Recent trends in U.S. flood risk

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- An edited version of this paper was published by AGU. Copyright 2016 American Geophysical Union. To view the published open abstract, go to http://dx.doi.org/10.1002/2016GL071199

Metadata Record: https://dspace.lboro.ac.uk/2134/23983

Version: Published

Publisher: © American Geophysical Union. Published by Wiley

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Recent trends in U.S. flood risk

Louise J. Slater¹ and Gabriele Villarini¹

¹IIHR-Hydroscience and Engineering, University of Iowa, Iowa City, Iowa, USA

Abstract: Flooding is projected to become more frequent as warming temperatures amplify the atmosphere’s water holding capacity and increase the occurrence of extreme precipitation events. However, there is still little evidence of regional changes in flood risk across the USA. Here we present a novel approach assessing the trends in inundation frequency above the National Weather Service’s four flood level categories in 2042 catchments. Results reveal stark regional patterns of changing flood risk that are broadly consistent above the four flood categories. We show that these patterns are dependent on the overall wetness and potential water storage, with fundamental implications for water resources management, agriculture, insurance, navigation, ecology, and populations living in flood-affected areas. Our findings may assist in a better communication of changing flood patterns to a wider audience compared with the more traditional approach of stating trends in terms of discharge magnitudes and frequencies.

1. Introduction

In water year 2016 alone (1 October 2015 to 30 September 2016), the United States witnessed major floods spanning Missouri, the South and Midwest, Texas, Oklahoma, West Virginia, Maryland, and Louisiana. Many of these floods were catastrophic, reaching unprecedented flood levels (Figure 1), and affecting hundreds of thousands of people. Such flood-related disasters are of global concern with the increasing concentration of population in urban settlements [Jha et al., 2012] and the heightened exposure of assets to flood damage. In the USA, the cost of flooding is extensive, with total insured values in the National Flood Insurance program reaching almost $1.28 trillion for the period of 2001–2012 [Kousky and Michel-Kerjan, 2015].

The National Weather Service (NWS) uses four categories to communicate the severity of water surface elevation: action (requiring mitigation action in preparation for more substantial flooding), minor (with minimal or no property damage, but possibly some public threat), moderate (with some inundation of structures and roads and evacuations of people and/or property transfer to higher ground), and major (with extensive inundation, significant evacuations, or property transfer) [NOAA National Weather Service, 2012]. These categories are numeric thresholds that are determined at each gauging site (Figure 1) to monitor and forecast flood risk. We define the risk as the interaction of the flood hazard (probability) and the vulnerability [e.g., Mez et al., 2012] based on the four NWS flood categories. Thus, we use historical gage height (i.e., stage) records from the U.S. Geological Survey (USGS) to understand how the local flood risk is changing over time: in the case of Louisiana, the frequency of flood days has been progressively increasing above action, minor, and moderate flood categories (Figure 1a), while in South Carolina it has actually been decreasing (Figure 1b).

Changes in the frequency of flooding may occur through a combination of factors, including shifts in atmospheric conditions (e.g., magnitude, type, seasonality, and phase of precipitation), snowmelt patterns, antecedent soil moisture, land use and land cover (e.g., urbanization and agriculture), and anthropogenic modifications of the water cycle (e.g., management, extractions, dams, and diversions). Flood trends are typically investigated using discharge as a proxy for flood magnitudes [Milly et al., 2002, 2005; Vogel et al., 2011; Hirsch and Ryberg, 2012; Mallakpour and Villarini, 2015; Slater et al., 2015; Archfield et al., 2016]. However, classic trend-detection approaches using annual flow maxima [e.g., Mallakpour and Villarini, 2015] or regional flow averages [e.g., Archfield et al., 2016] may actually conceal any changes in small or local flood events. Additionally, any changes in a river channel’s capacity (i.e., the depth, width, and roughness) may significantly alter the frequency of local flooding above set flood levels (even in the absence of any changes in discharge) [Slater et al., 2015], so trends in flood frequency that are measured using discharge and gage height time series can be markedly different [Slater, 2016], especially in locations that have experienced major changes in channel capacity due to urbanization/regulation [e.g., Stover and Montgomery, 2001].
Here we specifically use only the most recently obtained NWS flood categories (and not changing flood levels) to investigate how the frequency of flooding is changing at those given levels over time, irrespective of any changes in the capacity of the river channel [Slater et al., 2015; Slater, 2016]. By tying gage height with the NWS flood categories ($GH_{\text{action}}$, $GH_{\text{minor}}$, $GH_{\text{moderate}}$, or $GH_{\text{major}}$) it becomes possible to investigate the practical impacts of changes in precipitation on flooding at the local scale, across the continental USA. Further, corrected stage measurements are inherently more precise for this purpose than the discharge estimates that are obtained from the stage-discharge rating relationships [Carter and Davidian, 1968], particularly during high flows.

2. Materials and Methods

2.1. Gage Height Data

We listed all of the USGS stream gages [United States Geological Survey, 2015] for which NWS flood categories were documented and downloaded all available daily gage height data (1985–2015) from the National Water Information System (NWIS-web). When little or no data were available online, we contacted the USGS Water Science Centers and received provisional gage height data (Table S1 in the supporting information) in instantaneous form (15 to 60 min increments), with the caveat that they might contain erroneous values. The majority of these sites are streams, with some lakes, reservoirs, and tidally influenced sites, and a mix of regulated and unregulated sites. We obtained gage descriptions from online USGS Water Year Reports to detect any changes in the gage location or datum, and thus discard any data preceding a change in gage location (Table S2), and/or correct any changes in datum.
by shifting the time series prior to the date of the change. Any apparent shifts in gage height minima or in the gage height-discharge relationship were investigated as potential changes in datum or gage/measurement location. Any extreme, visibly erroneous minima and maxima were removed, and sites with data judged unreliable were discarded (Table S2). In many states, electronic gage height records have only been stored since 2000/2001. Therefore, we carried out sensitivity tests with a minimum of 14, 15, 20, and 25 complete (300+ daily measurements) water years and find similar patterns irrespective of record length (Figure S1 in the supporting information). Water year 1985 was chosen as the starting date because only a minority of gage height records extend before then (Figure S2). While these records may seem short in comparison with some discharge records, our aim is not to discuss or detect any long-term climate-related changes in flooding but to show how rapidly the local flood risk can change above NWS categories.

2.2. Trends in Flood Frequency
To assess flood frequency, we follow a peak over threshold approach and compute the number of days per water year in which the mean daily gage height exceeded or equaled the GHaction, GHminor, GHmoderate, or GHmajor thresholds. At each site, trends in these quantities are detected using a Poisson regression model because the data are discrete and bounded at zero (Figure 1b). Trends are considered significant at the 5% level; however, sensitivity tests performed for $p \leq 0.10$, $p \leq 0.01$, and $p < 0.001$ reveal very similar spatial patterns (Figure S3). In many locations, the NWS flood thresholds are exceeded only a few times over the entire record, so sensitivity analyses were also performed to determine if the results changed based on the number of years with events exceeding threshold levels. The results were very similar for all tests (Figure S4), so a minimum of 3 years of nonzero events (above a given threshold) was required before fitting a trend (Figure 2). Additionally, not all sites have all four flood categories, and this may affect the spatial patterns of trends (Figures 2, S1, and S3–S5).

For comparison with trends in the number of annual days above flood categories, we also calculate trends in the frequency of flood events above each NWS category. The contributing drainage area $A$ (in logarithmically transformed m$^2$), available from the USGS site inventory for 99% of our sites, is used as a proxy for the number of days required to produce a peak flow in the basin (hereafter referred to as “n-day window”), following an approach similar to Lang et al. (1999) (n-day window < 5 days + log(A)). We use log(A) following United States Water Resources Council (1981) to allow longer recession times in larger basins. The trend in the number of independent annual flood events (separated by at least one n-day window) is also computed using Poisson regression (Figure S5).

2.3. Precipitation Data
We use precipitation data to assess whether changes in flood risk are related to rainfall extremes or long-term wetness over the study period. Total daily precipitation data at ~4 km resolution were obtained from the PRISM Climate Group (2015) and averaged over the contributing basin boundaries from USGS Streamgage NHDRPLUS Version 1 Basins 2006. Time series were computed for 88% of sites (i.e., those located in the conterminous USA and with NHDRPLUS basin boundaries) and tailored to fit the same daily timespan as the gage height data. $P_{\text{max}}$ is the annual maximum precipitation occurring in an n-day window for each site (we use n-day = log(A) because precipitation is computed at the basin scale). To find $P_{\text{max}}$, the time series is split into n-day overlapping windows (beginning on each day of the year), in which the total precipitation is aggregated; the window with the largest precipitation total is retained for each year. $P_{\text{annual}}$ reflects total precipitation in each water year. The nonparametric Mann-Kendall test (Mann, 1945) is used to identify monotonic temporal trends in $P_{\text{max}}$ and $P_{\text{annual}}$ based on the Kendall rank correlation (accounting for any ties in the data) (McLeod, 2011). The association between precipitation and the frequency of exceedance of flood stage at each site is quantified by fitting a Poisson regression model between the annual time series of $P_{\text{annual}}$ (predictor) and that of GHminor (discrete response variable) if thresholds are exceeded at least 3 times in the entire annual time series (see sensitivity tests for exceedance in Figure S4).

2.4. Gravity Recovery and Climate Experiment Satellite Data
NASA’s Gravity Recovery and Climate Experiment (GRACE) satellites have been measuring Earth’s gravity field anomalies since 2002, providing information on both event flow (precipitation-driven) and base flow (stored
Thus, although GRACE time series are still relatively limited, the data may prove valuable for predicting the exceedance of NWS flood levels. We use the GRACE global monthly mass concentration blocks (mascons) version RL05m from the Jet Propulsions Laboratory, which have improved signal recovery in comparison with other processing methods (such as the spherical harmonic approach). The data were downloaded from https://podaac.jpl.nasa.gov/dataset/TELLUS_GRACE_MASCON_GRID_RL05_V1 and are described online at http://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/. These are 3° mascons gridded to a 0.5° grid; they have not been smoothed and thus are "blocky" in appearance. In Figure 2 we plot the linear trend using all available data (2002–2015), converted to cm/yr. For Figure 3, we compute the time series of basin-averaged GRACE data and extract the mean values in each water year (GRACEannual), requiring a minimum of 10 monthly values in each water year because of data incompleteness. For Figure 4, we require a minimum of nine complete water years to measure the strength of the association between GHminor and GRACEannual using Poisson regression.

2.5. Time-Varying Predictors of Flood Occurrence

To model the occurrence of individual flood events exceeding GHminor as a function of antecedent wetness and precipitation, we use Cox regression [Cox, 1972]. Cox regression can be used to model
Figure 3. Trends in precipitation and water storage. Maps indicate trends computed using the Mann-Kendall test in basin-averaged (a) $P_{\text{max}}$, (b) $P_{\text{annual}}$, and (c) GRACE$_{\text{annual}}$ (water years). The colors and symbology are the same as in Figure 2. Alaska, Hawaii, and Puerto Rico are not included because of the unavailability of the precipitation data for those locations.
Cox processes, which are a generalization of Poisson processes. While in a Poisson process all the events are considered independent, Cox processes are clustered, so bursts of activity alternate with more quiet periods. For more information and details about Cox processes and regression with time-dependent predictors, consult Cox [1972], Karr [1983], Smith and Karr [1983], Villarini et al. [2013a], and Section S1 in the supporting information.

Here we consider two covariates: short-term precipitation prior to a flood event ($P_{\text{event}}$) and long-term precipitation ($P_{\text{annual}}$, the sum of all precipitation occurring in the 365 days preceding $P_{\text{event}}$). To compute $P_{\text{event}}$, we first loop through all $GH_{\text{minor}}$ flood events and calculate the mean number of days that gage height stayed above flood stage before the flood peak ($n_{\text{days}}$, rounded up to the nearest whole integer). $P_{\text{event}}$ is the sum of all precipitation occurring in a window of $n_{\text{days}}$ plus 10 days preceding the flood event (different window sizes were tested, and we chose the largest value to best distinguish $P_{\text{event}}$ from $P_{\text{annual}}$; see Figures S6–S8). $P_{\text{annual}}$ is summed for each water year in Figures 4a–4j and computed prior to each event in Figures 4k–4o for the Cox modeling. The green colors indicate flood frequency trends computed using the number of flood days (first row) and number of flood events (second row) above $GH_{\text{minor}}$ every year. Fraction of sites where $P_{\text{event}}$ (blue) and $P_{\text{annual}}$ (magenta) are significant predictors of $GH_{\text{minor}}$, flood events using Cox regression (Figures 4k–4o). The numerical x axis variables were binned at approximately the 25th and 75th percentiles of the distributions.

Cox processes, which are a generalization of Poisson processes. While in a Poisson process all the events are considered independent, Cox processes are clustered, so bursts of activity alternate with more quiet periods. For more information and details about Cox processes and regression with time-dependent predictors, consult Cox [1972], Karr [1983], Smith and Karr [1983], Villarini et al. [2013a], and Section S1 in the supporting information. Here we consider two covariates: short-term precipitation prior to a flood event ($P_{\text{event}}$) and long-term precipitation ($P_{\text{annual}}$, the sum of all precipitation occurring in the 365 days preceding $P_{\text{event}}$). To compute $P_{\text{event}}$, we first loop through all $GH_{\text{minor}}$ flood events and calculate the mean number of days that gage height stayed above flood stage before the flood peak ($n_{\text{days}}$, rounded up to the nearest whole integer). $P_{\text{event}}$ is the sum of all precipitation occurring in a window of $n_{\text{days}}$ plus 10 days preceding the flood event (different window sizes were tested, and we chose the largest value to best distinguish $P_{\text{event}}$ from $P_{\text{annual}}$; see Figures S6–S8). $P_{\text{annual}}$ is representative of the overall wetness of the watershed.

Thus, at the resulting 2042 locations (949 published, 1093 provisional) we investigate (i) how the frequency of floods is evolving above the NWS’s four flood categories (Figures 1 and 2), (ii) how basin-averaged precipitation and water storage are changing at the same stream gages (Figure 3), and (iii) the factors affecting the trends in flood risk as well as the relationship between gage height and precipitation/antecedent basin wetness (Figure 4).
3. Results and Conclusions

3.1. Spatial Patterns of Trends in Flood Risk

Trends in the frequency of flooding above the four NWS flood thresholds computed using Poisson regression reveal a stark contrast between increases in flood risk around the upper Midwest/Great Lakes region and decreases on the Gulf Coastal Plain, the southeastern United States, and California (Figure 2). Broadly speaking, these patterns are similar to the changes in liquid water equivalent measured by NASA’s GRACE mission (GRACEannual, used here as a proxy for overall basin wetness); two-thirds of GHminor and GRACEannual trends exhibit the same sign (Figure S9). This cooccurrence of positive/negative trends in flooding and basin water storage suggests that progressive changes in basin wetness arising from combined climatic and anthropogenic influences precondition the local flood potential [see Reager et al., 2014].

Localized increases in flooding in the Gulf of Mexico, near New Orleans, and selected locations in Florida are also visibly in agreement with GRACE. Flood trends in the western half of the conterminous USA, Alaska, and Puerto Rico show mixed patterns of increases and decreases, while California stands out with almost uniform decreases in flood frequency (Figure 2).

The regional patterns of increasing and decreasing flood risk are broadly similar for all four flood categories but become less similar as the difference between flood levels increases (83% of GHaction and GHminor trends have the same sign, 73% of GHaction/GHmoderate and 66% of GHaction/GHmajor). For example, a site may have a significant increasing trend in GHaction, but a nonsignificant decreasing trend in GHmajor (see the southernmost sites in Florida). As the moderate and major flood categories are exceeded less frequently than action and minor levels, significant changes are detected at fewer sites (Figure 2), and longer records may be required before trends become significant. We also find a similar spatial distribution of increases versus decreases in flood risk, whether trends are computed using the number of independent annual flood events or with the number of flood days (Figures 2 and S5), and results are consistent across the published and the provisional gage height records (Figure 2). While there may be local differences in river systems that have experienced anthropogenic modifications, the changes in flood frequency that are measured using gage height are largely consistent with previously observed increases in the frequency and magnitude of streamflow across the Midwest [Mallakpour and Villarini, 2015; Slater et al., 2015], the North, and North-East [Milly et al., 2005; Collins, 2009; Slater et al., 2015], as well as decreases in runoff and flood magnitude in the West and Southwest [Milly et al., 2005; Hirsch and Ryberg, 2012] and coastal southeastern regions [Hirsch and Ryberg, 2012] of the United States.

3.2. The Relationship Between Changing Precipitation, Basin Wetness, and Flood Risk

Can these trends in flood risk be related to short-term changes in precipitation? While extreme flooding is mostly driven by very heavy rainfall [e.g., Groisman et al., 2001], flood flows often have multiple drivers [Ivancic and Shaw, 2015; Stephens et al., 2015; Berghuijs et al., 2016]; thus, we investigate the relationship between precipitation and stage, and the variables that may modulate it. Using daily precipitation data [PRISM Climate Group, 2015] averaged over each of the contributing basins for the same time periods as the daily gage height data, we computed trends in $P_{\text{max}}$ (the maximum annual precipitation occurring in an n-day flood-inducing moving window; see section 2) and $P_{\text{annual}}$ (the total annual precipitation) using the Mann-Kendall test. Spatial trends in $P_{\text{annual}}$ and $P_{\text{max}}$ are analogous to longer-term increases in extreme precipitation that have been measured over Central/Midwest/East [Janssen et al., 2014; Villarini et al., 2013b; Mallakpour and Villarini, 2015], Northeast, and Great Lakes [DeGaetano, 2009; Horton et al., 2014] regions, pointing to a generally heightened frequency of high-intensity precipitation days and events [Groisman et al., 2005, 2012; Guilbert et al., 2015]. Decreasing trends are more prominent in the southeastern coastal states of Texas, the Carolinas, and Georgia, and especially over California, where severe water shortages [Diffenbaugh et al., 2015] have resulted in the most severe drought that the U.S. West Coast has seen in centuries (Figures 3a and 3b).

While spatial precipitation trends broadly reflect changing patterns of flood risk, even more notable are the changes in basin-averaged water storage (Figure 3c). Using the same methods as for precipitation, but with shorter time series, we find that increases/decreases in GRACEannual are spatially consistent with $P_{\text{max}}$ and $P_{\text{annual}}$ and the distribution of trends in the frequency of flooding (Figures 2, 3, and S9). To quantify this association between basin wetness and flooding, Poisson regression models were fit between the yearly time series of total days exceeding minor flood stage (GHminor) and those of $P_{\text{annual}}$ and GRACEannual at each site.
The fraction of sites where $GH_{\text{minor}}$ flood events are significantly related to $P_{\text{annual}}$ is greater than with $GRACE_{\text{annual}}$ due to longer time series. However, the measured relationships between each of the predictors ($P_{\text{annual}}$ and $GRACE_{\text{annual}}$) and $GH_{\text{minor}}$ are remarkably similar (Figures 4a–4i) and suggest that broad shifts in basin wetness that occur at the land surface and subsurface over annual time scales (due to climate and human activity) are an important driver of flooding.

If the strength of the relationship between basin wetness (as measured by $P_{\text{annual}}$ and GRACE annual) and $GH_{\text{minor}}$ (or P-GH coupling) is indeed due to the influence of water storage within the catchment, then we would expect it to be affected by the physical properties that control water retention [Jiancic and Shaw, 2015; Stephens et al., 2015]. Significant P-GH coupling increases markedly in large and low-lying catchments, where there is better integration of flow paths reaching the channel (Figures 4a, 4b, 4f, and 4g). Locally, P-GH coupling is muted by the presence of impervious land surfaces which disconnect the channel from its floodplain (Figures 4c and 4h); however, at the catchment scale, anthropogenic alterations to water storage, which modify/sustain specific flow pathways, appear to have amplified the climate signal in river flows in comparison with less-modified sites (Figures 4d, 4i, and 4n) [as in Vogel et al. [2011]]. Last, we find that the strongest P-GH coupling occurs when climates are neither too wet nor too dry (Figures 4e, 4j, and 4o).

To further assess the dependence of flooding on basin wetness, we modeled the occurrence of individual $GH_{\text{minor}}$ flood events as a function of the precipitation occurring immediately before the flood ($P_{\text{event}}$) and of the annual wetness ($P_{\text{annual}}$) using Cox regression (Figures 4k–4o) [e.g., Villarini et al., 2013a]. While $P_{\text{event}}$ tends to be a significant predictor more than half of the time, and across all types of sites, $P_{\text{annual}}$ is also an important predictor, especially in large or dry catchments with unurbanized riparian areas (Figures 4, 57, and 58). In other terms, locations with greater potential water storage are more sensitive to year-to-year fluctuations in basin wetness, and it is to be expected that the inclusion of the $P_{\text{annual}}$ predictor will substantially improve the prediction of flooding in these locations.

In sum, in a warming world where changes in the spatial and temporal distributions of precipitation and scale climate indices are affecting the magnitude and frequency of high flows [Grosman et al., 2001; Karl et al., 2009; Mallekpour and Villarini, 2015, 2016], our findings reveal that there are actually strong regional patterns of changing flood risk which can be assessed and communicated from a practical standpoint in terms of the local threat to people and assets. These regional patterns are preconditioned by overall basin wetness, especially in low-lying areas with notable water storage. Thus, any projections of changes in flood risk above the action, minor, moderate, and major flood categories based on the observed amplification of precipitation extremes should most certainly take into account the short-term, concomitant changes in basin wetness resulting from broad-scale shifts in climate and water management.

Acknowledgments
We are extremely grateful to the USGS for collecting and providing us with the gage height data (see Table S1). This work would not have been possible without their efforts. This material is based upon work supported by the Broad Agency Announcement Program and the Engineer Research and Development Center—Cold Regions Research and Engineering Laboratory under Contract W9125E-16-C-0002, IHRR-Hydroscience and Engineering, the Iowa Flood Center, the USACE Institute for Water Resources (G.V.), and the National Science Foundation under CAREER Grant AGS-1349827 (G.V.). We gratefully acknowledge the suggestions and comments by Rhawn Denniston (Cornell College) and those made by anonymous reviewers. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.J.S. (louise-slater@uiowa.edu).

References
Mallakpour, I., and G. Villarini (2015), The changing nature of
Lang, M., T. B. M. J. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling,
Villarini, G., J. A. Smith, R. Vitolo, and D. B. Stephenson (2013a), On the temporal clustering of US
Villarini, G., J. A. Smith, and G. A. Vecchi (2013b), Changing frequency of heavy rainfall over the central United States,
Vogel, R. M., C. Yairdi, and M. Walter (2011), Nonstationarity: Flood magnification and recurrence reduction factors in the United States,
teleconnection patterns,
PRISM Climate Group (2015), Oregon State Univ. [Available at http://prism.oregonstate.edu.]
Reager, J. T., B. F. Thomas, and J. S. Famiglietti (2014), River basin flood potential inferred using GRACE gravity observations at several months lead time, Nat. Geosci., 7(8), 588–592, doi:10.1038/ngeo2203.