-432). Washington, DC: The National Academies Press. https://doi.org/10.17226/18714

Rice, K. C., & Hirsch, R. M. (2012). Spatial and temporal trends in runoff at long-term streamgages within and near the Chesapeake Bay Watershed. Scientific Investigations Report No. 2012–5151; p. 56. U.S. Geological Survey. http://pubs.usgs.gov/sir/2012/5151

Risbey, J. S., & Entekhabi, D. (1996). Observed Sacramento Basin streamflow response to precipitation and temperature changes and its relevance to climate impact studies. Journal of Hydrology, 184(3–4), 209–223. https://doi.org/10.1016/0022-1694(95)02984-2

Rootzén, H., & Katz, R. W. (2013). Design life level: Quantifying risk in a changing climate. Water Resources Research, 49(9), 5964–5972. https://doi.org/10.1002/wrcr.20425

Sankarasubramanian, A., & Vogel, R. M. (2003). Hydroclimatology of the continental United States. Geophysical Research Letters, 30(7). https://doi.org/10.1029/2002GL015937

Sankarasubramanian, A., & Vogel, R. M. (2002). Annual hydroclimatology of the United States. *Water Resources Research, 38*(6), 19-1. https://doi.org/10.1029/2001WR000619

Sankarasubramanian, A., Vogel, R. M., & Limbrunner, J. F. (2001). Climate elasticity of streamflow in the United States. *Water Resources Research, 37*(6), 1771-1781. https://doi.org/10.1029/2000WR900330

Schaake, J. C. (1990). From climate to flow. Climate Change and US Water Resources., 177–206.

Schaake, J. C., & Liu, C. Z. (1989). Development and application of simple water balance models to understand the relationship between climate and water resources. In *New Directions for Surface Water Modeling. Proceedings of a Symposium held in Baltimore, Maryland, May 1989. IAHS Publication* (No. 181).

Schultz, C., Ahmed, S., Mandel, R., & Moltz, H. (2014). Improvement in HSPF's Low-Flow Predictions by Implementation of a Power Law Groundwater Storage-Discharge Relationship. JAWRA Journal of the American Water Resources Association, 50(4), 909–927. https://doi.org/10.1111/jawr.12144

Serinaldi, F., & Kilsby, C. G. (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. Advances in Water Resources, 77, 17–36. https://doi.org/10.1016/j.advwatres.2014.12.013

Sharif, M., & Burn, D. H. (2006). Simulating climate change scenarios using an improved K-nearest neighbor model. Journal of Hydrology, 325(1–4), 179–196. https://doi.org/10.1016/j.jhydrol.2005.10.015

Sheer, D.P., & Flynn, K. (1983). Water-Supply. Civil Engineering, 53(6), 50–53.

Shenk, G. W., Bhatt, G., Tian, R., Cerco, C. F., Bertani, I., & Linker, L. C. (2021). Modeling Climate Change Effects on Chesapeake Water Quality Standards and Development of 2025 Planning Targets to Address Climate Change (CBPO Publication No. 328–21; p. 145). Chesapeake Bay Program.

Steinschneider, S., McCrary, R., Wi, S., Mulligan, K., Mearns, L. O., & Brown, C. (2015). Expanded decision-scaling framework to select robust long-term water-system plans under hydroclimatic uncertainties. Journal of Water Resources Planning and Management, 141(11), 04015023. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000536

Tang, Q., & Lettenmaier, D. P. (2012). 21st century runoff sensitivities of major global river basins. Geophysical Research Letters, 39(6). https://doi.org/10.1029/2011GL050834

Tang, Y., Tang, Q., & Zhang, L. (2020). Derivation of interannual climate elasticity of streamflow. *Water Resources Research, 56*(11), e2020WR027703. https://doi.org/10.1029/2020WR027703

Tebaldi, C., Hayhoe, K., Arblaster, J. M., & Meehl, G. A. (2006). Going to the extremes. Climatic Change, 79(3), 185–211. https://doi.org/10.1007/s10584-006-9051-4

Torbenson, M. C. A., & Stagge, J. H. (2021). Informing Seasonal Proxy-Based Flow Reconstructions Using Baseflow Separation: An Example From the Potomac River, United States. Water Resources Research, 57(2), e2020WR027706. https://doi.org/10.1029/2020WR027706

Toth, Z., & Kalnay, E. (1993). Ensemble forecasting at NMC: The generation of perturbations. Bulletin of the American Meteorological Society, 74(12), 2317–2330. https://doi.org/10.1175/1520-0477(1993)074%3C2317:EFANTG%3E2.0.CO;2

Trainer, F. W., & Watkins, F. A. (1975). Geohydrologic reconnaissance of the upper Potomac River basin. US Govt. Print. Off. https://doi.org/10.3133/wsp2035

Udall, B., & Overpeck, J. (2017). The twenty-first century Colorado River hot drought and implications for the future. Water Resources Research, 53(3), 2404–2418. https://doi.org/10.1002/2016WR019638

Vano, J. A., Das, T., & Lettenmaier, D. P. (2012). Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature. Journal of Hydrometeorology, 13(3), 932–949. https://doi.org/10.1175/JHM-D-11-069.1

Vogel, R. M., & Fennessey, N. M. (1995). Flow Duration Curves Ii: A Review of Applications in Water Resources Planning1. JAWRA Journal of the American Water Resources Association, 31(6), 1029–1039. https://doi.org/10.1111/j.1752-1688.1995.tb03419.x

Vose, R. S., Applequist, S., Squires, M., Durre, I., Menne, M. J., Williams, C. N., Fenimore, C., Gleason, K., & Arndt, D. (2014). Improved historical temperature and precipitation time series for US climate divisions. Journal of Applied Meteorology and Climatology, 53(5), 1232–1251. https://doi.org/10.1175/JAMC-D-13-0248.1

Watts, G., Christierson, B. von, Hannaford, J., & Lonsdale, K. (2012). Testing the resilience of water supply systems to long droughts. Journal of Hydrology, 414–415, 255–267. https://doi.org/10.1016/j.jhydrol.2011.10.038

Weinberger, K. Q., & Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research, 10(2).

Woodhouse, C. A., Pederson, G. T., Morino, K., McAfee, S. A., & McCabe, G. J. (2016). Increasing influence of air temperature on upper Colorado River streamflow. *Geophysical Research Letters, 43*(5), 2174-2181. https://doi.org/10.1002/2015GL067613

Xue, Z., & Ullrich, P. A. (2022). Changing Trends in Drought Patterns over the Northeastern United States Using Multiple Large Ensemble Datasets. Journal of Climate, 35(22), 3813–3833. https://doi.org/10.1175/JCLI-D-21-0810.1

Yang, Q., Tian, H., Friedrichs, M. A. M., Liu, M., Li, X., & Yang, J. (2015). Hydrological Responses to Climate and Land-Use Changes along the North American East Coast: A 110-Year Historical Reconstruction. JAWRA Journal of the American Water Resources Association, 51(1), 47–67. https://doi.org/10.1111/jawr.12232

Yates, D., Gangopadhyay, S., Rajagopalan, B., & Strzepek, K. (2003). A technique for generating regional climate scenarios using a nearest-neighbor algorithm. Water Resources Research, 39(7). https://doi.org/10.1029/2002WR001769

Zeff, H. B., Herman, J. D., Reed, P. M., & Characklis, G. W. (2016). Cooperative drought adaptation: Integrating infrastructure development, conservation, and water transfers into adaptive policy pathways. Water Resources Research, 52(9), 7327–7346. https://doi.org/10.1002/2016WR018771