BUREAU OF RECLAMATION

Stochastic Storm Transposition for Physically-Based Rainfall and Flood Hazard Analyses

Science and Technology Program
Research and Development Office
Final Report No. ST-2020-1735-1
Technical Memo No. ENV-2020-070

U.S. Department of the Interior **September 2020**

Mission Statements

The Department of the Interior (DOI) conserves and manages the Nation's natural resources and cultural heritage for the benefit and enjoyment of the American people, provides scientific and other information about natural resources and natural hazards to address societal challenges and create opportunities for the American people, and honors the Nation's trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated island communities to help them prosper.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

Disclaimer

Information in this report may not be used for advertising or promotional purposes. The data and findings should not be construed as an endorsement of any product or firm by the Bureau of Reclamation, Department of Interior, or Federal Government. The products evaluated in the report were evaluated for purposes specific to the Bureau of Reclamation mission. Reclamation gives no warranties or guarantees, expressed or implied, for the products evaluated in this report, including merchantability or fitness for a particular purpose.

Acknowledgements

The Science and Technology Program, Bureau of Reclamation, sponsored this research.

Stochastic Storm Transposition for Physically-Based Rainfall and Flood Hazard Analyses

Final Report No. ST-2020-1735-1
Technical Memo No. ENV-2020-070

Prepared by

Kathleen D. Holman, Ph.D., Meteorologist
Water Resources Engineering and Management Group
Technical Service Center
Bureau of Reclamation

Daniel B. Wright, Ph.D., Assistant Professor
Civil and Environmental Engineering Department
University of Wisconsin-Madison

Guo Yu, Graduate Student
Civil and Environmental Engineering Department
University of Wisconsin-Madison

Cover Photo Credit: Colorado Department of Transportation (2013)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-09-2020	Research	5/2017-9/2020

4. TITLE AND SUBTITLE
Stochastic Storm Transposition for Physically-Based Rainfall and Flood Frequency Analyses

5a. CONTRACT NUMBER
Z1735

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
1541 (S&T)

6. AUTHOR(S)
Kathleen D. Holman, Ph.D.
Daniel B. Wright, Ph.D.
Guo Yu

5d. PROJECT NUMBER
Final Report No. ST-2020-1735-1

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Bureau of Reclamation
U.S. Department of the Interior
Denver Federal Center
PO Box 25007, Denver, CO 80225-0007

University of Wisconsin-Madison
College of Engineering
1415 Engineering Dr, Madison, WI 53706

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Science and Technology Program
Research and Development Office
Bureau of Reclamation
U.S. Department of the Interior
Denver Federal Center
PO Box 25007, Denver, CO 80225-0007

10. SPONSOR/MONITOR'S ACRONYM(S)
Reclamation

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Final Report No. ST-2020-1735-1
Technical Memo No. ENV-2020-070

12. DISTRIBUTION/AVAILABILITY STATEMENT

Final Report may be downloaded from https://www.usbr.gov/research/projects/index.html

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Reclamation partnered with faculty from the University of Wisconsin-Madison to address a need for improved gridded technologies that support precipitation and flood frequency analyses. Research products developed in this study include a gridded, web-based stochastic storm transposition tool referred to as RainyDay and gridded flood-frequency estimates developed using a stochastic version of WRF-Hydro coupled to RainyDay output. These products were demonstrated within the Big Thompson watershed above Olympus Dam in Colorado.

15. SUBJECT TERMS
Stochastic storm transposition, precipitation-frequency analysis, flood-frequency analysis, hydrologic hazard analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
U	U	U			

Peer Review

Bureau of Reclamation
Research and Development Office
Science and Technology Program

Final Report No. ST-2020-1735-1
Technical Memo No. ENV-2020-070

Stochastic Storm Transposition for Physically-Based Rainfall and Flood Hazard Analyses

KATHLEEN HOLMAN
Digitally signed by KATHLEEN HOLMAN
Date: 2020.09.22 13:05:51 -06'00'

Prepared: Kathleen D. Holman, Ph.D.
Meteorologist, Bureau of Reclamation

Prepared: Daniel B. Wright, Ph.D.
Assistant Professor, University of Wisconsin-Madison

Prepared: Guo Yu
Graduate Student, University of Wisconsin-Madison

AMANDA STONE
Digitally signed by AMANDA STONE
Date: 2020.09.22 13:10:11 -06'00'

Peer Review: Amanda Stone, PE
Hydrologic Engineer, Bureau of Reclamation

"This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the Bureau of Reclamation. It does not represent and should not be construed to represent Reclamation's determination or policy."

Contents

Page

Executive Summary

Detailed precipitation inputs are needed for a variety of probabilistic flood hazard analyses. These precipitation inputs typically take the form of historical "storm templates" and a basin-average precipitation-frequency relationship. Storm templates are developed using historical data (e.g., point observations and reanalysis data) to capture the best representation of major storm events that actually occurred within the basin of interest. Most studies utilize a small number of individual storm templates (typically less than 20) to simulate hundreds to tens of thousands of annual or seasonal maximum flow events. The basin-average precipitation-frequency relationships are always based on historical point observations within and surrounding the basin of interest. These analyses do not take advantage of any gridded precipitation datasets, even those gridded precipitation datasets that incorporate estimates of uncertainty. There is a clear need for developing alternative precipitation tools for application in flood hazard analyses that are technically sound and independent from existing methods.

Reclamation partnered with faculty from the University of Wisconsin-Madison to address this need. The project was motivated by two main questions. First, can advanced gridded datasets and technologies be utilized to help develop precipitation inputs for flood hazard analyses? Second, can the application of those advanced gridded precipitation tools in conjunction with a gridded rainfall-runoff model be demonstrated within a basin of interest to Reclamation?

In order to answer the first question, we expanded software capabilities of RainyDay, an open-source Python-based Stochastic Storm Transposition (SST) implementation that uses probabilistic temporal resampling and spatial transposition of storm events drawn from a "catalog" of historical storms to effectively recreate the climatology of extreme rainfall frequency and intensity, to include a web-based version hosted by the University of Wisconsin-Madison. The web-based RainyDay includes two versions, basic and advanced, and provides users access to three different historical gridded precipitation datasets. This advanced probabilistic tool was demonstrated in the Big Thompson watershed above Olympus Dam, in Colorado. We answered the second question by combining output from RainyDay with a process-based rainfall-runoff model, WRF-Hydro, and Monte Carlo sampling procedures to develop peak discharge and inflow volume frequency relationships at Olympus Dam. Collectively, these new technologies represent major advances for future applications in hydrologic hazard analyses performed at Reclamation and beyond.

Main Report

Problem Statement

Reclamation's Water Resources Engineering and Management Group, housed in the Technical Service Center, produces probabilistic estimates of hydrologic hazards to support risk-informed decision making. The probabilistic estimates are often developed using a variety of methods, including streamflow observations, statistical analyses, paleoflood records, numerical modeling, and combinations thereof. Statistical approaches typically include fitting a probability distribution function to historical annual maximum streamflow records from within or around the basin of interest. Paleoflood records, when available, can provide flow thresholds and associated non-exceedance probabilities, which are critical to providing confidence of flow estimates at very rare return periods. Numerically-based methods involve developing and calibrating a rainfall-runoff model to a watershed of interest. The group employs multiple methods in an effort to better understand a range of floods that is difficult to predict using traditional flood-frequency estimates given the general lack of observed streamflow events for such rare return periods. Some of the most complex studies completed by the group employ all these methods in an effort to increase confidence in rare flood estimates.

The rainfall-runoff models employed by the group, which vary in complexity depending on the study needs, typically require detailed precipitation inputs in the form of storm templates and a site-specific (basin-average) precipitation-frequency relationship. For more detailed studies involving stochasticized rainfall-runoff simulations, together, the storm templates and precipitation-frequency analysis are used to force the rainfall-runoff model tens to hundreds of thousands of times to produce probabilistic magnitude-frequency estimates (Reclamation 2020a, b), such as peak inflow frequency and reservoir inflow volume frequency relationships. Many of the previous studies performed by the group utilize between five and 20 storm templates (e.g., Reclamation (2012) and (2016)) that produced precipitation and flow in the basin of interest and assume that such a small pool of historical events is sufficient to encompass the wide range of extreme storms that are possible in a basin at very low annual exceedance probabilities. Some of the simpler flood modeling studies also include the assumption of "AEP neutrality", whereby a precipitation event with a specified annual exceedance probability (AEP) can be converted to a runoff event of the same AEP. These studies do not typically utilize stochastic methods. The validity of these assumptions is questionable in many watersheds due to the complex space-time structure of extreme storms, the wide spectrum of possible soil moisture and snowpack conditions, and the joint role of these factors on flood wave propagation through river tributaries.

The purpose of the current project was to develop advanced, web-based stochastic software capable of combining gridded precipitation ensemble datasets, stochastic storm transposition methods, and physically-based hydrologic models to estimate precipitation and flood frequency relationships for use in hydrologic hazard investigations without the assumption of AEP neutrality, though the methods are applicable to such studies. The tools and methods developed

in this project will help identify strengths and weaknesses of watershed hydrologic methods and spatially distributed weather data, two priority areas in Reclamation's long term water resources planning objectives.

Research Activities and Results

The research presented herein is characterized by two main components. The first component involves developing a web-based tool for stochastic storm transposition (SST) using gridded precipitation datasets, while the second component involves utilizing the SST tool in conjunction with a gridded rainfall-runoff model to develop flood-frequency estimates for a watershed of interest to Reclamation. Both components are demonstrated for the Big Thompson watershed, located in Colorado, but the general methodologies are transferable to other Reclamation studies watersheds. The main approach and findings from each component are detailed below.

SST is a process that involves resampling and transposing (i.e., geographically moving in the zonal and meridional directions) storm events to generate hypothetical events from a collection of realistic events (Wright et al. 2020). The method was originally conceived as probabilistic alternative to probable maximum precipitation (PMP). SST "lengthens" the rainfall record by probabilistic temporal resampling and spatial transposition of storm events drawn from a "catalog" of historical storms to effectively recreate the climatology of extreme rainfall frequency and intensity. A detailed explanation of the methodology can be found in Wright et al. (2013). Wright et al. (2014) showed that when coupled with rainfall remote sensing datasets, SST can provide unique multi-scale flood hazard estimation capabilities that represent a fundamental improvement over existing design storm approaches, since they allow for realistic "decoupling" of rainfall and flood annual exceedance probabilities (i.e. not requiring an assumption of "AEP neutrality") and explicit consideration of the complex range of rainfall space-time structures. Accurate estimation of rainfall intensity-duration-frequency curve (IDF) and flood frequencies were demonstrated for return periods as high as 1,000 years using only ten years of radar rainfall data.

In this study, we expanded on previous research and software by developing a web-based version of RainyDay (https://her.cee.wisc.edu/rainyday-rainfall-for-modern-flood-hazard-assessment/). This web-based RainyDay collects user inputs (e.g., several parameters) from a web interface, runs the RainyDay software on High Performance Computers at the University of Wisconsin-Madison, and automatically sends the results back to the user via a Google Drive downloadable link (see Appendix A for details). This web-based RainyDay has been updated to the latest version of the RainyDay software, including some recently developed functions, like stochastic rescaling transposition (Appendix B; Wright and Holman, 2019).

The web-based version of RainyDay provides many benefits to Reclamation. First, the web version includes two different versions for variable user needs, basic and advanced. The advanced version allows the user to upload a shapefile of the basin of interest, producing watershed-specific precipitation-frequency results and a storm catalog. Another benefit of the web version is that it gives the user immediate access to three different gridded precipitation datasets with variable temporal coverage. The three gridded precipitation datasets allow the user

to produce three precipitation-frequency analyses that vary as a function of the underlying data source. Finally, and possibly most importantly, the web-based version of RainyDay represents an independent source of precipitation-frequency estimates that Reclamation users can use for comparisons in site-specific studies, including screening-level studies and more advanced hydrologic hazard analyses. This represents an improvement over the National Weather Service's (NWS's) Hydrometeorological Design Studies Center's (HDSC) NOAA Atlas 14, which includes point precipitation-frequency estimates out to return periods of 1,000 years across parts of the United States.

In many hydrologic analyses performed at Reclamation, precipitation-frequency estimates are combined with historical storm events within a rainfall-runoff model to produce magnitude-frequency relationships (e.g., three-day average flow frequency relationship) for a reservoir of interest. Current rainfall-runoff models in use at Reclamation include versions of the Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS; USACE 2016) and the Stochastic Event Flood Model (SEFM; MGS 2018). SEFM and HEC-HMS are typically run using lumped or semi-distributed configurations (e.g., Reclamation (2019)). Lumped rainfall-runoff models treat catchments as a series of homogeneous subbasins and are typically well suited for small, relatively homogeneous watersheds. These models simulate streamflow at outlet points, not flows within a catchment (EPA 2017). Semi-distributed models are similar to lumped models, though they typically treat the watershed as a series of (lumped) subbasins, simulating streamflow at outlet points for each subbasin. Gridded products can be used to develop inputs to lumped or semi-distributed models, though some resolution is typically lost in the transition.

Unlike lumped or semi-distributed models, gridded rainfall-runoff models account for spatial heterogeneity in model parameters and input fields. Fully-distributed models compute the hydrologic response on each grid cell within a catchment and incorporate impacts from neighboring grid cells. In this study, we developed a model of the Big Thompson watershed above Olympus Dam in Colorado using WRF-Hydro. WRF-Hydro is a large-scale hydrologic modeling system that combines a land surface model, Noah MP, with terrain routing, baseflow routing, and channel routing (Gochis et al. 2018). Both energy and water fluxes are simulated within the model. The model can be run offline as a hydrologic model or coupled with an atmospheric model, such as the Weather, Research, and Forecasting model (WRF).

The WRF-Hydro modeling framework began with the calibration and validation of the watershed between 1979 and 2018 using forcing data from the North American Land Data Assimilation System (NLDAS). The physical states from this long-term simulation as simulated by WRF-Hydro were saved and later used as initial conditions in the stochastic simulations. A storm catalog was developed for the basin using precipitation from NLDAS and SST technologies within RainyDay. Finally, 10 sets of 1,000 annual maximum streamflow events (for a total of 10,000) were simulated by sampling the RainyDay storm catalog and initial conditions from the long-term WRF-Hydro simulation. The 10 sets of annual maximum flow events can be used to estimate median and uncertainty estimates at return periods less than 1,000 years or to estimate a single frequency curve (without uncertainty) out to the 10,000-year return period. Results from this portion of the study are available in Appendix D, which represents a manuscript that will be submitted to *Water Resources Research*.

The gridded flood frequency analysis completed using WRF-Hydro represents an important advancement for Reclamation. This study and the associated technology transfer that took place on August 12, 2020 have provided employees with background, theoretical developments, and real applications of new technologies to a watershed of interest to Reclamation, the Big Thompson watershed above Olympus Dam. Furthermore, results from this modeling have been made available to members of the Water Resources Engineering and Management Group for further analysis.

Conclusions and Future Plans

Members of the TSC are tasked with producing probabilistic estimates of hydrologic hazards to support risk-informed decision making. The probabilistic estimates are often developed using a variety of methods, including streamflow observations, statistical analyses, paleoflood records, numerical modeling, and combinations thereof. The purpose of the current project was to develop advanced gridded tools that can be used to support probabilistic hydrologic hazard analyses within Reclamation. The first objective of this project was to develop an advanced, web-based SST tool for probabilistic precipitation analyses, including storm catalogs and precipitation-frequency analyses. The second objective was to combine SST technology with a physically-based rainfall-runoff model to develop flood frequency estimates at a basin of interest to Reclamation. These objectives have been successfully met through the development of the online version of RainyDay and through the development of a WRF-Hydro modeling system demonstrated for the Big Thompson watershed above Olympus Dam, in Colorado. Findings have been demonstrated through multiple peer reviewed papers and a webinar hosted on August 12, 2020 that was attended by more than 20 people within and outside of Reclamation.

Reclamation benefits from this project in a number of unique ways. For example, RainyDay represents a modern version of an age-old technique, stochastic storm transposition, to support probabilistic flood studies. RainyDay also provides users quick access to three, gridded precipitation datasets. These datasets can be used individually or collectively to develop storm catalogs and precipitation-frequency estimates for use in hydrologic hazard analyses. RainyDay also provides Reclamation with independent precipitation-frequency estimates across the western US. This is in contrast to NOAA Atlas 14, which provides precipitation-frequency estimates out to the 1,000-year return period across portions of the western US. Finally, with modifications to the underlying data, RainyDay can be used to understand impacts of climate variability and change on precipitation frequency estimates. Beyond RainyDay, the technologies from WRF-Hydro also represent a large step forward. Members of the research team provided Reclamation with background, theoretical developments, and a demonstration of new, gridded technologies to support flood frequency analyses in a watershed of interest to Reclamation.

Future applications of the RainyDay include canal safety projects, screening level studies (also referred to as comprehensive reviews), and larger, in-depth studies. Future WRF-Hydro applications include diagnosing the physical mechanism of flooding among the 10,000 synthetic annual maximum flood events using an artificial neural network, specifically the Self-Organizing Maps algorithm, for the broader purpose of understanding the impacts of mixed populations on flood frequency estimates. Additional ideas include utilizing RainyDay and WRF-Hydro to

develop historical and future flood frequency estimates for Reclamation basins using output from global climate model simulations.

References

Environmental Protection Agency (EPA), 2017. An Overview of Rainfall-Runoff Model Types. EPA/600/R-14/152. Accessed 07/30/2020.

Gochis, D., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, J. McCreight, J. Mills, L. Pan, A. Rafieei-Nasab, L. Read, K. Sampson, D. Yates, W. Yu, and Y. Zhang, 2018: The WRF-Hydro Modeling System Technical Description Version 5.0, UCAR/NCAR.

MGS Software LLC (MGS), 2018: Stochastic Event Flood Model (SEFM) Technical Support Manual. 168 pp.

Reclamation (Bureau of Reclamation), 2012. Altus Dam Hydrologic Hazard and Reservoir Routing for Corrective Action Study. W.C. Austin Project, OK. Great Plains Region.

Reclamation (Bureau of Reclamation), 2016. Island Park Dam Meteorology for Application in Hydrologic Hazard Analysis. Minidoka Project, ID. Pacific Northwest Region.

Reclamation (Bureau of Reclamation), 2019. Stochastic Flood Frequency Analysis at Folsom Dam: Rainfall-Runoff Modeling. Central Valley Project, California. Mid-Pacific Region. Technical Memorandum: ENV-2019-070.

Wright, D.B., J.A. Smith, G. Villarini, and M.L. Baeck. 2013. Estimating the Frequency of Extreme Rainfall Using Weather Radar and Stochastic Storm Transposition. *J. Hydrology*. **488**, 150-165.

Wright, D.B., J.A. Smith, and M.L. Baeck. 2014. Flood Frequency Analysis Using Radar Rainfall Fields and Stochastic Storm Transposition. *Water Resour. Res.,* **50**, 1592-1615. doi:10.1002/2013WR014224.

Appendix A – Research Products

A Web-Based Rainfall Hazard Assessment Tool (RainyDay)

RainyDay (Wright et al., 2017) is a framework for generating large numbers of realistic extreme rainfall scenarios based on remotely-sensed precipitation fields. It is founded on a statistical resampling concept known as stochastic storm transposition (SST). These rainfall scenarios can then be used to examine the extreme rainfall statistics for a user-specified region, or to drive a hazard model (usually a hydrologic model, but the method produces output that would also be useful for landslide models).

This web-based version of RainyDay is intended to make the software more accessible to a wide range of potential users (https://her.cee.wisc.edu/rainyday-rainfall-for-modern-flood-hazard-assessment/). This web tool allows users to type in the inputs on the website, runs the RainyDay simulation on the server, and sends the results back to users through an email with a downloadable Google Drive link (Figure 1). Users do not need to possess High Performance Computers, configure Python environment, download large amounts of input data, etc. We also created two tutorials on YouTube to facilitate potential users to understand how the SST method works and how to best use this web-based tool based on their specific application. (www.youtube.com/watch?v=H7v3RC1OqYg&list=PLmoYfOUvB7rUWkjxFkGQTSOT3lUUgzHe2) The user interface of the web-based RainyDay is shown in Figure 2.

The common outputs that the web-based RainyDay send to user include four main types of outputs:

- **Storm Catalogs:** A storm catalog is the collection of the most extreme storms within the user defined SST domain that will be used for subsequent resampling and transposition. The output is a NetCDF file.

- **Diagnostic plots:** These plots show different aspects of the rainfall fields and the storm catalog. A diagnostic plot of the storm catalog created for Big Thompson River watershed is shown in Figure 3.

- **FreqAnalysis file:** This file provides the results from a rainfall frequency analysis for various return periods for an area and a rainfall duration, both of which are specified by the user. The results in the .FreqAnalysis file are roughly equivalent to an Intensity-Duration-Frequency (IDF) curve, which many civil engineers and hydrologists are familiar with (Figure 4).

- **Spacetime rainfall scenarios:** These are scenarios that can serve as input to a hydrologic model. In order to keep the size of these scenarios manageable, they do not consider the entire transposition domain but only the rainfall within a certain user-defined area (usually either a rectangular “box” or a watershed boundary).

Figures

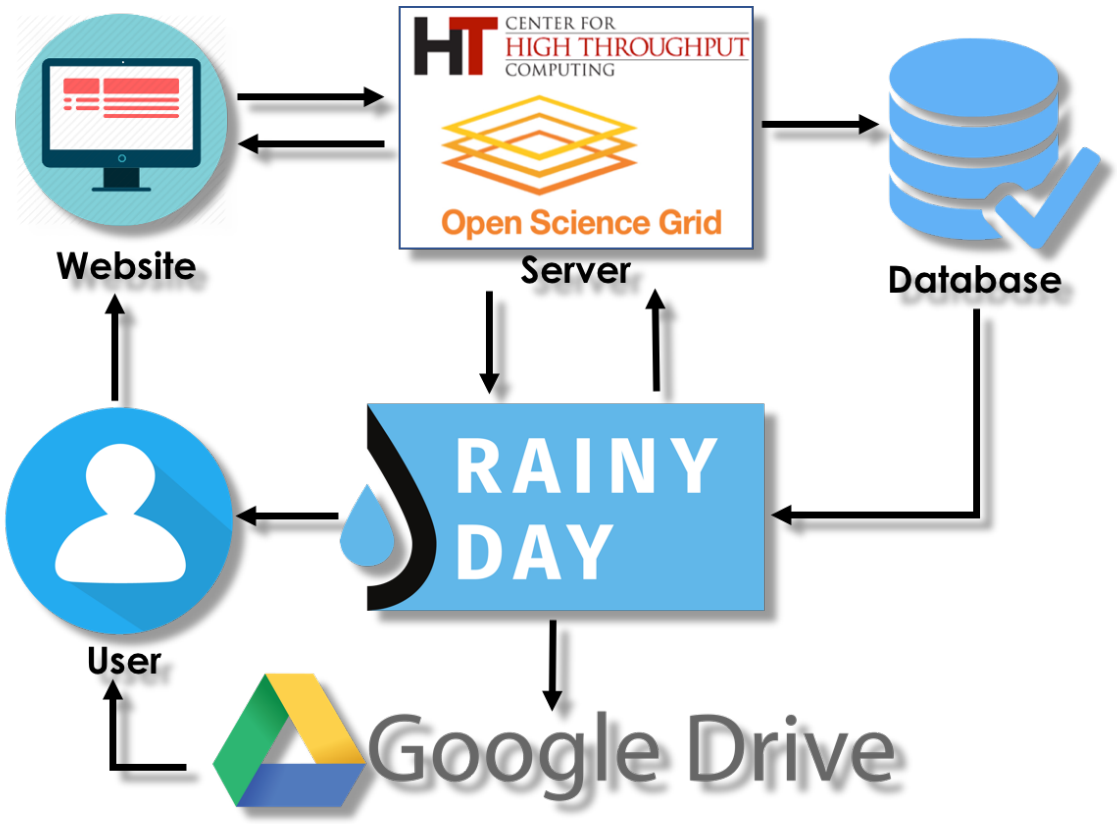

Figure 1. The conceptual flowchart of the Web-RainyDay.

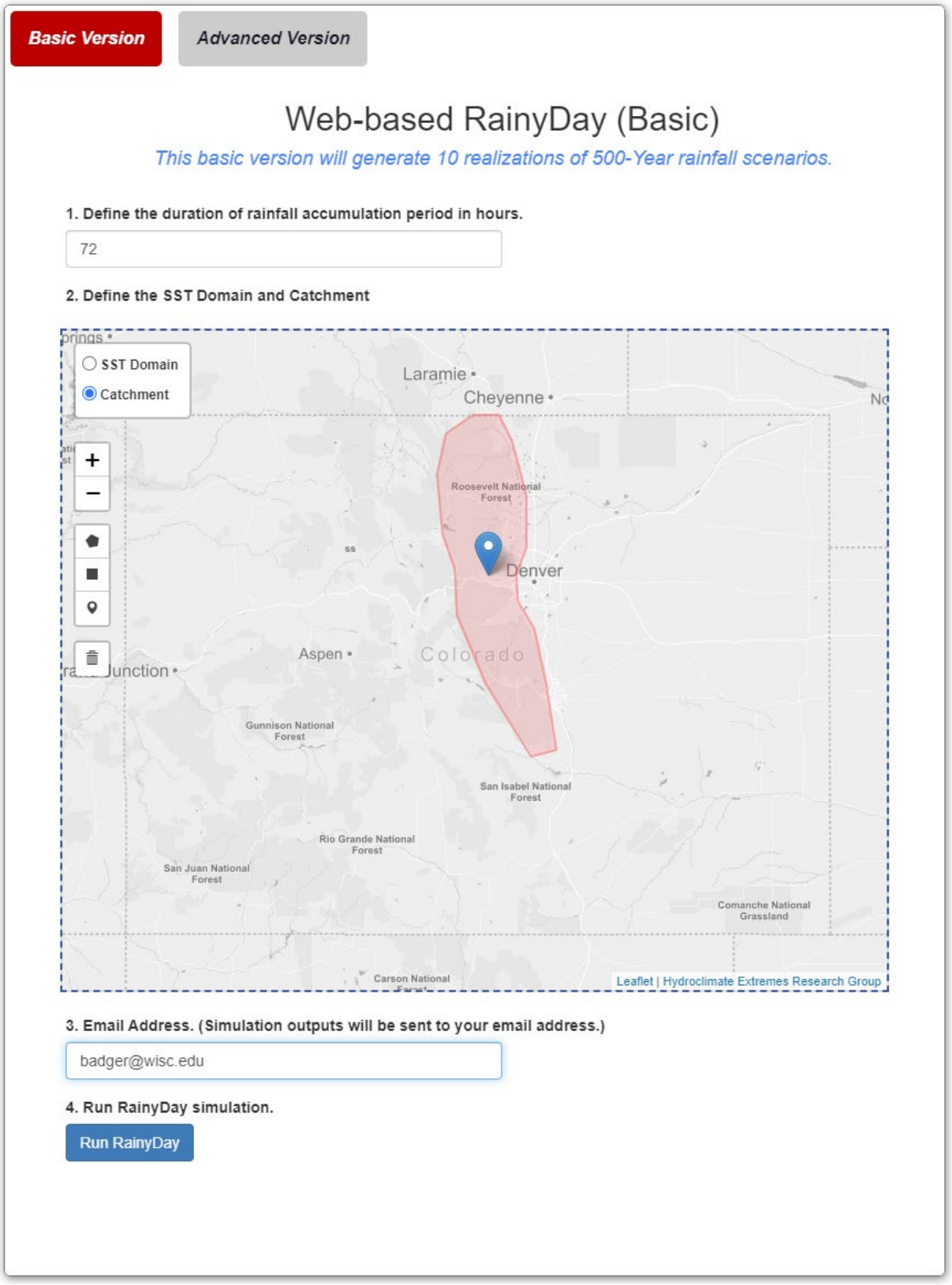

Figure 2. The user interface of the web-based RainyDay (Basic Version).

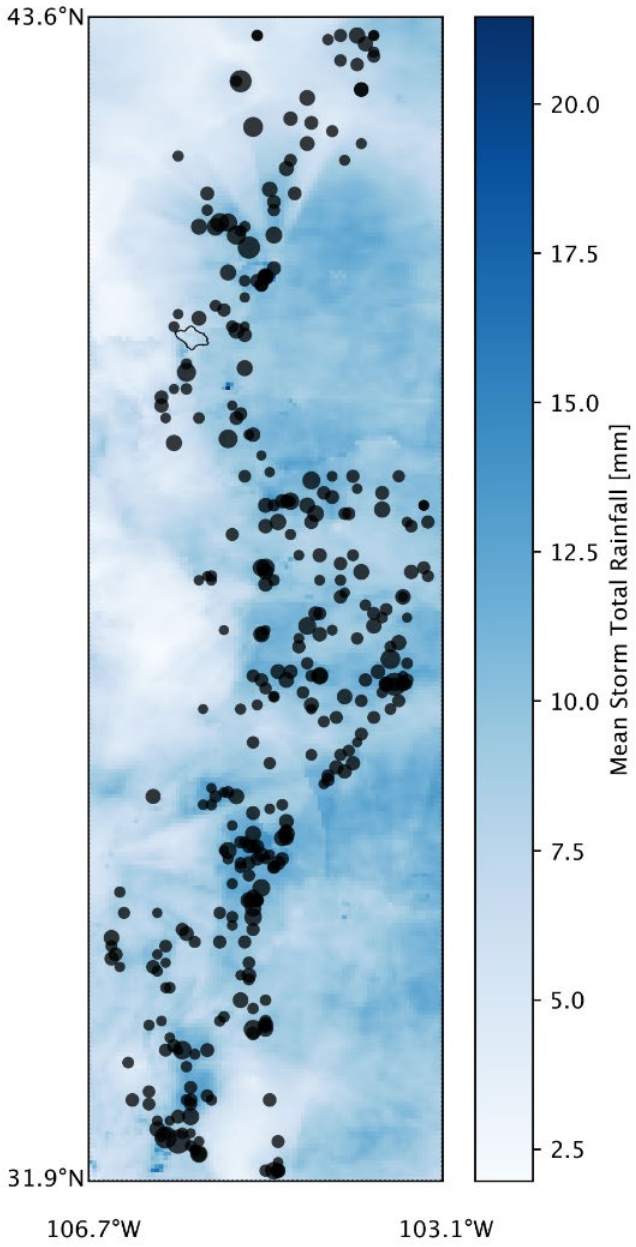

Figure 3. Average storm rainfall and location of storm centers. The black outline is the watershed (Big Thompson River upstream of Olympus Dam) used in this example.

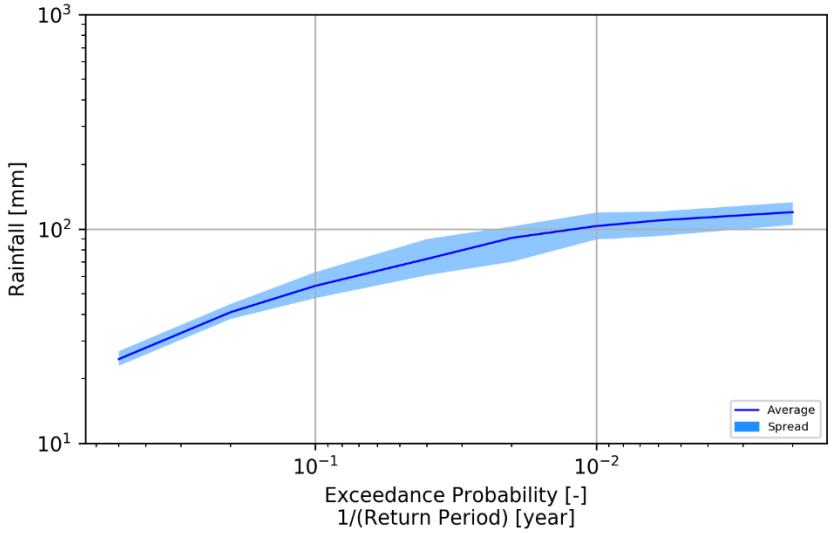

Figure 4. IDF Curve produced by RainyDay for the Big Thompson River watershed upstream of Olympus Dam, in northern Colorado.

Appendix B – Manuscript 1

Rescaling Transposed Extreme Rainfall Within a Heterogeneous Region

Daniel B. Wright, Ph.D.*
University of Wisconsin-Madison, Madison, Wisconsin

Kathleen D. Holman, Ph.D.
Bureau of Reclamation, Denver, Colorado

*danielb.wright@wisc.edu

Abstract

Geospatial transposition of observed storms can be useful for examining potential rainfall and flood hazards, and several recent software packages help facilitate transposition of remote sensing-based rainfall observations from radar or satellites. Two unanswered questions persist, however, in regions that exhibit heterogeneity in extreme rainfall properties: is transposition reasonable, and if so, how should it be performed? This note posits that the answer to the first question depends on the degree of heterogeneity, and attempts to answer the second question via "rescaling" transposed rainfall according to the degree of heterogeneity between two locations. General considerations for this rescaling are discussed and two rescaling methods are introduced. Both methods are illustrated using gage-corrected radar rainfall data from three locations in the Front Range of the Rocky Mountains in Colorado that exhibit moderate heterogeneity in extreme rainfall hydroclimate. The second of these methods, termed stochastic distribution ratio rescaling, is presented here for the first time, and has the advantage of straightforward uncertainty estimation.

1. Introduction

A challenging question in hydrologic engineering is how to best understand potential extreme rainfall and its impacts at a particular location using observations from nearby locations. For example, what lessons can be learned about possible rainfall extremes in Denver, Colorado based on rainfall observations from nearby areas during the "Great Colorado Flood" (see Friedrich et al., 2015; Gochis et al., 2014) of September 2013? Geospatial transposition of observed rainfalls is one approach to answer this question.

The best-known application of storm transposition is in the estimation of probable maximum precipitation (PMP), in which observed rainfall isohyetal patterns are transposed over a given watershed and "rescaled" via maximization of atmospheric moisture content (Hansen et al., 1982; Hansen, 1987; WMO, 2009). Storm transposition has also been increasingly used outside of PMP. Hayden et al. (2016), for example, transposed radar rainfall estimates to assess the efficacy of urban stormwater ordinances in Madison, Wisconsin. Both that study and PMP can be understood as "deterministic storm transposition," in which a rainfall observation or storm is moved from one location directly to the study location.

A second approach is stochastic storm transposition (SST), in which a sample of storms (henceforth referred to as a "storm catalog") is resampled and transposed probabilistically to generate a much larger set of possible rainfall scenarios of varying magnitudes and spatiotemporal properties. First introduced in Alexander (1963), SST was further expanded in Fontaine and Potter (1989), Foufoula-Georgiou (1989), Franchini et al. (1996) Gupta (1972), Wright et al. (2014, 2013), England et al. (2014), and Nathan et al. (2016). SST can be used to estimate rainfall intensity-duration-frequency (IDF) relationships and, when combined with rainfall-runoff models, flood frequency.

Both deterministic and stochastic storm transposition can benefit from rainfall remote sensing observations from ground-based radar or from satellites. These can provide detailed spatiotemporal depictions of severe storms that are generally not available from rain gages. Two recent software tools have been developed that facilitate the coupling of rainfall remote sensing and storm transposition. MetVue from the U.S. Army Corps of Engineers (HEC, 2017) supports editing and analysis of radar data, including transposition and adjustment of rainfall magnitudes (Benson, 2014). RainyDay is an open-source Python-based software and web-based platform aimed at facilitating rainfall and flood frequency analysis using SST and remote sensing data (Wright et al., 2017).

An important question arises in storm transposition: is it reasonable to transpose rainfall from one location to another, and if so, how? If the rainfall hydroclimates in two locations are very similar, transposition between them is straightforward. If there are differences, however, the answer is less clear. Heterogeneity in the precipitation hydroclimate can exist due to topographic features (e.g. Huff and Vogel, 1978; Javier et al., 2007; Kitchen and Blackall, 1992), land-water boundaries (e.g. Ntelekos et al., 2007; Smith et al., 2012), and urban impacts on rainfall (e.g. Shepherd, 2005). This note posits that transposition is reasonable if heterogeneity is not too great, but that it may be appropriate to "rescale" transposed rainfall magnitudes according to the degree of heterogeneity.

This note describes the problem of rescaling transposed rainfall based on differences in rainfall hydroclimatology, and presents two rescaling approaches, one of which is introduced here for the first time. Some important considerations are left to future work, including how to evaluate "acceptable levels" of heterogeneity for which transposition is reasonable and how to use rescaling methods in applications such as rainfall or flood frequency analysis. Basic rescaling considerations are discussed in Section 2. Section 3 introduces the study area, data, and the two rescaling methods. Section 4 shows brief results based on these methods. Concluding remarks follow in Section 5.

2. Rescaling Transposed Rainfall: General Considerations

Several general statements can be made regarding rescaling of transposed precipitation:

1. If the rainfall distributions at two locations *i* and *j* are identical, then no rescaling should be performed when rainfall is transposed from one location to the other. Rescaling should be modest if the rainfall distributions at two locations are similar, and more substantial when they are more different. Very dissimilar distributions, however, indicate that transposition among those two locations may not be appropriate.
2. The most obvious strength of storm transposition is the ability to leverage observations of extreme storms from one location to better understand potential impacts at another location. Thus, extreme rainfall should remain extreme upon transposition and rescaling. This means that the method commonly referred to as quantile mapping or cumulative distribution function (CDF) matching is unsuitable for storm transposition. Quantile matching can be expressed mathematically as:

$$\tilde{x}_j = F_j^{-1}\left(F_i(x_i)\right) \quad (1)$$

   where F_i^{-1} denotes the CDF at location *i*, x_i is the observed value at location *i*, and $\tilde{x}_j$ is the rescaled transposed rainfall. Equation 1 implies a unique mapping between F_i and F_j, and so transposition of x_i will produce a $\tilde{x}_j$ which is already known (via F_j). In other words, quantile mapping negates storm transposition, since transposed rainstorms will never be more extreme than the corresponding observed or estimated quantile at location *j*.
3. While rescaling could take on a variety of mathematical forms, a multiplicative "ratio" formulation is likely desirable, as shown in the two methods described Section 3. Other branches of precipitation science including for remote sensing error characterization (e.g. Tian et al., 2013) and hydrologic model calibration (e.g. Kavetski et al., 2006; Vrugt et al., 2008) show preference for multiplicative over additive approaches.

3. Methods

3.1 Study Area and Rainfall Data

Fig. 1a shows heterogeneity in topography and mean annual precipitation (MAP) across the state of Colorado in the central United States, highlighting three locations that are the focus of subsequent analysis. Heterogeneity in MAP results from influences of the Front Range of the Rocky Mountains on atmospheric circulation and storm tracks in addition to localized orographic effects (e.g. Javier et al., 2007).

The 9-16 September 2013 storm resulted in severe rainfall and flooding throughout the Front Range, though Denver escaped the worst impacts. September 12-13 was the period of most intense rainfall both north of Denver in Boulder and to the south in Fort Carson (Fig.1b). Though Denver was "spared," it is reasonable to surmise that there are lessons that can be learned there via transposition of nearby observations from the 2013 rainstorm or others.

A storm catalog was created using the RainyDay software (Wright et al., 2017) and the Stage IV rainfall dataset (Lin and Mitchell, 2005), which merges ground-based weather radar and rain gages. This storm catalog is comprised of 350 24-hour accumulations over the 2002-2017 period of the Stage IV data record for the single radar pixels centered on the three study locations. It should be emphasized, however, that the methods presented in this note are general and could be applied with point-scale rain gage data or areal estimates of rainfall from Stage IV or other rainfall remote sensing datasets.

CDFs derived from this storm catalog (Fig. 1c) highlight several things. First, the rainfall distributions for Denver and Boulder are virtually identical except for the extreme tail: the largest 24-hour rainfall accumulation in Denver is of 59 mm on 30-31 August 2016, much smaller than the 179 mm accumulation for 12-13 September 2013 in Boulder. We can conclude that Denver and Boulder have essentially the same rainfall hydroclimates, and that the difference in the extreme tail is likely an artifact of the relatively short (16-year) record length. Thus, it is likely reasonable to transpose the 2013 event from Boulder to Denver. Fort Carson experienced even heavier 24-hour rainfall (309 mm). The CDF for Fort Carson reveals a somewhat wetter rainfall hydroclimate, which casts some doubt on the validity of transposing rainfall to Denver, at least without rescaling.

3.2 Index Quantile Rescaling

One approach to rescaling rainfall transposed from location *i* to location *j* involves selecting an "index quantile" (I^q) and rescaling by its ratio:

$$\tilde{x}_j = \frac{I_j^q}{I_i^q} x_i. \qquad (2)$$

The quantile q could pertain to a given exceedance probability from a distribution of annual maxima (i.e. from the IDF curve) or to the "full" rainfall distribution. I_A^q / I_B^q depends on the choice of q and thus on the relative "shapes" of F_i and F_j. Furthermore, the ratio will be sensitive to errors in quantile estimates, which can be substantial. Nathan et al., (2016), for example, use the 50-year rainfall quantile, which they justify because of their focus on very rare events. It is well-known, however, that such quantile estimates are subject to substantial uncertainty, and it may be preferable to use less uncertain index quantiles such as the 2-year or 5-year rainfall.

3.2 Stochastic Ratio Distribution Rescaling

This section introduces an approach to stochastically rescale transposed rainfall, based on the concept of a ratio distribution, i.e. the distribution of $R = Y/X$, where X and Y are random variables. Commonly-used ratio distributions include the F-distribution (the ratio of two scaled independent Chi-squared distributions) and the Cauchy distribution (the ratio of two independent standard normal distributions). Other ratio distributions have been derived for certain marginals, parameter values, and correlations (e.g. Lee et al., 1979; Nadarajah, 2010; Nadarajah and Kotz, 2006). Few of these are suitable for describing rainfall, and parameter estimation can be challenging. An exception is if both X and Y are lognormally distributed (often a reasonable approximation for rainfall; e.g. Atlas et al., 1990; Shimizu, 1993; Zhang and Singh, 2007), in which case it is straightforward to derive the distribution of R for any parameter set and correlation. Let $X = X_i$ and $Y = X_j$, the lognormally-distributed rainfall at locations *i* and *j* with

linear correlation $\rho_{X,Y}$. Thus:

$$logR = log\left(\frac{Y}{X}\right) = logY - logX. \quad (3)$$

By definition, $logX$ and $logY$ are normally distributed with parameters $\{\mu_X, \sigma_X^2\}$ and $\{\mu_Y, \sigma_Y^2\}$. Via the additive properties of normal distributions, $logR$ is normally distributed with parameters $\mu_R = \mu_Y - \mu_X$ and $\sigma_R^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho_{X,Y}\sigma_X\sigma_Y$. Thus, $R \sim Lognormal(\mu_R, \sigma_R^2)$, $E[R] = \exp(\mu_R + \frac{1}{2}\sigma_R^2)$, and $Var[R] = \exp(2\mu_R + 2\sigma_R^2) - \exp(2\mu_R + \sigma_R^2)$. This ratio can then be used for rescaling:

$$\tilde{x}_j = Rx_i. \quad (4)$$

If *X* and *Y* are identically distributed then $E[R] = 1$, $Var[R] = 0$, and no rescaling is performed. If the distributions of *X* and *Y* are similar but not identical then $E[R]$ will be close to 1.0 and $Var[R]$ will be small. If the distributions are very different, then $E[R]$ may differ from 1.0 and $Var[R]$ may be relatively large. Second, unlike quantile mapping, rescaling "shifts" the distribution $\tilde{x}_j$ but leaves the extreme tail estimates intact, as shown in Section 4.

The derivation above is included for completeness. In the context of storm transposition, it is not necessary to compute marginal distributions for *X* and *Y*, nor their correlation, since the empirical distribution $\hat{R}$ can be estimated directly from their empirical distributions:

$$\hat{R}_k = \frac{y_k}{x_k}, k = 1..n \quad (4)$$

where y_k and x_k are the k^{th} ordered observations in the empirical CDF of *Y* and *X*, respectively. Any suitable continuous distribution can then be fitted to $\hat{R}$. Though this study uses the lognormal distribution, other distributions including the normal distribution could also be appropriate.

4. Rescaling Results

Estimates of $\hat{R}$ (Fig. 2) are computed from the CDFs shown in Fig. 1c. Results are illustrated only for single radar "pixels" centered on the three study locations, but the technique could be applied to larger geospatial storm patterns as well. Fig. 2 shows that $E[\hat{R}]$ is very close to 1.0 for the ratio of Denver to Boulder rainfall, consistent with the close resemblance in the CDFs shown in Fig. 1c. $E[\hat{R}]$ of the ratio of Denver to Fort Carson rainfall is 0.82, reflecting Fort Carson's wetter hydroclimate than Denver and thus rainfall from Fort Carson should be "scaled down" prior to transposition to Denver under this framework. The 90% interquantile range for these two ratio distributions are similar since, despite Fort Carson being wetter than the other two locations (Fig. 1c), this "shift" is relatively consistent across the range of rainfall values. If the shift varied significantly with rainfall magnitude, $Var[\hat{R}]$ would be greater. The distributions of $\hat{R}$ in Fig. 2 provide a basis for rescaling under conditions of repeated transposition (i.e. SST) or for bounding the rescaling of transposed rainfall under deterministic transposition, as demonstrated below.

Index quantile-based rescaling ratios calculated using estimates of the 50-year 24-hour rainfall from Atlas 14 (Bonnin et al., 2004) from the National Oceanic and Atmospheric Administration (NOAA) are also shown in Fig. 2. These ratios differ from those generated by the stochastic ratio

distributions: the index quantile method scales Boulder rainfall downward rather than upward ($I^{50-yr}_{Denver}/I^{50-yr}_{Boulder} = 0.86$, while $E[\hat{R}] = 1.02$). Differences are smaller for Fort Carson rainfall ($I^{50-yr}_{Denver}/I^{50-yr}_{Ft.\ Carson} = 0.90$; $E[\hat{R}] = 0.82$). Whether these discrepancies are due to some shortcoming of one or both methods, or due simply to differences in their derivations, is not clear.

NOAA Atlas 14 also publishes 90% confidence intervals (CIs) for IDF quantiles. If one assumes that these CIs are drawn from a normal distribution, it is possible to infer the distribution of I^q (note that the published CIs are not symmetric; this normality assumption is thus only approximate). If one further assumes that the quantile estimation uncertainties at locations i and j are independent, one can use parametric bootstrapping to estimate the distribution of I^q_i/I^q_j. These results are included in Fig. 2. Unlike the distributions of $\hat{R}$, the distributions of I^q_i/I^q_j bear little resemblance to the data, with unrealistically high variance. This is likely the result of the independence assumption between the quantile uncertainties—correlation between errors in I^q_i and I^q_j will reduce $Var(I^q_i/I^q_j)$. Such correlation likely exists, since nearby locations will be subject to similar sampling errors and thus their estimation uncertainties will be correlated. When regionalization is used (as in Atlas 14), the independence assumption becomes even more suspect. A more complex bootstrap approach on the original observations would likely be needed to properly estimate uncertainty in I^q_i/I^q_j.

The derived distributions of $\hat{R}$ shown in Fig. 2 are used to generate rescaled 24-hour rainfall distributions (Fig. 3). Results from index quantile rescaling are also shown. Consistent with Fig. 1, the rainfall distributions for Denver and Boulder (upper panels) are relatively similar, while the curve for Denver is below that of Fort Carson (lower panels) for all probabilities. The shaded areas indicate the 95% interval of stochastic rescaling achieved through 1000 random samples from the corresponding distribution of $\hat{R}$ applied to rescale the distribution of transposed rainfall. When used with the Boulder rainfall distribution (Fig. 3, top panels), index quantile rescaling appears to underestimate the Denver rainfall distribution for all but the highest quantiles, since $I^{50-}_{Denver}/I^{50-y}_{Boulder}$ scales the distribution downward. CIs are omitted for the index quantile rescaling approach due to the difficulty in producing realistic estimates (see Fig. 2 and related discussion).

Of particular interest is the tail behavior (i.e. above the 95th percentile), which has been emphasized in the right-hand panels of Fig. 3. The extreme tails of the transposed rainfall from both Boulder and Fort Carson differ substantially from that of Denver due to the 12-13 September 2013 rainfall. Stochastic ratio rescaling produces a mean 24-hour rainfall estimate of 183 mm for the 12-13 September rainfall transposed from Boulder to Denver, with the spread of the 1000 random draws ranging from 149-221 mm. The mean result is 254 mm for the 2013 rainfall from Fort Carson transposed to Denver, with a range from 188-325 mm. Index quantile rescaling results in 154 mm (279 mm) for the transposed Boulder (Fort Carson) rainfall. For reference, the 100-year (1000-year) 24-hour rainfall depth for Denver provided by NOAA Atlas 14 is 120 mm (175 mm), with a 90% CI of 90-155 mm (118-236 mm).

5. Conclusions

This note explores how extreme rainfall observations can be "rescaled" when transposed between two locations that have heterogeneous rainfall properties. The degree of heterogeneity between rainfall distributions should determine the degree of rescaling. Two rescaling approaches are introduced. Index Quantile Rescaling is feasible where rainfall IDF curves are available, but is subject to IDF estimation uncertainties and can be sensitive to the choice of index quantile. A new approach, based on ratio distributions, is introduced that considers differences in the means and variances of the rainfall at the two locations. This methodology assumes that there is a consistent "rescaling relationship" across a wide range of rainfall values, and that observed deviations from this assumption arise from limited samples sizes. Long rainfall records at both locations could allow one to test this assumption. While Section 4 presents results based on distributions of 350 rainfall values, it could also be performed on distributions of annual rainfall maxima. Unlike Index Quantile Rescaling, it is straightforward to produce uncertainty estimates under the ratio distribution approach. This feature is useful for stochastic applications of storm transposition.

Both approaches could prove useful in conjunction with the MetVue software from the U.S. Army Corps of Engineers, which facilitates transposition of radar rainfall observations for hydrologic engineering purposes. The authors intend to implement the latter approach within the RainyDay stochastic storm transposition software (Wright et al., 2017) to enhance the software's performance for rainfall and flood frequency analysis in heterogeneous regions.

A more fundamental question with storm transposition is: "should transpose extreme rainfall observations?". Rescaling transposed rainfall can be used to cope with some degree of heterogeneity, beyond which there is little credibility. It seems unreasonable, for example, to transpose rainfall from an arid region to a wet one or vice versa. It is likely that guidance on the degree of heterogeneity over which transposition and rescaling could be derived from the ratio distribution approach. This will be a topic of future investigation. Existing techniques for characterizing heterogeneity in rainfall extremes such as regionalized L-moments (e.g. Guttman, 1993) could also be brought to bear on questions of where and how to transpose.

Acknowledgements

Wright and Holman are partially supported by the Bureau of Reclamation Research and Development Office Project ID 1735: *Development of a Stochastic Storm Transposition Toolkit for Physically-based Rainfall and Flood Hazard Analysis*.

References

Alexander, G.N., 1963. Using the probability of storm transposition for estimating the frequency of rare floods. J. Hydrol. 1, 46–57. https://doi.org/10.1016/0022-1694(63)90032-5

Atlas, D., Rosenfeld, D., Short, D.A., 1990. The estimation of convective rainfall by area integrals: 1. The theoretical and empirical basis. J. Geophys. Res. Atmospheres 95, 2153–2160. https://doi.org/10.1029/JD095iD03p02153

Benson, S.A., 2014. Comparative analysis of flood frequency based on radar-based precipitation data and precipitation trends. University of Texas at Arlington.

Bonnin, G.M., Martin, D., Lin, B., Parzybok, T., Riley, D., 2004. NOAA Atlas 14: Precipitation-Frequency Atlas of the United States.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. 28, 2031–2064. https://doi.org/10.1002/joc.1688

England, J.F., Julien, P.Y., Velleux, M.L., 2014. Physically-based extreme flood frequency with stochastic storm transposition and paleoflood data on large watersheds. J. Hydrol. 510, 228–245. https://doi.org/10.1016/j.jhydrol.2013.12.021

Fontaine, T.A., Potter, K.W., 1989. Estimating Probabilities of Extreme Rainfalls. J. Hydraul. Eng. 115, 1562–1575. https://doi.org/10.1061/(ASCE)0733-9429(1989)115:11(1562)

Foufoula-Georgiou, E., 1989. A probabilistic storm transposition approach for estimating exceedance probabilities of extreme precipitation depths. Water Resour. Res. 25, 799–815. https://doi.org/10.1029/WR025i005p00799

Franchini, M., Helmlinger, K.R., Foufoula-Georgiou, E., Todini, E., 1996. Stochastic storm transposition coupled with rainfall—runoff modeling for estimation of exceedance probabilities of design floods. J. Hydrol. 175, 511–532. https://doi.org/10.1016/S0022-1694(96)80022-9

Friedrich, K., Kalina, E.A., Aikins, J., Gochis, D., Rasmussen, R., 2015. Precipitation and Cloud Structures of Intense Rain during the 2013 Great Colorado Flood. J. Hydrometeorol. 17, 27–52. https://doi.org/10.1175/JHM-D-14-0157.1

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., 2014. The Great Colorado Flood of September 2013. Bull. Am. Meteorol. Soc. 96, 1461–1487. https://doi.org/10.1175/BAMS-D-13-00241.1

Gupta, V.K., 1972. Transposition of Storms for Estimating Flood Probability Distributions.

Guttman, N.B., 1993. The Use of L-Moments in the Determination of Regional Precipitation Climates. J. Clim. 6, 2309–2325. https://doi.org/10.1175/1520-0442(1993)006<2309:TUOLMI>2.0.CO;2

Hansen, E.M., 1987. Probable maximum precipitation for design floods in the United States. J. Hydrol. 96, 267–278. https://doi.org/10.1016/0022-1694(87)90158-2

Hansen, E.M., Schreiner, L.C., Miller, J.F., 1982. Application of Probable Maximum Precipitation Estimates-United States East of the 105th Meridian. Washington, D. C.

Hayden, N.G., Potter, K.W., Liebl, D.S., 2016. Evaluating Infiltration Requirements for New Development Using Extreme Storm Transposition: A Case Study from Dane County, WI. JAWRA J. Am. Water Resour. Assoc. 52, 1170–1178. https://doi.org/10.1111/1752-1688.12441

HEC, 2017. Real-Time Simulation (HEC-RTS) User's Manual-Draft Version 3. U.S. Army Corps of Engineers, Davis, CA.

Huff, F.A., Vogel, J.L., 1978. Urban, Topographic and Diurnal Effects on Rainfall in the St. Louis Region. J Appl Meteor 17, 565–577. https://doi.org/10.1175/1520-0450(1978)017<0565:UTADEO>2.0.CO;2

Javier, J.R.N., Smith, J.A., England, J., Baeck, M.L., Steiner, M., Ntelekos, A.A., 2007. Climatology of extreme rainfall and flooding from orographic thunderstorm systems in the upper Arkansas River basin. Water Resour. Res. 43, n/a-n/a. https://doi.org/10.1029/2006WR005093

Kavetski, D., Kuczera, G., Franks, S.W., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resour. Res. 42, n/a-n/a. https://doi.org/10.1029/2005WR004368

Kitchen, M., Blackall, R.M., 1992. Orographic rainfall over low hills and associated corrections to radar measurements. J. Hydrol. 139, 115–134. https://doi.org/10.1016/0022-1694(92)90198-5

Lee, R.-Y., Holland, B.S., Flueck, J.A., 1979. Distribution of a Ratio of Correlated Gamma Random Variables. SIAM J. Appl. Math. 36, 304–320.

Lin, Y., Mitchell, K.E., 2005. The NCEP Stage II/IV hourly precipitation analyses: development and applications, in: Preprints, 19th Conf. on Hydrology, American Meteorological Society, San Diego, CA, 9-13 January 2005, Paper 1.2. pp. 2–5.

Nadarajah, S., 2010. Distribution properties and estimation of the ratio of independent Weibull random variables. AStA Adv. Stat. Anal. 94, 231–246. https://doi.org/10.1007/s10182-010-0134-1

Nadarajah, S., Kotz, S., 2006. ON THE PRODUCT AND RATIO OF GAMMA AND WEIBULL RANDOM VARIABLES. Econom. Theory 22, 338–344. https://doi.org/10.1017/S0266466606060154

Nathan, R., Jordan, P., Scorah, M., Lang, S., Kuczera, G., Schaefer, M., Weinmann, E., 2016. Estimating the exceedance probability of extreme rainfalls up to the probable maximum precipitation. J. Hydrol. 543, 706–720. https://doi.org/10.1016/j.jhydrol.2016.10.044

Ntelekos, A.A., Smith, J.A., Krajewski, W.F., 2007. Climatological analyses of thunderstorms and flash floods in the Baltimore Metropolitan region. J. Hydrometeorol. 8, 88–101. https://doi.org/10.1175/JHM558.1

Shepherd, J.M., 2005. A Review of Current Investigations of Urban-Induced Rainfall and Recommendations for the Future. Earth Interact. 9, 1–27. https://doi.org/10.1175/EI156.1

Shimizu, K., 1993. A Bivariate Mixed Lognormal Distribution with an Analysis of Rainfall Data. J. Appl. Meteorol. 32, 161–171. https://doi.org/10.1175/1520-0450(1993)032<0161:ABMLDW>2.0.CO;2

Smith, J.A., Baeck, M.L., Villarini, G., Welty, C., Miller, A.J., Krajewski, W.F., 2012. Analyses of a Long-Term High-Resolution Radar Rainfall Data Set for the Baltimore Metropolitan Region. Water Resour Res 48, W04504–W04504. https://doi.org/10.1029/2011WR010641

Tian, Y., Huffman, G.J., Adler, R.F., Tang, L., Sapiano, M., Maggioni, V., Wu, H., 2013. Modeling errors in daily precipitation measurements: Additive or multiplicative? Geophys. Res. Lett. 40, 2060–2065. https://doi.org/10.1002/grl.50320

Vrugt, J.A., ter Braak, C.J.F., Clark, M.P., Hyman, J.M., Robinson, B.A., 2008. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. Water Resour. Res. 44, n/a-n/a. https://doi.org/10.1029/2007WR006720

WMO, 2009. Manual on estimation of Probable Maximum Precipitation (PMP) (No. WMO-No. 1045). World Meteorological Organization.

Wright, D.B., Mantilla, R., Peters-Lidard, C.D., 2017. A remote sensing-based tool for assessing rainfall-driven hazards. Environ. Model. Softw. 90, 34–54. https://doi.org/10.1016/j.envsoft.2016.12.006

Wright, D.B., Smith, J.A., Baeck, M.L., 2014. Flood frequency analysis using radar rainfall fields and stochastic storm transposition. Water Resour. Res. 50, 1592–1615. https://doi.org/10.1002/2013WR014224

Wright, D.B., Smith, J.A., Villarini, G., Baeck, M.L., 2013. Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. J. Hydrol. 488, 150–165. https://doi.org/10.1016/j.jhydrol.2013.03.003

Zhang, L., Singh, V.P., 2007. Bivariate rainfall frequency distributions using Archimedean copulas. J. Hydrol. 332, 93–109. https://doi.org/10.1016/j.jhydrol.2006.06.033

Figures

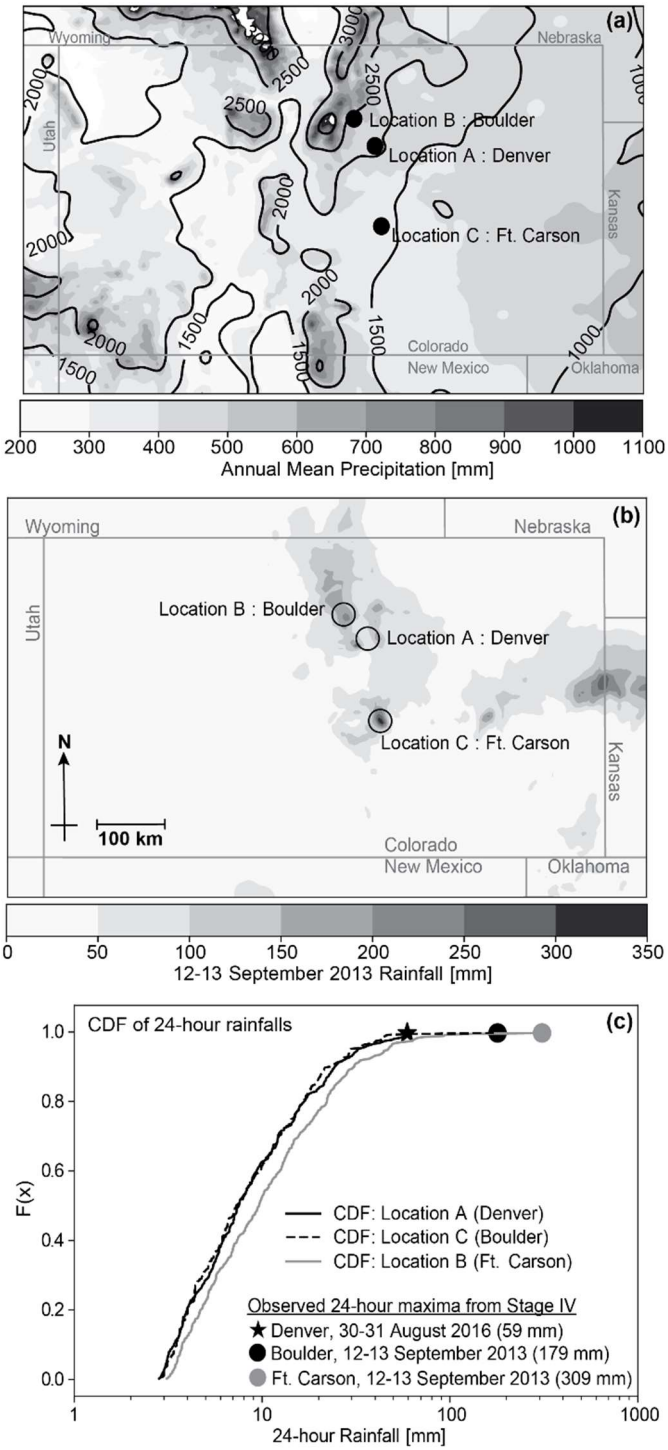

Fig. 1: (a) Annual mean precipitation (shading) and elevation (contour lines) over Colorado. The three locations of interest are also shown. Precipitation and elevation data are taken from the PRISM dataset (Daly et al., 2008) between 1981-2010. (b) Rainfall for 12-13 September 2013 derived from Stage IV radar rainfall dataset with the three study locations circled. (c) CDFs of the 350 largest 24-hour precipitation accumulations from Stage IV for 2002-2017 for the three study locations.

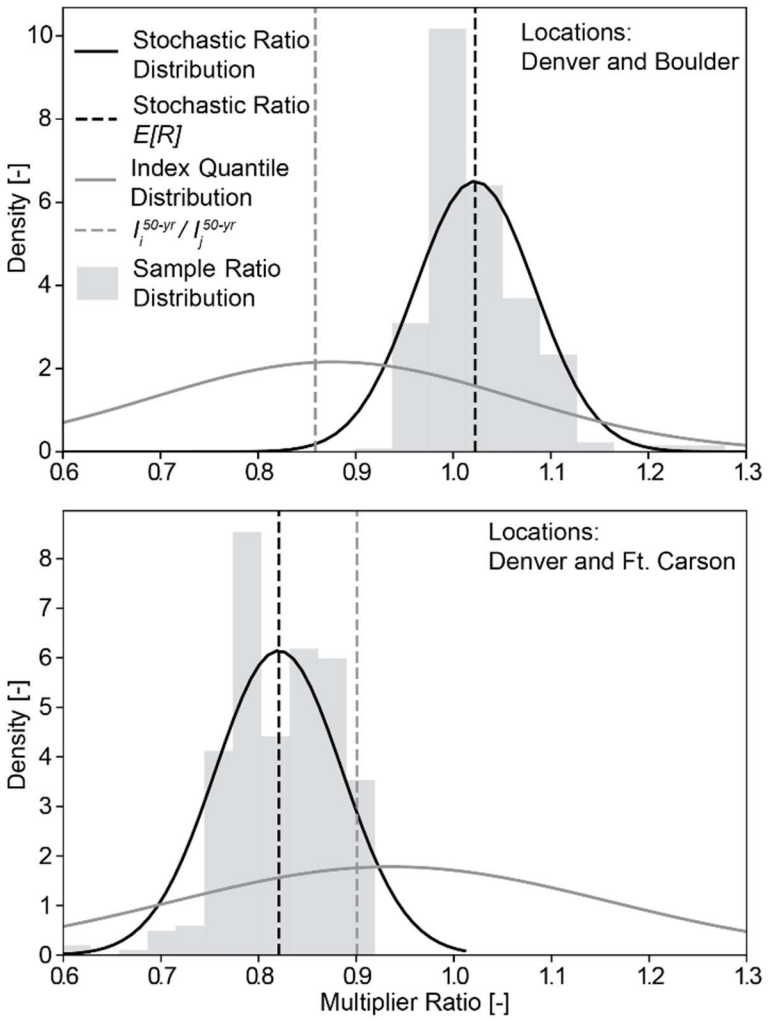

Fig. 2: Sample and theoretical PDFs for ratio distributions for Denver and Boulder (top) and locations Denver and Fort Carson (bottom). The ratios I_i^q / I_j^q calculated via index quantile rescaling based on the mean 50-year 24-hour rainfall estimates from NOAA's Atlas 14 as well as the distributions of I_i^q / I_j^q approximated via parametric bootstrapping are also shown.

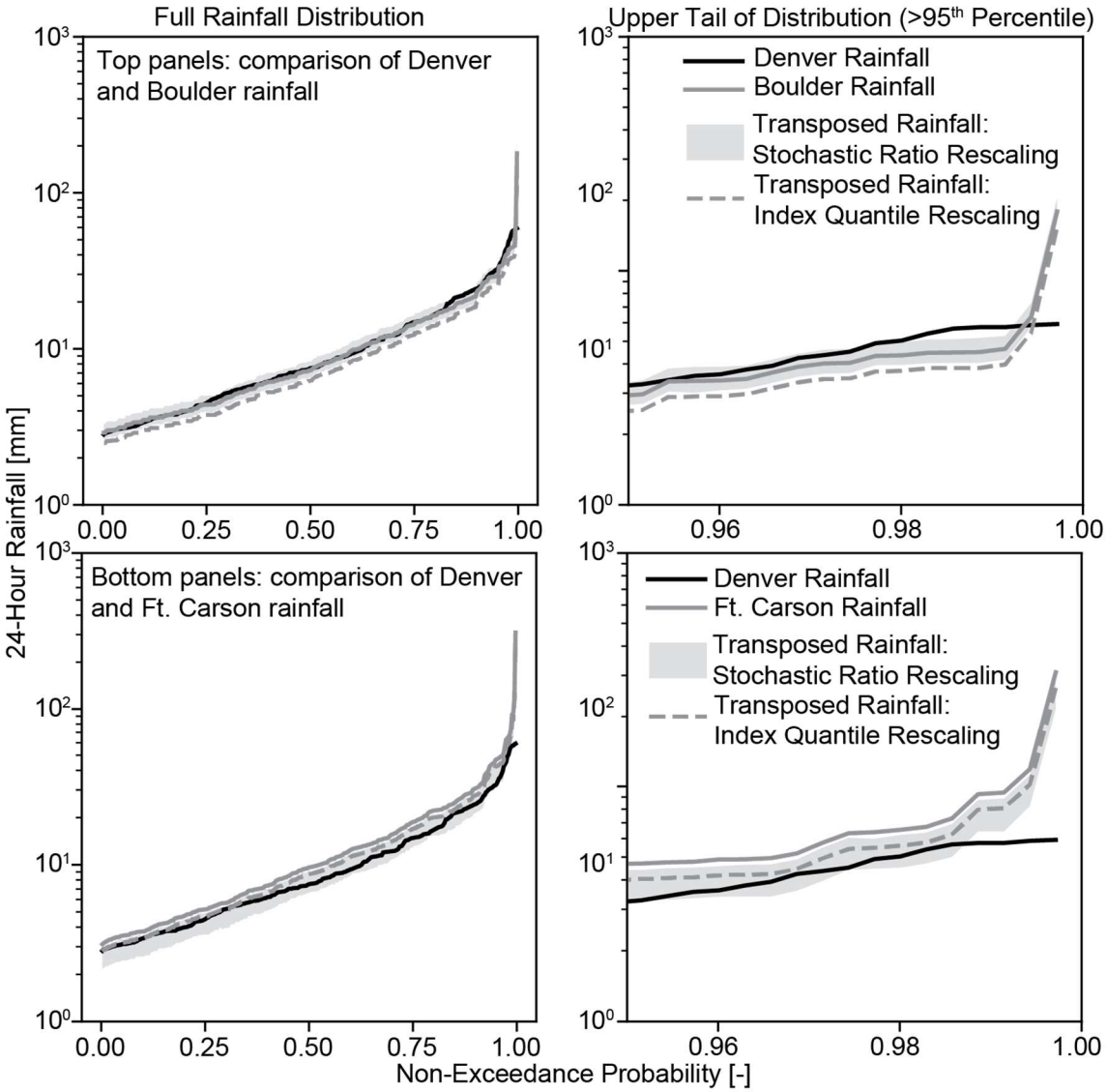

Fig. 3. Upper panels—24-hour rainfall distributions for Denver and Boulder as well as stochastic ratio rescaling and index quantile rescaling (using the 50-year quantile) of the Boulder distribution transposed to Denver. Lower panels—same as upper panels, but for Denver and Fort Carson. Left panels show the full distribution; right panels highlight the upper tail above the 95th percentile. Note that these are distributions of nonzero rainfall, rather than of annual rainfall maxima.

Appendix C – Manuscript 2

Process-Based Flood Frequency Analysis in an Agricultural Watershed Exhibiting Nonstationary Flood Seasonality

Guo Yu and Daniel Wright
University of Wisconsin-Madison, Madison, Wisconsin

Zhihua Zhu*
Sun Yat-sen University, Guangzhou, China

Cassia Smith
Carnegie Mellon University, Pittsburgh, Pennsylvania

Kathleen D. Holman
Bureau of Reclamation, Denver, Colorado

* zzhu264@wisc.edu

Abstract

Floods are the product of complex interactions among processes including precipitation, soil moisture, and watershed morphology. Conventional flood frequency analysis (FFA) methods such as design storms and discharge-based statistical methods offer few insights into these process interactions and how they "shape" the probability distributions of floods. Understanding and projecting flood frequency in conditions of nonstationary hydroclimate and land use requires deeper understanding of these processes, some or all of which may be changing in ways that will be undersampled in observational records. This study presents an alternative "process-based" FFA approach that uses stochastic storm transposition to generate large numbers of realistic rainstorm "scenarios" based on relatively short rainfall remote sensing records. Long-term continuous hydrologic model simulations are used to derive seasonally varying distributions of watershed antecedent conditions. We couple rainstorm scenarios with seasonally appropriate antecedent conditions to simulate flood frequency. The methodology is applied to the 4002 km2 Turkey River watershed in the Midwestern United States, which is undergoing significant climatic and hydrologic change. We show that using only 15 years of rainfall records, our methodology can produce accurate estimates of "present-day" flood frequency. We found that shifts in the seasonality of soil moisture, snow, and extreme rainfall in Turkey River exert

important controls on flood frequency. We also demonstrate that process-based techniques may be prone to errors due to inadequate representation of specific seasonal processes within hydrologic models. If such mistakes are avoided, however, process-based approaches can provide a useful pathway toward understanding current and future flood frequency in nonstationary conditions and thus be valuable for supplementing existing FFA practices.

1. Introduction

Riverine floods, among the most common natural disasters worldwide, are the product of complex interactions between heavy rainfall, watershed and river channel morphology, and antecedent (i.e. initial) conditions including soil moisture and snowpack. Their impacts are projected to increase in the future due to hydrometeorological factors (e.g. Hyndman, 2014) and increased human development in flood prone areas (e.g. Ntelekos et al., 2010; Ceola et al., 2014; Prosdocimi et al., 2015). Estimating the relationships between flood likelihood and severity is central to flood risk management and infrastructure design; these relationships are typically represented by flood frequency distributions (or curves), while the broad family of procedures used to derive them is termed flood frequency analysis (FFA). Most existing FFA methods belong to one of three approaches: statistical analysis of streamflow observations, design storms, and continuous simulation or other so-called "derived" or "process-based" methods. Each has strengths and shortcomings, which are briefly summarized in Sect. 2 (see Wright et al., 2014a for a distinct summary).

FFA is challenging even in stationary (i.e. unchanging) watershed and hydroclimatic conditions due to the scarcity of observations of large floods and the associated factors that generate them (Stedinger and Griffis, 2011). The role of soil moisture in flood frequency, for example, is very important (Berghuijs et al., 2016), but poorly understood due to a lack of long-term observations. Furthermore, the individual and joint flood causative factors will evolve as a watershed undergoes changes in land use or hydroclimate (Machado et al., 2015). Leading causes of change (i.e. nonstationarity) include human intervention through land use change or reservoir construction (Konrad and Booth, 2002; Schilling and Libra, 2003; Villarini et al., 2009), natural climate variability (Enfield et al., 2001; Jain and Lall, 2000) and anthropogenic climate change driven by increasing greenhouse gas concentrations (Milly et al., 2008; Hirsch and Ryberg, 2012). Combinations of these will lead to nonstationary flood frequency, a challenge for which the bulk of existing FFA methods are ill-suited (El Adlouni et al., 2007; Gilroy and McCuen, 2012).

In this study, we present an alternative FFA methodology that aims to "construct" the flood frequency curve through a combination of observations, stochastic methods, and hydrologic modeling that generates and combines the causative factors (i.e. processes) such as rainfall and soil moisture that produce floods. This concept is not new, and has traditionally be called "derived FFA" (e.g. Eagleson, 1972; Franchini et al., 2005; Haberlandt, 2008), though we prefer the more descriptive term "process-based FFA" (after Sivapalan and Samuel, 2009; see Clark et al., 2015a, 2015b and Lamb et al., 2016; who discuss somewhat similar techniques). Sivapalan and Samuel (2009) argue in favor of process-based approaches in the face of nonstationary conditions, though they do not actually lay out a specific FFA procedure.

We present such a process-based procedure, and apply it to an agricultural watershed in the Midwestern United States that is undergoing substantial seasonal hydroclimatic and hydrologic changes that have led to nonstationary flood frequency. We show that this procedure is useful for deciphering the underlying physical processes that drive flooding, as well as their changes in this watershed. Our methodology underscores the importance of seasonality in the joint contributions of rainfall, soil moisture, and snow to flood frequency. To our knowledge, this study is the first to explore the role that seasonal changes in hydroclimatic and hydrologic processes play in nonstationary flood frequency, though other studies have explored the importance of such processes in flood occurrence more generally (e.g. Berghuijs et al., 2016).

The structure of the paper is as follows: Section 2 briefly reviews the three aforementioned FFA approaches. Section 3 introduces the study region, watershed, and hydrometeorological data. Section 4 outlines the process-based FFA methodology used in this study, including the hydrologic model, the stochastic storm transposition (SST) procedure used to derive the synthetic rainfall scenarios, and elements of both continuous and event-based rainfall-runoff simulation. The nonstationary hydroclimate of the study watershed and trends in relevant hydrometeorological variables are analyzed in Sect. 5.1. Model validation is presented in Sect. 5.2. Process-based FFA results are presented and compared with "conventional" statistical estimates in Sect. 5.3. Simulated flood seasonality is explored in Sect. 5.4. The relationships between rainfall and simulated peak discharge quantiles are examined in Sect. 5.5. Section 6 includes a summary and concluding remarks.

2. Review of FFA Approaches

2.1 Discharge-based Statistical Approaches

Statistical FFA approaches involve fitting a statistical distribution to extreme discharge observations and extrapolating this distribution to estimate quantiles such as the 100-year or 500-year discharge. While these approaches utilize direct observations of flooding (e.g. peak discharge or volume), long streamflow records at or near the given river cross section are needed for reliable quantile estimates. Such records are lacking in many locations, even in developed countries. Statistical approaches are limited by the available observations; thus, the estimation distribution may not represent the "true" (unknown) distribution of possible outcomes (Linsley, 1986; Klemeš, 1986, 2000a, 2000b). In principle, regionalized FFA methods are able to improve quantile estimates both at gaged and ungauged locations (Dawdy et al., 2012); they make assumptions, however, regarding the transferability of regional information to specific locations and in doing so may neglect key geophysical processes that dominate the spatiotemporal variability of floods (Ayalew and Krajewski, 2017).

Though streamflow observations are the result of a range of complex factors including rainfall, soil moisture, and channel routing, without concurrent observations of these "upstream" variables, neither streamflow observations nor distributions fitted to them provide much insight into flood causes. Long-term records of such variables, particularly soil moisture, are virtually nonexistent. There have been numerous examples within the FFA literature pointing to situations in which discharge-based analyses can be inferior to those based on hydrologic modeling, including cases of basin storage "discontinuities" (Rogger et al., 2012), reservoirs (Ayalew et al., 2013), and land use change (Cunha et al., 2011).

Finally, most statistical FFA methods assume that the magnitude of extreme flood events and quantiles are stationary. This assumption conflicts with numerous examples in which hydrological records exhibit various types of nonstationarity (e.g. Potter, 1976; Villarini et al., 2009; Douglas et al., 2000; Franks and Kuczera, 2002). Though nonstationary statistical FFA techniques do exist (e.g. Cheng et al., 2014; Gilleland and Katz, 2016; Serago and Vogel, 2018), they face severe limitations extrapolating to future conditions (Luke et al., 2017; Sivapalan and Samuel, 2009; Stedinger and Griffis, 2011) since they rarely consider the fundamental physical causes of change.

2.2 Design Storm Approaches

Design storm (DS) approaches use idealized rainfall scenarios of a given return period as inputs to a hydrologic model to simulate flood peaks. DS is widely used in practice due to its simplicity (Cudworth, 1989; Kjeldsen, 2007; Ball et al., 2016). To some extent, the flood-producing physical processes are captured via the hydrologic model, which also provides a complete simulated flood hydrograph, as opposed to only the peak discharge or volume provided by statistical approaches. However, DS approaches rely on at least three major assumptions: (1) point-based rainfall intensity-duration-frequency (IDF) estimates (which are subject to some of the same aforementioned statistical and data availability issues as flood discharges) can be converted into hyetographs using dimensionless temporal rainfall distributions and into basin-averaged estimates using area reduction factors (e.g. Svensson and Jones, 2010); (2) IDF estimates, based on annual rainfall maxima, produce flood peaks which are quantiles of the distributions of flood annual maxima; and (3) there is a 1:1 equivalence between rainfall and simulated discharge quantiles (i.e. return periods or recurrence intervals), for example, a 100-year idealized rainfall event will produce a reasonable estimate of the 100-year peak discharge. The last of these assumptions discounts the possibility that watershed initial conditions such as soil moisture and snowpack can modulate the transformation of rainfall quantiles into discharge quantiles.

These assumptions are not without their shortcomings. Wright et al. (2014b), for example, showed significant disparities between observed point and basin-averaged rainfall extremes that cannot be captured using conventional ARF concepts. Using design storm in conjunction with a derived distribution approach, Viglione and Blöschl (2009) and Vigligone et al. (2009) demonstrated that the ratio of rainfall return period to flood peak return period is controlled by storm duration, a runoff coefficient (which is related to antecedent conditions), and a runoff threshold effect. Antecedent conditions can vary substantially by season, meaning that high soil moisture may only infrequently coincide with extreme rainfall. Wright et al. (2014a) discusses additional design storm shortcomings in greater detail, including time of concentration concepts, while also pointing out that design storm approaches (like other hydrologic model-based FFA) can incorporate future projections in land use and rainfall more explicitly than can statistical discharge-based methods.

2.3 Continuous Simulation and Process-Based FFA Approaches

Continuous simulation (CS) and process-based approaches to FFA leverage the potential benefits of hydrologic models while minimizing the simplifying assumptions of DS methods. CS approaches typically use long series of historical or stochastically generated rainfall, temperature, and occasionally other meteorological variables as model inputs, to simulate long discharge time series. Peak flows can be extracted from these series and the flood frequency distribution can be obtained. Thus, event rainfall return period and duration and antecedent conditions do not need to be specified and the equality between rainfall and discharge return period is not assumed (Calver et al., 1999, 2009). In addition, projections of future flood frequency can be developed by incorporating general circulation model (GCM) rainfall and temperature projections into the input meteorological series (Gilroy and McCuen, 2012; Rashid et al., 2017). On the other hand, CS approaches are limited by the general lack of reliable long-term time series of extreme rainfall and other meteorological data (Blazkova and Beven, 1997, 2002, 2009) and, in the case of sophisticated distributed approaches, by potentially high computational demands (Li et al., 2014; Peleg et al., 2017). Stochastic rainfall generation techniques typically struggle to produce the extremes that are critical for flooding (e.g. Cameron et al., 2000; Furrer and Katz, 2008), and training such models for locations with rainfall nonstationarities and strong seasonal variations is nontrivial. Camici et al. (2011) and Li et al. (2014) present process-based FFA approaches that couple long CS simulation results with event-based simulations.

One argument in favor of CS and process-based approaches is that the complex joint relationships between flood drivers such as rainfall and soil moisture are resolved within the modeling framework and thus do not rely on users' assumptions. We demonstrate that caution is needed in the representation of seasonality; to briefly summarize, it is critical that both seasonality in input variables as well as seasonally varying processes within the model be "correct." Without verifying this, process-based approaches may produce seemingly correct results as a result of incorrect methods.

3. Study Region and Data

The study watershed of Turkey River is situated in northeastern Iowa (Fig. 1a, 1b). The portion upstream of the US Geological Survey (USGS) stream gage at Garber (gage number 05412500) has a drainage area of 4002 km^2, with elevations ranging from approximately 426 m above sea level (masl) in the west to 197 masl at the stream gage (Fig. 1c). Streams in the upper part of the catchment have relatively mild slopes, while the channels and hillslopes in the lower part are steeper. Soils are mainly loams and silts (IFC, 2014). According to USGS 2012 National Land Cover Dataset (NLCD), the Turkey River watershed is predominantly agricultural, with less than 2% urban land cover (Fig. 1d). Comparisons of NLCD from 1992, 2001, 2006, and 2012 indicate that land uses have not evolved significantly over time (results not shown), though the hydrologic impacts of subsurface tile drainage, which has become ubiquitous throughout the region, are poorly understood and could exert meaningful influence on flooding (see, e.g. Schilling et al., 2014).

We use daily discharge observations for 84 years (1933-2016) from the USGS streamgage at Garber to understand the hydroclimatology of flooding and to validate our FFA results. Daily discharge observations for 69 years (1948-2016), in conjunction with Global Historical Climate Network (GHCN) daily temperature and snow data are used to configure, calibrate, and validate the hydrologic model, as described in Sect. 4.1. CPC US Unified (CPC-Unified; Chen et al., 2008) and Stage IV (Lin and Mitchell, 2005) precipitation data, available through the National Oceanic and Atmospheric Administration, are used for rainfall analyses. CPC-Unified provides daily, 0.25º rainfall estimates interpolated from rain gage observations, while Stage IV provides hourly, approximately 4 km estimates by merging data from rain gages and the National Weather Service Next-Generation Radar network (NEXRAD; Crum and Alberty, 1993). Analyses based on Stage IV use data from 2002-2016, while long-term analyses based on CPC-Unified use data from 1948-2016.

4. Methodology

The FFA approach presented in this study combines continuous simulation (CS), stochastic storm transposition (SST) using the RainyDay software, and event-based simulation. CS provides large samples of seasonally varying antecedent conditions, namely soil moisture and snowpack. SST produces large numbers of synthetic rainfall scenarios. Together, these drive event-based simulations to generate the synthetic flood peaks that are used to derive flood frequency distributions. The approach is illustrated schematically in Fig. 2 and summarized in the following subsections.

4.1 Hydrologic Model, Calibration, and Continuous Simulation

We used the lumped Hydrologiska Byråns Vattenavdelning (HBV) model (Bergström, 1992, 1995; Lindström et al., 1997). HBV has been widely used to study hydrologic response in United States (Vis et al., 2015; Niemeyer et al., 2017) and other regions of the world (Harlin and Kung, 1992; Osuch et al., 2015; Seibert, 2003; Chen et al., 2012). The "HBV-Light" version (henceforth referred to as HBV; Seibert and Vis, 2012) used in this study consists of four main routines: snowpack, soil moisture, catchment response, and runoff routine. HBV simulates daily discharges based on time series of precipitation and air temperature, as well as estimates of long-term daily potential evapotranspiration. A list of model parameters is shown in Table 1.

The process-based FFA methodology employed in this study could be coupled with other hydrologic models. A distributed model would allow for more realistic representation of important characteristics like changing land use, rainfall spatiotemporal structure, and flood wave attenuation in river channels, and could operate at higher (i.e. subdaily) temporal resolution. We selected HBV at the daily time step due to its simplicity, computational speed, and its ability to represent multiple watershed hydrological processes.

We calibrated separate HBV models using both CPC and Stage IV rainfall. Most parameter values were the same for CPC- and Stage IV-based models except for three snow routine parameters (TT, CFMAX, SFCF) and three recession coefficients (K0, K1, K2), allowing for the variability of model parameters for different climate conditions. For each model setup, we first calibrated the model with snowpack routine "turned off" (by setting TT parameter to a very low

value) to obtain parameters that can simulate summer floods adequately. Then, keeping these optimized non-snow routine parameters unchanged, we calibrated the snow routine parameters.

To determine the optimized model parameter sets in each procedures, we followed the Genetic Algorithm and Powell (GAP) optimization method as presented by Seibert (2000), which is briefly summarized here. First, 5000 parameter sets are randomly generated from a uniform distribution of the values of each parameter (Table 1), which were then applied to the HBV model in order to maximize Kling Gupta Efficiency (Gupta et al., 2009) of simulated daily discharge. After the GAP has finished, the optimized parameter set were fine-tuned using Powell's quadratic convergent method (Press, 1996) with 1000 additional runs. Lastly, the optimized parameter set was manually adjusted to improve the fits between observed and simulated annual peak flow (see Lamb, 1999). More elaborate calibration and uncertainty estimation procedures such as Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992; Beven, 1993; Beven and Binley, 2014) could be used, but are outside the scope of our study.

The two different HBV models were then used to perform CS with historical CPC and Stage IV rainfall and temperature data to derive long-term simulated soil moisture and snowpack values, which are usually difficult to obtain via measurement. We "pair" samples of these initial conditions with synthetic rainfall events to simulate hypothetical floods, as described in Sect. 4.2 and Sect. 4.3.

4.2 Stochastic Storm Transposition

Stochastic storm transposition (SST) is a bootstrap method to generate realistic probabilistic rainfall scenarios through temporal resampling and spatial transposing of observed storms from the surrounding region. SST effectively "lengthens" the rainfall record via "space-for-time substitution." Unlike rainfall IDF curves, SST can preserve observed rainfall space-time structure, and, unlike design storm methods, obviates the need to equate rainfall duration to catchment response time (Wright et al., 2013, 2014a, 2014b). Alexander (1963), Foufoula-Georgiou (1989), and Fontaine and Potter (1989) provide general descriptions of SST. Wilson and Foufoula-Georgiou (1990) apply the method for regional rainfall frequency analysis while Gupta (1972), Franchini et al. (1996), England et al. (2014) and Nathan et al. (2016) use it for FFA.

Wright et al. (2013) used SST with a 10-year high-resolution radar rainfall dataset to estimate spatial IDF relationships. Wright et al. (2014a) used this approach with a physics-based distributed hydrologic model for FFA in a heavily urbanized watershed, demonstrating its usefulness in evaluating multi-scale flood response.

RainyDay is an open-source, Python-based SST software that couples SST methods with rainfall remote sensing data. A more detailed description can be found in Wright et al. (2017); not all of its features are used in this study. The following steps describe how RainyDay is used here: We define a 6-degree (longitude) by 4-degree (latitude) geographic transposition domain (40° to 44° N, 90° to 96° W; blue dash line of Fig. 1 inset) which encompasses the Turkey River watershed. This same domain was used in Wright et al. (2017) and, importantly for the SST approach, extreme rainfall properties are roughly homogeneous within it.

The RainyDay software creates a "storm catalog" from 15 years of Stage IV (69 years of CPC) precipitation data that consists of the 450 (2070) most intense precipitation event within the transposition domain. These intense storms are in terms of 96-hour rainfall accumulation and have the same size, shape, and orientation of the Turkey River watershed, which is oriented roughly northwest-southeast and with an area of 4002 km2. In order to avoid overlapping storms, these selected events must be separated by at least 24 hours. Storms that exhibit "radar artifacts" such as major bright band contamination or beam blockage are excluded from subsequent steps. The RainyDay software generates a Poisson-distributed integer k that represents a "number of storms per year." The rate parameter » of this Poisson distribution is calculated by dividing the total number of rainfall events in the storm catalog by the number of years in the historical rainfall record ($e.g. \lambda = 450/15 = 30.0$ storms per year).

RainyDay randomly selects k storms from the storm catalog and transposes the associated rainfall fields within the transposition domain by an east-west distance Δx and a north-south distance Δy, where Δx and Δy are drawn from a two-dimensional Gaussian kernel density estimate based on the locations of the original storms in the storm catalog. For each of the k transposed storms, the time series of rainfall over the Turkey River watershed is computed. It must be noted that some of the k transposed storms may not "hit" Turkey River watershed, and thus their calculated watershed rainfall are zero. Steps 3 and 4 can be understood as temporal resampling of observed rainfall events to "synthesize" a hypothetical year of rainfall events over the transposition domain and, by extension, over the watershed. Although the rainfall events for the "synthetic" year do not form a continuous series, the dates associated with each observed storm event are recorded, thus facilitating seasonally-consistent flood simulations.

All k events within a synthetic year are assigned a new, randomly selected year from 1948-2016 (2002-2016) for CPC (Stage IV) rainfall data which used to select antecedent conditions. This ensures that the k rainfall events are all "embedded" within a single realistic annual representation of watershed conditions. This ensures that "wet" and "dry" years in terms of snowpack and soil moisture can potentially produce wet or dry years of flood response. Antecedent conditions are randomly selected from within seven days of the updated storm date to ensure realistic seasonality of storms and watershed conditions. A storm that occurred on 15 July, 2016, for example, could be paired with initial conditions selected from a date ranging between 8-22 July from a randomly selected year, while the remaining k-1 events would be paired with seasonally appropriate initial conditions from the same selected year.

RainyDay repeats Steps 3-5 500 times to create one realization of 500 synthetic years of rainfall events for Turkey River. Twenty such realizations of 500 synthetic years each are generated. Unlike in the existing version of RainyDay, all rainfall events within a synthetic year are retained for subsequent event-based flood simulations, since the modulating effects of antecedent conditions mean that the largest rainfall event in a given year does not necessarily produce that year's largest flood peak (this is explored in Sect. 5.4).

4.3 Event-Based Flood Simulation

Using the seasonally-consistent "paired" watershed initial conditions derived from CS (Sect. 4.1) and SST-based rainfall events (Sect. 4.2), HBV simulates the "event peak" (the maxima daily

discharge). The largest peak among the k events that comprise a synthetic year represents the simulated annual maximum daily streamflow. As mentioned in Step 5 of the SST procedure (Section 4.2), each synthetic rainfall event is randomly paired with seasonally-appropriate initial conditions (soil moisture, snowpack) and air temperature drawn from the continuous simulation (15 years in the case of Stage IV; 69 years for CPC). This creates combinations of initial conditions and forcing that in principle reflect the true variability of these processes. This procedure is repeated for all 500 synthetic years within each realization, resulting in 500 annual maximum streamflow values, which are then ranked in descending magnitude. The annual exceedance probability p_e (i.e. the probability in a given year that an event of equal or greater magnitude will occur) of each maximum streamflow are calculated by dividing its rank by 500 (the total number of simulated annual maximum daily streamflow). The twenty realizations provide estimates of variability for each flood quantile.

5. Results

5.1 Hydroclimatology and Nonstationarity

Four distinct time periods (Fig. 3a) are considered for analyzing the changing hydroclimatology in Turkey River: the USGS daily mean streamflow period of record (1933-2016), a more recent period of apparent elevated flood activity (1990-2016), the period of the Stage IV rainfall record (2002-2016), and the period of the CPC rainfall record (1948-2016). Results here and in subsequent subsections "align" with one or more of these time periods.

The hydroclimate of Turkey River is changing, as shown using the Mann-Kendall (MK) test for monotonic trends (Mann, 1945), a nonparametric method used to determine trend direction and significance (Table 2). Since 1948, annual precipitation and discharge show significant increases ($p<0.05$) and their variability has also increased, while annual maximum daily discharge has decreased, though not significantly. It is important to note, however, that there are two counteracting seasonal trends (see also Fig. 3a): annual daily discharge maxima have decreased significantly in March-April, but have increased somewhat in May-September. Thus, the lack of statistically significant change in annual maximum daily discharge in Turkey River masks changes in the seasonality of flooding.

We examine this flood seasonality, both in observations and in our continuous HBV simulations (Fig. 3b). The seasonal distribution of flood occurrence for 1948-2016 shows a March-April maximum, with elevated flood activity continuing through May and June. This is distinct from, though overlaps somewhat with the seasonality of both the four-day annual maxima of rainfall, which occur most frequently in the June-September period, and simulated daily annual maxima soil moisture, which only tend to occur in March-April. These results highlight that flood activity is the product of seasonal variations in both soil moisture and rainfall. (Four-day rainfall shown in Fig. 3b since it is used in SST; seasonality in one-day rainfall is similar; results not shown).

The March-April peak of flood occurrence corresponds with relatively high soil moisture associated with snowmelt, rain on or frozen soil, and frequent spring rains. The secondary peak of flood occurrence in May-June is associated with larger flood magnitudes (including the flood of record, in 2004) due to organized thunderstorm systems. Widespread flooding in Iowa in June 2008 showed that such thunderstorm systems make critical contributions to the upper tail of

flood peak distributions in the region (Smith et al., 2013). Although the frequent August-September heavy rainfall events evident in Fig. 3b have not triggered any recorded annual flood peaks in Turkey River, our process-based FFA demonstrates that they may still relevant to current and future flood frequency, as shown in Sect. 5.4.

The largest annual maxima (over 800 m3 s-1) occur in May-July (Fig. 3c), consistent with the broader climatology of flooding in Iowa (Smith et al., 2013; Villarini et al., 2011). Furthermore, both the seasonality and magnitude of flood peaks have shifted since approximately 1990 (Fig. 3a, 3c), with March-April (May-September) floods decreasing (increasing) in magnitude, leading to a shift in the seasonality of the overall distribution of annual maxima daily streamflow from a high in March prior to 1990 to a prolonged high from April to June post-1990. Although the small sample size of the annual maxima daily discharge during this elevated 1990-2016 late-spring and summertime flood period may affect the reliability of the derived distribution of flood occurrence, Park and Markus (2014) also reported a significant shift toward summertime flooding in the nearby Pecatonica River. Statistically based FFA (including nonstationary methods) based on annual maxima discharges may fail to capture the impact of this shifting seasonality on flood frequency.

5.2 Model Validation

We validated the performance of continuous HBV simulations with respect to flood seasonality, frequency of annual daily discharge maxima, and normalized peak flow (i.e. the simulated or observed daily discharge divided by the 2-year flood), using both Stage IV and CPC as precipitation inputs (Fig. 4). We also validated two model structures: one with and the other without the HBV snowpack module. The purpose for this latter validation effort is to highlight the importance of proper process representation (and subsequent validation) in process-based FFA.

Simulated flood seasonality varies substantially during the CPC period of record (1948-2016) depending on the inclusion of the snowpack routine (Fig. 4a). Differences are less for the Stage IV period of record (2002-2016), due to the decreasing role of snowpack in deriving the floods in recent years (Fig. 4b). In both cases, the seasonality of flooding simulated using HBV is improved with the inclusion of the snowpack module, with a higher (lower) frequency of springtime (summertime) floods which more closely resembles observations. Empirical (i.e. plotting position-based) distributions for the simulated annual daily discharge maxima are mostly within the 90% confidence interval (obtained by nonparametric bootstrap) of the observations (Fig. 4c, 4d). CPC-based simulation results differ considerably depending on the inclusion of the snowpack module for more common events, but differences in simulated maxima vanish as flood magnitude increases (e.g. AEP<0.1). This is because the most extreme flood events occur later in the season and are thus independent of snowpack or snowmelt processes. Differences are generally negligible between Stage IV-based simulations with and without snowpack, since floods in this more recent period are generally driven by summertime thunderstorms. These findings are consistent with the general understanding of the regional seasonality of flooding in the region, as discussed in Sect. 5.1.

We compared all simulated and observed flood peaks that can be associated with a USGS observed daily streamflow value that is at least three times the mean annual daily discharge

(Fig. 4e, 4f). When associating simulated and observed flood peaks, we look within a 2-day window to allow for modest errors in simulated flood peak timing. All peaks in Figs. 4e and 4f are normalized by the median annual (i.e. 2-year) flood, which, as a rule of thumb, can be considered as the "within bank" threshold. Again, HBV with the snowpack routine outperforms the model without it, especially for the small to modest flood events in CPC-based simulations. The model without snowpack underestimates small to modest flood events in two cases due to the neglect of potential snowmelt contributions. While modest scatter exists in the Stage IV-based simulated peaks, there is no obvious systematic bias with event magnitude when the snowmelt routine is included. The good performance of the Stage IV simulations suggests that, when focusing on the recent period of elevated flood activity, Stage IV may be a more suitable rainfall input than CPC-Unified. In addition, CPC rainfall is known to contain errors in the extreme tail, due to gage "undercatch", insufficient gage density to properly sample convective rain cells, and spatial averaging of such cells over large areas, which effectively reduces peak rainfall depths.

We also validate HBV's snowpack routine using observed GHCN daily snow depth for two simulation periods (Fig. 5a, 5b) and using USGS daily streamflow observations for Stage IV-based period (Fig. 5c). Because of their differing spatial resolutions and physical representations, point-scale GHCN daily snow depths cannot be directly compared to the watershed-scale snow water equivalent simulated by HBV. Instead, we validate snowpack simulations in terms of the snowpack occurrence, defined as the number of nonzero snowpack on a particular date divided by the total number of years in the historical or simulated record. For example, there are 50 days in the GHCN observations when snowpack is present on January 1st in the 69-year period from 1948-2016, thus the occurrence rate is 0.72 (50 divided by 69). The HBV model with the snowpack routine captures the central tendency of observed snowpack dynamics, showing that snowpack frequently exists from early November to mid-February, with frequency of snow decreasing from late February until disappearing in early April.

Model hydrograph validation is provided in Fig. 5c for the Stage IV period (2002-2016), when major flooding occurred throughout Iowa. Model performance shows no obvious evidence of systematic bias in the streamflow simulations (see also Fig. 4f). Although flood seasonality derived from Stage IV-based simulation differs slightly from observations (see also Fig. 4a), these mismatches are associated with flood events smaller than the median annual flood (blue dash line in Fig. 5c). Stage IV-based simulations do not show bias flood magnitude in late summer. In other words, remaining biases in terms of flood seasonality generally correspond with frequent, small-magnitude events that are typically of less interest in FFA. We therefore conclude that the HBV model with snowpack is generally suitable for subsequent process-based FFA.

5.3 Flood Frequency Analyses

RainyDay-based flood frequency distributions for Turkey River at Garber using both Stage IV and CPC precipitation are compared with the distribution based on statistical analyses of discharge observations using 1933-2016 USGS annual maxima daily streamflows (Fig. 6). The latter is estimated using the HEC-SSP software (Bartles et al., 2016), which implements methods from Bulletin 17B (Interagency Advisory Committee on Water Data, 1982) using "station skew" to fit the log-Pearson Type III distribution. Observed annual daily streamflow maxima from

1933 to 2016 are also shown, where plotting position (p_e) is estimated using the Cunnane plotting position (Cunnane, 1978). As mentioned above, different HBV parameters are used for the Stage IV and CPC-based simulations; this is necessary due to the differing time periods and error properties of these two precipitation datasets.

The Stage IV-based flood frequency curve agrees reasonably well with the Bulletin 17B results for $p_e > 0.3$ (left panel of Fig. 6), but yields higher estimates for rarer events. The CPC-based curve, on the other hand, matches closely with Bulletin 17B. The Stage IV analyses use shorter but more recent (2002-2016) meteorological and hydrological records than the other frequency curves. When streamflow observations are divided into two groups (1933-1989 and 1990-2016), it becomes clear that the recent peak flood observations align well with the Stage IV-based SST results (right panel of Fig. 6). This, along with the increasing trend of annual mean precipitation and discharge shown in the previous subsection, suggests that, despite the relatively short (15-year) rainfall record used, Stage IV- driven process-based FFA adequately reflects flood frequency in the wetter recent climate (a similar result is shown in Wright et al., 2017), while the CPC-based and Bulletin 17B methods, both based on much longer data records fail to do so.

The results shown in Fig. 6 suggest that the recent shift from spring to summer flood activity is accompanied by a substantial shift in the flood frequency distribution. The close agreement between process-based results using CPC and the statistically-based analysis using Bulletin 17B suggests that even in stationary situations with long records, statistical methods do not necessarily produce superior results to process-based approaches. Process-based FFA using CPC precipitation from 2002- 2016 closely resembles the Stage IV-based FFA (results not shown), suggesting that rainfall process nonstationarity, rather than differences between different input datasets, are the primary drivers of the differences in the CPC-based and Stage IV-based results in the left panel of Fig. 6.

5.4 Simulated Flood Seasonality

As shown in Sect. 5.1, the recent climatology of flooding in Turkey River watershed shows a peak in flood occurrence during March-April, with elevated activity (including high-magnitude events) continuing through July, reflective of the regional flood "mixture distribution" (e.g. Smith et al., 2011). March-April flooding is associated with springtime rains, high soil moisture, and potentially snowmelt processes, while May-July flooding results from warm-season organized thunderstorm systems. It is important that any process-based FFA approach capture the influence of this mixture on the flood frequency curve.

The seasonal distribution of simulated flood occurrence and magnitude using Stage IV- and CPC-based results show that most simulated floods in our process-based approach occur between March and June (Fig. 7), in accordance with observed annual maxima daily discharge (Fig. 3c). The peak of occurrence using Stage IV is shifted several weeks later than the CPC-based results, which agrees with the recent shift in seasonality of flood observations shown in the Fig. 3c. Although many simulated events still occur in April, our results show the largest peaks occur later, in May-September. This is consistent with Villarini et al. (2011), who showed that warm season organized convective systems are responsible for some of the largest peaks in Iowa.

Our process-based results show that August-September storms have the potential to cause severe flooding (Fig. 7), despite the lack of large floods during this time of year in the stream gage record. Stage IV- and CPC-based storm catalogs generated by RainyDay include major storms from the surrounding region, including several large late-summer events capable of producing substantial flood response, and which indeed to induce large floods within our process-based analysis. This suggests that the general lack of major late-summer floods in the watershed's observational record may not be a feature of the "true" (unknown) distribution of flooding in the watershed, but rather due to limited size of the observational record. This result is supported by regional analysis of floods (Villarini et al. 2011) and points to the potential for SST to improve understanding of flood frequency seasonality relative to discharge-based approaches alone.

To demonstrate that the discrepancies between the process-based FFA results generated using CPC and using Stage IV- are driven by changes in physical processes, rather than by differences in model structure (i.e. parameter values), we compared FFA results generated using CPC-based for 1948-2016 and 2002-2016, in terms of event rainfall, initial soil moisture, flood type and peak magnitude (Fig. 8). Compared with the 1948-2016 period (Figure. 8a), there are fewer flood events driven by snowmelt or rain-on-snow during 2002-2016 (Fig. 8b) but more driven by rainfall. This is particular true for flood events (larger than 1000 m3 s-1). In addition, some of the rainfall-driven floods from 2002-2016 were caused by relatively low rainfall but high initial soil moisture, in accordance with the significant increasing trend of annual precipitation and discharge (Table 2).

5.5 Comparison of rainfall and peak discharge quantiles

We examined the relationships between the return periods of 96-hour basin-averaged rainfall accumulations and simulated peak discharge for Turkey River at Garber using Stage IV-based results (Fig. 9; CPC-based results show similar patterns and thus are not shown here). Antecedent soil moisture for each simulated event is also shown. Similar to Wright et al. (2014a), Fig. 9 shows that simulated peak discharge quantiles can differ substantially from the rainfall quantiles of the rainfall that produce them. For instance, 500-year ($p_e = 0.002$) rainfall events can cause simulated peak discharges ranging from 11-year ($p_e = 0.091$) to 500-year ($p_e = 0.002$), corresponding to a range in peak discharge of 1072 to 2743 m3 s-1. Peak discharge quantiles are always larger (in terms of return period) than the quantiles of rainfall that produced them in wet antecedent soil moisture conditions, while the reverse is true in for dry conditions. These results also demonstrate that the DS assumption of 1:1 equivalency between rainfall and peak discharge quantiles does not hold in Turkey River. Rainfall spatial variability and drainage network structure, which are ignored in this study due to the lumped (i.e. non-distributed) nature of HBV, further complicate the relationship between rainfall and discharge quantiles.

We further examine the relationship between annual rainfall and annual flood peak maxima. In Sect. 2.2, we pointed out that DS methods utilize IDF curves, which are usually estimated using annual maxima from rain gage records and which depict quantiles from the distribution of annual rainfall maxima. DS methods use quantiles from this distribution to generate flood estimates, implicitly assuming that annual rainfall maxima produce annual discharge maxima. In our process-based FFA approach, we do not assume that annual discharge maxima are the result of

the largest rainfall event of the year. Rather, lower-magnitude rainfall events, combined with high soil moisture, could produce the highest discharge.

Table 3 shows the percentage of annual peak flow driven by annual maximum gains with increasing return period for both CPC-based and Stage IV-based results. For simulated peak flow with $p_e > 0.01$, a large portion of simulated annual peak flow is not caused by the annual maximum rainfall. For rarer peak flows ($p_e \leq 0.01$), over 90% of these flood events are driven by the annual maximum rainfall, pointing to the fact that the tail of flood peaks is driven by extreme rainfall, with antecedent conditions playing a modulating role.

6. Summary and Conclusions

Interactions between rainfall, land cover, river channel morphology, and watershed antecedent conditions are important drivers of flood response. Standard approaches to estimate extreme flood quantiles (termed flood frequency analysis; FFA), however, often take a superficial view of these interactions, as argued in Sect. 2. This study presents an alternative FFA framework that combines elements of observational analysis, stochastic rainfall generation, and continuous and event-scale hydrologic simulation. We apply the framework to Turkey River, an agricultural watershed in the Midwestern United States that is undergoing significant hydroclimatologic and hydrologic change which is increasing the magnitude of the largest flood events and shifting their occurrence from the spring to summer.

We use Stochastic Storm Transposition (SST) to create and resample from "storm catalogs" developed from both 15 years of high-resolution bias-corrected radar rainfall and from 69 years of gridded rain gage observations to produce large numbers of rainfall scenarios for Turkey River. These scenarios, when coupled with seasonally realistic watershed conditions, can help to reconstruct the seasonal and secular variations in meteorological and hydrological processes and their interactions, providing an alternative FFA approach which is well-suited to nonstationary environments (see also Sivapalan and Samuel, 2009). While statistical approaches can in principle be applied to investigate the impacts of seasonality on FFA (e.g. Ouarda et al., 2006), such methods still do not directly provide process-level understanding of the factors that "shape" flood frequency. Unlike design storm approaches to FFA, the synthetic rainfall scenarios derived by the SST-based procedure do not require any assumptions regarding the spatial and temporal structure of rainfall, since they are driven by the structure and variability of historical observed storms.

Our analyses show that using the most recent 15 years of rainfall can produce realistic "present-day" flood quantile estimates that reflect the nonstationarities in rainfall and watershed conditions. The use of longer records, both within our procedure and conventional statistical FFA methods, leads to underestimates of current flood frequency due to their inability to represent recent shifts in flood activity in Turkey River. Our results challenge some common FFA assumptions, including the design storm presumption that rainfall annual maxima produce discharge annual maxima and the assumption of 1:1 equivalence in rainfall and flood quantiles. We paint a more complex picture in Turkey River, in which the shifting seasonality in rainfall and watershed conditions combine to shape the flood frequency. Spatial variability in rainfall structure, soil moisture, land use and watershed morphology, which are ignored in this study due

to the use of a lumped hydrologic model, add further complexity to the flood-generating processes. The proposed framework can be employed with more sophisticated distributed hydrologic models, thus facilitating the examination of rainfall spatial variability and its interactions with other factors (e.g. heterogeneous watershed characteristics and river network processes; Zhu et al., 2018; Viglione et al., 2010b, 2010a). This coupling may prove particularly useful for FFA in large watersheds in which there is a practically infinite number of different combinations of such spatially and temporally varying processes that could produce floods—a population that is almost certain to be undersampled in stream gage records and poorly served by design storm assumptions.

A number of issues remain that make broader usage of our process-based framework challenging. Perhaps the biggest limitation of process-based approaches is the necessity of discharge observations, which are central to both identifying hydrologic changes and to calibrate and validate the hydrologic model. Thus, usage of the approach in ungaged basins may not produce satisfactory results. This issue is fundamental to other FFA techniques as well. Statistically-based discharge analyses, for example, similarly rely on streamflow observations, while design storm approaches also require hydrologic model calibration.

We also note that caution is needed when attempting to employ process-based FFA. We were able to produce very similar flood frequency distributions using our approach, regardless of whether or not the HBV hydrologic model's snowpack routine was "turned on" or off (results omitted for brevity), despite very different simulated seasonality of flooding. This highlights that process-based frequency analyses can be influenced by poor model process representation that can lead to seemingly "correct" results for the wrong reasons. This implies that the modeler must have sufficient data and experience to recognize such issues. It also illustrates a key issue in FFA using both statistical approaches and process-based methods: flood quantiles, though the product of interactions between physical processes, reveal relatively little about those underlying processes that produce them. This is particularly problematic in changing hydroclimatic or watershed conditions, because nonstationary behavior is likely the result of seasonal shifts in one or more processes that may affect flooding in ways that are not well-reflected in observational records. Our results showing that major floods could occur in Turkey River in the late summer under current hydroclimatic conditions, despite their absence in the instrumental record, is one example of this. Failure to recognize and model such shifts could lead to results for past or present flood conditions that appear to be correct, but that may lead to incorrect inferences about future conditions.

In summary, our framework and results highlight the opportunity and challenge with process-based FFA approaches; namely, that progress on understanding and estimating flood frequency and how it is evolving in an era of unprecedented changes in land use and climate requires better understanding of how the underlying physical processes, and the interactions between them, are changing. Poor model representation of key hydrological processes, however, can lead to incorrect conclusions about present or future flood frequency. Despite the challenge, we share the view of Sivapalan and Samuel (2009) that process-based approaches hold great potential for advances in FFA research and practice, particularly in projecting future flood hazards in conjunction with data and modeling advances in the climate science community. We do not propose that process-based approaches should necessarily supplant more conventional discharge-

based analyses, and acknowledge that discharge observations are essential in such studies. Rather, we anticipate a gradual "merging" of statistical and process-based stochastic simulation techniques as well as of the associated observations and synthetic data.

References

Alexander, G. N.: Using the probability of storm transposition for estimating the frequency of rare floods, J. Hydrol., 1, 46–57, 1963.

Ayalew, T., Krajewski, W. and Mantilla, R.: Exploring the Effect of Reservoir Storage on Peak Discharge Frequency, J. Hydrol. Eng., 18(12), 1697–1708, doi:10.1061/(ASCE)HE.1943-5584.0000721, 2013.

Ayalew, T. B. and Krajewski, W. F.: Effect of River Network Geometry on Flood Frequency: A Tale of Two Watersheds in Iowa, J. Hydrol. Eng., 22(8), 06017004, doi:10.1061/(ASCE)HE.1943-5584.0001544, 2017.

Ball, J., Babister, M., Nathan, R., Weeks, W., Weinmann, E. and Retallick, M.: Australian Rainfall and Runoff: A Guide to Flood Estimation, edited by I. Testoni, Commonwealth of Australia (Geoscience Australia)., 2016.

Bartles, M., Brunner, G., Fleming, M., Faber, B. and Slaughter, J.: HEC-SSP Statistical Software Package Version 2.1, Computer Program Documentation, US Army Corps of Engineers, Institute for Water Resources Hydrologic Engineering Center (HEC), 609 Second Street Davis, CA 95616-4687., 2016.

Berghuijs, W. R., Woods, R. A., Hutton, C. J. and Sivapalan, M.: Dominant flood generating mechanisms across the United States: Flood Mechanisms Across the U.S., Geophys. Res. Lett., 43(9), 4382–4390, doi:10.1002/2016GL068070, 2016.

Bergström, S.: The HBV Model: Its Structure and Applications, Swedish Meteorological and Hydrological Institute (SMHI), Hydrology, Norrköping., 1992.
Bergström, S.: The HBV model (Chapter 13), in Computer Models of Watershed Hydrology, edited by V. P. Singh, pp. 443–476, Water Resources Publications, Highlands Ranch, Colorado, USA., 1995.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Adv. Water Resour., 16(1), 41–51, doi:10.1016/0309-1708(93)90028-E, 1993.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. Process., 6(3), 279–298, doi:10.1002/hyp.3360060305, 1992.

Beven, K. and Binley, A.: GLUE: 20 years on: GLUE: 20 YEARS ON, Hydrol. Process., 28(24), 5897–5918, doi:10.1002/hyp.10082, 2014.

Blazkova, S. and Beven, K.: Flood frequency prediction for data limited catchments in the Czech Republic using a stochastic rainfall model and TOPMODEL, J. Hydrol., 195(1–4), 256–278, 1997.

Blazkova, S. and Beven, K.: Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty), Water Resour. Res., 38(8), 14-1-14–14, doi:10.1029/2001WR000500, 2002.

Blazkova, S. and Beven, K.: A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, Water Resour. Res., 45(12), doi:10.1029/2007WR006726, 2009.
Calver, A., Lamb, R. and Morris, S. E.: River flood frequency estimation using continuous runoff modelling, Proc. Inst. Civ. Eng. - Water Marit. Energy, 136(4), 225–234, 1999.

Calver, A., Stewart, E. and Goodsell, G.: Comparative analysis of statistical and catchment modelling approaches to river flood frequency estimation: River flood frequency estimation, J. Flood Risk Manag., 2(1), 24–31, doi:10.1111/j.1753-318X.2009.01018.x, 2009.

Cameron, D., Beven, K. and Tawn, J.: An evaluation of three stochastic rainfall models, J. Hydrol., 228(1–2), 130–149, doi:10.1016/S0022-1694(00)00143-8, 2000.

Camici, S., Tarpanelli, A., Brocca, L., Melone, F. and Moramarco, T.: Design soil moisture estimation by comparing continuous and storm-based rainfall-runoff modeling: DESIGN SOIL MOISTURE ESTIMATION, Water Resour. Res., 47(5), doi:10.1029/2010WR009298, 2011.

Ceola, S., Laio, F. and Montanari, A.: Satellite nighttime lights reveal increasing human exposure to floods worldwide, Geophys. Res. Lett., 41(20), 7184–7190, doi:10.1002/2014GL061859, 2014.

Chen, H., Xu, C.-Y. and Guo, S.: Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff, J. Hydrol., 434–435, 36–45, doi:10.1016/j.jhydrol.2012.02.040, 2012.

Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R. and Janowiak, J. E.: Assessing objective techniques for gauge-based analyses of global daily precipitation, J. Geophys. Res., 113(D4), doi:10.1029/2007JD009132, 2008.

Cheng, L., AghaKouchak, A., Gilleland, E. and Katz, R. W.: Non-stationary extreme value analysis in a changing climate, Clim. Change, 127(2), 353–369, doi:10.1007/s10584-014-1254-5, 2014.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J. and Rasmussen, R. M.: A unified approach for process-based hydrologic modeling: 1. Modeling concept: A unified approach for process-based hydrologic modeling, Water Resour. Res., 51(4), 2498–2514, doi:10.1002/2015WR017198, 2015a.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Gochis, D. J., Rasmussen, R. M., Tarboton, D. G., Mahat, V., Flerchinger, G. N. and Marks, D. G.: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies: A unified approach for process-based hydrologic modeling, Water Resour. Res., 51(4), 2515–2542, doi:10.1002/2015WR017200, 2015b.

Crum, T. D. and Alberty, R. L.: The WSR-88D and the WSR-88D operational support facility., Bull. Am. Meteorol. Soc., 74, 1669–1687, 1993.

Cudworth, A. G.: Flood hydrology manual, US Dept. of the Interior, Bureau of Reclamation, Denver Office., 1989.

Cunha, L. K., Krajewski, W. F., Mantilla, R. and Cunha, L.: A framework for flood risk assessment under nonstationary conditions or in the absence of historical data, J. Flood Risk Manag., 4(1), 3–22, doi:10.1111/j.1753-318X.2010.01085.x, 2011.

Cunnane, C.: Unbiased plotting positions-a review, J. Hydrol., (37), 205–222, 1978.

Dawdy, D. R., Griffis, V. W. and Gupta, V. K.: Regional Flood-Frequency Analysis: How We Got Here and Where We Are Going, J. Hydrol. Eng., 17(9), 953–959, doi:10.1061/(ASCE)HE.1943-5584.0000584, 2012.

Douglas, E. M., Vogel, R. M. and Kroll, C. N.: Trends in floods and low flows in the United States: impact of spatial correlation, J. Hydrol., 16, 2000.

Eagleson, P. S.: Dynamics of flood frequency, Water Resour. Res., 8(4), 878–898, doi:10.1029/WR008i004p00878, 1972.

El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R. and Bobée, B.: Generalized maximum likelihood estimators for the nonstationary generalized extreme value model, Water Resour. Res., 43(3), doi:10.1029/2005WR004545, 2007.

Enfield, D. B., Mestas-Nuñez, A. M. and Trimble, P. J.: The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S., Geophys. Res. Lett., 28(10), 2077–2080, doi:10.1029/2000GL012745, 2001.

England, J. F., Julien, P. Y. and Velleux, M. L.: Physically-based extreme flood frequency with stochastic storm transposition and paleoflood data on large watersheds, J. Hydrol., 510, 228–245, doi:10.1016/j.jhydrol.2013.12.021, 2014.

Fontaine, T. A. and Potter, K. W.: Estimating probabilities of extreme rainfalls, J. Hydraul. Eng., 115(11), 1562–1575, 1989.

Foufoula-Georgiou, E.: A probabilistic storm transposition approach for estimating exceedance probabilities of extreme precipitation depths, Water Resour. Res., 25, 799–815, 1989.

Franchini, M., Helmlinger, K. R., Foufoula-Georgiou, E. and Todini, E.: Stochastic storm transposition coupled with rainfall—runoff modeling for estimation of exceedance probabilities of design floods, J. Hydrol., 175(1–4), 511–532, 1996.

Franchini, M., Galeati, G. and Lolli, M.: Analytical derivation of the flood frequency curve through partial duration series analysis and a probabilistic representation of the runoff coefficient, J. Hydrol., 303(1–4), 1–15, doi:10.1016/j.jhydrol.2004.07.008, 2005.

Franks, S. W. and Kuczera, G.: Flood frequency analysis: Evidence and implications of secular climate variability, New South Wales, Water Resour. Res., 38(5), 20-1-20–7, doi:10.1029/2001WR000232, 2002.

Furrer, E. M. and Katz, R. W.: Improving the simulation of extreme precipitation events by stochastic weather generators, Water Resour. Res., 44(12), n/a-n/a, doi:10.1029/2008WR007316, 2008.

Gilleland, E. and Katz, R. W.: extRemes 2.0: An Extreme Value Analysis Package in R, J. Stat. Softw., 72(8), doi:10.18637/jss.v072.i08, 2016.

Gilroy, K. L. and McCuen, R. H.: A nonstationary flood frequency analysis method to adjust for future climate change and urbanization, J. Hydrol., 414–415, 40–48, doi:10.1016/j.jhydrol.2011.10.009, 2012.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Gupta, V. K.: Transposition of Storms for Estimating Flood Probability Distributions, Colo. State Univ. Hydrol. Pap., 59, 35, 1972.

Haberlandt, U.: A space-time hybrid hourly rainfall model for derived flood frequency analysis, Hydrol. Earth Syst. Sci., 15, 2008.

Harlin, J. and Kung, C.-S.: Parameter uncertainty and simulation of design floods in Sweden, J. Hydrol., 137(1–4), 209–230, 1992.

Hirsch, R. M. and Ryberg, K. R.: Has the magnitude of floods across the USA changed with global CO2 levels?, Hydrol. Sci. J., 57(1), 1–9, doi:10.1080/02626667.2011.621895, 2012.

Hyndman, D. W.: Impacts of Projected Changes in Climate on Hydrology, in Global Environmental Change, edited by B. Freedman, pp. 211–220, Springer Netherlands, Dordrecht., 2014.

IFC: Hydrologic Assessment of the Turkey River Watershed (DRAFT), Iowa Flood Center, 100 C. Maxwell Stanley Hydraulics Laboratory Iowa City, Iowa 52242., 2014.

Interagency Advisory Committee on Water Data (IACWD): Guidelines for Determining Flood Flow Frequency, Bulletin 17B, Reston, VA., 1982.

Jain, S. and Lall, U.: Magnitude and timing of annual maximum floods: Trends and large-scale climatic associations for the Blacksmith Fork River, Utah, Water Resour. Res., 36(12), 3641–3651, doi:10.1029/2000WR900183, 2000.

Kjeldsen, T. R.: The revitalised FSR/FEH rainfall-runoff method, NERC/Centre for Ecology & Hydrology., 2007.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrol. Sci. J., 31(1), 13–24, doi:10.1080/02626668609491024, 1986.

Klemeš, V.: Tall tales about tails of hydrological distributions. I, J. Hydrol. Eng., 5(3), 227–231, 2000a.

Klemeš, V.: Tall tales about tails of hydrological distributions. II, J. Hydrol. Eng., 5(3), 232–239, 2000b.

Konrad, C. P. and Booth, D. B.: Hydrologic Trends Associated with Urban Development for Selected Streams in the Puget Sound Basin, Western Washington, , 48, 2002.

Lamb, R.: Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation, Water Resour. Res., 35(10), 3103–3114, 1999.

Lamb, R., Faulkner, D., Wass, P. and Cameron, D.: Have applications of continuous rainfall-runoff simulation realized the vision for process-based flood frequency analysis?: Process-based Flood Frequency: Have Applications Realized the Vision?, Hydrol. Process., 30(14), 2463–2481, doi:10.1002/hyp.10882, 2016.

Li, J., Thyer, M., Lambert, M., Kuczera, G. and Metcalfe, A.: An efficient causative event-based approach for deriving the annual flood frequency distribution, J. Hydrol., 510, 412–423, doi:10.1016/j.jhydrol.2013.12.035, 2014.

Lin, Y. and Mitchell, K. E.: 1.2 the NCEP stage II/IV hourly precipitation analyses: Development and applications, in 19th Conf. Hydrology, American Meteorological Society, San Diego, CA, USA., 2005.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201(1–4), 272–288, 1997.

Linsley, R. K.: Flood Estimates: How Good Are They?, Water Resour. Res., 22(9S), 1986.

Luke, A., Vrugt Jasper, A., AghaKouchak, A., Matthew, R. and Sanders, B. F.: Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the United States, Water Resour. Res., 53(7), 5469–5494, doi:10.1002/2016WR019676, 2017.

Machado, M. J., Botero, B. A., López, J., Francés, F., Díez-Herrero, A. and Benito, G.: Flood frequency analysis of historical flood data under stationary and non-stationary modelling, Hydrol. Earth Syst. Sci., 19(6), 2561–2576, doi:10.5194/hess-19-2561-2015, 2015.

Mann, H. B.: Nonparametric tests against trend, Econometrica, 13, 245–259, 1945.

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

Nathan, R., Jordan, P., Scorah, M., Lang, S., Kuczera, G., Schaefer, M. and Weinmann, E.: Estimating the exceedance probability of extreme rainfalls up to the probable maximum precipitation, J. Hydrol., 543, 706–720, doi:10.1016/j.jhydrol.2016.10.044, 2016.

Niemeyer, R. J., Link, T. E., Heinse, R. and Seyfried, M. S.: Climate moderates potential shifts in streamflow from changes in pinyon-juniper woodland cover across the western U.S., Hydrol. Process., 31(20), 3489–3503, doi:10.1002/hyp.11264, 2017.

Ntelekos, A. A., Oppenheimer, M., Smith, J. A. and Miller, A. J.: Urbanization, climate change and flood policy in the United States, Clim. Change, 103(3–4), 597–616, doi:10.1007/s10584-009-9789-6, 2010.

Osuch, M., Romanowicz, R. J. and Booij, M. J.: The influence of parametric uncertainty on the relationships between HBV model parameters and climatic characteristics, Hydrol. Sci. J., 60(7–8), 1299–1316, doi:10.1080/02626667.2014.967694, 2015.

Ouarda, T. B. M. J., Cunderlik, J. M., St-Hilaire, A., Barbet, M., Bruneau, P. and Bobée, B.: Data-based comparison of seasonality-based regional flood frequency methods, J. Hydrol., 330(1–2), 329–339, doi:10.1016/j.jhydrol.2006.03.023, 2006.

Park, D. and Markus, M.: Analysis of a changing hydrologic flood regime using the Variable Infiltration Capacity model, J. Hydrol., 515, 267–280, doi:10.1016/j.jhydrol.2014.05.004, 2014.

Peleg, N., Blumensaat, F., Molnar, P., Fatichi, S. and Burlando, P.: Partitioning the impacts of spatial and climatological rainfall variability in urban drainage modeling, Hydrol. Earth Syst. Sci., 21(3), 1559–1572, doi:10.5194/hess-21-1559-2017, 2017.

Potter, K. W.: Evidence for nonstationarity as a physical explanation of the Hurst Phenomenon, Water Resour. Res., 12(5), 1047–1052, doi:10.1029/WR012i005p01047, 1976.

Press, W. H., Ed.: FORTRAN numerical recipes, 2nd ed., Cambridge University Press, Cambridge [England]/; New York., 1996.

Prosdocimi, I., Kjeldsen, T. R. and Miller, J. D.: Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models, Water Resour. Res., 51(6), 4244–4262, doi:10.1002/2015WR017065, 2015.

Rashid, M. M., Beecham, S. and Chowdhury, R. K.: Simulation of extreme rainfall and projection of future changes using the GLIMCLIM model, Theor. Appl. Climatol., 130(1–2), 453–466, doi:10.1007/s00704-016-1892-9, 2017.

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

Schilling, K. E. and Libra, R. D.: INCREASED BASEFLOW IN IOWA OVER THE SECOND HALF OF THE 20TH CENTURY, J. Am. Water Resour. Assoc., 39(4), 851–860, doi:10.1111/j.1752-1688.2003.tb04410.x, 2003.

Schilling, K. E., Gassman, P. W., Kling, C. L., Campbell, T., Jha, M. K., Wolter, C. F. and Arnold, J. G.: The potential for agricultural land use change to reduce flood risk in a large watershed, Hydrol. Process., 28(8), 3314–3325, doi:10.1002/hyp.9865, 2014.

Seibert, J.: Multi-Criteria calibration of a conceptual runoff model using a genetic algorithm, Hydrol. Earth Syst. Sci., 4, 215–224, 2000.

Seibert, J.: Reliability of model predictions outside calibration conditions, Hydrol. Res., 34(5), 477–492, 2003.

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrol. Earth Syst. Sci., 16(9), 3315–3325, doi:10.5194/hess-16-3315-2012, 2012.

Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, J. Am. Stat. Assoc., 63(324), 1379, doi:10.2307/2285891, 1968.

Serago, J. M. and Vogel, R. M.: Parsimonious nonstationary flood frequency analysis, Adv. Water Resour., 112, 1–16, doi:10.1016/j.advwatres.2017.11.026, 2018.

Sivapalan, M. and Samuel, J. M.: Transcending limitations of stationarity and the return period: process-based approach to flood estimation and risk assessment, Hydrol. Process., 23(11), 1671–1675, doi:10.1002/hyp.7292, 2009.

Smith, J. A., Villarini, G. and Baeck, M. L.: Mixture Distributions and the Hydroclimatology of Extreme Rainfall and Flooding in the Eastern United States, J. Hydrometeorol., 12(2), 294–309, doi:10.1175/2010JHM1242.1, 2011.

Smith, J. A., Baeck, M. L., Villarini, G., Wright, D. B. and Krajewski, W.: Extreme Flood Response: The June 2008 Flooding in Iowa, J. Hydrometeorol., 14(6), 1810–1825, doi:10.1175/JHM-D-12-0191.1, 2013.

Stedinger, J. R. and Griffis, V. W.: Getting from here to where? Flood frequency analysis and climate, JAWRA J. Am. Water Resour. Assoc., 47(3), 506–513, 2011.

Svensson, C. and Jones, D. A.: Review of methods for deriving areal reduction factors: Review of ARF methods, J. Flood Risk Manag., 3(3), 232–245, doi:10.1111/j.1753-318X.2010.01075.x, 2010.

Viglione, A. and Blöschl, G.: On the role of storm duration in the mapping of rainfall to flood return periods, Hydrol. Earth Syst. Sci., 13(2), 205–216, 2009.

Viglione, A., Merz, R. and Blöschl, G.: On the role of the runoff coefficient in the mapping of rainfall to flood return periods, Hydrol. Earth Syst. Sci., 13(5), 577–593, 2009.

Viglione, A., Chirico, G. B., Woods, R. and Blöschl, G.: Generalised synthesis of space–time variability in flood response: An analytical framework, J. Hydrol., 394(1–2), 198–212, doi:10.1016/j.jhydrol.2010.05.047, 2010a.

Viglione, A., Chirico, G. B., Komma, J., Woods, R., Borga, M. and Blöschl, G.: Quantifying space-time dynamics of flood event types, J. Hydrol., 394(1–2), 213–229, doi:10.1016/j.jhydrol.2010.05.041, 2010b.

Villarini, G., Serinaldi, F., Smith, J. A. and Krajewski, W. F.: On the stationarity of annual flood peaks in the continental United States during the 20th century, Water Resour. Res., 45(8), doi:10.1029/2008WR007645, 2009.

Villarini, G., Smith, J. A., Baeck, M. L. and Krajewski, W. F.: Examining Flood Frequency Distributions in the Midwest U.S.1, JAWRA J. Am. Water Resour. Assoc., 47(3), 447–463, doi:10.1111/j.1752-1688.2011.00540.x, 2011.

Vis, M., Knight, R., Pool, S., Wolfe, W. and Seibert, J.: Model Calibration Criteria for Estimating Ecological Flow Characteristics, Water, 7(5), 2358–2381, doi:10.3390/w7052358, 2015.

Wilson, L. L. and Foufoula-Georgiou, E.: Regional rainfall frequency analysis via stochastic storm transposition, J. Hydraul. Eng., 116(7), 859–880, 1990.

Wright, D. B., Smith, J. A., Villarini, G. and Baeck, M. L.: Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition, J. Hydrol., 488, 150–165, doi:10.1016/j.jhydrol.2013.03.003, 2013.

Wright, D. B., Smith, J. A. and Baeck, M. L.: Critical examination of area reduction factors, J. Hydrol. Eng., 19(4), 769–776, 2014a.

Wright, D. B., Smith, J. A. and Baeck, M. L.: Flood frequency analysis using radar rainfall fields and stochastic storm transposition, Water Resour. Res., 50(2), 1592–1615, doi:10.1002/2013WR014224, 2014b.

Wright, D. B., Mantilla, R. and Peters-Lidard, C. D.: A remote sensing-based tool for assessing rainfall-driven hazards, Environ. Model. Softw., 90, 34–54, doi:10.1016/j.envsoft.2016.12.006, 2017.

Zhu, Z., Wright, D. B. and Yu, G.: The Impact of Rainfall Space-Time Structure in Flood Frequency Analysis, Water Resour. Res., 54(11), 8983–8998, doi:10.1029/2018WR023550, 2018.

Tables

Table 1. Overview of HBV model parameters and upper and lower parameter limits used for calibration

Parameter	Description	Units	Min value	Max value
Snow Routine				
TT	Threshold temperature for liquid and solid precipitation	°C	-3	3
CFMAX	Degree-day factor	mm d^{-1}°C^{-1}	0.5	4
SFCF	Snowfall correction factor	-	0.5	1.2
CFR	Refreezing coefficient	-	0.01	0.1
CWH	Water holding capacity of the snow storage	-	0.1	0.3
Soil Moisture Routine				
FC	Maximum soil moisture storage (field capacity)	mm	100	550
LP	Relative soil water storage below which AET is reduced linearly	-	0.3	1
BETA	Exponential factor for runoff generation	-	1	5
Response Routine				
PERC	Maximum percolation from upper to lower groundwater box	mm d^{-1}	0	10
UZL	Threshold of upper groundwater box	mm	0	50
K0	Recession coefficient 0	d^{-1}	0.5	0.9
K1	Recession coefficient 1	d^{-1}	0.15	0.5
K2	Recession coefficient 2	d^{-1}	0.01	0.15
Routing Routine				
MAXBAS	Length of triangular weighting function	d	1	2.5

Table 2. Mann-Kendall trend test (two sided) for hydrological variables. p-values are given in parentheses; bolded values are significant at the 5% level. Analyses of trends in variances examine changes in the absolute values of residuals obtained from a linear regression using the Thiel-Sen estimator (Sen, 1968)

Data	Time Range	Trend
Annual Discharge	**1933-2016**	**‘ (0.001)**
Annual Max. Daily Discharge	1933-2016	“ (0.447)
Variance of Annual Max. Daily Discharge	1933-2016	‘ (0.056)
Annual Max. Daily Discharge in March-April	**1933-2016**	**“ (0.002)**
Annual Max. Daily Discharge in May-September	1933-2016	‘ (0.089)
Annual Precipitation	**1948-2016**	**‘ (0.003)**
Annual Max. Daily Precipitation	1948-2016	‘ (0.362)
Annual Max. 4-day Precipitation	1948-2016	‘ (0.419)
Annual Mean Temperature	1948-2016	“ (0.462)
March-May Mean Temperature	1948-2016	‘ (0.443)

Table 3. Percentages of simulated annual maxima daily flows driven by 96-hour rainfall annual maximum

Return Period	Driven by Annual Maximum Rainfall	
	CPC-based results	Stage IV-based results
1-2	24%	37%
2-5	32%	45%
5-10	39%	67%
10-20	48%	77%
20-50	60%	80%
50-100	72%	84%
100-200	77%	85%
200-500	93%	95%

Figures

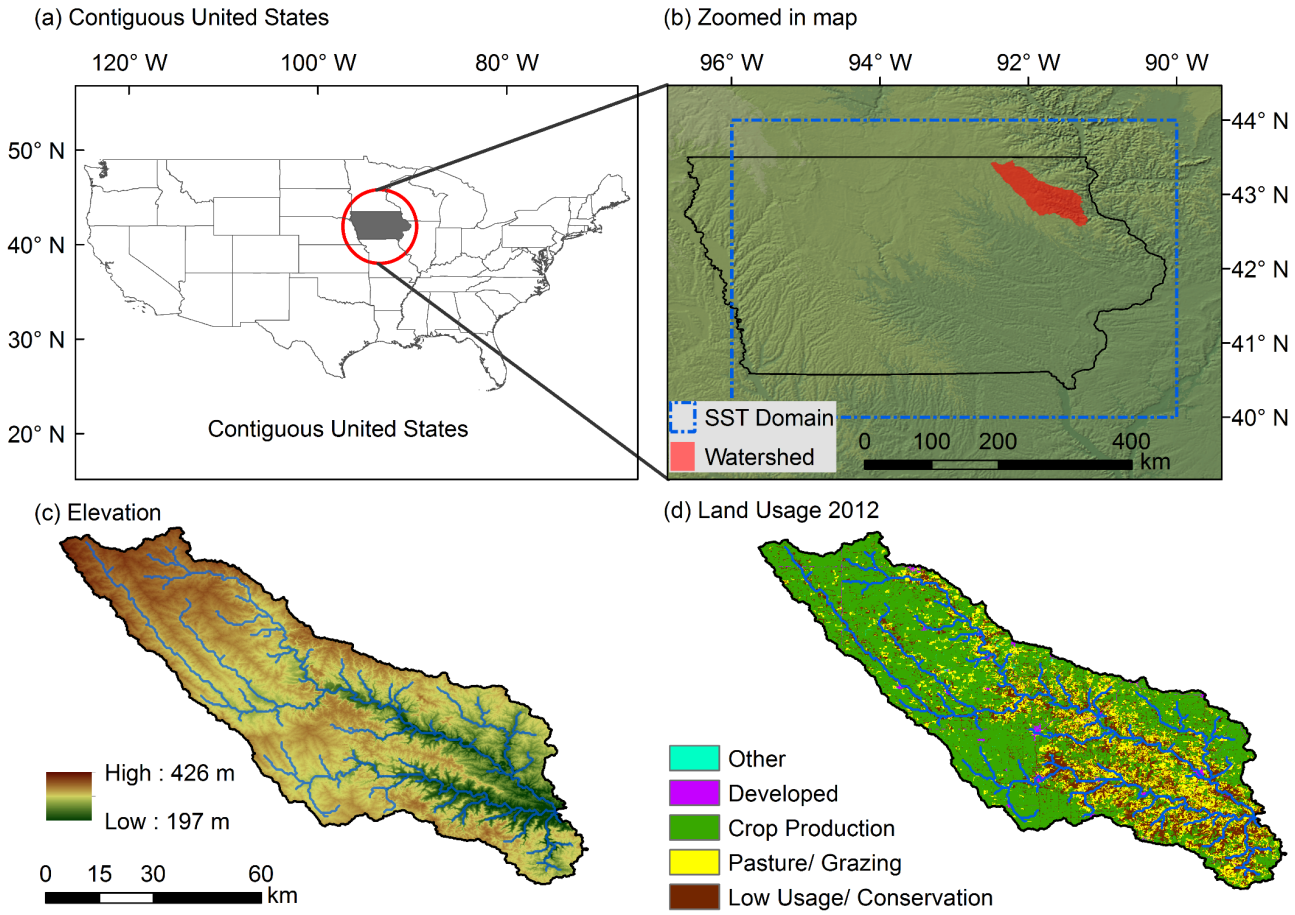

Figure 2. Study region. (a) Contiguous United States with the state of Iowa highlighted in grey. (b) Zoomed-in map showing Iowa (black outline) and the Turkey River watershed (red) and the extent of the stochastic storm transposition region (blue dash line). (c, d) Turkey River watershed showing land surface elevation (based on the USGS National Elevation Dataset) and land use (based on the USGS 2012 NLCD), respectively.

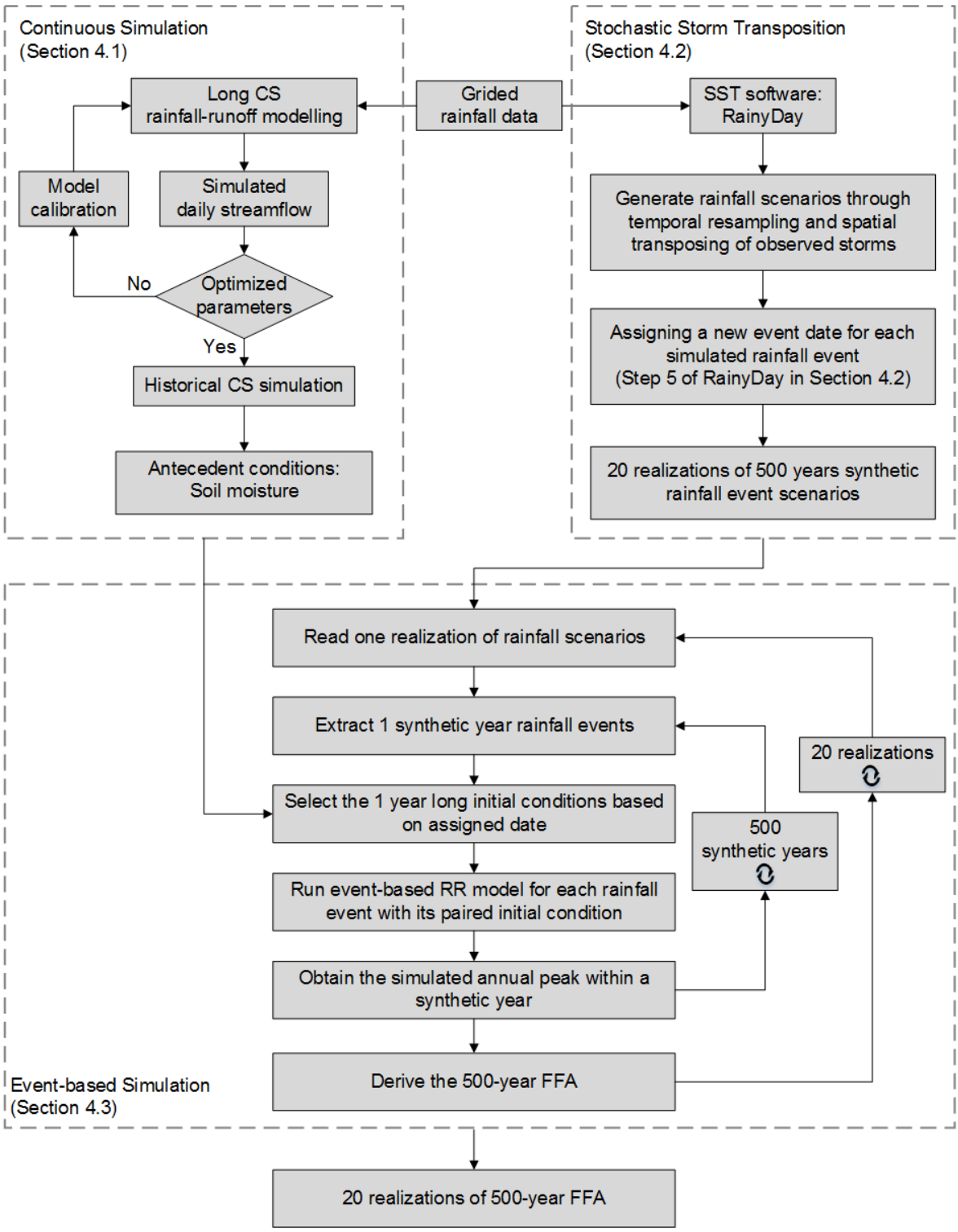

Figure 3. Flow chart showing the process-based FFA approach. Dotted outlines delineate components associated with subsections 4.1, 4.2 and 4.3.

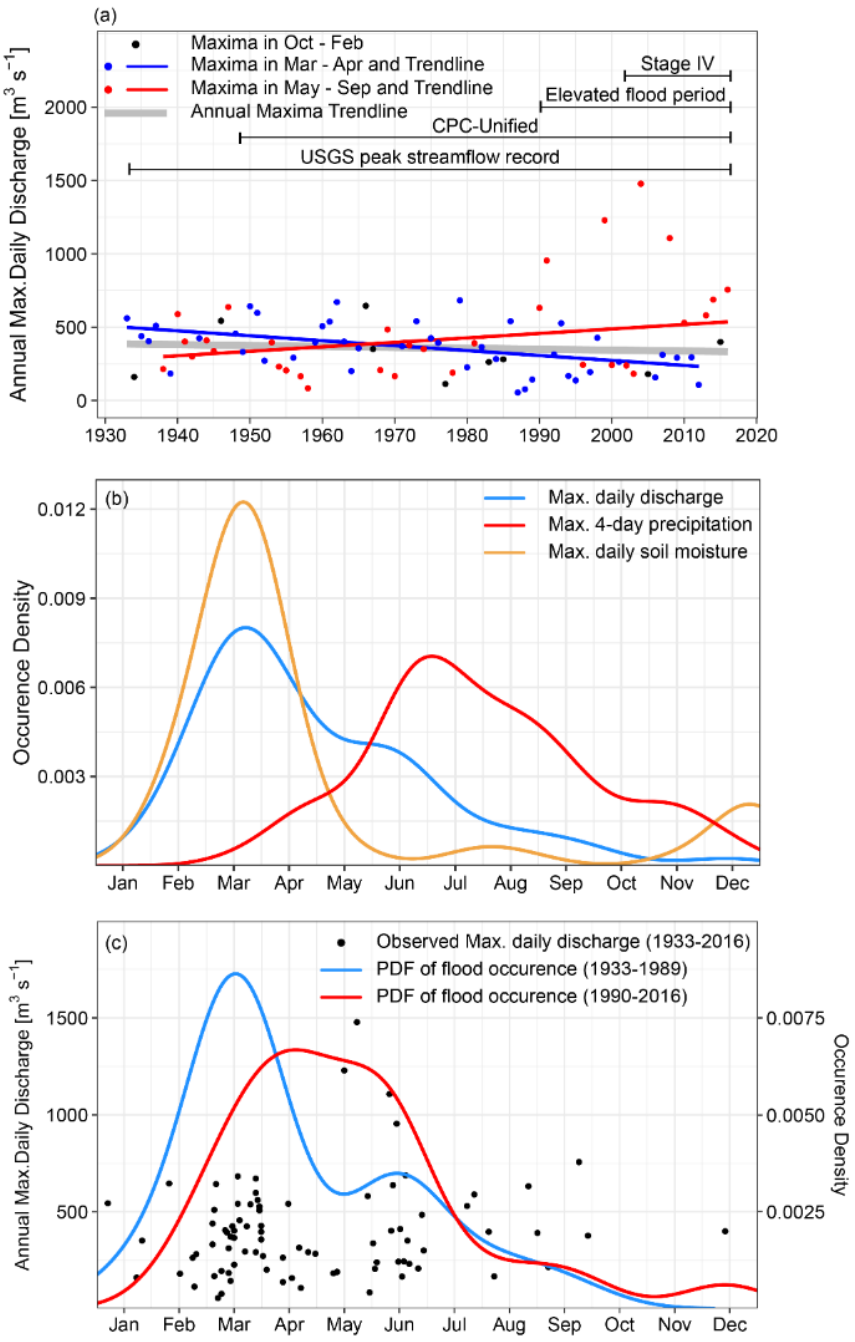

Figure 4. (a) Linear trends for two groups of annual maxima daily discharge: March-April floods (blue) and May-September floods (red) using the nonparametric Thiel-Sen estimator (Sen, 1968). The October-February maxima daily discharge are in black dots and its trend line is not calculated because only nine annual maxima occur during this period. The four critical time ranges are shown in black lines.
(b) Occurrence densities of the date during the year for the observed annual daily maxima discharge, observed annual 4-day maxima precipitation, and simulated annual daily maxima soil moisture in Turkey River watershed from 1948 to 2016. (c) The magnitude and date during the year for annual flood peaks (black dots)and sample probability density functions (PDFs) for flood in different periods (1933-1989, 1990-2016). In this study, all probability densities for occurrence date are estimated using Gaussian kernel smoothing.

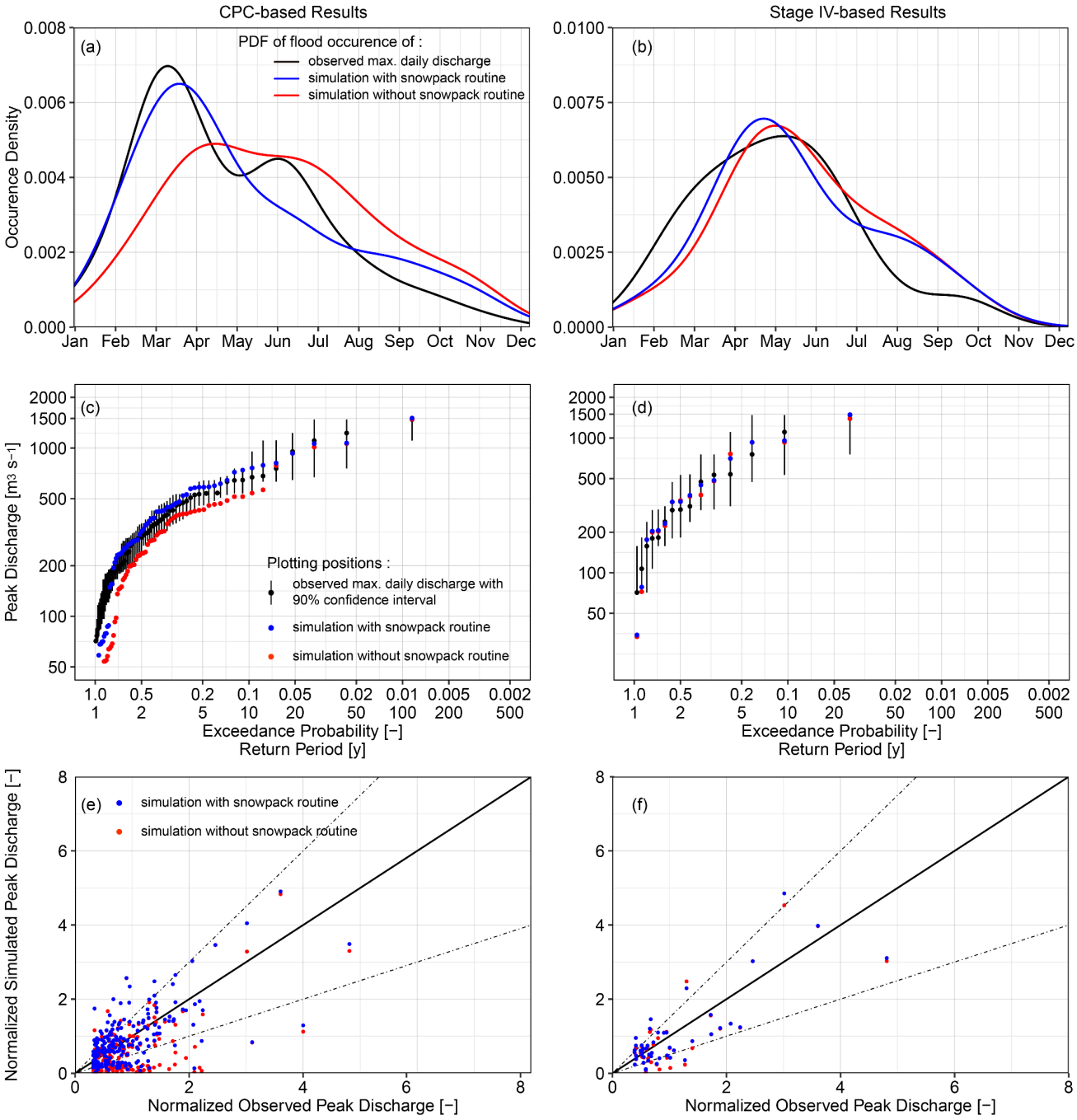

Figure 5. HBV model validation for flood seasonality (a, b), frequency of annual max. daily discharge (c, d) and normalized peak flow (e, f) for CPC and Stage IV-based continuous simulations. Model validation is performed for HBV simulations with and without using CPC for 1948-2016 (panels a, c, e) and Stage IV for 2002-2016 (panels, b, d, f). The 90% confidence intervals for the empirical distributions of observed maximum daily discharges (c, d) are derived using nonparametric bootstrapping. Flood peak discharge in (e) and (f) is defined as a data point with USGS observed value that is at least three times the average observations. Peak discharges are normalized by the median of annual daily discharge maxima (i.e. the 2-year flood). Straight solid black lines indicate 1:1 correspondence, while dashed lines denote an envelope within which the modeled values are within 50% of observed.

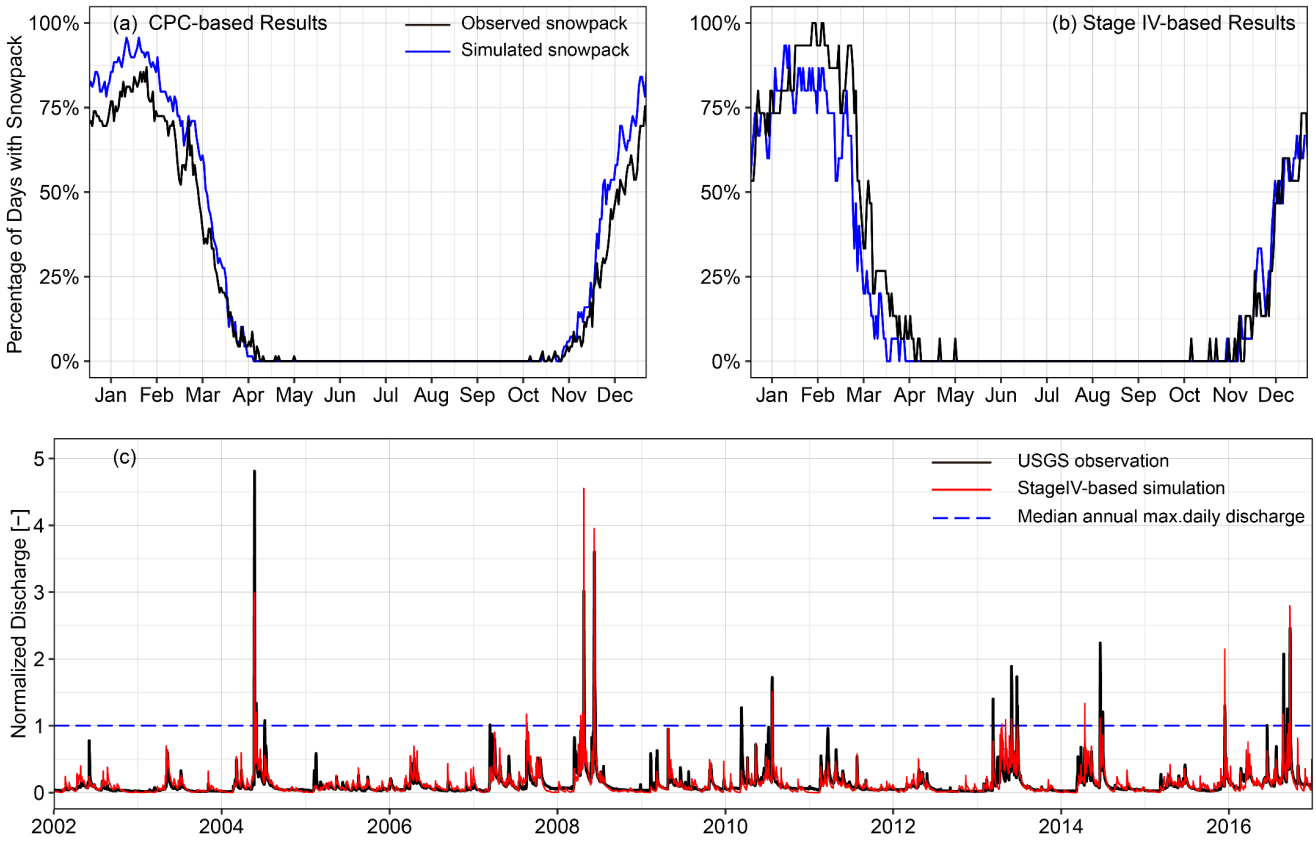

Figure 6. Percentage of days with nonzero snowpack present in observations and simulations (a, b) and hydrograph validation for Stage IV-based simulation (c). For each day within a year, the percent with nonzero snowpack is calculated as the ratio of the number of years in which snowpack is present on that day to the total years (69 years for CPC and 15 years for Stage IV). Observed and simulated hydrographs are normalized by the median annual flood, which is indicated by the dashed blue line.

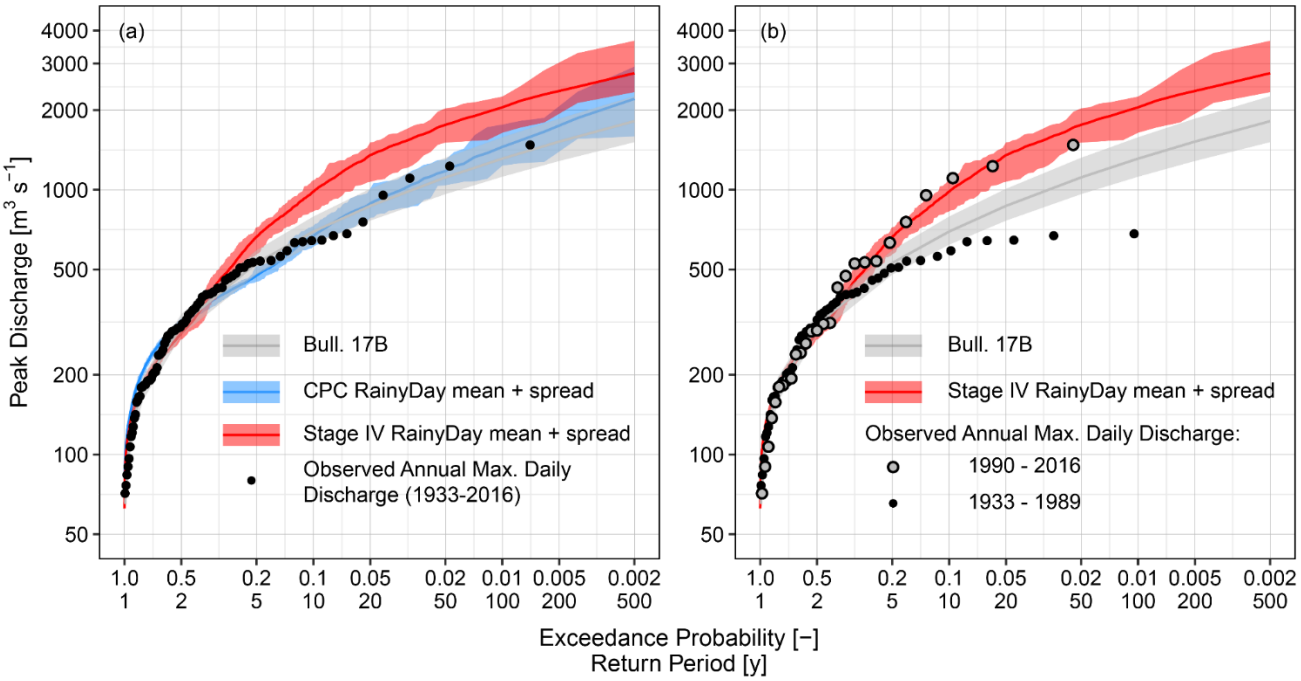

Figure 8. Peak discharge analyses for Turkey River at Garber, IA. (a) RainyDay with Stage IV (2002-2016) and CPC-(1948-2016) rainfall and USGS frequency analyses using Bulletin 17B methods. All observed USGS annual maxima daily streamflow from 1933 to 2016 are also shown. Shaded areas denote the ensemble spread (RainyDay-based results) and the 90% confidence intervals (Bulletin 17B-based analysis), respectively. (b) Same as (a), but with the USGS observations divided into pre-1990 and post-1990 groups, and replotted to highlight recent changes in flood frequency.

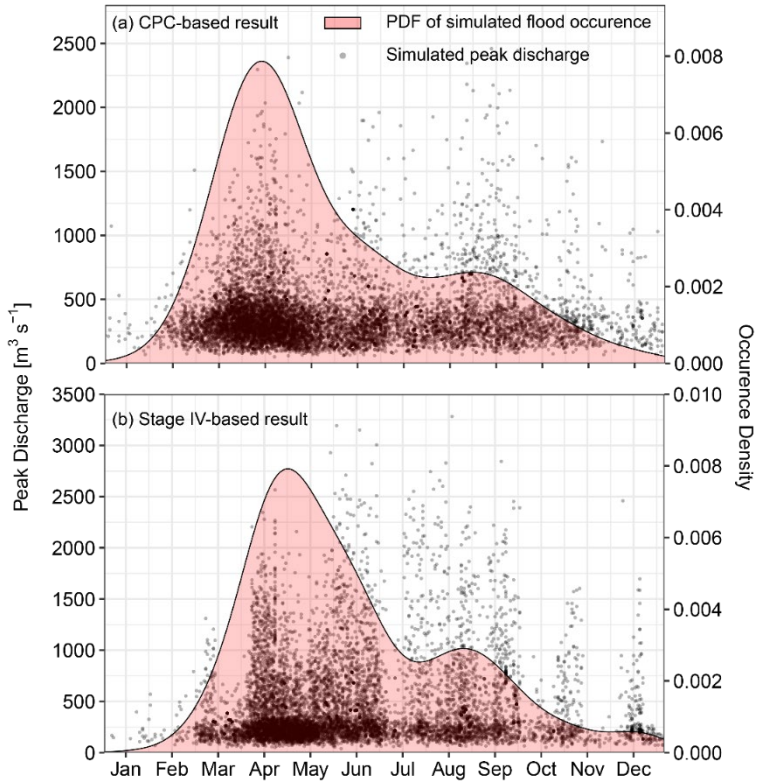

Figure 7. Time of occurrence during the year for simulated peak discharge in Turkey River at Garber using (a) CPC and (b) Stage IV.

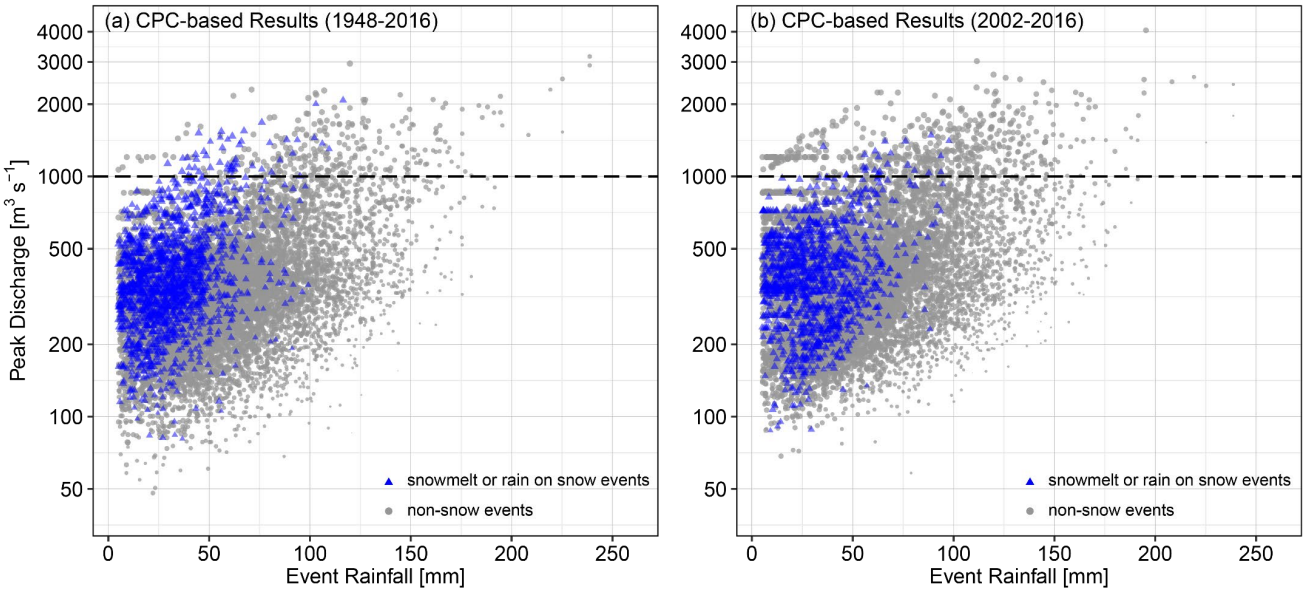

Figure 9. The simulated flood magnitude using CPC rainfall during 1948-2016 (a) and 2002-2016 (b) periods, and corresponding antecedent conditions. The blue triangles denote the snow related flood events (e.g. snowmelt was nonzero in the simulation) and grey dots represent the non-snow related flood events (e.g. rainfall driven). The size of the triangles or dots indicate the antecedent soil moisture with higher value in larger shape. The black dash line indicates the 1000 m3 s-1 flood magnitudes.

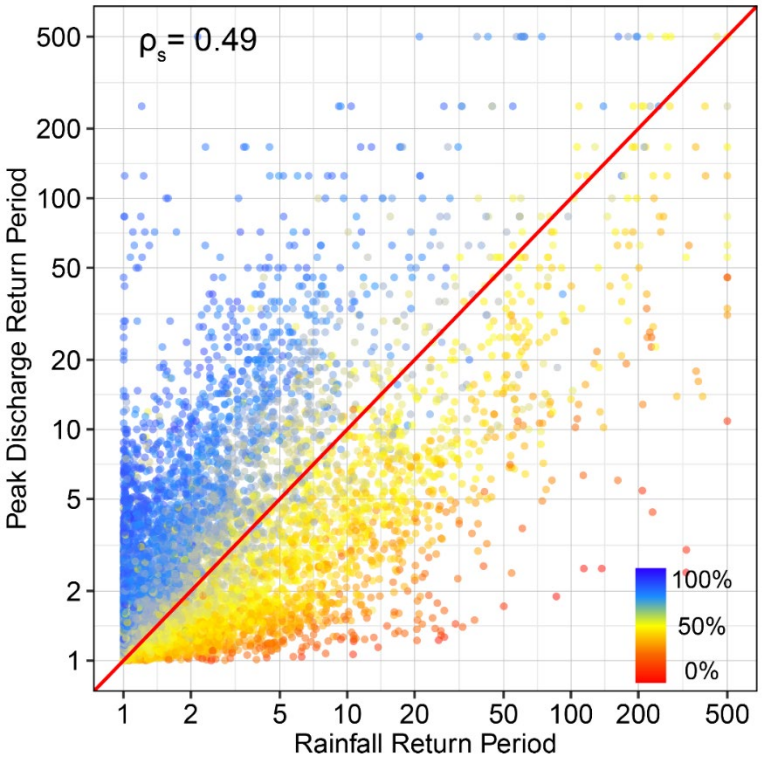

Figure 9. Relationships between rainfall and simulated peak discharge return periods estimated via our process-based method using Stage IV rainfall data. Spearman rank correlation ρ_s is given. Color indicates the normalized modeled antecedent soil moisture value calculated as $\frac{soil\ moisutre - min.soil\ moisture}{field\ capacity - min.soil\ moisture} * 100\%$.

Appendix D – Manuscript 3

Process-based Understanding of Flood Hydrometeorology and its Utilization in Flood Frequency Analysis for a Mountainous Watershed in the Colorado Front Range

Guo Yu* and Daniel Wright
University of Wisconsin-Madison, Madison, Wisconsin

Kathleen D. Holman
Bureau of Reclamation, Denver, Colorado

* gyu29@wisc.edu

Abstract

Understanding and deriving the frequency distributions for flood peaks and volumes in a mountainous watershed is extreme difficult due to complexities in topography, orographic precipitation, and flood regimes. Conventional flood frequency analysis (FFA) methods typical neglect these difficulties and thus offer few insights into complex interactions of physical flood drivers. Understanding flood drivers and how they "shape" the flood frequency distributions is as important as generating the distribution curves. This study presents an alternative "process-based" FFA approach that couples stochastic storm transposition, physics-based hydrological model, and long-term continuous simulations simulate frequency distributions for flood peaks and volumes. We apply these approaches with 36-year NLDAS-2 forcings to derive and examine frequency distributions for flood peaks and volumes up to at least the 1,000-year recurrence interval in a mountainous watershed in the Colorado Front Range, United States. We found that watershed antecedent conditions and snowmelt drive the frequent flood peaks and long-term flood volumes while the upper tail of the flood quantiles are determined by the heavy rainfall or rain on snow events. Findings highlight both the potential and limitations of SST and physics-based hydrological model to help understand the relationships between different flood drivers and their corresponding flood peaks and volumes.

1. Introduction

The September 2013 Colorado floods, which delivered 7-day (9-15 September 2013) total rainfall that exceeded 300 mm over a broad region across the Colorado Front Range, destroyed at least 1,882 structures and caused total property damage more than $2 billion (Gochis et al., 2015; NWS, 2013, 2014). Gochis et al. (2015) provided a detailed review of the factors that caused

widespread heavy rainfall and resulting catastrophic floods through climatological, meteorological, and hydrological analyses. The probability or likelihood of a flood/rainfall event is expressed as annual exceedance probability (AEP; the probability in a given year that an event of equal or greater magnitude will occur), or as its inverse, the average recurrence interval. According to NOAA Hydrometeorological Design Studies Center, several locations in the Front Range, including the city of Denver and Boulder, received 7-day (9-15 September 2013) rainfall accumulations equivalent to, at least, the 1,000-year event. Though numerous streamgages were destroyed during the event, the peak flows at several river networks were indirectly estimated by US Geological Survey (USGS) and their respective AEPs range from 50 to 500 years.

The procedures used to estimate AEPs corresponding to the magnitude of flood peak discharges or discharge volumes are referred to as flood frequency analysis (FFA). Given the full range of AEPs (from higher to lower), the notional classes of floods (e.g., ANCOLD, 2000; USBR, 1999) can be characterized into frequent ($AEP \geq 0.5$), rare ($AEP \geq 0.01$), very rare ($AEP \geq 0.001$), extreme ($AEP < 0.001$), and until the absolute upper limit of flood magnitude, Probable Maximum Flood (PMF), whose AEP cannot be readily assigned. (Frequency indicators inside parenthesis are intended as a guide; each situation is different and should be assessed individually.) PMF, a widely-used deterministic design approach, provides a single, limiting value of flood that could conceivably occur but is not well suited to risk-based analysis; thus it is beyond the scope of this study. On the other hand, the estimation of extreme floods, despite its difficulty, is central to many aspects of hydrologic engineering, including floodplain management, safety design, and maintenance and operation of dams. For instance, United States Bureau of Reclamation (USBR), the nation's largest wholesale water supplier and owner of over 350 storage dams in the western U.S., is committed to providing decision makers with risk-based information (i.e., FFA) for public safety practices (USBR, 2011a, 2011b).

Very rare to extreme floods (AEP between 10-2 and 10-7) lie in the range of creditable limit of extrapolation (i.e., justifiable estimates that are beyond the range of observations), which is dependent upon available data at the site of interest and FFA approaches (ANCOLD, 2000; USBR, 1999). For example, if a site has 50 years of streamflow records, then its typical creditable limit of extrapolation might be an AEP of 1 in 100. The ways to credibly extend the limit of extrapolation can be interpreted within the context of Bayes' Theorem (i.e., improve the belief by incorporating additional information): using paleoflood data, regional precipitation and flood data, and hydrological simulation to improve estimates for very rare and extreme (i.e., low AEP) flood frequency.

Existing approaches (tools) that incorporate these additional information (e.g., paleoflood and regional flood data) into estimates of extreme floods with AEPs include the Mixed-Population Graphical Approach (England et al., 2001), Expected Moments Algorithms (i.e., Bulletin 17C; see England et al.(2019) for a review), FLDFRQ3 (O'Connell, 1999; O'Connell et al., 2002), GRADEX (Naghettini et al., 1996), and Australian Rainfall-Runoff method (ARR; Nathan and Weinmann, 2000; James et al., 2019). In Australia, the ARR method, which couples extreme rainfall frequency analysis with rainfall-runoff modelling, is used to estimate flood frequency curves for very rare to extreme events ($AEP<0.001$). In the U.S., USBR provides seven methods, including all the aforementioned ones, to develop flood frequency curves and recommends utilizing a combination of them to reduce uncertainties in estimating FFA for the site of interest

(USBR, 2004, 2006). For instance, USBR utilized paleoflood data and several corresponding methods to conduct a hydrologic hazard analysis for dam safety evaluations for Los Banos Dam in California (Weghorst and Klinger, 2002) and Fresno Dam in Montana (Wickenberg, 1985).

Since the estimation of flood frequency with very low AEP ($\leq 10^{-3}$) is still a challenging area, innovative science, engineering and tools in extreme flood hydrology are always encouraged. England et al. (2014) estimated flood frequency up to 10^5-year recurrence intervals for the Arkansas River watershed (12,000 km^2) above Pueblo Dam, Colorado in the western U.S. using stochastic storm transposition (SST) technique and a distributed physics-based rainfall-runoff model. SST, a bootstrap resampling technique, can be used with relatively short (i.e., 10 to 20 years) gridded precipitation records to generate a large number of synthetic rainfall scenarios; when SST is combined with a hydrologic model, they can be used to perform a FFA out to at least 500-year recurrence intervals (e.g., England et al., 2014; Wright et al., 2014, 2017; Yu et al., 2019). England et al. (2014) also compared SST-based results with statistical FFA using at-site flood peak observations, estimates of several historical floods, and paleoflood data, highlighting the usage of multiple sources of probabilistic flood estimates to increase credibility in the results.

This study follows the direction of "integration of collaborative work in hydrometeorology, flood hydrology, and paleoflood hydrology" in England et al. (2014) for examining FFA with very low AEP. We focus on a small (357 km^2), mountainous watershed in the Colorado Front Range, in which snowmelt- and rainfall-driven floods coexist. We also attempt to extend England et al. (2014) by incorporating an improved SST model, implementing a more sophisticated hydrologic model, and utilizing stochastic watershed antecedent conditions in driving the event-based flood simulation. It is worth noting that this study also addresses two priority areas (PA) in USBR's strategic goals for long-term water resources planning and management (Brekke, 2011): PA 4.01 focuses on determining strengths and weaknesses of watershed hydrologic methods to support scoping decisions; PA 4.04 focuses on identifying strengths and weaknesses of spatially distributed weather data for hydrologic model development.

The central objectives of this study are: (1) develop, calibrate and validate a physics-based hydrologic model that can represent the long-term watershed runoff generating mechanisms and flooding behavior in the high-elevation basin of interest; (2) create SST-based storm catalogs from regional precipitation that provide extreme rainfall scenarios; (3) couple SST with a physics-based hydrologic model to derive process-based FFA (and flood volume analysis) up to 10^4-year recurrence intervals and compare results with statistical FFA using paleoflood data.

2. Study region and data

The study watershed of the Big Thompson River above Lake Estes (henceforth referred to as Big Thompson River watershed) is situated in northcentral Colorado (Figure 10a). An approximate 15,500 km^2, stochastic storm transposition (SST; Section 3.3) domain (red dashed line of Figure 10a), which encompasses the Big Thompson River watershed, stretches along the Colorado Front Range. The portion upstream of the US Geological Survey (USGS) stream gage above Lake Estes (gage number 06733000; watershed outlet) has a drainage area of 357 km^2, with elevations ranging from approximately 4300 m above the sea level (masl) in the west to 2300 masl near the Lake (Figure 10b). Mountainous area in the west of this watershed (above

3000 masl) can be characterized as intermittent snow-cover zones with peak snowpack accumulation during March and April; the area on the east part (below 3000 masl) seldom develops snowpack. According to the USGS 2016 National Land Cover Dataset, the land use in the watershed is 53% forest, 23% shrub, 14% pasture, and 10% other types including developed, barren land, open water and wetlands (Figure 10c).

The historical North American Land Data Assimilation System version 2 (NLDAS-2; Mitchell, 2004) hourly meteorological forcings from 1979 to 2018 are used for both long-term continuous and event-based hydrological simulation. Eight NLDAS-2 forcings are used, including precipitation, specific humidity, air temperature, atmospheric pressure, incoming shortwave and longwave radiation, and horizontal wind speed (zonal and meridional directions). These forcings are downscaled for the study watershed except precipitation, which are extracted for the entire SST domain. The measured Snow Telemetry (SNOTEL) snow water equivalent (SWE) at two stations as well as simulated basin-averaged SWE from Noah Land Surface Model (LSM) driven by NLDAS-2 (referred to here as NLDAS-Noah; Chen et al., 1996; Ek et al., 2003) is used for understanding snow processes. We also use hourly discharge observations for the USGS stream gage above Lake Estes (i.e., outlet) and annual peak flows for the outlet and two inner USGS stream gages (USGS 06732500 and 402114105350101; 103 and 102 km^2) to calibrate and validate the hydrologic model (Figure 10b).

3. Methodology

The process-based FFA framework presented in this study was also applied in Yu et al. (2019), except a more sophisticated, physics-based hydrologic model is used here (Section 3.1). This framework comprises continuous and event-based hydrological simulation combined with stochastic storm transposition (SST) using the RainyDay software. The framework is illustrated schematically in Figure 11 and briefly summarized in the following subsections.

3.1 Hydrologic model, calibration and continuous simulation

The Weather Research and Forecasting Hydrological modeling system (WRF-Hydro; Gochis et al., 2018), a physics-based, distributed hydrologic model, is used in this study. WRF-Hydro couples a land surface model (i.e., Noah-MP; Niu et al., 2011), with terrain routing (e.g., overland and subsurface flow) and channel routing (i.e., Muskingum-Cunge; Cunge, 1969) modules, and has been widely used in operational hydrological practices (e.g., Senatore et al., 2015; Yucel et al., 2015; Lin et al., 2018). WRF-Hydro was calibrated using a combination of manual and automated (i.e., DDS; Tolson and Shoemaker, 2007) approaches, targeted on capturing multiple aspects (e.g., long-term water balance, snowpack processes, runoff generation and routing dynamics) of hydrological processes (See Section 4.2).

Calibration and validation of WRF-Hydro aims to improve the overall performance with respect to multiple aspects of the dominant hydrological processes rather than to merely fit hydrographs of any particular events. The calibrated (and validated) WRF-Hydro model was run with NLDAS-2 forcings as inputs to a historical continuous simulation for the period from 1979 to 2018. The first 4 years were treated as "warm up" and the simulated results from 1983 to 2018 (36 years) were preserved at a daily time step and were subsequently used as watershed antecedent conditions for the event-based flood simulations.

3.2 Stochastic Storm Transposition

SST, a bootstrap resampling technique, was used to generate a large number of realistic precipitation scenarios for the study watershed via temporal resampling and spatial transposition of observed precipitation from the surrounding region (see Figure 10a for comparing the size between watershed and SST domain). Two main advantages of coupling SST into a process-based FFA framework are: (1) the spatiotemporal structures of observed precipitation are preserved, which is critical to flood simulation; (2) it enables the stochastic use of a hydrologic model that produces equally likely model output, accommodating sensitivity or uncertainty analyses. Wright et al. (2020) reviewed the development and various applications of SST since its origin.

In this study, we couple RainyDay (Wright et al., 2017), an open-source SST software, with 36 years (1983-2018) of NLDAS-2 precipitation to derive hypothetical but realistic precipitation events. The following steps describe how RainyDay is implemented in this study:

1. The SST domain (Figure 10a) refers to a homogeneous region over which the observed precipitation behaves statistically similarly. The SST domain in this study was initially sketched by the identifying a cluster of Global Historical Climatology Network (GHCN) stations characterized by similar values of L-kurtosis, L-skewness, the ratio of L-kurtosis to L-skewness, and discordancy (Hosking and Wallis, 1993, 1997). Then an SST domain encompassing these selected GHCN stations was delineated by extending the domain north and south of the Big Thompson River watershed along the Front Range of the Rockies, thus avoiding transposing storms from substantially different elevations.

2. RainyDay used 36 years of NLDAS-2, April-October (i.e., flood season) precipitation record to create a "storm catalog" that consists of the 360 most intense, 72-h precipitation events within the SST domain. These intense 72-h storms are in terms of rainfall accumulation over an area that has the similar size and shape of the Big Thompson River watershed.

3. RainyDay randomly selected k storms from the catalog to represent a synthetic year. The number of storm "arrivals", k, was calculated by RainyDay using the Poisson distribution with the rate parameter λ, defined as the ratio of total number of events in the storm catalog to the number of years in the record (i.e., $\lambda = 360/36 = 10$ storms yr^{-1}).

4. RainyDay uniformly transposed the selected k storms within the SST domain and computed the precipitation field over the Big Thompson River watershed, comprising a synthetic year of precipitation scenarios for the study watershed. If some of the k-transposed storms do not "hit" the watershed, their calculated rainfall is zero.
5. RainyDay can repeat steps 3-4 n times to create a single realization of n synthetic years of precipitation. These precipitation scenarios can be used directly to derive RFA or coupled with a hydrologic model to derive FFA. In this study, we created 10 realizations of 1000 synthetic year, 72-h rainfall scenarios.

3.3 Event-Based Flood Simulation

We selected five sets of watershed antecedent conditions drawn from the CS outputs to pair with the largest five out of the *k* transposed storms for each synthetic year to constitute event-based flood simulations. To ensure both realistic seasonality of floods and interannual variability in watershed conditions (i.e., dry or wet year), two criteria are required: (1) all five sets of watershed states were randomly selected from the same year from the CS; (2) their day of year are within ± 14 days of the corresponding date time of the precipitation events. Each paired event was run for 15 days with 72-h precipitation occur in the first 3 days and the largest peak and n-day volumes among five events represent the simulated annual flood peak and n-day maximum volumes. Thus, the event causes (instantaneous) flood peak can be different from the one results in 15-day peak volumes, obviating the "nested (or concurrent)" requirement.

We repeated this procedure for all 1000 synthetic years for each RainyDay-based precipitation realization to produce 1000 annual peak flows and n-day peak volumes. Their associated annual exceedance probabilities (AEPs) are calculated using the Weibull plotting position (Weibull, 1939). Ten realizations of simulated 1,000 annual peak flows can provide estimates of variability in FFA with recurrence intervals up to 1,000-year, or be treated as a single realization of 10,000 simulated annual peak flows to produce FFA with recurrence intervals up to 10^4-year. It also has to be noted that the former does not represent the "whole" population of flood outcomes while the later does not account for the uncertainties in the flood generating mechanisms associated with these flood magnitudes.

4. Results

4.1 The seasonality and magnitudes of regional extreme precipitation

The RainyDay selected 360 most extreme 72-h precipitation events are partitioned into rainfall and snowfall based on their corresponding air temperature (i.e., if the 72-h mean air temperature is below 273 K, this event is defined as snowfall). The seasonality (Figure 12a) and spatial distribution (Figure 12b) of these regional extreme precipitation events are shown. Temporally, the occurrence of these extreme precipitation events is, relatively, uniformly distributed between April and October, with two peaks clustering in late April (rainfall and snowfall) and early August (rainfall), respectively (Figure 12a). Spatially, the intensity and locations of these extreme precipitation events are also, relatively, uniformly distributed within the SST domain, revealing no relationship between precipitation and elevation within the SST domain (Figure 12b).

The total precipitation accumulation of most events range from 25 mm to 75 mm, with a median value of 40 mm (Figure 12a). These magnitudes are relatively small compared to the 72-h events from lower elevation in the eastern foothills (results not shown), consistent with Jarret and Costa (1988) that more moisture available at the lower than higher elevation can feed the convective storms. However, it also has to be noted that the most extreme rainfall event (i.e., the September 2013 event) occurred in the late summer (i.e., low occurrence density) resulted in an intensity (213 mm) 5 times as large as the median value.

4.2 WRF-Hydro Validation

In this study, WRF-Hydro is calibrated for a10-year period (2007-2016) and validated for the entire simulation period from 1983 to 2018. Long-term water balance derived from observations

(i.e., NLDAS-2 precipitation and USGS streamflows) shows that 40% of annual precipitation contributes to streamflow in average, which is well captured by the WRF-Hydro long-term CS (Figure 13a). The WRF-Hydro simulated annual mean evapotranspiration (ET; including sublimation) is 477 mm (i.e., 60% of precipitation contributes to ET), consistent with the estimated ratio of ET to precipitation from a nearby region in the Colorado Front Range (Sanford and Selnick, 2013; Sexstone et al., 2018). In addition, the inter-annual variabilities in precipitation (i.e., "gain") resembles the simulated variabilities in the sum of annual streamflows and ET (i.e., "loss"; error bars in Figure 13a).

Seasonality of snow water equivalent (SWE), represented by the watershed-averaged daily mean value from 1983 to 2018, are compared between WRF-Hydro and NLDAS-Noah simulations (Figure 13b). Though WRF-Hydro produces higher SWE than the NLDAS-Noah simulation in March and April, two model results show a consistent intra-annual variation, with the peak SWE in early April and a dominant snowmelt season in May and June. We also compared WRF-Hydro simulated SWE at two LSM grid cells (1km) against corresponding gage-based SNOTEL observations; WRF-Hydro simulated SWE for two grid cells were evaluated based on annual percent bias and long-term seasonality. The results for 2011-2015 period are consistent with simulated SWE in Sexstone et al. (2018), who used a process-based snow model with eddy covariance measurements.

WRF-Hydro simulated daily streamflows for Big Thompson River above Lake Estes adequately resemble the USGS daily observation (Figure 13c), obtaining a Kling-Gupta Efficiency (Gupta et al., 2009) of 0.87. WRF-Hydro simulated annual peak flows at three locations are compared with corresponding USGS observations for their available recording period (36 years for USGS 06733300, 9 years for USGS 06732500, and 21 years for USGS 402114105350101); no obvious bias in simulated annual peak flows is detected despite a slight underestimation for USGS gage 06732500 (Figure 13d). WRF-Hydro simulated 1-day flood volumes for Big Thompson River above Estes Park (the other two inner gages do not have long-term daily streamflow observations) also resemble USGS observed values (Figure 13e). Comparison for 3-, 5-, 7-, and 15-day flood volumes for the same gage (i.e., watershed outlet) between the WRF-Hydro simulation and USGS observations also shows consistent model performance (results are not shown). This "process-oriented" validation demonstrates the model credibility in capturing multiple aspects of the watershed hydrological processes and thus can be used for deriving the subsequent FFA.

4.3 Process-based understanding of historical annual peaks

A long-term (i.e., 36 years) CS using a physics-based hydrologic model and observed forcings can provide several hydrological variables (e.g., soil moisture, snowpack and ground temperature) that are always difficult to be measured on a watershed scale. (Even if measurements exist, they often do not represent model state variables.) These variables can be used as additional information other than precipitation and streamflows for understanding the flood regimes of historical events. To examine flood seasonality, we used observed USGS annual peak flows for Big Thompson River above Lake Estes, NLDAS-2 basin-averaged daily maximum precipitation, and simulated, basin-averaged, daily mean SWE and soil moisture (0-40 cm) for the 1983-2018 period. These historical flood events have been partitioned into snowmelt- and rainfall-driven events based on their simulated snowmelt accumulation during the

event. If the snowmelt is larger than 5 mm, this flood event is deemed to be snowmelt-driven; this criteria is applied through the study.

The maximum daily precipitation is roughly uniformly distributed along the year (Figure 14), consistent with the characteristics of the regional extreme precipitation (Figure 12). The dominant flood season is between May and June, coinciding with the high soil moisture period due to snowmelt (Figure 14). Flood events in this period are a mixture of snowmelt and rain-on-snow events (also referred to as snowmelt in this study); their resulting flood peaks cluster in a range between 15 and 50 $m^3 s^{-1}$ (Figure 14). However, the record flood (i.e., September 2013 event) at this watershed, surprisingly, occurred in a dry soil moisture period with no snowpack existing; a prolonged heavy rainfall resulted in a flood peak of 84 $m^3 s^{-1}$, which is roughly three times as large as the annual mean flood magnitude (i.e., 29 $m^3 s^{-1}$) and approximately 1.6 times as large as the previous record peak (i.e., 54 $m^3 s^{-1}$).

One snowmelt- and one rainfall-driven (i.e., the record flood) flood event are diagnosed, respectively, using detailed observations and the WRF-Hydro simulation (Figure 15). This snowmelt-driven flood event has a prononced diurnal cycle, stemming from the snowmelt dynamics as snowpack melts during the day time (i.e., higher temperature) and refreezes during the night (Figure 15a). This diurnal cycle, along with the snowmelt dynamics, implies that the magnitude of the corresponding flood peak is mainly determined by the snowmelt rate in a single day (i.e., energy limited). Furthermore, sparse rainfall, a small fraction (~20%) of snowcover and relative high antecedent soil mositure, exacerbate the river streamflows, increasing both the peak and volume. This snowmelt event is representative of all other snomelt-driven flood peaks (blue circles in Figure 14), thus reflecting WRF-Hydro well captures the snowmelt dynamics and the coressponding runoff generation.

The September 2013, rainfall-driven event occurred in a relative dry period with low antecednt soil moisutre but resulted in a dramatic increases in streamflows from around 2 to over 80 $m^3 s^{-1}$ within two days (Figure 15b) due to the intense, persistent, and widespread nature of the precipitation. The simulated top-layer soil moisture shows a consistent increase with the rainfall, reflecting high infiltration into the soil layer at the beginning of this event (Figure 15b). This is consistent with the runoff generation processes in this region, where the saturation excess ("Dunne") flows dominate (e.g., Miller et al., 2014, 2016; Rumsey et al., 2015). These unusual features, including dry soil moisture, extreme rainfall, and no snowmelt, set the September 2013 event apart from all other historical flood peaks, implying the existence of two different populations of floods in this watershed (also see Section 5).

We also noticed that WRF-Hydro underestimate both flood peak and volume for the September 2013 event, which is mainly due to the low bias in the NLDAS-2 precipitation. The 7-day watershed rainfall accumulation is 157 mm using NLDAS-2, which lies in the 50-year recurrence interval estimated by NOAA atlas 14. However, NOAA Hydrometeorological Design Studies Center estimated most parts of this watershed received 7-day rainfall accumulation equivalent to, at least, 100-year event, and 1,000-year for the lower watershed. In addition, Colorado Climate Center and National Weather Service estimated the 7-day rainfall accumulation between 203 and 254 mm for this watershed. In the meantime, this underestimation shows that our process-based calibration approach does not force the model to obtain a "correct answer" (i.e., similar

hydrograph) for the "wrong reasons" (i.e., compensating model parameters for the low biases in precipitation).

4.4 Flood frequency analysis

Process-based estimates of flood frequency distribution for Big Thompson River above Lake Estes using NLDAS-2 forcings from 1983 to 2018 are compared with the log-Pearson Type 3 (LP3) distribution fitted to 1983-2018 USGS observed annual peak flows (Figure 16). The LP3 estimated flood frequency curve is shown with 90% confidence interval. The process-based flood frequency curve is shown in different ways. Up to the 1,000-year recurrence interval, we show a range of flood frequency distributions based on 10 realizations representing uncertainty. Between the two- and 10,000-year recurrence intervals, we also show a single flood frequency distribution based on LP3 curve fit to historical observations. Finally, USGS observed annual peak flows from 1983 to 2018 are plotted in using the Cunnane plotting position (Cunnane, 1978).

Both process- and LP3-based flood frequency curves agree adequately well with USGS observed peak flows for $AEP \geq 0.05$, despite a slight underestimation in process-based estimates for $AEP \geq 0.33$. The process- and statistical-based estimates of flood frequency distribution relatively resemble each other until $AEP = 0.03$, beyond which the process-based FFA yields higher estimates. Though both process- and statistical-based approaches slightly underestimate the record flood, the process-based estimates tend to be closer to its magnitude. The recurrence interval for the record flood is estimated around 250-years using process-based approach while 2500-years using the statistical approach. It is important to note that the record flood was the result of persistent moisture transport to the north and west supported by a blocking ridge over the Canadian Rockies and a cutoff cyclonic circulation to the west of that, which lasted for days (Gochis et al., 2015). This mixture population challenges the validity of the statistical FFA approaches, which assume that the sampled data are representative of the same population (Cudworth, 1989; will be further discussed in Section 5)

The discrepancy between process- and statistical-based FFA for $\text{AEP} \leq 0.01$(i.e. upper tail; Figure 16) can be attributed to the fact that the different flood generating mechanisms are shaping the lower (i.e., common floods) and upper part of the flood frequency distributions. In Big Thompson River above Lake Estes, the FFA with lower AEP (i.e., frequent events) is dominant by snowmelt or rain on snow events while the upper tail of the FFA is controlled by extreme rainfall-driven events. However, the statistical FFA approach assume the sampled observations belong to the same population, thus neglecting the disparity in flood regime. In this watershed, the largest natural flow is estimated between 85 to 140 m^3/s during the last 8,000 to 10,000 year (Jarret and Costa, 1988; Jarrett, 1989), due to glacial melting during the post-glacial times. Understanding the flood regime (i.e., glacial melt) of this paleoflood partly explains why its nonexceedance bound lies within the 90% confidence interval of LP3 estimated FFA (Figure 16): it may represent the upper bound of snowmelt- or energy-driven flood events.

We also examined the range of rainfall- and snowmelt (rain on snow)-driven floods, respectively, by applying four deterministic flood simulations: two rainfall events and each paired with two different (i.e., wet and dry) watershed antecedent conditions. These four flood events are: (1-2) the transposed September 2013 rainfall event (212 mm) paired with summer

time high (low) soil moisture condition in early July (late September) generates flood magnitude of 273 (68) $m^3 s^{-1}$ (green error bar in Figure 16); (3-4) the transposed June 1987 rainfall event (41 mm) paired with late spring high (low) SWE in early June generates flood magnitude of 147 (53) $m^3 s^{-1}$ (blue error bar in Figure 16). The deterministic flood simulation reflects that the rainfall-driven event has a higher potential in governing the upper tail of the FFA than does rain on snow event and watershed antecedent condition plays a critical role in modulating flood peaks.

The upper bound of the deterministic rainfall-driven floods is almost twice larger than the maximum flood in stochastic simulation (i.e., flood frequency curve) due to an "unconfined criteria" for pairing rainfall with initial condition is applied. In stochastic simulation, the September 2013 rainfall event is paired with a watershed initial condition within a 28-day window while in deterministic simulation an approximate 90-day window (July-September) is used. The long "pairing" window enables this severe thunderstorm to fall in a relative high soil moisture in early July (i.e., right after the snowmelt season), reflecting the earlier a thunderstorm occurs in summer the larger impact it could result in.

To understand the flood regimes of observed and simulated flood peaks, the relationships between peak discharge and each physical drivers (i.e., precipitation, snowmelt and antecedent soil moisture) are investigated, respectively (Figure 17). For these historical flood peaks, their flood magnitude is from USGS observation and physical drivers are simulated via long-term CS (Section 4.3). In general, snowmelt plays a dominant role in driving most observed and simulated (i.e., synthetic) annual flood peaks (Figure 17c) and their derived flood magnitudes are higher than the rainfall-driven floods due to "wet" watershed antecedent condition. However, extreme rainfall, in terms of both 72-h accumulation and maximum hourly intensity, determines the magnitudes of severe floods (> 100 $m^3 s^{-1}$; Figure 17a and b). Furthermore, watershed antecedent soil moisture modulates both rainfall- and snowmelt-driven floods except for some extreme rainfall events which can cause large flood magnitudes even in the relative dry soil moisture condition (Figure 17d). The relationships between peak flows and the underlying drivers derived by the process-based simulation are generally consistent with the results based on observation, albeit with a far larger sample (34 events vs. 10,000 simulations; Figure 17)

4.5 Flood volume analysis

In addition to flood peaks, flood volumes of different durations are also the necessary aspects of FFA, which can be used for sizing the reservoir storage or evaluating the spillway adequacy for attenuating the inflows. The empirical flood frequency for n-day (i.e., 1-, 3-, 5-, 7-, and 15-day) flood volumes are derived, respectively, using USGS observed daily discharge from 1949 to 2018 and Cunnane plotting position (Cunnane, 1978; Figure 18). Six events with the largest 1-day flood volumes were examined for determining whether they possess the coincidence relationship or "equal severity" (Beard and Fredrich, 1975). A coincidence flood refers to an event that has the same recurrence interval for its 1-day and n-day flood volumes (Balocki and Burges, 1994). For example, if 100-year 1-, 3-, 5-, 7-, and 15-day flood volumes are resulted from the same event, this event has the coincidence of frequency. (The reason for examining the six largest floods was to include the September 2013 event.)

The six largest 1-day flood volumes occurred in June-1965, June-1995, June-1957, June-1949, June-2010 and September-2013, respectively; none of them shows the coincidence relationship between recurrence interval and n-day flood volumes (Figure 18a). For example, the 1949 flood event increases from the 4th largest 1-day flood volume to the largest 15-day flood volume. Conversely, four events out of six show a (roughly) consistent drop in recurrence intervals from 1- to 15-day flood volumes. The recurrence intervals of the 2013 event, which is the record flood with respect to peak flows, drops from over 10-year for 1-day to 2-year for 15-day volumes. Recall from Section 4.3 and Figure 15b that we found the 2013 flood event occurred in a relative dry period and no additional snowmelt for contributing the flood volume. In this watershed, the lack of coincidence relationship challenge the conventional design flood hydrograph that assumes 1:1 equivalence of recurrence intervals among n-day flood volumes.

The process-based flood frequency analysis for 1-, 3-, 5-, 7-, and 15-day flood volumes are derived, respectively (Figure 18b). The sources of these derived flood volumes include 360-h precipitation and watershed antecedent conditions (mainly the SWE). The process-based estimates of flood volume frequency distributions resemble the observations (i.e., empirical distribution) for all durations except a slight underestimation for the 15-d volume for the small return periods (Figure 18b). The process-based estimates for n-day flood volumes resemble the observation-based results because our approach accounts for snowmelt, which is the non-negligible contributors to long-term flood volumes. In the meantime, the process-based flood volume analysis also obviates the procedures for obtaining long-term flood hydrograph, which is an inevitable steps in design flood approach to temporally distribute flood volumes of a given recurrence interval.

5. Discussion

The September 2013 flood event, or the record flood, in the Big Thompson River watershed above Lake Estes demonstrated how the flood generating mechanisms of the extreme events are different from the common ones. However, how much information can be extracted from this record flood in terms of FFA, to some extent, depends on the methods applied. Statistical FFA approaches simply treat this event as one additional sample, neglecting the fact that this record flood belongs to a different population of flood agents (severe rainfall-driven) compared to all other observed events (snowmelt-driven). Consequently, their derived flood frequency distribution, especially for the upper tail, are substantially affected by the common snowmelt-driven floods, which is hydrologically inconsistent (Klemes, 1986). In addition, mixing different populations of flood agents violates the underlying assumption of statistical FFA approaches that require the sampled data represent the same population. It also has to be noted that the unusual characteristics of the record floods and their markedly influence on flood frequency distribution actually occurred in a variety of watersheds in the United States (Smith et al., 2018).

On the contrary, process-based FFA approaches utilize the “upstream” information of historical floods (i.e., physical drivers), and attempt to restructure the full spectrum of hydrologic responses, from which the flood frequency distribution is derived. Therefore, the upper tail of the process-based FFA is reliant on the physical process and watershed states that control floods. According to the 10,000 derived synthetic annual peak flows (Figure 16 and 7), the upper tail of flood peaks (i.e., AEP ≤ 0.01) can be reached by two different combinations: (1) extreme summertime rainfall (somewhat) regardless of antecedent soil moisture; (2) late spring rainfall

with high hourly intensity combined with wet soil moisture or snowmelt condition. The former is the similar process driving the record flood in Big Thompson River watershed while the latter is nonexistent in observed flood peaks, probably stemming from limited sample size (34 events).

The September 2013 flood event also seems to be an exception of Jareet's 2,300 m-hypothesis (Jareet, 1988), which argued that 2,300 m is the upper limit of the occurrence of extreme orographic rainfall-driven flood events. (The elevation of the Lake Estes, the outlet of our study watershed, is approximate 2,300 m.) This remarkable exception also provides insights into one particular aspect of nonstationary flood frequency-the changing flood generating mechanisms. Though future studies are needed to evaluate changes in extreme flood agents, we believe this event should not be simply treated as an "outlier".

6. Summary and Conclusions

In this study, we used regional precipitation, stochastic storm transposition (SST), and a physics-based hydrological model to provide an alternative approach for analyzing frequency distributions of flood peaks and volumes. We applied this process-based FFA framework to Big Thompson River, a small mountainous watershed in the Colorado Front Range that is exposing two different flood regimes in shaping the flood frequency distributions. Key findings include:

1. Though the Big Thompson River watershed is located in the complex mountainous terrain, we delineated an SST domain encompassing the watershed which can be interpreted as a meteorologically homogeneous region. This delineated SST domain is roughly 45 times larger than the watershed in size but has a relative narrow elevational bands ranging from 1,500 m to 4,000 m. We detected no obvious elevation-precipitation relationship using annual maximum 72-h precipitation for all NLDAS-2 grids within the domain; the L-moments-based discordancy and heterogeneity tests further improves the credibility in transposing precipitation within this SST domain. Validating the regional homogeneity is an indispensable step for applying SST or other regional rainfall analysis tools into a mountainous region.

2. We utilized a "process-oriented" approach to calibrate the WRF-Hydro model using 10-year continuous simulation and cross validate it over the entire 1983-2018 (36-year) period. The calibrated and validated WRF-Hydro adequately captures long-term water balances, and the seasonality of SWE; the model simulated long-term hydrographs, annual peak flows and flood volumes for different durations well matches the results based on USGS observations. The well-performed WRF-Hydro enables the generation of historical watershed states (e.g., soil moisture, SWE, streamflow and ponding water depth), which are usually difficult to be field measured, via long-term continuous simulation using observed NLDAS-2 forcings. These hydrological variables can provide model antecedent conditions for performing event-based flood simulation and deriving FFA.

3. We also demonstrated that these long-term continuous simulated watershed states can help to better understand the flood generating mechanisms associated with historical annual peaks. Conventionally, the time of year is taken as an expedient method to qualitatively understand the dominant flood drivers of historical events (e.g., Cudworth,

1989). Our "process-based" diagnoses can explicitly show how snow dynamics and its diurnal cycle affecting the snowmelt-driven floods and how extreme convective storm shaping the rainfall-driven floods regardless the dry watershed antecedent conditions. The WRF-Hydro model is able to reconstruct historical annual peaks with respect to their climatological, meteorological, and hydrological processes.

4. Coupling SST-based rainfall scenarios, with a physics-based hydrological model and derived watershed antecedent conditions can perform process-based frequency analysis of flood peaks and volumes. The derived frequency distributions for flood peaks and volumes can reach the AEP as low as 10^{-4}. In addition, process-based frequency distributions for both flood peaks and volumes resemble the empirical distribution using USGS observations and Cunnane plotting position. The upper tail of process-based flood quantiles (i.e., 10^4-year) is higher than largest natural flood flow during the last 10,000 years using paleoflood data due to different flood regimes.

5. Process-based FFA provides an opportunity to understand the generating mechanisms of the flood quantiles, especially for the upper tail, which conventional FFA fails to do. In this watershed, the simulated upper tail flood quantiles can be caused by extreme rainfall regardless of initial soil moisture or snowpack (similar to record flood), or by intermediate rainfall combined with snowmelt and high soil moisture (never happened in the past). Process-based FVA obviates the coincidence of floods occur when the flood peaks, 1-day and n-day flood volumes have the same recurrence intervals, which is always assumed by design floods. The flood characteristics of the study watershed violate this assumption since the snowmelt-driven floods control the flood volume with a long duration while the extreme rainfall-driven floods govern the short-term flood volume and flood peaks.

References

Australian National Committee on Large Dams (ANCOLD): Guidelines on Selection of Acceptable Flood Capacity for Dams. [online] Available from: https://www.ancold.org.au/?product=guidelines-on-selection-of-acceptable-flood-capacity-for-dams-2000 (Accessed 12 May 2020), 2000.

Balocki, J. B. and Burges, S. J.: Relationships between n -Day Flood Volumes for Infrequent Large Floods, J. Water Resour. Plan. Manag., 120(6), 794–818, doi:10.1061/(ASCE)0733-9496(1994)120:6(794), 1994.

Beard, L. R. and Fredrich, A. J.: Hydrologic Engineering Methods For Water Resources Development, U.S. Army Corps of Engineers. Institute of Water Resources. Hydrologic Engineering Center, Davis, CA. [online] Available from: https://www.hec.usace.army.mil/publications/IHDVolumes/IHD-3.pdf, 1975.

Brekke, L. D.: Addressing Climate Change in Long-Term Water Resources Planning and Management: User Needs for Improving Tools and Information Title, Technical Report, U.S. Army Corps of Engineers Bureau of Reclamation, U.S. Department of the Interior., 2011.

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., Duan, Q. Y., Ek, M. and Betts, A.: Modeling of land surface evaporation by four schemes and comparison with FIFE observations, J. Geophys. Res. Atmospheres, 101(D3), 7251–7268, doi:10.1029/95JD02165, 1996.

Cudworth, A. G.: Flood Hydrology Manual, U.S. Department of the Interior, Bureau of Reclamation, Denver Office., 1989.

Cunge, J. A.: On The Subject Of A Flood Propagation Computation Method (Musklngum Method), J. Hydraul. Res., 7(2), 205–230, doi:10.1080/00221686909500264, 1969.

Cunnane, C.: Unbiased plotting positions — A review, J. Hydrol., 37(3–4), 205–222, doi:10.1016/0022-1694(78)90017-3, 1978.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G. and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res. Atmospheres, 108(D22), 2002JD003296, doi:10.1029/2002JD003296, 2003.

England, J. F., Klinger, R. E., Camrud, M. and Klawon, J. E.: Guidelines for preparing preliminary flood frequency analysis reports for comprehensive facility reviews, Bureau of Reclamation, Denver, Colorado., 2001.

England, J. F., Julien, P. Y. and Velleux, M. L.: Physically-based extreme flood frequency with stochastic storm transposition and paleoflood data on large watersheds, J. Hydrol., 510, 228–245, doi:10.1016/j.jhydrol.2013.12.021, 2014.

England, J. F., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas Jr., W. O., Veilleux, A. G., Kiang, J. E. and Mason, Jr., R. R.: Guidelines for determining flood flow frequency—Bulletin 17C, USGS Numbered Series, U.S. Geological Survey, Reston, VA., 2019.

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. and Brown, B.: The Great Colorado Flood of September 2013, Bull. Am. Meteorol. Soc., 96(9), 1461–1487, doi:10.1175/BAMS-D-13-00241.1, 2015.

Gochis, D., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., McCreight, J., Mills, J., Pan, L., RafieeiNasab, A., Read, L., Sampson, K., Yates, D., Yu, W. and Zhang, Y.: The WRF-Hydro modeling system technical description, (Version 5.0), UCAR/NCAR., 2018.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, Water Resour. Res., 29(2), 271–281, doi:10.1029/92WR01980, 1993.

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

James, B., Mark, B., Rory, N., William, W., Erwin, W., Monique, R. and Isabelle, T., Eds.: Australian Rainfall and Runoff: A Guide to Flood Estimation., 2019.

Jarret, R. D. and Costa, J. E.: Evaluation of the flood hydrology in the Colorado Front Range using precipitation, streamflow, and paleoflood data for the Big Thompson River basin., 1988.

Jarrett, R. D.: Hydrology and paleohydrology used to improve the understanding of flood hydrometeorology in Colorado, Des. Hydraul. Struct., 89, 9–16, 1989.

Lin, P., Hopper, L. J., Yang, Z.-L., Lenz, M. and Zeitler, J. W.: Insights into Hydrometeorological Factors Constraining Flood Prediction Skill during the May and October 2015 Texas Hill Country Flood Events, J. Hydrometeorol., 19(8), 1339–1361, doi:10.1175/JHM-D-18-0038.1, 2018.

Miller, M. P., Susong, D. D., Shope, C. L., Heilweil, V. M. and Stolp, B. J.: Continuous estimation of baseflow in snowmelt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach, Water Resour. Res., 50(8), 6986–6999, doi:10.1002/2013WR014939, 2014.

Miller, M. P., Buto, S. G., Susong, D. D. and Rumsey, C. A.: The importance of base flow in sustaining surface water flow in the Upper Colorado River Basin: UCRB BASE FLOW DISCHARGE TO STREAMS, Water Resour. Res., 52(5), 3547–3562, doi:10.1002/2015WR017963, 2016.

Mitchell, K. E.: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res., 109(D7), D07S90, doi:10.1029/2003JD003823, 2004.

Naghettini, M., Potter, K. W. and Illangasekare, T.: Estimating the upper tail of flood-peak frequency distributions using hydrometeorological information, Water Resour. Res., 32(6), 1729–1740, doi:10.1029/96WR00200, 1996.

Nathan, R. and Weinmann, E.: Book VI:Estimation of large and extreme floods, In Australian Rainfall and Runoff: A Guide to Flood Estimation, Engineers Australia, Canberra., 2000.

National Weather Service (NWS): Exceedance Probability Analysis for the Colorado Flood Event, 9 - 16 September 2013, Hydrometeorological Design Studies Center, Silver Spring, Maryland. [online] Available from: https://www.nws.noaa.gov/oh/hdsc/aep_storm_analysis/8_Colorado_2013.pdf, 2013.

National Weather Service (NWS): The Record Front Range and Eastern Colorado Floods of September 11–17, 2013, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland., 2014.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res., 116(D12), D12109, doi:10.1029/2010JD015139, 2011.

O'Connell, D. R. H.: FLDFRQ3: Three-parameter maximum likelihood flood-frequency estimation with optional probability regions using parameter grid integration, Denver, Colorado., 1999.

O'Connell, D. R. H., Ostenaa, D. A., Levish, D. R. and Klinger, R. E.: Bayesian flood frequency analysis with paleohydrologic bound data: BAYESIAN FLOOD FREQUENCY, Water Resour. Res., 38(5), 16-1-16–13, doi:10.1029/2000WR000028, 2002.

Rumsey, C. A., Miller, M. P., Susong, D. D., Tillman, F. D. and Anning, D. W.: Regional scale estimates of baseflow and factors influencing baseflow in the Upper Colorado River Basin, J. Hydrol. Reg. Stud., 4, 91–107, doi:10.1016/j.ejrh.2015.04.008, 2015.

Sanford, W. E. and Selnick, D. L.: Estimation of Evapotranspiration Across the Conterminous United States Using a Regression With Climate and Land-Cover Data 1: Estimation of Evapotranspiration Across the Conterminous United States Using a Regression with Climate and Land-Cover Data, JAWRA J. Am. Water Resour. Assoc., 49(1), 217–230, doi:10.1111/jawr.12010, 2013.

Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N. and Kunstmann, H.: Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales: FULLY COUPLED ATMOSPHERE-HYDROLOGY MODEL, J. Adv. Model. Earth Syst., 7(4), 1693–1715, doi:10.1002/2015MS000510, 2015.

Sexstone, G. A., Clow, D. W., Fassnacht, S. R., Liston, G. E., Hiemstra, C. A., Knowles, J. F. and Penn, C. A.: Snow Sublimation in Mountain Environments and Its Sensitivity to Forest Disturbance and Climate Warming, Water Resour. Res., 54(2), 1191–1211, doi:10.1002/2017WR021172, 2018.

Tolson, B. A. and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration: DYNAMICALLY DIMENSIONED SEARCH ALGORITHM, Water Resour. Res., 43(1), doi:10.1029/2005WR004723, 2007.

U.S. Bureau of Reclamation (USBR): A framework for developing extreme flood inputs for dam safety risk assessments, Draft., 1999.

U.S. Bureau of Reclamation (USBR): Hydrologic Hazard Curve Estimating Procedures., 2004.

U.S. Bureau of Reclamation (USBR): Guidelines for Evaluating Hydrologic Hazards., 2006.

U.S. Bureau of Reclamation (USBR): Dam Safety Public Protection Guidelines, Denver, Colorado., 2011a.

U.S. Bureau of Reclamation (USBR): Dam Safety Public Protection Guidelines – Examples of Use, Denver, Colorado., 2011b.

Weghorst, K. M. and Klinger, R. E.: Preliminary Hydrologic Loadings for Los Banos and Little Panoche Dams, California, Flood Hydrology Group, Bureau of Reclamation, Denver, Colorado., 2002.

Weibull, W.: A statistical theory of strength of materials, IVB-Handl, (151), 1–45, 1939.

Wickenberg, D. F.: Fresno Dam Probable Maximum Flood, Bureau of Reclamation, Great Plains Regional Office, Billings, Montana., 1985.

Wright, D. B., Smith, J. A. and Baeck, M. L.: Flood frequency analysis using radar rainfall fields and stochastic storm transposition, Water Resour. Res., 50(2), 1592–1615, doi:10.1002/2013WR014224, 2014.

Wright, D. B., Mantilla, R. and Peters-Lidard, C. D.: A remote sensing-based tool for assessing rainfall-driven hazards, Environ. Model. Softw., 90, 34–54, doi:10.1016/j.envsoft.2016.12.006, 2017.

Wright, D. B., Yu, G. and England, J. F.: Six decades of rainfall and flood frequency analysis using stochastic storm transposition: Review, progress, and prospects, J. Hydrol., 585, 124816, doi:10.1016/j.jhydrol.2020.124816, 2020.

Yu, G., Wright, D. B., Zhu, Z., Smith, C. and Holman, K. D.: Process-based flood frequency analysis in an agricultural watershed exhibiting nonstationary flood seasonality, Hydrol. Earth Syst. Sci., 23(5), 2225–2243, doi:10.5194/hess-23-2225-2019, 2019.

Yucel, I., Onen, A., Yilmaz, K. K. and Gochis, D. J.: Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall, J. Hydrol., 523, 49–66, doi:10.1016/j.jhydrol.2015.01.042, 2015.

Figures

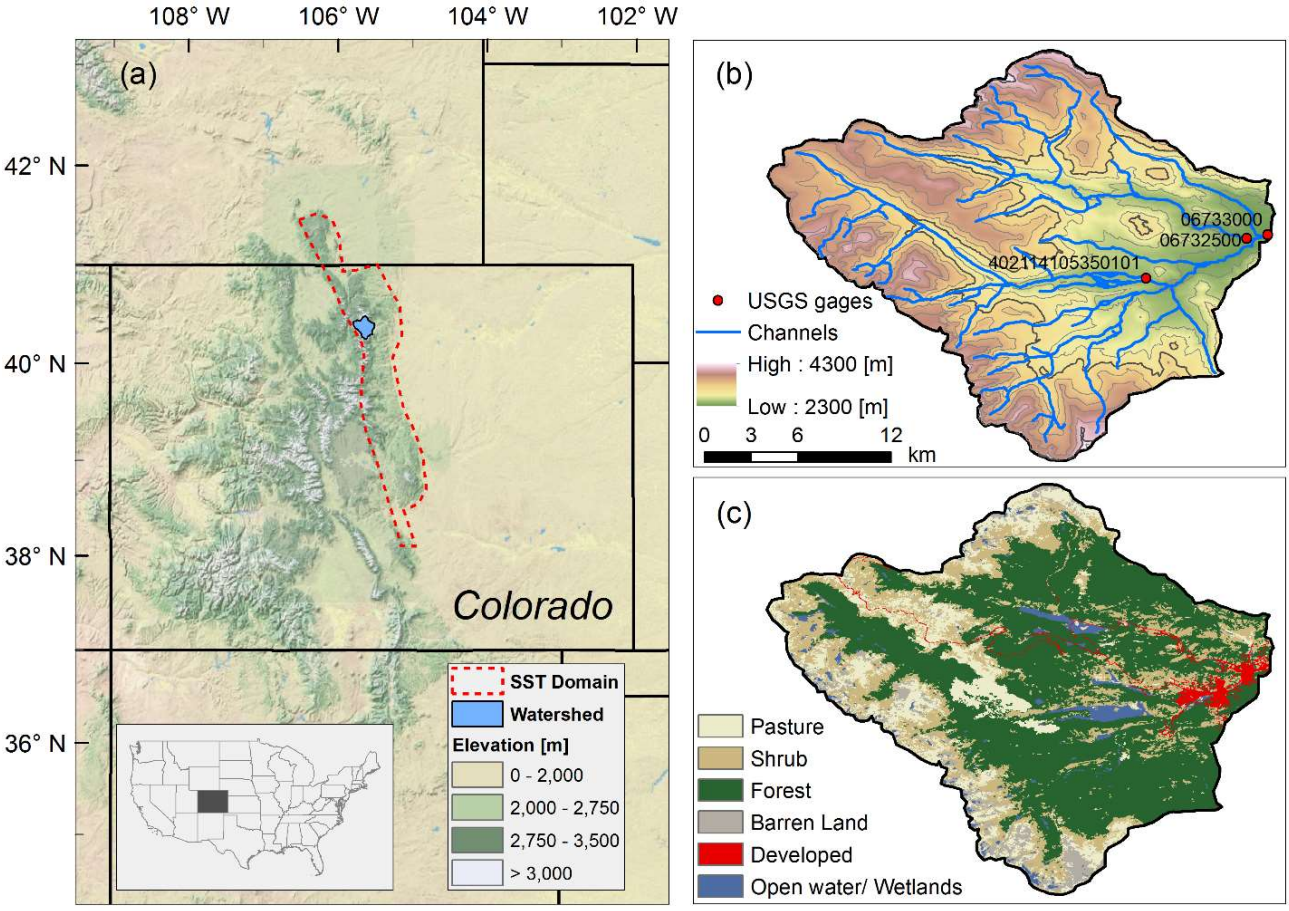

Figure 10. (a) Study region showing Colorado, the Big Thompson River watershed, and the extent of the stochastic storm transposition domain. The inset in (a) shows conterminous United States with the state of Colorado highlighted in grey. The Big Thompson River watershed showing land surface elevation and river channels (b), and NLCD 2016 land use (c).

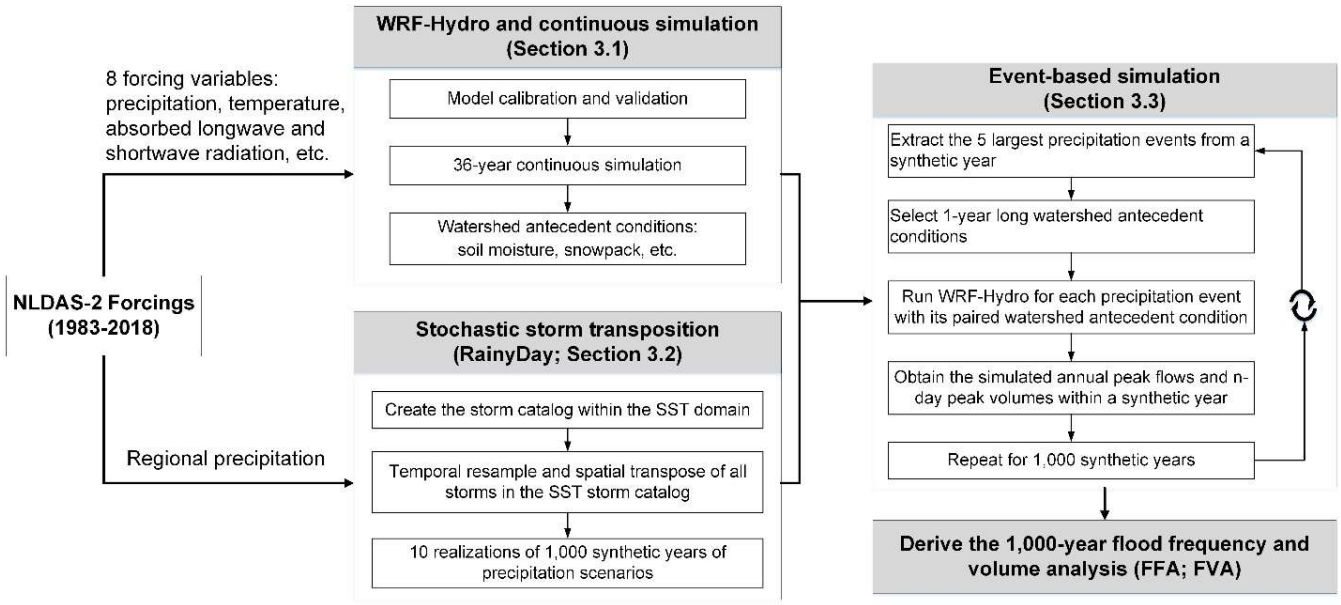

Figure 11. Flowcharts of the process-based flood frequency and volume analysis approach.

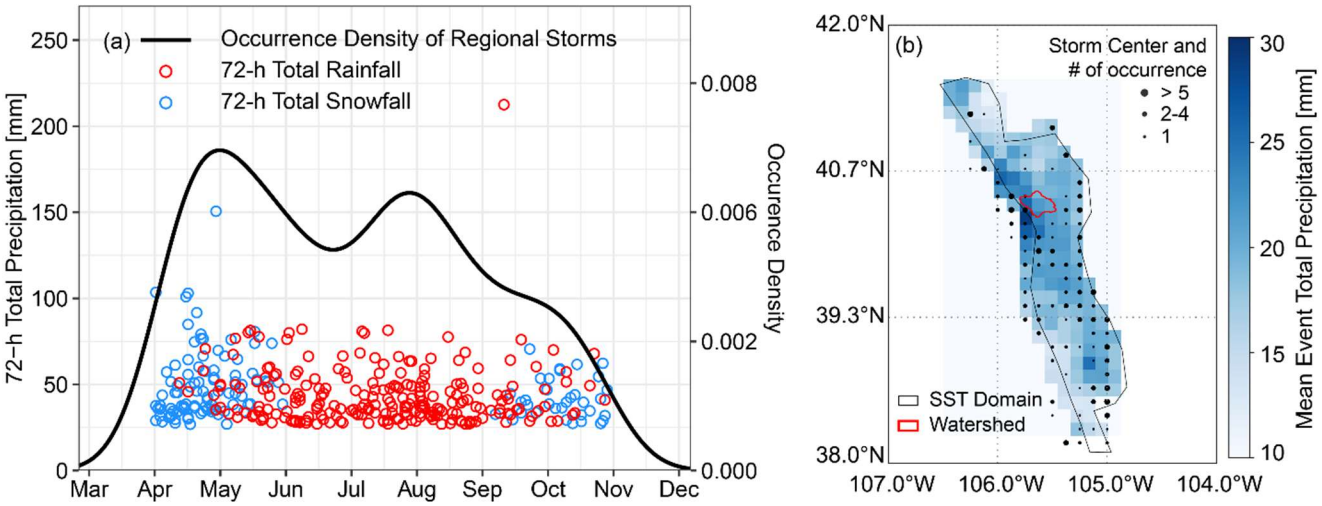

Figure 12. The seasonality (a) and spatial distribution (b) of the extreme precipitation in the SST derived storm catalog. Black dots in (b) are the locations of the centers of 360 storms and their size represent the number of occurrence.

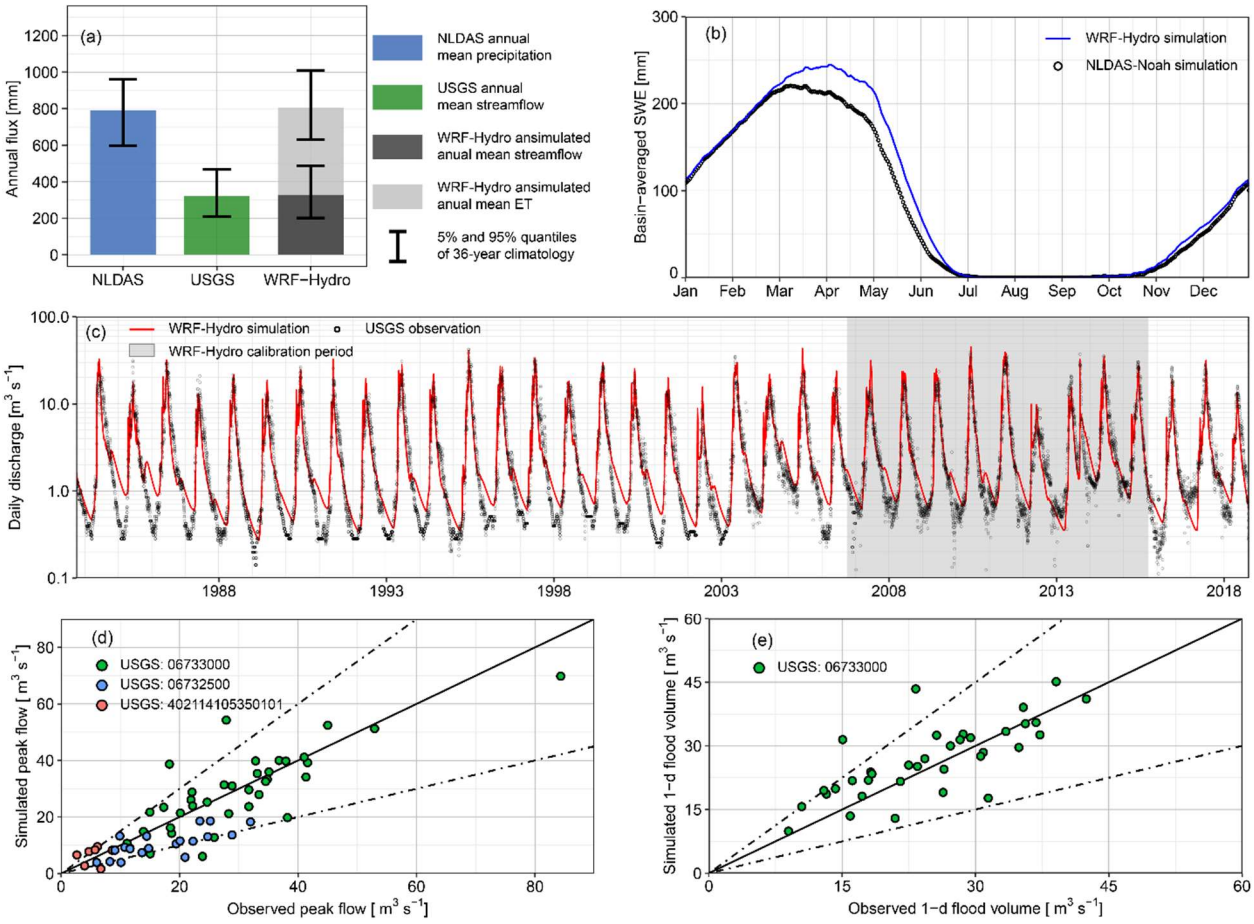

Figure 13. WRF-Hydro model validation in terms of long-term (36 year) water balance (a), SWE climatology (b), daily hydrograph (c), annual peak flows (d) and maximum 1-day flood volume (e). The validation period is from water year 1983 to 2018, which overlaps calibration period, 2007-2016. The 1-day flood volume in d is represented as daily mean discharge.

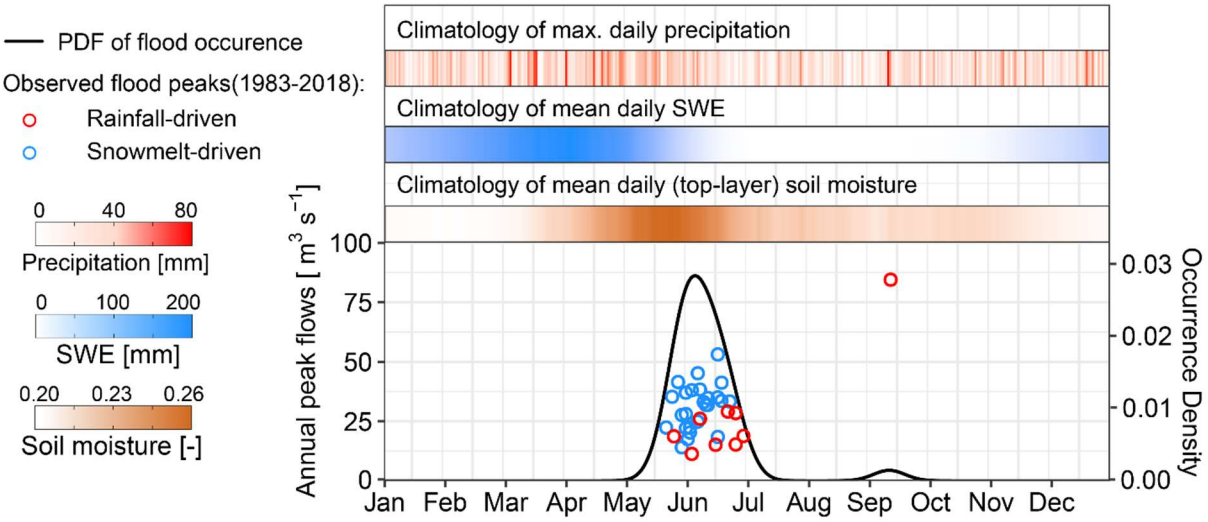

Figure 14. The seasonality of annual flood peaks and the climatology of maximum daily precipitation, mean daily snowmelt rate and mean daily soil moisture.

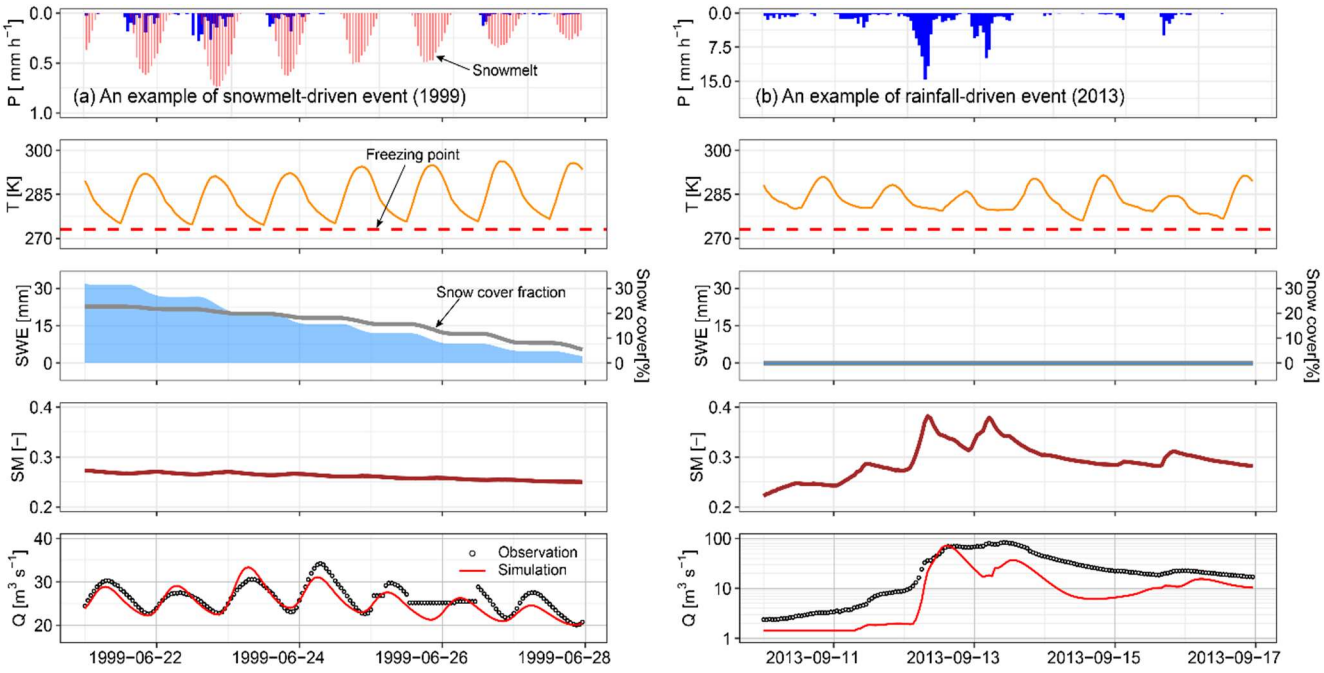

Figure 15. Assessment of two historical peak floods using observed (USGS) streamflow and (NLDAS-2) precipitation and simulated basin-averaged snowmelt, SWE, snowcover fraction, ground temperature, streamflow, soil moisture, and streamflows at the watershed outlet. The 1999 flood event (a) was driven by snowmelt while the 2013 flood event was extreme rainfall-driven.

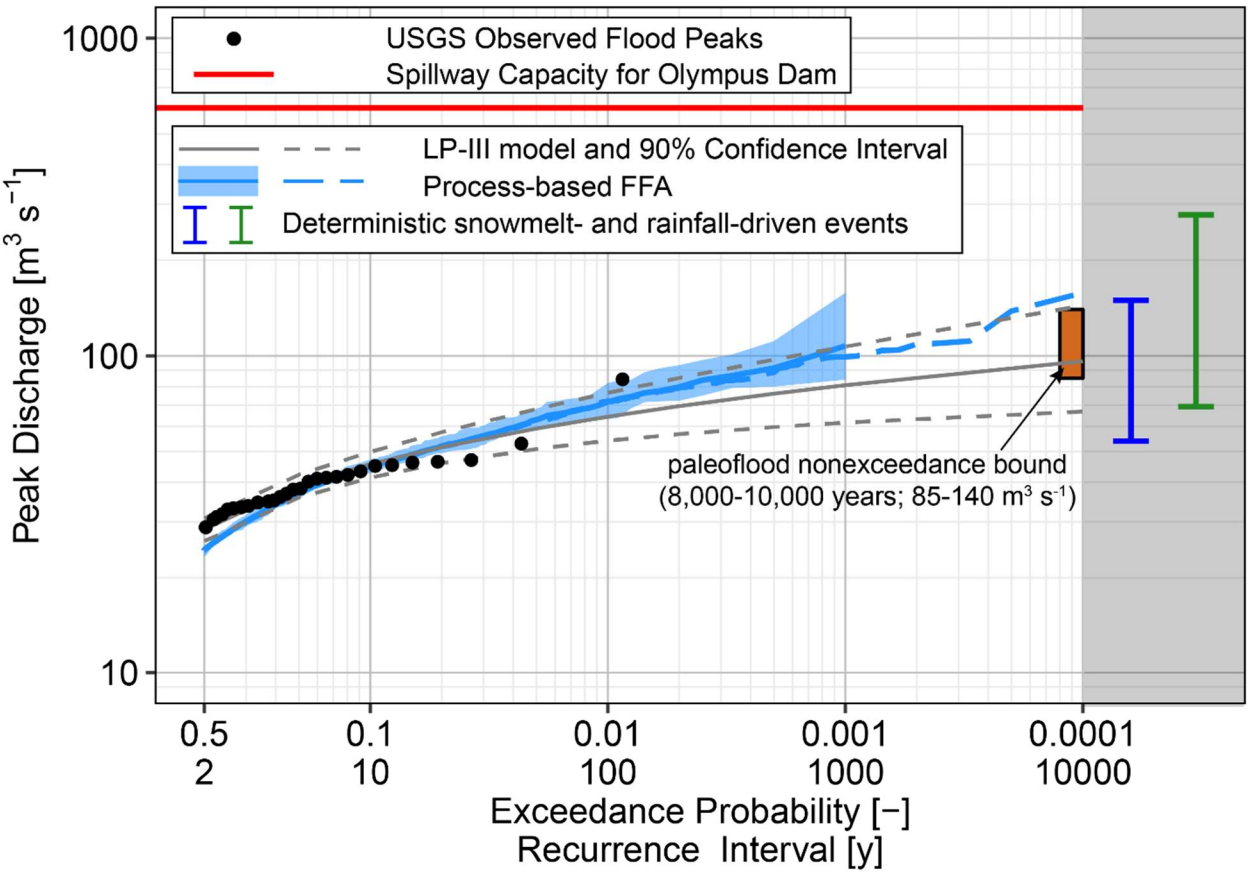

Figure 16. Flood frequency analyses for Big Thompson River above Lake Estes using statistical and process-based approaches. The process-based FFA are shown in two ways: 1,000 year recurrence interval with 10 realizations (blue solid line and shaded areas); 10,000 year recurrence interval with single realization (blued dashed line). Gray dashed lines represent 90% confidence intervals for LP3-based flood frequency curves. Paleoflood nonexceedance bound data is shown as filled gray rectangles. Blue (green) error bar represents the one snowmelt (rainfall)-driven flood range using the deterministic approach.

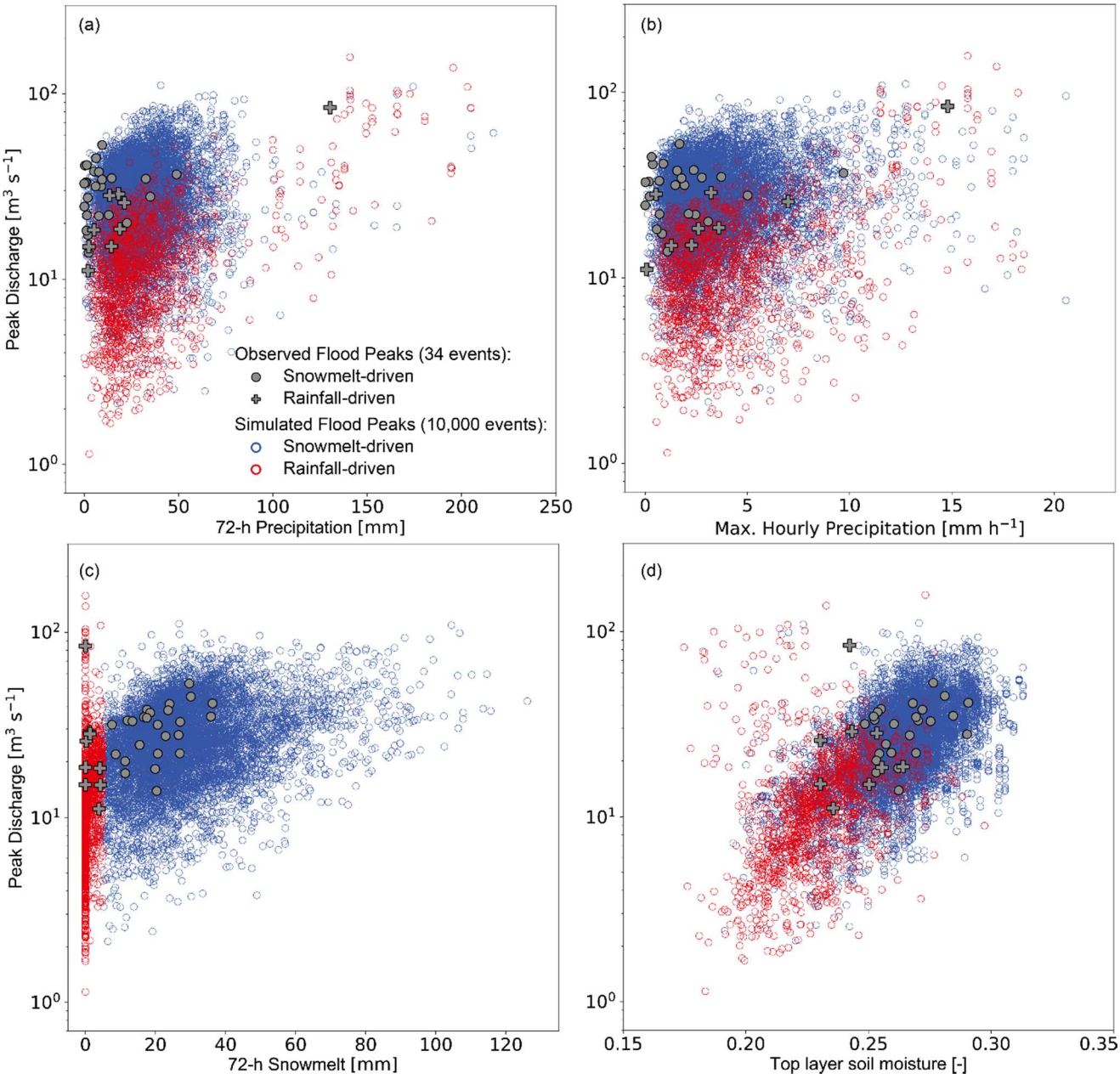

Figure 17. The relationships between peak discharge and 72-h precipitation (a), maximum hourly precipitation rate (b), 72-h snowmelt (c) and antecedent top-layer soil moisture (d). Both observed and simulated annual peak discharge are shown here and grouped into rainfall- and snowmelt-driven events.

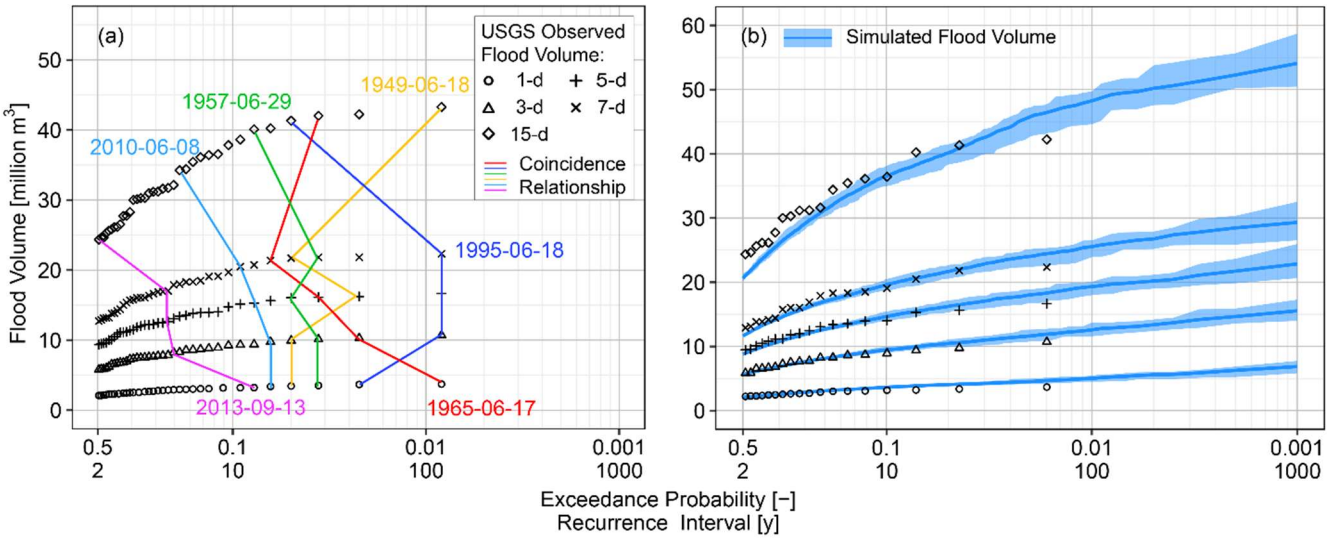

Figure 18. Empirical ploting position for 1-, 3-, 5-, 7-, and 15-day flood volumes are compired with coincidence relationship of flood volumes (a) and process-based FFA (b). The conincidence relationship examined in (a) are for the six largest 1-day flood volumes.