Mapping a Novel Metric for Flash Flood Recovery Using Interpretable Machine Learning

ANIL KUMAR[a], MANABENDRA SAHARIA[a,b] AND PIERRE KIRSTETTER[c]

[a] *Department of Civil Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India*
[b] *Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India*
[c] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 20 November 2023, in final form 29 May 2024, accepted 22 August 2024)

ABSTRACT: Flash floods are one of the most devastating natural disasters, yet many aspects of their severity and impact are poorly understood. The recession limb is related to postflood recovery and its impact on communities, yet it remains less documented than the rising limb of the hydrograph to predict the peak discharge and timing of floods. This work introduces a new metric called flash flood recovery or recoveriness, which is the potential for recovery of a watershed to preflood conditions. Using a comprehensive database of 78 years and supervised machine learning algorithms, flash flood recovery is mapped in the conterminous United States. A suite of geomorphological and climatological variables is used as predictors to provide probabilistic estimates of recoveriness. Slope index, river basin area, and river length are found to be the most significant predictors to predict recoveriness. Several new localized hotspots were identified, such as the western slopes of the Appalachians consisting of Kentucky, Tennessee, and West Virginia and the interlinked areas of western Montana and northern Idaho. This new metric can be useful for prioritizing relief and rehabilitation efforts as well as precautionary measures for disaster risk reduction.

KEYWORDS: Flood events; Hydrometeorology; Uncertainty; Decision trees; Machine learning

1. Introduction

Accounting for one-third of all global geophysical disasters, floods cause substantial damage to agriculture, infrastructure, human life, and the socioeconomic system (Berz 2000; Douben 2006; Rentschler et al. 2022). Although total flood damage caused nationally in the United States varies from year to year, there has been a statistically notable increasing trend of 2.92% per year in the twentieth century (Pielke and Downton 2000). In the water year 2021–22 alone, 102 flood-related deaths and direct flood damages of $2.8 billion have been reported in the United States. A comprehensive large-sample characterization of floods in the United States revealed that the fastest-responding flash floods were caused by intense monsoon thunderstorms and steep terrain in the arid southwestern United States (Saharia et al. 2017a). Globally, streamflow extremes are characterized by various hydrogeomorphic factors such as annual precipitation, precipitation of wettest month and quarter, basin magnitude, first-order stream length, basin perimeter, and drainage area (Kuntla et al. 2022). A new variable called flashiness was proposed for quantifying flash flood severity of such severe events using a multitude of geophysical and climatological variables (Saharia et al. 2017b). Most prognostic efforts in hydrology are dedicated toward understanding and predicting peak discharge and rise time of flooding, while very little research is available on the recession limb of the hydrograph, including nonavailability of metrics describing the recovery of a watershed to preflood conditions.

Understanding and predicting flood hydrograph is the overarching goal of the field of hydrology. A flood hydrograph consists of a rising limb and a falling limb. At the event scale, the rising limb rate of a watershed was found to be correlated with the maximum rainfall intensity and the elapsed time of rainfall centroid (Shuster et al. 2008). Potdar et al. (2021) found that the spatial organization of rainfall influences the basin response on par with the geomorphology and climatology of flash-flood-generating basins. The rising limb of the flood hydrograph is related to extreme events such as flash floods, while the recession part of the curve is concerned with water drainage (Shorr 2000). The recession curve is typically longer than 50% of an entire flooding event due to a copious supply of water resources during the entire flood event (Ahmad et al. 2014; Liu et al. 2015). In operational flood forecasting, the crossing of different river stages (action/minor/moderate) activates different types of procedural responses from local responders, and restoration of a watershed to preflood conditions is dependent on how fast the river flow can return below the breached river stage. Despite this, much of hydrology research is focused on the rising limb, especially predicting the peak discharge and timing of a flood, while the recession period of floods, although being crucial for relief and recovery of communities, postflooding health, and environmental concern, has gained less attention. A multitude of metrics have been developed to assess the severity of floods (Saharia et al. 2017b; Baker et al. 2004; Diakakis et al. 2020; Alfieri and Thielen 2015), but based on our current knowledge, no such metric exists for recovery of a watershed to preflood conditions, i.e., flood recovery. Li et al. (2023b) introduced a new metric called flashiness–intensity–duration–frequency (F-IDF), quantifying flash flood intensity across 3722 stream gauge locations in the United States. This work was extended in Li et al. (2023a), where the study presented two distributed products, a machine learning–based F-IDF

Corresponding author: Manabendra Saharia, msaharia@iitd.ac.in

DOI: 10.1175/JHM-D-23-0196.1

product and a physics-based F-IDF product, and compared their effectiveness in identifying flash-flood-prone regions. It is worth mentioning that there do exist metrics based on performance for flood resilience which have been in use in urban water systems and management (Lee and Kim 2017; Mugume et al. 2015). In the work by Wang et al. (2023), they describe the "flood susceptibility index" as a quantitative measure to assess the resilience of urban areas against floods at the grid-cell level. This index incorporates a performance-based metric derived from flood duration and magnitude. The index values were computed using high-resolution data, which made it possible to assess resilience against flood, in great detail across different urban cells.

A few researchers have explored the relationship between the recession period and physical predictors. According to Khaleghi et al. (2011), in mountainous regions, the two main predictors responsible for the flood recovery period are 1) rainfall and 2) underlying surface conditions. Ground surface conditions also affect the flood recession process (Chang and Feng 2017; Costa et al. 2003). Conversely, the transformation of grassland or cultivated land into urban areas would lead to a rise in flood peak flow, thereby increasing the recovery period. Another factor affecting the flood recovery period is terrain slope. For example, in the slope experiment presented by Shixiang and Shaowen (1991), the impact on the recession flow drops rapidly with a slope exceeding 9°, and greater impact is expected when the slope value is below 9°, suggesting that 9° is a critical value. Studies such as Ye et al. (2019) discuss how the river cross section affects the recession period and flow. Amit et al. (2002) found that the main factors affecting the recession curve in perennial springs are the aquifer lithology and the geometry of the water conduits. Biswal and Marani (2010) suggested a link between recession curves and the topology of a river network. Bhaskar et al. (2000) defined an index based on flood hydrograph shape characteristics such as magnitude ratio, gradient of the rising curve, and the response time of the flash flood to differentiate between floods and flash floods. Evaluation of various techniques for baseflow and recession analyses was done by Nathan and McMahon (1990) and Chapman (1999), where they showed that the linear storage model can be used as a very good approximation in most cases by comparing algorithms for streamflow recession.

Though the literature on baseflow recession curves is extensive, none of the metrics and techniques developed are based on the actual definition of floods used by operational agencies such as the U.S. National Weather Service (NWS). NWS has predefined flood stages at thousands of gauging stations across the country, established in cooperation with local public officials. Once the river reaches the flood stage, NWS declares a minor, moderate, or major flood, with each category defined based on property damage and public threat. Thus, a metric for flood recovery tied to operational definitions of floods and flash floods has the potential for wide usage. Apart from developing this metric at designated gauge stations, this metric is required at every watershed so that flood recovery can be monitored in ungauged locations.

In this paper, we have introduced a new index for flash flood recovery, called "recoveriness," based on a large observational dataset. Further, we have mapped recoveriness to every watershed in the United States on a continuous grid using a machine learning–based approach on a multitude of predictors related to basin climatology and geomorphology. Further, an interpretable machine learning framework called Shapley additive explanations (SHAP) has been used to quantify the relative impact of the causative hydrogeomorphic factors behind recoveriness. The study is organized as follows: Section 2 describes the data sources and methods that have been applied; section 3 describes the machine learning methods and interpretation methodology; section 4 characterizes recoveriness based on a multitude of variables and maps it over the conterminous United States; and finally, section 5 provides a summary of the findings and concluding remarks.

2. Study area and data source

a. Datasets

This study uses the unified flash flood database (Gourley et al. 2013), which is a compilation of carefully selected data from various sources such as the USGS streamflow data, NWS flash flood reports, and a storm event database. The unified flash flood database is further enhanced by public survey response data collected during the Severe Hazards Analysis and Verification Experiment (SHAVE) described by Ortega et al. (2009) and the event-scale rainfall spatial variability (Saharia et al. 2021).

The USGS conducts an automated collection of instantaneous streamflow information at every 5–60-min interval for 10 106 gauges distributed over the United States. A coordinated approach by NWS with the USGS and local stakeholders has resulted in defining various categories of flooding such as minor, moderate, and major in addition to establishing crucial flood level stages like the action stage. This comprehensive approach is applied to the streamflow gauges at 3490 locations. Well-defined threshold levels have been defined for a specific set of gauges within the USGS network, which forms a significant input for various applications including modeling. The term "action stage" denotes the level at which the NWS could initiate measures for a possible adverse hydrologic event. The stage mostly appears with conditions indicative of the bank full cases.

Interestingly, around 41% of the USGS stations display the same values for the action and bank full stages with a mean variation of 1.3%. Data reliability is further ensured with USGS providing regulations for the stations vulnerable to human alterations and diversions leaving us with 70 596 flooding events from 1649 unique monitoring stations. By convention, the USGS characterizes flood events when the streamflow exceeds a predefined action stage for each gauge in its network. Further, events must be separated by at least 24 h to be counted as independent events in the unified flash flood database. Essential details are already specified in the core database. For each gauge in the network, the following are specified: unique identifier (ID) for each gauge, the latitude-

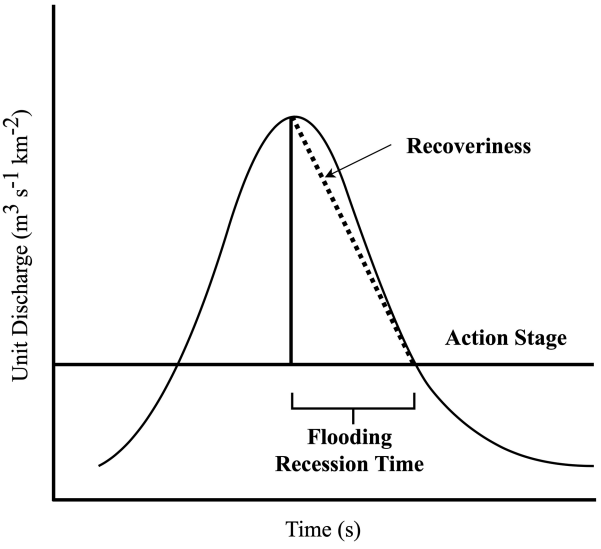

FIG. 1. Graphical representation of recoveriness—the part right of the peak of the graph.

and longitude-based geographical coordinates, initiation time (UTC) that marks the exceedance of the flow above the action stage threshold, conclusion time (UTC) that marks the recession below the threshold, the peak flow magnitude ($m^3\ s^{-1}$), the time at which the peak was attained (UTC), and the flood rise time, i.e., the interval between the discharge when it surpassed the threshold and the discharge when it attained the peak (hours).

b. *Recoveriness as a metric of flood recovery*

According to the NWS, flash floods are said to have occurred when there is an intense rise in water level in an area that typically remains arid or when the water level surpasses the predefined flood threshold in the case of a stream or a creek. This can happen within 6 h from the beginning of situations like dam break, heavy rainfall, and water flow obstruction due to ice.

This work introduces a novel metric called recoveriness for measuring flood recovery. We define recoveriness in Eq. (1) to be the ratio of the difference between the peak discharge and the action stage discharge, divided by the time it takes for the flooding to return to the action stage and the basin area. As can be seen in Fig. 1, the metric characterizes the rate of decline of the hydrograph from its peak position, consequently capturing both the magnitude and timing components. Higher values signify accelerated recovery to preflood conditions. If ψ denotes recoveriness, the gauging station numbers as S, and the number of events as N_i (for gauge i), where i ranges from 1 to S, then the recoveriness for a jth event can be expressed as in Eq. (1):

$$\psi_{ij} = \frac{Q_{ij}^{(p)} - Q_{ij}^{(a)}}{A_i T_{ij}}. \tag{1}$$

The term $Q^{(p)}$ in the above equation represents the peak discharge, $Q^{(a)}$ represents the action stage discharge, A represents the basin area, and T represents the recession time. To normalize these values between 0 and 1, an empirical cumulative distribution function (ECDF) is utilized as depicted in Eq. (2). The standardized version of recoveriness ϕ is given by

$$\tilde{\phi}_{ij} = \frac{1}{\sum_{i=1}^{S} N_i} \sum_{i=1}^{S} \sum_{j=1}^{N_i} I_{(\phi_{ij} \le t)}, \tag{2}$$

where t is the ranked value of recoveriness and $I_{(E)}$ is the indicator function yielding 1 if the condition E is true and 0 otherwise. Event-level recoveriness is calculated by Eq. (2). The standardized recoveriness for a given basin i and the median value of computed recoveriness can be given using Eq. (2) for all events N_i observed at that gauge station.

TABLE 1. Recoveriness indicator.

$Q_{\text{diff.}}$ ($m^3\ s^{-1}$)	Area (km^2)	Time (months)	Recoveriness	Interpretation
0.65 Low	150.0 Low	1.0 Low	0.66	Moderate
0.65 Low	150.0 Low	6.0 High	0.33	Moderate
0.65 Low	500.0 High	6.0 High	0.16	Slow
11.50 High	150.0 Low	1.0 Low	1.00	Rapid
11.50 High	150.0 Low	6.0 High	0.83	Rapid
11.50 High	500.0 High	6.0 High	0.50	Moderate

High recoveriness values are a fingerprint of rapid (transient) hydrological processes that are efficient at transferring water across the watershed. It translates also into high capability to concentrate water at the watershed outlet which is associated with high flashiness (Saharia et al. 2017b). The definition of recoveriness involves the discharge terms $Q^{(p)}$ and $Q^{(a)}$ in the numerator and the area and recession time in the denominator. Table 1 illustrates how recoveriness varies as a function of discharge, area, and time and how it can be interpreted. The highest recoveriness is achieved when $Q_{\text{diff.}}$ is high, while area and time are low. Conversely, if $Q_{\text{diff.}}$ is low while area and recession times are high, we expect the slowest watershed recovery. A recoveriness value of 0–0.25 reflects slow recovery, 0.25–0.75 reflects moderate recovery, and 0.75–1 reflects rapid recovery.

In simple terms, recoveriness can be thought of as the system's capacity to rebound from a flood event to preflood conditions. The index bears a nonlinear relationship with its 14 predictors shown in Table 2. For instance, the $Q^{(p)}$ term readily encapsulates the effect of precipitation and rainfall as they can directly escalate its value which consequently manifests in the computation of recovery. Similarly, temperature can affect the melting times of snow again affecting $Q^{(p)}$. Water retention and runoff are determined by slope index and rock volume; drainage efficiency is determined by the outlet

TABLE 2. Description of the variables.

Sr. No.	Hydrogeomorphic predictor	Description
1	Area (area)	Watershed area contributing to runoff
2	Elongation length (el)	How long the watershed is
3	Shape factor (k)	A dimensionless quantity which is equal to the watershed area divided by the square of the channel length
4	River length (rl)	The river length can be defined as the distance measured along the line that connects the watershed outlet to the point where the watershed's boundary meets the river's main channel
5	Relief ratio (rr)	The relief ratio is defined as the elevation difference between a watershed's lowest point and its highest point to the watershed's length, directly affecting both the runoff and the speed of flood rising speed
6	Slope index (si)	A quantity calculated from the slopes at 10% and 85% of the length of the main channel in the upstream direction starting from the mouth of the watershed (Saharia et al. 2017b)
7	Slope outlet (slopeoutlet)	Computed slope at a distance of 1 km from the watershed outlet
8	Precipitation (precip)	Precipitation
9	Temperature (temp)	Temperature

slope and the relief ratio; elevation and area are concerned with water input and flow; soil infiltration and transport are affected by the river length; and curve number, rock depth, and texture determine the runoff rates. The collective effect of all these variables is embedded in the hydrograph, and the recoveriness index facilitates a quantity that is easy to interpret.

As per the definition provided, observed recoveriness is derived across the United States as shown in Fig. 2. Several regional hotspots can be visually identified with high recoveriness: 1) the West Coast, 2) Arizona, 3) the Front Range, 4) Flash Flood Alley, 5) the Missouri Valley, and 6) the Appalachians. As recoveriness directly depends on the geomorphology and topology of the watershed, high recoveriness in these six regions can be explained by examining those predictors. First, in the West Coast, the high recoveriness values could be due to the ridgetops and steep slopes on the western side of the Sierra Nevada range which obstruct the northern Pacific moisture-laden air moving eastward. Second, high recoveriness values are observed in the basins in Arizona ranging from the northern higher terrain plateau to the lower deserts in the southeastern Arizona. This is due to the high slopes of the front range of the Rocky Mountain range capturing the moisture traveling from the Gulf of Mexico, e.g., during the American monsoon. High recoveriness is also observed in the urban corridor covering Austin, San Antonio, and Waco along the Balcones Escarpment in Texas (Flash Flood Alley). A secondary cluster can be observed near the Gulf Coast close to Houston. The recoveriness increases as one moves farther toward the northeast following the Ozarks regions and maximizes in and around the center of Missouri. Also, in the eastern United States, from Georgia to Maine in and around the Appalachians, high recoveriness values are also observed.

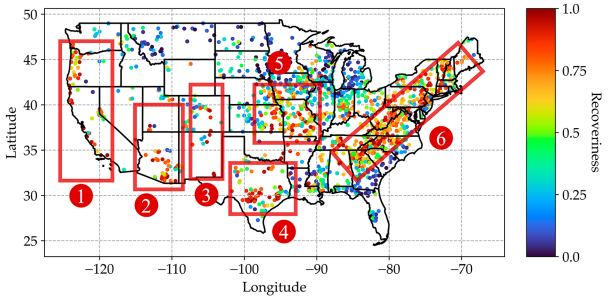

FIG. 2. Map of the United States showing the spatial distribution of recoveriness. The bounded boxes denote the following regions: 1) the West Coast, 2) Arizona, 3) Front Range, 4) Flash Flood Alley, 5) Missouri Valley, and 6) the Appalachians.

3. Methodology

In this study, the random forest quantile regression (RFQR) has been used as the supervised learning algorithm for regression between recoveriness and the multitude of hydrogeomorphic predictors. The complete methodology of the study covering data preparation, train–test split, hyperparameter tuning, performance evaluation, and model interpretation is explained in Fig. 3. The dataset is divided into a training set and a testing set (4:1 ratio) for a supervised learning analysis.

Grid search is a widely used technique for hyperparameter tuning where the training algorithm is tested for its performance at different combinations (grid) of hyperparameter values. The scheme naturally avoids any exhaustive search step while providing the best combination from the set. Therefore, for tuning the algorithm, a grid search is performed, of which we only report the final and optimal set of hyperparameters of the model here. The mean-square error and coefficient of determination metrics are used to evaluate model performance. Model interpretation is performed using SHAP, which is a game-theoretic approach to explain the output of any machine learning model (Lundberg and Lee 2017).

a. *Hydrogeomorphic predictors*

A total of 14 physical variables have been derived from various datasets at the watershed scale and used as predictors in this study. The list of predictors and their description in the context of floods is provided in Table 2.

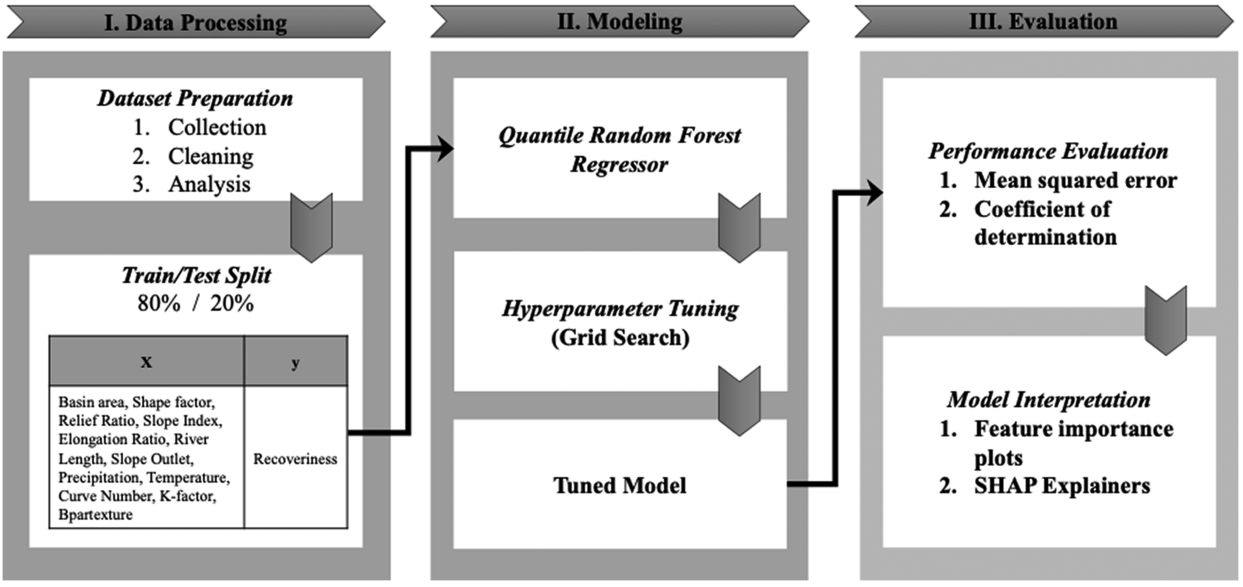

FIG. 3. Methodology for the computation of observed and predicted recoveriness. See Table 2 for a description of variables.

- The area (area) of the watershed is important as larger areas usually collect more water/runoff.
- The elongation length (el) and shape factor (k) affect the time and speed with which water is channeled into nearby water bodies like rivers.
- River length (rl) can play a dual role—while elongated regions may appear more prone to flooding, they may also allow more time for assessment.
- Relief ratio (rr), slope at the outlet (slopeoutlet), and slope index (si) are indicative of steepness of the watershed and terrain.
- Climatologic variables such as precipitation (precip) and temperature (temp) also impact floods. The risk is high with higher rainfall and higher temperatures, and the snowmelt is faster.
- Infiltration into the ground and runoff are determined by soil and rock predictors such as the k factor (k-fact), rock depth (rd), and rock volume (rv).

The soil texture (bpartexture) and the curve number (cnbasin) are indicative of infiltration rates and direct runoff, respectively.

As a first-hand examination of data, the Pearson correlation coefficient matrix is computed for the entire data. The formula for the Pearson correlation coefficient is given by

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \quad (3)$$

In Eq. (3), r_{xy} denotes the correlation coefficient between x and y. The means of the two variables are denoted by $\bar{x}, \bar{y}$. Application of the above formula on the input data produces (Fig. 4) a heatmap indicative of the correlation coefficient among all pairs of variables. Apart from a high correlation with themselves, the following variable pairs show a correlation higher than 0.5—(area, river length), (relief ratio, slope outlet), (relief ratio, slope index), and (elongation, shape factor). Regarding the output variable, recoveriness (mr.ecdf), the following variables are least correlated—curve number, temperature, and Bpar texture. The blue and green regions in the heatmap represent zero to negatively correlated pairs. The dendrogram represents the closeness among different variables. The dendrogram above the heatmap shows the clusters formed among combinations of variables. Predictor pairs that branch in together show closer association with similar information.

b. *Shapely values*

Originating from the cooperative game theory, the "shapely" values are a way to assign value to individual contributor from the total value created from the coalition. The SHAP is a global feature interpretation method based on shapely values which can be used to interpret the relative impact of predictors on model outputs in machine learning (Lundberg and Lee 2017). The following formula is used to compute the SHAP values:

$$\phi_i(x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!}[f_x(S \cup \{i\}) - f_x(S)]. \quad (4)$$

In the above formula, ϕ_i is the SHAP value for ith predictor and x instance, N is the complete predictor set, S is the set of predictors excluding the ith predictor, and $f_x(S \cup \{i\})$ is the prediction of the model including the ith predictor in subset S for the x input. The term $f_x(S)$ is the prediction of the model when only including predictors in a subset of S for x input. SHAP has been widely used for explaining the output of machine learning models, and in the context of hydrology, Pradhan et al. (2023) used them to explain the outputs of

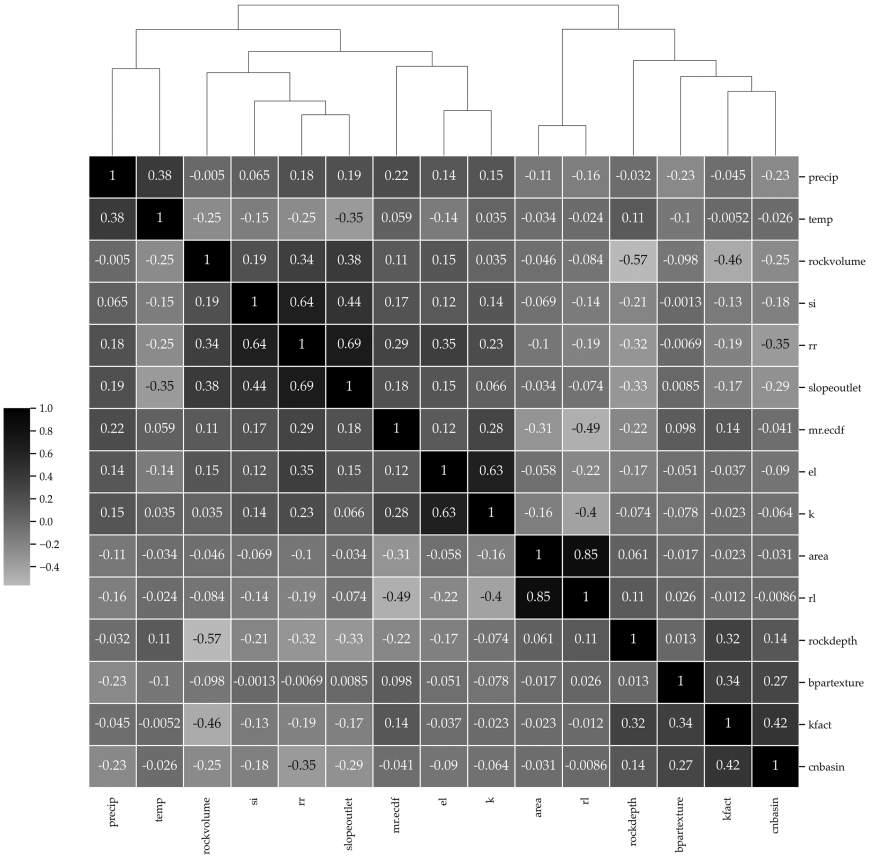

FIG. 4. Pearson's correlation coefficient showing the correlation among different input variables with the numerical values indicating the *r* value.

the deep convolutional neural networks for flood susceptibility mapping; Ekmekcioğlu and Koc (2022) used SHAP, the algorithm to analyze the impact of hazard conditioning factors for floods and landslides; Madhushani et al. (2024) used SHAP to explain the output of the extreme gradient boosting (XGB) model in streamflow predictions; and Yang and Chui (2021) used SHAP in identifying the contribution of each input feature (such as rainfall) to the model's output (such as runoff), even revealing instances, thus helping realize where the model's output did not align with the physical reality.

c. *RFQR*

The relationship between a hydrograph variable (recoveriness) and hydrogeomorphic controls is complex and nonlinear. Tree-based methods have been particularly successful in modeling such relationships in hydrology since their versatility can be attributed to their capability of handling nonlinearity (Ao et al. 2019; Babagoli et al. 2019; Huan et al. 2020), explainability (Gimeno et al. 2023; Islam et al. 2020; Joshi et al. 2023; Nanfack et al. 2022), and robustness toward outliers in data (Buschjäger et al. 2022; John 1995; Marcuzzi et al. 2022; Panjei et al. 2022). Since our objective is to model recoveriness in observed locations, map it to ungauged locations, and provide an estimate of uncertainty, we selected the RFQR algorithm (Meinshausen and Ridgeway 2006). RFQR is an extension of the RF algorithm that can model the statistical distribution of the target variable, allowing us to provide quantile estimates of the predicted values. Interval estimates in RFQR provide quantiles that help gain a more complete picture of the potential outcomes in the target. Markov chain Monte Carlo–based methods also exist when it comes to making probabilistic estimates about the dependent variable (Amaya et al. 2022; Kumar et al. 2020; Yan et al. 2020). However, scaling them to large datasets poses an issue; therefore, we opt for the random forest quantile regression algorithm.

During the training stage, while the conventional RFs only store summary statistics such as the mean and mode of the training samples, RFQR stores all the samples reaching the terminal nodes. At the prediction stage, RFQR uses these stored samples to construct an ECDF which is later used to predict the quantiles for new data points. Thus, any quantile can be queried based on these CDFs.

The algorithmic steps of RFQR can be described as follows:

- Bootstrap sampling: Create bootstrap samples from the original data.
- Build decision tree: For each of such bootstrap partitions, build a decision tree using the mean-square criterion (for regression problems).

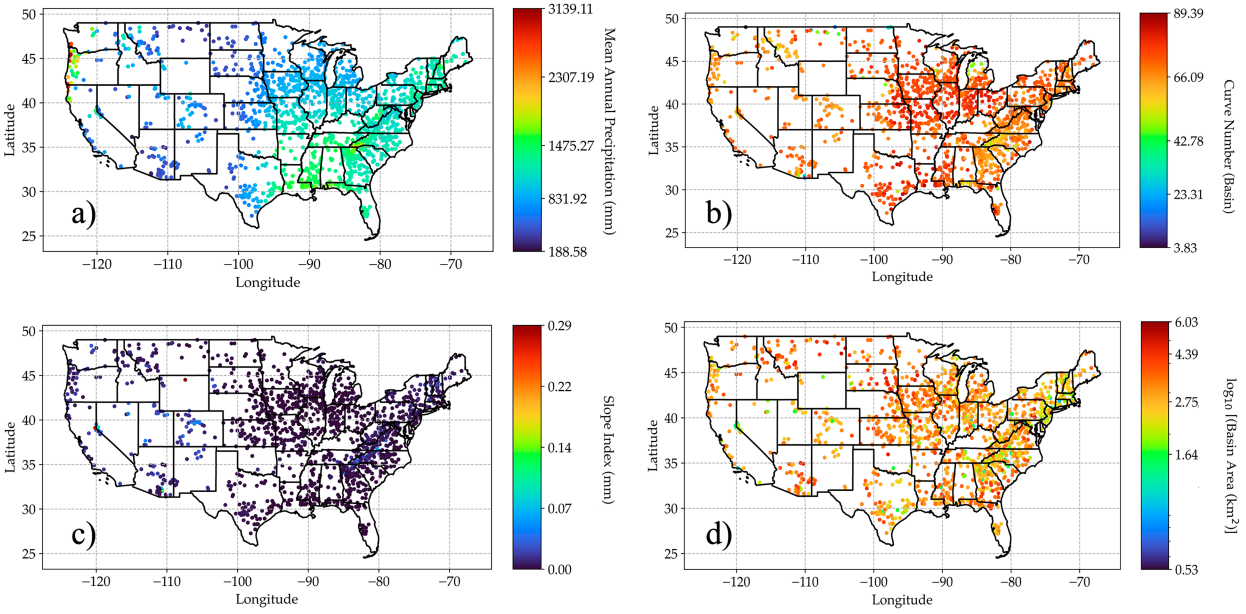

FIG. 5. Scatterplot showing the distribution of (a) mean annual precipitation, (b) CN, (c) slope index, and (d) logarithm of watershed area across the CONUS.

- Leaf node prediction: Store all the data points reaching the leaf node to enable quantile computation.
- Voting: During prediction, get the votes from each tree to predict the quantile estimate for the new data point.

As the recoveriness variable bears a nonlinear relationship with the predictors (Table 2, Fig. 4), it is important to model the entire statistical distribution of the target variable and not just summary statistics. It is worth mentioning that the recoveriness follows the same pattern as flashiness, and these are consistent with the geographical patterns of the fatalities, and therefore, there is a geographical consistency (Saharia et al. 2017a).

4. Results and discussion

a. Hydrogeomorphic controls on recoveriness

Recoveriness is a highly complex and nonlinear function of geomorphological and climatological predictors of a watershed. The various hydrogeomorphic controls include annual precipitation, watershed area, ruggedness, k factor, rock depth, and curve number of the watershed. The spatial distributions of important climatological and geomorphological variables are summarized in Fig. 5. By comparison with the recoveriness spatial distribution in Fig. 2, with some notable exceptions in the extreme west coast of Washington and Oregon County (marked red), the area is characterized by high recoveriness in the Pacific Northwest and toward the southeast near the Appalachians (Fig. 5a). Arizona, although being more arid than other flashy regions, exhibits several high recoveriness areas.

The runoff response to excess rainfall can be characterized by the basin curve number. Many predictors affect this variable, including land cover and hydrologic soil group that help approximate runoff generation basing infiltration, vegetative interception, and soil moisture retention. It varies in the range of 30–100, with higher values indicating higher runoff potential. In Fig. 5b, we observe a high runoff potential in the Flash Flood Alley and Missouri. It is expected to be highly correlated with the recoveriness due to the presence of a combination of steep terrains and impermeable soils. In Fig. 5c, the slope index is representative of the DEM-derived slope computed along the basin length (Costa 1987).

Regions like the Appalachians, Sierra Nevada, and Arizona show high recoveriness due to the presence of higher slopes with a few exceptions where high values are observable in flat regions as well. In Fig. 5d, some notable clusters of gauged catchments can be observed on the border of Nevada and California near the Tahoe Lake, as well as in some midwestern areas like Indianapolis and Saint Louis and extending along the Appalachian regions toward densely populated northeastern United States. Due to the presence of small catchments, the watersheds are bound to possess high recoveriness values—as evident upon comparison with Fig. 2.

The influence of geomorphological and climatological predictors such as basin area, average annual rainfall, steepness of the terrain or slope index, and curve number on recoveriness has been examined using quantile plots (Fig. 6). In the current context, a quantile plot gives information about the correlation and variability in the relationship between the predictor variable and recoveriness. Each plot shows the 1st through 99th quantiles for specific ranges of the predictor. The first-order dependency is revealed by the conditional median or the 50th quantile; the remaining ones can serve as estimates of the uncertainty in the relationship. Extreme values of recoveriness can be determined by excluding the interquartile range between the 10th and 90th quantiles.

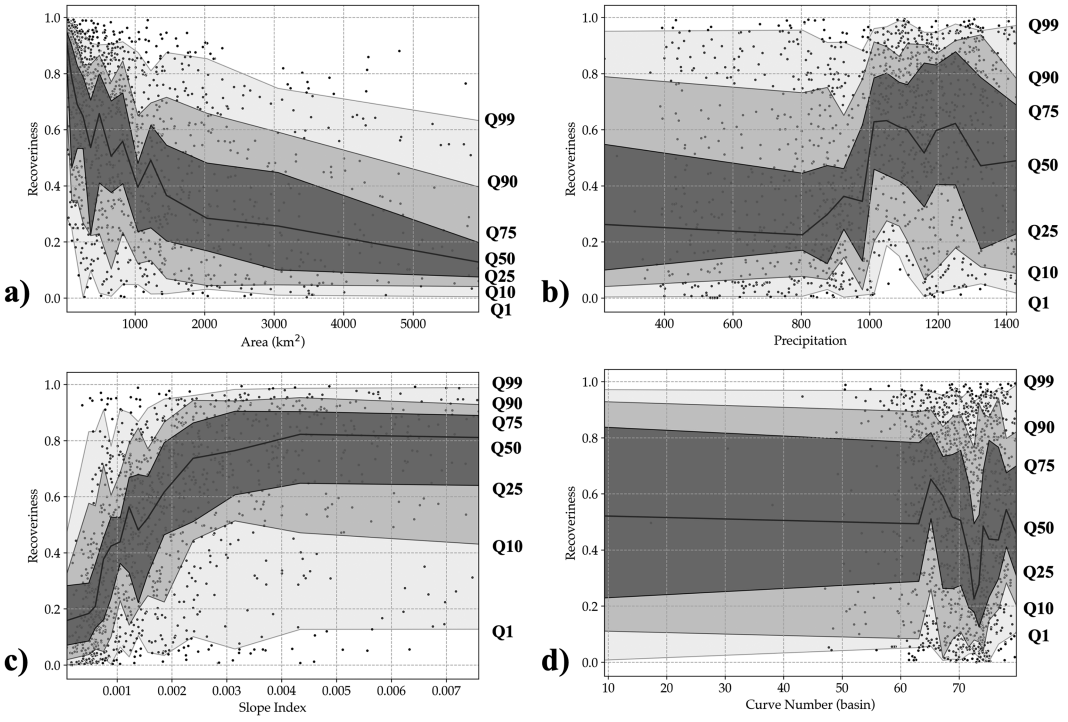

FIG. 6. Plots showing recovery percentiles from the first to the ninety ninth, for the following variables: (a) basin area, (b) mean annual precipitation, (c) slope index, and (d) CN. The percentiles are obtained by a simple partitioning of the data.

Figure 6a shows the variation of recoveriness with respect to the watershed area. One can observe that there is a gradual decrease in recoveriness with increasing area, as expected since larger basins are expected to display slower dynamics than smaller basins. Figure 6b shows the variation of recoveriness with mean annual precipitation. Clearly, there is a conditional relationship between recoveriness and mean annual precipitation which appears to be constant below 1000 mm, after which it increases slightly and then decreases. This is suggestive of a threshold effect. Figure 6c shows the variation of recoveriness against the slope index, which displays a clear increasing relationship till a threshold value. Recoveriness increases with slope index for slope index values below 0.3 (higher slope is associated with faster transfer of water), and for higher values, recoveriness bears an almost constant relation with this geomorphologic variable. The relation of the basin curve number (CN) and recoveriness is shown in Fig. 6d. Recoveriness shows a varying degree of correlation with the four predictors which are deemed important in most hydrological data modeling. A positive correlation in the case of slope index is indicative that basins with steeper topography generally experience quick recovery, hence high recoveriness with greater maximum flow rates and quicker decline times. Similarly, the basin area shows an inverse trend of a similar magnitude. The plot also explains the fact that basins with smaller catchment areas are expected to recover faster. Though the quantile curves provide valuable information about the relationship recoveriness exhibits with individual predictors, flood processes in reality result from complex interactions between a large number of variables. Thus, we adopt a multidimensional modeling approach using random forest quantile regressor to uncover the collective and individual impact of this large number of predictors on recoveriness.

b. *Multidimensional modeling of recoveriness*

The RFQR model consists of a total of 10 decision trees—Fig. 7 shows two participating decision trees that can vote for the outcome during prediction. We can see that the slope index (si) variable happens to occupy the root node making it the most important variable. The random forest algorithm evaluates this predictor's values to be most informative in forming the bifurcation at the root level. The threshold value of "si" is computed to be 0.0. The "true" branch leads to the "area" node, and the "false" branch leads to the precipitation node. Upon analysis, it was found that 9 out of 10 decision trees preferred the structure shown in Fig. 7a, helping it to discover the relationship (nonlinear) among the different predictors.

The deeper branches still model more complex relationships among the predictors. We only show the first three levels of the trees here.

Figure 7b shows a decision tree structure that was preferred just once out of 10 trees in the forest. We learn from the figures that area, precipitation, and river length are second in importance (revealed by their occurrence at the second level). At the third level, two new predictors are included—the

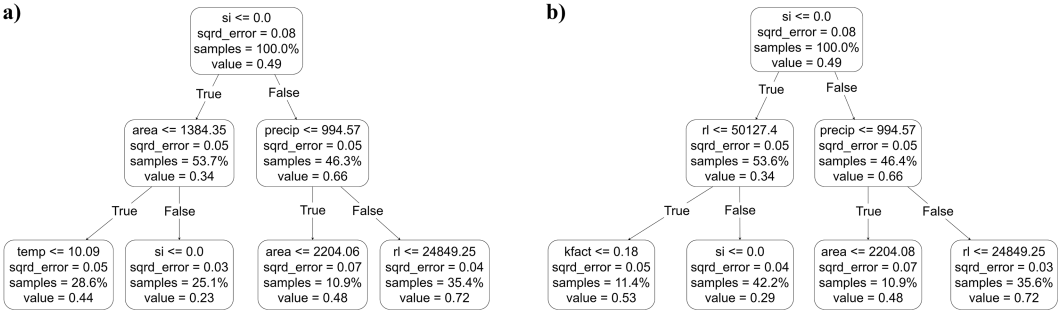

FIG. 7. Nonlinear dependence of recoveriness on predictors: The RF consisted of 10 decision trees. The bifurcation in the predictor space tends to follow the above two structures: (a) This structure occurs 9/10 times, and (b) this structure occurs 1/10 times, showing that slope index (si) is the most important variable followed by area, precipitation, and river length. The depth of the trees has been truncated to level 3 although deeper layers were involved in the model.

temperature and the *k* factor, making them the third most important variables. The slope index, area, and river also accompany them, again signifying their importance in the model.

The R^2 achieved by the prediction algorithm was >88% for all the testing sessions. Multiple training and testing sessions helped check the robustness of the algorithm, and it was found that the accuracy was consistent at 88% or more. The unchanging decision tree structures throughout the testing sessions further attested to the dependency among the predictor variables. Table 3 shows the result of training and testing of the RFQR on the dataset.

c. *SHAP interpretation and predictor importance*

In Fig. 8, we see a bee swarm plot showing the strength of various input predictors in predicting the output variable. The *y* axis lists the different predictors, while the corresponding area in the graph shows how strongly (via SHAP values), the predictor contributed toward the predicted target. For every predictor, we have 1000 dots corresponding to the 1000 test data points. The adjoining color bar gives a qualitative description of the magnitude of the target (recoveriness) variable, and the *x* axis gives the numerical value (SHAP) of the input predictor. It can be seen that slope index (si) has a wider spread and occurs at the top in the *y* axis. The spread indicates that this predictor is more informative of the target. Toward the bottom, in the axis, is the basin texture which is least informative as evident from its concentrated occurrences for all the 1000 test data points.

Figure 8 gives us the idea of the predictors in decreasing order of informativeness. The contribution of basin texture input is negligible with the rock volume being just marginally better than the previous. The temperature, area, and river length are the second most important input predictors after the slope index in predicting recoveriness.

TABLE 3. Training and testing results.

	MSE	R^2
Training	0.010	0.866
Testing	0.011	0.865

A grid search of hyperparameters was done for the random forest algorithm. The result of the best-performing model has been reported in this section. The trained model is used to make predictions of recoveriness over the CONUS. The three hyperparameters of the regression algorithm were as follows: 1) max_depth, which determines the maximum levels allowed during the bifurcation of the decision space; 2) max_features, a parameter whose value must be less than the total number of predictors, which determines the maximum number of predictors that can be used for building the bootstrap dataset; and 3) number_of_trees, which determines the total number of weak estimators to build the ensemble model. The max_depth hyperparameter was specified to restrict the decision tree formation with only seven levels (half the total predictors) in the

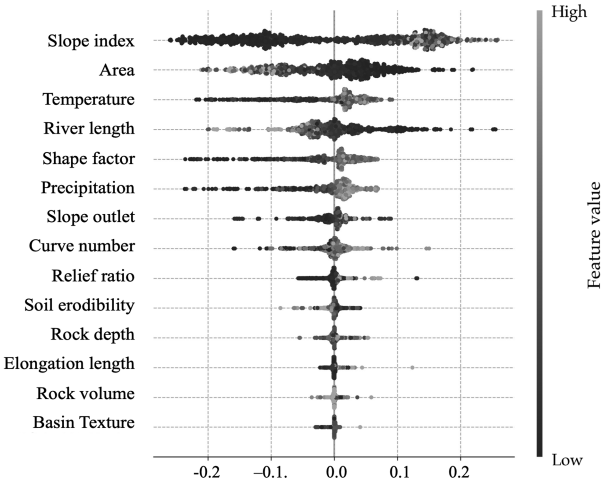

FIG. 8. Bee plot showing the strength of different predictors in predicting 1000 random values of recoveriness. Each predictor in the *y* axis has 1000 dots in the graph area. The color bar shows the SHAP value of the predictor used to predict the target value to achieve an overall R^2 of 89%. Again, the slope index (si) happens to be the most important variable followed by temperature, area, and river length.

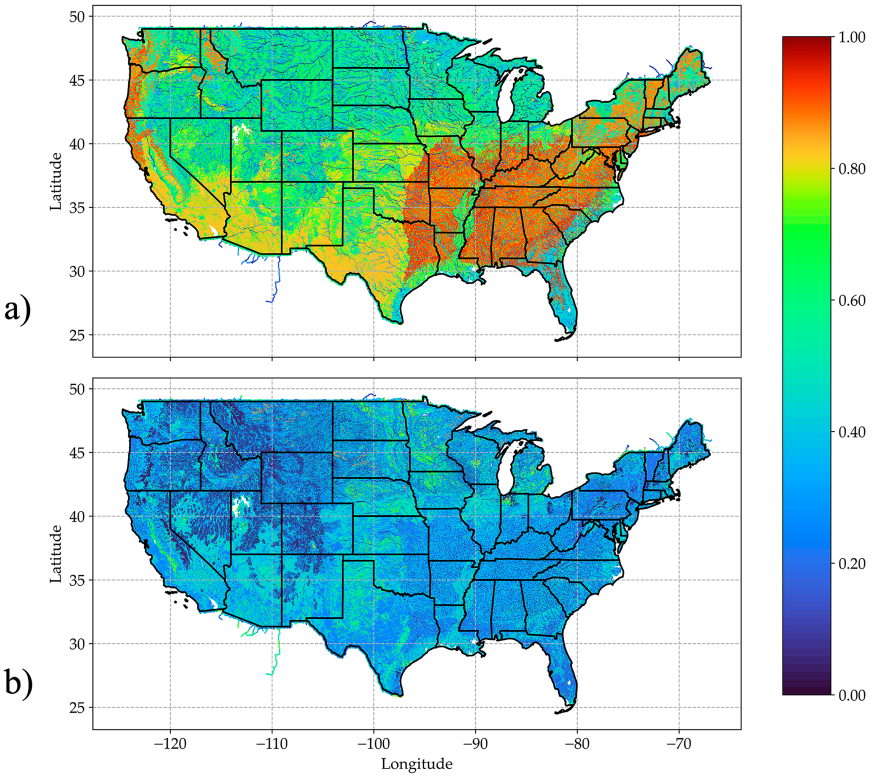

FIG. 9. (a) High-density prediction of recoveriness across the United States. The red regions mark the areas with short recovery times, while the green and blue regions indicate areas with longer recovery times. (b) 95% prediction interval width for quantification of the uncertainty in predicted recoveriness: The blueish parts indicate regions with high uncertainty; the darker parts include more precise predictions.

hierarchy. The max_features is set to the total number of predictors, i.e., 14, effectively allowing the algorithm to consider all the predictors but with only seven levels of complexity in relationship. The tree plot in Fig. 7 shows only the first three levels of the hierarchy. The number_of_trees is fixed at value 10 so as not to promote overcomplicated models.

The spatial resolution for the prediction dataset used to estimate recoveriness over the United States is 1 km. A plot of the predicted recoveriness is shown in Fig. 9. Recoveriness values range from 0 to 1. Regions of high recoveriness are identified as red; the green and blue regions indicate regions of low recovery (bad). The southern part in the map shows high values of recoveriness. Referring to Fig. 2, the regions with high recovery can be identified as the Appalachians, the Missouri Valley, and the Flash Flood Alley. The northern parts of Arizona and the Front Range are mostly characterized by low recovery. The eastern parts of the West Coast show similarly low recovery regions.

Figure 9b shows the associated uncertainty associated with the estimated recoveriness. The uncertainty is defined as the "prediction interval width," where the width is given using Eq. (5):

$$\text{Predicted interval width} = 95\%\ \text{quantile} - 5\%\ \text{quantile}. \tag{5}$$

Uncertainty is indicated by the bottom plot in Fig. 9b. The lighter regions indicate high uncertainty; in other words, the predicted interval of recoveriness values is wider in those regions. The uncertainty does not seem to show any pattern except some dark regions at the center and southwest direction. Both the plots (Figs. 9a and 9b) help us to identify the hotspots of high and low recoveriness in ungauged regions, inclusive of the associated prediction uncertainty.

High recovery regions show specific patterns—clinging to the western shoreline and running along the inlands of Sierra Nevada are situated in the highlands, consistent with Fig. 2. The same potential for recoveriness can be associated with Arizona. High recovery potential can also be observed in the region extending from the southeast Arizona to the Mogollon Rim.

The predictions in the map clearly indicate high recoveriness zones extending from southwest Texas to Kansas, Arkansas, Oklahoma, and Missouri. Worth noting in the predictions is the revelation of high recovery regions of the western slopes of the Appalachians consisting of Kentucky, Tennessee, and West Virginia.

d. Correlation between flashiness and recoveriness

Figure 10a displays the relationship between recoveriness and flashiness, whose mapping consistency over the United

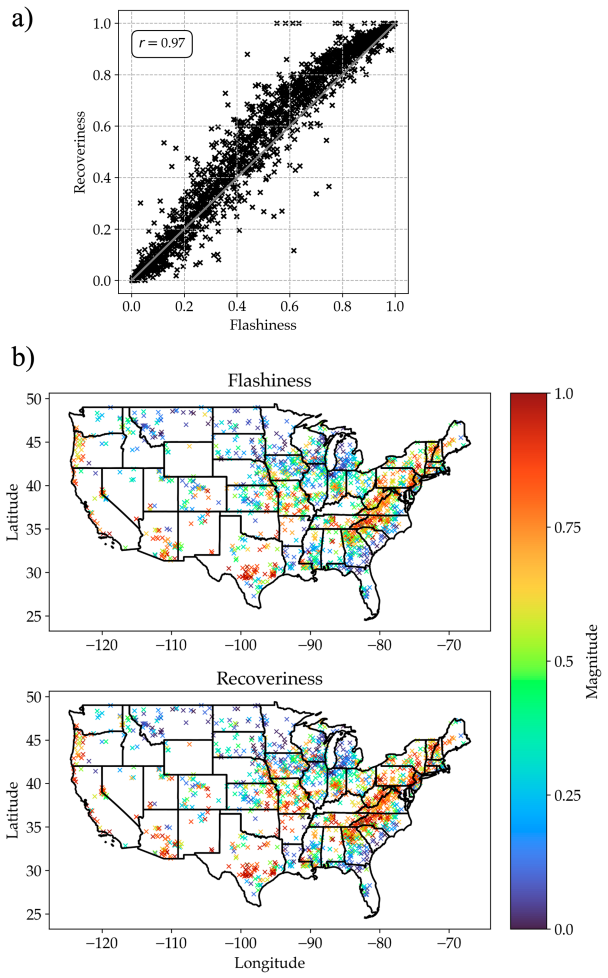

FIG. 10. (a) A plot showing the correlation between flashiness and recoveriness. The relation is nonlinear although there exists a high degree of correlation as indicated by the high Pearson correlation coefficient and the r value; (b) spatial distribution of flashiness and recoveriness over the CONUS—high magnitude indicates high flashiness and quick recovery times and vice versa.

States has been indirectly validated by Saharia et al. (2017b). One can note that the two variables display a strong relationship, suggesting that similar hydrologic processes are at stake for the two variables.

For flashiness values less than 0.2, there is a linear rise in recoveriness. For higher values greater than 0.2, a second-degree polynomial would be a better fit. Fitting that gives the following equation:

$$y = -0.34x^2 + 1.38x - 0.03, \quad (6)$$

where y represents recoveriness and x represents flashiness. The equation indicates a weakly nonlinear relationship between the two. Figure 10b gives the spatial plot of recoveriness, where higher values indicate quick recovery, while lower values are indicative of delayed recovery. Basins that can quickly concentrate flows and result in high flashiness are also associated with high recoveriness. The consistency of the flashiness mapping over the United States has been qualitatively established in Saharia et al. (2017b).

5. Conclusions

This study focuses on a recession period of a flood, a less explored area of research as compared to the rising flood period. The dataset used in this study contains a number of climatological and geomorphological information along with details of flooding events spanning 78 years from the gauged region. The dependency of geomorphological variables and climatology on recovery period has been explored in this work following which a new variable called recoveriness is introduced in this paper, which measures the flood recovery times and is a combination of peak and recession period. We have found that recoveriness is a highly nonlinear variable and bears a complex relationship with the geomorphological and climatological predictors. The results are supported by suitable machine learning models that predict recoveriness at every ungauged location across the CONUS. The main findings of this study are as follows:

1) There exists a high correlation between the flood recovery times and the basin properties and topology. For example, the slope of the basin happens to be highly correlative with recovery, and this is quite reasonable acknowledging the fact that steep slopes tend to cause faster runoffs needing lesser recovery times. Similarly, basins with a smaller area would require less time because there is less water to disperse.
2) Six hotspots with high recoveriness spanned the United States: Arizona, the Front Range, the West Coast, the Missouri Valley, the Flash Flood Alley, and the Appalachian Mountains.
3) Machine learning model was used to predict the recoveriness values using highly dense data containing ungauged region over the CONUS. Several new localized hotspots were identified: the western slopes of the Appalachians consisting of Kentucky, Tennessee, and West Virginia and the interlinked areas of eastern Montana and northern Idaho.

This paper develops a comprehensive climatological perspective on flood recovery. In future work, we see the potential to establish an event-level recoveriness metric that could facilitate a deeper understanding of the postflood recovery process. Overall, this new metric has the potential to be adopted for disaster risk reduction efforts.

Acknowledgments. This research was conducted in the HydroSense Lab (https://hydrosense.iitd.ac.in/) of IIT Delhi, and the authors acknowledge the IIT Delhi High Performance Computing facility for providing computational and storage resources. Dr. Manabendra Saharia gratefully acknowledges financial support for this work through grants from ISRO Space Technology Cell (STC0374/RP04139), MoES Monsoon Mission III (RP04574), and IC-IMPACTS (RP04558). The authors declare that they have no known

competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement. The data that support the findings of this study are openly available in the repository titled "Mapping a Novel Metric for Flash Flood Recovery Using Interpretable Machine Learning" at https://zenodo.org/doi/10.5281/zenodo.10090794.

REFERENCES

Ahmad, A., A. El-Shafie, S. F. M. Razali, and Z. S. Mohamad, 2014: Reservoir optimization in water resources: A review. *Water Resour. Manage.*, **28**, 3391–3405, https://doi.org/10.1007/s11269-014-0700-5.

Alfieri, L., and J. Thielen, 2015: A European precipitation index for extreme rain-storm and flash flood early warning. *Meteor. Appl.*, **22**, 3–13, https://doi.org/10.1002/met.1328.

Amaya, M., N. Linde, and E. Laloy, 2022: Hydrogeological multiple-point statistics inversion by adaptive sequential Monte Carlo. *Adv. Water Resour.*, **166**, 104252, https://doi.org/10.1016/j.advwatres.2022.104252.

Amit, H., V. Lyakhovsky, A. Katz, A. Starinsky, and A. Burg, 2002: Interpretation of spring recession curves. *Groundwater*, **40**, 543–551, https://doi.org/10.1111/j.1745-6584.2002.tb02539.x.

Ao, Y., H. Li, L. Zhu, S. Ali, and Z. Yang, 2019: The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. *J. Pet. Sci. Eng.*, **174**, 776–789, https://doi.org/10.1016/j.petrol.2018.11.067.

Babagoli, M., M. P. Aghababa, and V. Solouk, 2019: Heuristic nonlinear regression strategy for detecting phishing websites. *Soft Comput.*, **23**, 4315–4327, https://doi.org/10.1007/s00500-018-3084-2.

Baker, D. B., R. P. Richards, T. T. Loftus, and J. W. Kramer, 2004: A new flashiness index: Characteristics and applications to midwestern rivers and streams. *J. Amer. Water Resour. Assoc.*, **40**, 503–522, https://doi.org/10.1111/j.1752-1688.2004.tb01046.x.

Berz, G., 2000: Flood disasters: Lessons from the past—Worries for the future. *Proc. Inst. Civil Eng.: Water Marit. Eng.*, **142**, 3–8, https://doi.org/10.1680/wame.2000.142.1.3.

Bhaskar, N. R., M. N. French, and G. K. Kyiamah, 2000: Characterization of flash floods in eastern Kentucky. *J. Hydrol. Eng.*, **5**, 327–331, https://doi.org/10.1061/(ASCE)1084-0699(2000)5:3(327).

Biswal, B., and M. Marani, 2010: Geomorphological origin of recession curves. *Geophys. Res. Lett.*, **37**, L24403, https://doi.org/10.1029/2010GL045415.

Buschjäger, S., P.-J. Honysz, and K. Morik, 2022: Randomized outlier detection with trees. *Int. J. Data Sci. Anal.*, **13**, 91–104, https://doi.org/10.1007/s41060-020-00238-w.

Chang, C., and P. Feng, 2017: The impact of land use/land cover changes and hydraulic structures on flood recession process. *J. Water Climate Change*, **8**, 375–387, https://doi.org/10.2166/wcc.2017.100.

Chapman, T., 1999: A comparison of algorithms for stream flow recession and baseflow separation. *Hydrol. Processes*, **13**, 701–714, https://doi.org/10.1002/(SICI)1099-1085(19990415)13:5<701::AID-HYP774>3.0.CO;2-2.

Costa, J. E., 1987: A comparison of the largest rainfall-runoff floods in the United States with those of the People's Republic of China and the world. *J. Hydrol.*, **96**, 101–115, https://doi.org/10.1016/0022-1694(87)90146-6.

Costa, M. H., A. Botta, and J. A. Cardille, 2003: Effects of large-scale changes in land cover on the discharge of the Tocantins River, southeastern Amazonia. *J. Hydrol.*, **283**, 206–217, https://doi.org/10.1016/S0022-1694(03)00267-1.

Diakakis, M., G. Deligiannakis, Z. Antoniadis, M. Melaki, N. K. Katsetsiadou, E. Andreadakis, N. I. Spyrou, and M. Gogou, 2020: Proposal of a flash flood impact severity scale for the classification and mapping of flash flood impacts. *J. Hydrol.*, **590**, 125452, https://doi.org/10.1016/j.jhydrol.2020.125452.

Douben, K.-J., 2006: Characteristics of river floods and flooding: A global overview, 1985–2003. *Irrig. Drain.*, **55**, S9–S21, https://doi.org/10.1002/ird.239.

Ekmekcioğlu, Ö., and K. Koc, 2022: Explainable step-wise binary classification for the susceptibility assessment of geo-hydrological hazards. *Catena*, **216**, 106379, https://doi.org/10.1016/j.catena.2022.106379.

Gimeno, M., K. Sada del Real, and A. Rubio, 2023: Precision oncology: A review to assess interpretability in several explainable methods. *Briefings Bioinf.*, **24**, bbad200, https://doi.org/10.1093/bib/bbad200.

Gourley, J. J., and Coauthors, 2013: A unified flash flood database across the United States. *Bull. Amer. Meteor. Soc.*, **94**, 799–805, https://doi.org/10.1175/BAMS-D-12-00198.1.

Huan, J., H. Li, M. Li, and B. Chen, 2020: Prediction of dissolved oxygen in aquaculture based on gradient boosting decision tree and long short-term memory network: A study of Chang Zhou fishery demonstration base, China. *Comput. Electron. Agric.*, **175**, 105530, https://doi.org/10.1016/j.compag.2020.105530.

Islam, S. R., W. Eberle, and S. K. Ghafoor, 2020: Towards quantification of explainability in explainable artificial intelligence methods. *Proc. 33 Int. Florida Artificial Intelligence Research Society Conf.*, North Miami Beach, FL, AAAI Press, 75–81, https://aaai.org/ocs/index.php/FLAIRS/FLAIRS20/paper/view/18410.

John, G. H., 1995: Robust decision trees: Removing outliers from databases. *KDD-95 Proc. First Int. Conf. on Knowledge Discovery and Data Mining*, Vol. 95, Montreal, QC, Canada, AAAI Press, 174–179, https://dl.acm.org/doi/10.5555/3001335.3001364.

Joshi, R. M., A. Kumar, and K. H. Singh, 2023: A simulation and data informed approach to porosity partitioning. *Geoenergy Sci. Eng.*, **229**, 212044, https://doi.org/10.1016/j.geoen.2023.212044.

Khaleghi, M. R., V. Gholami, J. Ghodusi, and H. Hosseini, 2011: Efficiency of the geomorphologic instantaneous unit hydrograph method in flood hydrograph simulation. *Catena*, **87**, 163–171, https://doi.org/10.1016/j.catena.2011.04.005.

Kumar, A., R. K. Shrivastava, and K. H. Singh, 2020: Bayesian inference of material properties in disordered media using sound characteristics. *Europhys. Lett.*, **129**, 24001, https://doi.org/10.1209/0295-5075/129/24001.

Kuntla, S. K., M. Saharia, and P. Kirstetter, 2022: Global-scale characterization of streamflow extremes. *J. Hydrol.*, **615**, 128668, https://doi.org/10.1016/j.jhydrol.2022.128668.

Lee, E. H., and J. H. Kim, 2017: Development of resilience index based on flooding damage in urban areas. *Water*, **9**, 428, https://doi.org/10.3390/w9060428.

Li, Z., and Coauthors, 2023a: Distributed flashiness-intensity-duration-frequency products over the conterminous US. ESS

Open Archive, https://doi.org/10.22541/essoar.169272223.38642667/v1.

——, S. Gao, M. Chen, J. Zhang, J. J. Gourley, Y. Wen, T. Yang, and Y. Hong, 2023b: Introducing Flashiness-Intensity-Duration-Frequency (F-IDF): A new metric to quantify flash flood intensity. *Geophys. Res. Lett.*, **50**, e2023GL104992, https://doi.org/10.1029/2023GL104992.

Liu, P., L. Li, S. Guo, L. Xiong, W. Zhang, J. Zhang, and C.-Y. Xu, 2015: Optimal design of seasonal flood limited water levels and its application for the Three Gorges Reservoir. *J. Hydrol.*, **527**, 1045–1053, https://doi.org/10.1016/j.jhydrol.2015.05.055.

Lundberg, S. M., and S.-I. Lee, 2017: A unified approach to interpreting model predictions. *NIPS'17: Proc. 31st Int. Conf. on Neural Information Processing Systems*, Long Beach, CA, Curran Associates Inc., 4768–4777, https://dl.acm.org/doi/10.5555/3295222.3295230.

Madhushani, C., K. Dananjaya, I. U. Ekanayake, D. P. P. Meddage, K. Kantamaneni, and U. Rathnayake, 2024: Modeling stream-flow in non-gauged watersheds with sparse data considering physiographic, dynamic climate, and anthropogenic factors using explainable soft computing techniques. *J. Hydrol.*, **631**, 130846, https://doi.org/10.1016/j.jhydrol.2024.130846.

Marcuzzi, F., C. Lucchese, and S. Orlando, 2022: Filtering out outliers in learning to rank. *Proc. 2022 ACM SIGIR Int. Conf. on Theory of Information Retrieval*, Madrid, Spain, Association for Computing Machinery, 214–222, https://dl.acm.org/doi/10.1145/3539813.3545127.

Meinshausen, N., and G. Ridgeway, 2006: Quantile regression forests. *J. Mach. Learn. Res.*, **7**, 983–999.

Mugume, S. N., D. E. Gomez, G. Fu, R. Farmani, and D. Butler, 2015: A global analysis approach for investigating structural resilience in urban drainage systems. *Water Res.*, **81**, 15–26, https://doi.org/10.1016/j.watres.2015.05.030.

Nanfack, G., P. Temple, and B. Frénay, 2022: Constraint enforcement on decision trees: A survey. *ACM Comput. Surv.*, **54** (10s), 1–36, https://doi.org/10.1145/3506734.

Nathan, R. J., and T. A. McMahon, 1990: Evaluation of automated techniques for base flow and recession analyses. *Water Resour. Res.*, **26**, 1465–1473, https://doi.org/10.1029/WR026i007p01465.

Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. G. Kolodziej, and J. J. Gourley, 2009: The severe hazards analysis and verification experiment. *Bull. Amer. Meteor. Soc.*, **90**, 1519–1530, https://doi.org/10.1175/2009BAMS2815.1.

Panjei, E., L. Gruenwald, E. Leal, C. Nguyen, and S. Silvia, 2022: A survey on outlier explanations. *VLDB J.*, **31**, 977–1008, https://doi.org/10.1007/s00778-021-00721-1.

Pielke, R. A., Jr., and M. W. Downton, 2000: Precipitation and damaging floods: Trends in the United States, 1932–97. *J. Climate*, **13**, 3625–3637, https://doi.org/10.1175/1520-0442(2000)013<3625:PADFTI>2.0.CO;2.

Potdar, A. S., P.-E. Kirstetter, D. Woods, and M. Saharia, 2021: Toward predicting flood event peak discharge in ungauged basins by learning universal hydrological behaviors with machine learning. *J. Hydrometeor.*, **22**, 2971–2982, https://doi.org/10.1175/JHM-D-20-0302.1.

Pradhan, B., S. Lee, A. Dikshit, and H. Kim, 2023: Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model. *Geosci. Front.*, **14**, 101625, https://doi.org/10.1016/j.gsf.2023.101625.

Rentschler, J., M. Salhab, and B. A. Jafino, 2022: Flood exposure and poverty in 188 countries. *Nat. Commun.*, **13**, 3527, https://doi.org/10.1038/s41467-022-30727-4.

Saharia, M., P.-E. Kirstetter, H. Vergara, J. J. Gourley, and Y. Hong, 2017a: Characterization of floods in the United States. *J. Hydrol.*, **548**, 524–535, https://doi.org/10.1016/j.jhydrol.2017.03.010.

——, ——, ——, ——, ——, and M. Giroud, 2017b: Mapping flash flood severity in the United States. *J. Hydrometeor.*, **18**, 397–411, https://doi.org/10.1175/JHM-D-16-0082.1.

——, ——, ——, ——, I. Emmanuel, and H. Andrieu, 2021: On the impact of rainfall spatial variability, geomorphology, and climatology on flash floods. *Water Resour. Res.*, **57**, e2020WR029124, https://doi.org/10.1029/2020WR029124.

Shixiang, F., and H. Shaowen, 1991: Testing research on the effects of land surface slopes upon surface runoff. *Bull. Soil Water Conserv.*, **11**, 6–10.

Shorr, N., 2000: Early utilization of flood-recession soils as a response to the intensification of fishing and upland agriculture: Resource-use dynamics in a large Tikuna community. *Hum. Ecol.*, **28**, 73–107, https://doi.org/10.1023/A:1007079721915.

Shuster, W. D., Y. Zhang, A. H. Roy, F. B. Daniel, and M. Troyer, 2008: Characterizing storm hydrograph rise and fall dynamics with stream stage data. *J. Amer. Water Resour. Assoc.*, **44**, 1431–1440, https://doi.org/10.1111/j.1752-1688.2008.00249.x.

Wang, Y., C. Zhang, A. S. Chen, G. Wang, and G. Fu, 2023: Exploring the relationship between urban flood risk and resilience at a high-resolution grid cell scale. *Sci. Total Environ.*, **893**, 164852, https://doi.org/10.1016/j.scitotenv.2023.164852.

Yan, B., F. Ren, M. Cai, and C. Qiao, 2020: Bayesian model based on Markov chain Monte Carlo for identifying mine water sources in submarine gold mining. *J. Cleaner Prod.*, **253**, 120008, https://doi.org/10.1016/j.jclepro.2020.120008.

Yang, Y., and T. F. M. Chui, 2021: Modeling and interpreting hydrological responses of sustainable urban drainage systems with explainable machine learning methods. *Hydrol. Earth Syst. Sci.*, **25**, 5839–5858, https://doi.org/10.5194/hess-25-5839-2021.

Ye, R., Y. He, S. Yu, and Z. Song, 2019: Effects of recent morphodynamic evolution on flood regimes in the Pearl River Delta. *Nat. Hazards*, **96**, 1091–1119, https://doi.org/10.1007/s11069-019-03592-6.