Introducing Flashiness-Intensity-Duration-Frequency (F-IDF): A New Metric to Quantify Flash Flood Intensity

Zhi Li[1], Shang Gao[2], Mengye Chen[1], Jiaqi Zhang[1], Jonathan J. Gourley[2], Yixin Wen[3], Tiantian Yang[1], and Yang Hong[1*]

[1]School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, OK, USA

[2]NOAA/National Severe Storms Laboratory, Norman, OK, USA

[3]Department of Geography, University of Florida, Gainesville, FL, USA.

Corresponding author: Yang Hong (yanghong@ou.edu)

Key Points:

- We introduce the Flashiness-Intensity-Duration-Frequency curve to quantify flash flood intensity
- The CONUS-wide Flashiness-Intensity-Duration-Frequency values are provided at 3,722 stream gage locations
- The relations between 59 basin attributes and flashiness values are explored
- The Flashiness-Intensity-Duration-Frequency curves have implications for hydrologic modelers, decision-makers, and emergency responders

Abstract (150 words)

Flash flooding is one of the most damaging weather types, yet it remains challenging to quantify its severity. We propose a novel development – the Flashiness-Intensity-Duration-Frequency (F-IDF) curve – to quantify and spatially analyze flash flood intensity based on the frequency and duration of the event. As a proof-of-concept, we mapped Contiguous US (CONUS)-wide F-IDF values at 3,722 stream gage locations and explored their relations with 59 basin attributes. It is found that (1) Climatological precipitation amounts exhibit the most positive correlation with flashiness while a basin's drainage area is the most negatively correlated; (2) Correlation of flashiness with basin attributes decreases with increasing F-IDF return periods and shorter event durations. Both aspects are attributable to the rainfall signal overwhelming the underlying basin attributes as the intensities become more extreme. This new term can have implications for hydrology, especially for hydrologic modelers, decision-makers, and emergency responders.

Plain Language Summary

Flash floods are among the most devasting natural hazard types that can cause severe property damage and loss of life. However, it's challenging to measure and quantify the severity. This study proposes a new way of quantifying flash flood intensity using a newly developed Flashiness-Intensity-Duration-Frequency (F-IDF) curve. It links flash flood severity with how often they happen and how long they last. We mapped F-IDF values across the United States and found that certain areas are more prone to flash floods than others. The amount of rain and the size of the basin area are the most important factors in determining how severe a flash flood is. This new quantification tool can help experts better identify and respond to flash flood risks.

1 Introduction

Flash floods, by definition, are a type of flood that occur within minutes to several hours of heavy rainfall or other causes (Doswell III, 2015; Gourley et al., 2013; Hong et al., 2013). In recent years, fatalities and damage caused by flash flooding have been increasing worldwide, making it one of the most destructive weather types (Ashley & Ashley, 2008).

To identify flash flood risks, researchers have sought various approaches. One of the most common practices for flash flood warning over the US and the world is the Flash Flood Guidance (FFG) methodology (Georgakakos et al., 2022). It has been adopted as the operational early-warning systems for flash flooding by the US National Weather Service since the 1970s (Georgakakos, 1986). FFG is defined as an estimate of total rainfall that causes bankful flow. As it suggests, this method does not take into account the full contiuum of land surface responses to extreme rainfall and river routing processes. Beyond FFG, there are other attempts to quantify flash flood risks. We generalize them into event-dependent and event-independent approaches. An event-dependent approach directly calculates flash flood risks based on archived flash flood events (Alipour et al., 2020) or a flashiness index (Gannon et al., 2022; Li et al., 2022; Saharia et al., 2017, 2021; Smith & Smith, 2015). The term flashiness index was introduced to measure how quickly and how high streamflow rises in response to an event (Baker et al., 2004). Among variants of flashiness index, the Richards-Baker Flashiness Index (RBI) is one of the earliest indices, denoted by the time derivative of daily streamflow (Baker et al., 2004). Gannon et al. (2022) evaluated the RBI at daily time scales and found little or no correspondence between basin responses and watershed area. This result differs with Saharia et al. (2017) who revealed a significant relationship of increasing flashiness with smaller watersheds, with the discrepancy being attributed to the latter study's use of sub-hourly stream gage data instead of daily. Since it is event-dependent, this approach presumably delivers accurate and precise results. However, it is heavily based on a dense observational network. Alternatively, an event-independent approach seeks a statistical model that relates climate variables and basin physiography to flash flood risk (Lin et al., 2020; Ma et al., 2019). In doing so, this approach bypasses the requirement for observations, which is particularly useful in ungauged basins or rural regions. Its validity, however, requires particular attention.

Given the dense stream gage network in the US, we propose a new method using the flashiness index applied to specific events. Although the definition of flashiness is diverse, this study adopts the approach of estimating the slope of the rising limb of the hydrograph to reflect the flood rising rate (Baker, 2004; Li et al., 2022; Saharia et al., 2017; Smith & Smith, 2015).

The flashiness index used in previous studies is only a static quantity that is irrespective of event frequency and duration (Li et al., 2022; Saharia et al., 2017, 2021; Smith & Smith, 2015). Weather forecasters, emergency responders, and the public are particularly concerned about the degree of severity of a flash flood event, which needs to be quantified by frequency. Additionally, we particularly value the representativeness of this index with respect to simplicity and reproducibility. In light of these concerns, we adopt the idea from the Rainfall Intensity-Duration-Frequency (R-IDF) curve that encapsulates three-dimensional information of a rainfall event (Perica et al. 2013), and apply it to quantify a flash flood event. Hence, we introduce the Flashiness-Intensity-Duration-Frequency (F-IDF) curve for the first time. Similar to the R-IDF curve, the F-IDF curve describes the intensity (based on flashiness values), duration, and frequency of flash flood events. We envision such a measure has practical implications in flash flood forecasting and risk management. The aim of this article is threefold: (1) introducing the F-IDF curve; (2) mapping F-IDF values for all US stream gages; and (3) investigating geographical and hydrometeorological factors associated with F-IDF values. The newly introduced F-IDF curve can be applied to observed or simulated hydrographs, meaning that it can be integrated into any flood forecast system. We discuss how this new method can benefit hydrologic science, hydrologic modelers, emergency responders, and city planners.

2 Materials and Methods

2.1 Flashiness-Intensity-Duration-Frequency

The F-IDF curves in this study are computed as follows: (1) Find the maximum rising (positive) slope S of a hydrograph using a recursive moving time window (i.e., D=1 hour, 2 hours, 3 hours, 4 hours, 5 hours, and 6 hours) over the available period of streamflow record; (2) Extract the annual maxima for each duration D; (3) Fit the annual maxima into a general extreme value distribution (GEV) and logPearson Type III distribution (LP3); (4) Find an optimal fit based on the Bayesian Information Criterion; and (5) Return flashiness values for different frequencies (i.e., 1-in-2-years, 1-in-5-years, 1-in-10-years, 1-in-25-years, 1-in-50-years, and 1-in-

100-years). The resulting flashiness value F is a measure of rapidness and magnitude changes over the time window and is represented in Eq.1. An illustrative example is given in Fig. 1a.

$$F = \frac{\max\{O_t - O_{t-1}, O_t - O_{t-2}, \dots O_t - O_{t-d}\}}{FAC \times d}, \tag{1}$$

where O_t is the observed streamflow time series at time t, d is the duration, FAC is the drainage area (km^2). The unit of F is dependent on the observation but is generally expressed in units of [L/T^2]. We standardize the unit of flashiness value to be measured in mm/h^2. In this study, we use the USGS stream gage record at a 15-minute time interval, so a conversion factor 0.4078 is applied to convert $ft^3/s/km^2/15\text{-min}$ to mm/h^2.

Repeating the process of calculating flashiness values at different durations and different frequencies, we can depict the F-IDF curve as shown in Fig. 1b for one site. The shape of the F-IDF curve is similar to the rainfall IDF curve, where intensity decreases with longer duration but increases with event rarity.

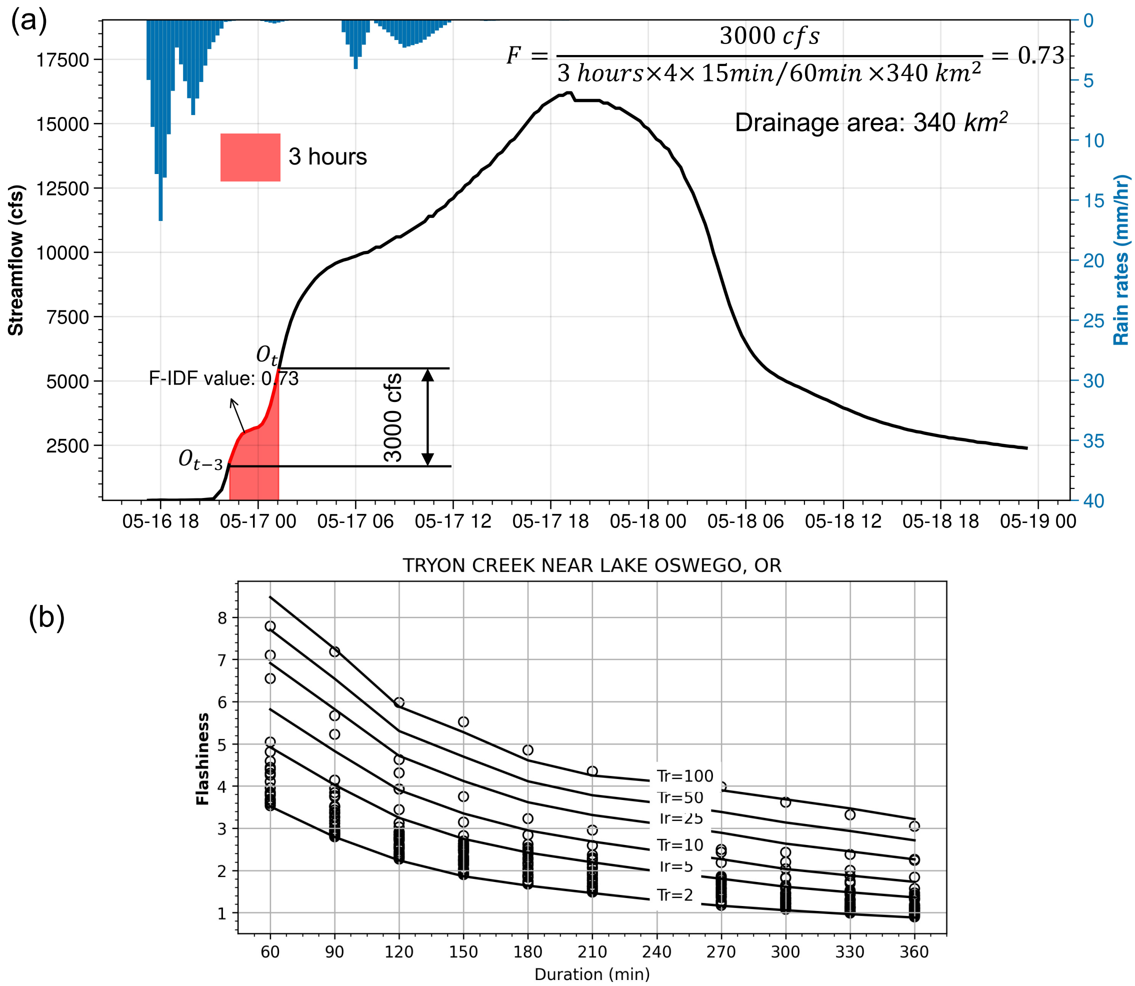

Figure 1. (a) An illustrative example of calculating Flashiness-Intensity-Duration-Frequency values. The figure is produced with the Python Matplotlib library; (b) The empirical F-IDF plot and points are real events that surpass 2-year flashiness values.

There are several noteworthy points in calculating F-IDF values. First, because flash floods typically occur within 6 hours of the causative rainfall (Li et al., 2022), we did not consider events with durations greater than six hours. Second, we select two extreme value distributions in this study: (1) LP3 distribution and (2) GEV distribution. The LP3 distribution is a common approach in hydrologic frequency analysis, recommended by the US Water Resources Council (Singh, 1998). The GEV is an alternative approach that harmonizes the type I, type II, and type III extreme value distributions into a single family to allow a continuous range of possible shapes. Wallis & Wood (1985) compared two methods and found the goodness-of-fit for the two methods varied across different sites. Third, given the short gage record length (22.3

years), we only extrapolate return periods to 100 years; otherwise, there are large uncertainties associated with the fitted GEV model (details refer to Section 3.1).

3 Data

3.1 CONUS-wide streamflow

We intended to collect 15-min streamflow time series data for all stream gages over the CONUS from 1950 to 2020. However, not all gauge sites have such data length, especially for sub-hourly instantaneous values. A map of stream gage data length distribution is shown in Fig. S1. We filter out gages that have available data of less than 20 years to ensure enough data samples for fitting the extreme value distributions. There are 3,722 gages left after filtering. Next, we harmonize an equal time interval of 15 minutes for all stream gages by using linear interpolation because some gages have an interval of 30 minutes. The linear interpolation method is often used to fill in gaps in streamflow data (Pestrone et al., 2010). After preprocessing, those data are analysis-ready to feed into the pipeline described in Section 2.1.

3.2 Catchment attributes

To analyze the flashiness values with basin characteristics, we use the basin attributes from the HydroATLAS dataset (Linke et al., 2019). These attributes include eight sections: Hydrology (i.e., annual runoff, precipitation, natural discharge, inundation extent, groundwater table, river area, and river volume), Physiography (i.e., channel slope, catchment slope, elevation, and drainage area), Climate (i.e., annual precipitation, potential evaporation, actual evaporation, climate moisture index, aridity index, air temperature, snow cover), Soils & Geology (i.e., soil water content, clay fraction, silt fraction, sand fraction, karst fraction, soil erosion), Human (i.e., road density, urban density, population), Land Cover (i.e., area extent of trees, shrubs, herbaceous, cultivated land, water bodies, snow, and artificial lands), Natural Vegetation (i.e., evergreen, deciduous, savanna, grassland, tundra, desert), and Wetland (i.e., lake reservoir, river, and peatland). There are 59 basin attributes in total used in this study. We spatially join these attributes to the catchments of all stream gages and use the values representing the total watershed upstream of the gage. A detailed description of these attributes is provided in Linke et al. (2019).

4 Results

4.1 Mapping CONUS-wide F-IDF values

After iterating through steps 1-5 in Section 2.1 for each stream gage, we can map the CONUS-wide F-IDF values. Figure 2 shows the one-hour flashiness values at six return periods (2-year, 5-year, 10-year, 25-year, 50-year, and 100-year) as an example. Maps for other durations (i.e., 2-hour, 3-hour, 4-hour, 5-hour, and 6-hour) can be found in Figs. S2-6 in the Supplementary File. **A general observation for these maps as indicated in Fig. 1b is that F-IDF values decrease with frequency and duration, in a similar manner as with R-IDF values**. We can easily identify flashy regions in the CONUS by clustering stream gages that have flashiness values larger than 1 (shown in Fig. 2b). Those five regions are (1) West Coast, (2) Missouri Valley, (3) the Appalachians, (4) Flash Flood Alley, and (5) Southwest. The results agree well with Saharia et al. (2017) and Li et al. (2022), despite slight differences in defining the flashiness variable. We also compared our results with real flash flood events from 1970 to 2020 in a newly developed US flood database (Fig. S7; Li et al., 2021). These flash flood events were verified by the US National Weather Service. Our identified regions also emerge, except for the Pacific Northwest region, which has a low incidence of flash flood reports. A similar finding is reached by Smith & Smith (2015), who reported the differences are in nature due to different measures.

The main drivers for flash floods are region-dependent. On the West Coast, the main atmospheric agent for flash flooding is atmospheric rivers, which transport considerable moisture from the tropics to mid-latitudes. Even though atmospheric rivers produce long-duration winter rainfall and snowfall, the steeply sloped terrain and compact watersheds can generate fast-rising runoff (Saharia et al., 2017; Smith & Smith, 2015). Further inland, the contributions of warm-season thunderstorms to flash flood occurrences start to dominate, especially for the Missouri Valley (Region 2) and Flash Flood Alley region (Region 4). The destructive flash floods in 2022 in these two regions were the result of training thunderstorms that produced several record-setting flood events. Flash Flood Alley also bears frequent tropical cyclones and hurricanes off the Gulf Coast. The Appalachians (Region 3) are another known hot spot for flash flooding, extending from Georgia up to Maine. Besides the hilly terrain, extratropical cyclones are the synoptic weather types that frequently hit this region and result in a sequence of flood events (Li et al., 2021). The Southwest (Region 5) is renowned for its hot and dry environment that initiates convective thunderstorms during the North American monsoon season (Smith et al., 2019). Besides the atmospheric forcings, land surface conditions such as impervious area ratio, antecedent soil moisture, groundwater level, catchment drainage density, etc., jointly determine flash flood severity.

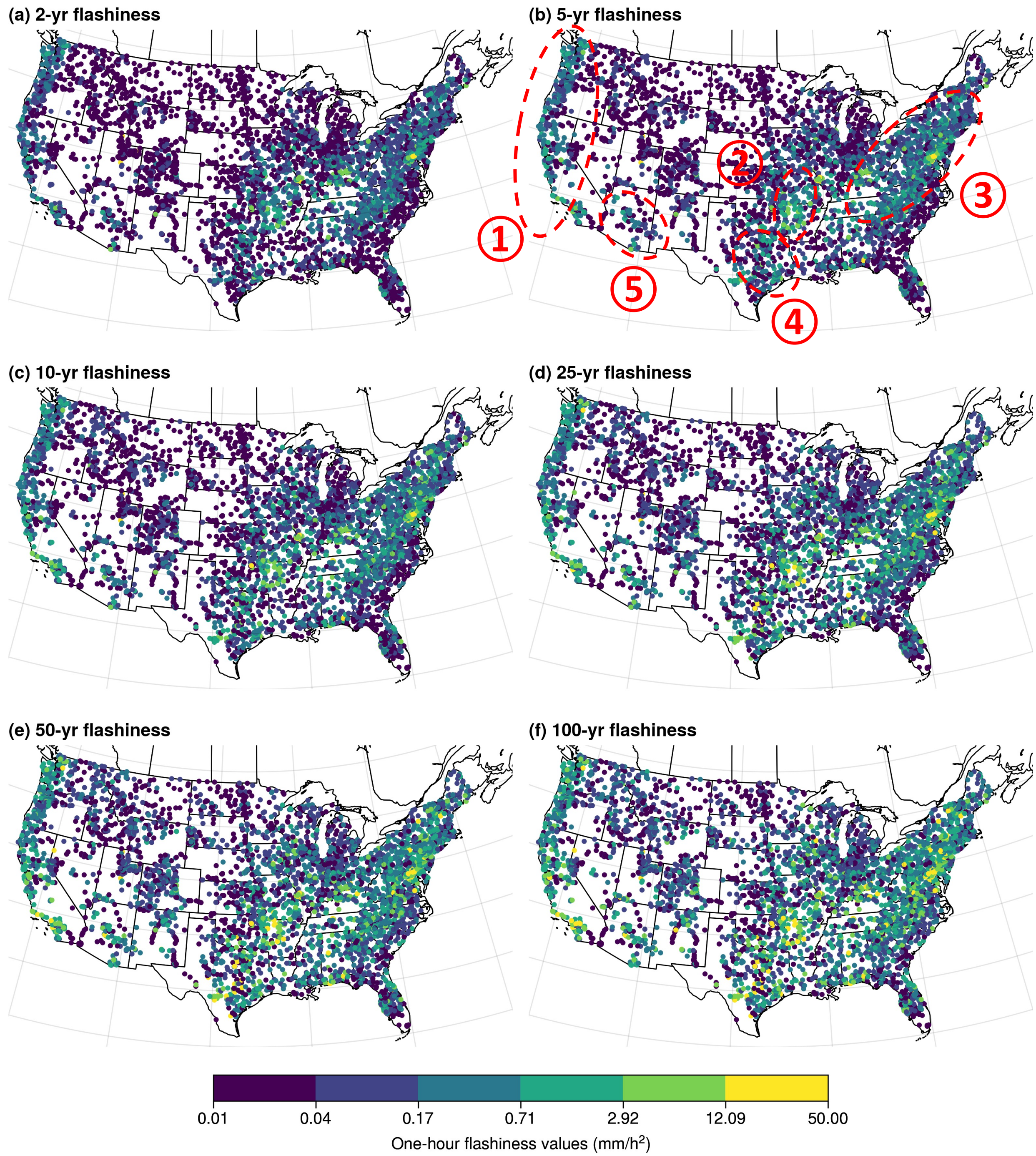

Figure 2. Maps of F-IDF values at 1-hour duration. Highlighted (numbers from 1-5) regions are clustered flashy regions in the CONUS. 1: West Coast; 2: Missouri Valley; 3: the Appalachians; 4: Flash Flood Alley; and 5: Southwest.

4.2 Factors determining flashiness values

We present a comprehensive view of factors determining flashiness values by utilizing 59 basin attributes and analyzing their correlation with flashiness. Figure 3 depicts the Spearman Correlation Coefficient (CC) between flashiness values and 59 basin attributes across 3,722 gage sites. For each site, we have CCs for six event durations and six return periods, but only the minimum, median, and maximum values are taken in the table and grouped into Hydrology, Physiography, Climate, Soils & Geology, Human, Land Cover, Natural Vegetation, and Wetland. Overall, **climate** exerts the most positive correlation with flashiness values, with annual precipitation ranked 1st place (Median CC=0.42), followed by actual evaporation and moisture index (CCs=0.4), aridity index (CC=0.39), and air temperature (CC=0.28). It's worth noting that the aridity index is positively related to the amount of moisture in the land. In other words, the lower the aridity index, the drier the land is. **Hydrologic variables** are mostly negatively correlated with flashiness in decreasing order: natural discharge (CC=-0.20), degree of regulation (CC=-0.27), river volume (CC=-0.32), river area (CC=-0.35). The exception is land surface runoff which has positive CC of 0.38. **Physiographic variables** exhibit a negative correlation with flashiness, with elevation (CC=-0.28) and drainage area (CC=-0.43) being the most significant factors. The **soils & geology** group has a relatively weak association with flashiness. Soil water content has the greatest CC of 0.39 within this class, followed by clay fraction (CC=0.19), silt fraction (CC=0.09), and sand fraction (CC=-0.16). The **human** group shows positive correlations with road density (CC=0.32) and urban density (CC=0.23) being the most significant ones. The notable features in the **land cover** group are deciduous trees (CC=0.25), artificial surface (CC=0.16), herbaceous (CC=-0.25), and deciduous shrubs (CC=-0.38). Similar to land cover, the **natural vegetation** group shows the temperate deciduous region has a positive correlation (CC=0.24) with flashiness, while grassland (CC=-0.34), open shrub (CC=-0.31), boreal evergreen (CC=-0.25), and boreal deciduous (CC=-0.23) have negative correlations. The **wetland** group does not exhibit a significant positive correlation.

The controlling factors above can be summarized as follows. First, small river reaches tend to have higher flashiness values, as the negative correlations between river area, volume, and natural discharge testify this point. Second, flood defense infrastructures impede flash flood generation, as indicated by the negative impact of the degree of regulation. Third, flashiness is highly related to wetness or annual precipitation. Fourth, flash floods are typically not snowmelt-driven processes as seen with the weakly negative correlations to snow cover. Fifth, regarding soil types, the degrees of soil types contributing to flashiness are ranked as: clay>silt>sand, which is a reversed order of permeability. Sixth, wet soils, urban density, and road density help generate flash floods by impeding soil infiltration. Lastly, dense vegetation and land cover (e.g., shrub and grassland) increase surface roughness and thus negatively correlate with flashiness.

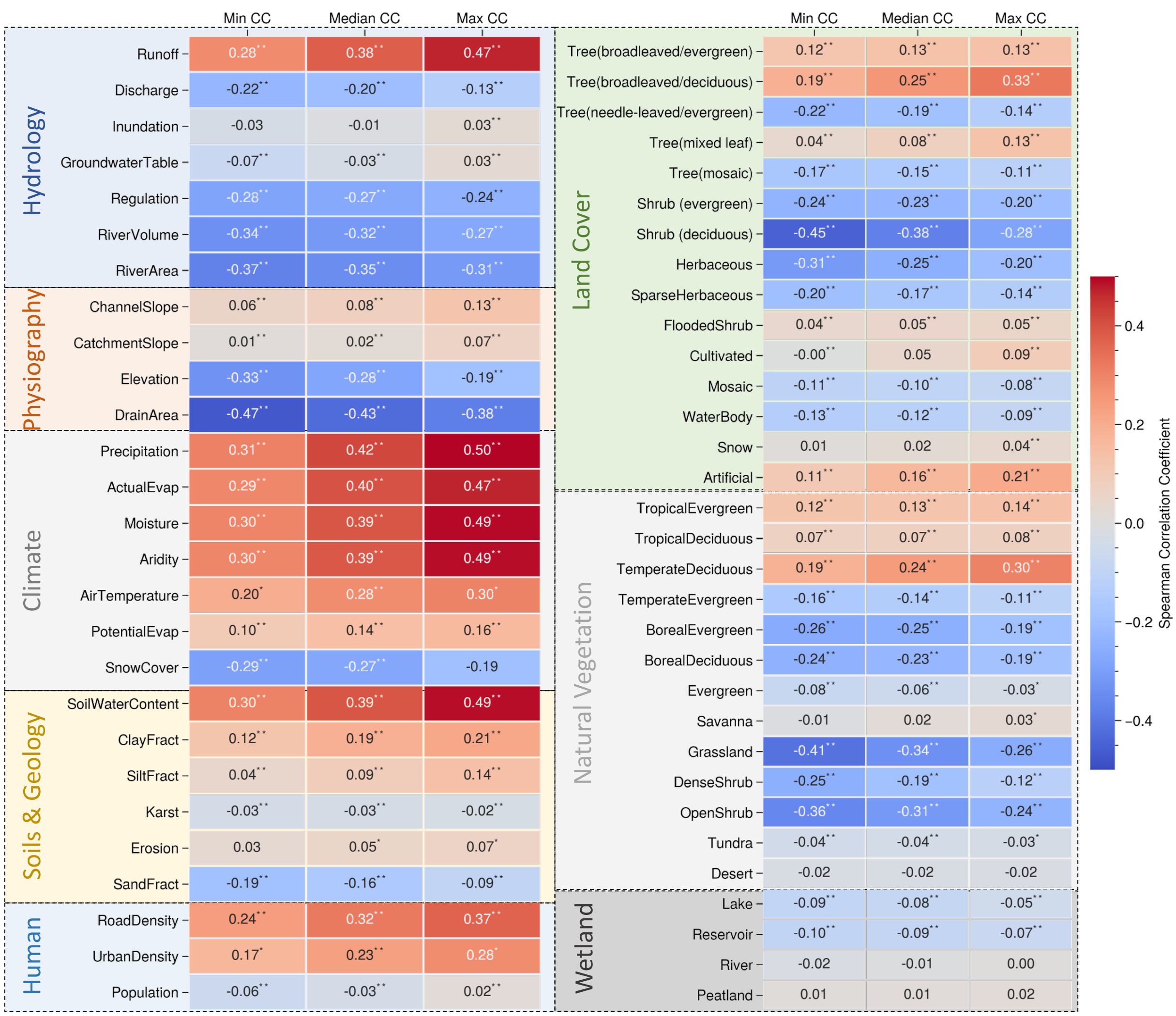

Figure 3. A table of Spearman Correlation Coefficients between flashiness and 59 basin attributes. A single asterisk (*) indicate 95% confidence level, and two asterisks (**) indicate 99% confidence level to reject a null hypothesis (zero correlation).

A unique finding in this study is the correlation of flashiness to basin attributes changes with regard to flash flood frequency and duration. We divide the 59 factors into positive correlation and negative correlation and plot their respective changes with regard to return periods and durations in Fig. 4. The significance of each slope is tested against a zero slope with the general linear F-statistics. As the occurrence of flash flood events becomes less frequent (i.e., larger return period), the absolute correlation coefficient decreases. When reaching higher levels of intensity (i.e., 100-year event), the event flashiness is less influenced by basin attributes as the causative rainfall emerges as the primary driver. The correlation coefficients increase with the duration of the event (see Figs. 4b and 4d). Likewise, correlation increases with longer-duration events, as shown in the F-IDF curve in Fig.1b, and becomes more influenced by basin attributes.

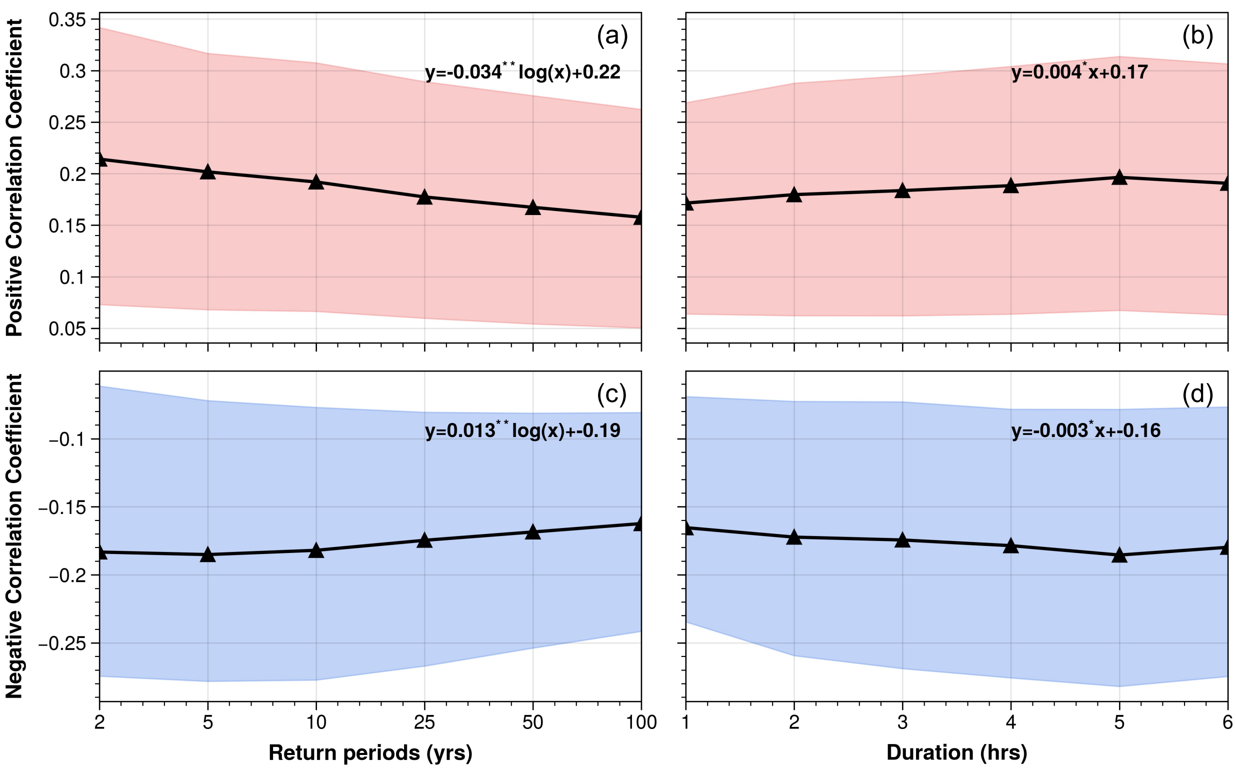

Figure 4. Plots of positive and negative correlation coefficients (by aggregating respective variables) with respect to return periods and duration. The black dotted line shows the mean correlation coefficient while the band shows the interquartile range from Q25 to Q75. The significance of the slope is tested against a zero slope using the general linear F-statistic with the

fitted regression model (equation). A single asterisk (*) indicates 95% confidence level and two asterisks (**) indicate 99% confidence level.

5 Discussion

5.1 The representativeness of flashiness index

In this study, we choose the maximum sub-hourly time derivative of streamflow over a time window as the basis to build the F-IDF curves. First, using data collected at a time scale appropriate for the application requires consideration. For investigations of flash flooding, the time step needs to be sub-hourly. Acknowledging many other variants of flashiness indices (Gannon et al., 2022; Kim & Choi, 2011; Saharia et al., 2017, 2021; Smith & Smith, 2015), this approach has several benefits. First, it is fairly simple and reproducible. The most important factors we consider the new index is the simplicity and reproducibility as it is easy to adopt and comprehend by people. Second, it represents both the flood magnitude and flood rising limb well, which is the nature of the term "flashiness" introduced by Baker et al. (2004). The first point highlights the advantage of our method, compared to previous studies. For instance, Smith & Smith (2015) fitted the discharge into a Generalized Pareto distribution (GPD), and use the shape parameter to represent flashiness. This approach generally assumes a good fit of peak flow with GPD, and it is not straightforward. Saharia et al. (2017, 2021) used a similar approach to this study, but they rescaled the flashiness index into the range of 0-1 with an empirical cumulative distribution function (ecdf). This approach prevents reproducibility since the number of gages used in rescaling will affect the final results.

This study only considers flash flood events with durations less than six hours which is a common definition for flash flooding (Clark et al., 2014). But for large basins (where the time of concentration is long) or long-duration storm, this duration of F-IDF can be further extended to 12 and 24 hours by tuning the time window parameter in Eq.1.

5.2 Correlation with basin attributes

We calculated the Spearman Correlation Coefficient of flashiness index against 59 basin attributes acquired from the HydroATLAS. As noted, the CC values are generally low (CC<0.6) for those factors. That is mainly because flash flooding, by nature, is a dynamic weather-driven phenomenon that is challenging to predict by static features (such as basin slope and annual precipitation). Similarly, Smith & Smith (2015) found that most of the CCs of number of flash flood peaks with basin attributes are lower than 0.6. Second, the CC values are calculated with uni-variate analysis, but we expect a higher value if we choose a multi-variate analysis, such as regression models and/or machine learning models. Since the main focus of this study is to provide a proof-of-concept of flashiness index, we will explore the predictability of a statistical model in a future work.

5.3 Implications for hydrologic science and flash flood response

Our proposed new metrics – F-IDF curve, has implications not only for hydrologic science but also for flash flood preparedness and responses. For the first time, this study quantifies the frequency of flash floods based on the flashiness variable computed from observed streamflow data, which provides a metric of the rapidity and severity of flooding. The same variable and associated analysis can be applied to streamflow simulations from hydrologic models. Then, the forecast flashiness and its associated frequency for a given duration can be provided ahead of time. Weather forecasters can then use such metrics to guide the issuance of flash flood warnings. Additionally, it is worth noting that the implementation of F-IDF curves is model-agnostic, meaning that it can be integrated into any flood forecast system. In the U.S., such a system may include NOAA's FLASH system, the National Water Model, etc. Second, for hydrologic modelers, the F-IDF curve provides a means of identifying flash flood events. Prior to this study, the identification of a flash flood event was vague and subjective. A common definition – a flood that occurs within six hours of a rainfall event – was too obscure for modelers to identify the start and end date of an event. However, with the help of the F-IDF curve, one can easily establish a quantitative threshold to determine a flash flood event. For instance, in a flood study, a two-year streamflow return period has often been used as a threshold to identify a flood event, given that this threshold approximately corresponds to an overbank flow rate (Li et al., 2022). Similarly, we can use a two-year flashiness value at a particular duration to sift through flash flood events. Third, for city planners and decision-makers, the existing F-IDF values can inform them of the risks of flash floods in the local area. Mitigation strategies such as green infrastructure, low-impact development, and flood defenses can help reduce flash flood risks. Fourth, assessing the risk of flash floods and planning accordingly is crucial for emergency responders. In the US, it is common practice to block flooded roads to prevent drivers from entering the water. However, this response requires proper guidance on when and how quickly road barriers should be put in place. With the help of our F-IDF curves, responders can access crucial information, such as the relationship between rate of action and the flood rising rate. This information supports their decision-making processes, enabling them to take timely actions that mitigate the risk associated with flash floods. There are undoubtedly other applications beyond those mentioned here. In summary, this newly introduced metric has implications not only for the scientific community but also for its potential role in the science-informed, policy-making process.

6 Conclusions

This article introduces a new tool – the F-IDF curve – to quantify the intensity, duration, and frequency of flash floods adopting a similar concept of the rainfall IDF curve. The F-IDF curves are quantified for 3,722 US stream gages that have at least 20 years of observation of sub-hourly streamflow. Additionally, the correlation of flashiness with regard to 59 basin attributes is also explored and discussed. Lastly, the application of F-IDF curves is demonstrated to a recent, devastating flash flood event – the 2021 Tennessee flooding. The conclusions are drawn as follows:

1. F-IDF curves are capable of revealing the spatial variability of flashy basins across the US and the following regions are identified as prone to flash flooding: the West Coast, Missouri Valley, Appalachians, Flash Flood Alley in Texas, and the Southwest.
2. Among the explored geographical and hydrometeorological factors, mean annual precipitation is the most positively correlated with flashiness while the basin's drainage area is the most negatively correlated variable.
3. The correlations weaken with increasing return periods and shorter event durations. This is attributable to the extremity of the rainfall overwhelming the influence from underlying basin attributes.

Similar to flood studies, predicting flashiness values in ungauged basins is a grand challenge that warrants scientific exploration. We plan to integrate F-IDF curves into flash flood forecast models over the US and beyond in a future work.

Acknowledgments

The authors declare that they have no known competing financial interests or personal relations that could have appeared to influence the work reported in this paper.

Open Research

The F-IDF values with joined basin attributes at US stream gages are available at https://doi.org/10.5281/zenodo.7806694 with a Creative Commons Attribution 4.0 International license (Li, 2023). The basin attributes are retrieved from https://www.hydrosheds.org/products/hydrobasins. The USGS 15-min streamflow time series is downloaded using the "dataretrieve" Python package.

References

Baker, D.B., Richards, R.P., Loftus, T.T. and Kramer, J.W. (2004), A NEW FLASHINESS INDEX: CHARACTERISTICS AND APPLICATIONS TO MIDWESTERN RIVERS AND STREAMS. *Journal of the American Water Resources Association*, 40: 503-522. doi:10.1111/j.1752-1688.2004.tb01046.x

Doswell III, C.A. (2015). Hydrology, Floods and Droughts: Flooding. In Encyclopedia of Atmospheric Sciences (Second Edition), 201-208. doi: 10.1016/B978-0-12-382225-3.00151-1.

Clark, R. A., J. J. Gourley, Z. L. Flamig, Y. Hong, and E. Clark, 2014: CONUS-Wide Evaluation of National Weather Service Flash Flood Guidance Products. *Weather Forecasting*, **29**, 377–392, https://doi.org/10.1175/WAF-D-12-00124.1.

Gannon, J., Kelleher, C., & Zimmer, M. (2022). Controls on watershed flashiness across the continental US. *Journal of Hydrology*, *609*, 127713. https://doi.org/10.1016/j.jhydrol.2022.127713

Georgakakos, K. P. , 1986. On the design of national real-time warning systems with capability for site-specific, flash-flood forecasts. *Bulletin of the American Meteorological Society*, 67, 1233–1239

Georgakakos, K. P., Modrick, T. M., Shamir, E., Campbell, R., Cheng, Z., Jubach, R., Sperfslage, J. A., Spencer, C. R.& Randall, B. (2022). Flash Flood Guidance System Implementation Worldwide: A successful Multidecadal Research-to-Operation Effort. *Bulletin of the American Meteorological Society*, 103(1), E1-E22. doi:10.1175/BAMS-D-20-0032.1

Gourley, J. J. (2013). A Unified Flash Flood Database across the United States. *Bulletin of the American Meteorological Society*, 94(6), 799-805. doi:10.1175/BAMS-D-12-00198.1

Gourley, J. J., Flamig, Z. L., Vergara, H., Kirstetter, P., Clark, R. A., III, Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J. M., Hong, Y., & Howard, K. W. (2017). The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States, *Bulletin of the American Meteorological Society*, 98(2), 361-372. doi: 10.1175/BAMS-D-15-00247.1

Hong, Y., Adhikari, P., Gourley, J.J. (2013). Flash Flood. In: Bobrowsky, P.T. (eds) Encyclopedia of Natural Hazards. Encyclopedia of Earth Sciences Series. Springer, Dordrecht. doi:10.1007/978-1-4020-4399-4_136

Kim, E. and Choi, H. (2011). Assessment of Vulnerability to Extreme Flash Floods in Design Storms, International Journal of Environmental Research and Public Health,

Li, Z., 2023. F-IDF values at USGS stream gage sites (v1.0) [Data set]. Zenodo. doi:10.5281/zenodo.7806694

Li, Z., Chen, M., Gao, S., Gourley, J. J., Yang, T., Shen, X., Kolar, R., and Hong, Y. (2021). A multi-source 120-year US flood database with a unified common format and public access, *Earth System Science Data*, 13, 3755–3766. doi:10.5194/essd-13-3755-2021.

Li, Z., Gao, S., Chen, M. Gourley, J., Liu, C., Prein, A.F., Hong, Y. (2022). The conterminous United States are projected to become more prone to flash floods in a high-end emissions scenario. *Communications Earth and Environment*, 3, 86. doi:10.1038/s43247-022-00409-6

Lin, K., Chen, H., Xu, C., Yan, P., Lan, T., Liu, Z., & Dong, C. (2020). Assessment of flash flood risk based on improved analytic hierarchy process method and integrated maximum likelihood clustering algorithm. *Journal of Hydrology*, *584*, 124696. doi:10.1016/j.jhydrol.2020.124696

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Scientific Data* 6: 283. doi: 10.1038/s41597-019-0300-6

Ma, M., Liu, C., Zhao, G., Xie, H., Jia, P., Wang, D., Wang, H., & Hong, Y. (2019). Flash Flood Risk Analysis Based on Machine Learning Techniques in the Yunnan Province, China. *Remote Sensing*, 11(2), 170. https://doi.org/10.3390/rs11020170

Maddox, R. A., Chappell, C. F., & Hoxit, L. R. (1979). Synoptic and Meso-α Scale Aspects of Flash Flood Events. *Bulletin of the American Meteorological Society*, 60(2), 115–123. doi:10.1175/1520-0477-60.2.115

Perica, S., Martin, D., Pavlovic, S., Roy, I., Laurent, M.S., Trypaluk, C.A.R.L., Unruh, D., Yekta, M. and Bonnin, G., 2013. Noaa atlas 14 volume 9 version 2, precipitation-frequency atlas of the united states, southeastern states. NOAA, National Weather Service, Silver Spring, MD, 18.

Saharia, M., Kirstetter, P. E., Vergara, H., Gourley, J. J., Hong, Y., & Giroud, M. (2017). Mapping Flash Flood Severity in the United States. *Journal of Hydrometeorology*, 18(2), 397-411.

Saharia, M., Kirstetter, P.-E., Vergara, H., Gourley, J. J., Emmanuel, I., & Andrieu, H. (2021). On the impact of rainfall spatial variability, geomorphology, and climatology on flash floods. *Water Resources Research*, 57, e2020WR029124. doi:10.1029/2020WR029124

Singh, V.P. (1998). Log-Pearson Type III Distribution. In: Entropy-Based Parameter Estimation in Hydrology. Water Science and Technology Library, vol 30. Springer, Dordrecht. doi:10.1007/978-94-017-1431-0_15

Smith, B. K., & Smith, J. A. (2015). The Flashiest Watersheds in the Contiguous United States. *Journal of Hydrometeorology*, 16(6), 2365-2381. doi:10.1175/JHM-D-14-0217.1

Smith, J. A., Baeck, M. L., Yang, L., Signell, J., Morin, E., & Goodrich, D. C. (2019). The Paroxysmal Precipitation of the Desert: Flash Floods in the Southwestern United States. *Water Resources Research*, 55(12), 10218-10247. doi:10.1029/2019WR025480

Titley, H. A., Cloke, H. L., Harrigan, S., Pappenberger, F., Prudhomme, C., Robbins, J. C., Stephens, E. M., & Zsótér, E. (2021). Key factors influencing the severity of fluvial flood hazard from tropical cyclones. *Journal of Hydrometeorology*, 22(7), 1801-1817. doi:10.1175/JHM-D-20-0250.1

Wallis, J. R., & Wood, E. F. (1985). Relative Accuracy of Log Pearson III Procedures. *Journal of Hydraulic Engineering*, 111(7), 1043-1057.