Six Decades of Rainfall and Flood Frequency Analysis Using Stochastic Storm Transposition: Review, Progress, and Prospects

Daniel B. Wright, PhD[1,*]

Guo Yu[1]

John F. England, PhD[2]

[1]University of Wisconsin-Madison, Madison, Wisconsin, USA

[2]U.S. Army Corps of Engineers, Denver, Colorado, USA

[*]Corresponding Author: danielb.wright@wisc.edu

Submitted to *Journal of Hydrology*

Abstract

Stochastic Storm Transposition (SST) involves resampling and random geospatial shifting (i.e. transposition) of observed storm events to generate hypothetical but realistic rainstorms. Though developed as a probabilistic alternative to probable maximum precipitation (PMP) and sharing PMP's storm transposition characteristic, SST can also be used in more typical rainfall frequency analysis (RFA) and flood frequency analysis (FFA) applications. This paper explains the method, discusses its origins and linkages to both PMP and RFA/FFA, and reviews the development of SST research over the past six decades. Discussion topics includes: the relevance of recent advances in precipitation remote sensing to frequency analysis, numerical weather prediction, and distributed rainfall-runoff modeling; uncertainty and boundedness in rainfall and floods; the flood frequency challenges posed by climatic and land use change; and the concept of multi-scale flood frequency. Recent literature has shown that process-based multiscale FFA, in which the joint distributions of flood-producing meteorological and hydrological processes are synthesized and resolved using distributed physics-based rainfall-runoff models, provides a useful framework for translating nonstationary hydroclimatic conditions into flood frequency estimates. SST pairs well with the process-based approaches. This pairing is promising because it can leverage advances from other branches of hydrology and hydrometeorology that appear to be difficult to integrate into better-known RFA and FFA approaches. The paper closes with several recommendations for future SST research and applications.

Keywords: Extreme rainfall; floods; rainfall frequency analysis; flood frequency analysis; rainfall remote sensing; stochastic hydrology

Highlights:

1. SST is a storm-based alternative to statistical rainfall and flood frequency analysis
2. The SST methodology and six decades of related research are reviewed
3. SST is able to leverage developments in related fields
4. SST and rainfall runoff modeling can address nonstationary flood frequency
5. Recommendations for future SST research are provided

1. Introduction

The estimation of flood flows emerged in the early twentieth century as a core challenge in hydrology, spurred by a dam-building boom which in the U.S. ran from roughly 1910 to the 1970s. Beyond dam spillway design, these estimates are used for the sizing of stormwater and flood control infrastructure and for floodplain mapping. The objective is usually to determine a flood quantile, i.e. the peak discharge or volume associated with a particular annual exceedance probability (*AEP*); a flood with *AEP* = 0.01 corresponds to the "100-year flood." The quantile estimation process is referred to as flood frequency analysis (FFA); the related practice for rainfall is referred to as rainfall frequency analysis (RFA).

There are two widely-known approaches to FFA (National Research Council, 1988). The first, flood-based statistical FFA, involves fitting a probability distribution to extreme values, typically annual maxima, of multi-decadal streamflow records. The desired quantile can then be obtained from that distribution. The second uses representations of one or more rainstorms as input to a rainfall-runoff hydrologic model to produce simulated flood peaks or hydrographs. The most common starting point for model-based approaches is intensity-duration-frequency (IDF) information, which describes the probability distribution of extreme rainfall depths or rates and is generated using similar methods to statistical FFA. Idealized or observed spatial or temporal patterns are often used to disaggregate a rainfall quantile into a more realistic hypothetical storm. The resulting design storm is then used as input to a rainfall-runoff model which has been initialized using a prescribed soil moisture condition. This combination of IDF and a rainfall-runoff model with an assumed initial soil moisture is referred to as the "design storm method" (e.g. Curtis et al., 2013a,b; Packman and Kidd, 1980).

Both IDF estimation and flood-based statistical FFA are “station-based,” in that the observations and resulting predictions are made at individual, fixed locations—rain gages or stream gages (National Research Council, 1994). While a rain gage measurement may reflect the passage of a rainstorm and a streamflow measurement may reflect the result of that rainstorm’s interactions with hillslopes and river networks, neither observation explicitly considers the complex space-time rainfall structure and its interactions with watershed feature such as varying terrain and river channels.

IDF estimation and flood-based statistical FFA face certain limitations: 1.) long-term records of rainfall or flood extremes may not yield accurate *AEP* estimates for current or future conditions due to climatic and land use changes; 2.) their station-based nature offers limited insight into the joint meteorological and hydrologic processes, highly variable in space and time, that cause floods and that dictate their probability of occurrence; 3.) these station-based methods are formulated such that it is difficult to integrate recent advances from adjacent branches of hydrology and meteorology such as precipitation remote sensing, numerical weather prediction, and (in the case of FFA) distributed hydrologic modeling.

A separate class of methods has also evolved for high-risk infrastructure such as large dams and nuclear power facilities: Probable Maximum Precipitation (PMP) and Probable Maximum Flood (PMF). PMP/PMF methods differ from typical RFA/FFA in two ways: 1.) they do not yield exceedance probabilities, but rather theoretical or practical upper bounds of rainfall and floods, and 2.) they are “storm-based,” rather than station-based. The largest conceivable rainstorm for a watershed of interest, the PMP, is developed based on theoretical arguments, regional-scale observations, and assumptions (Hansen, 1987). PMP estimation explicitly uses rainfall spatiotemporal structure, often in the form of rainfall fields (i.e. rainfall maps at regular time steps). PMF, by extension, considers this rainfall's interaction with watershed features by routing it through a rainfall-runoff model.

Rainstorm structure, including fine-scale variability and motion, is an important determinant of flood response (e.g. Arnoud et al., 2002; Meierdiercks et al., 2010; Mejia and Moglen, 2010; Morin et al., 2006; Norbiato et al., 2007; Ramos et al., 2005; Sivapalan et al., 1987; Smith et al., 2005, 2002; Yang et al., 2013). Explicit consideration of rainfall structure means that storm-based methods such as PMP, unlike station-based methods, can incorporate advances in both meteorological understanding and observations such as radar, satellites, and numerical weather prediction (National Research Council, 1994; see Abbs, 1999 and Ohara et al., 2011 for PMP examples). Important limitations of PMP/PMF are: 1.) the use of single values without an exceedance probability makes them unsuitable for hydrologic risk analyses (Ball et al., 2019; Swain et al., 2006; USBR and USACE, 2018); and 2.) developing the largest conceivable rainstorm necessarily involves the analyst's subjectivity (e.g. Dawdy and Lettenmaier, 1987).

Both the distinctions between statistical RFA/FFA and PMP/PMF and their limitations raise two questions: can we leverage spatiotemporal observations of extreme rainstorms probabilistically to perform storm-based RFA? And can we combine such an approach with rainfall-runoff modeling for FFA? In this paper, we examine stochastic storm transposition (SST), a technique developed to answer these questions. We review developments over 60 years that indicate SST can address the limitations of both RFA/FFA and PMP/PMF. Incidentally, SST is as old as the Journal of Hydrology—though the term wasn't coined until Fontaine and Potter (1989), Alexander (1963) introduced the concept in Volume 1, Issue 1 of this journal.

Historical and conceptual links of SST to other FFA methods and to PMP/PMF are discussed in Section 2. Section 3 describes the SST methodology. In Section 4, we review more than six decades of peer-reviewed SST research. Some important considerations, limitations, and useful properties of SST are discussed in Section 5. We conclude in Section 6 with some recommendations for future directions of SST research and applications.

2. Historical Background

Early researchers noted that rainfall records tended to be more numerous and often longer than those of flood flows (Miami Conservancy District, 1917). This implied that the estimation of extreme flood flows, probabilistic or otherwise, could be improved by considering extreme rainfall observations. Nonetheless, records of extreme rainfall over individual watersheds even today tend to be limited to at most a handful of notable events, making it difficult to characterize the upper tail of rainfall and flood hazard using these records alone. As a response, two ways have emerged to use rainfall and flood observations from a wider region to support FFA.

The first, "regionalized frequency analysis," involves leveraging nearby rainfall or streamflow observations to increase the robustness of statistical parameter or quantile estimates at a specific location or to produce estimates at ungaged locations. In the streamflow case, this is referred to as regional flood frequency analysis (RFFA); we use the term regional rainfall frequency analysis (RRFA) for rainfall applications. These contrast with "at-site" FFA or RFA, which only use local observations at the site of interest. RFFA and RRFA are still station-based—they use point observations and produce point predictions.

Fuller (1914) introduced what is believed to be the first formula for estimating flood quantiles:

$$Q_T = \bar{Q}(1 + 0.8 \log_{10} T) \tag{1}$$

where $\bar{Q}$ is the mean annual flood peak, T is the return period (the reciprocal of the AEP), and Q_T is the flood peak estimate corresponding to T. Fuller calculated the empirical coefficient 0.8 using flood observations from across the non-arid U.S. In other words, the first flood frequency formula was an RFFA formula. Research has continued ever since (see Requena et al., 2019 and Stephenson et al., 2016 for recent examples), and both RFFA and RRFA are commonplace in applications. We point readers to Dawdy et al. (2012) and Svensson and Jones (2010) for further information on RFFA and RRFA, respectively.

The second approach to using regional information is flood or storm transposition. Rather than use nearby observations to support statistical parameter or quantile estimation, transposition involves "moving" storm or flood observations to the watershed of interest and evaluating the result. Fuller (1914) solved Equation 1 for T using available annual maxima flood observations and found that the largest of these observations yielded estimates of T in excess of 1000 years. When preparing a dam spillway design in a new location, he therefore advocated using $T \geq 1000$ years, since evidence of such floods was available in observational records. Myers (1969) points out that Fuller was thus implicitly recommending transposing a property (T) of observed floods to new locations.

While the transposition of flood events poses theoretical and practical challenges, transposition of rainstorms is more straightforward (Myers, 1966). Myers (1969) summarized the advantages: "(a) rainfall is much less dependent on the underlying topography than is peak discharge; its transposition is therefore more physically realistic and accurate. (b) Precipitation records are in many instances longer and more comprehensive than discharge records… (c) the isohyets of a storm may be centered precisely over a basin, or in a number of different positions over a basin, a flexibility not available in discharge transposition." These factors contributed to deterministic storm transposition being proposed as an element of flood estimation (Meyer, 1917; Woodward, 1920) and becoming an integral part in the evolution of PMP/PMF methods (Bernard, 1936; Showalter and Solot, 1942). By the 1940s, PMP/PMF, rather than the probabilistic methods of Fuller and his successors, had become the preferred approach for spillway design in the United States (Myers, 1969) and these methods were also adopted by nascent nuclear power industry (England, 2011).

Meanwhile, probabilistic approaches continued to be widely used for applications in which less extreme return periods (e.g. 10^{-1} to 10^{-3}) were relevant. Hershfield (1961) and Miller (1964), for example, provided nationwide rainfall IDF maps for return periods up to 100 years, while the National Flood Insurance Program, started in 1968, focused on risk management within 100-year floodplains (Knowles and Kunreuther, 2014). Interest also renewed in probabilistic estimation of extremely rare storms and floods (National Research Council, 1994, 1988), motivated by unease at the level of subjectivity in PMP estimation, which could potentially result in costly designs or retrofitting (Alexander, 1963; Dawdy and Lettenmaier, 1987; YEAC, 1984). Some federal agencies in the U.S. (England, 2011; Swain et al., 1998) and elsewhere (Ball et al., 2019; Wilson et al., 2011) now use a combination of deterministic and probabilistic approaches for dam safety risk analysis and risk-informed decision making (USBR, 2013).

It was one such probabilistic foray into the realm of $AEP < 10^{-3}$ that conceived SST (Alexander, 1969, 1963; see Section 4.1). It should be noted that other probabilistic approaches, generally Monte Carlo in their nature, have also emerged (e.g. Beran, 1973; Charalambous et al., 2013; Muzik, 1993; Rahman et al., 2002; Schaefer and Barker, 2002; Stephens et al., 2016). Thorough review of these other approaches is beyond the scope of this study. Many of the considerations and challenges that we explore in this review, however, are also relevant to those techniques. Examples include storm spatiotemporal structures and "pairing" them with seasonally-varying probabilistic watershed initial conditions for flood simulation.

3. SST Methodology

3.1 The Basics

SST includes the following key elements: defining a transposition domain; developing an extreme storm catalog; randomly transposing storms in a region over a watershed; and estimating rainfall or flood probabilities. The concept, shown schematically in Fig. 1, can be briefly summarized: observed storms are transposed at random within a transposition domain of area A_D in such a way that new unobserved realizations of extreme rainfall over the domain are produced. In doing so, new realizations of extreme rainfall are created for a watershed of area A_W that resides within this domain. The space-time structure of rainstorms, including intensities, areas, and movement is preserved. SST can be understood as a bootstrap method, in which resampling from a catalog of observed storms is followed by random transposition of the newly-generated sample of storms. The details are described in Sections 3.2 through 3.5.

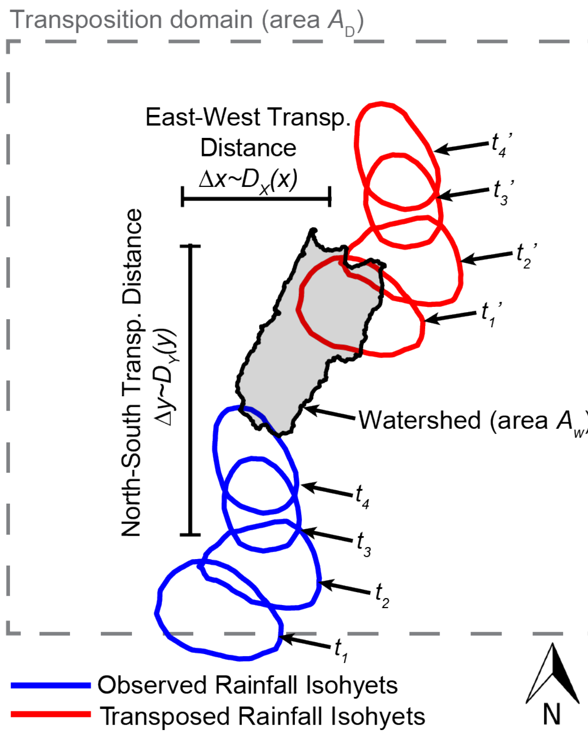

Fig. 1: Schematic of SST procedure for a single storm. Specific rainfall isohyets for the observed and transposed storms are shown for four time periods. Entire rainfall fields, as opposed to isohyets, can be transposed. All rainfall fields or isohyets are transposed by a north-south distance Δy and east-west distance Δx, which are randomly selected from distributions of D_y and D_x, respectively. In some SST efforts, D_y and D_x have been assumed to be uniform; in other cases, they have been estimated based on the locations of historical storms. Adapted from Wright et al., (2013).

3.2 Defining a Meteorologically Homogeneous Transposition Domain

Storm transposition is only defensible insofar as the storm could have occurred at that location with some nonzero probability. Chow (1964) defined a "homogeneous region," also referred to as a "transposition domain," as "the area surrounding the given river basin in which storm-producing factors are substantially comparable; i.e., the general area within which meteorological influences and topography are sufficiently alike." The transposition domain shown in Fig. 1 is square and centered around the watershed, but this need not be the case—see, for example, Fig. 2a, which shows an example transposition domain for Hurricane Harvey and New Orleans, Louisiana. Gupta (1972) argued that a transposition domain could "include a very large geographic area in the eastern half of the United States where (topographic) relief is generally moderate and it may include relatively small areas in the western United States where extreme topography is encountered." The issue of regional homogeneity is not unique to SST: RFFA and RRFA must also wrestle with it (e.g. Hosking and Wallis, 1993).

For SST to be of value, the transposition domain must be sufficiently large that it includes multiple observed extreme rainstorms. If the domain is very large relative to the size of the watershed of interest, however, the probability of transposing one of these rainstorms over the watershed is small. Alexander (1963) introduced a simple equation that shows this:

$$p_{tr} = A_w / A_D \tag{2}$$

where p_{tr} is the probability that the centroid of a storm from a transposition domain A_D will be transposed over a watershed of area A_w (see Fig. 1). Though Equation 2 is an incomplete description of the true process (since a storm could produce nonzero rainfall over a basin even if its centroid falls outside the basin boundary) it can be nonetheless instructive. Some implications of Equation 2 are discussed in Sections 3.4 and 5.2.

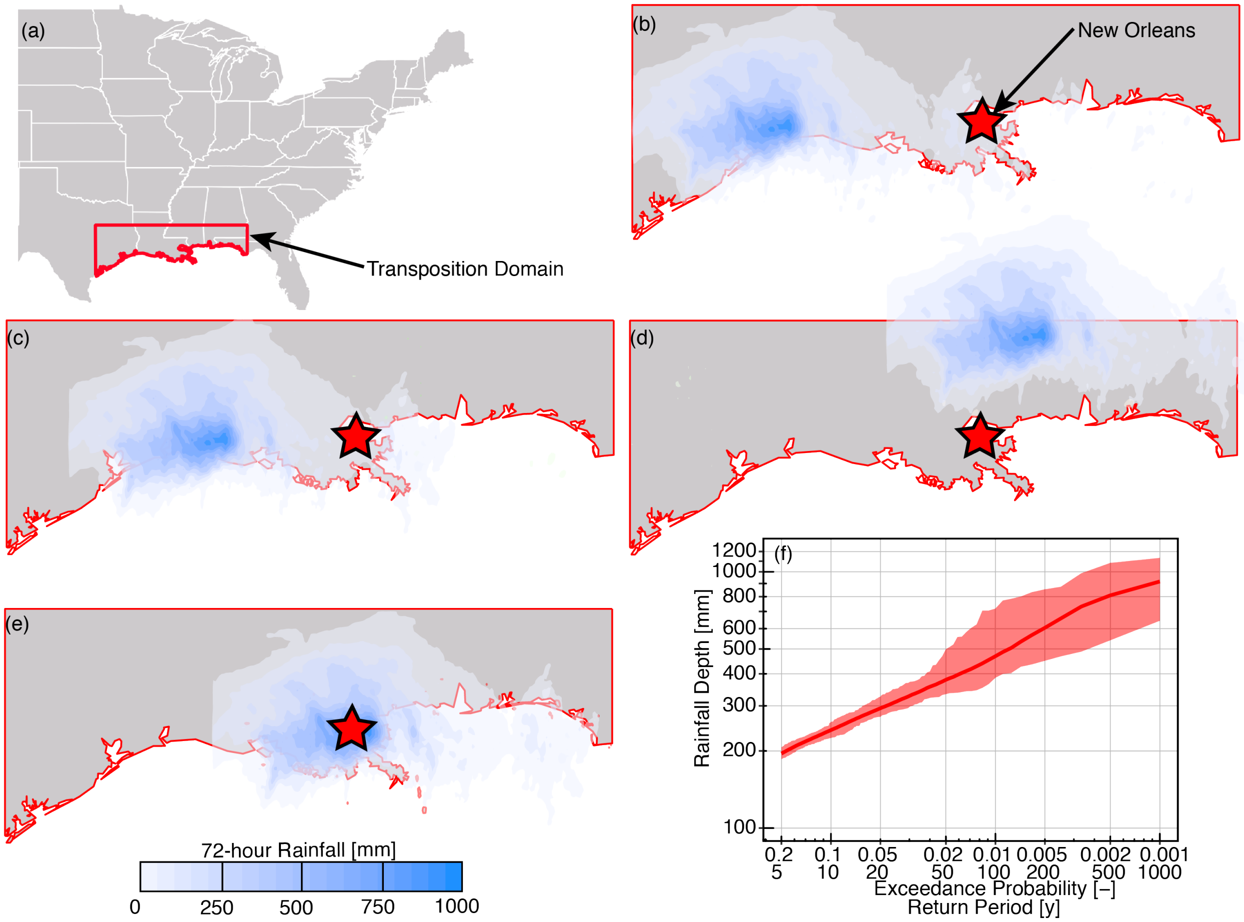

Fig. 2: (a) Transposition domain for a region surrounding New Orleans, Louisiana along the southern United States Gulf Coast. (b) Peak 72-hour rainfall map for Hurricane Harvey in August 2017, based on Stage IV gage-corrected radar rainfall data (Lin, 2011). (c) to (e) three possible random transpositions of Hurricane Harvey rainfall which produce little, no, and extreme rainfall over New Orleans, respectively. (f) Example 72-hour IDF curve for New Orleans generated using the RainyDay software (Section 4.4); shaded area portrays the spread of 100 distinct realizations, each consisting of 1000 annual rainfall maxima.

3.3 Creating a storm catalog from spatial rainfall observations

SST considers multiple storms that have occurred within the transposition domain for the watershed of interest. This set of storms is henceforth referred to as a storm catalog. Most SST studies have used rainfall observations from rain gage networks, often expressed as depth-area-duration (DAD) curves or tables that depict rainfall depth or rate as a function of averaging area and duration. By using DAD information along with assumptions of storm geometry, the labor of drawing or digitizing paper-based rainfall maps and then transposing them can be avoided. A useful source of both DAD tables and rainfall maps has been USACE (1973), which includes information on nearly 600 major U.S. storms starting in the 1880s. Though impressive in length and level of detail, this volume nonetheless has shortcomings—it has relatively fewer storms in the western U.S., and the evolution of rain gages over that time period means that the record does not provide a consistent picture of major storm activity even in the eastern part of the country. Such inconsistencies pose potential problems for SST (Foufoula-Georgiou, 1989; National Research Council, 1988), since the resampling described in Section 3.4 implicitly assumes that the storm catalog reflects the “true” extreme rainfall hydroclimate within the transposition domain.

Advances in rainfall remote sensing using ground-based radar and satellites offer alternative data sources for storm catalog creation. An example of the spatially detailed depiction of regional rainfall provided by radar remote sensing is shown in Fig. 2b for Hurricane Harvey, which struck the southeastern Texas coast in August 2017. Sections 4.4 and 5.1 discusses some potential strengths of these new data sources for SST.

3.4 Storm Resampling and Transposition

The main objective of SST, to estimate rainfall or flood *AEP*s, is achieved by resampling from the storm catalog to generate large numbers of realizations of extreme rainstorms over the watershed of interest. These realizations should synthesize new realistic annual patterns of rainstorms over the transposition domain and, by extension, over the watershed. To do this, a random number of storms k is generated by modeling the annual "arrival process" of storms over the domain.

The storm arrival process is usually assumed to follow either a Bernoulli or Poisson distribution with an arrival rate $\lambda = m/n$ storms per year, where m is the number of storms in the storm catalog and n is the length of the record (in years) from which the catalog was generated. Note the restrictions $0 < \lambda < 1$ and $k = \{0,1\}$ if a Bernoulli distribution is used, meaning that $m < n$. This makes the Bernoulli arrival process suitable only if small values of m are used. As a consequence, p_{tr} (Equation 2) will be relatively low and many realizations of transposed rainfall will thus be small or zero. This can be seen in Fig. 2c-d, in which little and no rainfall is produced for New Orleans, Louisiana for two possible transpositions of Hurricane Harvey. Fig. 2e, meanwhile, shows a transposition that produces extreme rainfall over New Orleans. This feature makes the Bernoulli arrival model suitable only for estimation of very low *AEP*s such as those needed for spillway design; the magnitude of more common events will be greatly underestimated. This restriction is lifted if a Poisson distribution is used, though, as discussed in Section 5.2, underestimation of the magnitude of more common events can nonetheless result if m is not sufficiently large. No SST study to date has sought to model the temporal sequencing of these k storms, meaning that, to use the parlance of rainfall-runoff modeling, SST has thus far only been event-based (e.g. Chu and Steinman, 2009).

The k storms are randomly sampled from the storm catalog and randomly transposed within the transposition domain. A single transposition is shown schematically in Fig. 1; three possible transpositions of 72-hour rainfall from Hurricane Harvey are shown in Fig. 2c-2e. All rainfall fields or isohyets that constitute a storm are transposed by a north-south distance Δy and an east-west distance Δx which are randomly selected from distributions D_y and D_x, respectively. These distributions jointly describe the spatial probability of storm occurrence.

The notion of a homogeneous transposition domain implies that the probabilities of random placement of transposed storms should be equal throughout the transposition domain, i.e. that D_y and D_x are uniform. Relaxation of this stricture allows larger values of A_D and thus m, but could introduce bias since rainfall properties would not be strictly homogeneous. This issue was discussed in very general terms by Alexander (1963), who wrote that "the basic problem is to preserve only the essential statistical features of the area by discarding those which may be ascribed to sampling errors." Several more recent schemes have provided ways to either modify the storm transposition probability or storm magnitude to limit these biases (Nathan et al., 2016; Wilson and Foufoula-Georgiou, 1990; Wright et al., 2017; Zhou et al., 2019).

Once transposed, the k resulting rainfall amounts over the watershed are computed. The largest of these can be understood as a "synthetic annual rainfall maxima," which form the basis of *AEP* estimation.

3.5 Estimating Rainfall and Flood Quantiles

Unlike statistical FFA and RFA, it is not necessary to fit distributions to the synthetic annual maxima in order to obtain *AEP*s. Rather, *AEP*s can be estimated directly from the ranked synthetic annual maxima using plotting position formulae. For example, 1,000 annual maxima generated through SST would facilitate direct estimation of *AEP*s as low as 10^{-3}. Uncertainty bounds can be obtained by generating multiple such "sets" of realizations. Fig. 2f, for example, shows an SST-based IDF curve for New Orleans, in which 100 sets of 1,000 annual maxima each were generated. The shaded area denotes the spread among these 100 sets.

Obtaining flood quantiles requires the use of a rainfall-runoff model, but flood *AEP*s are otherwise computed in the same manner. There need not be a 1:1 correspondence between rainfall and flood *AEP* if rainfall spatiotemporal structure is considered or if watershed initial conditions such as soil moisture are treated as random variables (Franchini et al., 1996; Gupta, 1972; Wright et al., 2017; Yu et al., 2019).

A brief graphical summary of the step-by-step procedure described in Section 3 is shown in Fig. 3.

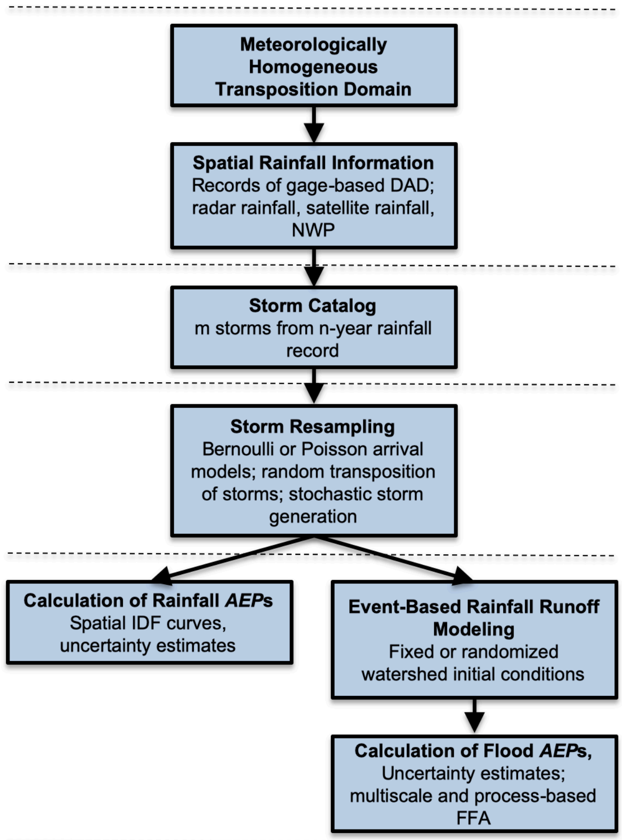

Fig. 3: Schematic of the SST methodology.

4. Review of SST Developments and Applications

4.1 Early Theoretical Development (1963-1972)

SST originated in Australia, where preference for PMP/PMF was not as strongly-rooted as in the United States (Alexander, 1969). Alexander's (1963) initial study focused mainly on the probability of transposing a storm event over a watershed given their respective spatial scales and the size of the transposition domain; e.g. Equation 2. Though some empirical properties of DAD were mentioned, no results were provided. The paper outlined the basic concept of a homogeneous region, identified relevant spatial and temporal scales, and described the Bernoulli storm arrival process. Shape and orientation of storms were neglected. Alexander (1969) followed by discussing additional aspects including the meteorological rationale used to select a transposition domain in the Australian state of Victoria. This second study also showed results of rainfall depth vs. *AEP* as low as 10^{-4} for an unspecified watershed in Victoria.

Gupta (1972) provided a more detailed treatment, including discussion (though not results) of its extension to the frequency analysis of flood volumes. He introduced the Poisson arrival model and suggested that "second-order" storm properties (orientation, shape, within-storm temporal distribution, etc.) were also important considerations. His procedure included stochastic rotation of transposed rainfall fields, which were then run through a simple rainfall-runoff model. The storm catalog consisted of five storms in the Midwestern U.S., and the rainfall-runoff model considered seasonally-appropriate initial conditions. Though these initial conditions were not treated as random variables, it was recommended that future studies do so, echoing other contemporary work on non-SST probabilistic FFA (e.g. Beran, 1973). He proposed that stochastic simulation of storms could be used to augment the limited size of a storm catalog. Notably, his thesis also presented "a practical, computer-oriented methodology of transposing storms of a region and a historical sample to any river basin inside that region," though the lack of probabilistic results implies that this methodology stopped somewhat short of a fully-realized SST software—something that would not emerge until much later.

4.2 Theoretical and Practical Advances (1984-1996)

Interest in SST renewed in the 1980s. This was spurred by a technical study from the Yankee Atomic Electric Company (YEAC, 1984). That study used a variant of SST to assess the likelihood that Harriman Dam in Vermont in the Northeastern U.S. would be overtopped, thus threatening a downstream nuclear facility. It was motivated by disagreement over PMP estimates for the 520 km^2 watershed, which ranged from 14 inches to over 22 inches in 24 hours. This study showed that the DAD-based SST procedure produced more credible extreme rainfall frequency estimates than conventional RFA, and demonstrated how reservoir initial conditions could be randomized based on historical records.

The YEAC (1984) report featured prominantly in a comprehensive report on rare flood quantile estimation published by the National Research Council (1988), which went so far as to propose its own SST approach. Fontaine and Potter (1989) compared the YEAC (1984) and National Research Council (1988) approaches for *AEP*s approaching 10^{-6} using a catalog of four major storms in the Midwestern U.S. The authors argued that “there is a need to develop a more formal theoretical framework,” but also that requirements of strict homogeneity should be relaxed to maximize the potential size of the transposition domain. They also argued that uncertainty estimation including errors in rainfall observations should be considered, and echoed Gupta's (1972) recommendation that stochastic simulation of storm events could be useful for enlarging the storm catalog.

This latter recommendation was picked up in Foufoula-Georgiou (1989), who introduced a stochastic model to simulate elliptical storms using DAD and geometric information from USACE (1973). This model could produce arbitrary numbers of storms, including ones with higher rainfall magnitudes, larger spatial extents, etc. The random variables used in the storm model were maximum rainfall depth at the storm center, orientation of the ellipsis’ major axis, the ratio of major to minor axes, and two parameters describing the decay of rainfall depth with distance from the storm center. This formulation thus precluded non-elliptical or multi-cell storms, and did not incorporate temporal rainfall structure or motion. The author evaluated several aspects related to the probability of catchment rainfall by transposing elliptical representations of 18 observed storms from the midwestern U.S. over idealized watershed shapes and geometries.

This stochastic rainfall model was fully implemented in Wilson and Foufoula-Georgiou (1990) and was parameterized using a larger set of 65 storms. They also linked this model to a nonhomogeneous point process that jointly modeled storm occurrence location (Fig. 4) and peak rainfall depth, easing the need for a homogeneous transposition domain. They estimated rainfall *AEP*s lower than 10^{-12} over hypothetical circular 100 mi^2 catchments, showing that the combination of the stochastic storm and point process models could qualitatively reproduce the spatial patterns exhibited by existing PMP studies in the region. They also assessed sensitivity of the method to storm eccentricity and to incomplete storm catalogs.

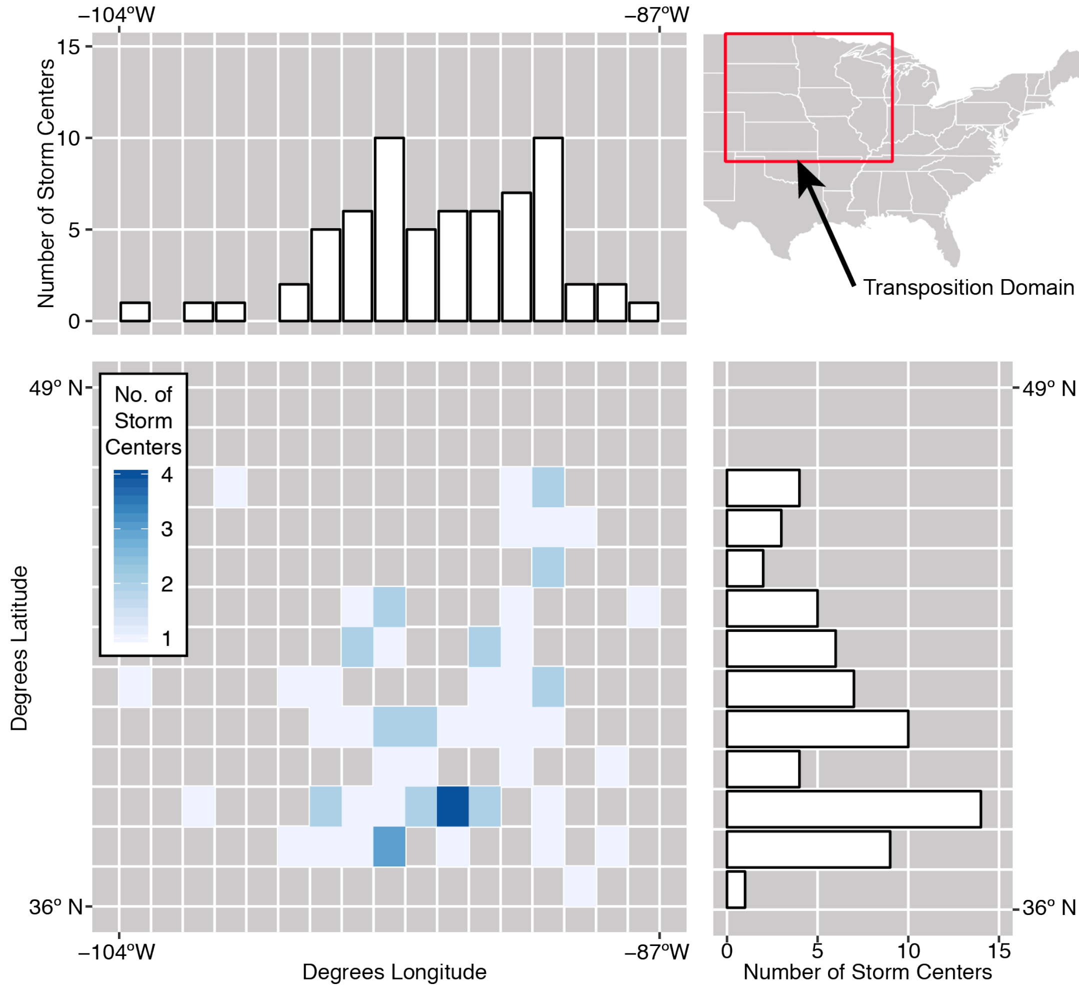

Fig. 4: Transposition domain and locations of the centers of 65 storms from USACE (1973) used to parameterize the point-process arrival and stochastic storm models in Franchini et al. (1996) and Wilson and Foufoula-Georgiou (1990), from which this figure is adapted. Note the heterogeneity in observed storm occurrence across the domain, which is accounted for in the arrival process.

Franchini et al. (1996) applied the method of Wilson and Foufoula-Georgiou (1990) to estimate the frequency of peak flows using a lumped rainfall-runoff model. To incorporate rainfall temporal structure, they probabilistically disaggregated 24-hour storm totals to finer time scales using dimensionless temporal distributions. They examined the sensitivity of FFA to initial soil moisture, though a lack of observations precluded its representation as a random variable. Even for very extreme events, this sensitivity was high: the *AEP* associated with the 400 m^3/s (1800 m^3/s) flood peak ranged from 10^{-1} (10^{-3}) for saturated conditions to roughly 10^{-3} ($<10^{-5}$) for dry conditions. They also showed importance of rainfall temporal distribution in determining flood peaks, thus highlighting that flood magnitude and thus flood frequency strongly depend on characteristics beyond simply rainfall depth.

4.3. Recent SST Efforts for Rare Storms and Floods

Nathan et al. (1999) and Agho et al. (2000) investigated the potential of SST in southeastern Australia. The authors documented challenges addressing homogeneity while transposing storms in the region. This line of work continued to evolve in Nathan et al. (2016) , who used SST and another stochastic method to estimate the *AEP* of existing PMP values for two watersheds in southeastern Australia. Their *AEP* estimates ranged from 10^{-5} and 10^{-6}, depending on the method and watershed. They used a 114-year gridded rain gage dataset to generate elliptical storm events. A key contribution was the elaboration of "dimensionless SST," in which storm rainfall R_i transposed from location i to location j is rendered dimensionless by dividing by an "index quantile" (I_i^q) and then "rescaled" by I_j^q to obtain a transposed rainfall value $\tilde{R}_j$:

$$\tilde{R}_j = \frac{I_j^q}{I_i^q} R_i. \tag{3}$$

Dimensionless SST provides a straightforward means for relaxing the strictures of transposition domain selection, since transposed rainfall is rescaled according to a measure of the local extreme rainfall climatology. This allowed them to use a very large transposition domain that covered much of southeastern Australia. Dimensionless SST may render nonuniform transposition probabilities unnecessary, at least in practice, since it implicitly accounts for spatial variability in storm frequency as well as intensity.

The ratio I_j^q / I_i^q will be sensitive to errors in quantile estimates. The authors used at-site estimates of the 50-year rainfall quantile to compute this ratio; such a high quantile may be prone to substantial sampling error. This can also be said for earlier efforts (Agho et al., 2000; Nathan et al., 1999), which used areal PMP estimates to rescale R_i. Nathan et al. (2016) state that dimensionless SST deserves further research. This could include evaluation of less uncertain index quantiles such as the 2-year or 5-year rainfall (Wright and Holman, 2019). It should be noted that rescaling by an index quantile is common practice in RFFA and RRFA, specifically in the 'index flood" approach (e.g. Stedinger et al., 1993).

England et al. (2014) examined FFA in the 12,000 km^2 Arkansas River watershed upstream of Pueblo, Colorado in the western U.S. by combining SST-based rainfall events with a distributed physics-based rainfall-runoff model. SST-based results are presented alongside flood-based statistical FFA, estimates of several historical floods, and paleoflood data (Fig. 5). This "integration of collaborative work in hydrometeorology, flood hydrology, and paleoflood hydrology" is noteworthy for several reasons: 1.) it was the first to generate SST-based FFA estimates for a relatively large watershed in mountainous terrain; 2.) it was among the first to combine SST with a physics-based distributed hydrologic model (Wright et al., 2014 appeared in the same year), 3.) it highlighted that transposition domain selection and watershed interaction in complex terrain can have a major influence on results (this contrasts with Wright et al., 2013b; 2017, who reported only modest sensitivities in less topographically-complex regions); and 4.) its usage of multiple sources of probabilistic flood estimates provided an interesting demonstration uncertainty estimation through (to use a legal term) “preponderance of evidence,” which can “increase the credibility and resulting confidence in the results” (Swain et al., 1998).

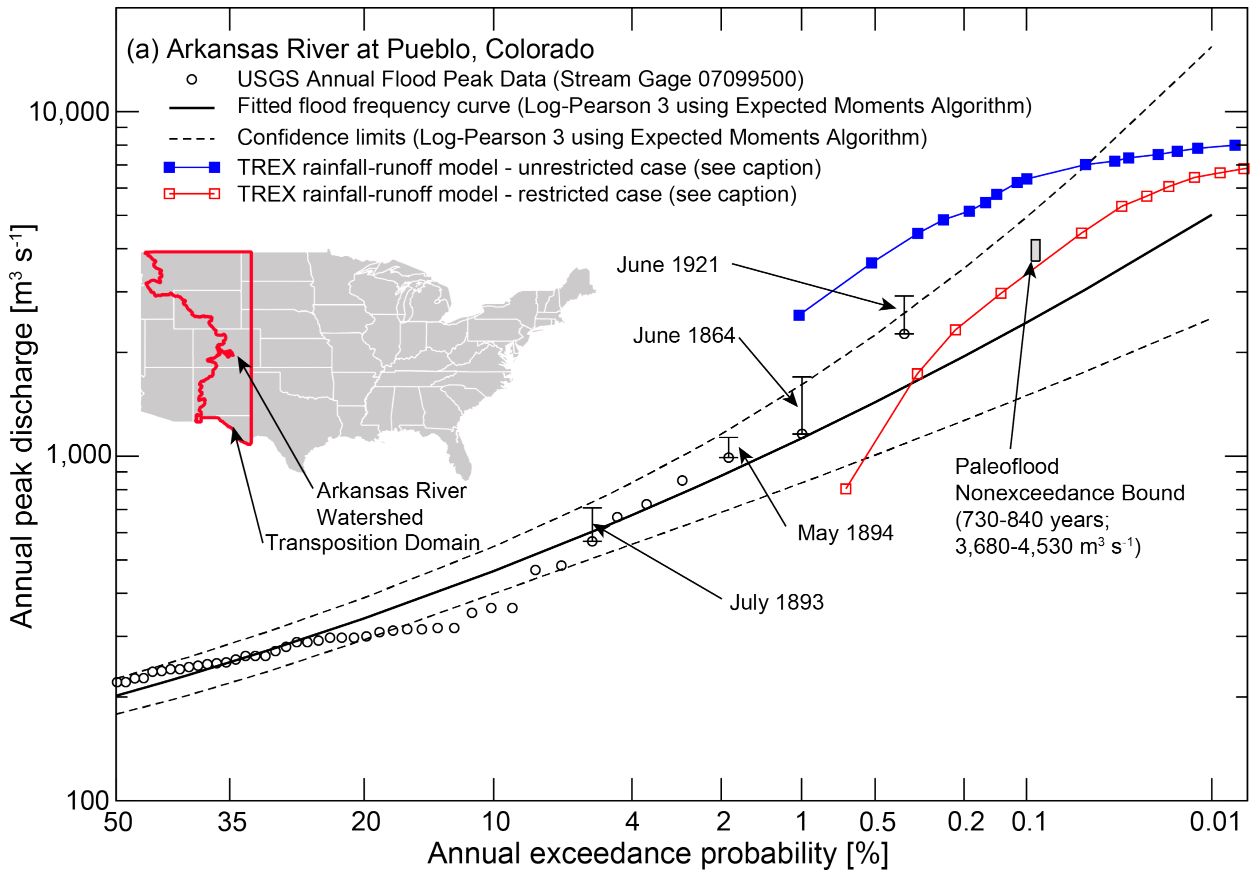

Fig. 5: Flood frequency curves for the Arkansas River at Pueblo, Colorado based on statistical flood-based FFA (black lines), SST (red and blue lines and squares), and observed flood peaks (open black circles). Error bars indicate the estimated magnitudes of four historical floods. Estimates of the paleoflood nonexceedance bound is shown in grey. “Restricted” SST-based results refer to those generated when storm transposition was limited to areas where orographically-enhanced heavy rainfall is prevalent in relation to th watershed; “unrestricted” refers to results generated when storm transposition could occur anywhere in the domain. Figure is adapted from England et al. (2014), which contains further details.

4.4 *Rainfall Remote Sensing, Multi-scale FFA, and RainyDay*

The SST studies reviewed thus far sought to estimate rare rainfall and flood probabilities, mainly for dam safety. Dams can be understood as "point-scale" features, since the concern is generally a single distribution of rainfall or flood discharge upstream of the facility. Levee systems, transportation networks, and stormwater systems, on the other hand, are "multi-scale"—flood distributions may be needed at many locations along the drainage network. Many applications in such contexts are focused on more common events, e.g. $AEP > 10^{-3}$.

Wright et al. (2013b) introduced an SST methodology for estimating rainfall IDF curves using a high-resolution 10-year radar rainfall dataset. Results were shown for watersheds from 2.5 to 240 km^2 and for durations from 1 to 12 hours in Charlotte, North Carolina, with IDF curves "tailored" to the specific watershed size, shape, and orientation. These IDF curves reflected the effects of storm structure and motion, and the authors showed that the temporal and spatial rainfall averages represented by these IDFs masked considerable variability in the spatiotemporal properties of the transposed rainstorms used to derive them. They argued that rainfall remote sensing is the key to assessing this variability, since few rain gage networks are dense enough to sample fine-scale rainfall structure. SST's storm-based nature also allowed the authors to examine how storm hydroclimate influenced IDF estimates— finding that for short durations and small areas, tropical cyclones are insignificant contributors to extreme rainfall distributions in Charlotte, while their importance grows with duration and watershed size.

Wright et al. (2014) extended this analysis to multiscale FFA for an urbanized 110 km^2 watershed using a detailed physically-based distributed hydrologic model. When combined with high-resolution rainfall scenarios, the interactions of second-order rainfall properties with land cover, river channels, and the urban storm drain network could be translated into FFA. This circumvents the design storm assumption of 1:1 equivalency of rainfall and runoff return period, eliminates the need to carefully link design storm duration to watershed characteristics such as the time of concentration, and captures the intra-event *AEP* variability that can occur across the river network. They also argued that the combination of SST with physically-based models opens opportunities for modeling future flood frequencies in nonstationary land use or climate conditions.

Wright et al. (2017) detailed RainyDay, an open-source Python-based SST software (see Wright, 2019 for source code). RainyDay is based on the methodology of Wright et al. (2013b; 2014). Subsequent software updates include the dimensionless SST of Nathan et al. (2016) and an alternative “rescaling” approach outlined in Wright and Holman (2019). Wright et al. (2017) demonstrated RFA and FFA results using ground-based radar rainfall and several satellite-based datasets. The distributed Hillslope-Link Model (HLM; Krajewski et al., 2017; Mantilla and Gupta, 2005) was used in conjunction with RainyDay in Turkey River, a 4000 km^2 watershed in the Midwestern U.S. that has exhibited flood nonstationarity in recent decades. RainyDay FFA estimates were more consistent with recent peak discharge observations than RFFA results published by the US Geological Survey (Fig. 6). Wright et al. (2017) was also the first SST study to treat antecedent watershed conditions, specifically soil moisture and channel flow, as random variables. (It should be noted that other non-SST probabilistic FFA studies had previously considered this issue—see, for example, Charalambous et al., 2013; Muzik, 1993; Rahman et al., 2002; Schaefer and Barker, 2002).

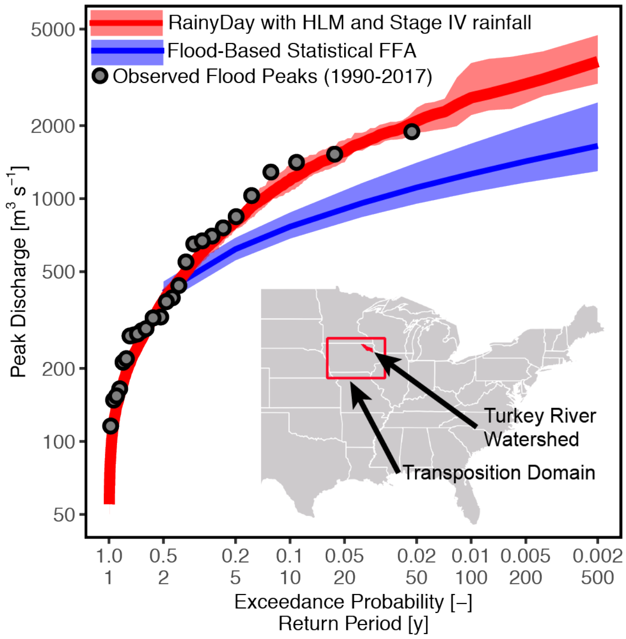

Fig. 6: Flood frequency curves for Turkey River at Garber, Iowa generated by SST using RainyDay combined with the HLM distributed rainfall-runoff model and Stage IV radar rainfall and by the U.S. Geological Survey (USGS; see Eash et al., 2013) using Bulletin 17B RFFA methodology (IACWD, 1982). The watershed and transposition domain are shown in the inset map. Red shaded area shows the spread of 10 distinct realizations of 500 years each using SST. Blue shaded area shows the 90% confidence interval on the Bulletin 17B estimates. Adapted from Wright et al. (2017).

Zhou et al. (2019) present an investigation of rainfall heterogeneity and its consequences for RFA in the region surrounding Baltimore, Maryland in the U.S. Mid-Atlantic. The rainfall hydroclimate there is influenced by the Chesapeake Bay to the southeast, the topographic gradient leading to the Appalachian mountain ranges to the northwest, and localized urban rainfall modification (Smith et al., 2012). They used the RainyDay software to examine the impact of this heterogeneity on the extreme storm hydroclimatology and hydrometeorology and on IDF estimates. To account for regional heterogeneities in extreme rainfall, rainfall R_i transposed from location i to location j was "rescaled" according to the ratio:

$$\tilde{R}_j = \frac{\bar{R}_j}{\bar{R}_i} R_i \quad (4)$$

where $\bar{R}_i$ is the mean rainfall of all storms in the storm catalog from location i. A stochastic generalization of Zhou et al.'s (2019) approach was introduced in Wright and Holman (2019), who compared it against the index quantile ratio method (Equation 3) from Nathan et al. (2016).

Zhu et al. (2018) examined the importance of rainfall spatiotemporal structure in flood frequency using the combination of RainyDay and HLM from Wright et al. (2017). They also examined how this importance is modulated by soil moisture. A set of 10,000 high-resolution rainfall scenarios were used to simulate flood frequency at 5,000 subwatersheds of the 4,000 km^2 Turkey River watershed, as were three additional sets generated by "downsampling" to coarser spatial and temporal resolutions. Complex relationships were found between rainfall structure, watershed scale, and initial soil moisture. Their results suggest that rainfall structure is an important control on real-world flood frequency, and that FFA efforts that simplify this structure may underestimate flood risk, especially for smaller-scale watersheds. This confirms that high-resolution rainfall remote sensing datasets are valuable when combined with SST and distributed hydrologic models.

Yu et al. (2019) introduced a more thorough approach to SST-based FFA for the same watershed evaluated in Wright et al. (2017) and Zhu et al. (2018). A "library" of seasonally-varying soil moisture and seasonal snowpack conditions was created via a long-term continuous hydrologic simulation. Initial conditions were then sampled from this library, paired with RainyDay-based transposed storms from the same season, and run through a simple lumped conceptual rainfall-runoff model (though their framework is generalizable to more complex models). Seasonally realistic joint distributions of rainfall, soil moisture, and snowpack were thus preserved and translated into FFA results. They refer to this approach as "process-based FFA," which is discussed further in Section 5.6. They showed that seasonal shifts in soil moisture, snowpack, and extreme rainfall explain the recent upward trend in large-scale floods and downward trend in more common floods in Turkey River.

Using the same combination of RainyDay SST, radar rainfall, and HLM as Wright et al. (2017), Perez et al. (2019) conducted a "synthetic analysis" of commonly-used statistical at-site and regional FFA methods. 10,000 simulated annual maximum flood peaks at 5,000 locations within Turkey River formed a synthetic population of flood peaks. This combination of tools allowed for explicit simulation of the interactions between rainfall spatiotemporal variability with watershed and river channel morphology. Subsets of this population were then created, representing samples of various sizes (e.g. 10 years, 30 years, etc.) and, in the case of RFFA methods, various numbers of sites. The robustness of several commonly-used FFA/RFFA methods were then compared as a function of sample size and number of sites. This contrasts with previous efforts to assess statistical RFFA methods, which have relied on simplistic assumptions regarding the regional variability of flood peaks. Perez et al. (2019) found that the difficulty of estimating distributional shape or skew, a major challenge in RFFA, can be partially explained by the river network structure and the orientation of the watershed relative to prevailing storm directions. The approach shows how SST can facilitate the use of recent advances in other branches of hydrology and other fields (e.g. distributed models, rainfall remote sensing) to better understand and potentially improve or replace existing FFA methods.

5. Discussion

5.1 Rainfall Data and SST

Sparse rain gage networks can fail to adequately sample fine-scale rainfall variability such as localized precipitation events and the locations of peak storm intensity (e.g. Curtis, 2007; Foufoula-Georgiou, 1989). This lack of detail may limit rain gages' direct applicability to multi-scale FFA. The recent SST studies reviewed in Section 4.4 show the merit of coupling SST with weather radar data, due to the latter's depiction of rainfall variability at high spatial and temporal resolution. On the other hand, a lack of long-term records radar records and the need for careful bias correction places limits on radar's broad usefulness. To estimate rainfall and flood *AEP*s 10^{-1} to 10^{-3}, it appears that at least one decade of radar rainfall data is needed; estimation of rarer *AEP*s may require longer records. Other remote sensing-based RFA attempts exist using station-based methods (see Faridzad et al., 2018, McGraw et al., 2019, and references therein); those methods appear to be more sensitive to the relative shortness of remote sensing records than SST.

The potential of satellites and numerical weather prediction (NWP) models for SST have received little and no attention, respectively, but these data sources are currently showing improvements in terms of resolution, accuracy, and record length. The latest satellite datasets, for example, offer relatively long records (2+ decades) at higher resolution and improved accuracy relative to previous generations. Convection-permitting regional climate models (RCMs) are now able to be run for decadal periods at higher spatial and temporal resolutions (e.g. Prein et al., 2015). Recent multi-decadal atmospheric reanalysis datasets have resolutions sufficient for some RFA and FFA applications (e.g. Toride et al., 2018). RCMs and reanalyses have the potential to perform better than radar or satellites in complex terrain (e.g. Wright, 2018). Both satellite datasets and NWP provide the potential for SST in regions such as developing countries where both rain gage networks and weather radar are lacking.

Though most SST studies have listed data quality as critical, its influence on SST has only partially been evaluated (Foufoula-Georgiou, 1989; Wilson and Foufoula-Georgiou, 1990; Wright et al., 2013b). While rainfall measurement using any instrument is nontrivial (National Research Council, 1994), radar, satellite, and NWP-based estimates are typified by large errors. Presently, all the sources of rainfall measurements mentioned above are heavily reliant on rain gages to identify and eliminate biases. Thus, the rain gage networks that were critical for early SST work are as essential now. The current decline in rain gage networks worldwide (e.g. Stokstad, 1999) thus poses a threat to the future of SST and RFA/FFA more generally.

Wright et al. (2017) pointed out that SST and other regional methods are able to “improve” more quickly than at-site frequency analysis as new extreme events are observed, since these observations need not have occurred over the watershed of interest. This logic extends backwards as well; notwithstanding challenges data homogeneity, the prospect of combining recent rainfall records with older isohyetal maps and DAD observations holds the allure of providing very long rainfall records for SST. England et al. (2014) shows an example of this. This prospect is perhaps not far off, at least in the U.S., since USACE (1973) and other records of old storm information have been or are currently being digitized.

5.2 AEP Uncertainty and "Boundedness" in SST

Suppose the SST resampling and transposition described in Section 3.4 is repeated to create 100,000 annual maxima. When converting these to *AEP* estimates (Section 3.5), the analyst can rank these maxima to obtain probabilities as low as 10^{-5}, or subdivide these into 100 samples and thus obtain 100 estimates for each *AEP* down to 10^{-3}. Studies focusing on very rare floods (Sections 4.1-4.3) have done the former, while the recent studies in Section 4.4 do the latter. Clearly, the former does not give the "true" population of all possible rainstorms and floods, while the latter does not represent the full "real-world" uncertainty about these quantities. The side-by-side comparison of multiple FFA methods by England et al. (2014) showed one practical pathway to better grapple with such uncertainty.

SST using observed rainstorms will have an upper bound associated with the largest rainstorm in the storm catalog, transposed to maximize the rainfall over the watershed of interest (Wright et al., 2013b; 2017). This upper bound is the result of sampling error, i.e. incomplete knowledge of the true population of extreme rainstorms.

Upper bounds in SST should not be confused with upper boundedness in real rainfall and flood processes, which has been a topic of lively debate in the hydrologic community for nearly a century (see Smith and Baeck, 2015 and references therein). An interesting question, however, is whether SST can bring new insights to that debate. Linkages between generation and sampling from storm catalogs and the asymptotic behavior of extrema that is the focus of extreme value statistics (e.g. Davison and Huser, 2015) may be worth exploring.

Quite separate from the question of upper bounds is the existence of a lower bound or bias, which is an artifact of the SST procedure: there is a nonzero probability that a transposed storm will cause no rainfall at all over the watershed of interest (e.g. Fig. 2c). This is especially true if, via Equation 2, A_w/A_D is small. Under a Bernoulli arrival model, this would translate to a synthetic rainfall maxima of zero; under a Poisson model, zero rainfall remains a potential outcome. Nonzero but still small synthetic annual maxima are relatively likely in both models and are clearly unrealistic. This issue is unimportant when the objective is the estimation of very rare events. For the studies reviewed in Section 3.4, however, this is a major problem, since more common AEPs are an objective. This can be remedied by using a large storm catalog, which increases the Poisson arrival rate. Wright et al. (2017) recommended a catalog with at least $m = 10n$ storms, while the current version of the RainyDay software defaults to $m = 20n$. Bias associated with this effect can be seen in Fig. 7 in SST-based results for $AEP \geq 0.5$, which used $m = 14.3n$.

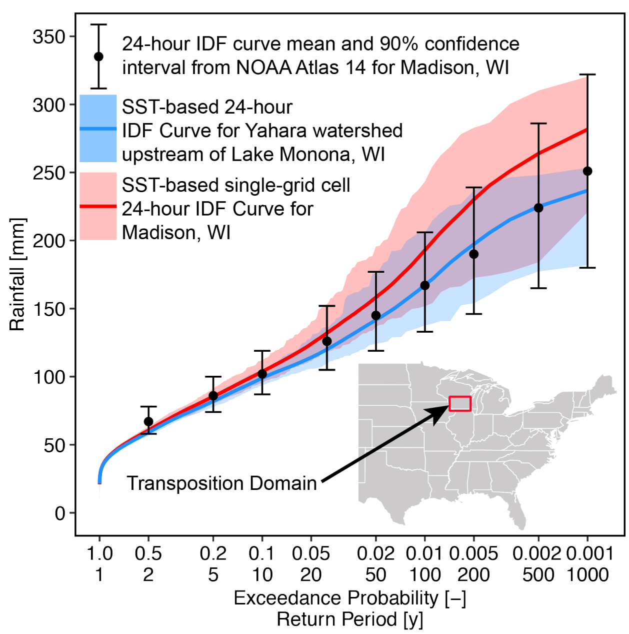

Fig. 7: 24-hour IDF curves for the Madison, Wisconsin, USA area based on NOAA Atlas 14 (Bonnin et al., 2006) and the RainyDay SST software. RainyDay transposition domain is shown in inset map. The

RainyDay-based IDF curve (blue) is for the size, shape, and orientation of the 728 km^2 Yahara River watershed upstream of Lake Monona, Wisconsin. Deviations between the single-grid cell RainyDay-based IDF curve and Atlas 14 (black points and confidence intervals) are related to extreme storms that occurred near Madison in August 2018 and are not reflected in the older Atlas 14 estimates. Shaded areas for RainyDay-based estimates portray the spread of 100 distinct realizations, each consisting of 1000 annual rainfall maxima.

5.3 Spatial Considerations of Rainfall and Flood Extremes

Beyond very small catchment areas, the spatial dimensions of storm events become increasingly important in flooding (e.g. Marston, 1924). Taken alone, rain gage-based IDF curves are therefore unsuitable for FFA (aside from in small catchments) since the statistics of point rainfall can differ dramatically from those of watershed-scale rainfall. To compensate for this, analysts can employ area reduction factors (ARFs) to estimate areal rainfalls from point-scale IDF information. Wright et al. (2013a) showed that typical ARF methods are conceptually flawed in ways that are not easily remedied. Estimation of spatial IDF curves is often challenging because of the required density of long-term contemporaneous rain gage observations. SST's storm-based nature provides an alternative method for estimating rainfall IDFs for spatial scales beyond a single rain gage (Fig. 7) and obviate the usage of ARFs. These scales can be relatively small, down to a single rainfall remote sensing grid cell, e.g. <1 km^2 for ground-based weather radar to 100+ km^2 for satellite or reanalysis datasets. Despite SST's ability to produce IDF estimates, such estimates need not be an intermediate step in SST-based FFA. Rather, transposed rainfall fields should be used directly as inputs to rainfall-runoff modeling, preserving observed space-time structure and enabling multi-scale FFA.

While SST holds distinct advantages for RFA and FFA at spatial scales beyond small watersheds, there is likely an upper limit to the areas over which it can be applied. Gupta (1972) pointed out that “large basins can have pronounced differences in meteorological and topographic influences” which may make the transposition of storms unrealistic. He suggested that this might constrain SST to watersheds smaller than “1,000 square miles” (2,600 km^2).

Recent SST studies, particularly England et al. (2014), underline that this limit might be larger than Gupta speculated. Nonetheless, SST is almost certainly not applicable for continental-scale river basins such as the Mississippi or Amazon without modification. Floods in continental river basins are generally not the result of individual storms but rather of spatially and temporally clustered events occurring over periods of weeks to months and covering spatial scales as high as 10^6 km^2. These storms may occur in climatologically distinct regions, rendering the notion of a single transposition domain unworkable. One could in principal create a storm catalog comprised of the most rainy multi-week periods and conduct SST using these. The number of available “events” would be small—if month-long storm periods were used to create a storm catalog, for example, the number of “events” would be at most 12 times the length of the rainfall record in years. Few of these storm catalog entries would be particularly important from a flood-generation standpoint. Second, transposition of a month-long rainy period may not yield sufficient realizations of rainfall spatiotemporal variability, since the individual storm elements within this period are fixed in their relative locations and timing. The “reshuffling” of individual storm elements in space and time to create new realizations may be possible but poses conceptual and practical challenges.

5.4 Other Event-Based and Continuous Rainfall Generators

An alternative to SST for generating large numbers of hypothetical rainfall events is to develop a point- or basin-scale IDF curve, randomly sample quantiles from this curve, and combine these quantiles with spatial or temporal rainfall patterns from observed storms. Examples of this include Charalambous et al. (2013) and Schaefer and Barker (2002). This requires a collection of observed storm events, as well as the ability to estimate either areal-averaged IDFs or ARFs (not necessarily easy to do; see Section 5.3). The approach implicitly assumes that the annual rainfall maxima provided by IDF curves lead to annual flood maxima. This may be adequate for estimating rare flood quantiles; it is problematic for more common events (Yu et al., 2019). As far as we are aware, existing examples of the approach have assumed that rainfall quantiles and spatiotemporal structures can be considered to be independent.

Stochastic rainfall generators (SRGs) are another way to generate hypothetical realizations of rainfall. They attempt to simulate rainfall variability, so that one could generate potentially many thousands of years of rainfall (Sharma and Mehrotra, 2010). While SST involves transposing observed storms, SRGs rely on rainfall observations for "training" and validation. The stochastic storm model of Foufoula-Georgiou (1989) and Wilson and Foufoula-Georgiou (1990) is one example, but most generate continuous sequences, rather than individual storms. Continuous SRGs have been used for FFA (e.g. Blazkova and Beven, 2002; Cameron et al., 1999; Peleg et al., 2017). While most SRGs have been either point-based or provide areal-averaged rainfall, more detailed models provide high detail in both space and time (e.g. Paschalis et al., 2013; Peleg and Morin, 2014). Spatially-explicit SRGs are preferable for FFA, since fine-scale variability is an important determinant of flood response (Paschalis et al., 2014; Zhu et al., 2018). Like SST, these SRGs also facilitate multi-scale FFA (e.g. Peleg et al., 2017) and process-based FFA.

The key test for both SST and SRGs is whether they can faithfully represent the extreme tail of rainfall magnitude and associated spatiotemporal variability. This test is arguably easier passed in the SST paradigm, which relies directly on observed rainfall properties. SRGs must depict these processes using space-time statistics and have historically struggled to reproduce extremes (Furrer and Katz, 2008; Willems et al., 2012). To date, no detailed comparison of SST and an SRG for RFA or FFA has been conducted. Furthermore, no effort since Wilson and Foufoula-Georgiou, (1990) has sought to merge progress in SRGs with SST.

5.5 Process-Based FFA and Flood Nonstationarity

To our knowledge, the term "process-based FFA" was first introduced in Sivapalan and Samuel, (2009), but the general approach has been previously referred to as "derived FFA" (e.g. Eagleson, 1972; Franchini et al., 2005; Haberlandt et al., 2008) and "physically-based FFA" (e.g. Díaz-Granados et al., 1984; Muzik, 1993; Shen et al., 1990). Briefly introduced in Section 4.4, process-based FFA merits further discussion. Essentially, it aims to "construct" distributions of flood outcomes by recreating the joint distributions of the relevant flood-producing processes: rainfall, watershed states such as soil moisture and snowpack, surface and subsurface flow, and river channel routing. A key aspect of these joint distributions is seasonality. In the Midwestern U.S., for example, snowpack usually persists only into March, while soil moisture peaks in March-April and peak rainfall rates occur in June-August (Yu et al., 2019). This means that floods usually occur sometime in March-August, but the specific combinations of processes that cause them can vary dramatically over this timeframe.

Nonstationarity, particularly climatic change, manifests as changes in seasonality of one or more processes. This implies that process-based approaches may be necessary for translating nonstationarity in the hydrologic cycle into nonstationarity in flood frequency. The results of Wright et al. (2017) and Yu et al., (2019) show the potential of SST combined with process-based approaches in one watershed exhibiting hydrologic nonstationarity. Here, we demonstrate the opportunity to examine the relative roles of rainfall and antecedent soil in determining flood peak variability (Fig. 8) using RainyDay with Stage IV gage-corrected radar rainfall and the WRF-Hydro distributed hydrologic model (Gochis et al., 2018).

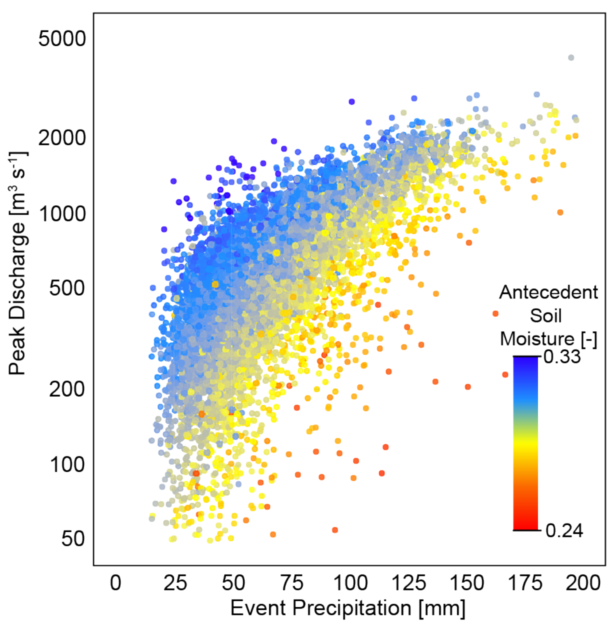

Fig. 8: Relationships between accumulated event rainfall over watershed and simulated peak discharge generated using RainyDay-based SST, Stage IV rainfall, and the WRF-Hydro rainfall runoff model for the Turkey River at Garber, Iowa. Results are based on 20 distinct realizations, each consisting of 500

simulated annual flood peaks. Color indicates the randomized antecedent basin-averaged volumetric water content of the 2 m soil column, used to initialize simulations. The watershed and transposition domain are shown in Fig. 6.

In contrast with process-based FFA, flood-based statistical methods "do not explicitly consider the physical processes that produce floods" (Wright et al., 2014). Rather, the sample of flood observations used in those methods represent outcomes of combinations of these processes. Whether this sample space adequately describes the population of possible floods that could result from different combinations is unknowable. Furthermore, how these combinations change with spatial scale and with climate or land use change is not clear from streamflow records alone. Thus, it is not obvious how to generate estimates of future flood quantiles, even using nonstationary statistical models (Sivapalan and Samuel, 2009; Stedinger and Griffis, 2011).

The rainfall-runoff models used in design storm methods can explicitly represent important process interactions to a degree. Design storms require strong assumptions regarding the joint distributions of these processes, however, including: 1.) an annual maximum flood peak result from rainfall annual maxima, even though a smaller rainstorm plus wetter initial conditions could produce an annual maximum; 2.) a single antecedent soil moisture value is sufficient to understand flood outcomes, even though real-world soil moisture is variable; and 3.) a rainfall IDF quantile translates directly to the same flood quantile (i.e. a 100-year storm produces a 100-year flood). The consequences of these assumptions are not well understood (Adams and Howard, 1986; Curtis et al., 2013a,b; Packman and Kidd, 1980; Wright et al., 2014, 2013a, 2013b).

Hydrologic practice has been well-served by both statistically-based and design storm-based FFA, due to their strengths and also to healthy doses of conservatism in engineering design practices. Both are starting to "show their age," however. Published IDF estimates have been shown to be out-of-date, often severely underestimating the frequency of extreme rainfalls in the U.S. (Chin and Ross, 2018; Wright et al., 2019) and elsewhere (e.g. Madsen et al., 2014, 2009). Though the issue of flood nonstationarity is less clear (e.g. Sharma et al., 2018), recent work has argued that future progress on RFFA will require more careful consideration of watershed geomorphology (Ayalew and Krajewski, 2017) and hydrometeorology (Smith et al., 2018). Consistent with the conclusions of the National Research Council (1994), storm-based approaches such as SST, together with process-based FFA, are better positioned than conventional station-based approaches to leverage advances in distributed modeling and in rainfall remote sensing and numerical simulation.

5.6 Multiscale FFA

The concept of multi-scale FFA was briefly introduced in Section 4.4. There are applications in which flood quantiles across an entire drainage network are needed, rather than for an individual river reach or gage location. Such applications include mapping of floodplains and the design of levee and stormwater systems. In addition, it can be useful to be able to model flood scenarios that resolve the spatial distribution of impacts within an individual flood event. Applications that can benefit from this capability include probabilistic risk assessment, "stress-testing" of actively managed hydrologic infrastructure such as reservoirs, or evaluation of transportation network response to flood impacts on bridges or roadways.

In such applications, depiction of a wide range of storm events with realistic spatiotemporal distribution of rainfall across the watershed, and the ability to resolve the resulting flood response using distributed rainfall-runoff modeling, is essential (e.g. Pilgrim, 1986). SST and some other methods (see Section 5.4) are able to do these things, provided that adequate input rainfall data are available, a well-performing rainfall runoff model is used, and suitable approaches are employed to account for variability in initial conditions. Thus, the concepts of multiscale FFA and process-based FFA are linked—the former requires the latter, while the latter requires the former if scale-dependent interactions in flood generating mechanisms are to be properly considered.

We demonstrate several examples of process-based SST, generated using RainyDay together with Stage IV rainfall and WRF-Hydro, including the distribution of specific flood quantiles across the river network (Fig. 9a,b). Also shown are two random storm transpositions with very different rainfall patterns (Fig. 9c,d) and spatial patterns of resulting flood peak discharges (Fig. 9e,f), despite having both produced 100-year peak discharges at the watershed outlet at Garber, Iowa.

5.7 Rainfall-Runoff Models for FFA

Both design storm methods and the process-based and multiscale FFA described in Sections 5.5 and 5.6 rely on rainfall-runoff hydrologic models. A common criticism is that parameter and structural uncertainty in such models is unacceptably large. While these can indeed be problematic, we provide several counterarguments. First, the confidence intervals provided in flood-based statistical FFA methods understate the true uncertainty inherent in such methods by ignoring the potentially major role of rating curve errors (Potter and Walker, 1985). When these errors are considered, uncertainty has been shown to balloon dramatically (Steinbakk et al., 2016). Second, there have been numerous demonstrations that rainfall-runoff FFA methods can perform as well or better than statistical methods in cases of basin storage “discontinuities” (Rogger et al., 2012), reservoirs (Ayalew et al., 2013), and land use or climatic change (Cunha et al., 2011). Third, at least some “blame” for poor model performance lies on the precipitation and other meteorological inputs. The role of rainfall-runoff model error has not yet been studied in SST-based FFA, and would presumably yield wider uncertainty bounds than those shown in existing studies. The non-SST FFA work of Blazkova and Beven (2009, 2002) and Cameron et al. (1999) provide a possible roadmap for considering model errors.

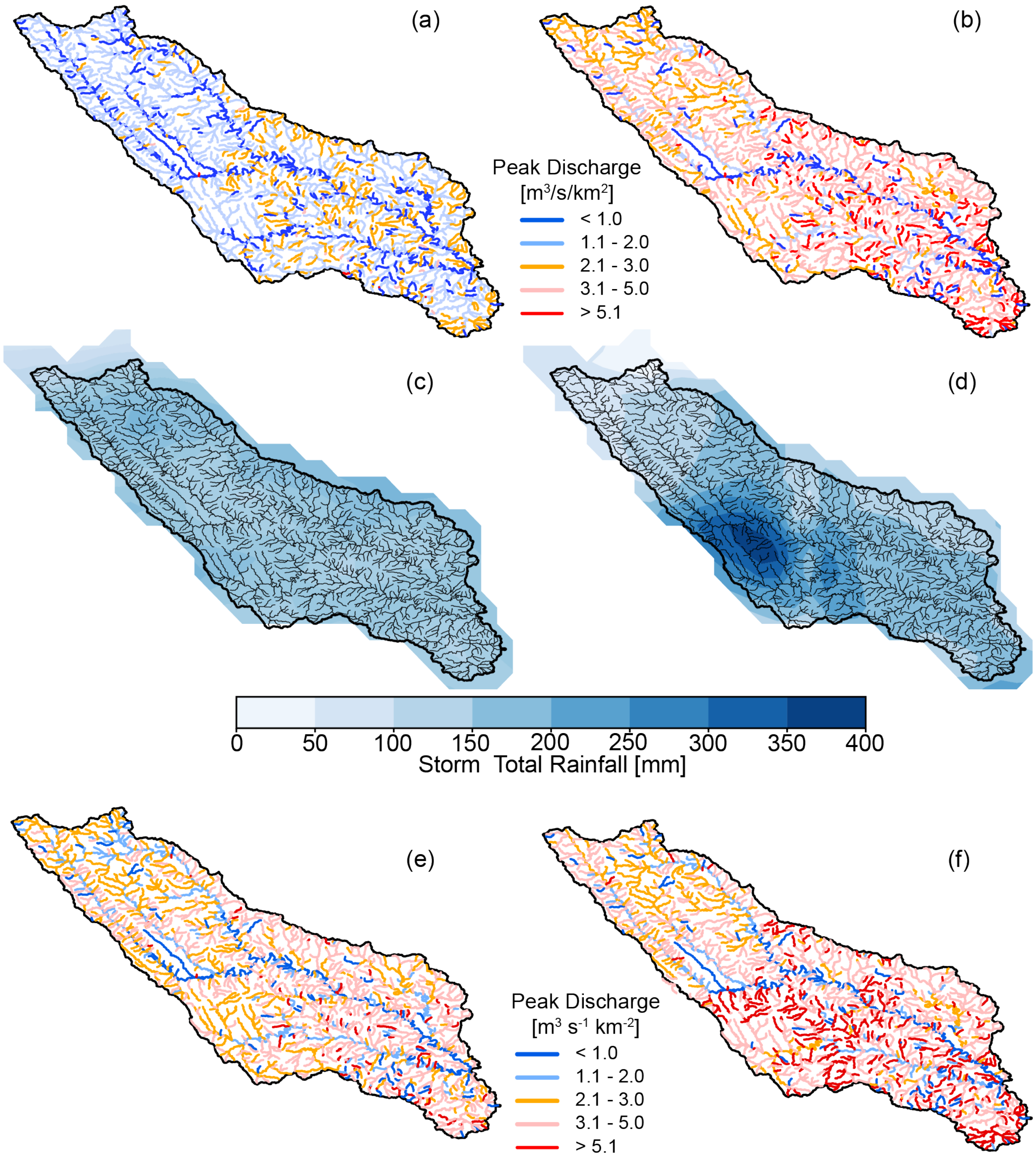

Fig. 9: Multiscale FFA results over the Turkey River watershed in northeastern Iowa using RainyDay-based SST, Stage IV rainfall, and the WRF-Hydro rainfall runoff model. Results are based on 20 distinct realizations, each consisting of 500 simulated annual flood peaks. (a) and (b) show the median of the 10-year and 100-year flood peak magnitude, respectively, based on the 20 realizations. (c) and (d) show rainfall maps for two random transpositions which produced 100-year peak discharges at the watershed

outlets, but featured very different rainfall spatial distributions. (e) and (f) show the resulting peak discharges produced by the transposed rainstorms shown in (c) and (d). The watershed and transposition domain are shown in Fig. 6.

6. Summary and Recommendations

In this review, we summarize the origins of Stochastic Storm Transposition (SST) in the context of three better-known forms of rainfall and flood hazard estimation: regionalized rainfall and flood frequency analysis, design storms, and PMP/PMF. We briefly explain the methodology, review existing research, discuss some of SST's properties, strengths, and limitations, and contrast it with other methods.

In the six decades since Alexander's 1963 description of SST in Issue 1 of the Journal of Hydrology, the problem of estimating the likelihood and magnitude of floods has not been solved. Indeed, climate change, economic growth, and urbanization mean that risks have and will likely continue to grow (Kundzewicz et al., 2014). At the same time, recent experiences have identified weaknesses in longstanding methods for RFA, FFA, and PMP/PMF; climatic and land cover changes are particularly challenging.

As a "storm-based" approach built explicitly around the spatiotemporal variability of rainfall, SST holds promise to address these issues. Critically, it is able to leverage advances from other branches of hydrology and from related fields such as meteorology—including distributed hydrologic modeling and remote sensing and numerical simulation of extreme rainfall. Example RFA and FFA applications for watersheds in Louisiana, Iowa, Wisconsin, Maryland, Colorado, and southeastern Australia have demonstrated the practical utility of SST for floodplain management and dam safety.

While six decades of research into the concept suggests that SST is a viable complement to existing approaches, important questions remain. We conclude with five areas for future work:

1. A chief criticism of SST is the subjectivity involved in defining the transposition domain. Definition of transposition domains based on climatological characteristics combined with modest rescaling of storms represent a reasonable compromise between the desire for large domains and the need for approximate homogeneity. Previous SST work has begun to address this (Nathan et al., 2016; Wilson and Foufoula-Georgiou, 1990; Wright et al., 2019; Zhou et al., 2019), but more is needed. Methods from related fields such as regionalized L-moments (Hosking and Wallis, 1997) should be considered.
2. Errors from rainfall measurements, extreme storm sampling, and rainfall-runoff modeling should be examined to understand the propagation of such errors through to SST-based FFA estimates. These errors should be compared side-by-side with those resulting from flood-based statistical FFA to better understand the relative strengths and limits of each method. The Bayesian approach used in Steinbakk et al. (2016) both highlights the importance of error propagation in FFA and provides a roadmap for studying it.
3. England et al. (2014) showed that SST can be used alongside other methods to develop a “preponderance of evidence” approach to FFA, while Perez et al. (2019) showed that SST-based FFA can help to evaluate other methodologies. More explicit “merging” of SST with other RFA/FFA approaches and rainfall estimation techniques is likely to prove valuable. An example of the former could be the usage of SST to help estimate the skewness of rainfall and

flood distributions, a major challenge in conventional FFA. Examples of the latter could include integration of SST and modern high-resolution stochastic rainfall generators and explicit coupling of SST with numerical weather prediction models that can explicitly simulate rainfall in complex terrain.

4. Process-based and multi-scale FFA concepts can connect flood processes to flood distributions, including in nonstationary conditions. Though previous studies have begun to explore these connections, rigorous interpretation of results in terms of hydrologic processes such as runoff generation mechanisms and channel routing has been lacking. A key question is whether a return to the simplified storm and watershed representations of early SST work could prove illuminating, in this effort, or if real-world process complexities would limit the value of such idealizations.
5. Finally, while SST research has been generally confined to the United States and Australia, the need for rainfall and flood frequency estimation is widespread. While global precipitation estimates using satellites and atmospheric reanalyses are improving to the point that they may be useful in RFA and FFA applications, validation of these results in ungaged regions remains a challenge. Nonetheless, numerous regions around the globe spanning diverse hydroclimatic and socioeconomic conditions have at least some high-quality rainfall and flood observations. Partnerships with researchers and end users in those regions would help to explore the potential for SST to "go global."

Acknowledgements

The authors thank Stephen J. Burges and James A. Smith for their help in identifying pioneering early rainfall and flood studies. Daniel Wright's and Guo Yu's efforts were supported by the U.S. National Science Foundation Hydrologic Sciences Program CAREER project EAR-1749638 and by the U.S. Department of the Interior Bureau of Reclamation Research and Development Office Project 1735. The authors would like to acknowledge the support of the various funding organizations that have facilitated the SST research that is reviewed in this paper.

References

Abbs, D.J., 1999. A numerical modeling study to investigate the assumptions used in the calculation of probable maximum precipitation. Water Resour. Res. 35, 785–796. https://doi.org/10.1029/1998WR900013

Adams, B.J., Howard, C.D.D., 1986. Design Storm Pathology. Can. Water Res J. 11, 49–55.

Agho, K., Kuczera, G., Green, J., Weinmann, E., Laurenson, E., 2000. Estimation of Extreme Rainfall Exceedance Probabilities: Nondirnensional Stochastic Storm Transposition, in: Proceedings. Presented at the International Hydrology and Water Resources Symposium PerthInternational Hydrology and Water Resources Symposium Perth, Engineers Australia, Canberra, pp. 60–65.

Alexander, G., 1969. Application of probability to spillway design flood estimation, in: Proceedings of the Leningrad Symposium on Floods and Their Computation, August 1967. International Association of Scientific Hydrology, Gentbrugge, Belgium.

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

Arnoud, P., Bouvier, C., Cisneros, L., Dominguez, R., 2002. Influence of rainfall spatial variability on flood prediction. J. Hydrol. 260, 216–230. https://doi.org/10.1016/S0022-1694(01)00611-4

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

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

Ball, J., Babister, M., Nathan, R., Weeks, W., Weinmann, E., Retallick, M., Testoni, I., 2019. Australian Rainfall and Runoff: A Guide to Flood Estimation. Commonwealth of Australia (Geoscience Australia).

Beran, M., 1973. Estimation of design floods and the problem of equating the probability of rainfall and runoff. Presented at the Symposium on the Design of Water Resources Projects with Inadequate Data, Madrid, Spain, pp. 33–50.

Bernard, M.M., 1936. The unit-hydrograph method and storm transposition in flood problems relating to great storms In the Eastern and Central United States, in: Studies of Relations of Rainfall and Run-off in the United States, U.S. Geological Survey Water-Supply Paper. United States Government Printing Office, Washington, DC, p. 301.

Blazkova, S., Beven, K., 2009. 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, n/a-n/a. https://doi.org/10.1029/2007WR006726

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

Bonnin, G.M., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D., 2006. NOAA Atlas 14: Precipitation-frequency atlas of the United States Volume 2. National Weather Service, National Oceanic and Atmospheric Administration, Silver Spring, Maryland.

Cameron, D.S., Beven, K.J., Tawn, J., Blazkova, S., Naden, P., 1999. Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). J. Hydrol. 219, 169–187. https://doi.org/10.1016/S0022-1694(99)00057-8

Charalambous, J., Rahman, A., Carroll, D., 2013. Application of Monte Carlo Simulation Technique to Design Flood Estimation: A Case Study for North Johnstone River in Queensland, Australia. Water Resour. Manag. 27, 4099–4111. https://doi.org/10.1007/s11269-013-0398-9

Chin, D.A., Ross, E., 2018. Canonical Rainfall Distributions in the United States. J. Irrig. Drain. Eng. 144, 04018031. https://doi.org/10.1061/(ASCE)IR.1943-4774.0001350

Chow, V.T., 1964. Frequency Analysis, in: Handbook of Applied Hydrology: A Compendium of Water-Resources Technology. McGraw-Hill, New York, NY.

Chu, X., Steinman, A., 2009. Event and Continuous Hydrologic Modeling with HEC-HMS. J. Irrig. Drain. Eng. 135, 119–124. https://doi.org/10.1061/(ASCE)0733-9437(2009)135:1(119)

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

Curtis, D.C., 2007. Evaluation of the Spatial Structure of Storms and the Development of Design Storms. Presented at the World Environmental and Water Resources Congress 2007: Restoring Our Natural Habitat, pp. 1–10.

Curtis, D.C., Boucher, M., Prakash, O., Martinez, B., Shurtz, K., 2013a. Benchmarking DARF, Design Storms, and Temporal Distribution Procedures for Hydrologic Design, in: World Environmental and Water Resources Congress 2013. pp. 117–128. https://doi.org/10.1061/9780784412947.011

Curtis, D.C., Martinez, B., Prakash, O., 2013b. Antecedent Precipitation Analysis for Radar-Based Design Storm Development in Phoenix, AZ, in: World Environmental and Water Resources Congress 2013. pp. 77–86. https://doi.org/10.1061/9780784412947.008

Davison, A.C., Huser, R., 2015. Statistics of Extremes. Annu. Rev. Stat. Its Appl. 2, 203–235. https://doi.org/10.1146/annurev-statistics-010814-020133

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

Dawdy, D.R., Lettenmaier, D.P., 1987. Initiative for Risk-Based Flood Design. J. Hydraul. Eng. 113, 1041–1051. https://doi.org/10.1061/(ASCE)0733-9429(1987)113:8(1041)

Díaz-Granados, M.A., Valdes, J.B., Bras, R.L., 1984. A Physically Based Flood Frequency Distribution. Water Resour. Res. 20, 995–1002. https://doi.org/10.1029/WR020i007p00995

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

Eash, D.A., Barnes, K.K., Veilleux, A.G., 2013. Methods for estimating annual exceedance-probability discharges for streams in Iowa, based on data through water year 2010: U.S. Geological Survey Scientific Investigations Report 2013-5086.

England, J.F., 2011. Flood Frequency and Design Flood Estimation Procedures in the United States: Progress and Challenges. Australas. J. Water Resour. 15, 33–46. https://doi.org/10.1080/13241583.2011.11465388

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

Faridzad, M., Yang, T., Hsu, K., Sorooshian, S., Xiao, C., 2018. Rainfall frequency analysis for ungauged regions using remotely sensed precipitation information. J. Hydrol. 563, 123–142. https://doi.org/10.1016/j.jhydrol.2018.05.071

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

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

Fuller, W.E., 1914. Flood Flows. ASCE Trans 77, 567–617.

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

Gochis, D.J., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., McCreight, J., Mills, J., RafieeiNasab, A., Read, L., Sampson, K., Yates, D.N., Yu, W., 2018. The WRF-Hydro model technical description and user's guide, version 5.0 (NCAR Technical Note).

Gupta, V.K., 1972. Transposition of Storms for Estimating Flood Probability Distributions. Colorado State University, Fort Collins, CO.

Haberlandt, U., Ebner von Eschenbach, A.-D., Buchwald, I., 2008. A space-time hybrid hourly rainfall model for derived flood frequency analysis. Hydrol Earth Syst Sci 12, 1353–1367. https://doi.org/10.5194/hess-12-1353-2008

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

Hershfield, D.M., 1961. Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1-100 Years (No. Technical Paper 40). Weather Bureau, Department of Commerce, Washington, D.C.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511529443

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

Interagency Advisory Committee on Water Data (IACWD), 1982. Guidelines for determining flood flow frequency, Bulletin 17B. Reston, VA.

Knowles, S.G., Kunreuther, H.C., 2014. Troubled waters: The national flood insurance program in historical perspective. J. Policy Hist. 26, 327–353.

Krajewski, W.F., Ceynar, D., Demir, I., Goska, R., Kruger, A., Langel, C., Mantilla, R., Niemeier, J., Quintero, F., Seo, B.-C., Small, S.J., Weber, L.J., Young, N.C., 2017. Real-Time Flood Forecasting and Information System for the State of Iowa. Bull. Am. Meteorol. Soc. 98, 539–554. https://doi.org/10.1175/BAMS-D-15-00243.1

Kundzewicz, Z.W., Kanae, S., Seneviratne, S.I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L.M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G.R., Kron, W., Benito, G., Honda, Y., Takahashi, K., Sherstyukov, B., 2014. Flood risk and climate change: global and regional perspectives. Hydrol. Sci. J. 59, 1–28. https://doi.org/10.1080/02626667.2013.857411

Lin, Y., 2011. GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data. Version 1.0.

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

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., Kjeldsen, T.R., 2014. Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. J. Hydrol. 519, 3634–3650. https://doi.org/10.1016/j.jhydrol.2014.11.003

Mantilla, R., Gupta, V.K., 2005. A GIS numerical framework to study the process basis of scaling statistics in river networks. Geosci. Remote Sens. Lett. IEEE 2, 404–408.

Marston, F.A., 1924. The Distribution of Intense Rainfall and Some Other Factors in the Design of Storm-Water Drains. Trans ASCE 87, 56.

McGraw, D., Nikolopoulos, E.I., Marra, F., Anagnostou, E.N., 2019. Precipitation frequency analyses based on radar estimates: An evaluation over the contiguous United States. J. Hydrol. 573, 299–310. https://doi.org/10.1016/j.jhydrol.2019.03.032

Meierdiercks, K.L., Smith, J.A., Baeck, M.L., Miller, A.J., 2010. Analyses of Urban Drainage Network Structure and its Impact on Hydrologic Response. JAWRA J. Am. Water Resour. Assoc. 46, 932–943. https://doi.org/10.1111/j.1752-1688.2010.00465.x

Mejia, J.M., Moglen, G.E., 2010. Impact of the spatial distribution of imperviousness on the hydrologic response of an urbanizing basin. Hydrol. Process. 24, 3359–3373. https://doi.org/10.1002/hyp.7755

Meyer, A.F., 1917. The Elements of Hydrology, 1st ed. John Wiley and Sons, New York, NY.

Miami Conservancy District, 1917. Storm Rainfall of Eastern United States (No. 5), Technical Reports. Dayton, OH.

Miller, J.F., 1964. Two-to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States (No. Technical Paper 49). Weather Bureau, Department of Commerce, Washington, D. C.

Morin, E., Goodrich, David.C., Maddox, R.A., Gao, X., Gupta, H.V., Sorooshian, S., 2006. Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. Adv. Water Resour. 29, 843–860. https://doi.org/10.1016/j.advwatres.2005.07.014

Muzik, I., 1993. Derived, physically based distributions of flood probabilities. IAHS Publ. 183–183.

Myers, V.A., 1969. The estimation of extreme precipitation as the basis for design floods—Resume of practice in the United States, in: Proceedings of the Leningrad Symposium on Floods and Their Computation, August 1967. International Association of Scientific Hydrology, Gentbrugge, Belgium.

Myers, V.A., 1966. Criteria and Limitations for the Transposition of large storms over various size watersheds, in Symposium on Consideration of Some Aspects of Storms and Floods in Water Planning, Texas Water Development Board Report 33.

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

Nathan, R.J., Weinmann, E., Minty, L., 1999. Estimation of the annual exceedance probability of probable maximum precipitation in South-East Australia. Aust. J. Water Resour. 3, 143–153.

National Research Council, 1994. Estimating Bounds on Extreme Precipitation Events: A Brief Assessment. National Academy Press, Washington, D.C.

National Research Council, 1988. Estimating probabilities of extreme floods: methods and recommended research. National Academy Press, Washington, D.C.

Norbiato, D., Borga, M., Sangati, M., Zanon, F., 2007. Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood. J. Hydrol. 345, 149–166. https://doi.org/10.1016/j.jhydrol.2007.07.009

Ohara, N., Kavvas, M.L., Kure, S., Chen, Z.Q., Jang, S., Tan, E., 2011. Physically Based Estimation of Maximum Precipitation over American River Watershed, California. J. Hydrol. Eng. 16, 351–361. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000324

Packman, J., Kidd, C., 1980. A logical approach to the design storm concept. Water Resour. Res. 16, 994–1000.

Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S., Burlando, P., 2014. On the effects of small scale space-time variability of rainfall on basin flood response. J. Hydrol. 514, 313–327. https://doi.org/10.1016/j.jhydrol.2014.04.014

Paschalis, A., Molnar, P., Fatichi, S., Burlando, P., 2013. A stochastic model for high-resolution space-time precipitation simulation. Water Resour. Res. 49, 8400–8417. https://doi.org/10.1002/2013WR014437

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

Peleg, N., Morin, E., 2014. Stochastic convective rain-field simulation using a high-resolution synoptically conditioned weather generator (HiReS-WG). Water Resour. Res. 50, 2124–2139. https://doi.org/10.1002/2013WR014836

Perez, G., Mantilla, R., Krajewski, W.F., Wright, D.B., 2019. Using Physically Based Synthetic Peak Flows to Assess Local and Regional Flood Frequency Analysis Methods. Water Resour. Res. n/a. https://doi.org/10.1029/2019WR024827

Pilgrim, D.H., 1986. Bridging the gap between flood research and design practice. Water Resour. Res. 22, 165S-176S. https://doi.org/10.1029/WR022i09Sp0165S

Potter, K.W., Walker, J.F., 1985. An Empirical Study of Flood Measurement Error. Water Resour Res 21, 403–406. https://doi.org/10.1029/WR021i003p00403

Prein, A.F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N.P.M., Leung, R.,

2015. A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. Rev. Geophys. 53, 323–361. https://doi.org/10.1002/2014RG000475

Rahman, A., Weinmann, P.E., Hoang, T.M.T., Laurenson, E.M., 2002. Monte Carlo simulation of flood frequency curves from rainfall. J. Hydrol. 256, 196–210. https://doi.org/10.1016/S0022-1694(01)00533-9

Ramos, M.H., Creutin, J.D., Leblois, E., 2005. Visualization of storm severity. J. Hydrol. 315, 295–307. https://doi.org/10.1016/j.jhydrol.2005.04.007

Requena, A.I., Burn, D.H., Coulibaly, P., 2019. Pooled frequency analysis for intensity–duration–frequency curve estimation. Hydrol. Process. 33, 2080–2094. https://doi.org/10.1002/hyp.13456

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

Schaefer, M.G., Barker, B.L., 2002. Stochastic Event Flood Model, Chapter 20, in: Singh, V.P. and Frevert, D.K. Eds., Mathematical Models of Small Watershed Hydrology and Applications. Water Resources Publication.

Sharma, A., Mehrotra, R., 2010. Rainfall generation, in: Washington DC American Geophysical Union Geophysical Monograph Series. American Geophysical Union, Washington, DC, pp. 215–246.

Sharma, A., Wasko, C., Lettenmaier, D.P., 2018. If Precipitation Extremes Are Increasing, Why Aren’t Floods? Water Resour. Res. 54, 8545–8551. https://doi.org/10.1029/2018WR023749

Shen, H.W., Koch, G.J., Obeysekera, J.T.B., 1990. Physically Based Flood Features and Frequencies. J. Hydraul. Eng. 116, 494–514. https://doi.org/10.1061/(ASCE)0733-9429(1990)116:4(494)

Showalter, A.K., Solot, S.B., 1942. Computation of maximum possible precipitation. Eos Trans. Am. Geophys. Union 23, 258–274. https://doi.org/10.1029/TR023i002p00258

Sivapalan, M., Beven, K., Wood, E.F., 1987. On hydrologic similarity: 2. A scaled model of storm runoff production. Water Resour. Res. 23, 2266–2278. https://doi.org/10.1029/WR023i012p02266

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

Smith, J.A., Baeck, M.L., 2015. “Prophetic vision, vivid imagination”: The 1927 Mississippi River flood. Water Resour. Res. 51, 9964–9994. https://doi.org/10.1002/2015WR017927

Smith, J.A., Baeck, M.L., Meierdiercks, K.L., Nelson, P.A., Miller, A.J., Holland, E.J., 2005. Field studies of the storm event hydrologic response in an urbanizing watershed. Water Resour Res 41, W10413–W10413. https://doi.org/10.1029/2004WR003712

Smith, J.A., Baeck, M.L., Morrison, J.E., Sturdevant-Rees, P., Turner-Gillespie, D.F., Bates, P.D., 2002. The Regional Hydrology of Extreme Floods in an Urbanizing Drainage Basin. J. Hydrometeorol. 3, 267–282. https://doi.org/10.1175/1525-7541(2002)003<0267:TRHOEF>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

Smith, J.A., Cox, A.A., Baeck, M.L., Yang, L., Bates, P., 2018. Strange Floods: The Upper Tail of Flood Peaks in the United States. Water Resour. Res. 54, 6510–6542. https://doi.org/10.1029/2018WR022539

Stedinger, J.R., Griffis, V.W., 2011. Getting From Here to Where? Flood Frequency Analysis and Climate. JAWRA J. Am. Water Resour. Assoc. 47, 506–513. https://doi.org/10.1111/j.1752-1688.2011.00545.x

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events, in: D. R. Maidment (Ed.), Handbook of Hydrology. McGraw-Hill, New York.

Steinbakk, G.H., Thorarinsdottir, T.L., Reitan, T., Schlichting, L., Hølleland, S., Engeland, K., 2016. Propagation of rating curve uncertainty in design flood estimation. Water Resour. Res. 52, 6897–6915. https://doi.org/10.1002/2015WR018516

Stephens, D., Nathan, R., Hill, P., Scorah, M., 2016. Incorporation of snowmelt into joint probability event based rainfall-runoff modelling. Presented at the 37th Hydrology & Water Resources Symposium 2016: Water, Infrastructure and the Environment, Engineers Australia, p. 532.

Stephenson, A.G., Lehmann, E.A., Phatak, A., 2016. A max-stable process model for rainfall extremes at different accumulation durations. Weather Clim. Extrem. 13, 44–53. https://doi.org/10.1016/j.wace.2016.07.002

Stokstad, E., 1999. Scarcity of Rain, Stream Gages Threatens Forecasts. Science 285, 1199. https://doi.org/10.1126/science.285.5431.1199

Svensson, C., Jones, D.A., 2010. Review of rainfall frequency estimation methods. J. Flood Risk Manag. 3, 296–313. https://doi.org/10.1111/j.1753-318X.2010.01079.x

Swain, R.E., Bowles, D., Ostenaa, D., 1998. A framework for characterization of extreme floods for dam safety risk assessments, in: Proceedings of the 1998 USCOLD Annual Lecture, Buffalo, New York.

Swain, R.E., England, J.F., Bullard, K.L., Raff, D.A., 2006. Guidelines for evaluating hydrologic hazards. US Department of the Interior, Bureau of Reclamation, Denver, CO.

Toride, K., Cawthorne, D.L., Ishida, K., Kavvas, M.L., Anderson, M.L., 2018. Long-term trend analysis on total and extreme precipitation over Shasta Dam watershed. Sci. Total Environ. 626, 244–254. https://doi.org/10.1016/j.scitotenv.2018.01.004

USACE, U.S.A.C. of E., 1973. Storm Rainfall in the United States. Office of the Chief of Engineers, Washington, D.C.

USBR, 2013. Design Standards No. 14 Appurtenant Structures for Dams (Spillways and Outlet Works). U.S. Department of Interior Bureau of Reclamation, Denver, CO.

USBR, USACE, 2018. Dam Safety Risk Analysis Best Practices Training Manual. U.S. Bureau of Reclamation and U.S. Army Corps of Engineers.

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

Willems, P., Olsson, J., Arnbjerg-Nielsen, K., Beecham, S., Pathirana, A., Gregersen, I.B., Madsen, H., Nguyen, V.-T.-V., 2012. Impacts of Climate Change on Rainfall Extremes and Urban Drainage Systems. IWA Publishing. https://doi.org/10.2166/9781780401263

Wilson, D., Fleig, A.K., Lawrence, D., Hisdal, H., Pettersson, L.-E., Holmqvist, E., 2011. A review of NVE's flood frequency estimation procedures (No. 9). Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Wilson, L.L., Foufoula-Georgiou, E., 1990. Regional Rainfall Frequency Analysis via Stochastic Storm Transposition. J. Hydraul. Eng. 116, 859–880. https://doi.org/10.1061/(ASCE)0733-9429(1990)116:7(859)

Woodward, S.M., 1920. Hydraulics of the Miami flood control project (No. VII), Technical Reports. Miami Conservancy District, Dayton, OH.

Wright, D.B., 2019. RainyDay Rainfall Hazard Analysis System for Python 3. https://doi.org/10.5281/zenodo.3478516

Wright, D.B., 2018. Rainfall Information for Global Flood Modeling, in: Global Flood Hazard: Applications in Modeling, Mapping, and Forecasting (Eds G. J. Schumann, P. D. Bates, H. Apel and G. T. Aronica). Wiley.

Wright, D.B., Bosma, C.D., Lopez-Cantu, T., 2019. U.S. Hydrologic Design Standards Insufficient Due to Large Increases in Frequency of Rainfall Extremes. Geophys. Res. Lett. 46, 8144–8153. https://doi.org/10.1029/2019GL083235

Wright, D.B., Holman, K.D., 2019. Rescaling Transposed Extreme Rainfall within a Heterogeneous Region. J. Hydrol. Eng. 24, 06019001. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001781

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., Baeck, M.L., 2013a. Critical examination of area reduction factors. J. Hydrol. Eng. 19, 769–776.

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

Yang, L., Smith, J.A., Wright, D.B., Baeck, M.L., Villarini, G., Tian, F., Hu, H., 2013. Urbanization and Climate Change: An Examination of Nonstationarities in Urban Flooding. J Hydrometeor.

YEAC, 1984. Probability of extreme rainfalls and the effect on the Harriman Dam. Yankee Atmoic Electric Company, Framingham, MA.

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

Zhou, Z., Smith, J.A., Wright, D.B., Baeck, M.L., Yang, L., Liu, S., 2019. Storm Catalog-Based Analysis of Rainfall Heterogeneity and Frequency in a Complex Terrain. Water Resour. Res. 55, 1871–1889. https://doi.org/10.1029/2018WR023567

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