An Areal Based Approach for Improving Estimates of Extreme Precipitation Values

Peng, Gao, Greg Carbone, Diansheng Guo, and Junyu Lu

Department of Geography, University of South Carolina (gaop; carbone; guod@mailbox.sc.edu, jlu@email.sc.edu)

INTRODUCTION & OBJECTIVE

Understanding the probability of heavy rainfall events aids the design and operation of infrastructure that could mitigate losses from hydrologic hazards (Zhu, 2013). Estimating the probability of heavy precipitation faces two challenges.

Challenge One: Insufficient sample sizes. Regional Frequency Analysis (RFA) could solve the problem by "trading space for time". However, it is difficult to identify homogenous sets of stations that have similar extreme value characteristics (Willems et al. 2012). RFA also creates abrupt changes at the borders of adjacent regions.

Challenge Two: Areal representation of precipitation. NOAA's Atlas 14 provides point-based precipitation frequency estimates and uses Areal Reduction Factors (ARF) to convert point-based estimates to areal estimates. But ARF is limited to small areas and short rainfall durations.

This study provides precipitation frequency estimates based on the *total volume of precipitation* received in differently-sized United States Geological Survey (USGS) Hydrologic Units and addresses the two challenges by taking the advantage of high-resolution gridded precipitation data and adopting a bootstrapping approach that substitutes space for time.

DATA, HEAVY RAINFALL EVENTS & WATERSHEDS

Data: PRISM (Parameter-elevation Regressions on Independent Slopes Model) that provides the gridded daily precipitation with 4 km spatial resolution.

Table 1 Investigated heavy rainfall events and watersheds

Time	Location	Watersheds	Area (km^2)	Hydrologic Unit Codes
Oct. 2015	South Carolina	Gills Creek	193	0305011002
		Cooper	3,276	03050201
Mar. 2016	Texas & Louisiana	Bayou	1,275	11140208
Jun. 2016	West Virginia	Gauley	3,276	05050005

REFERENCES

Anderson, T. W., and D. A. Darling, 1954: A Test of Goodness of Fit. *J Am Stat Assoc*, **49,** 765-769.
Gao, P., G. J. Carbone, and D. Guo, 2015: Assessment of NARCCAP model in simulating rainfall extremes using a spatially constrained regionalization method. *Int J Climatol*.
Guo, D., 2008: Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *Int J Geogr Inf Sci*, **22,** 801-823.
Hosking, J. R. M., and J. R. Wallis, 2005: *Regional frequency analysis: an approach based on L-moments.* Cambridge University Press.
Kupfer, J. A., P. Gao, and D. Guo, 2012: Regionalization of forest pattern metrics for the continental United States using contiguity constrained clustering and partitioning. *Ecol Inform*, **9,** 11-18.
Pettitt, A. N., 1976: A two-sample Anderson-Darling rank statistic. *Biometrika*, **63,** 161-168.
Willems, P., 2012: *Impacts of climate change on rainfall extremes and urban drainage systems.* IWA Publishing.
Zhu, J., 2013: Impact of Climate Change on Extreme Rainfall across the United States. *J Hydrol Eng*, **18,** 1301-1309.

METHODS

The estimation of reoccurrence intervals of 4-day total rainfall depth in Gills Creek is chosen for illustration. Other watersheds and 1- and 2- day totals are reported in the results section. Total rainfall depth in a watershed is standardized by the number of PRISM grids within the watershed (i.e., the average rainfall depth of each grid) to facilitate comparison with observations at weather stations.

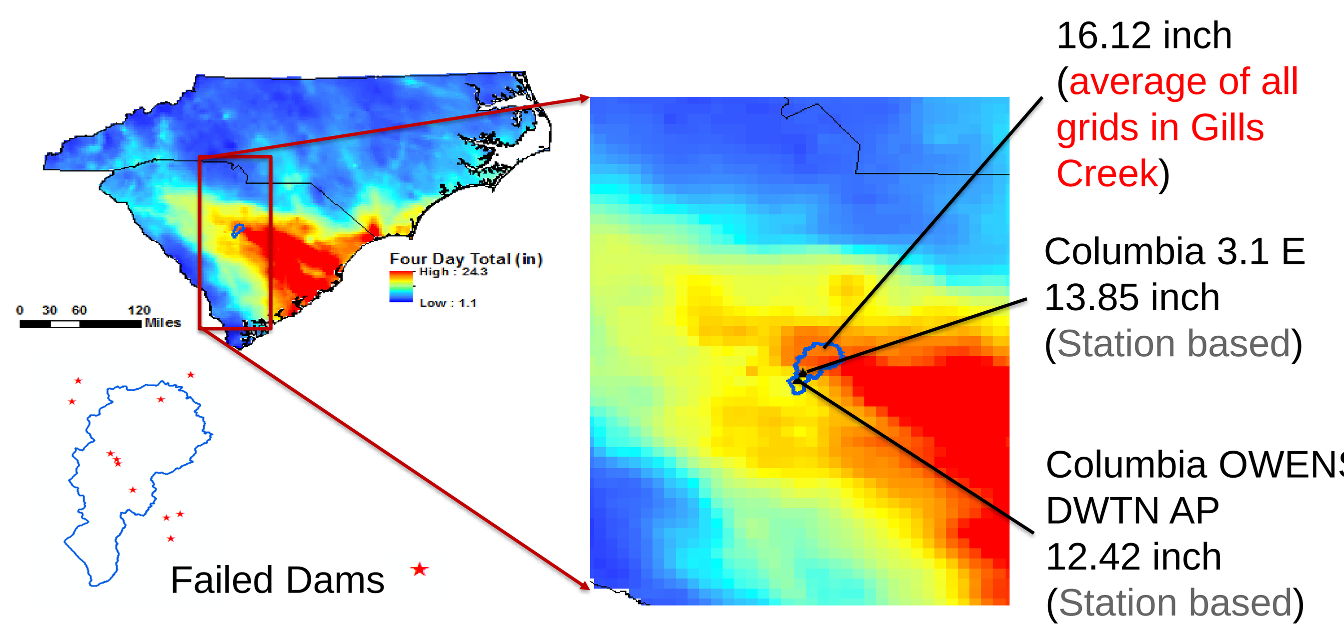

Figure 1 Four day total (2-5 Oct 2015) rainfall depth from PRISM and at stations within Gills Creek, South Carolina

Regionalization

REDCAP (Regionalization with Dynamically Constrained Agglomerative Clustering And Partitioning; Guo, 2008) was employed to delimit regions with relatively homogenous annual maximum 1-day, 2-day, and 4-day rainfall statistical properties separately from PRISM. For each grid in PRISM, annual maximum 1-day, 2-day, and 4-day rainfall totals were extracted. The dissimilarity of the pairwise grids was defined by the Anderson-Darling (AD) distance for annual maximum 1-day, 2-day, and 4-day rainfall totals. It disproportionately weights observations in the tails of the distribution (Anderson and Darling, 1954; Pettitt, 1976).

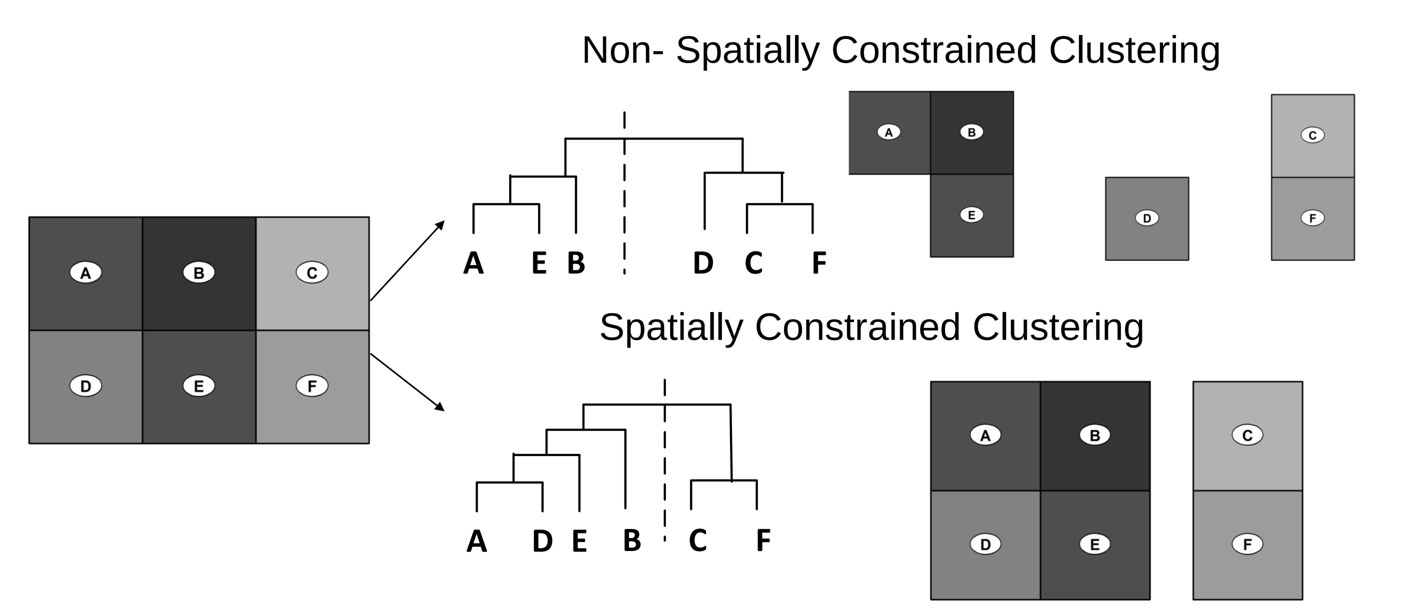

Figure 2 A hypothetical example of spatially constrained clustering (i.e. REDCAP). Contrasts of grey shading between grids represent the AD distance between them. A standard non-spatial method yields two clusters: a region that contains grids A, B, and E and a disjointed cluster that includes grids C, D, and F. Spatially constrained clustering requires that every cluster at each hierarchical level be spatially contiguous. In this example, it would create two regions which contain, respectively grids A, B, D, and E and grids C and F (Adapted from: Kupfer et al., 2012; Gao et al. 2015).

ACKNOWLEDGEMENTS

This research was partially supported by the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office (grant no. NA060AR4310007) to the Carolinas Integrated Sciences and Assessments (CISA) and partly based on work funded by NSF grant no. 0748813.

METHODS

Bootstrap Sampling

A layer that contains 100 non-overlapping polygons was generated in each year from 1981 to 2015, assuming these extreme events have relatively similar odds of occurrence anywhere within the region (Figure 3). The annual maximum of 1-day, 2-day, and 4-day totals were extracted from these randomly created polygons. In this way, additional samples were created to estimate GEV (Generalized Extreme Value) parameters. This method substitutes space for time, to account for the limited period of record.

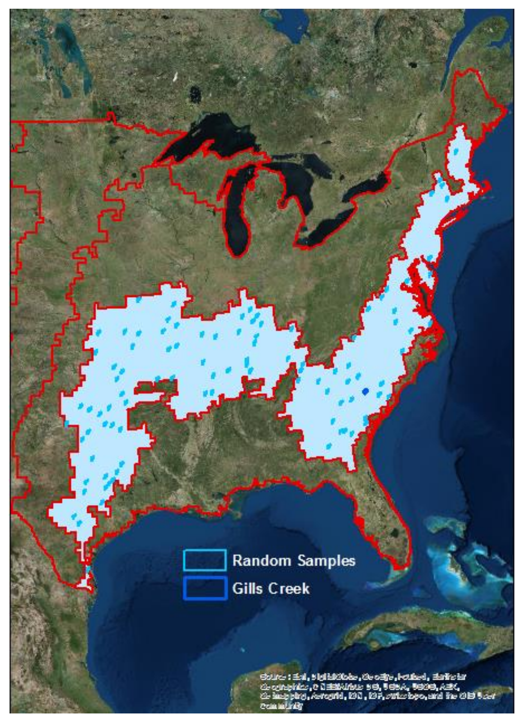

Figure 3 Bootstrap approach for Gills Creek in the region of four day total delineated by REDCAP

Probability Estimation

The annual maximum from the Gill Creek and random samples were used together to fit GEV curves. The location, scale, and shape parameters of the GEV distribution and the intensity of 2-, 5-, 10-, 25-, 50-, 100-, 200-, 400-, 600-, 800-, and 1000-year return periods were estimated using L-moments (Hosking and Wallis, 2005). To avoid bias that might be caused by randomly generating the layers, the procedure was repeated ten times by permutating the layers used for sampling in each year. GEV curves were fitted using annual maximum extracted from the randomly created datasets and the Gills Creek each time. The intensity at each reoccurrence interval was averaged across the ten sample sets.

RESULTS

Table 2 Return periods (years) at stations in Gills Creek and in the investigated watersheds

Stations	1-day	2-day	4-day	Reference Source
Columbia 3.1E	500 – 1000	500 – 1000	500 - 1000	Atlas 14
Columbia Owens Dwtn Ap	50 – 100	200 - 500	200 - 500	
Watersheds	**1-day**	**2-day**	**4-day**	**Approach**
Gills Creek	> 1000	> 1000	> 1000	Bootstrap
Cooper	> 1000	> 1000	In progress	
Bayou	80-90	≈ 400	≈ 500	
Gauley*		≈ 100		

* The event in Gauley was a 2-day event, so only 2-day total precipitation was investigated.

Summary

Our areal-based approach (bootstrapping) showed longer return periods for 1-, 2-, and 4- day totals than the point based method that compared observations at stations against estimates from Atlas 14, exceeding 1000 year (probability lower than 0.1%) in Gills Creek which had the highest number of failed dams during the heavy rainfall and flooding event in South Carolina in October 2015. Though Atlas 14 is an engineering standard, it is point-based and has limitations for areal estimation. Our approach more appropriately measured the severity of the event.