Incorporating Climate Non-stationarity and Snowmelt Processes in Intensity-Duration-Frequency Analyses with Case Studies in Mountainous Areas

Zhangshuan Hou[a*], Huiying Ren[a], Ning Sun[a], Mark S. Wigmosta[a], Ying Liu[b], L. Ruby Leung[b], Hongxiang Yan[a], Richard Skaggs[a], and Andre Coleman[a]

[a]Hydrology, Pacific Northwest National Laboratory, Richland, WA USA

[b]Earth Systems Analysis and Model, Pacific Northwest National Laboratory, Richland, WA USA

To be submitted to *Journal of Hydrometeorology*

*corresponding author email: zhangshuan.hou@pnnl.gov

ABSTRACT

Downscaled high resolution climate simulations were used to provide inputs to the physics-based Distributed Hydrology Soil Vegetation Model (DHSVM), which accounts for the combined effects of snowmelt and rainfall processes, to determine spatially distributed available water for runoff (AWR). After quasi-stationary time windows were identified based on model outputs extracted for two different mountainous field sites in Colorado and California, intensity-duration-frequency (IDF) curves for precipitation and AWR were generated and evaluated at each numerical grid to provide guidance on hydrological infrastructure design.

Impacts of snowmelt are found to be spatially variable due to spatial heterogeneity associated with topography according to geostatistical analyses. AWR extremes have stronger spatial continuity compared to precipitation. Snowmelt impacts on AWR is more pronounced at the wet California site than at the semi-arid Colorado site. The sensitivities of AWR and precipitation IDFs to increasing greenhouse gas emissions are found to be localized and spatially variable. In sub-regions with significant snowfall, snowmelt can result in an AWR (e.g., 6-hour 100-year events) that is 70% higher than precipitation. For comparison, future greenhouse gas emissions may increase 6-hour 100-year precipitation and AWR by up to 50% and 80%, respectively, towards the end of this century.

Keywords: Intensity-duration-frequency: non-stationarity; snowmelt; available water for runoff; geostatistics; spatial heterogeneity; RESM; DHSVM

1. Introduction

Standards for designing civil engineering infrastructure (e.g., stormwater management facilities, erosion and sediment control structures, flood protection structures) usually involves statistical analysis of historic precipitation events, particularly in terms of intensity, duration, and frequency (IDF) [Chow et al., 1988; McCuen, 1998]. The traditional design paradigm makes several significant assumptions such as climate stationarity and neglects snowmelt-driven runoff, and implicitly assumes that the reoccurrence interval of a precipitation event produces a runoff event of the same interval. However, given the potential impacts of greenhouse gas emissions, these limiting assumptions can substantially increase infrastructure development risks, because structures designed to meet traditional criteria (e.g., 100-year 6-hour storm) could be over- or under-designed leading to issues of safety and/or unnecessary expenses. Meanwhile, changes in extreme precipitation may vary significantly from site to site, presenting further challenges for managing flood risk in the future [Chow et al., 1988; Guo, 2006; Kao and Ganguly, 2011; Kharin et al., 2007; Koutsoyiannis and Baloutsos, 2000; Mirhosseini et al., 2013; Peck et al., 2012; Ragno et al., 2017; Teegavarapu, 2013].

In previous studies, some attention has been given to non-stationarity in extreme precipitation characteristics [Haddad and Moravej, 2015; Mailhot et al., 2007; Patel et al., 2015], because when the series are non-stationary (e.g., with trending or changing auto-covariance), the stationary assumption delivers IDF curves that can substantially underestimate extreme events [Cheng et al., 2014]. Some work has been done to evaluate the location and scale parameters in the extreme value distributions but assuming the shape parameter is constant [Bracken et al., 2016; Cooley et al., 2007; Lima and Lall, 2010; Yan and Moradkhani, 2015]. The Mann-Kendall and linear regression trend test, von Neumann independence test, Wald-Wolfowitz stationarity test, and Mann-Whitney homogeneity test have been applied to the precipitation annual maximum series (AMS) of standard durations to inspect the presence of monotonic trends and evaluate the independency, stationarity, and homogeneity of AMS of standard durations[Haddad and Moravej, 2015].

In addition to non-stationarity, spatial heterogeneity of extreme precipitation distribution is also a critical issue [*Ghosh et al.*, 2012]; *Mailhot et al.* [2007] showed that for a given duration, spatial correlations of precipitation extremes will decrease in a future climate, suggesting that annual extreme precipitation events may result more often from convective (and thus more localized) rather than synoptic-scale weather systems. Quantifying the spatial heterogeneity in historical precipitation extremes is difficult because weather stations are sparsely distributed, especially in mountainous regions[*Bales et al.*, 2006; *López-Moreno et al.*, 2009]. This issue becomes more serious when only stations with an adequately long period of record are considered. This issue is similarly critical regarding climate projections, given the spatial resolution of commonly available global climate projections (e.g., from CMIP5 [*Taylor et al.*, 2012]) (~100 – 200 km) and dynamically or statistically downscaled scenarios (~10 – 50 km) relative to the much smaller spatial scales of extreme precipitation events. A reliable spatial statistics based scheme, such as a geostatistical approach [*Deutsch and Journel*, 1998] involving spatial random functions and variogram models, is helpful to model the spatial correlation patterns of extreme precipitation events.

Yet another important issue for reliable IDF analysis is the impact of snowmelt processes on runoff generation, particularly in snow-dominated regions, where much of the precipitation is stored as snowpack till springtime when it melts and produces runoff. Therefore, available water for runoff (AWR) in snow-dominated environments relates more directly to the magnitude and timing of runoff from melting snow and rain than that of precipitation. Although large flood events are often caused by extreme rain events over short time periods, a large number of significant flood events in snow-dominated regions are attributable to snowmelt from deep snowpack especially during rain-on-snow (ROS) events [Bookhagen and Burbank, 2010; Fang et al., 2014; Kampf and Richer, 2014; Kattelmann, 1997]. With local measurements of meteorology and Snow Water Equivalent (SWE) at 376 Snowpack Telemetry (SNOTEL) stations across the Western United States, researchers estimated the extreme runoff events at each SNOTEL station [*Yan et al.*, 2018; *Yan et al*., 2019a; *Yan et al.*, 2019b]. Compared to AWR-based IDFs, traditional precipitation-based IDFs led to under-design at 45% of the SNOTEL stations, many with significant underestimation of 100-year extreme events. Improvements on the traditional IDF using snowmelt-incorporated AWR are readily achievable with land surface hydrologic models such as the Distributed Hydrology Soil Vegetation Model (DHSVM) [Wigmosta et al., 1994] that represent the integral snow processes.

In this study, we address the aforementioned issues in traditional IDF development with high-resolution climate projections [Hurrell et al., 2013; Kay et al., 2015; L. R. Leung and Ghan, 1999; L. Ruby Leung et al., 2004; Voisin et al., 2013a; Voisin et al., 2013b; Wang et al., 2004] coupled with the DHSVM model [Wigmosta et al., 1994]. DHSVM has been extensively applied in simulating snow and hydrological processes in mountainous snow environments due to its detailed and spatially explicit representation of the physical processes that govern the energy and mass exchange between the atmosphere, (overstory) canopy, snowpack, and ground surface. Climate projections from a high-resolution regional earth system model are available for the period 1975-2100 over the western United States under two future greenhouse gas emission scenarios. The bias-corrected climate model outputs (e.g., temperature, wind speed, relative humidity, shortwave and longwave radiation, precipitation) were used to provide atmospheric forcing to DHSVM to predict spatially distributed AWR at two mountainous sites. IDF analyses were performed together with comprehensive evaluations of temporal stationarity and spatial heterogeneity of the simulated climate and hydrologic extremes.

2. Study sites

Two mountainous sites with different climate conditions are considered, as shown in Figure 1. One site is near the Mountain Warfare Training Center (MWTC) located in Pickel Meadows on California State Route 108 at 2,100 m above sea level in the Toiyabe National Forest, about 34 km northwest of Bridgeport, California. It features a humid continental climate (Dsb) with cold, relatively snowy winters and dry summers with very warm days and cold mornings. The average annual rainfall is 9.41 inches or 239.0 millimeters. The wettest "rain year" was from July 1968 to June 1969 with 20.76 inches (527.3 mm) and the driest was from July 1959 to June 1960 with 4.37 inches (111.0 mm) [http://w2.weather.gov]. The most precipitation in one month was 7.69 inches (195.3 mm) during January 1969. The most precipitation in 24 hours was 2.59 inches or 65.8 millimeters on January 31, 1963. Average annual snowfall is 49.5 inches or 1.26 meters. The most snowfall in one year was 174.06 inches or 4.42 meters between 1915 and 1916, including 121.0 inches or 3.07 meters in January 1916 [https://wrcc.dri.edu]. The maximum snow cover was 51 inches or 1.30 meters on February 25, 1969.

The other site is Fort Carson (FC), located south of Colorado Springs in Colorado. The FC site features cold semi-arid climate (BSk) with dry winters and wet summers. The summer thunderstorms drive a number of peak runoff events. The FC site gets ~16 inches (406 mm) of rain per year mostly in summer, which is lower than the US average of 39 inches (990 mm); and it gets average snowfall of 54 inches (1372 mm) mostly in March. A total snowfall of 61 inches (1550 mm) was recorded in 1984. The number of days with any measurable precipitation is 42. The most precipitation in 24 hours was 3.02 inches (77 mm) on July 10, 1996.

Both sites have significant amount of snow, but the MWTC area is characterized by thicker snowpack and more frequent rain-on-snow events [Yan et al., 2018]; therefore, the two selected sites allow us to evaluate snowmelt impacts under different hydro-climatic conditions.

3. Methods

a. Regional climate and hydrological models

Regional climate models have been used to study regional climate processes and provide dynamical downscaling of global climate projections (e.g., [Giorgi et al., 1990]) for the past three decades. The Regional Earth System Model (RESM) developed at PNNL is based on the Weather Research and Forecasting (WRF) model [Skamarock et al., 2008] for the atmosphere and the Community Land Model (CLM) [Lawrence et al., 2011] for the land surface, coupled through the flux coupler (CPL7) of the Community Earth System Model (CESM) [Gent et al., 2011] that facilitates exchange of fluxes in a conservative manner. RESM was applied to a North America domain, at a 20-km grid spatial resolution, with lateral boundary conditions and sea surface temperature and sea ice data provided by CESM. The CESM simulations are part of the Coupled Model Intercomparison Project Phase 5 (CMIP5) archive [Taylor et al., 2012]. For the current climate, downscaling was performed for 1975 – 2004 using boundary conditions from a CESM historical run. For the future, two simulations were performed for 2005 – 2100 using boundary conditions from two CESM ensemble members for the Representative Concentration

Pathways RCP4.5 and RCP8.5 scenarios, which are the most widely used emissions scenarios that capture more plausible pathways for the future with mitigation (RCP4.5) and business-as-usual scenario (RCP8.5).

Bias correction was applied to the RESM output following the Bias-Correction Spatial Disaggregation (BCSD) method described by Wood et al. [2004]. In brief, quantile mapping was used to remove biases in the simulated monthly mean temperature and precipitation based on monthly data from North America Land Data Assimilation System (NLDAS-2) [Xia et al., 2012], with a grid spacing of 1/8 degree, which is comparable to that of the RESM simulations (20km). The monthly bias correction factors from the BCSD were then applied to the hourly RESM model outputs. This method of bias correction does not correct for biases of temperature and precipitation at daily and hourly scales. In other words, the co-variability of temperature and precipitation due to weather systems and storms simulated by the regional model is preserved. To bias correct the future climate simulations, a linear trend was fitted to the surface temperature time series between 2005 and 2100 using linear regression and quantile mapping was applied to the residuals after removing the linear trend for each grid cell. No linear trend was removed from the precipitation data because the trend of monthly mean precipitation is generally very small. Although trends may be more apparent for extreme precipitation at hourly and daily time scales, our bias-correction method does not apply to hourly-to-daily scale. The bias-correction approach has been widely used for correcting climate simulations used in previous studies of regional climate change impacts (see [*Maraun*, 2016] and the references therein). The bias-corrected hourly data used in this study was also used in the studies of future changes in regional water stress [*Hejazi et al.*, 2015] and drought [*Wan et al.*, 2017] in the contiguous U.S.

Besides surface temperature and precipitation, DHSVM also requires additional atmospheric forcing such as downward radiation fluxes and humidity. The solar radiation fluxes were bias corrected by subtracting the long-term mean bias based on comparison with the NLDAS-2 data. Humidity was bias corrected by multiplying the RESM simulated relative humidity with the saturation humidity estimated based on the bias-corrected surface temperature. The full set of hourly bias-corrected atmospheric forcing at 1/8 degree resolution for the historical (1975 – 2004) and future (2005 – 2100) RCP8.5 scenario were used to generate high-resolution spatial distributions of meteorological time series at the two selected demonstration sites in mountainous regions with different hydroclimatic conditions and where elevated warming and snowmelt changes in the future may post larger challenges. The bias-corrected time series were used to provide atmospheric forcing for DHSVM.

DHSVM is a physics-based spatially distributed hydrologic model that simulates the effects of soil, vegetation, and topography on the movement of water at and near the land surface. DHSVM models the processes associated with snowpack morphology in the open or under canopy [Storck, 2000; Storck and Lettenmaier, 1999], using a two-layer snowpack representation of snow accumulation and melt, governed by coupled mass and energy balance. The energy balance components of the model address snowmelt, refreezing, and changes in snowpack heat content, while the mass-balance equations address the change of mass during snow accumulation and ablation, transformations in the snow water equivalent, and snowpack water yield [Wigmosta et al., 2002]. The model can account for the effects of topography and vegetation cover on energy and mass exchange at the snow surface, including topographic and canopy shading effects on radiative input to snowpack. The canopy impacts have not been quantified and calibrated over the study sites; therefore the canopy effect was not included in this study, which focused on topography driven spatial heterogeneity in precipitation and AWR. Meteorological inputs required by DHSVM include hourly precipitation, air temperature, wind speed, relative humidity, and downward shortwave and longwave radiations, provided by the downscaled and bias-corrected RESM outputs.

At every grid location, the amount of the water available for runoff (AWR) was calculated as: $AWR = P - \Delta SWE - S$, at hourly temporal resolution, where P is the hourly precipitation from meteorological input, ΔSWE is the change in snowpack water content over the hourly time step, and S is the change in snow mass due to condensation (negative) or evaporation/sublimation. We calibrated DHSVM and evaluated its simulated SWE at the nearby SNOTEL locations where long-term continuous observations of daily SWE records are available during the historical period 1975–2004.

b. *IDF development*

RESM and DHSVM simulated precipitation and AWR are then used for IDF development. An IDF curve presents the probability of a given precipitation intensity and duration expected to occur at a particular location. Standards have been developed for designing infrastructures based on IDF curves [Wolcott et al., 2009]. Given local precipitation data, IDF curves are developed using frequency analysis by first determining the annual maximum precipitation intensity of the selected duration from n years of historical data and then fitting extreme value distributions to the annual maximum series of event intensity for given durations (e.g., 6-hour, 24-hour). Generalized extreme value (GEV) distribution consists of Gumbel, Fréchet, and Weibull distribution families [*De Haan and Ferreira*, 2007]. In this study, we adopt the Gumbel distribution [*Gumbel*, 2012; *Peck et al.*, 2012; *Shaw*, 2005], which is also called the type I extreme value (EV1) distribution (see Figure S1 in the Supplementary for the goodness of fit evaluation, which supports the choice of the Gumbel distribution). The IDF derivation procedure is used to develop IDF curves for both AWR and precipitation.

c. *Non-stationarity evaluation*

IDFs are developed for stationary series. However, changes in extreme precipitation events can lead to a revision of standards for designing civil engineering infrastructures to prevent water management infrastructures from performing below the designated guidelines in the future [*Prodanovic and Simonovic*, 2007; *Simonovic and Peck*, 2009], and the changes very likely will result in non-stationary series, which invalidate the IDF computation by violating data stationarity [*Cunderlik and Burn*, 2003].

To obtain reliable IDFs, two solutions can be implemented: (1) identify quasi-stationary time windows from the time series of interest [Appel and Brandt, 1983; Michelangeli et al., 1995] and compute the IDF curves using data for the corresponding time windows; (2) introduce a parameter representing the trend in the means of the extreme value distributions [*AghaKouchak et al.*, 2012; *Cheng and AghaKouchak*, 2014; *Cheng et al.*, 2014; *Patel et al.*, 2015; *Ren et al.*, 2019]. Here we adopt the first approach using standard IDF calculation but with systematic evaluation of the extremes time series to identify stationary or quasi-stationary time windows. The metrics used include Sen's slope [Sen, 1968] and MK test[Kendall, 1975; Mann, 1945]. The MK test is based entirely on ranks and hence is robust to non-normality. By dividing the entire 1975-2100 simulation time period into many smaller time windows with a fixed size of n (e.g., 30) years, we can perform the MK test on each window, and compute the MK test failure rate as the number of "failed" small windows relative to the total number of n-year-long time windows.

In order to have adequate number of time windows for reliable estimates of failure rates, the time windows are allowed to have up to 50% overlap with adjacent windows. The failure rate evaluation is done for 10-, 15-, 20-, 25-, 30-, 50-year time windows, and low failure rates correspond to quasi-stationary time windows.

d. Geostatistical modeling and mapping

IDFs curves were developed at each numerical grid cell. To illustrate the spatial distribution and patterns of precipitation intensity and AWR for given frequencies and durations, we used geostatistical approaches to produce spatial maps of precipitation/AWR intensity and/or their temporal changes (e.g., differences between future RCP8.5 and historical). The geostatistical models capture the spatial heterogeneity and correlation patterns in the properties associated with spatial locations. Such heterogeneity can be attributed to many factors including topography or microclimates via various physical processes such as temperature lapse rates and snow melt [*Lisi et al.*, 2015; *Nijzink et al.*, 2016; *Sohrabi et al.*, 2019; *Sun et al.*, 2018; *Sun et al.*, 2019]. These impacts have been represented in the physics-based numerical modules in the regional climate models and DHSVM.

Intensity of design storms at different spatial points are treated as spatial random functions (SRFs), $Z(\boldsymbol{x})$, and spatial variogram models $\gamma(\boldsymbol{x}, \boldsymbol{x}')$ are fitted. The variogram characterizes the spatial continuity and correlation patterns of a field with a mathematical formulation $\gamma(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{2}E(\{Z(\boldsymbol{x}) - Z(\boldsymbol{x}')\}^2)$, which can be approximated with experimental variogram in practice as $\hat{\gamma}(h) = \frac{1}{2N(h)}\sum_{\{(i,j)||\boldsymbol{x}_i - \boldsymbol{x}_j| \approx h\}}(z_i - z_j)^2$, where z_i and z_j are the realizations of Z at spatial locations $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, respectively, and h is the desired lag distance. In this study, we fitted the experimental variogram using the exponential model in the form of $\gamma(h) = C_0[1 - \exp(-\frac{h}{I})]$, where C0 and I are the variance and correlation length parameters to be fitted [Deutsch and Journel, 1998]. The fitted model quantifies the variability and spatial continuity, and was used for spatial mapping of the SRF property (e.g., intensity of design storms or AWR).

4. Results

a. Quasi-stationary time windows

The Mann-Kendall (MK) test failure rate was used to identify quasi-stationary time windows in the study period, as explained above. Figure 2 (a) and (b) show that for annual maxima precipitation given three different durations at the MWTC site, the failure rate (aggregated across 36 grid cells) is the lowest when the time window size is 30 years; that is, the annual maxima precipitation series for 30-year time windows are more likely to be stationary than those with a different time window size. The failure rate evaluation was done for annual maxima AWR series as well (Figure 2 (c) and (d)), where it is more convincing that 30-year time window is the best choice to achieve quasi-stationary data for reliable IDF development. The optimal time windows are found to be consistent for both representative concentration pathway (RCP4.5 and RCP8.5) scenarios. Temperature is usually not subject to IDF analysis, but when simulated temperature is considered, we found that in order to achieve quasi-stationary time windows, the maximum window size is about 30 years under RCP4.5 scenario, and about 15 years under RCP8.5 scenario due to a strong positive trend in temperature.

MK tests were also performed at the 30 grid cells at the FC site (Figure 2 (e), (f), (g) and (h)). The summarized MK test failure rates do not have clear patterns with respect to the time window size, but generally a 30-year time window is the best choice since the corresponding failure rate is lower than 5% for both precipitation and AWR extreme events of the selected durations under either RCP4.5 or RCP8.5 scenarios.

Based on the above evaluation, a 30-year time period at the two study sites provide appropriate precipitation and AWR annual maxima data for IDF evaluations with minimal violation of the stationarity assumption.

b. *IDF curves*

Extreme events characteristics near the end of the 21st century is the focus of many scientific studies involving future climate projections. Therefore, data from the 1975-2004 and 2071-2100 time periods, each 30 years long, were used to fit extreme value Gumbel distributions used to derive IDF curves and evaluate the impacts of snowmelt as well as increasing greenhouse gas emissions on the spatial and temporal changes of extreme events of precipitation and AWR.

Figure 3 shows the derived precipitation IDFs for three selected individual grid cells at different elevations in the FC study area, for the historical 1975-2004 time period, and for the future 2071-2100 time period under RCP4.5 and RCP8.5 scenarios. In general, the curves shift upward from historical to future time period indicating more frequent and more intense precipitation. But the degree of increase in the extreme event frequency and intensity is dependent on the durations, and it varies remarkably from cell from cell. IDFs and temporal changes in IDFs are more different between regions on the mountain and those in the valley. The derived AWR IDF curves (not shown) are very similar, which generally show higher intensity for given duration and frequency from historical to RCP4.5 to RCP8.5, and the amount of intensity increase also varies spatially.

c. Geostatistical modeling

Geostatistical analyses were performed for AWR and precipitation intensity for given durations and frequencies, and fitted variogram models were used to generate spatial intensity maps of these events. The fitted variogram models for 6-hour 100-year precipitation intensity are shown in Figure 4. Events of different durations have been analyzed, but only figures for the 6-hour events are shown hereafter for brevity. Increasing greenhouse gas emissions seem to have resulted in much shorter spatial correlation ranges (i.e., weaker spatial continuity) and larger total variance, both indicating stronger spatial heterogeneity. Under climate change, stronger contrast in precipitation intensity within a short distance is expected in the future.

d. Spatiotemporal distributions of precipitation and AWR extremes

Figure 5 shows the spatial distribution of precipitation intensity (top row) at the FC site, of 6-hour 100-year events against the distributions of AWR (bottom row) for the same duration and frequency, during the historical (1975-2004) and future (2071-2100) time periods. The figure enables us to evaluate the impacts of both snowmelt processes (by comparing AWR vs. precipitation intensity) and increasing greenhouse gas emissions (by comparing historical vs. RCP4.5&8.5). In the spatial distribution map of AWR extremes, the local anomalies are reduced compared to the precipitation distributions, so AWR events exhibit smoother spatial patterns, particularly along the gradient between the mountain and the adjacent valley. The smoothing pattern corresponds to stronger spatial continuity of AWR compared to precipitation. At some locations, snowmelt might have coincide with rainfall, resulting a higher AWR than precipitation; for example, on the slope to the west of Colorado Springs, the 6-hour 100-year

AWR is about 5%-10% higher than 6-hour 100-year precipitation intensity during 1975-2004. But in some areas, AWR is lower than precipitation; for example, in the middle-west region (near Victor) of the study area, the 6-hour 100-year AWR is about 15% lower than precipitation, during 2071-2100, under RCP8.5. This is because precipitation is accumulated in the snowpack and not immediately available for runoff, so the annual maxima of (rainfall + snowmelt) are smaller than that of (rainfall + snowfall). This situation can occur in inland mountains where the temperature remains below freezing even in the warmer future so snow accumulation reduces available water for immediate runoff during winter storms. These findings show that if precipitation IDFs are used, the hydrologic infrastructure under- or over-design can be up to 15% at the FC site, resulting in potential snowmelt related flood risk. By using the AWR IDFs as guidance, the over- or under-design issues can be alleviated.

By comparing the results for historical and future time periods, the impact of increasing greenhouse gas emissions can be evaluated and quantified. The local high of historical precipitation intensity and AWR of about 8.5 mm/hr seems to shift northeastward and reach a local high of above 10.5 mm/hr under RCP4.5 and above 12.0 mm/hr under RPC8.5, representing more than 20% and 40% increase respectively. There are great differences between precipitation and AWR under historic conditions, and these differences are reduced under RCP4.5 and 8.5. This could be due to a shift from rain/rain-on-snow peaks under current climate to rain dominated peaks in the future (i.e., AWR is essentially rainfall in the future).

Figure 6 shows the spatial distributions of precipitation intensity and AWR, but for the MWTC area. Similar to Figure 5, for either historical or future events (particularly under RCP8.5), the spatial contrast is weaker in AWR compared to that in precipitation, with some local highs/lows reduced or removed. One difference between the two sites is that snowmelt dominates AWR at MWTC more compared to FC, possibly because of thicker snowpack and more frequent rain-on-snow events in the MWTC area [Yan et al., 2018]. Atmospheric rivers are responsible for a majority of flooding events in the Sierra Nevada [Ralph et al., 2006], where atmospheric river conditions occur during 17% of all winter precipitation events, but they are associated with 50% of rain-on-snow events, because atmospheric river conditions are on average 2^{o}C warmer than the average conditions [Guan et al., 2016]. Near the southwest region of the MWTC area, AWR is about 50% higher than precipitation, for both the historical time period and the future period under the RCP4.5 scenario, suggesting that snowmelt and rainfall might have comparable contributions to AWR extremes. Under the RCP8.5 scenario, however, the snowmelt contribution is smaller (about 10%-20% of AWR) in the southwest region. This is because DHSVM accounts for the change of energy input to snowpack, warmer air temperature in RCP8.5 certainly affects snowpack dynamics, including snowpack duration, ablation rate, and timing of melt. In addition, the phase of precipitation (falling as snow or rain) is determined by air temperature. Because of the significant temperature increase under the RCP8.5 scenario, it is very likely that a larger fraction of precipitation falls as rain so snowpack on the ground becomes thinner and the frequency of rain-on-snow also becomes smaller [L. Ruby Leung et al., 2004]. In this case, melt becomes less important, so does AWR comparing to precipitation.

In the northwest region, on the other hand, AWR is 5%-20% lower than precipitation intensity, such that rainfall + snowfall extremes are greater than rainfall + snowmelt extremes. This is likely due to longer and slower release of meltwater than snowfall. Snowmelt has relatively weak contributions in other sub-regions of the MWTC area under RCP8.5.

The impacts of increasing greenhouse gas emissions are outstanding in the northeast region, where the historical local high of precipitation intensity of about 12-17 mm/hr increases to more than 17 mm/hr (RCP4.5) and 22 mm/hr (RCP8.5), representing more than 20% and 50% increase respectively. Similar to precipitation, the corresponding AWR increases from 12 mm/hr to 17mm/hr and 22 mm/hr, representing more than 40% and 80% increases respectively. The southwest MWTC area has a similar amount of increases in precipitation intensity (about 20% and 70%) and AWR (about 25% and 60%) in the future.

Longer duration extreme events (e.g., 24-hour duration precipitation and AWR) are also studied for completeness. Figure 7 shows the spatial intensity distributions of 24-hour 100-year events of precipitation (top row) against AWR (bottom row), for historical (1975-2004) and future (2071-2100) time periods, at the snow-dominated MWTC site.

The impacts of snowmelt and warming on 24-hour 100-year events are somewhat different compared to those on 6-hour 100-year events. One observation is that the contribution of snowmelt is more pronounced – AWR is up to 70% higher than precipitation intensity near the southwest region for the historical and future RCP4.5 cases, although for the future RCP8.5 case, the snowmelt contribution remains about 10%-20% of AWR. The impacts of warming are outstanding in the northwest region, with precipitation intensity increasing by up to 20% (RCP4.5) and 60% (RCP8.5), and with AWR increasing by up to 30% (RCP4.5) and 60% (RCP8.5) respectively. Figure S2 in the Supplementary shows that the levels of trend and spatial heterogeneity are comparable at the FC site, but with heterogeneous behaviors associated with local topography as expected.

In summary, snowmelt and increasing greenhouse gas emissions have comparable impacts on AWR IDFs for 6-hour and 24-hour duration events, particularly in sub-regions with significant snowfall and snowmelt. At both the MWTC and FC sites, the snowmelt impact is consistent and significant, especially in snow-dominated areas. Due to the influence of topography, the impacts of both snowmelt and warming are very likely to be localized and spatially heterogeneous. The spatial variations of extreme events' intensity are comparable to their temporal variations averaged across the study areas.

5. Discussion and Conclusions

Based on the assessment of stationarity (MKtest failure rate), two 30-year quasi-stationary time periods, 1975-2004 and 2071-2100, were considered for IDF computation and geospatial mapping of precipitation intensity and AWR for 6- and 24-hour duration events. It is interesting to examine the entire 125-year period to fully understand how strong the temporal trends in precipitation and AWR might be during the 21st century. Sen's slope is used to quantify the trends, as shown in Figure 8, where warmer colors represent positive Sen's slopes and positive trends. Figure 8 shows that under RCP8.5 scenario, the northwest region of the FC area has a weak negative trend in 6-hour AWR during the period 1975-2004, and weak positive trend after 2005, while the eastern and southern regions have weak positive trends throughout the four periods. At places with relatively strong positive/negative trends, a time window shorter than 30-years might be preferable to achieve quasi-stationarity of the annual maxima data for IDF development. The positive trend is slightly stronger and spatially more consistent in the MWTC site (see Figure S3 in the Supplementary).

The Sen's slope analyses were also performed for other durations, for both AWR and precipitation intensity, for both RCP scenarios, and for both sites. Overall, the analyses confirmed that quasi-stationarity can be achieved by using 30-year time windows, and they also illustrated that spatial heterogeneity exists in the temporal trends in both precipitation intensity and AWR during historical and future time periods.

In this study, we integrated high-resolution bias-corrected climate simulations with the spatially distributed snow-hydrology model DHSVM to provide more accurate and reliable AWR estimates for IDF analyses. AWR takes into account both rainfall and snowmelt, and is more directly linked to peak runoff than precipitation in mountainous areas with significant snowfall and snowmelt; the corresponding AWR IDFs can be used to reduce over- or under-design of hydrological infrastructure for cost-savings or risk reduction. In the study areas, AWR tends to have stronger spatial continuity compared to precipitation intensity, although both have local areas exhibiting relatively strong temporal trends, or areas that are sensitive to factors such as snowmelt or warming.

Non-stationarity in precipitation intensity and AWR for events with various durations was evaluated using the Mann-Kendall tests, and quasi-stationary time windows were identified such that the subsequent IDF derivations do not violate the stationarity assumption. Impacts of snowmelt and increasing greenhouse gas emissions are found to be spatially variable due to spatial heterogeneity particularly related to topography. In the sub-areas with significant snowfall (e.g., southwest MWTC), snowmelt has a bigger contribution than rainfall does to AWR, and the snowmelt impact is comparable to the impact of increasing greenhouse gas emission on AWR extremes under the RCP8.5 scenario.

Tremendous efforts have been made in developing precipitation IDFs at local scales based on historical data or future climate projections from one GCM or an ensemble of GCMs or regional climate models. Use of ensemble model outputs is important for characterizing uncertainties in the climate projections. Multi-model projections are available from CMIP5 and large ensembles of climate projections by single models to capture internal variability are also available [Kay et al., 2015]; however, statistically downscaled data driven by multiple GCMs are only available at daily time step and lack a number of important atmospheric forcing variables such as radiation and winds, which are important for simulating snowpack [e.g., Pierce et al. 2015]. In contrast, our regional climate simulations provide a large number of physically consistent variables at hourly time step that are better suited for distributed hydrologic modeling important for simulating runoff driven by both rain and snowmelt. This enabled us to fully evaluate and distinguish the impacts of three factors: snowmelt, increasing greenhouse gas emissions, and spatial heterogeneity on IDFs. However, uncertainty in the methods used to bias correct the dynamically downscaled climate simulations should also be recognized. In particular, the bias-corrected dataset used in this study ignored bias correction at hourly and daily scale, which may introduce biases and uncertainties in the hydrologic simulations driven by the atmospheric forcing. As noted in [Maraun 2016], bias correction of higher quantiles can produce noisy results and run into overfitting and implausible applications. This may be particularly true in mountainous regions where higher frequency data are not commonly available so there is a need for further improvement of bias-correction methodology to improve the use of dynamically downscaled simulations. Furthermore, impacts of ensemble model uncertainty on snowmelt-incorporated IDFs will be addressed in future work. In addition to using ensemble models for reducing parametric and model structural uncertainties, alternative non-stationary IDF analysis with uncertainty bounds will be integrated to enable reduction and quantification of the overall IDF estimation uncertainty.

Acknowledgement

This study is supported by the Department of Defense (DOD) Strategic Environmental Research and Development Program (SERDP) (Grant no. RC-2546).

Additional Information

Competing interests: The authors declare no competing financial and/or non-financial interests.

References

AghaKouchak, A., D. Easterling, K. Hsu, S. Schubert, and S. Sorooshian (2012), Extremes in a changing climate: detection, analysis and uncertainty, Springer Science & Business Media.

Appel, U., and A. V. Brandt (1983), Adaptive sequential segmentation of piecewise stationary time series, Information sciences, 29(1), 27-56.

Bales, R. C., N. P. Molotch, T. H. Painter, M. D. Dettinger, R. Rice, and J. Dozier (2006), Mountain hydrology of the western United States, Water Resour Res, 42(8).

Bookhagen, B., and D. W. Burbank (2010), Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge, Journal of Geophysical Research: Earth Surface, 115(F3).

Bracken, C., B. Rajagopalan, L. Cheng, W. Kleiber, and S. Gangopadhyay (2016), Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain, Water Resour Res, 52(8), 6643-6655.

Cheng, L., and A. AghaKouchak (2014), Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate, Scientific reports, 4, 7093.

Cheng, L., A. AghaKouchak, E. Gilleland, and R. W. Katz (2014), Non-stationary extreme value analysis in a changing climate, Climatic Change, 127(2), 353-369.

Chow, V. T., D. R. Maidment, and L. W. Mays (1988), Applied Hydrology, McGraw-Hill, New York.

Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme precipitation return levels, J Am Stat Assoc, 102(479), 824-840.

Cunderlik, J. M., and D. H. Burn (2003), Non-stationary pooled flood frequency analysis, J Hydrol, 276(1), 210-223.

De Haan, L., and A. Ferreira (2007), Extreme value theory: an introduction, Springer Science & Business Media.

Deutsch, C. V., and A. G. Journel (1998), Geostatistical software library and user's guide, 369 pp., Oxford University Press, New York.

Fang, S., L. Xu, H. Pei, Y. Liu, Z. Liu, Y. Zhu, J. Yan, and H. Zhang (2014), An integrated approach to snowmelt flood forecasting in water resource management, IEEE transactions on industrial informatics, 10(1), 548-558.

Gent, P. R., et al. (2011), The Community Climate System Model Version 4, Journal of Climate, 24(19), 4973-4991, doi: 10.1175/2011JCLI4083.1.

Ghosh, S., D. Das, S.-C. Kao, and A. R. Ganguly (2012), Lack of uniform trends but increasing spatial variability in observed Indian rainfall extremes, Nature Climate Change, 2(2), 86.

Giorgi, F., M. R. Marinucci, and G. Visconti (1990), Use of a limited‐area model nested in a general circulation model for regional climate simulation over Europe, Journal of Geophysical Research: Atmospheres, 95(D11), 18413-18431.

Guan, B., D. E. Waliser, F. M. Ralph, E. J. Fetzer, and P. J. Neiman (2016), Hydrometeorological characteristics of rain‐on‐snow events associated with atmospheric rivers, Geophys Res Lett, 43(6), 2964-2973.

Gumbel, E. J. (2012), Statistics of extremes, Courier Corporation, Massachusetts.

Guo, Y. (2006), Updating rainfall IDF relationships to maintain urban drainage design standards, Journal of Hydrologic Engineering, 11(5), 506-509.

Haddad, O. B., and M. Moravej (2015), Discussion of “trend, independence, stationarity, and homogeneity tests on maximum rainfall series of standard durations recorded in turkey” by tefaruk haktanir and hatice citakoglu, Journal of Hydrologic Engineering, 20(10), 07015016.

Hejazi, M. I., N. Voisin, L. Liu, L. M. Bramer, D. C. Fortin, J. E. Hathaway, M. Huang, P. Kyle, L. R. Leung, and H.-Y. Li (2015), 21st century United States emissions mitigation could increase water stress more than the climate change it is mitigating, Proceedings of the National Academy of Sciences, 112(34), 10635-10640.

Hurrell, J. W., et al. (2013), The Community Earth System Model: A Framework for Collaborative Research, B Am Meteorol Soc, 94(9), 1339-1360, doi: 10.1175/bams-d-12-00121.1.

Kampf, S. K., and E. E. Richer (2014), Estimating source regions for snowmelt runoff in a Rocky Mountain basin: tests of a data ‑ based conceptual modeling approach, Hydrological Processes, 28(4), 2237-2250.

Kao, S. C., and A. R. Ganguly (2011), Intensity, duration, and frequency of precipitation extremes under 21st ‑ century warming scenarios, Journal of Geophysical Research: Atmospheres, 116(D16).

Kattelmann, R. (1997), Flooding from rain-on-snow events in the Sierra Nevada, IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 239, 59-66.

Kay, J., C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. Arblaster, S. Bates, G. Danabasoglu, and J. Edwards (2015), The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability, B Am Meteorol Soc, 96(8), 1333-1349.

Kendall, M. (1975), Rank Correlation Measures [M], London: Charles Griffin.

Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl (2007), Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, Journal of Climate, 20(8), 1419-1444.

Koutsoyiannis, D., and G. Baloutsos (2000), Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences, Natural Hazards, 22(1), 29-48.

Lawrence, D., et al. (2011), Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model, Journal of Advances in Modeling Earth Systems, 3(Art. #1), 27 pp., doi: 10.1029/james.2011.3.1.

Leung, L. R., and S. J. Ghan (1999), Pacific Northwest Climate Sensitivity Simulated by a Regional Climate Model Driven by a GCM. Part II: $2\times CO_2$ Simulations, Journal of Climate, 12(7), 2031-2053, doi: doi:10.1175/1520-0442(1999)012<2031:PNCSSB>2.0.CO;2.

Leung, L. R., Y. Qian, X. Bian, W. M. Washington, J. Han, and J. O. Roads (2004), Mid-Century Ensemble Regional Climate Change Scenarios for the Western United States, Climatic Change, 62(1), 75-113, doi: 10.1023/b:clim.0000013692.50640.55.

Lima, C. H., and U. Lall (2010), Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow, J Hydrol, 383(3-4), 307-318.

Lisi, P. J., D. E. Schindler, T. J. Cline, M. D. Scheuerell, and P. B. Walsh (2015), Watershed geomorphology and snowmelt control stream thermal sensitivity to air temperature, Geophys Res Lett, 42(9), 3380-3388.

López-Moreno, J., S. Goyette, and M. Beniston (2009), Impact of climate change on snowpack in the Pyrenees: Horizontal spatial variability and vertical gradients, J Hydrol, 374(3-4), 384-396.

Mailhot, A., S. Duchesne, D. Caya, and G. Talbot (2007), Assessment of future change in intensity–duration–frequency (IDF) curves for Southern Quebec using the Canadian Regional Climate Model (CRCM), J Hydrol, 347(1), 197-210.

Mann, H. B. (1945), Nonparametric tests against trend, Econometrica: Journal of the Econometric Society, 245-259.

Maraun, D. (2016), Bias correcting climate change simulations-a critical review, Current Climate Change Reports, 2(4), 211-220.

McCuen, R. H. (1998), Hydrologic design and analysis, Prince Hall, New Jersey.

Michelangeli, P.-A., R. Vautard, and B. Legras (1995), Weather regimes: Recurrence and quasi stationarity, Journal of the atmospheric sciences, 52(8), 1237-1256.

Mirhosseini, G., P. Srivastava, and L. Stefanova (2013), The impact of climate change on rainfall Intensity–Duration–Frequency (IDF) curves in Alabama, Regional Environmental Change, 13(1), 25-33.

Nijzink, R. C., L. Samaniego, J. Mai, R. Kumar, S. Thober, M. Zink, D. Schäfer, H. H. Savenije, and M. Hrachowitz (2016), The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological models, Hydrology and Earth System Sciences, 20(3), 1151-1176.

Patel, A. P., P. Timbadiya, and P. Patel (2015), Development of IDF curves under non-stationary meteorological condition, IAHR World Congress 36.

Peck, A., P. Prodanovic, and S. P. Simonovic (2012), Rainfall intensity duration frequency curves under climate change: city of London, Ontario, Canada, Canadian Water Resources Journal, 37(3), 177-189.

Prodanovic, P., and S. P. Simonovic (2007), Development of rainfall intensity duration frequency curves for the City of London under the changing climate, Department of Civil and Environmental Engineering, The University of Western Ontario.

Ragno, E., A. AghaKouchak, C. Love, F. Vahedifard, L. Cheng, and C. Lima (2017), Intensity-Duration-Frequency Curves for US Cities in a Warming Climate, EGU General Assembly Conference Abstracts, 19, 18246.

Ralph, F. M., P. J. Neiman, G. A. Wick, S. I. Gutman, M. D. Dettinger, D. R. Cayan, and A. B. White (2006), Flooding on California's Russian River: Role of atmospheric rivers, Geophys Res Lett, 33(13).

Ren, H., Z. J. Hou, M. Wigmosta, Y. Liu, and L. R. Leung (2019), Impacts of Spatial Heterogeneity and Temporal Non-Stationarity on Intensity-Duration-Frequency Estimates—A Case Study in a Mountainous California-Nevada Watershed, Water, 11(6), 1296.

Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's tau, J Am Stat Assoc, 63(324), 1379-1389.

Shaw, E. (2005), Hydrology in practice, CRC Press.

Simonovic, S. P., and A. Peck (2009), Updated rainfall intensity duration frequency curves for the City of London under the changing climate, Department of Civil and Environmental Engineering, The University of Western Ontario.

Skamarock, W., J. Klemp, J. Dudhia, D. Gill, D. Barker, M. Dudha, X. Huang, W. Wang, and Y. Powers (2008), A description of the advanced research WRF Version 3, NCAR technical note, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA.

Sohrabi, M. M., D. Tonina, R. Benjankar, M. Kumar, P. Kormos, D. Marks, and C. Luce (2019), On the role of spatial resolution on snow estimates using a process ‐ based snow model across a range of climatology and elevation, Hydrological Processes, 33(8), 1260-1275.

Storck, P. (2000), Trees, snow and flooding: An investigation of forest canopy effects on snow accumulation and melt at the plot and watershed scales in the Pacific Northwest*Rep.*, Washington State University Department of Civil and Environmental Engineering.

Storck, P., and D. P. Lettenmaier (1999), Predicting the effect of a forest canopy on ground snow accumulation and ablation in maritime climates, Proceedings of 67th Western Snow Conference, Colo. State Univ. Fort Collins, 1-12.

Sun, N., M. Wigmosta, T. Zhou, J. Lundquist, S. Dickerson‐Lange, and N. Cristea (2018), Evaluating the functionality and streamflow impacts of explicitly modelling forest‐snow interactions and canopy gaps in a distributed hydrologic model, Hydrological processes, 32(13), 2128-2140.
Sun, N., H. Yan, M. S. Wigmosta, L. R. Leung, R. Skaggs, and Z. Hou (2019), Regional Snow Parameters Estimation for Large‐Domain Hydrological Applications in the Western United States, Journal of Geophysical Research: Atmospheres, 124(10), 5296-5313.
Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, B Am Meteorol Soc, 93(4), 485-498.
Teegavarapu, R. S. (2013), Climate change‐sensitive hydrologic design under uncertain future precipitation extremes, Water Resour Res, 49(11), 7804-7814.
Voisin, N., H. Li, D. Ward, M. Huang, L.-Y. R. Leung, and M. Wigmosta (2013a), On an improved sub-regional water resources management representation for integration into earth system models, Hydrology and Earth System Sciences Discussions, 10, 3501-3540, doi: doi:10.5194/hessd-10-3501-2013.
Voisin, N., H. Li, D. Ward, M. Huang, M. Wigmosta, and L. Leung (2013b), On an improved sub-regional water resources management representation for integration into earth system models, Hydrology and Earth System Sciences, 17(9), 3605-3622.
Wan, W., J. Zhao, H. Y. Li, A. Mishra, L. Ruby Leung, M. Hejazi, W. Wang, H. Lu, Z. Deng, and Y. Demissisie (2017), Hydrological drought in the Anthropocene: impacts of local water extraction and reservoir regulation in the US, Journal of Geophysical Research: Atmospheres, 122(21).
Wang, Y., L. R. Leung, J. L. McGREGOR, D.-K. Lee, W.-C. Wang, Y. Ding, and F. Kimura (2004), Regional climate modeling: progress, challenges, and prospects, Journal of the Meteorological Society of Japan. Ser. II, 82(6), 1599-1628.
Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier (1994), A distributed hydrology‐vegetation model for complex terrain, Water Resour Res, 30(6), 1665-1679.
Wigmosta, M. S., B. Nijssen, P. Storck, and D. Lettenmaier (2002), The distributed hydrology soil vegetation model, Mathematical models of small watershed hydrology and applications, 7-42.
Wolcott, S., M. Mroz, and J. Basile (2009), Application of Northeast Regional Climate Center Research Results for the Purpose of Evaluating and Updating Intensity-Duration-Frequency (IDF) Curves—Case Study: Rochester, New York, World Environmental and Water Resources Congress 2009: Great Rivers, 1-9.
Wood, A. W., L. R. Leung, V. Sridhar, and D. Lettenmaier (2004), Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change, 62(1-3), 189-216.
Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, and J. Meng (2012), Continental‐scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS‐2): 1. Intercomparison and application of model products, Journal of Geophysical Research: Atmospheres, 117(D3).
Yan, H., and H. Moradkhani (2015), A regional Bayesian hierarchical model for flood frequency analysis, Stochastic environmental research and risk assessment, 29(3), 1019-1036.
Yan, H., N. Sun, M. Wigmosta, R. Skaggs, Z. Hou, and R. Leung (2018), Next‐Generation Intensity‐Duration‐Frequency Curves for Hydrologic Design in Snow‐Dominated Environments, Water Resour Res, 54(2), 1093-1108.
Yan, H., N. Sun, M. Wigmosta, R. Skaggs, Z. Hou, and L. R. Leung (2019a), Next-Generation Intensity–Duration–Frequency Curves to Reduce Errors in Peak Flood Design, Journal of Hydrologic Engineering, 24(7), 04019020.
Yan, H., N. Sun, M. Wigmosta, R. Skaggs, L. R. Leung, A. Coleman, and Z. Hou (2019b), Observed spatiotemporal changes in the mechanisms of extreme water available for runoff in the western United States, Geophys Res Lett, 46(2), 767-775.

List of Figures