Journal of Hydrology 530 (2015) 15–23

Contents lists available at ScienceDirect

Journal of Hydrology

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

A spatial model to examine rainfall extremes in Colorado's Front Range

Mari R. Tye [a,*], Daniel Cooley [b]

[a] National Center for Atmospheric Research, Boulder, CO, USA
[b] Department of Statistics, Colorado State University, Fort Collins, CO, USA

ARTICLE INFO

Article history:
Received 9 April 2015
Received in revised form 4 September 2015
Accepted 8 September 2015
Available online 15 September 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Ashish Sharma, Associate Editor

Keywords:
Extreme Value Analysis
Spatial hierarchical model
Gaussian process
Covariate-defined space

SUMMARY

Between 9th and 16th September 2013, northeast Colorado received some of its most extreme rainfall on record. The event affected 6 major rivers and their tributaries and 14 counties, breaking observed records for accumulations from sub-daily through to annual total. NOAA's rainfall atlases indicated that this event had an anticipated return period of 1000 years.

We use the rainfall that led to the 2013 Colorado floods as a case study in order to explore how a large event can affect the generalized extreme value (GEV) parameter estimates often used by designers and planners. We employ daily rainfall observations, with at least 30 years of data, from stations across Colorado's Front Range of the Rocky Mountains to develop a spatial statistical model for annual maximum daily rainfall. We produce estimates of relatively rare events such as the 1% Annual Exceedance Probability (AEP) level and of extremely rare events such as the return period associated with Boulder's 2013 observation. To explore sensitivity, we compare estimates including and excluding data from 2013, and both using only individual station data and our model which borrows strength across multiple stations. We compute the uncertainty associated with all of our estimates, and find large uncertainties associated with extremely rare events.

Our statistical model is a spatial hierarchical model and we employ a two-stage approach for inference which can be implemented by practitioners. Additionally, the spatial model allows us to interpolate spatially and estimate the GEV parameters at unobserved locations. A further development of the model makes use of an alternatively defined space in terms of elevation and a climate variable, rather than geographical space defined by longitude and latitude, which seems to better account for orographic effects.

In addition to producing AEP level and return period estimates to the annual maximum data, we investigate sensitivity to the choice of block length. We find point estimates indicate the tail to be much heavier when a longer block length is used, but the uncertainty associated with this parameter is such that one cannot say the difference is significant. To describe the spatial extent of severe storms, we also investigate the amount of data dependence between station locations. We find evidence in the record for storms with large spatial extent, although an extremal dependence parameter estimate indicates that this dependence is relatively weak.

1. Introduction

Between 9th and 16th September 2013, an area of northeast Colorado received some of its most extreme rainfall on record. Observational records were broken at many stations, most notably at Boulder where 230.6 mm (approx. 9.08″) fell in 24 h, approximately double the previous record from 1919 (Hamill, 2014; Lavers and Villarini, 2013). In total, 460.5 mm fell in Boulder over the course of 7 days, also breaking the monthly recorded total of May 1965 (Gochis et al., 2015). A State of Federal Emergency was declared in 18 counties along the Front Range (Office of The Press Secretary, 2013, Accessed February 2015), for the second most expensive natural disaster in Colorado after the 1965 floods (Lukas et al., 2014). The floods resulted in 8 fatalities; 18,000 people evacuated; 1500 houses destroyed and 19,000 damaged. Nearly 500 miles of state highways were damaged, isolating many of the northern Front Range mountain communities; 30 highway bridges destroyed and 20 damaged; 20 state dams damaged; and 150 miles of railway damaged (National Weather Service (NWS), 2014; Gochis et al., 2015; Kim et al., 2014). Rehabilitation is still taking place into 2015 (The City of Boulder, 2014). Annual Exceedance Probability (AEP) Estimates of the rainfall contributing to the flood were estimated from NOAA's Precipitation Frequency Atlas

* Corresponding author.
E-mail address: maritye@ucar.edu (M.R. Tye).

http://dx.doi.org/10.1016/j.jhydrol.2015.09.023
0022-1694/© 2015 The Authors. Published by Elsevier B.V.

(Perica et al., 2013), and range from around 10% AEP (10 year return period) for the hourly total to 0.1% AEP (1000 year return period) for the 24 h total rainfall at Boulder (National Weather Service (NWS), 2014).

While the prolonged rainfall, its spatial extent, and critical location over the headwaters of two major rivers was unusual (Gochis et al., 2015), it was not unprecedented with similar events occurring in September 1938, June 1965 and May 1969 (Lukas et al., 2014). Mahoney et al. (2014) analyzed the climatology of rainfall extremes (including snow, hail and graupel) across the Front Range, finding that while the event was similar in character to those associated with the North American Monsoon, it was later than the usual June-August timing. An event with very similar spatial extent occurred in the locality in September 1938 (Boulder Area Sustainability Information Network (BASIN), 2008; Lukas et al., 2014) and multiple climate simulations to recreate September 2013's conditions generated a few more intense cases (Hoerling et al., 2014), highlighting that this event may not be as rare as suggested. However, as noted by Mahoney et al. (2014), the lower density of observation records in population sparse areas and intermittent gauge readings prior to station automation mean that not all historic extremes are captured in the data record and thus cannot be included in a data-based analysis.

In this study we use the rainfall that led to the 2013 Northern Colorado flood as a case study to examine how the occurrence of an extreme event can affect estimation of the probabilities of rare events. To this end, we use annual maximum daily rainfall data to fit a spatial hierarchical extremes model which borrows strength across locations when producing parameter estimates, and which allows interpolation at unobserved locations. Our inference method for this model is relatively straightforward, and we believe it can be readily implemented by practitioners. The focus of this article is on the extremity of the 24 h rainfall total, and does not consider the other factors that affect a flood response, such as antecedent wetness conditions and exposure to the flood hazard. We use this model to compare estimates which include and exclude data from the 2013 event. To assess the influence of borrowing strength, we also compare estimates made at individual stations to ones produced by the spatial model. Additionally for all approaches, we quantify uncertainty associated with return level estimates and with estimates the rarity of September 2013's rainfall total. We assume temporal stationarity for this spatial model as this study is retrospective, thus we assume that the climate has been stationary enough during the observational record that a stationary model is adequate.

Drawing strength across space to improve estimates of annual exceedance probabilities is not a new concept. Regional Frequency Analysis (RFA) (e.g. Dalrymple, 1960; Institute of Hydrology, 1973) is an approach which pools data after it has been normalized by a local quantity often termed an 'index flood'. RFA has its roots in modeling stream flows, but the statistical approach is equally appropriate for rainfall (e.g. Jones et al., 2013), and NOAA currently uses a variant of the approach to produce official rainfall atlases (Bonnin et al., 2006). RFA pools data by pre-defining homogeneous regions, and the homogeneity of a region's data can be tested statistically (Hosking and Wallis, 1993; Stedinger et al., 1993). A related approach considers the Region of Influence (ROI) surrounding each station, rather than within a predefined region (Burn, 1990). In contrast to RFA and ROI, which do not explicitly build a spatial model, hierarchical modeling approaches have been applied to spatially model extreme behavior over a region (Cooley et al., 2007; Sang and Gelfand, 2009; Dyrrdal et al., 2015). Our approach is hierarchal, but unlike the aforementioned references which employ a Bayesian approach and obtain inference via Markov Chain Monte Carlo, we employ a two stage, non-Bayesian inference method. We believe that this inference approach may make hierarchical modeling more applicable to those practitioners who wish to try an alternative approach with the best available information. Both RFA and hierarchical approaches primarily aim to characterize how the *marginal* behavior of extreme rainfall varies over a study region. This differs fundamentally from approaches which aim to characterize spatial dependence in the data and which may use max-stable processes (Ribatet, 2013; Padoan et al., 2010) or copulas (e.g. Fuentes et al., 2013).

Although daily data tell only part of the story for Colorado's 2013 flood, daily precipitation totals are commonly analyzed, are regularly used for structural design purposes, and serve our purpose for this study. The various methods we explore could be applied to any the accumulation period of interest below 24 h. However, the model does not account for temporal dependence between stations experiencing a storm that transits across locations, e.g. for storms >24 h duration. Given our quality control criteria, we have only used annual daily rainfall maxima from 71 observation stations across the Front Range, with annual maxima that are verifiable against neighboring stations, or from contemporary news sources.

This article is laid out as follows. Section 2 describes the annual maximum rainfall data we analyze and also the available covariate information. Section 3 outlines the statistical approach, technical details of which are relegated to an appendix. Results are presented in Sections 3 and 4 discusses the further implications of the study.

2. Data

Daily precipitation time series were obtained from the Global Historical Climatology Network data set (Menne et al., 2012) for 71 stations across the Front Range. While a substantial data set of records from the volunteer community exists (e.g. the Community Collaborative Rain, Hail & Snow Network, CoCoRaHS), we opted not to include these to avoid the need for additional quality control to validate measurements or inconsistencies in observation timing (Mahoney et al., 2014). We selected stations that had been operational for at least 30 years, removing years where more than 5 days were missing in a month or 20 days during the year, and rejecting stations with fewer than 15 years remaining in their record. Where possible, annual maxima were verified against neighboring stations and other data sources (e.g. Colorado Climate Center, 2014) and included in the annual maxima series.

As we are interested in daily rainfall, the annual maxima were selected from days with daily minimum and maximum temperatures >5 °C and with no record of solid precipitation. The resultant data comprises 71 stations with a range of 18 to >120 annual maxima; station elevation, latitude and longitude.

In addition to the annual maximum data which we model, we employ a gridded mean seasonal rainfall product as a covariate which is also used to identify climatologically similar stations. Monthly "normals" for April–October (1981–2010) were obtained for a 4 km resolution grid over the study region from the Performance Reporting Information System archive (PRISM) (PRISM Climate Group at Oregon State University, 2014). The monthly "normals" represent the average seasonal conditions over the 30 year period, rather than anomalies from a baseline mean, and more appropriate for identifying climatological similarity. Importantly, this seasonal mean rainfall covariate as well as covariates of longitude, latitude, and elevation, are available not only at the station locations, but also every location at which we interpolate.

The study region is shown in Fig. 1. Although we only spatially interpolate on the indicted subregion, we use data from all the indicated stations, including the 5 stations which lie outside the subregion.