The Role of Snowmelt Temporal Pattern in Flood Estimation for A Small Snow-Dominated Basin in the Sierra Nevada

Hongxiang Yan[1]; Ning Sun[1]; Mark S. Wigmosta[1,2]; Zhuoran Duan[1,2]; Ethan D. Gutmann[3]; Bert Kruyt[3]; and Jeffrey R. Arnold[4]

[1] Earth Systems Science Division, Pacific Northwest National Laboratory, Richland, WA, USA

[2] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA

[3] National Center for Atmospheric Research, Boulder, CO, USA

[4] MITRE Corporation, McLean, VA, USA

Correspondence:
Hongxiang Yan, hongxiang.yan@pnnl.gov

Key Points:

- Standard rainfall hyetographs substantially underestimate floods in a small snow-dominated basin in the Sierra Nevada

- Snowmelt hyetograph shows a more rapid rise (i.e., higher intensity) compared to the standard rainfall hyetographs used in hydrologic design

- A general method to develop probabilistic hyetographs that represent the underling flood-generation mechanism is described

Abstract

Prior research confirmed the substantial bias from using precipitation-based intensity-duration-frequency curves (PREC-IDF) in design flood estimates and proposed next-generation IDF curves (NG-IDF) that represent both rainfall and snow processes in runoff generation. This study improves the NG-IDF technology for a snow-dominated test basin in the Sierra Nevada. A well-validated physics-based hydrologic model, the Distributed Hydrology Soil Vegetation Model (DHSVM), is used to continuously simulate snowmelt and streamflow that are used as benchmark datasets to systematically assess the NG-IDF technology. We find that, for the studied small snow-dominated basin, the use of standard rainfall hyetographs in the NG-IDF technology leads to substantial underestimation of design floods. Thus, we propose probabilistic hyetographs that can represent unique patterns of events with different underlying mechanisms. For the test basin where flooding events are generated entirely by snowmelt, we develop a hyetograph that characterizes snowmelt temporal patterns, which greatly improves the performance of NG-IDF technology in design flood estimates. In contrast to the standard rainfall hyetographs characterized by a symmetrically peaked, bell-shaped curve, the snowmelt hyetograph displays a more rapid rise (i.e., greater intensity) and a distinct diurnal pattern influenced by solar energy input. The results also show that the uncertainty of hyetography plays an important role in design flood estimation and can have important implications for future flood projections.

Plain Language Summary

In recent years, flood hazards have gained increasing attention from national and international homeland security communities. Accurately assessing floods is crucial for many hydrologic applications, including infrastructure design, planning, and renewal, as well as the national flood insurance program. This research focuses on evaluating and enhancing the next-generation flood design technology, which is an improvement over the traditional rainfall-based method that does not account for snow processes in flood generation. Our study reveals a significant underestimation of floods when using standard rainfall temporal pattern in a small snow-dominated basin. To address this issue, we propose probabilistic curves that consider the temporal patterns of snowmelt, resulting in a considerable reduction in flood estimation errors. In contrast to the standard rainfall temporal pattern characterized by a symmetrically peaked, bell-shaped curve, the snowmelt temporal pattern displays a more rapid rise (i.e., greater intensity) and a distinct diurnal pattern influenced by solar energy input. The results demonstrate that the next-generation flood design technology has the potential to complement the traditional method for hydrologic design in snow-dominated regions, providing a consistent design approach in both rain-dominated and snow-dominated areas.

1 Introduction

The repeated recurrence of high-profile flood events (e.g., California in 2017, Michigan in 2020, Germany and Belgium in 2021, Yellowstone National Park in 2022) has resulted in major public safety concerns and motivated U.S. Department of Homeland Security communities to explore new sources and tools for designing proper infrastructure and facilities (ASCE, 2018; ESTCP, 2018). Traditionally, engineers use statistics of observed extreme precipitation, referred to as precipitation-based intensity-duration-frequency (PREC-IDF) curves (Chow et al., 1988), for infrastructure design to withstand extreme flooding events, such as the National Oceanic and Atmospheric Administration (NOAA) Atlas 14 (Perica et al., 2013). This PREC-IDF approach assumes the phase of precipitation as rainfall that immediately starts the rainfall-runoff process. In the mountainous regions of the western United States where snowmelt or rain-on-snow (ROS) is the dominant flood-generating mechanism, the use of the PREC-IDF approach can lead to significant biases (i.e., largely underestimation) of design basis events that subsequently propagate into infrastructure design (Hamlet, 2018; Hou et al., 2019; Yan et al., 2019a, 2020a). For instance, Cho and Jacobs (2020) utilized gridded snow water equivalent (SWE) data for the conterminous United States (CONUS) to calculate design snowmelt values. They then compared these values to the NOAA Atlas 14 for the 44 U.S. states. Their findings indicate that standard design values are surpassed by design snowmelt values in 23% of the total extent.

In snow-dominated regions of the United States, there is a lack of consistent and coordinated surface water design manuals (Yan et al., 2018). Instead, different methods are employed, varying from a basic "blind approach" that solely utilizes the PREC-IDF curves to a "tuning factor approach" that involves augmenting the PREC-IDF values with a snowmelt factor, and more advanced techniques such as utilizing physics-based hydrologic modeling. For example, Snohomish County in Washington State, which experiences a range of hydrologic conditions (from rain-dominant to transitional rain-snow to snow-dominant), recommends the use of NOAA PREC-IDF curves for designing facilities like wetpool treatment facilities (SCDM, 2016). In contrast, Chelan County in the same state, with snow-dominated watersheds, uses the tuning factor approach outlined in the Stormwater Management Manual for Eastern Washington (SWMMEW, 2019). However, this approach is backed by data from only nine sites and rests on the premise that the observed December–February average daily snow depth will melt during a 72h ROS event. The federal Unified Facilities Criteria (UFC) suggests using National Resource Conservation Service (NRCS) Technical Release 55 (TR-55) for small watershed design and the Storm Water Management Model (Rossman, 2004) for large, high-risk infrastructure design projects (UFC, 2013).

Considering the significant expense associated with the use of advanced physically based models, including staff proficiency, manpower, and computational resources, along with the local regulations discussed previously, it's reasonable to anticipate that IDF-based technology will remain crucial in hydrologic design in the foreseeable future, particularly for small-scale infrastructure projects. Furthermore, agencies and regulators are more likely to adopt modifications of the current IDF technology instead of a complete technological shift to a physically based hydrologic modeling approach. Yan et al. (2018) proposed next-generation IDF (NG-IDF) curves to overcome the deficiency of PREC-IDF and provided a consistent IDF design approach for both rain- and snow-dominated regions. Briefly, the NG-IDF curves used the concept of “water available for runoff (W)” rather than “rainfall” to capture the actual water reaching the land surface from combined or individual effects of rainfall, snowmelt, and/or ROS. Based on the observations from nearly 400 Snowpack Telemetry (SNOTEL) stations across the western United

States, Yan et al. (2019a) compared the design flood estimates by using PREC-IDF and NG-IDF curves coupled with the NRCS TR-55 single-event rainfall-runoff model (Cronshey et al., 1986). They showed that about 70% of the stations in the western United States were subject to underdesign with the use of PREC-IDF curves, for which the PREC-IDF curves generated lower design floods by 324%. The lower estimation is due to the fact that in snow-dominated regions, precipitation falls as snow during winter, and the subsequent spring snowmelt or ROS events exhibit higher intensities.

Follow-up research in the NG-IDF context extended the application of NG-IDF beyond SNOTEL to cover the CONUS (Sun et al., 2019, 2022a). However, there are several remaining challenges associated with NG-IDF that we address explicitly in this research. First, an outstanding issue with the use of NG-IDF technology in practice is the choice of hyetograph (i.e., W temporal pattern), which is used to temporally distribute the W magnitude over the selected duration (e.g., 24-h). Yan et al. (2020b) found that by assuming a uniform W hyetograph, the design flood estimates from NG-IDF technology were consistently underestimated in snow-dominated regions of the western United States. Despite that the choice of hyetograph has a significant impact on design flood estimates (Huff, 1990; Hettiarachchi et al., 2018), currently, all standard hyetographs are developed for rainfall cases only such as triangular/NRCS hyetographs (McCuen, 1998; Perica et al., 2013) and we are unaware of a hyetograph developed or studied for W with different underlying mechanisms.

In regions where snowmelt and ROS events are the dominant flood-generating mechanisms, the use of the standard 24h symmetrically peaked, bell-shaped rainfall hyetograph may lead to underestimation of W intensity. This is due to the tendency of snowpack to freeze overnight and melt rapidly during periods of high net solar radiation (Gleason et al., 2013; Musselman et al., 2017). Moreover, the W hyetograph may become nonstationary in the future due to global warming-induced changes in precipitation phase, flood-generating mechanisms, and snowmelt rate. For example, mountainous regions that were previously dominated by snow in March and April during the late 20th century will likely experience increased rainfall frequency in those months during the mid-21st century (Klos et al., 2014; USGCRP, 2018; Cho et al., 2021). In a warmer climate, snowpack decline is expected to decrease ROS events at lower elevations while increasing them at higher elevations due to a transition from snowfall to rain (Musselman et al., 2018; Li et al., 2019). Additionally, snowpack will melt earlier and at a lower rate due to reduced energy availability (Musselman et al., 2017; Yan et al., 2019b). Land use changes, such as afforestation activities, may increase ROS frequency (Mooney and Lee, 2022), while postfire land may increase snowmelt rate (Gleason et al., 2013).

Second, NG-IDF technology has not undergone validation at a basin scale that is more relevant for hydrological design compared to the point or the hillslope scale used in previous studies (Yan et al., 2019a, 2020b). Previous NG-IDF studies also assumed uniform bare soil cover conditions, neglecting the complexity of land surface and associated processes that drive runoff in real design problems. Lastly, the uncertainties associated with the W hyetograph selection on design flood estimates are unknown. The shape of hyetographs is subject to uncertainties due to the inherent natural variability of the climate. For instance, the standard NRCS rainfall hyetographs may differ from the corresponding temporal distribution curves derived from analyzing a specific local storm. Additionally, the shape of hyetograph can also vary in response to atmospheric temperature fluctuations. For instance, Wasko and Sharma (2015) indicate that higher temperatures, irrespective of the climatic region or season, tend to result in steeper temporal patterns. Consequently, it is crucial to accurately quantify the impacts of hyetograph shape uncertainties on the estimation of design floods. This paper aims to address these challenges. Specifically, we aim to

1. Propose a general method to develop a hyetograph of W considering snowmelt and/or ROS processes to illustrate the role of snowmelt temporal patterns in estimating design hydrologic extremes,
2. Evaluate performances of NG-IDF technology in design flood estimates at a design basin scale following the current TR-55 hydrologic design guideline, and
3. Quantify the contribution of W hyetograph selection to the uncertainty in NG-IDF design flood estimates.

2 Methodology

In this section, we introduce the assessment framework of the NG-IDF technology in design flood estimates, followed by detailed descriptions of the study area, data sources, and each framework component.

2.1 Assessment Framework

The framework for quantitative assessment of design flood estimates from the NG-IDF technology is presented in Figure 1. First, we generate continuous simulations of streamflow for the basin of interest using the DHSVM model. DHSVM is forced by 15min meteorological inputs. Using the Mountain Microclimate Simulation Model (Hungerford et al., 1989), daily Livneh meteorological data (Livneh et al., 2013) are disaggregated into 15min time steps. Simulations from process-based models are used because long-term flow measurements are generally scarce at high-latitude locations and especially for headwater streams that fit the small-scale engineering design basin, due to inherent difficulties of access (Bales et al., 2006; Curran et al., 2016; Lundquist et al., 2016). With no observations found within the basin of interest, DHSVM is calibrated and validated against the available snow and streamflow observations at a higher level of the hydrologic unit (which contains the small design basin). Second, the NG-IDF curves are developed based on the DHSVM simulated annual maximum water available for runoff (W) following the methods detailed in section 2.5. Meanwhile, flood frequency statistics (e.g., the 10-year flood) are also derived directly from the DHSVM annual maximum streamflow data, which is used as the benchmark to assess design floods estimated using the NG-IDF technology in the last step. Third, based on the DHSVM W time series, hyetographs of W are studied and proposed for NG-IDF curves. Fourth, with a selection of W hyetograph, the derived NG-IDF curves are used to drive the TR-55 event-based model to estimate the associated design floods. Last, the design floods from TR-55 are compared to the benchmark estimates of design floods from DHSVM continuous streamflow simulation with uncertainty quantification. In our frequency analysis of DHSVM streamflow and NG-IDF curves, we utilize the Monte Carlo method to measure the uncertainty present in the sample data. Further information regarding the quantification of uncertainty is provided in section 2.7.

[Place Figure 1 here]

2.2 Study Area

The Upper West Walker Basin (UWWB) is a sub-basin of the West Walker Basin (Hydrologic Unit Code: 16050302) located in the eastern Sierra Nevada Mountains, California (Figure 2a). The UWWB has a drainage area of about 633 km^2 and encompasses an elevation range of about 1,500 m. The primary source of streamflow is derived from the spring snowmelt, which originates from winter precipitation in the form of snow. Runoff from snowmelt flows through the DoD Marine Corps Mountain Warfare Training Center (MCMWTC) to Walker Lake with peak streamflow occurring between April and June (Hatchett et al., 2016). In this study, the assessment of NG-IDF technology focuses on a selected small test basin that drains into MCMWTC, rather than considering the entire UWWB (Figure 2b). We focus on infrastructure safety at the DoD MCMWTC site due to its critical role in force training to defend U.S. national security interests. The test basin has an area of 2.36 km^2, which aligns with the TR-55 design scale for small watersheds (Cronshey et al., 1986). The test basin has uniform bedrock-derived soil but contains multiple land covers. The dominant land cover is grassland (32%) and forest (31%). The data sources used to derive this information will be introduced in the next section.

[Place Figure 2 here]

2.3 DHSVM Hydrological Simulation

DHSVM (Wigmosta et al., 1994) is a physics-based, spatially distributed hydrologic model that explicitly solves water and energy balances for each model grid cell with a spatial resolution ranging from 10 m to 150 m. DHSVM includes a two-layer canopy submodel that represents canopy processes such as canopy snow processes and evapotranspiration, a two-layer energy-balance submodel for snow accumulation and melt, a four-layer soil submodel, and three-dimensional surface and saturated subsurface flow routing submodels (Wigmosta et al., 2002). DHSVM was originally developed for simulating hydrologic response in mountainous terrain and subsequent developments have enhanced the snow submodel to better represent climate-forest-snow interactions (Sun et al., 2018, 2022b), extended the model capability for simulating urban hydrology and water quality (Cao et al., 2016; Yan et al., 2021; Fullerton et al., 2022), and parallelized the model structure for large-scale application at high-performance computing infrastructure (Perkins et al., 2019).

The required meteorological forcing data for DHSVM include precipitation, air temperature, wind speed, relative humidity, downward solar, and downward longwave radiation. In this study, we ran the DHSVM at a 90 m scale and 15min time step for 33 years from water years 1981–2011 with a 3-year spin-up period. The 15min meteorological forcing data were generated by disaggregating the daily Livneh meteorological data (Livneh et al., 2013) using the Mountain Microclimate Simulation Model (MTCLIM) (Hungerford et al., 1989). Subdaily precipitation assumes daily precipitation occurred at a uniform rate throughout the day. Subdaily air temperatures were estimated using third-order Hermite polynomials spline based on daily minimum and maximum air temperatures. Wind speed is assumed to be constant throughout the day. Relative humidity was calculated based on subdaily temperatures, with the assumption that the dew point is equal to the daily minimum temperature. Downward shortwave radiation was calculated based on daily temperature range and dewpoint temperature using the Thornton and Running (1999) algorithm, where dewpoint temperature was estimated based on daily minimum temperature and precipitation. Downward longwave radiation was calculated using the method described by Prata (1996) and is dependent on subdaily temperatures. For more information on the MTCLIM algorithm, refer to Hungerford et al. (1989). In this study, because the meteorological forcing, SWE, and flow measurements are only available at 24h scale, we developed IDF curves solely at a 24h scale and then utilized the hyetograph to disaggregate the 24h precipitation/W into subdaily scales for TR-55 modeling. Moreover, all other meteorological factors influencing snowmelt within this basin exhibit variations occurring at 15-minute intervals. Given that this basin is primarily influenced by snowmelt rather than rainfall, the effect of subdaily fluctuations in precipitation on peak flow simulations in the DHSVM model will be negligible.

Other data used for DHSVM model input and parameterization include the U.S. Geological Survey (USGS) digital elevation model (DEM) terrain data (Danielson and Gesch, 2011), the NRCS Soil Survey Geographic Database (SSURGO) soil data (http://soils.usda.gov/survey/geography/), and the Multi-Resolution Land Characteristics Consortium National Land Cover Database (https://www.mrlc.gov/). Because no snow or flow measurement exists within or nearby the test basin, we calibrated and validated the DHSVM model for the UWWB, using daily streamflow data from two USGS gauges (ID 10296500 and 10296000) and daily SWE data from the Sonora SNOTEL site (Figure 2a). Data from the SNOTEL site were first screened following a rigorous three-stage SNOTEL quality control filter (Yan et al., 2018) and subsequently bias corrected for snowfall undercatch (Sun et al., 2019). The resulting SNOTEL data is referred to as bias-corrected quality-controlled (BCQC) SNOTEL data and is available at https://climate.pnnl.gov/. DHSVM calibration and validation use the common period of available streamflow, SWE, and Livneh meteorological data from the water year 1984 to 2011.

The performance of the DHSVM for simulating daily streamflow and SWE is evaluated using three statistical metrics including the root-mean-square-error (RMSE), Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), and Kling-Gupta efficiency (KGE) (Gupta et al., 2009). The metrics RMSE and NSE focus on the modeling skills of high flow, while the KGE metric is a multi-objective metric that takes into account the water balance, flow variability, and correlation. The value of NSE varies from $-\infty$ to 1 and a value of 1 indicates a perfect fit between observations and simulations. The KGE metric addresses several shortcomings in NSE (e.g., underestimation of the variability) and is now increasingly used for hydrologic model calibration and evaluation (Knoben et al., 2019; Mizukami et al., 2019; Clark et al., 2021). Like NSE, the value of KGE ranges between $-\infty$ and 1, and a value of 1 indicates a perfect agreement between simulations and observation.

2.4 TR-55 Event-based Rainfall-Runoff Modeling

The NRCS TR-55 guideline (Cronshey et al., 1986) provides a standard procedure for hydrologic design at small-scale watersheds. If the watershed is not divided and the channel routing is not taken into account, it is advisable to refrain from using TR-55 for basins larger than 250 km^2 (Ponce and Hawkins, 1996). The model described in TR-55 assumes a rainfall amount uniformly imposed on the watershed over a specified duration. A design storm depth per unit area (selected from IDF curves with a predefined design hyetograph) is converted to a runoff depth using the runoff curve number (CN) approach, which estimates runoff as a function of the antecedent moisture condition (AMC) and watershed physical characteristics (e.g., soil type, vegetation cover). Runoff is then transformed into a hydrograph by using the unit hydrograph (UH) routing method (Mockus, 1957) that depends on the runoff travel time through segments of the watershed (i.e., time of concentration). The dimensionless NRCS UH has two parameters, peak flood and time to peak, which are empirically estimated using the basin area and time of concentration.

The basin average CN is set to 90 which corresponds to the wet AMC based on the CN table of the TR-55 guideline (Table S1). The physical explanation behind the use of wet AMC is that snowmelt events usually last for days to weeks and are more like to infiltrate soils (except when the soil is frozen), therefore producing a high AMC for runoff generation (Jencso et al., 2009; Yan et al., 2019a). The test basin time of concentration (t_c) is estimated to be 44min following the TR-55 procedure (Cronshey et al., 1986). The critical design duration approach (Rogger et al., 2012; Yan et al., 2020b) is used here to identify potential peak design flood to make a fair comparison with design flood estimates from DHSVM continuous simulations. More specifically, the critical design duration refers to the duration that produces the largest peak flow. In this study, for each selected average recurrence interval (ARI), such as the 50-year event, we calculate the corresponding flood peaks for 24h, 48h, 72h, and 96h durations. The duration that yields the highest flood peak is identified as the critical design duration and utilized for the assessments.

2.5 NG-IDF Curves vs. DHSVM Design Floods

To estimate design floods from NG-IDF curves, we follow the following steps. First, from DHSVM continuous simulations, we construct the basin mean time series of W with a 15min interval through mass balance as $W = P - \Delta SWE + S$, where P is precipitation, S indicates condensation or evaporation/sublimation of snowpack, and ΔSWE is the change in SWE. We then aggregated the 15min W time series for constructing basin-scale NG-IDF curves at selected durations varying from 1–4 days (Perica et al., 2013). We did not include the subdaily duration because the input precipitation data has no diurnal variability, and DHSVM was calibrated for daily flow and SWE observations (i.e., to reduce uncertainties in estimated NG-IDF curves). For each duration, the annual maximum (water year) W data set was extracted using a moving window approach. As an illustration, when considering a 24h period, a moving window with a size of 96 is utilized to extract 96 sets of 15min data points. The window advances 15min at a time to estimate the maximum 24h W for a given year.

Following the NOAA Atlas 14 (Bonnin et al., 2011), the generalized extreme value (GEV) distribution was fit to the annual maximum W data set based on L-moments statistics (Hosking and Wallis, 1997). Before the frequency analysis, we used the nonparametric Mann-Kendall test

(Mann, 1945; Kendall, 1975) to examine the stationarity assumption (Milly et al., 2008) of the annual maximum W data set. To investigate the independence and stationarity, we additionally utilized the nonparametric Wald-Wolfowitz test (Wald and Wolfowitz, 1943). The NG-IDF curves were then developed for four selected exceedance probabilities: 0.2, 0.1, 0.04, and 0.02, which correspond to extreme events with ARIs of 5, 10, 25, and 50 years. The ARI was cut off at 50-year to reduce uncertainties in NG-IDF curves from extrapolating longer return periods from 28 years of simulations. For design floods estimated from continuous DHSVM simulations, the approach is Similar to the frequency analysis procedure for NG-IDF curve development. We first extracted the DHSVM continuously simulated annual maximum streamflow (also based on water years) and examined the stationarity assumption using the Mann-Kendall test. We then used the same GEV distribution to fit the data set using the L-moments statistics and estimated the benchmark design floods for the ARIs of 5, 10, 25, and 50 years. It is noted that other methods like peaks-over-threshold (Coles, 2001) or r-largest order statistics (Smith, 1986) can expand the dataset and reduce uncertainties in frequency analysis. To eliminate any discrepancies that could arise from dissimilar sample sizes or subjective threshold values in frequency analysis and focus solely on the snow process, we utilized the same GEV distribution for both NG-IDF, PREC-IDF, and flood frequency analysis. Furthermore, our evaluation of floods adhered to standard hydrologic design methods, like NOAA Atlas 14, which employs the GEV distribution. Considering that DHSVM basin-wide continuous flow simulations provide more reliable estimates of design floods, the difference between the two design flood estimates (DHSVM vs. NG-IDF) is a good indication of the limitations underlying the NG-IDF technology, which can include simplified physical hydrologic process, assumption of equal ARI between design storm and design flood, and selection of hyetograph. As an example, many studies (Viglione and Blöschl, 2009; Viglione et al., 2009) have shown that flood ARI can be more or less frequent than the corresponding storm ARI depending on the storm duration, watershed time of concentration, and antecedent moisture condition. As previously mentioned in the Introduction, although the IDF approach has its limitations, it is still sensible to expect that this method will continue to be essential in hydrologic design, particularly for smaller infrastructure projects, in the near future.

Besides a deterministic assessment of the relative difference between the two design flood estimates, we further provide a probabilistic assessment to test if these differences are statistically significant. A Monte Carlo (MC) simulation suggested by Hosking & Wallis (1997) and NOAA Atlas 14 (Bonnin et al., 2011) was used to quantify the sample data uncertainty (i.e., the uncertainty of GEV parameters). A total of 1,000 synthetic ensembles were generated to quantify the GEV parameter uncertainties associated with NG-IDF curves and DHSVM flood frequency analysis. In this study, we used the "lmom" package (version 2.6) (Hosking, 2017) in "R" (version 3.4.3) to perform all L-moments and MC analyses. The Z statistic (Mikkelsen et al., 2005; Madsen et al., 2009; Ganguli and Coulibaly, 2017; Yan et al., 2020b) was used to test the statistically significant differences of the design flood estimates between the NG-IDF technology (q_{ng}) and DHSVM continuous simulation method (q_{dhsvm}):

$$Z = \frac{q_{dhsvm} - q_{ng}}{\sqrt{0.5\left(s_{dhsvm}^2 + s_{ng}^2\right)}} \tag{1}$$

where q_{dhsvm} is the design flood estimated from DHSVM flood frequency analysis, q_{ng} is the design flood obtained from the NG-IDF technology, s_{dhsvm} is the DHSVM design flood standard deviation estimated from the 1,000 DHSVM design flood ensemble, and s_{ng} is the standard deviation of the NG-IDF derived design flood, which is estimated from the 1,000 design flood ensemble generated from running TR-55 using each of the 1,000 NG-IDF ensemble members.

2.6 Hyetograph of Water Available for Runoff

We developed and compared the W hyetographs generated from two approaches. The first approach follows the NOAA Atlas 14 that develops standard hyetographs for rainfall cases only (i.e., assuming all W events are rainfall events), while the second and new approach introduced here develops hyetographs based on the dominant mechanism of W events.

The NOAA Atlas 14 method (Bonnin et al., 2011) was modified from the methodology originally proposed by Huff (1990). We computed W accumulation for specific periods (i.e., 1–4 days) to be consistent with the durations used in the NG-IDF curves. For each selected duration, the following steps were repeated. First, a moving window approach was used to estimate W accumulation over the selected duration based on the 15min DHSVM basin mean W time series. The largest three W accumulations were then obtained for each month over the entire simulation period. Following the NOAA Atlas 14, the 2-year ARI W magnitude was used as the minimum threshold to select large W events for developing hyetographs. Different thresholds were evaluated, including the 25-year ARI, and found that the results were comparable to the 2-year ARI. As a result, the 2-year ARI was selected to generate more samples for the development of probabilistic hyetographs. Each event was then converted into a ratio of the cumulative 15min W to the total W for that duration (i.e., percent of total W), and a ratio of the cumulative time to the total time (i.e., percent of duration). Thus, the last value of the summation ratios is always equal to 100% in the hyetograph. The obtained ensemble large W events were further subdivided into quartiles based on where in the hyetograph (i.e., temporal distribution) most W occurred to provide more specific information on the observed varying hyetographs (Bonnin et al., 2011). For example, the 1st-quartile data consists of events where the greatest percentage of the total W fell during the 1st-quartile of the duration, i.e., the first 6 hours of a 24h period. For the W events classified for each quartile, we can then estimate the cumulative probability of occurrences (i.e., quantiles). For example, the 10% hyetograph curve (90% quantile) indicates that 10% of the corresponding W events have temporal distributions above the curve.

Contrary to the NOAA Atlas 14 hyetograph method that considers rainfall events only, the new approach developed here takes into account the generating mechanism of W events, including rainfall, snowmelt, and ROS (Yan et al., 2018). As will be shown in section 3.2, the hyetograph shapes of W driven by different mechanisms show substantial differences. For example, the temporal distribution of snowmelt-driven W events shows an explicit diurnal cycle associated with the diurnal variability of solar radiation, while the temporal distribution of rainfall events presents a symmetrically peaked, bell-shaped curve. Thus, it is questionable to develop a unified W hyetograph that simply combines the temporal patterns of all large W events generated from different mechanisms. A better approach to developing the W hyetograph is first to identify the dominant W generating mechanism for the study basin, and then generate the W hyetograph for events with the same dominant mechanism. Following Yan et al. (2019b) and Sun et al. (2022a), we classified the daily W time series into three mechanism classes:

1. Rainfall dominated: daily W of at least 10 mm and contains less than 20% snowmelt,
2. Snowmelt dominated: daily W of at least 10 mm and the W contains less than 20% P, and
3. ROS dominated: daily P of at least 10 mm falling on snowpack of at least 10 mm SWE, and W contains at least 20% snowmelt.

2.7 Uncertainty in Design Floods

Besides proposing a hyetograph of W, another goal here is to disentangle and quantify uncertainty contributions of the W sample data and W hyetograph selection on the design flood estimates through the standard TR-55 procedure. The individual uncertainty contribution of an NG-IDF sample set or W hyetograph is estimated using a sequential sampling procedure similar to that used by Schewe et al. (2014) and Samaniego et al. (2017). For example, assume we quantify the W hyetograph uncertainty using 50 ensemble members and NG-IDF sample data uncertainty using 1,000 ensemble members. The component of the NG-IDF sample data uncertainty is characterized by calculating the range of design floods across all 1,000 NG-IDF MC samples separately for each selected W hyetograph, which is then averaged over all 50 W hyetograph ensembles. The component of the W hyetograph uncertainty is estimated in a similar fashion. We first calculate the design flood range across all 50 W hyetograph ensembles for each NG-IDF sample set and then average them over 1,000 NG-IDF samples. The above procedure was applied separately to each selected duration (i.e., 1–4 days) and ARI (i.e., 5–50 years). The range statistic is used here to understand the full range of the dispersion. The statistically significant difference between the two averaged range statics is tested using the aforementioned Z statistic.

3 Results and Discussion

In the following, the analyses performed for NG-IDF technology and DHSVM continuous simulation are reported. We first report the results of the DHSVM evaluation, the development of NG-IDF curves, and DHSVM flood frequency estimations in section 3.1. Second, we discuss the water available for runoff (W) hyetographs and compare the design flood estimates derived from NG-IDF technology to the corresponding DHSVM benchmark in sections 3.2 and 3.3. Last, we disentangle and quantify the uncertainty contributions of the W sample data and W hyetograph to TR-55 design flood estimates in section 3.4.

3.1 DHSVM Evaluation and NG-IDF Curves

The historical records for both streamflow and SWE over water years 1984–2011 were split into two periods of 19 and 9 years in length: 1984–2002 for DHSVM calibration and 2003–2011 for validation. Model calibration was conducted manually by comparing the daily simulations with observations of SWE and streamflow, sequentially. Figure 3 presents the DHSVM model performance in streamflow and SWE simulations for both calibration and validation periods. Except for one year (1997) when the observed streamflow peak was significantly greater than the simulated value at the USGS gauge 10296000, other periods had a very good agreement between the simulations and observations. In the calibration period, statistical comparisons of measured versus simulated daily values resulted in KGE values of 0.84, 0.82, and 0.74, NSE values of 0.73, 0.72, and 0.77, and RMSE values of 6.25 m^3/s, 6.25 m^3/s, and 131 mm for the USGS gauges 10296500, 10296000, and the Sonora SNOTEL site, respectively. The performance of the calibrated model on the validation data set had slightly lower skill, with KGE values of 0.63, 0.82, and 0.76, NSE values of 0.53, 0.68, and 0.78, and RMSE values of 9.28 m^3/s, 7.94 m^3/s, and 141 mm for the USGS gauges 10296500, 10296000, and the Sonora SNOTEL site, respectively. The January 1997 observed peak flow in the Walker River was greater than that of previous and subsequent floods. The January 1997 flood was caused by ROS resulting from unseasonably warm rain in the Sierra Nevada. Accurate simulation of ROS flooding is challenging due to various factors such as rain intensity and amount, prevailing freezing level, and spatial distribution of snow cover. Either of these uncertainties could contribute to bias in predicting the peak flow during ROS events (Fehlmann et al., 2019). Although SNOTEL simulation indicated a good match in SWE simulation in 1997, one SNOTEL site may not represent all mountain ranges, and the SNOTEL network is biased towards specific types of terrain and vegetation (Mote et al., 2016). Interpolated precipitation data at higher mountains may be subject to bias due to limited gauge coverage or gauge undercatch (Groisman and Legates, 1994; Serreze et al., 2001; Lundquist et al., 2019), which could also lead to underestimation of the 1997 flood. We also examined the 28-year annual maximum time series of daily streamflow and W from DHSVM simulations and observations. The mean absolute relative differences for the annual maximum streamflow (AM-S) are 20.2% and 19.0% at the USGS gauges 10296500 and 10296000, for the annual maximum W (AM-W) is 19.6% at the Sonora SNOTEL, respectively. Given the hydrologically challenging context, it is good in practice if errors in estimated flood peaks are within 20% of the value derived from validation data (Calver et al., 2009; Yan et al., 2020b). In summary, model calibration and validation results gave satisfactory and comparable performances on both streamflow and snow simulations. This validated DHSVM model was used as the benchmark for the following assessment of NG-IDF technology.

[Place Figure 3 here]

After extracting the basin mean AM-W time series, we used the nonparametric Mann-Kendall test to examine the stationarity assumption for frequency analysis. For each selected duration (24–96 hours), no statistically significant trend was identified (p-value>5%). The supplementary Wald-Wolfowitz test (p-value>5%) verified the assumption of stationarity and independence for frequency analysis as well. Figure 4 presents the basin-scale NG-IDF curves for the four selected durations varying from 24h to 96h. The associated basin-scale PREC-IDF curves are provided in Figure S1. The shaded areas characterized the 90% confidence intervals associated with the W sample data uncertainty in frequency analysis. It is observed that extreme events with longer ARIs had larger uncertainties; it comes as no surprise because we used 28 years of data to extrapolate the 50-year events. For example, for a 5-year 24h event (i.e., an event with an ARI of 5 years and duration of 24 hours), the range of the 90% confidence intervals was 6.2 mm; while for a 50-year 24h event, the range of 90% confidence intervals was 27.7 mm. Figure 5 presents the DHSVM simulated hydrograph and AM-S time series for the test basin. About 93% of AM-S data occurred between February–May, indicating the snowmelt-dominant flood-generating mechanism for the test basin. Based on the nonparametric Mann-Kendall and Wald-Wolfowitz tests, no statistically significant trend was identified for the AM-S time series. After confirming the stationary assumption, we estimated the DHSVM design flood benchmark in addition to their associated uncertainties for NG-IDF assessment.

[Place Figures 4–5 here]

3.2 Water Available for Runoff Hyetograph

Based on the classification criteria described in section 2.6, we first identified the dominant mechanism for the selected large daily W events. During the 28-year simulation period, we found 125 snowmelt events, 3 rainfall events, and 0 ROS events. For the 3 large rainfall events, we confirmed that their occurrence dates (e.g., July 24, 1998) were far away from their associated AM-S dates (e.g., April 22, 1998), indicating that the large rainfall events did not lead to an AM-

S. Thus, the W hyetograph developed in this study was generated from the snowmelt mechanism only.

Following the procedure used by the NOAA Atlas 14, for each selected duration, we extracted large W accumulations (i.e., > 2-year ARI event) estimated from the basin mean 15min W time series. Any large W accumulation that contained the 3 rainfall events was removed for the following hyetograph development. As a result, a total of 55, 46, 42, and 34 large W events generated from the snowmelt mechanism only were retained for the 24, 48, 72, and 96h duration, respectively. Figure 6 illustrates the ensemble temporal distributions of W events (from snowmelt only), in terms of percent of total W versus percent of duration, at each selected duration. For illustration purposes only, Figure S2 presents the ensemble temporal distributions of W events including the 3 large rainfall events for the 24h duration case. Contrary to the rainfall temporal distribution, it is observed that the snowmelt temporal distribution showed a more rapid rise (i.e., higher intensity) and an explicit diurnal pattern controlled by solar energy input. For instance, at nighttime, there is no change in the W temporal distribution. Among the 3 large rainfall events, the largest percentage of the total rainfall fell during the 1st-quartile of the duration (i.e., first 6 hours of 24h duration) was only about 60% (Figure S2); while for snowmelt events, this value was 100% (i.e., all snowpack melted in the first 6 hours). After acquiring all ensemble temporal distributions of W, we further divided each distribution by quartiles based on where in distribution the most W occurred. Analysis shows that the 24h and 48h events were dominated in the 1st-quartile while the 72h and 96h events were dominated in the 2nd-quartile. About 78% and 50% of the 55 and 46 large W events occurred in the 1st-quartile for the 24h and 48h durations, respectively; and about 43% and 50% of the 42 and 34 large W events occurred in the 2nd-quartile for the 72h and 96h durations, respectively.

[Place Figure 6 here]

Based on the obtained ensemble temporal distributions at each selected duration, we then develop a probabilistic hyetograph of W by estimating the exceedance probabilities of W occurrence (i.e., quantiles) at each time step. Contrary to the probabilistic rainfall hyetograph developed in the NOAA Atlas 14 that combines all large rainfall cases, the unique diurnal pattern associated with the snowmelt-dominant W hyetograph requires special attention. Taking a 24h duration event for example, if all obtained W temporal distributions are divided evenly into the 1st-quartile and 4th-quartile (e.g., the horizontal lines occur evenly at the 10% or 90% y-axis value), the estimated median (50% exceedance probability) curve will be close to the uniform distribution over the 24h duration (e.g., a 1:1 line), which substantially underestimates the snowmelt intensity. To address this issue, we developed the probabilistic W hyetograph based only on the large W events that occurred in the dominant quartile, similar to the alternative quartile-based rainfall hyetograph in the NOAA Atlas 14.

Figure 7 presents the probabilistic hyetographs of W for the test basin over the four selected durations. The graph represents the cumulative probability of occurrence at 20% increments and a moving window smoothing technique was performed on each curve (Bonnin et al., 2011). For the 24h and 48h durations, the probabilistic hyetographs were developed using ensemble W events where most W occurred in the 1st-quartile; for the 72h and 96h durations, they were developed using the W events where most W occurred in the 2nd-quartile. For each duration, the 10% hyetograph curve indicates that 10% of the corresponding W events had temporal distributions that fell above the curve (i.e., 10% exceedance probability); the 50% curve represents the median temporal distribution. The broad range between these curves represents the broad uncertainties associated with the W hyetograph in design flood estimates. In this study, we tested the median (50% curve) hyetograph and an optimized hyetograph method in the NG-IDF modeling. In the optimized hyetograph method, all curves varying from 10% to 90% (at a 10% increment) were used in the NG-IDF modeling and the best results (i.e., closest to the DHSVM benchmark) were retained. The median hyetograph was used to test whether a hyetograph under the "average" condition can lead to acceptable results; the optimized hyetograph method was used to evaluate what are the best results we can achieve with the use of W hyetographs in the NG-IDF modeling. Note that our objective is not to overfit the model in order to make NG-IDF estimation align with DHSVM benchmark through optimized hyetograph selection. Rather, our aim is to measure the level of uncertainty in hyetograph selection during flood estimation and to emphasize the significance of taking hyetograph uncertainty into account when conducting flood risk analysis, as detailed in section 3.4.

[Place Figure 7 here]

3.3 NG-IDF Modeling

Before using the developed snowmelt W hyetographs, we first tested the standard rainfall hyetographs in the NG-IDF modeling for comparisons. In our previous studies, we extensively investigated the difference between PREC-IDF and NG-IDF curves (Yan et al., 2018, 2019a, 2020b). However, in this study, our focus is solely on NG-IDF modeling. Specifically, we analyze NG-IDF modeling using a rainfall hyetograph in comparison to a developed snowmelt hyetograph. Figure 8 compared the developed snowmelt W hyetograph (using 24h median hyetograph as a proof-of-concept) against five standard rainfall hyetographs: uniform hyetograph proposed in the rational method and four types of NRCS rainfall hyetographs widely used over the U.S. Results suggest that all rainfall hyetographs substantially underestimate the W intensity in the snow-dominated test basin. In the following comparisons, we used two rainfall hyetographs − the uniform hyetograph and the NRCS Type IA hyetograph that is recommended for hydrologic design in the Sierra Nevada and Cascade mountains (McCuen, 1998). The uniform hyetograph is included because only daily precipitation data are available, and we uniformly disaggregated the precipitation data over 24 hrs.

[Place Figure 8 here]

Figure 9a compares the design flood estimates from the DHSVM continuous simulations (q_{dhsvm}) and NG-IDF modeling (q_{ng}) with the use of uniform hyetograph and NRCS Type IA hyetograph, respectively. Note that we used the critical design duration approach to identify potential peak design flood in the NG-IDF technique. In addition to the deterministic estimates, sample uncertainties associated with DHSVM flood and NG-IDF frequency analysis were also quantified and shown as the 90% confidence intervals (i.e., error bars) in Figure 9a. As described in section 2.5, the Z statistic was used to test the statistically significant differences in the design flood estimates between the two methods. The associated p-values of the pairwise comparison between DHSVM and NG-IDF estimates were also shown in Figure 9a. It is observed that compared to q_{dhsvm}, q_{ng} with the use of uniform and NRCS hyetographs both showed statistically significant underestimations of design floods (p-value < 1%) for all four ARIs, even though the wet AMC and critical design duration were used to reduce the potential underestimation of the design flood. In sum, these results suggested that standard rainfall hyetographs can lead to a substantial underestimate of flood risk and it is necessary to develop new hyetographs to enhance NG-IDF performance. Note that the larger confidence interval associated with the NG-IDF method are due to the TR-55 nonlinear rainfall-runoff process used to convert W magnitude into flood magnitude, whereas DHSVM directly generates design flood magnitude through frequency analysis of annual maximum flood. Similar results are found in Yan et al. (2019a).

[Place Figure 9 here]

Figure 9b compares the design flood estimates from q_{dhsvm} and q_{ng} with the use of median and optimized W hyetographs, respectively. To select the optimized hyetographs, various curves ranging from 10% to 90% were tested (with a 10% increment) in the NG-IDF modeling. The optimized hyetographs were then chosen based on their similarity to the DHSVM benchmark, with the best-fit curve being selected. Similar to Figure 9a, the error bars represent the 90% confidence intervals of design flood estimates associated with the sample data uncertainties in the NG-IDF curves; the p-values of the pairwise comparison indicate the Z statistics. It is observed that compared to the results with the use of standard rainfall hyetographs (Figure 9a), the use of either median or optimized W hyetographs substantially improved the design flood estimates. When using the median W hyetograph, the absolute relative errors between q_{dhsvm} and q_{ng} were about 23%, 8%, 9%, and 20% for the 5-, 10-, 25-, and 50-year events, respectively. The absolute relative errors were further reduced to about 5%, 1%, 2%, and 0.5% for the 5-, 10-, 25-, and 50-year events with the use of optimized W hyetograph, respectively. When considering the sample uncertainties in the NG-IDF curves, we found statistically insignificant differences (i.e., p-value >

5%) between the q_{dhsvm} and q_{ng} for all four selected ARIs with the use of median and optimized W hyetographs, respectively. These results suggested that W hyetographs offer good performance for NG-IDF technology in hydrologic design and the median W hyetograph may be appropriate. However, the broad range in the W hyetograph shown in Figure 7 can result in a broad range of design peak flow or volume estimates, leading to large uncertainties associated with the W hyetograph selection. We argue that the probabilistic W hyetograph should be used in a way that reflects the goals of the user, and no single W hyetograph (e.g., median curve) works the best under all design conditions. A practical path forward is to quantify the uncertainty contribution of W hyetograph selection in NG-IDF design flood estimates and then perform a design risk analysis that depends on the risk tolerance of a particular asset or project. To illustrate, a wide variety of design floods or volumes can be estimated. To minimize the risk of failure in the design of critical infrastructure, it is advisable for users to focus on temporal distributions that are more likely to produce higher peaks instead of median cases. Additionally, a decision-analytic cost-loss-ratio model (Murphy, 1977) can be utilized to determine the total cost associated with a specific hyetograph. Furthermore, users should evaluate whether utilizing results from one of the quartiles instead of all samples would yield more suitable outcomes for their particular circumstances.

It should be noted that when conducting frequency analysis on annual maximum W and streamflow, an important assumption is made that all samples are stationary, independent, and identically distributed. In the case of the small basin examined in this study, this assumption holds true because all samples are generated from the snowmelt process and no statistically significant trends have been observed. However, for larger basins that involve multiple flood-generation processes, such as flood mixtures, additional attention is required when estimating flood quantiles. Previous studies (Murphy, 2001; Barth et al., 2019) have highlighted the need to account for such flood mixtures. Furthermore, with the expected impact of climate change, flood-generation processes may undergo changes, such as a decrease in ROS events at lower elevations and an increase at higher elevations (Musselman et al., 2018; Li et al., 2019). Consequently, it is crucial to employ a physically informed approach for flood frequency analysis, considering both the historical period and future projections. For instance, Barth et al. (2017) employed the mechanism of atmospheric rivers to partition annual peak flows in the western U.S. They estimated flood quantiles by considering a mixed population. Additionally, Yu et al. (2022) demonstrated that neglecting mixture effects can lead to substantial uncertainties in estimating the magnitudes of extreme flood statistics. As explained in section 2.6, the classification of W into rainfall, snowmelt, and ROS events can assist in partitioning flood peaks and conducting flood frequency analysis using mixed flood populations.

3.4 Sample and Hyetograph Uncertainty Contributions on Design Floods

In this NG-IDF technology assessment, both methods used simulations from validated DHSVM, and therefore the two major uncertainty sources associated with the NG-IDF technology assessment were the data sample uncertainty in NG-IDF frequency analysis and W hyetograph uncertainty in the TR-55 modeling. Following the procedure described in section 2.7, we quantified the uncertainty contribution of each component to the mean range of design flood estimates using the range statistic in Figure 10.

[Place Figure 10 here]

Using the standard Z test, the differences between the two mean ranges were all statistically significant (i.e., p-value < 5%) except for the 72h, 5-year event as shown in Figure 10. The magnitudes of both mean ranges tend to increase with ARI, suggesting larger uncertainty associated with less frequent extreme events. For events with lower ARIs such as 5-year or 10-year, both the data sample and W hyetograph uncertainties are important to the design flood uncertainties (relative difference <50%); for events with higher ARIs such as 50-year, the data sample uncertainty dominates (relative difference >100%), due to extrapolating to large ARIs using short- to modest-length records (i.e., 28 years). This is mainly because the shape parameter of extreme value distribution determines the nature of the tail of the distribution and its value has significant impacts on the severity of large ARI events. Unfortunately, no matter what estimation method is selected, the shape parameter is difficult to estimate and sensitive to extreme events, especially given short records (Cooley et al., 2007; Cooley and Sain, 2010). These findings are also consistent with the latest federal flood frequency guideline Bulletin 17C (England et al., 2018), which recommended using regional information to reduce data sample uncertainty, particularly when records are short (less than 30 years).

It is worth noting that the contribution of the W hyetograph uncertainty, although smaller than the data sample uncertainty at larger ARIs, cannot be neglected, considering that the associated mean range of design flood estimates exceeded the absolute value of the DHSVM benchmark as shown in Figure 9. For example, for the 24h, 50-year event, the mean range of design flood contributed from the W hyetograph uncertainty was 4.66 m^3/s, exceeding the absolute value of the DHSVM benchmark of 3.74 m^3/s. Especially for small infrastructures such as drainage system design, many local/federal surface water design manuals (UFC, 2013; SCDM, 2016; SWMMEW, 2019) recommend ARIs of 5-year to 25-year. The sample data uncertainty versus the

W hyetograph uncertainty is actually the *epistemic uncertainty* versus *aleatory uncertainty* which involves a lack of knowledge about the response data or random natural variability (Beven and Smith, 2015). The *epistemic uncertainty* can be suitably resolved by improving the model while the *aleatory uncertainty* cannot be diminished. For example, we can use the peaks-over-threshold method, the regionalization method, or extend the simulation period to reduce the sample data uncertainty in NG-IDF curves; however, the W hyetograph uncertainty cannot be reduced because of the natural variability. These results support that we must consider W hyetograph uncertainty in NG-IDF design and nonstationary changes of W hyetograph in future flood projections, and the developed W hyetograph can provide information using a probabilistic approach to risk assessment that is adjustable based on design cost and changing operating conditions. As mentioned in the Introduction section, the IDF design approach necessitates not only a storm magnitude but also a predetermined hyetograph, with the shape of the hyetograph differing depending on the underlying mechanism, such as the standard rainfall hyetograph versus the snowmelt hyetograph illustrated in Figure 8. Numerous studies (Cheng and AghaKouchak, 2014; Ragno et al., 2018; Hou et al., 2019; Schlef et al., 2023) have examined the potential changes in IDF curves resulting from global warming. However, climate change will cause different regions to experience varying shifts in their dominant flood-generating mechanisms, with higher mountains potentially experiencing more ROS events.

4 Conclusions

In this study, we assessed the performance of the recently developed NG-IDF technology in design flood estimates for hydrologic design at a snow-dominated small basin located in the eastern Sierra Nevada Mountains, California where the snowmelt flows through the U.S. DoD's MCMWTC.

Based on evaluations of NG-IDF technology, we proposed a new probabilistic W hyetograph that explicitly represents the temporal patterns of snowmelt and quantified the contribution of W hyetograph uncertainty in design flood estimates.

Based on the results of this study, we have four major conclusions:

1) The standard rainfall hyetographs such as uniform distribution proposed in the rational method, or NRCS temporal distributions lead to substantial underestimation of design floods and therefore are inappropriate for the small snow-dominated basin in the Sierra Nevada. There is an emerging need to systematically develop W hyetograph over the CONUS.

2) The median W hyetograph generates acceptable flood estimates but the probabilistic W hyetograph represents a broad range of uncertainty or *aleatory uncertainty* which is caused by random natural variability and cannot be diminished, resulting in a broad range of variability in design peak flow or volume estimates. A full risk analysis that includes W hyetograph uncertainty is recommended for risk-based hydrologic design and future climate change impact studies on flood risk (e.g., changes in hyetograph versus changes in storm magnitude).

3) Instead of simply "pooling" all ensembles of large W events together in developing a probabilistic W hyetograph, it is important to investigate the underlying mechanism (e.g., rainfall, snowmelt, ROS) to gain a physical understanding of their behaviors. For example, the test basin in this study was dominated by the snowmelt mechanism and therefore the W hyetograph presented an explicit diurnal pattern controlled by solar energy input.

4) Sample data uncertainty in the NG-IDF frequency analysis and W hyetograph uncertainty in the TR-55 modeling equally contribute to the design flood uncertainties at smaller ARIs such as 5-year event (which is commonly used in hydrologic design such as culverts); sample data uncertainty dominates the W hyetograph uncertainty at larger ARIs such as 50-year event.

The sample data uncertainty or *epistemic uncertainty* can be suitably resolved by improving the model such as the regionalization method or extending the DHSVM simulation period. Nevertheless, it should be acknowledged that these findings are derived from a small snow-dominated basin in the Sierra Nevada, and as such, the outcomes may differ for larger snow-dominated areas with varying basin sizes, climates, and vegetation.

Despite the promising results with the use of the developed W hyetograph, we acknowledge that further assessments are still necessary to make NG-IDF technology ready for practicing design. The conclusion reached in this study pertains solely to the small study basin in the Sierra Nevada Mountains. Obviously, more case studies are necessary to develop regional W hyetographs generated from different physical mechanisms and validate them in different hydroclimate regions. The characteristics of a basin, including its size, topography, climate, and vegetation type, can have an impact on the shape of the hyetograph and the estimation of design flood. For instance, in a large basin with rainfall- and snow-dominated regions, the standard rainfall hyetograph may be suitable if summer thunderstorms are the primary cause of flooding instead of spring snowmelt (Gochis et al., 2015). However, at lower elevations where snow has historically dominated, the standard rainfall hyetograph may become applicable in the future due to a shift in precipitation phase from snowfall to rain. Conversely, at higher elevations, the snowmelt hyetograph may currently be effective but could underestimate design flood estimation in the future due to an increase in the frequency of ROS events under climate change (Musselman et al., 2018). Additionally, the vegetation type in a basin can also affect the hyetograph shape, with evergreen forests having a higher occurrence of ROS events than open spaces (Mooney and Lee, 2022), and postfire lands potentially accelerating the rate of snowmelt (Gleason et al., 2013). Next, we plan to use the well-validated DHSVM model with the developed regionally coherent snow parameters (Sun et al., 2019) to construct regional W hyetographs over the CONUS and test them against different hillslope configurations and selected validation basins using the general method presented in this study.

Open Research

The DHSVM source code is available at (Perkins et al., 2023).

The BCQC SNOTEL data used in this study are available at https://climate.pnnl.gov/?category=Hydrology&card=2520c45b90be6207c5b68aef3817da52

The USGS and the National Geospatial-Intelligence Agency Global Multi-resolution Terrain Elevation Data 2010 are available at (Danielson and Gesch, 2011).

The Livneh meteorological forcing data are available at (Livneh et al., 2015).

The National Land Cover Database is available at https://www.mrlc.gov/data?f%5B0%5D=category%3Aland%20cover&f%5B1%5D=region%3Aconus

The NRCS Soil Survey SSURGO soil data are available at (Soil Survey Staff, 2022).

The USGS streamflow observations data for gauge 10296500 are available at https://waterdata.usgs.gov/nwis/dv?cb_00060=on&format=rdb&site_no=10296500&legacy=&referred_module=sw&period=&begin_date=1984-01-01&end_date=2013-12-31

The USGS streamflow observations data for gauge 10296000 are available at https://waterdata.usgs.gov/nwis/dv?cb_00060=on&format=rdb&site_no=10296000&legacy=&referred_module=sw&period=&begin_date=1984-01-01&end_date=2013-12-31

Acknowledgments

The authors declare that they have no known competing financial interests. This material is based upon work supported by the Strategic Environmental Research and Development Program under Contract No. RC-2546 and Environmental Security Technology Certification Program under Contract No. EW21-5140. Battelle Memorial Institute operates the Pacific Northwest National Laboratory (PNNL) for the U.S. Department of Energy under contract DE-AC06-76RLO-1830. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 using NERSC award BER-ERCAP0023966.

References

ASCE, 2018. Climate-Resilient Infrastructure: Adaptive Design and Risk Management. Committee on Adaptation to a Changing Climate. Manual of Practice (MOP) 140, American Society of Civil Engineers, Reston, VA. https://doi.org/10.1061/9780784415191

Bales, R.C., Molotch, N.P., Painter, T.H., Dettinger, M.D., Rice, R., Dozier, J., 2006. Mountain hydrology of the western United States. Water Resour. Res. 42, W08432. https://doi.org/10.1029/2005WR004387

Barth, N.A., Villarini, G., Nayak, M.A., White, K., 2017. Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. Water Resour. Res. 53, 257–269. https://doi.org/10.1002/2016WR019064

Barth, N.A., Villarini, G., White, K., 2019. Accounting for Mixed Populations in Flood Frequency Analysis: Bulletin 17C Perspective. J. Hydrol. Eng. 24. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001762

Beven, K., Smith, P., 2015. Concepts of Information Content and Likelihood in Parameter Calibration for Hydrological Simulation Models. J. Hydrol. Eng. 20, A4014010. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000991

Bonnin, G.M., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D., 2011. Precipitation-Frequency Atlas of the United States, NOAA Atlas 14. vol. 1, version 5.0, U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

Calver, A., Stewart, E., Goodsell, G., 2009. Comparative analysis of statistical and catchment modelling approaches to river flood frequency estimation. J. Flood Risk Manag. 2, 24–31. https://doi.org/10.1111/j.1753-318X.2009.01018.x

Cao, Q., Sun, N., Yearsley, J., Nijssen, B., Lettenmaier, D.P., 2016. Climate and land cover effects on the temperature of Puget Sound streams. Hydrol. Process. 30, 2286–2304. https://doi.org/10.1002/hyp.10784

Cheng, L., AghaKouchak, A., 2014. Nonstationary precipitation Intensity-Duration-Frequency curves for infrastructure design in a changing climate. Sci. Rep. 4, 7093.

https://doi.org/10.1038/srep07093

Cho, E., Jacobs, J.M., 2020. Extreme Value Snow Water Equivalent and Snowmelt for Infrastructure Design Over the Contiguous United States. Water Resour. Res. 56. https://doi.org/10.1029/2020WR028126

Cho, E., McCrary, R.R., Jacobs, J.M., 2021. Future Changes in Snowpack, Snowmelt, and Runoff Potential Extremes Over North America. Geophys. Res. Lett. 48, e2021GL094985. https://doi.org/10.1029/2021GL094985

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

Clark, M.P., Vogel, R.M., Lamontagne, J.R., Mizukami, N., Knoben, W.J.M., Tang, G., Gharari, S., Freer, J.E., Whitfield, P.H., Shook, K.R., Papalexiou, S.M., 2021. The Abuse of Popular Performance Metrics in Hydrologic Modeling. Water Resour. Res. 57, e2020WR029001. https://doi.org/10.1029/2020WR029001

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values, Springer. London.

Cooley, D., Nychka, D., Naveau, P., 2007. Bayesian Spatial Modeling of Extreme Precipitation Return Levels. J. Am. Stat. Assoc. 102, 824–840. https://doi.org/10.1198/016214506000000780

Cooley, D., Sain, S.R., 2010. Spatial Hierarchical Modeling of Precipitation Extremes From a Regional Climate Model. J. Agric. Biol. Environ. Stat. 15, 381–402. https://doi.org/10.1007/s13253-010-0023-9

Cronshey, R., McCuen, R.H., Miller, N., Rawls, W., Robbins, S., Woodward, D., 1986. Urban Hydrology for Small Watersheds—TR-55. U.S. Department of Agriculture, Washington, D.C.

Curran, J., Barth, N., Veilleux, A.G., Ourso, R., 2016. Estimating flood magnitude and frequency at gaged and ungaged sites on streams in Alaska and conterminous basins in Canada, based on data through water year 2012. U.S. Geological Survey Scientific Investigations Report 2016–5024. https://doi.org/10.3133/sir20165024

Danielson, J.J., Gesch, D.B., 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010). [Dataset]. https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-multi-resolution-terrain-elevation.

England, J.F.J., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O.J., Veilleux, A.G., Kiang, J.E., Mason, R.R.J., 2018. Guidelines for determining flood flow frequency—Bulletin 17C. U.S. Geological Survey Techniques and Methods, book 4, chap. B5, 148 p. https://doi.org/10.3133/tm4B5

ESTCP, 2018. Nonstationary Weather Patterns and Extreme Events: Informing Design and Planning for Long-Lived Infrastructure. Project RC-201591. Retrieved from https://www.serdp-estcp.org/News-and-Events/Blog/Nonstationary-Weather-Patterns-and-Extreme-Events-Workshop-Report.

Fehlmann, M., Gascón, E., Rohrer, M., Schwarb, M., Stoffel, M., 2019. Improving Medium-Range Forecasts of Rain-on-Snow Events in Prealpine Areas. Water Resour. Res. 55, 7638–7661. https://doi.org/10.1029/2018WR024644

Fullerton, A.H., Sun, N., Baerwalde, M.J., Hawkins, B.L., Yan, H., 2022. Mechanistic Simulations Suggest Riparian Restoration Can Partly Counteract Climate Impacts to Juvenile Salmon. JAWRA J. Am. Water Resour. Assoc. 58, 525–546. https://doi.org/10.1111/1752-1688.13011

Ganguli, P., Coulibaly, P., 2017. Does nonstationarity in rainfall require nonstationary intensity--duration--frequency curves? Hydrol. Earth Syst. Sci. 21, 6461–6483. https://doi.org/10.5194/hess-21-6461-2017

Gleason, K.E., Nolin, A.W., Roth, T.R., 2013. Charred forests increase snowmelt: Effects of burned woody debris and incoming solar radiation on snow ablation. Geophys. Res. Lett. 40, 4654–4661. https://doi.org/10.1002/grl.50896

Gochis, D., Schumacher, R., Friedrich, K., Doesken, N., Kelsch, M., Sun, J., Ikeda, K., Lindsey, D., Wood, A., Dolan, B., Matrosov, S., Newman, A., Mahoney, K., Rutledge, S., Johnson, R., Kucera, P., Kennedy, P., Sempere-Torres, D., Steiner, M., Roberts, R., Wilson, J., Yu, W., Chandrasekar, V., Rasmussen, R., Anderson, A., Brown, B., 2015. The Great Colorado Flood of September 2013. Bull. Am. Meteorol. Soc. 96, 1461–1487. https://doi.org/10.1175/BAMS-D-13-00241.1

Groisman, P.Y., Legates, D.R., 1994. The Accuracy of United States Precipitation Data. Bull. Am. Meteorol. Soc. 75, 215–227. https://doi.org/10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO;2

Gupta, H. V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hamlet, A.F., 2018. New Observed Data Sets for the Validation of Hydrology and Land Surface Models in Cold Climates. Water Resour. Res. 54, 5190–5197. https://doi.org/10.1029/2018WR023123

Hatchett, B.J., Boyle, D.P., Garner, C.B., Kaplan, M.L., Putnam, A.E., Bassett, S.D., 2016. Magnitude and frequency of wet years under a megadrought climate in the western Great Basin, USA. Quat. Sci. Rev. 152, 197–202. https://doi.org/10.1016/j.quascirev.2016.09.017

Hettiarachchi, S., Wasko, C., Sharma, A., 2018. Increase in flood risk resulting from climate change in a developed urban watershed – the role of storm temporal patterns. Hydrol. Earth Syst. Sci. 22, 2041–2056. https://doi.org/10.5194/hess-22-2041-2018

Hosking, J.R.M., 2017. Package “lmom.” https://CRAN.R-project.org/package=lmom.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, Cambridge, U. K.

Hou, Z., Ren, H., Sun, N., Wigmosta, M.S., Liu, Y., Leung, L.R., Yan, H., Skaggs, R., Coleman, A., 2019. Incorporating climate nonstationarity and snowmelt processes in intensity–duration–frequency analyses with case studies in mountainous areas. J. Hydrometeorol. 20. https://doi.org/10.1175/JHM-D-19-0055.1

Huff, F.A., 1990. Time Distributions of Heavy Rainstorms in Illinois. Illinois State Water Survey, Champaign, 173, 17 pp.

Hungerford, R.D., Nemani, R.R., Running, S.W., Coughlan, J.C., 1989. MTCLIM: A Mountain Microclimate Simulation Model. U.S. Department of Agriculture, Ogden, Utah, Utah.

Jencso, K.G., McGlynn, B.L., Gooseff, M.N., Wondzell, S.M., Bencala, K.E., Marshall, L.A., 2009. Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale. Water Resour. Res. 45, W04428. https://doi.org/10.1029/2008WR007225

Kendall, M.G., 1975. Rank Correlation Methods. Griffin, London.

Klos, P.Z., Link, T.E., Abatzoglou, J.T., 2014. Extent of the rain-snow transition zone in the western U.S. under historic and projected climate. Geophys. Res. Lett. 41, 4560–4568. https://doi.org/10.1002/2014GL060500

Knoben, W.J.M., Freer, J.E., Woods, R.A., 2019. Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23, 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Li, D., Lettenmaier, D.P., Margulis, S.A., Andreadis, K., 2019. The Role of Rain-on-Snow in Flooding Over the Conterminous United States. Water Resour. Res. 55, 8492–8513. https://doi.org/10.1029/2019WR024950

Livneh, B., Bohn, T.J., Pierce, D.W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D.R., Brekke, L., 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013 [Dataset]. Sci. Data 2, 150042. https://doi.org/10.1038/sdata.2015.42

Livneh, B., Rosenberg, E.A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K.M., Maurer, E.P., Lettenmaier, D.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P., Nijssen, B., 2013. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. J. Clim. 26, 9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1

Lundquist, J., Hughes, M., Gutmann, E., Kapnick, S., 2019. Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. Bull. Am. Meteorol. Soc. 100, 2473–2490. https://doi.org/10.1175/BAMS-D-19-0001.1

Lundquist, J.D., Roche, J.W., Forrester, H., Moore, C., Keenan, E., Perry, G., Cristea, N., Henn, B., Lapo, K., McGurk, B., Cayan, D.R., Dettinger, M.D., 2016. Yosemite Hydroclimate Network: Distributed stream and atmospheric data for the Tuolumne River watershed and surroundings. Water Resour. Res. 52, 7478–7489. https://doi.org/10.1002/2016WR019261

Madsen, H., Arnbjerg-Nielsen, K., Mikkelsen, P.S., 2009. Update of regional intensity–duration–frequency curves in Denmark: Tendency towards increased storm intensities. Atmos. Res. 92, 343–349. https://doi.org/https://doi.org/10.1016/j.atmosres.2009.01.013

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica 13, 245. https://doi.org/10.2307/1907187

McCuen, R.H., 1998. Hydrologic analysis and design, 2nd ed. Prentice-Hall, Upper Saddle River, N.J.

Mikkelsen, P.S., Madsen, H., Arnbjerg-Nielsen, K., Rosbjerg, D., Harremoës, P., 2005. Selection of regional historical rainfall time series as input to urban drainage simulations at ungauged locations. Atmos. Res. 77, 4–17. https://doi.org/https://doi.org/10.1016/j.atmosres.2004.10.016

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 (80-.). 319, 573–574. https://doi.org/10.1126/science.1151915

Mizukami, N., Rakovec, O., Newman, A.J., Clark, M.P., Wood, A.W., Gupta, H. V., Kumar, R., 2019. On the choice of calibration metrics for “high-flow” estimation using hydrologic models. Hydrol. Earth Syst. Sci. 23, 2601–2614. https://doi.org/10.5194/hess-23-2601-2019

Mockus, V., 1957. Use of storm and watersheds characteristics in synthetic hydrograph analysis and application. U.S. Department of Agriculture, Soil Conservation Service.

Mooney, P.A., Lee, H., 2022. Afforestation affects Rain-On-Snow climatology over Norway. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/ac6684

Mote, P.W., Rupp, D.E., Li, S., Sharp, D.J., Otto, F., Uhe, P.F., Xiao, M., Lettenmaier, D.P., Cullen, H., Allen, M.R., 2016. Perspectives on the causes of exceptionally low 2015 snowpack in the western United States. Geophys. Res. Lett. 43, 10980–10988. https://doi.org/10.1002/2016GL069965

Murphy, A.H., 1977. The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation. Mon. Weather Rev. 105, 803–816. https://doi.org/10.1175/1520-

0493(1977)105<0803:TVOCCA>2.0.CO;2

Murphy, P.J., 2001. Evaluation of Mixed-Population Flood-Frequency Analysis. J. Hydrol. Eng. 6, 62–70. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:1(62)

Musselman, K.N., Clark, M.P., Liu, C., Ikeda, K., Rasmussen, R., 2017. Slower snowmelt in a warmer world. Nat. Clim. Chang. 7, 214–219. https://doi.org/10.1038/nclimate3225

Musselman, K.N., Lehner, F., Ikeda, K., Clark, M.P., Prein, A.F., Liu, C., Barlage, M., Rasmussen, R., 2018. Projected increases and shifts in rain-on-snow flood risk over western North America. Nat. Clim. Chang. 8, 808–812. https://doi.org/10.1038/s41558-018-0236-4

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Perica, S., Martin, D., Pavlovic, S., Roy, I., Laurent, M. St., Trypaluk, C., Unruh, D., Yekta, M., Bonnin, G., 2013. Precipitation-Frequency Atlas of the United States, NOAA Atlas 14. vol. 8, version 2.0, U.S. Dep. of Commer., National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD.

Perkins, W.A., Duan, Z., Sun, N., Wigmosta, M.S., Richmond, M.C., Chen, X., Leung, L.R., 2019. Parallel Distributed Hydrology Soil Vegetation Model (DHSVM) using global arrays. Environ. Model. Softw. 122, 104533. https://doi.org/10.1016/j.envsoft.2019.104533

Perkins, W.A., Sun, N., Duan, Z., 2023. Lizzy0Sun/DHSVM: DHSVM v3.3 (v3.3). Zenodo. https://doi.org/10.5281/zenodo.8339549.

Ponce, V.M., Hawkins, R.H., 1996. Runoff Curve Number: Has It Reached Maturity? J. Hydrol. Eng. 1, 11–19. https://doi.org/10.1061/(ASCE)1084-0699(1996)1:1(11)

Prata, A.J., 1996. A new long-wave formula for estimating downward clear-sky radiation at the surface. Q. J. R. Meteorol. Soc. 122, 1127–1151. https://doi.org/10.1002/qj.49712253306

Ragno, E., AghaKouchak, A., Love, C.A., Cheng, L., Vahedifard, F., Lima, C.H.R., 2018. Quantifying Changes in Future Intensity-Duration-Frequency Curves Using Multimodel Ensemble Simulations. Water Resour. Res. 54, 1751–1764. https://doi.org/10.1002/2017WR021975

Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., Merz, R., Blöschl, G., 2012. Runoff models and flood frequency statistics for design flood estimation in Austria – Do they tell a consistent story? J. Hydrol. 456–457, 30–43. https://doi.org/10.1016/j.jhydrol.2012.05.068

Rossman, L.A., 2004. Storm water management model user manual. Version 5. U.S. Environmental Protection Agency.

Samaniego, L., Kumar, R., Breuer, L., Chamorro, A., Flörke, M., Pechlivanidis, I.G., Schäfer, D., Shah, H., Vetter, T., Wortmann, M., Zeng, X., 2017. Propagation of forcing and model uncertainties on to hydrological drought characteristics in a multi-model century-long experiment in large river basins. Clim. Change 141, 435–449. https://doi.org/10.1007/s10584-016-1778-y

SCDM, 2016. Snohomish County Drainage Manual Volume III Hydrologic Analysis and Flow Control BMPs. Available online at: https://snohomishcountywa.gov/1130/Drainage-Manual.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N.W., Clark, D.B., Dankers, R., Eisner, S., Fekete, B.M., Colón-González, F.J., Gosling, S.N., Kim, H., Liu, X., Masaki, Y., Portmann, F.T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., Kabat, P., 2014. Multimodel assessment of water scarcity under climate change. Proc. Natl. Acad. Sci. 111, 3245–3250.

https://doi.org/10.1073/pnas.1222460110
Schlef, K.E., Kunkel, K.E., Brown, C., Demissie, Y., Lettenmaier, D.P., Wagner, A., Wigmosta, M.S., Karl, T.R., Easterling, D.R., Wang, K.J., François, B., Yan, E., 2023. Incorporating non-stationarity from climate change into rainfall frequency and intensity-duration-frequency (IDF) curves. J. Hydrol. 616, 128757. https://doi.org/10.1016/j.jhydrol.2022.128757
Serreze, M.C., Clark, M.P., Frei, A., 2001. Characteristics of large snowfall events in the montane western United States as examined using snowpack telemetry (SNOTEL) data. Water Resour. Res. 37, 675–688. https://doi.org/10.1029/2000WR900307
Smith, R.L., 1986. Extreme value theory based on the r largest annual events. J. Hydrol. 86, 27–43. https://doi.org/10.1016/0022-1694(86)90004-1
Soil Survey Staff, 2022. Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. [Dataset]. https://www.nrcs.usda.gov/resources/data-and-reports/soil-survey-geographic-database-ssurgo. Accessed [1/10/2022].
Sun, N., Wigmosta, M., Zhou, T., Lundquist, J., Dickerson-Lange, S., Cristea, N., 2018. Evaluating the functionality and streamflow impacts of explicitly modelling forest-snow interactions and canopy gaps in a distributed hydrologic model. Hydrol. Process. 32, 2128–2140. https://doi.org/10.1002/hyp.13150
Sun, N., Yan, H., Wigmosta, M.S., Coleman, A.M., Leung, L.R., Hou, Z., 2022a. Datasets for characterizing extreme events relevant to hydrologic design over the conterminous United States. Sci. Data 9, 154. https://doi.org/10.1038/s41597-022-01221-9
Sun, N., Yan, H., Wigmosta, M.S., Leung, L.R., Skaggs, R., Hou, Z., 2019. Regional Snow Parameters Estimation for Large-Domain Hydrological Applications in the Western United States. J. Geophys. Res. Atmos. 124, 5296–5313. https://doi.org/10.1029/2018JD030140
Sun, N., Yan, H., Wigmosta, M.S., Lundquist, J., Dickerson-Lange, S., Zhou, T., 2022b. Forest Canopy Density Effects on Snowpack Across the Climate Gradients of the Western United States Mountain Ranges. Water Resour. Res. 58, e2020WR029194. https://doi.org/10.1029/2020WR029194
SWMMEW, 2019. Stormwater Management Manual for Eastern Washington (SWMMEW). Washington State Department of Ecology.
Thornton, P.E., Running, S.W., 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agric. For. Meteorol. 93, 211–228. https://doi.org/10.1016/S0168-1923(98)00126-9
UFC, 2013. Surface Drainage Design (150/5320-5D). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration. Available online at: https://www.faa.gov/documentLibrary/media/Advisory_Circular/150_5320_5d.pdf.
USGCRP, 2018. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II. Washington, DC. https://doi.org/10.7930/NCA4.2018
Viglione, A., Blöschl, G., 2009. On the role of storm duration in the mapping of rainfall to flood return periods. Hydrol. Earth Syst. Sci. 13, 205–216. https://doi.org/10.5194/hess-13-205-2009
Viglione, A., Merz, R., Blöschl, G., 2009. On the role of the runoff coefficient in the mapping of rainfall to flood return periods. Hydrol. Earth Syst. Sci. 13, 577–593. https://doi.org/10.5194/hess-13-577-2009
Wald, A., Wolfowitz, J., 1943. An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation. Ann. Math. Stat. 14, 378–388.
Wasko, C., Sharma, A., 2015. Steeper temporal distribution of rain intensity at higher temperatures

within Australian storms. Nat. Geosci. 8, 527–529. https://doi.org/10.1038/ngeo2456

Wigmosta, M.S., Nijssen, B., Storck, P., 2002. The distributed hydrology soil vegetation model, in: Singh, V.P. (Ed.), Mathematical Models of Small Watershed Hydrology and Applications. Water Resources Publication, Highlands Ranch, Colorado, Colorado, pp. 7–42.

Wigmosta, M.S., Vail, L.W., Lettenmaier, D.P., 1994. A distributed hydrology-vegetation model for complex terrain. Water Resour. Res. 30, 1665–1679. https://doi.org/10.1029/94WR00436

Yan, H., Sun, N., Chen, X., Wigmosta, M.S., 2020a. Next-Generation Intensity-Duration-Frequency Curves for Climate-Resilient Infrastructure Design: Advances and Opportunities. Front. Water 2, 545051. https://doi.org/10.3389/frwa.2020.545051

Yan, H., Sun, N., Fullerton, A., Baerwalde, M., 2021. Greater vulnerability of snowmelt-fed river thermal regimes to a warming climate. Environ. Res. Lett. 16, 054006. https://doi.org/10.1088/1748-9326/abf393

Yan, H., Sun, N., Wigmosta, M., Leung, L.R., Hou, Z., Coleman, A., Skaggs, R., 2020b. Evaluating next-generation intensity–duration–frequency curves for design flood estimates in the snow-dominated western United States. Hydrol. Process. 34, 1255–1268. https://doi.org/10.1002/hyp.13673

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Hou, Z., Leung, L.R., 2019a. Next-Generation Intensity–Duration–Frequency Curves to Reduce Errors in Peak Flood Design. J. Hydrol. Eng. 24, 04019020. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001799

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Hou, Z., Leung, R., 2018. Next-Generation Intensity-Duration-Frequency Curves for Hydrologic Design in Snow-Dominated Environments. Water Resour. Res. 54, 1093–1108. https://doi.org/10.1002/2017WR021290

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Leung, L.R., Coleman, A., Hou, Z., 2019b. Observed Spatiotemporal Changes in the Mechanisms of Extreme Water Available for Runoff in the Western United States. Geophys. Res. Lett. 46, 767–775. https://doi.org/10.1029/2018GL080260

Yu, G., Wright, D.B., Davenport, F. V., 2022. Diverse Physical Processes Drive Upper-Tail Flood Quantiles in the US Mountain West. Geophys. Res. Lett. 49. https://doi.org/10.1029/2022GL098855

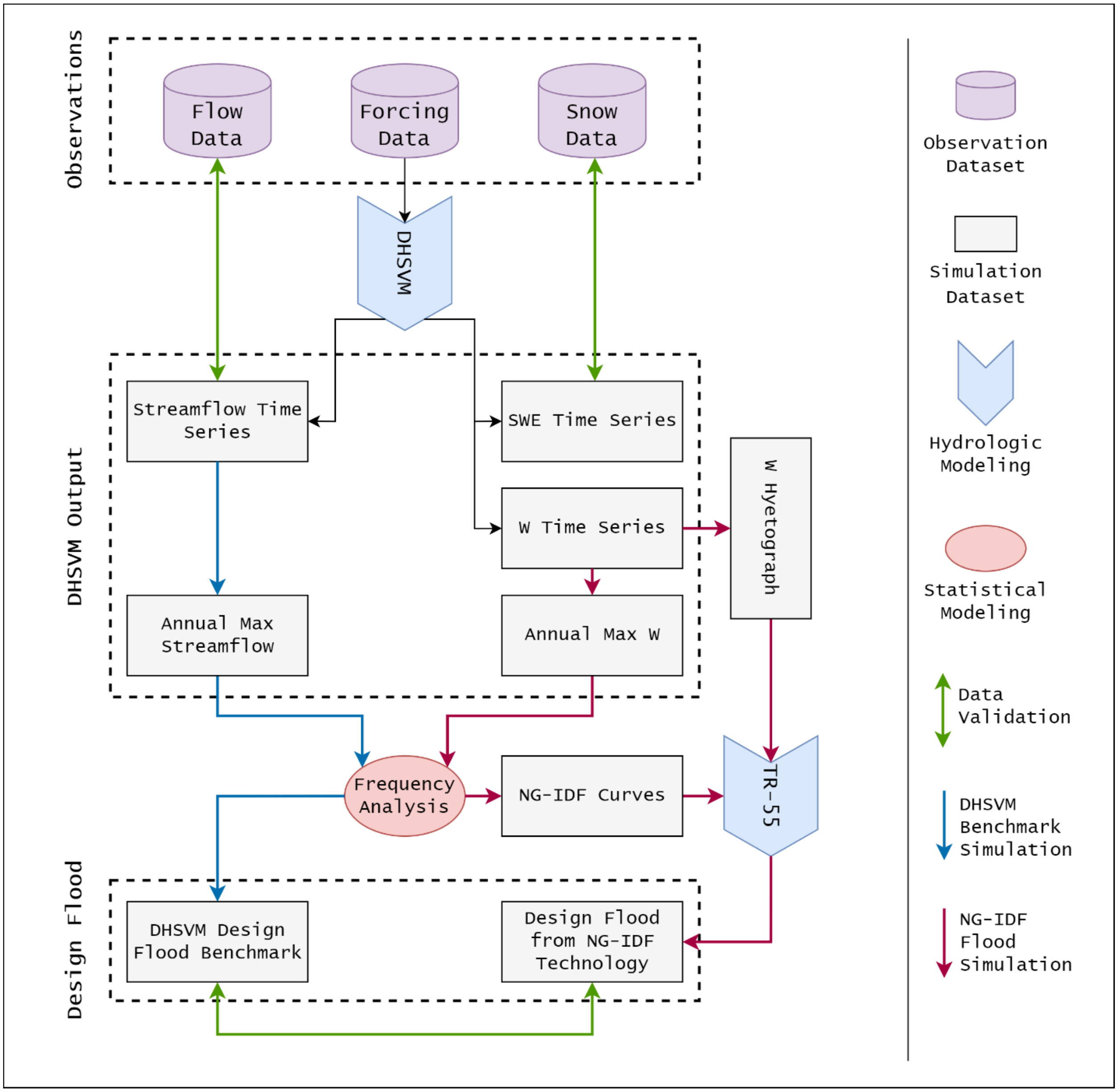


Figure 1. The framework for developing water available for runoff hyetograph and evaluating the NG-IDF technology in design flood estimates using DHSVM continuous simulations.

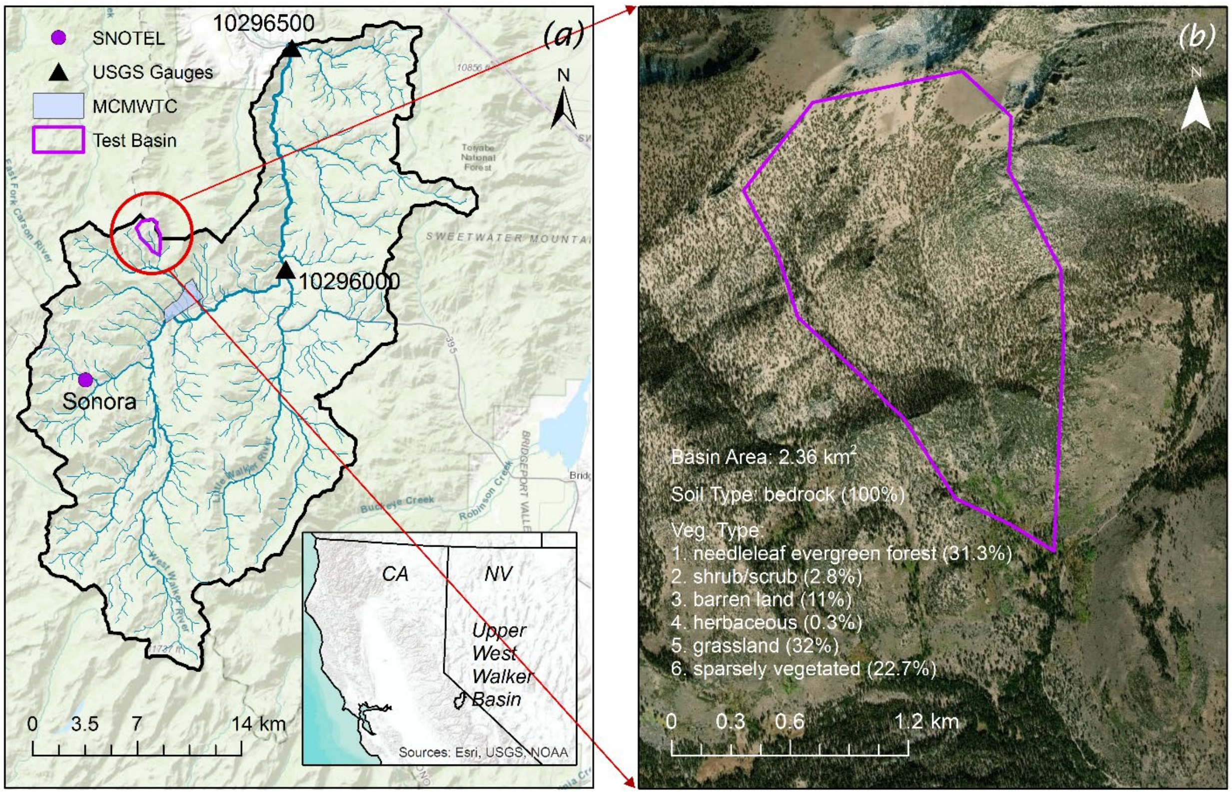


Figure 2. The location of Upper West Walker Bain and the selected small snow-dominated basin that flows into the U.S. Department of Defense (DoD) Marine Corps Mountain Warfare Training Center (MCMWTC).

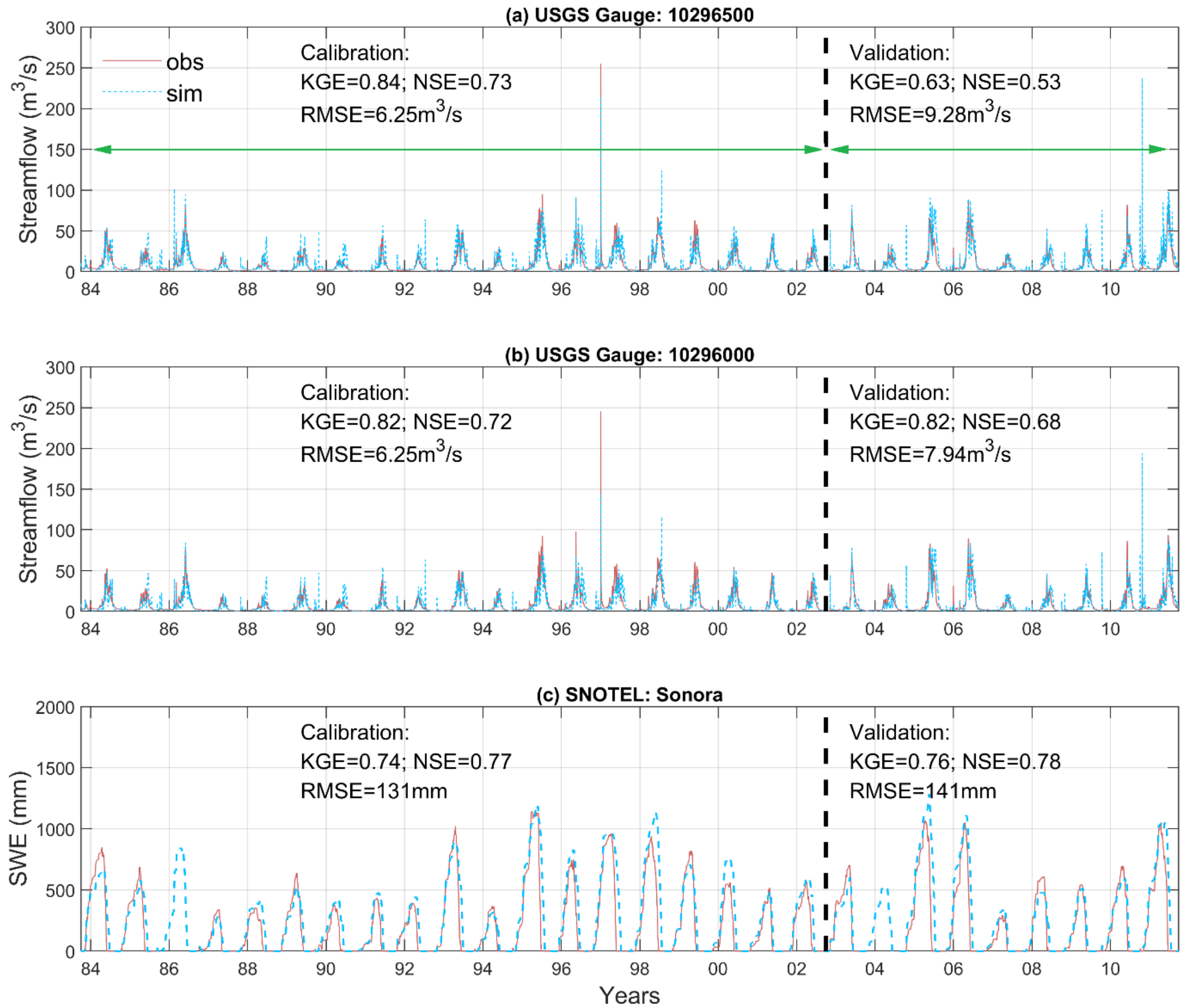


Figure 3. Comparison of measured and simulated daily streamflow (a and b) and snow water equivalent (c) for the Upper West Walker Bain.

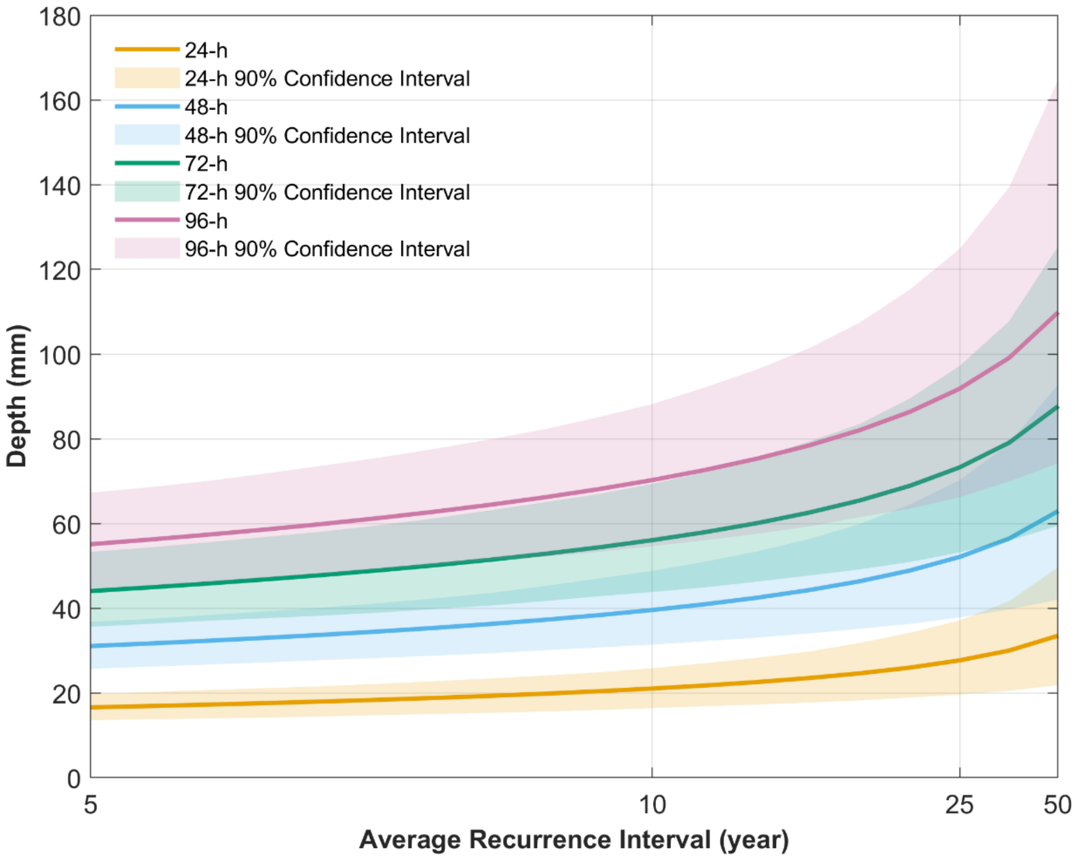


Figure 4. Basin-scale NG-IDF curves with the associated 90% confidence intervals for the four selected durations varying from 24h to 96h.

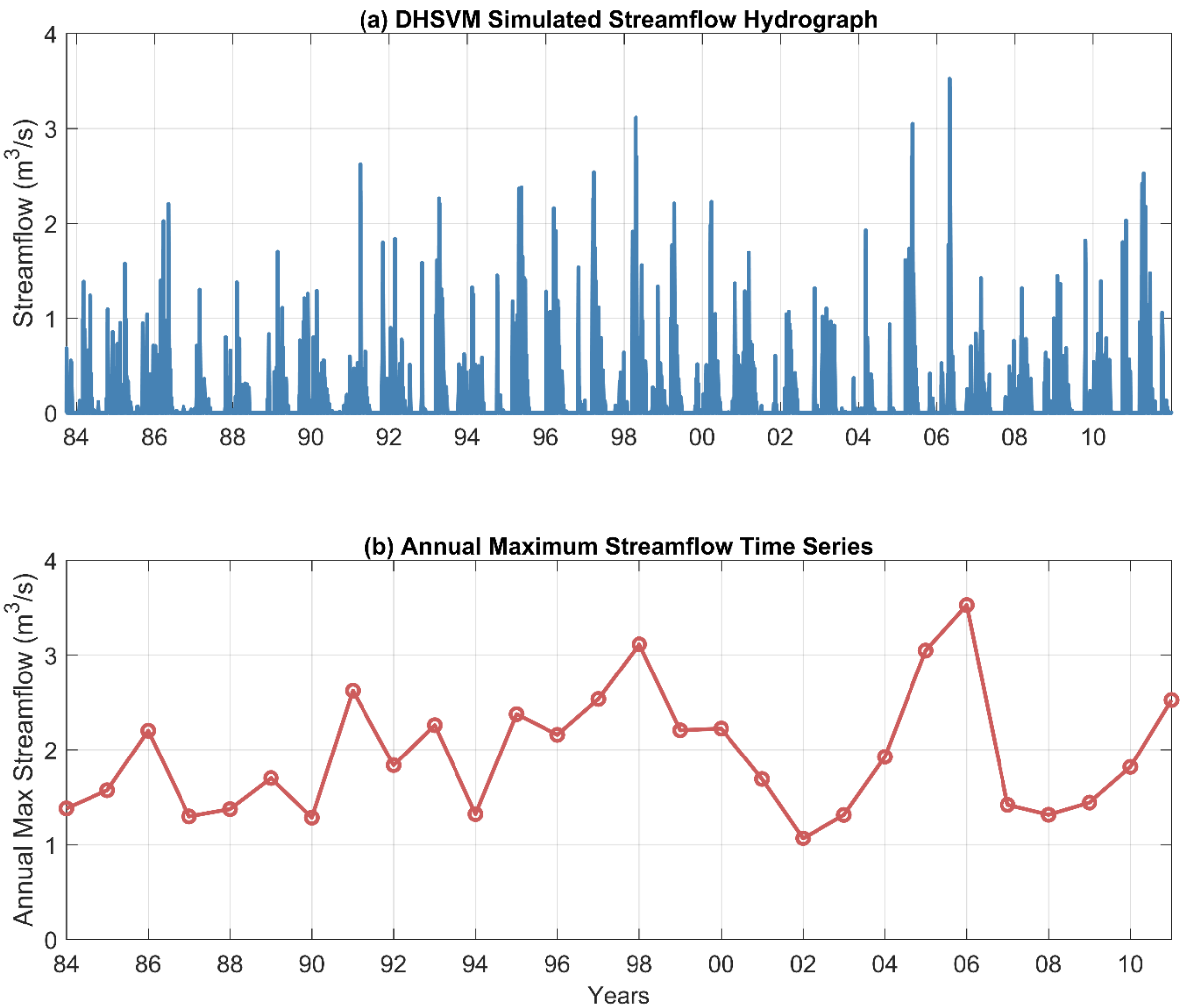


Figure 5. (a) DHSVM simulated 15min hydrograph for the small test basin. (b) Annual maximum (water year) flood time series (obtained from the 15min hydrograph) for the small snow-dominated basin.

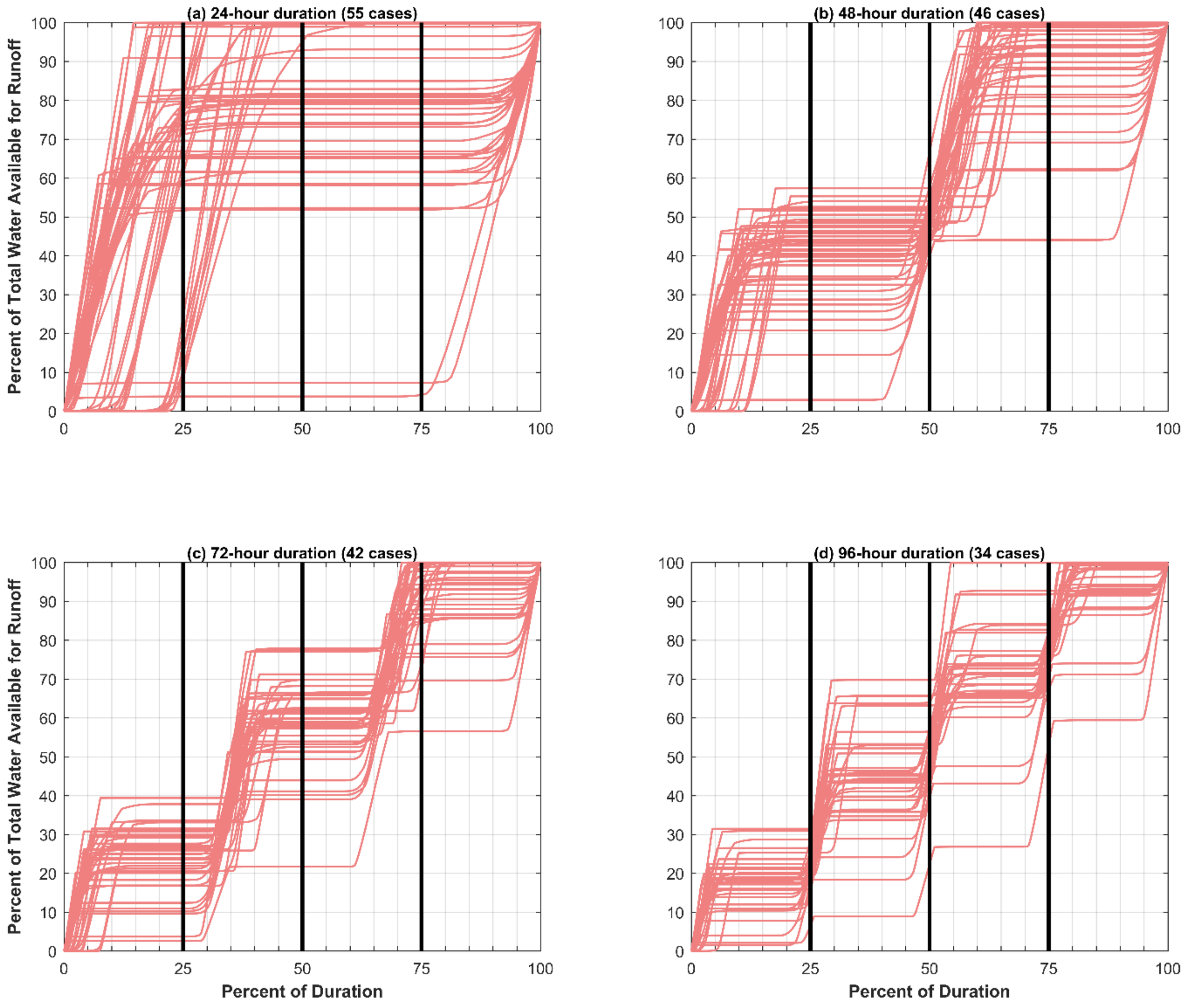


Figure 6. Ensemble hyetographs of water available for runoff (W) for each selected duration. All hyetographs were derived from large snowmelt events only.

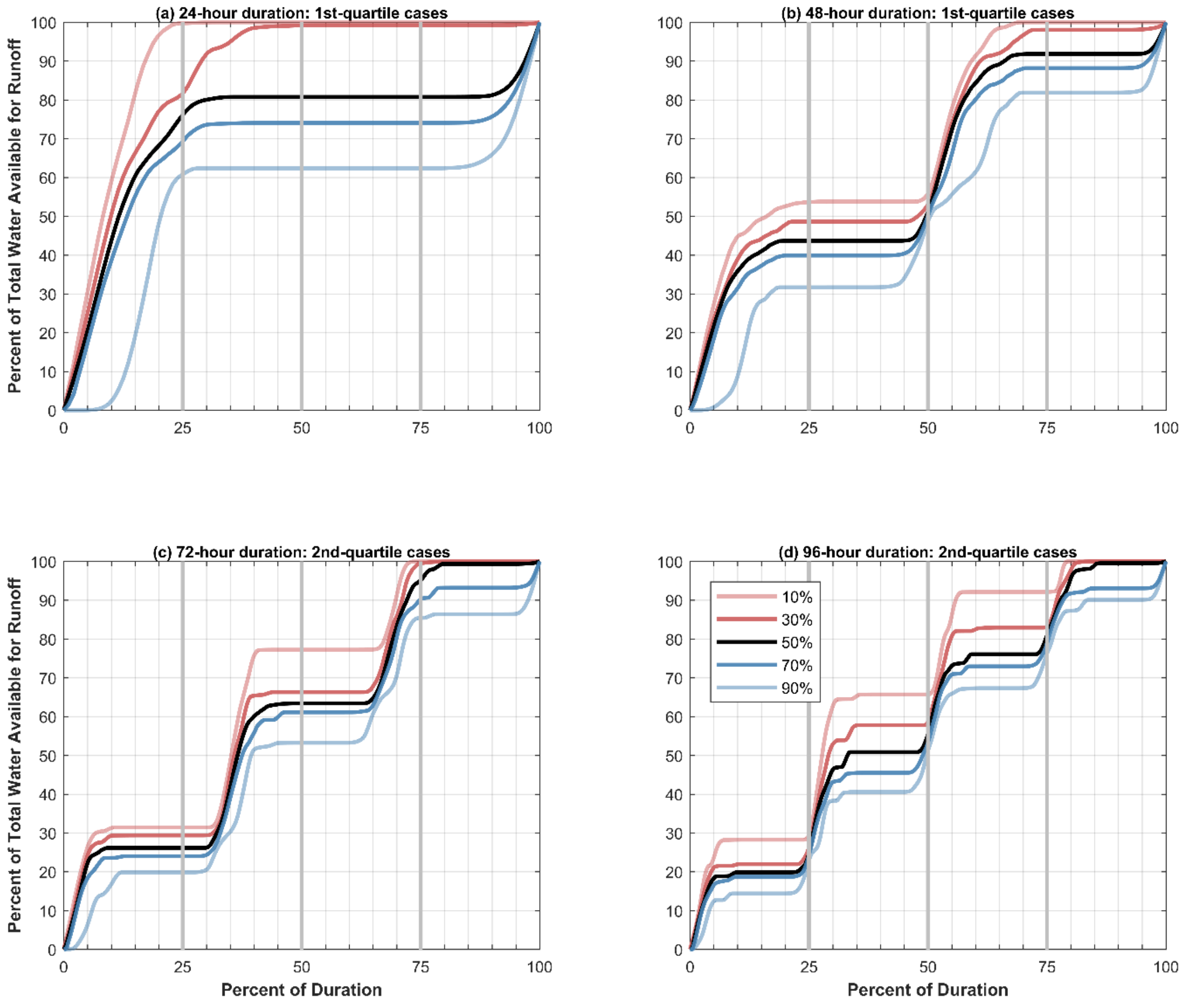


Figure 7. Probabilistic water available for runoff (W) hyetograph for the test basin. The graph represents the cumulative probability of occurrence at 20% increments based only on the W events that belonged to the dominant quartile. Using the 24h duration as an example, the 1st-quartile is the dominant quartile which means that most of the W events had their greatest percentage of the total W fell during the 1st-quarter, i.e., the first 6h.

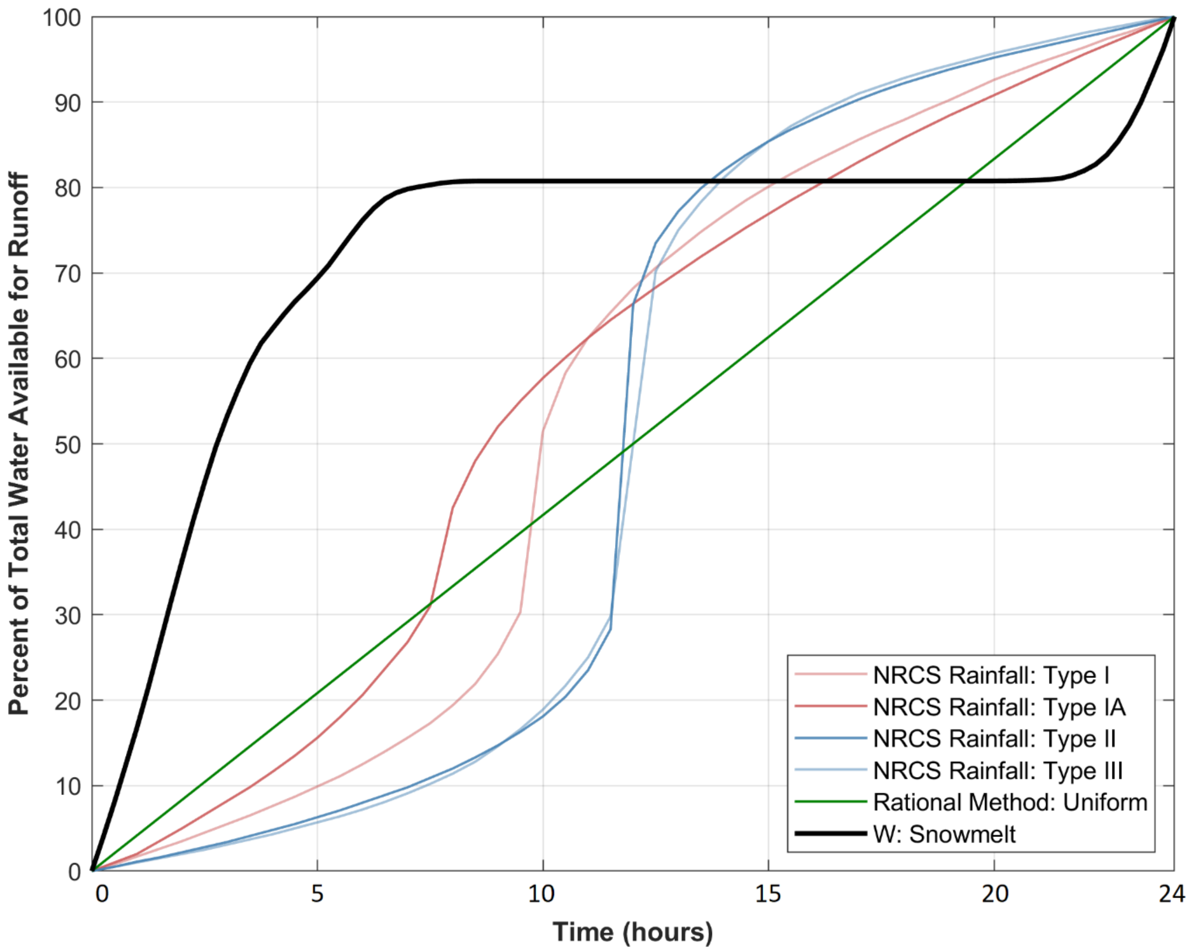


Figure 8. The standard NRCS rainfall hyetographs and uniform rainfall hyetograph versus the snowmelt hyetograph.

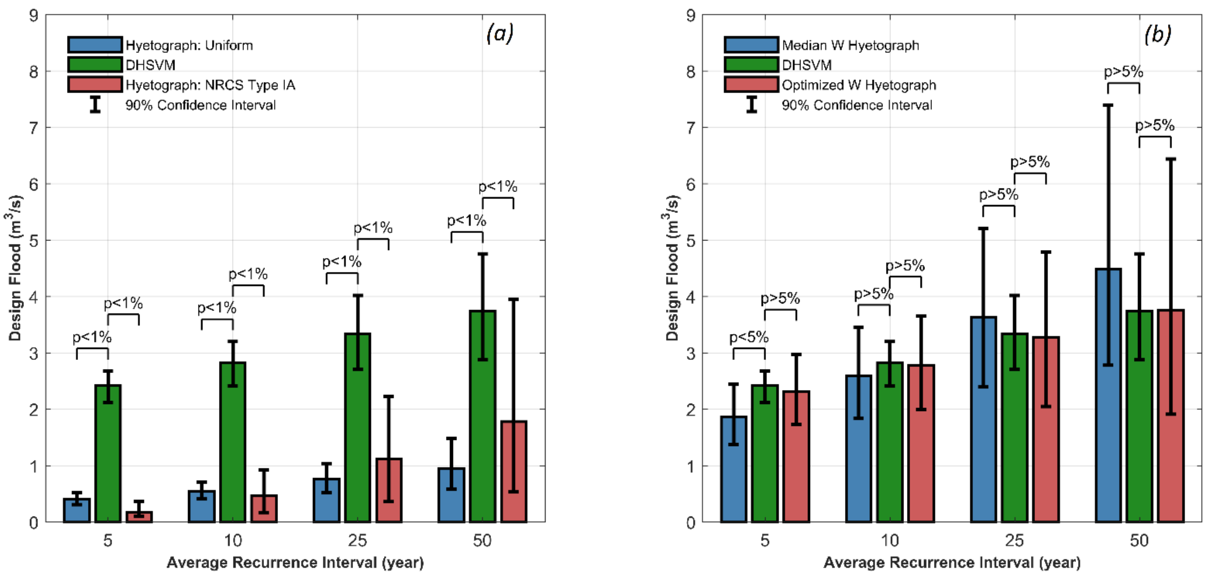
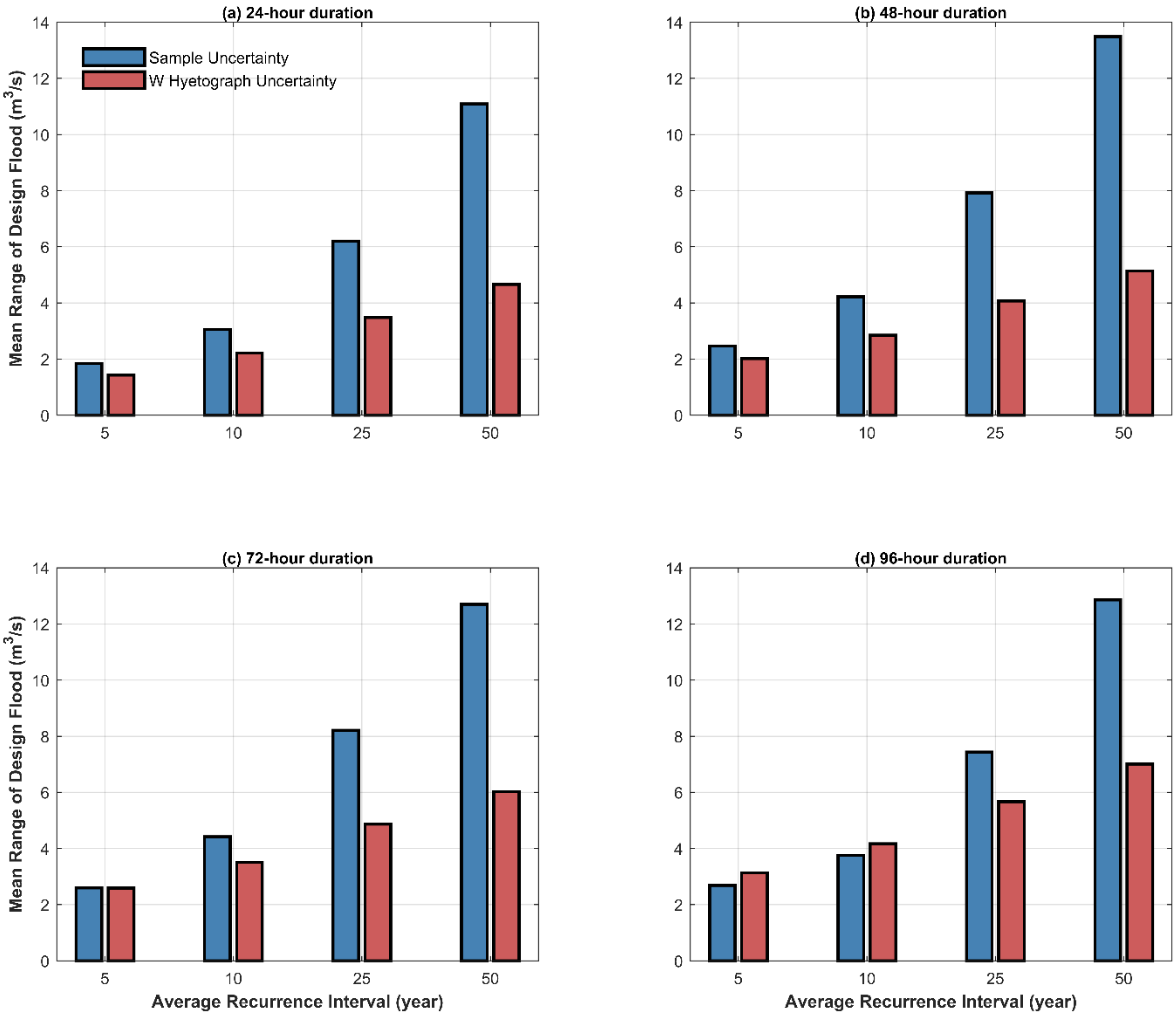


Figure 9. (a) Comparison of design flood estimations from DHSVM and NG-IDF technology in which two standard rainfall hyetographs were used. The error bar represents the associated 90% confidence intervals due to sample data uncertainty. The pair bracket presents the p-value from Z statics in the significance test. (b) Similar to (a) but the developed snowmelt hyetographs were used in the NG-IDF technology. Median W hyetograph indicates the 50% hyetographs extracted from Figure 7, separately for each duration. Optimized W hyetograph indicates that we ran the TR-55 with different W hyetographs (e.g., varying from 10% curve to 90% curve) and reported the best design flood estimates we had achieved.



1249
1250
1251 **Figure 10.** Mean range of design flood estimates depicting the uncertainty contribution of the
1252 sample data and water available for runoff hyetograph inTR-55 modeling for the four selected
1253 durations and ARIs.