Journal of Hydrology 562 (2018) 648–663

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

How extreme was the October 2015 flood in the Carolinas? An assessment of flood frequency analysis and distribution tails

R.C. Phillips[a,b], S.Z. Samadi[a,*], M.E. Meadows[a]

[a] Department of Civil and Environmental Engineering, University of South Carolina, Columbia, SC, USA
[b] Davis & Floyd, Inc., SC, USA

ARTICLE INFO

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Gokmen Tayfur, Associate Editor

Keywords:
Flood frequency analysis
Return period
Distribution tails
The carolinas

ABSTRACT

This paper examines the frequency, distribution tails, and peak-over-threshold (POT) of extreme floods through analysis that centers on the October 2015 flooding in North Carolina (NC) and South Carolina (SC), United States (US). The most striking features of the October 2015 flooding were a short time to peak (T_p) and a multi-hour continuous flood peak which caused intensive and widespread damages to human lives, properties, and infrastructure. The 2015 flooding was produced by a sequence of intense rainfall events which originated from category 4 hurricane Joaquin over a period of four days. Here, the probability distribution and distribution parameters (i.e., location, scale, and shape) of floods were investigated by comparing the upper part of empirical distributions of the annual maximum flood (AMF) and POT with light- to heavy- theoretical tails: Fréchet, Pareto, Gumbel, Weibull, Beta, and Exponential. Specifically, four sets of U.S. Geological Survey (USGS) gauging data from the central Carolinas with record lengths from approximately 65–125 years were used. Analysis suggests that heavier-tailed distributions are in better agreement with the POT and somewhat AMF data than more often used exponential (light) tailed probability distributions. Further, the threshold selection and record length affect the heaviness of the tail and fluctuations of the parent distributions. The shape parameter and its evolution in the period of record play a critical and poorly understood role in determining the scaling of flood response to intense rainfall.

1. Introduction

The October 3–5, 2015 historic rains caused by hurricane Joaquin released more than 500 mm of rain in South Carolina (SC) and North Carolina (NC), United States (US). The flood peak of many U.S. Geological Survey (USGS) gauges, including those located in the center of SC, were almost twice the previous maximum from a record of over 65 years. The spatial extent of flooding in this portion was also unprecedented, with more record flood peaks at USGS stream gauging stations in urban areas such as Columbia, the capital of SC, than for any other rural catchments.

Such an extraordinary flood lies within the fundamental issue of infrastructure safety and raises the crucial question of how to proceed if this event is not visible for a given dataset and if it is too rare for design applications. Although recently significant progress has been made to predict short-term flood for operational purposes (e.g., Pourreza-Bilondi et al., 2017), long-term prediction, on which infrastructure design is based, is difficult in deterministic terms (e.g., Papalexiou and Koutsoyiannis, 2013). Thus, it is common to treat this event in a probabilistic manner (i.e., as a random variable) that is governed by a distribution law. Such a distribution enables the modeler to capture the probability of exceedance and assign a return period to any flood event, the procedure called flood frequency analysis (FFA) in design hydrology.

Assessment of flood probability has been an active research topic, yet a less understood concept. However, the analysis is well rooted in an extensive literature dating back to the work of Nicolaus Bernoulli three centuries ago (mentioned in Gumbel, 1958). Extreme value theory (EVT) was the first and widely accepted method for FFA that has rapidly evolved and found applications in engineering hydrology. Fuller's (1914) study was probably the first application of extreme value distributions. Some recent studies, such as Papalexiou and Koutsoyiannis (2013) and Serinaldi and Kilsby (2014, 2015), expanded the EVT concepts for hydrological design applications. Specifically, EVT has stimulated an extensive investigation to estimate the parent distribution (e.g., Michele and Rosso, 2001; Bernardara et al., 2008) and (upper) tail behaviors of flood properties (Papalexiou and Koutsoyiannis, 2013; Serinaldi and Kilsby, 2015), just to mention a few recent studies.

* Corresponding author.
E-mail address: samadi@cec.sc.edu (S.Z. Samadi).

https://doi.org/10.1016/j.jhydrol.2018.05.035
Received 27 December 2017; Received in revised form 24 April 2018; Accepted 14 May 2018
Available online 19 May 2018
0022-1694/

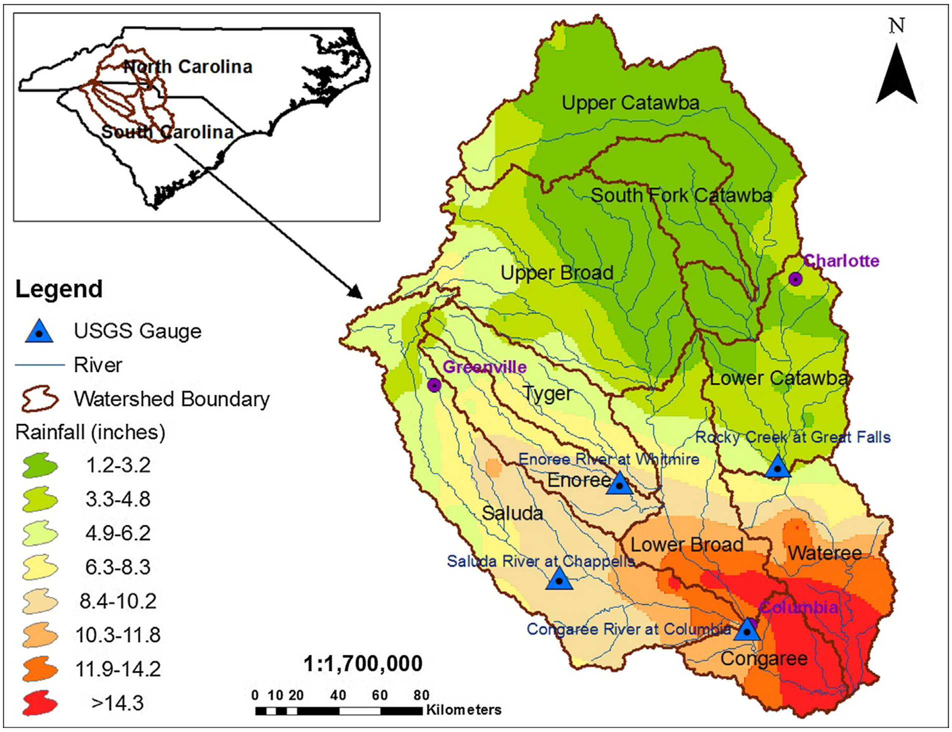

Fig. 1. The October 02–05, 2015 total rainfall (inches) in the study area along with four USGS monitoring gages: (i) Rocky Creek at Great Falls, SC, USGS 02147500; (ii) Enoree River at Whitmire, SC, USGS 02160700; (iii) Saluda River at Chappells, SC, USGS 02167000; and (iv) Congaree River at Columbia, SC, USGS 02169500.

Focusing on EVT and referring to Renard et al. (2013) and Martinkova (2013) for a recent review of the EVT applications in hydrology, this theory captures the asymptotic distributional behavior of two types of data, namely, the so-called block maxima (BM) and peak-over-threshold (POT). BM extracts the maximum values from subsets (i.e., blocks) of observations, whilst POT performs observations exceeding a certain threshold. When the size of the blocks approaches infinity, the distribution of BM converges to three types of extreme value distribution families (Gumbel, Fréchet, and reverse Weibull (Fisher and Tippett, 1928; Gnedenko, 1943)) where the parameters scale with the information dimension. These three types of extreme families can be described by the so-called generalized extreme value (GEV) distribution with the location, scale, and shape parameters (e.g., Coles, 2001) as defined by the unified von Mises-Jenkinson parameterization (Jenkinson, 1955).

If the threshold of exceedance increases, the GEV then converges to the so-called generalized Pareto (GP) distribution as described by the Pickands-Balkema-de Haan theorem (Pickands III, 1975; Balkema and de Haan, 1974). In many cases, GP yields a more accurate approximation to the distribution of absolute and relative excesses, as well as distribution tails. In addition, it represents distribution tails obliquely, but rigorously, by *"letting the data decide the function"*. In practice, a way to verify the validity of GP is to check whether the estimates of the shape parameter are stable when the model is fitted to excesses over a range of thresholds. From a theoretical point of view, absence of the stability can be explained by a slow rate of convergence in the Pickands–Balkema–de Haan theorem. The fitted model can then be used to compute any tail-related risk measure, such as tail probabilities, tail quantiles (or value-at-risk), etc. There is an established link between GP and GEV in the EVT modeling. In practice, if block maxima follow a GEV distribution, then the threshold excesses have a corresponding approximate distribution within the GP family (e.g., Coles, 2001) and vice versa GEV parameterization can be estimated using GP such as Poisson distribution for the occurrence frequency of the POT (e.g., Goda, 2011).

Recently, the probabilistic fitting of these extreme distributions to hydrological variables signifies major progress in design hydrology as it quantifies risk and disputes arbitrary notions (e.g., Koutsoyiannis, 2004). Although, in spite of the extensive literature on EVA model fitting and goodness-of-fit testing, only few studies have recently tackled the practical problems of flood frequency analysis facing real time application and uncertainty (e.g., Vogel et al., 2011; Stedinger and Griffis, 2011; Rootzén and Katz, 2013; Papalexiou and Koutsoyiannis, 2013; Obeysekera and Salas, 2014; Serinaldi and Kilsby, 2015; Mondal and Mujumdar, 2015). The application of extreme theory on various real-world applications is essential for risk assessment and water resources planning, which demand long time horizons with no other rational scientific basis than probability. Therefore, the aim of this paper is to compute FFA and return periods for annual and instantaneous floods in the center of the Carolinas with special attention to the POT approach and to place the October 2015 flood in a flood frequency analysis context. An important class of questions addressed in this study concerns the impact of peak rates and thresholds on the upper tails of flood distributions. The goal was to investigate the distribution fitting model and the upper- tail distribution of maxima and to provide a better answer to the question of "how extreme was the October 2015 flood in the Carolinas?"

To address aforementioned question, four different applications across the Carolinas were used to infer various procedures and to relate these analyses to properties of the October 2015 flooding. Spatial and temporal variability of flood events and the uncertainty associated with flood properties were also addressed during the period of analysis. The underlying parent distributions were also re-assessed with the inclusion of the 2015 flood event in order to characterize distributional changes associated with the fitting parameters. This study quantified the sampling uncertainty via confidence intervals (CIs) in the EVT framework to highlight its fundamental role for a fair comparison between models and a fair assessment of the output reliability.

This paper is organized as follows. In Section 2, the study region and flood data used in this study are explained. The theoretical concept and mathematical structures of probability distributions, distribution parameters, and POT are explained in Section 3. These methodologies were