

## Spatial-temporal characteristics and driving factors of flash floods in Shaanxi Province considering regional differentiation

Han Zhang, Jungang Luo\*, Jingyan Wu and Mengjie Yu

State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Xi'an University of Technology, Xi'an, Shaanxi 710048, China

\*Corresponding author. E-mail: jgluo@xaut.edu.cn

### ABSTRACT

Flash floods show strong regional differentiation in spatial-temporal distribution and driving forces, thereby hindering their effective prevention and control. This study analyzed the spatiotemporal characteristics of flash floods in Shaanxi Province, China, differentiated among the northern Shaanxi (NS), Guanzhong (GZ), and southern Shaanxi (SS) regions based on the Mann-Kendall, Theil-Sen Median, and standard deviation ellipse methods. The main factors driving disasters and their interactions in each region were then identified within the three categories of precipitation factor (PPF), surface environment factor, and human activity factor (HAF) based on a geographical detector. Finally, the differences in flash flood characteristics among the NS, GZ, and SS regions were analyzed. The results showed that flash floods in Shaanxi Province are greatly affected by the PPF and the HAF, although the spatial-temporal characteristics and disaster-causing factors were significantly different in each region. The regions were ranked according to the number and growth trends of flash floods as follows: SS > GZ > NS. Furthermore, flash floods were affected by multiple factors, with the interaction between factors acting as a driving force of flash floods. The results of this study can provide a reference for the management of flash floods under regional differentiation.

**Key words:** driving analysis, flash floods, geographical detector, regional differentiation, spatial-temporal characteristics

### HIGHLIGHTS

- Taking Shaanxi Province as a typical region of regional differentiation, the temporal and spatial characteristics of flash floods are evaluated. The main disaster-causing factors and their interaction of flash floods in each region are discussed. Results are helpful to put forward the flash flood prevention and control measures according to local conditions in the same administrative region and reduce the potential risks.

### INTRODUCTION

Flash floods are considered to be the type of natural disasters of extremely high risk due to their characteristics of sudden occurrence, rapid formation, and strong destructiveness. Many countries globally are affected by flash floods (Trigo *et al.* 2016; Ngo *et al.* 2021). Over 60,000 flash flood disasters have been reported in China since 1949, which resulted in great casualties and economic losses (Liu *et al.* 2018). Flash floods have strong geographical differentiation, with their characteristics and disaster-causing factors varying from region to region (Hou *et al.* 2020; Xiong *et al.* 2020). Therefore, revealing the characteristics of flash floods in different regions is of great significance for preventing and mitigating hazards resulting from flash floods.

Flash floods are affected by multiple factors, including precipitation, surface environment, and human activities (Arora *et al.* 2021). Quantitatively and accurately revealing the characteristics and drivers of flash floods present a great challenge due to regional differentiation (Adnan *et al.* 2019). The analysis of the spatial and temporal variation and driving factors of flash floods is the basis for scientific understanding of the distribution and mechanism of the formation of flash floods and is further necessary for the management of flash floods under regional differentiation.

The Mann-Kendall (M-K) test is widely used to explore temporal and spatial distribution characteristics and is used in trend analysis and mutation point tests due to its insensitivity to abnormal values (Seenu & Jayakumar 2021). The Theil-Sen Median (Sen) estimation method is also a commonly used nonparametric trend test method. The combined M-K and Sen (MK + Sen) test is an important method used to assess the trend of a time series since this method is insensitive to data errors and has a solid statistical theoretical basis for testing the significance level (Phuong *et al.* 2020; Shi *et al.* 2021). In addition, the standard deviation ellipse (SDE) is an important method for reflecting spatial characteristics, such as the central

trend, discrete trend, and directional trend (Yang *et al.* 2021). Shi *et al.* (2018) used the SDE and MK + Sen to analyze the spatiotemporal characteristics of PM<sub>2.5</sub> (particulate matter with a diameter of  $\leq 2.5 \mu\text{m}$ ) from 1999 to 2014. Wang *et al.* (2021) used the transfer matrix in combination with the SDE and spatial autocorrelation analysis to analyze the influence of land use and land cover on the spatiotemporal pattern of precipitation.

The geographical detector is a new statistical method used to reveal the driving forces of a certain spatial distribution. The method was first proposed by Wang Jinfeng and can detect both numerical and qualitative data (Wang *et al.* 2016). Using the geographical detector, the power of a single independent variable to explain the variation in a dependent variable as well as the power of the coupling of the spatial distributions of two independent variables to explain the variation in the dependent variable can be detected. Recent studies have used geographical detectors to analyze the driving forces of flash floods (Xiong *et al.* 2020).

Realizing the effective prevention and control of flash floods under regional differentiation in the same administrative region requires certain prerequisites (Singh *et al.* 2008). These include analyzing the changes in the properties of flash floods before and after sudden change, identifying changes in the intensity and evolution of flash floods at different spatial scales, and detecting the driving forces of flash floods and the synergistic effects between them. Therefore, an in-depth analysis of the temporal and spatial distributions of flash floods under regional differentiation and the quantification of disaster-causing factors are necessary.

Shaanxi Province, China, shows a strong regional differentiation in climatic conditions, with a semi-arid area in northern Shaanxi (NS), a semi-humid area in Guanzhong (GZ), and a humid area in southern Shaanxi (SS). This present study applied the M-K, MK + Sen, and SDE methods to the historical data of flash flood disaster events in Shaanxi Province from 1949 to 2015 to analyze the spatial and temporal distributions of flash floods in the NS, GZ, and SS regions. Second, the present study quantified the influences of the precipitation factor (PPF), the surface environment factor (SEF), and the human activity factor (HAF) and their interactions on flash floods in the NS, GZ, and SS regions using the geographical detector. Finally, the differences in flash flood disaster characteristics and the explanatory power of different factors under regional differentiation were analyzed. Figure 1 shows the research framework followed under the current study. The results of the present study can provide a reference for the prevention and mitigation of flash floods in Shaanxi Province.

## METHODOLOGY

### M-K statistical test

The M-K statistical test is applicable for both trend analysis and assessing mutation in a time series. Since the M-K is insensitive to abnormal values and human factors, it has been widely used in time-series analysis within hydrology and meteorology (Nyikadzino *et al.* 2020). The M-K test can be implemented through the following steps:

**Step 1:** Constructing a order sequence of time series for  $k$  samples:

$$S_k = \sum_{i=1}^k \sum_{j=i+1}^{k-1} \text{sign}(X_i - X_j) \quad (1)$$

$$\text{sign}(X_i - X_j) = \begin{cases} 1 & X_i > X_j \\ 0 & X_i = X_j \\ -1 & X_i < X_j \end{cases} \quad (2)$$

In Equations (1) and (2),  $\text{sign}$  is the express sign function,  $X_i$  is the express random variable and is composed of  $i$  elements ( $i = 1, i + 2, \dots, k$ ), and  $X_j$  represents other random variables comprising  $j$  elements ( $j = 2, \dots, k - 1$ ).

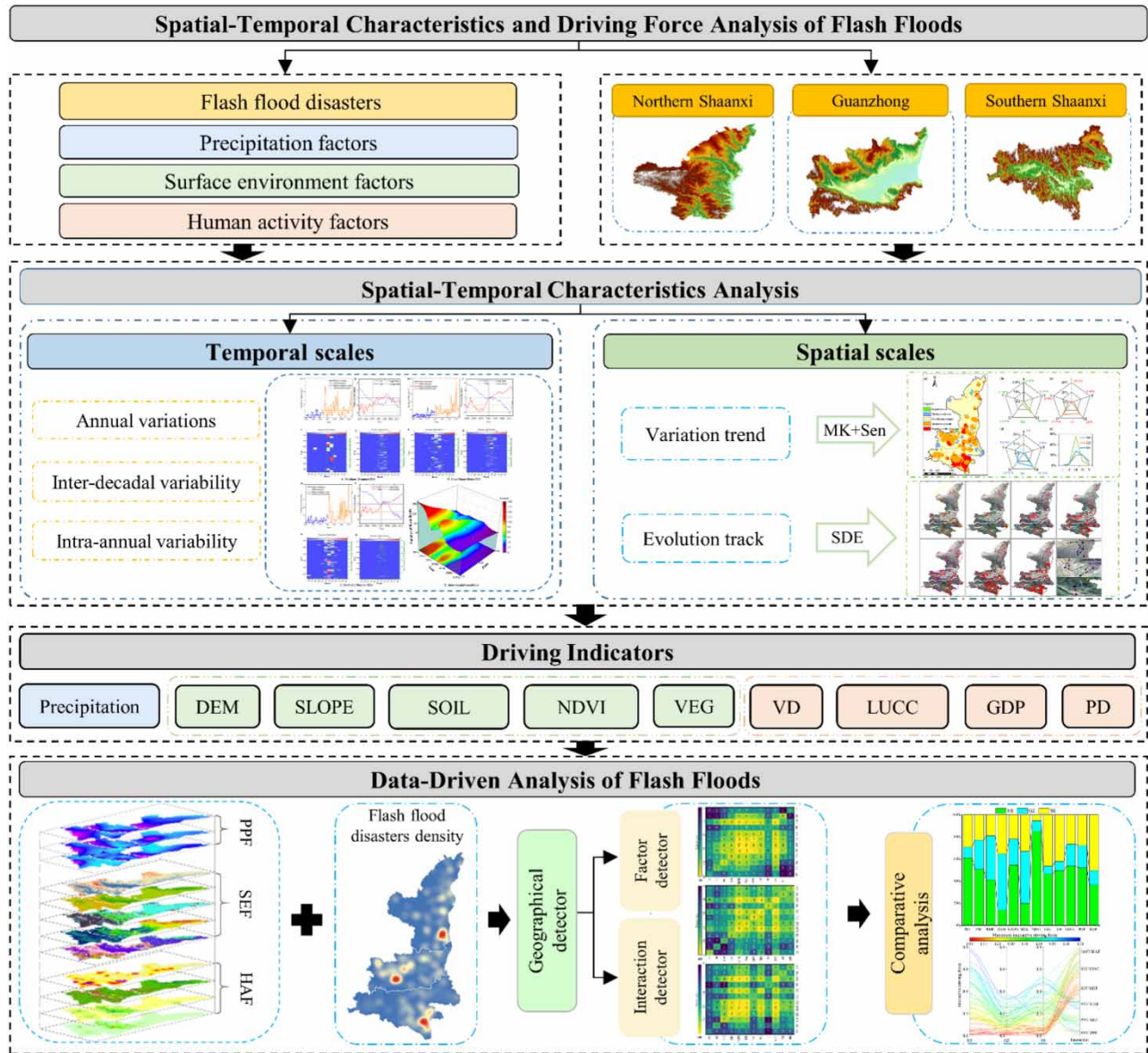
**Step 2:** Assuming a random time series, the definition of  $UF_k$  is:

$$UF_k = \frac{|S_k - E(S_k)|}{\sqrt{\text{Var}(S_k)}} \quad (3)$$

$$E(S_k) = \frac{k(k-1)}{4} \quad (4)$$

$$\text{Var}(S_k) = \frac{k(k-1)(2k+5)}{72}, \quad (k = 2, 3, \dots, n) \quad (5)$$

In Equations (3)–(5),  $E(S_k)$  and  $\text{Var}(S_k)$  represent the mean and variance of  $S_k$ .



**Figure 1** | Research framework of the present study.

**Step 3:** Composing the value of  $UF_k$  into a curve  $UF$ .

The processes outlined in **step 2** are repeated for the inverse sequence, and the curve  $UB$  can be obtained with  $UF = -UB$ . The  $UF$  and  $UB$  refer to the coordinate system. The testing of trends is conducted at the specific  $t_0$  significance level ( $t_0 = 95\%$  was chosen in the present study, the null hypothesis of no trend was rejected if  $|UF_k| > 1.96$ ). A curve  $UF$  within  $t_0$  indicated a significant trend in the time series, with  $UF > 0$  and  $UF < 0$  indicating an upward and downward trend in the time series, respectively. The intersection of the curves  $UF$  and  $UB$  in the confidence interval indicated a possible mutation point.

### Theil-Sen Median

The Theil-Sen Median trend analysis (Sen) is a robust nonparametric trend calculation method, which can reduce the influence of data outliers. Sen assesses the trend and degree of the sequence by calculating the median of the slope of the data combination. Sen can be expressed using the following formula (Zhang *et al.* 2021):

$$S = \text{Median}\left(\frac{x_k - x_i}{k - i}\right), \quad k > i \quad (6)$$

In Equation (6),  $S$  is the variation trend of the flash flood,  $x_k$  and  $x_i$  are the sequence values for  $k$  and  $i$  periods, respectively, and  $\text{Median}(\cdot)$  is the median function.  $S > 0$  and  $S < 0$  indicate an upward and downward trend in the time series, respectively.

The M-K test can be combined with the Sen (MK + Sen) to assess the trend of time series, with this approach often used in hydrology and meteorology (Esmailpour *et al.* 2021; Syed *et al.* 2021). The present study used the MK + Sen to detect the variations in flash floods over time.

### SDE analysis

The SDE analysis can express the spatial characteristics of a variable by calculating the standard distance from a set discrete point to the average (Peng *et al.* 2016; Shi *et al.* 2018). Therefore, the present study used the SDE to identify the change in the center position and the moving trend of flash floods. The specific expression is:

$$\text{SDE}_x = \sqrt{\frac{\sum_{i=1}^n (x_i - x_{mc})^2}{n}} \quad (7)$$

$$\text{SDE}_y = \sqrt{\frac{\sum_{i=1}^n (y_i - y_{mc})^2}{n}} \quad (8)$$

In Equations (7) and (8),  $(x_i, y_i)$  is the coordinate of variable  $i$ ,  $(x_{mc}, y_{mc})$  is the arithmetic mean center of the variable, and  $n$  is equal to the total number of variables.  $(\text{SDE}_x, \text{SDE}_y)$  is the coordinate of the center of the ellipse.

The angle of rotation is calculated as follows:

$$\tan \theta = \frac{A + B}{C} \quad (9)$$

$$A = \left( \sum_{i=1}^n xx_i^2 - \sum_{i=1}^n yy_i^2 \right) \quad (10)$$

$$B = \sqrt{\left( \sum_{i=1}^n xx_i^2 - \sum_{i=1}^n yy_i^2 \right)^2 + 4(xx_iyy_i)^2} \quad (11)$$

$$C = 2 \sum_{i=1}^n xx_iyy_i \quad (12)$$

In Equations (9)–(12),  $xx_i$  and  $yy_i$  are the deviations from the  $x$ -axis and  $y$ -axis to the mean center, respectively.

### Geographical detector

The geographical detector method is based on the geographical spatially stratified heterogeneity theory. The method is able to interpret the degrees to which a different impact factor  $X$  explains  $Y$  by detecting the relationship between  $X$  and the spatial change in  $Y$ . The geodetector method uses four different detectors (Wang *et al.* 2010). The present study applied the factor detector and the interaction detector to detect the main factors influencing flash floods.

### Factor detector

The factor detector was used to detect the degrees of explanation between impact factor  $X$  and the spatial differentiation of  $Y$  according to the rate of local variance to the global variance. The expression is:

$$q = 1 - \frac{1}{N\sigma^2} \sum_{k=1}^l N_k \sigma_k^2 \quad (13)$$



In Equation (13),  $q$  is the degree to which the impact factor  $X$  explained the spatial differentiation in  $Y$  ( $0 \leq q \leq 1$ ),  $k = 1, \dots, l$  represents the number of partitions of the impact factors  $X$ ,  $N_k$  is the number of units in the  $k$ th partition of the impact factors  $X$ ,  $N$  is the number of all units of  $X$ ,  $\sigma^2$  is the global variance of  $Y$ , and  $\sigma_k^2$  is the local variance of  $Y$  in the  $k$ th partition.

### Interaction detector

The interaction detector can assess the explanatory power of the interaction factor to  $Y$  through space superposition technology. Here,  $X1$  and  $X2$  are different influencing factors,  $X1 \cap X2$  represents the spatial superposition of  $X1$  and  $X2$ , and  $q(X1 \cap X2)$  represents the degree to which the interaction factor explains the spatial differentiation  $Y$ . As shown in Table 1, the interaction relationship of factors was determined by comparing the explanatory powers of  $q(X1)$ ,  $q(X2)$ , and  $q(X1 \cap X2)$ .

The power of each driving factor to explain the variance in flash floods can be identified through the following steps: (1) sampling points are established for each partition; (2) continuous driving forces are divided into 10 classes according to the natural break method; and (3) the density and driving factors of flash floods for each region are extracted by sampling points and the powers of each driving factor and each combination of interacting driving factors were calculated by the geographical detector. In addition, the average value of each driving factor is used for driving factors with multi-year data to calculate the power of each driving factor to explain flash floods.

## OVERVIEW OF THE STUDY AREA AND DATA PREPARATION


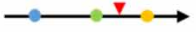
### Study area

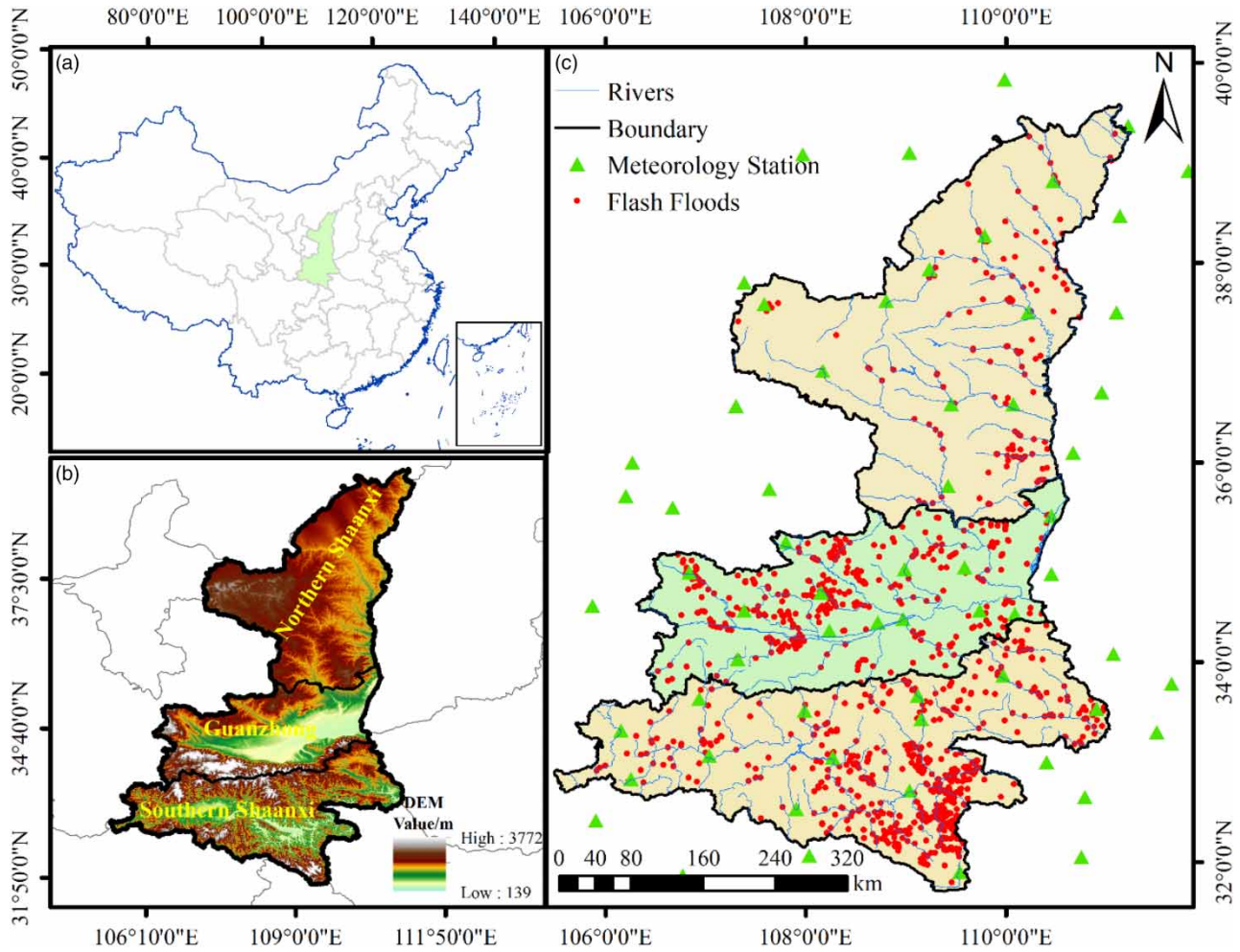
Shaanxi Province (105°29'–111°15'E, 31°42'–39°35'N) is in the northern part of central China and has a total area of  $2.056 \times 10^7$  km<sup>2</sup>. The altitude of the province decreases from north to south, showing obvious spatial zoning. The 'North Mountain' (a series of limestone-based rocky hills in the transition zone between the southern margin of the NS Loess Plateau and the GZ Basin) and Qinling Mountains can act as a boundary dividing Shaanxi Province into the following three natural regions: (1) the NS Loess Plateau; (2) the GZ Basin; and (3) the SS Qinba Mountain.

The different regions of Shaanxi Province show considerable differences in topography, geomorphology, climate, population distribution, and so on. The NS Loess Plateau is characterized by an altitude, average annual temperature, average annual precipitation, and population of 900–1,900 m; 7–12 °C; 400–550 mm; and ~5.9 million, respectively, whereas the corresponding characteristics of the GZ Basin are 190–850 m, 12–14 °C, 500–800 mm, and ~25.9 million, respectively, and that of the SS Qinba Mountain are 500–3,767 m; 14–16 °C; 700–900 mm; and ~7.7 million, respectively.

These regional differences in characteristics result in spatial-temporal and hazard-level differences in flash floods among the different regions. Therefore, an in-depth analysis of the distribution and driving factors of flash floods in the three regions of Shaanxi Province is required. Figure 2 shows an overview of the study area and the spatial distribution of historical flash floods.

**Table 1** | Types of interaction detectors

Graphical representation	Description	Interaction	Legend
	$q(X1 \cap X2) < \text{Min}(q(X1), q(X2))$	Weaken, nonlinear	
	$\text{Min}(q(X1), q(X2)) < q(X1 \cap X2) < \text{Max}(q(X1), q(X2))$	Weaken	• : $\text{Min}(q(X1), q(X2))$
	$q(X1 \cap X2) > \text{Max}(q(X1), q(X2))$	Enhance	• : $\text{Max}(q(X1), q(X2))$
	$q(X1 \cap X2) = q(X1) + q(X2)$	Independent	• : $q(X1) + q(X2)$
	$q(X1 \cap X2) > q(X1) + q(X2)$	Enhance, nonlinear	• : $q(X1 \cap X2)$



**Figure 2 |** Overview of Shaanxi Province, China: (a) the geographical position of Shaanxi Province in China; (b) the three regions of Shaanxi Province: Northern Shaanxi (NS), Guanzhong (GZ), and Southern Shaanxi (SS); and (c) the spatial distribution of historical flash floods in Shaanxi Province.

### Data preparation

The data used for Shaanxi Province in the present study included historical flash flood disaster data, the PPF, the SEF, and the HAF. The datasets were mainly obtained from the Flash Flood Investigation and Evaluation Dataset of Shaanxi Province (FFIEDSP), the National Meteorological Information of China (CMA, <http://www.resdc.cn>), the Geospatial Data Cloud (GDC, <http://www.gscloud.cn>), and the Resource and Environment Science and Data Center (RESDC, <http://www.nmic.cn>) (Table 1).

The present study used daily precipitation data as monitored by meteorological stations in Shaanxi Province and its surrounding areas from 1949 to 2015. The PPF was calculated by the following equation (Liu *et al.* 2017):

$$P_i = \frac{\sum_{y=1}^m \sum_{d=1}^n \left( \frac{R_d - R_i}{R_i} \times C_d \right)}{m}, \quad C_d = \begin{cases} 1, & R_d > R_i \\ 0, & R_d \leq R_i \end{cases} \quad (14)$$

In the above equation,  $m$  is the number of years and  $y$  is the index of each year,  $n$  is the number of precipitation days in a year,  $R_d$  is the daily precipitation on day  $d$ ,  $R_i = 25, 50$ , or  $100$  mm, corresponding with the threshold of the  $i$ th precipitation index of daily precipitation, and  $P_i$  is the precipitation indicator under  $R_i$  for each meteorological station.

The SEF included the digital elevation model (DEM, unit: m), the slope (SLOPE, unit: °), the soil type (SOIL, unitless), the normalized difference vegetation index (NDVI, unitless), and the vegetation type (VEG, unitless). The DEM, SLOPE, and NDVI were used to reflect topographic relief features, steepness of the surface, and vegetation cover, respectively, whereas the SOIL and VEG were closely related to the precipitation runoff process. Every indicator has a relationship with flash floods.

The HAF included the population density (PD; unit: person/km<sup>2</sup>), the density of gross domestic product (GDP; unit: 10<sup>4</sup> yuan/km<sup>2</sup>), the density of villages along the river point (VD; unit: village/km<sup>2</sup>), and the land use/land cover change (LUCC; unitless). The LUCC was included in the HAF since the LUCC is mainly affected by human activities (Chen *et al.* 2020; Yu *et al.* 2020).

## RESULTS AND ANALYSIS

### Spatial-temporal characteristics of flash floods

The present study analyzed the inter-annual, intra-annual, and inter-decadal variation of flash floods in different regions of Shaanxi Province. Spatial analysis of flash floods was based on the M-K, Sen, and SDE methods during which the trends in spatial variation and variation in the center of gravity of flash floods were analyzed, respectively.

#### Temporal scales

Figure 3 shows the temporal variation of historical flash floods in different regions of Shaanxi Province from 1949 to 2015.

#### Annual variations

Figure 3A(a), B(a), and C(a) shows that there was a fluctuating increasing trend in the number of flash floods in the NS, GZ, and SS regions. The time of mutation of the number of flash floods in each region was obtained using the M-K method. Figure 3A(b) shows that the mutation points of NS were in 1972 and 1999 (the intersection point of the *UF* and *UB*). A comparative analysis showed that a more obvious average value of abrupt change occurred after 1972 (Figure 3A(a)). Therefore, the mutation point of NS was 1972. Similarly, Figure 3B(b) and C(b) shows the mutation points of GZ and SS in 1981.

A comparative analysis of the number of flash floods before and after the mutation year was performed in each region to further explore the evolution of flash floods over time. As shown in Figure 3A(a), B(a) and C(a), the mean numbers of flash floods before and after the mutation year in the NS were 1 and 3, respectively, whereas those in the GZ were 3 and 9 and those in the SS were 3 and 16, respectively. There was a significant change in the mean number of flash floods in each region before and after the mutation year. The number of flash floods after the mutation exceeded that before the mutation by a factor of 3–5. The rank of the regions in terms of the acceleration in the growth rate in flash floods was SS > GZ > NS.

#### Intra-annual variability in flash floods

The present study analyzed the concentration of flash floods in each year. As shown in Figure 3A(c, d), B(c, d), and C(c, d), flash floods in Shaanxi Province were mainly concentrated from May to October. There was no significant change in the intra-annual distribution of flash floods before and after the mutation in each region. The flash floods in NS were mainly concentrated during July to August, whereas those in GZ and SS were from June to September and June to October, respectively. However, there was a significant difference in the number of flash floods, and the annual cumulative maximum after the mutation exceeded that before the mutation by a factor of 3–6. The cumulative maximum number of flash floods in NS and SS after the mutation year was 60 and 136, respectively, and occurred during July. The highest numbers of flash floods in GZ occurred in July and August at 98 and 100, respectively.

#### Inter-decadal variability in flash floods

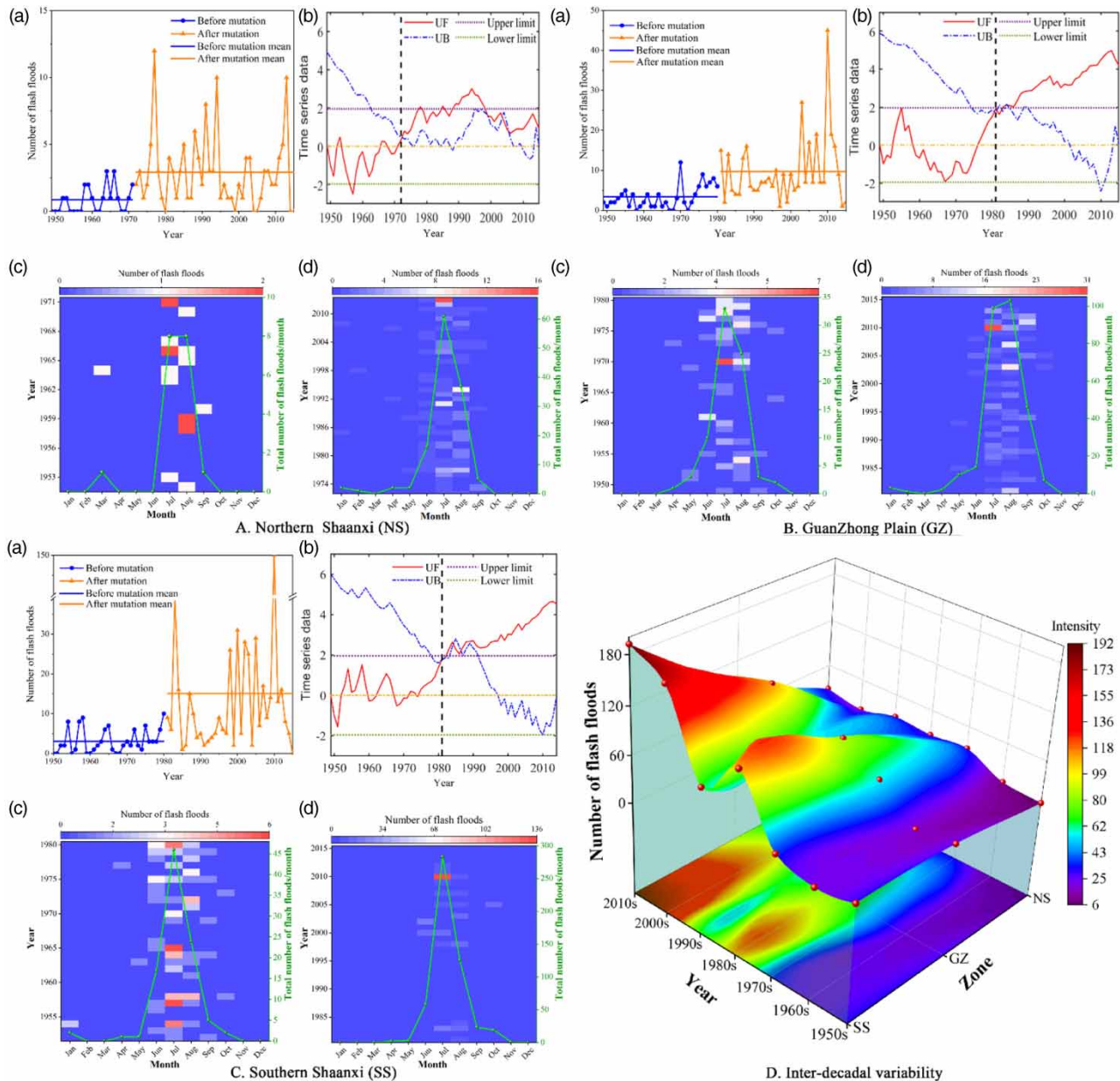
Figure 3(d) shows the inter-decadal variability of flash floods in the NS, GZ, and SS regions. The number of flash floods showed an increasing trend during the seven decadal periods from the 1950s to 2010s. In particular, the number of flash floods showed a significant increasing trend after the 1980s. The highest increase in flash floods occurred in SS. In addition, the frequency of flash floods in NS was lower than that in SS and GZ during each decadal period.

#### Spatial scales

##### Trend in spatial variation

The spatial distribution of flash floods in Shaanxi Province was analyzed based on the MK + Sen method. The trends in variations in *S* were divided into three categories: (1) a stable trend (from −0.0001 to 0.0001); (2) a decreasing trend (< −0.0001);





**Figure 3** | Temporal variation of flash floods in different regions of Shaanxi Province, China: (A) NS; (B) GZ; (C) SS; and (D) inter-decadal variability. (a) The annual number of flash floods and their multi-year averages before and after mutation; (b) flash flood mutation analysis using the Mann-Kendall test; (c) and (d) are the monthly number of flash floods before and after the mutation, respectively.

and (3) an increasing trend ( $>0.0001$ ). At the same time, values which did not show a variation trend (absolute value of the M-K  $< 1.96$ ) and other values that represented significant variation trends were noted. Table 2 shows a summary of the five categories of spatial variation trends.

The proportions of flash floods following within the different classifications in each region were calculated. Figure 4 shows the distribution of spatial trends in each region. There was a ‘no obvious change’ trend in flash flood disaster in  $>80\%$  of the area in NS, whereas 12.44% of the area showed a ‘moderate growth’, mainly in the northeast. In GZ, 30.74% of the area showed a ‘moderate growth’ in flash flood disaster, whereas 28.97 and 5.3% of the area showed ‘no obvious change’ and ‘rapid growth’, respectively. There was a ‘no obvious change’ in flash flood disaster in 42.16% of the area of SS, whereas ‘moderate growth’ and ‘rapid growth’ collectively accounted for  $\sim 42.46\%$  of the total area, mainly in the southeast. Overall, there was an increasing trend in flash floods in Shaanxi Province, with a relatively large proportion of area showing a ‘moderate



**Table 2** | List of data used and sources

Data category	Parameters	Scale or resolution	Time	Date sources
Historical flash flood data	Flash floods	1:50,000	1949–2015	FFIEDSP
PPF	Daily precipitation	–	1951–2015	CMA
SEF	DEM	30 m × 30 m	2009	GDC
	SLOPE	30 m × 30 m	2009	
	SOIL	1:1,000,000	2010	RESDC
	NDVI	1 km × 1 km	2000–2015	
	VEG	1:1,000,000	2001	
HAF	PD	1 km × 1 km	1990, 1995, 2000, 2005, 2010, 2015	
	GDP	1 km × 1 km	1990, 1995, 2000, 2005, 2010, 2015	
	LUCC	1 km × 1 km	2010	
	VD	1:500,000	2014	FFIEDSP

DEM, digital elevation model; NDVI, normalized difference vegetation index; VEG, vegetation type; PD, population density; GDP, gross domestic product; LUCC, land use/cover change; VD, density of villages along the river point; FFIEDSP, Flash Flood Investigation and Evaluation Dataset of Shaanxi Province; CMA, National Meteorological Information of China; RESDC, Resource and Environment Science and Data Center; GDC, the Geospatial Data Cloud.

**Table 3** | Classification of spatial variation trends of flash floods in Shaanxi Province, China

Classification	Abbreviation	M-K	Sen
Rapid decline	I	$UF_k < -1.96$	$S < -0.0001$
Moderate decline	II	$-1.96 < UF_k < 1.96$	$S < -0.0001$
No obvious change	III	$-1.96 < UF_k < 1.96$	$-0.0001 < S < 0.0001$
Moderate growth	IV	$-1.96 < UF_k < 1.96$	$S > 0.0001$
Rapid growth	V	$UF_k > 1.96$	$S > 0.0001$

M-K, Mann-Kendall test; Sen, Theil-Sen Median trend analysis.

growth' in flash floods, mainly concentrated in the GZ, whereas 'rapid growth' was mainly concentrated in the SS. There was a gradual increase in flash floods in the NS, in which the risk of flash floods was relatively small compared with that in the other two regions.

### Evolution of flash floods

The spatial patterns in flash floods in each region from the 1950s to 2010s were analyzed using the SDE method. As shown in Figure 5, the orientation angles in the NS, GZ, and SS regions ranged from 142.92° to 9.83°, 78.83° to 71.73°, and 68.66° to 100.23°, respectively. The track of the gravity center indicted a shift in the gravity center from the northwest to the southeast in NS, west to the east in GZ, and north to the south in SS. The oblateness of the ellipse of flash floods during each period in the three regions indicated that the directionality transfer of flash floods in NS was the most significant, followed by that in SS, whereas that in GZ was the least. The SDE area in the NS, GZ, and SS regions ranged from 63,393 to 32,336 km<sup>2</sup>; 23,506 to 35,183 km<sup>2</sup>; and 35,183 to 20,069 km<sup>2</sup>, respectively. The results showed that the SDE area of flash floods in each region decreased, indicating that the frequency of flash floods was higher and more concentrated.

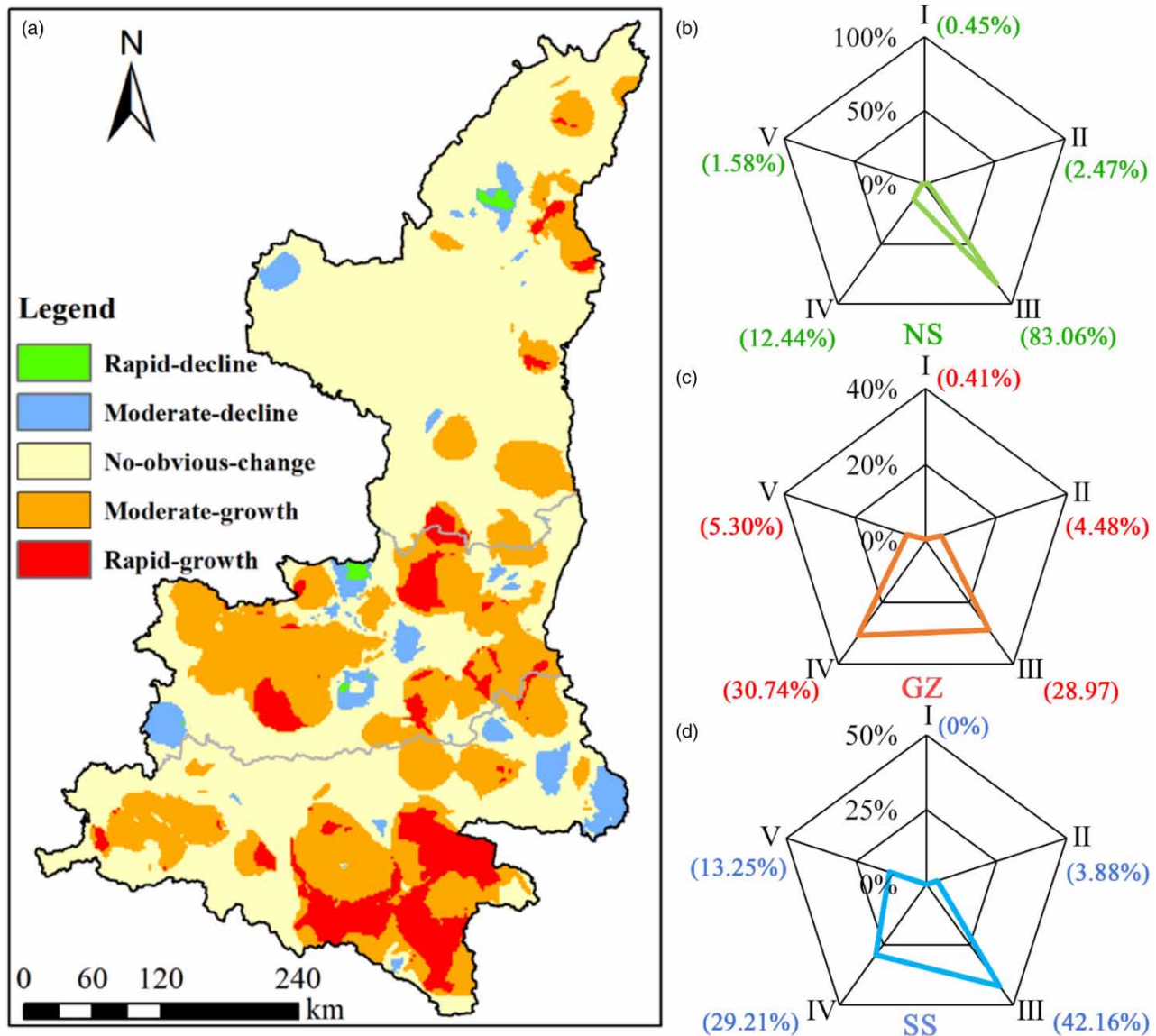
### Analysis of the factors driving flash floods

There are 12 factors driving flash floods, categorized into the PPF, SEF, and HAF. The current study calculated the power of each driving factor and interactions between driving factors to explain flash floods after mutation in the different regions.

### Driving factors

#### Precipitation factors

Mutations in the flash floods occurred at different time scales among the different regions. Therefore, the present study analyzed precipitation data after the mutation. Three precipitation indicators, P25, P50, and P100, were calculated according to Equation



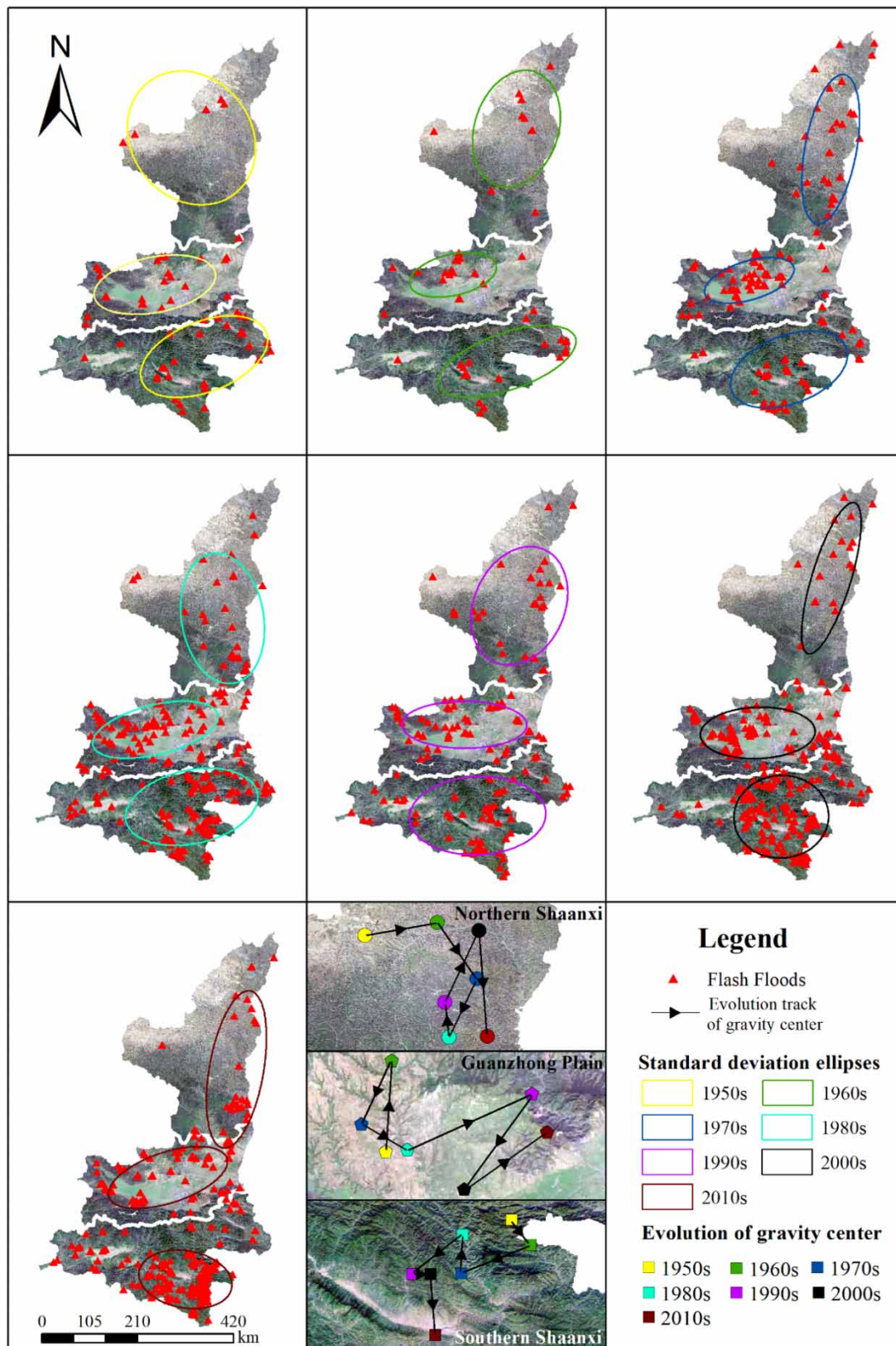
**Figure 4** | (a) Trends in the spatial variation of flash floods in Shaanxi Province, China; (b)–(d) represent the proportions of area showing different categories of trends in variation in flash floods in NS, GZ, and SS, with the actual values given in parentheses.

(14) to characterize the driving effect of moderate rain, heavy rain, and rainstorms on flash floods considering the actual situation and precipitation classification standard. The Kriging interpolation method was used to obtain the spatial distribution of continuous data of different precipitation intensities in each region. As shown in Figure 6, the average root mean square error values of different precipitation intensities in the NS, GZ, and SS regions were 0.23, 0.41, and 0.39, respectively. These results indicated that the interpolation results can accurately reflect the spatial distribution of precipitation intensities in different regions.

As shown in Figure 6, precipitation intensity increased from NS to SS. Higher values of P25 were concentrated mainly at the boundary of each region, whereas higher values of P50 were distributed in the central part of the NS, southern GZ, and SS regions, respectively. Higher values of extreme weather (P100) showed a lower distribution throughout Shaanxi Province, mainly concentrated in SS.

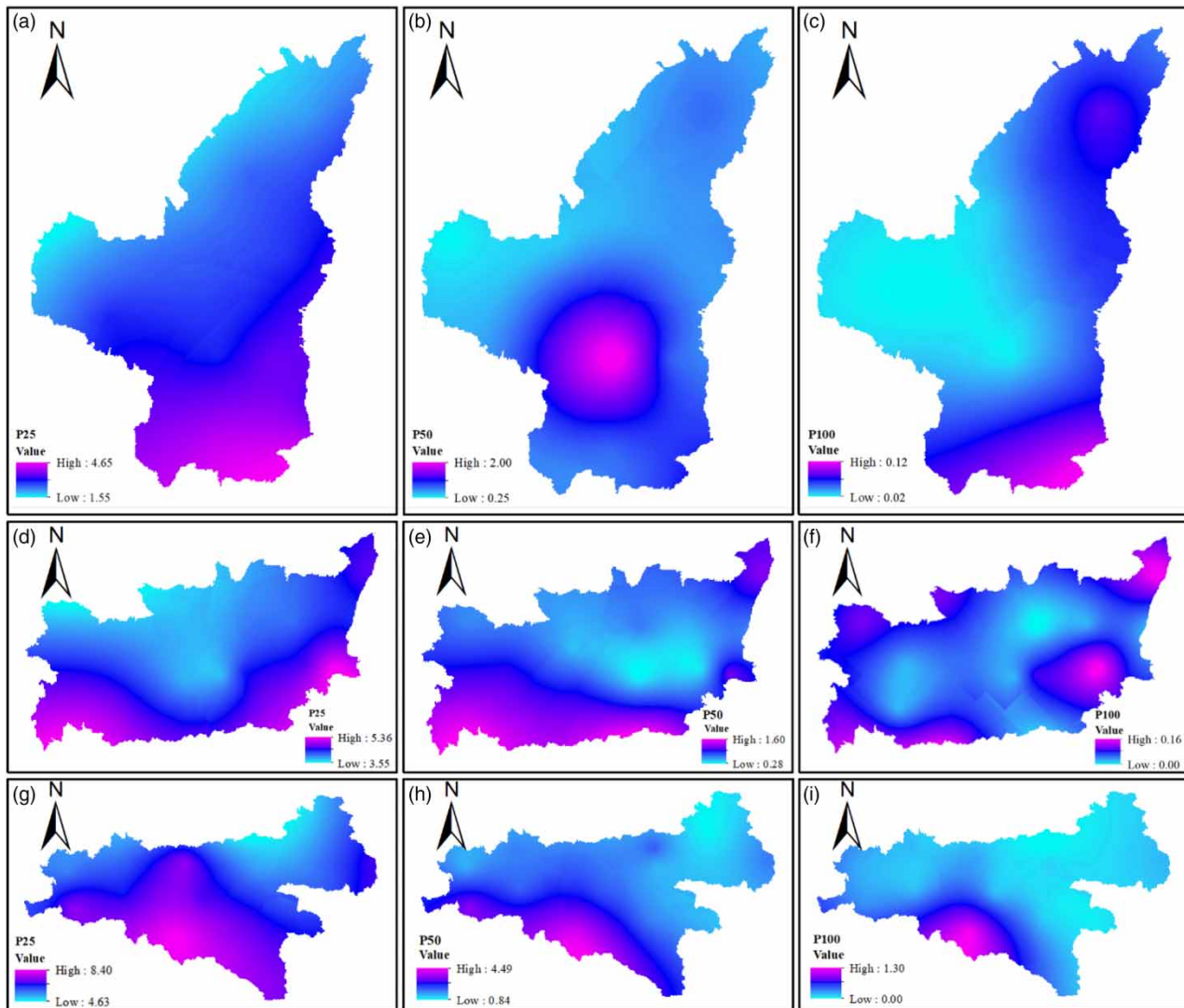
#### Surface environmental factors

The DEM, SLOPE, SOIL, NDVI, and VEG were selected as the SEF. Figure 7 shows the spatial distribution of all indicators. As shown in Figure 7, the indicators in Shaanxi Province displayed obvious regional differences. The elevation of Shaanxi



**Figure 5** | Distribution and evolution of the gravity center of flash floods from the 1950s to the 2010s in the northern Shaanxi (NS), Guanzhong (GZ), and southern Shaanxi (SS) regions in Shaanxi Province, China.





**Figure 6** | Precipitation factors of Shaanxi Province, China. (a)–(c) represent the P25, P50, and P100 in NS, respectively; (d)–(f) represent the P25, P50, and P100 in GZ, respectively; (g)–(i) represent the P25, P50, and P100 in SS, respectively.

Province is high in the north and low in the center. Initial soil, semi-luvisols, and luvisols are the main soil types in the NS, GZ, and SS regions, respectively. As SS contains many types of vegetation, vegetation cover in this region is relatively complex. In contrast, NS and GZ contain relatively simple vegetation cover, with cultivated vegetation accounting for most vegetation. Within the regions, the NDVI increases from NS to SS.

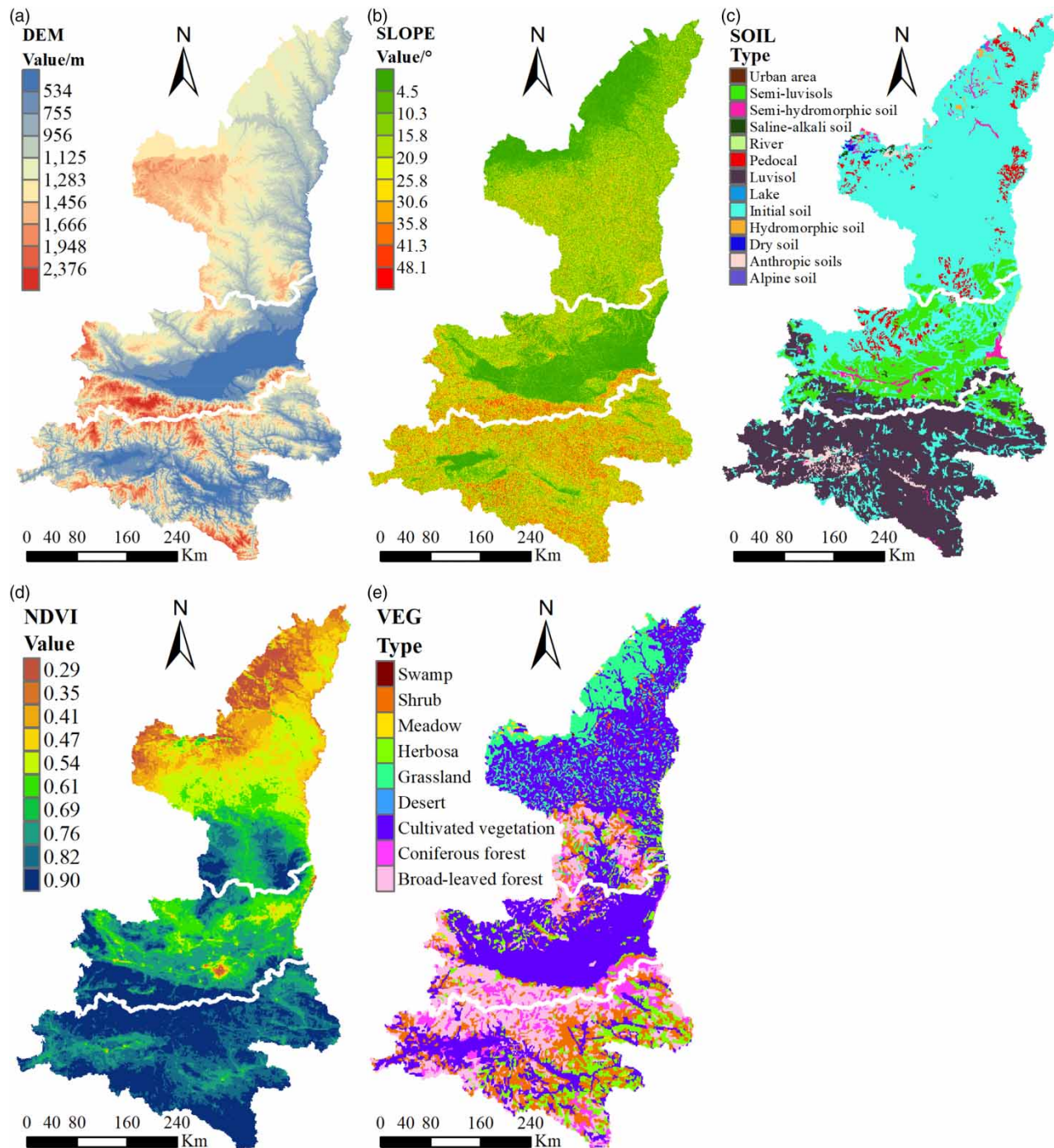
### Human activity factors

The HAF not only increase the risk of flash floods but also act as an indicator of the degree of potential damage resulting from flash floods. Therefore, the present study selected VD, LUCC, PD, and GDP as the HAF. Figure 8 shows the categories and distributions of all the indicators. As shown in Figure 8, there were a clear regional difference in the HAF, with the VD along the river in NS and SS exceeding that in GZ, whereas PD and GDP were mainly concentrated in the GZ Plain. Cultivated land accounts for a large proportion of the LUCC in Shaanxi Province, thereby providing a favorable environment for the occurrence of flash floods.

### Powers of the driving factors for explaining flash floods

The power of each driver and that of their interactions to explain flash floods in the different regions were obtained based on the geographic detector method. The  $P$ -values obtained were  $<0.01$ , indicating that the power of each factor to explained



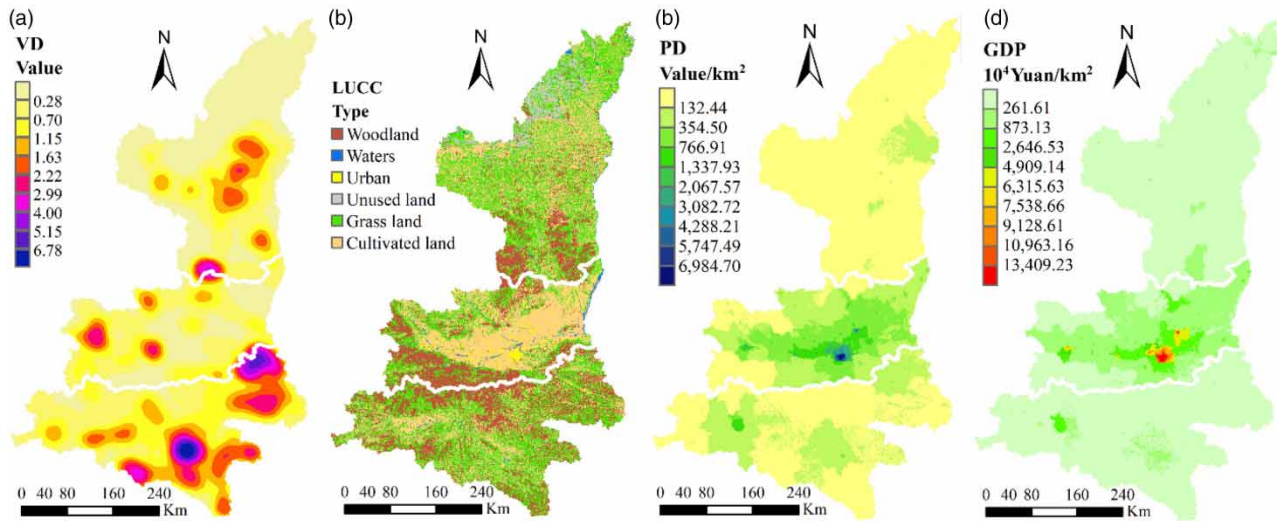


**Figure 7** | Surface environmental factors of Shaanxi Province, China: (a) digital elevation model (DEM); (b) slope (SLOPE); (c) soil type (SOIL); (d) normalized difference vegetation index (NDVI); and (e) vegetation type (VEG).

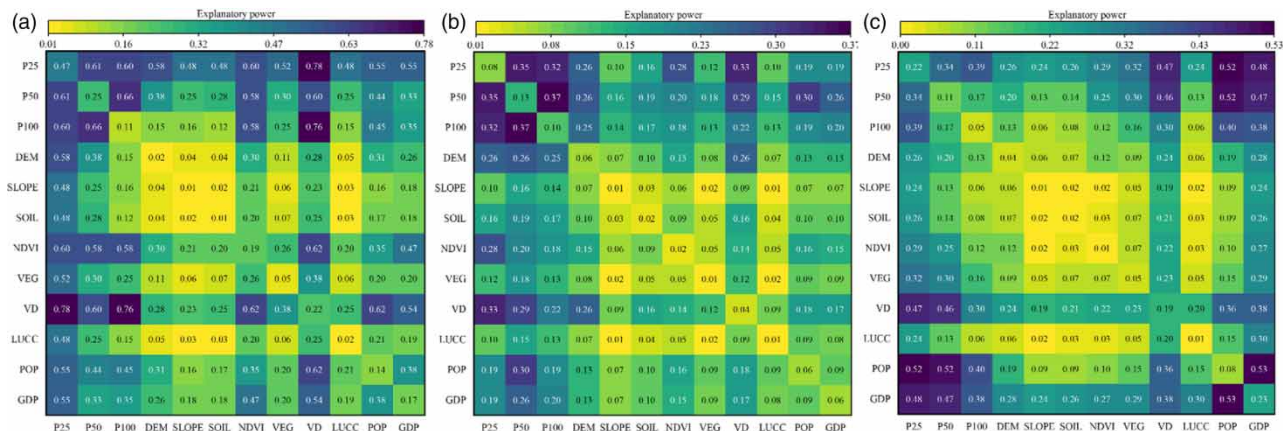
flash floods was significant. The powers of the interacting driving factors to explain the spatial patterns of flash floods in the NS, GZ, and SS regions are shown in Figure 9.

#### Northern Shaanxi

As shown in Figure 9(a), flash floods in NS were most affected by precipitation. The power of rainfall factors to explain the spatial pattern of flash floods exceeded 25%, with P25 playing a major role. The NDVI was the driving factor in the SEF, with the highest



**Figure 8** | Human activity factors of Shaanxi Province, China: (a) village density (VD); (b) land use/land cover change (LUCC); (c) population density (PD); and (d) density of gross domestic product (GDP).



**Figure 9** | Powers of interactions between various driving factors to explain flash floods in Shaanxi Province, China: (a) NS; (b) GZ; and (c) SS.

power to explain flash floods. This could mainly be attributed to the position of NS in the Loess Plateau in which serious soil erosion occurs. The VD was the driving force in the HAF, with the most power to explain flash floods. This result could be attributed to the direct relationship between the VD and the spatial pattern of flash floods. The results of the detection of the power of interactions between factors to explain flash floods showed that the powers of  $PPF \cap SEF$  and  $PPF \cap HAF$  to explain flash floods were significantly higher than those of  $SEF \cap HAF$ . The powers of  $P25 \cap VD$ ,  $P100 \cap VD$ , and  $P50 \cap P100$  to explain flash floods exceeded those of others' interactions, indicating that precipitation has a greater impact on flash floods in NS.

### Guanzhong

As shown in Figure 9(b), the rank of rainfall factors in terms of the power to explain flash floods was as follows:  $P50 > P100 > P25$ . The maximum explanatory power of P50 was 0.13. The rank of surface factors to explain flash floods was as follows:  $DEM > SOIL > NDVI > VEG > SLOPE$ . This result could be attributed to flash floods mainly being concentrated in the junction area of mountains and plains and along the river. During flash floods, areas of low elevation are more vulnerable to flooding. Among the HAF, the GDP and POP showed strong powers to explain flash floods. This result could be attributed to the large PD of GZ, resulting in human activities having a strong impact on flash floods. Among the interactions



in driving factors,  $P50 \cap P100$  ( $q = 0.37$ ) showed the strongest power to explain flash floods, indicating that heavy rainfall is the key factor driving flash floods in the GZ area. In addition, there were relatively strong interactions between impact factors for  $P25 \cap VD$  ( $q = 0.33$ ),  $P25 \cap NDVI$  ( $q = 0.27$ ), and  $DEM \cap VD$  ( $q = 0.26$ ). In general, the powers of interactions between the PPF and SEF and between the PPF and HAF to explain flash floods exceeded that of the interaction between the SEF and HAF. An analysis of the drivers of flash floods showed that historical flash floods in GZ were mainly as a result of the joint action of multiple factors, with these interactions increasing the risk of flash floods.

### Southern Shaanxi

As shown in Figure 9(c), among the PPF, the power of P25 to explain flash floods exceeded that of P50 and P100. The DEM and VEG were the main driving factors of flash floods among the SEF. The HAF showed a high power to explain flash floods in SS, with the explanatory powers of the GDP and VD being 0.23 and 0.19, respectively. This result could mainly be attributed to the population being the main factor subjected to disaster during a flash flood.

The results showed that among the interactions between driving factors, those between the POP and GDP had the greatest interactive power to explain the flash floods of 0.53. The results showed that flash floods in SS have the potential to result in considerable losses. Besides, the interactions between various driving factors showed a high potential for driving flash floods. Therefore, flash floods in SS are relatively complex and are affected by multiple factors.

## DISCUSSION

### Spatial-temporal variation in flash floods

The present study showed the spatial-temporal distributions of flash floods in three regions of Shaanxi Province. There were increasing trends in the numbers of flash floods in each region. The numbers of flash floods increased in NS after the 1970s and in GZ and SS after the 1980s, consistent with the previous research results (Liu *et al.* 2018). The possible reasons for this result include the following: (1) the original underlying surface features changed due to increased human activities; (2) the increases in the frequency of extreme precipitation events; and (3) a lack of records for flash floods before the mutation periods. There were no significant differences in the intra-annual distributions of flash floods among the NS, GZ, and SS regions, with flash floods concentrated during the flood season. The frequency of flash floods was the highest during July, indicating a strong correlation between flash floods and precipitation.

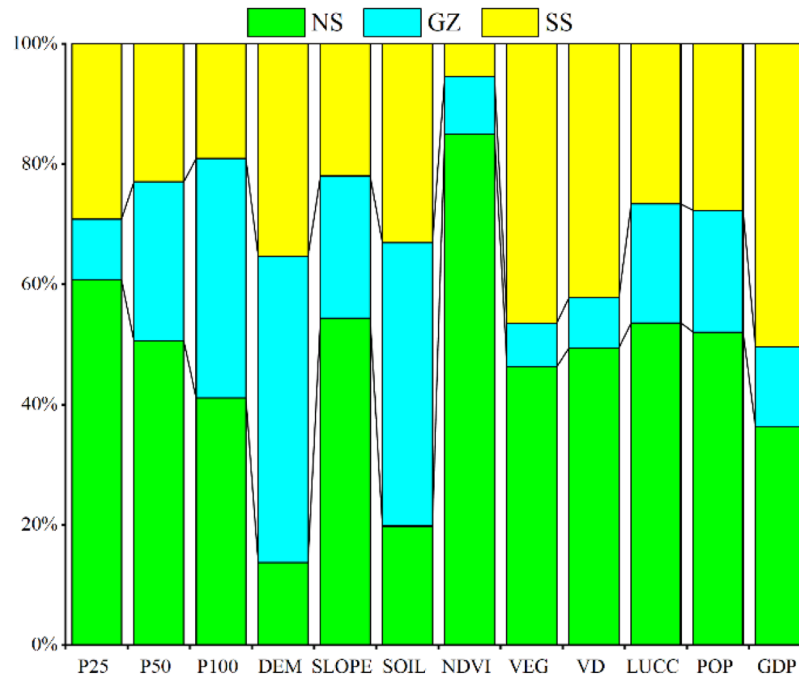
The increasing trends in flash floods were different among the different regions. The increase in flash floods was most obvious in SS, followed by GZ, with that in NS the weakest. These differences between regions could be attributed to precipitation being the main driver of flash floods. The significant differences in the spatial and temporal distributions of precipitation between each region could be attributed to the unique climatic conditions and geographical environment of each region. Consequently, there were differences in the spatial and temporal distribution of precipitation among each region, leading to a gradual increase of precipitation in Shaanxi Province from north to south. The results of the SDE indicated that the distributions of flash floods in each region showed different directions among the different periods, with the patterns of distribution increasingly obvious and more concentrated from the 1950s to 2010s. High-intensity precipitation and human activities were the characteristics of areas in which flash floods were concentrated. Therefore, it can be concluded that the main drivers of flash floods are precipitation and human activities.

### Assessment of driving forces

#### Factor detection

A further exploration of the main drivers of and differences in flash floods between the NS, GZ, and SS regions was required to increase the prevention and control of flash floods in Shaanxi Province. Figure 10 shows a comparison of driving factors identified for the three regions. As shown in Figure 10, the PPF played a major role in the occurrence of flash floods. Flash floods in NS and SS were mainly driven by moderate precipitation, whereas GZ was greatly affected by an intense precipitation. This result could be mainly attributed to a large area of GZ being located in the plain area in which the runoff rate is slow when precipitation intensity is weak.

Among the SEF, the NDVI and SLOPE showed the greatest powers to explain flash floods in NS. This result could be mainly attributed to NS being in the Losses Plateau in which there is less vegetation coverage and a large slope. Precipitation tends to result in soil erosion, which promotes the occurrence of flash floods. There were large differences between SS and NS in terms of surface factors explaining flash floods, and beside for the NDVI, other surface indicators showed relatively high



**Figure 10** | Powers of different factors to explain flash floods among the different regions in Shaanxi Province, China.

powers to explain flash floods. This result could be attributed to the high, complex, and diverse vegetation coverage in SS. Flash floods are affected by the comprehensive effects of the SEF.

Among the HAF, the VD was the factor indicating the greatest vulnerability to flash floods. The POP and GDP also showed strong powers to explain flash floods in areas in which flash floods have a great impact on casualties and property losses. It is noteworthy that the GDP and POP in GZ were significantly higher than those in the other two regions. However, Figure 10 shows that the powers of these two indicators to explain flash floods were weak, whereas those of the DEM and SOIL were strong. This result could be mainly attributed to flash floods in GZ mostly occurring at the junction of the mountains and plains in which the PD is relatively low.

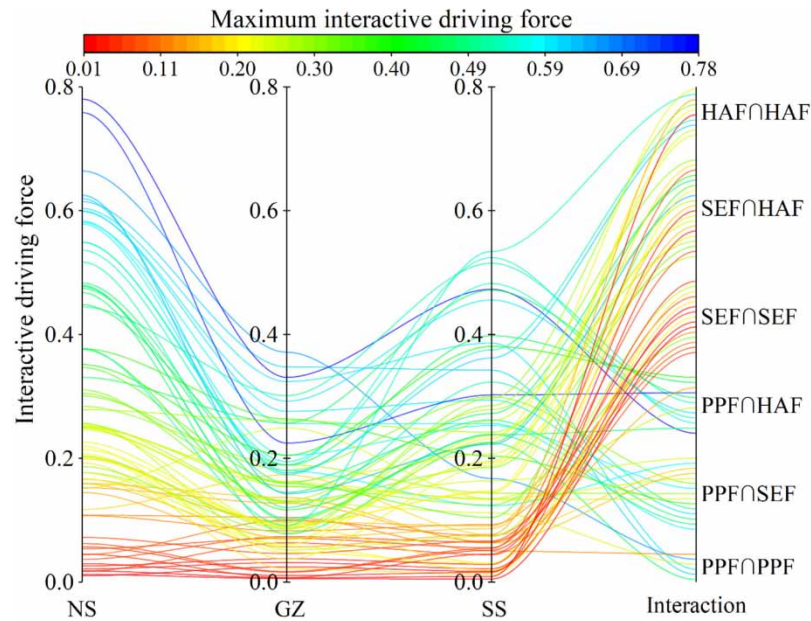
### Detection of interactions

The interaction detector allowed the powers of interactions among different factors to enhance flash floods in each region to be identified. The proportions of nonlinear enhancement in factor interaction in the NS, GZ, and SS regions were 73, 97, and 91%, respectively. This result showed that flash floods in GZ and SS were affected by multiple factors, presenting a more complex situation than that in NS. Figure 11 shows the comparative result of the interaction detector for the three regions. NS showed a stronger interaction among driving factors, which could mainly be attributed to the relatively large power of a single indicator to explain flash floods, after which the enhancement by interaction resulted in the strengthening of the explanatory power. Although there were weak interactions between factors in GZ, their powers to explain flash floods were relatively concentrated. The interactions between factors in SS were similar to that in GZ, although the powers of interactions to explain flash floods were larger in SS.

The above analysis showed differences in the main forces driving flash floods among the different regions. However, the flash floods in Shaanxi Province are greatly impacted by the combined effects of precipitation and human activities. Precipitation is required for the formation of flash floods, whereas human activities indicate vulnerability to flash floods. Therefore, these factors have strong correlations with the occurrence of flash floods. In addition, population growth and urban expansion have resulted in changes to the surface environment since the 1980s, thereby shortening the time taken for runoff formation and increasing the incidences of flash floods (Cortes *et al.* 2018).

Therefore, certain strategies should be implemented to control and mitigate flash floods according to the main factors driving flash floods in different regions. These include strengthening the forecast of rainfall of different intensity, proving an early





**Figure 11** | Comparison of interaction detection results under regional differentiation.

warning of flash floods, and implementing the timely evacuation of residents of villages along the river. In addition, the vegetation coverage in NS should be strengthened to reduce soil erosion.

### Limitations of the present study

The present study analyzed the spatial-temporal distributions and flash flood driving factors and their explanatory power in Shaanxi Province. However, insufficiencies in the current study remain. First, the daily precipitation data used in the current study do not accurately reflect the impact of short-term heavy precipitation on flash floods. Second, there may be strong correlations between the different drivers of flash floods. Therefore, there is a need to further explore the sensitivity of flash floods to each driving factor. Despite these shortcomings, the present study provides a preliminary reference for the spatial and temporal distribution pattern of flash floods in Shaanxi Province and can improve public awareness of flash floods.

### CONCLUSIONS

This present study quantitatively explored the spatial and temporal distributions of historical flash floods in the NS, GZ, and SS regions of Shaanxi Province from 1949 to 2015 based on the regional heterogeneity in flash floods. The powers of the main PPF, SEF, and HAF and their interactions to explain flash floods in each region were revealed. The differences in the distribution and driving factors of flash floods in three regions were compared and analyzed. The following conclusions were drawn:

- There were obvious increasing trends in flash floods in Shaanxi Province among the different regions, with flash floods mainly concentrated in July and August. The rank of the regions in terms of the number of disasters and the growth trend was as follows: SS > GZ > NS. Moreover, disasters become increasingly concentrated after mutation.
- The forces driving flash floods in the NS, GZ, and SS regions were mainly related to precipitation intensity and human activities, although there were some differences among the regions. The flash floods in GZ and SS were the result of multiple factors, with the superposition effect of nonlinear enhancement exceeding 90%. However, individual driving factors had the greatest power to explain flash floods in NS and were mainly related to moderate rain intensity and the NDVI.
- Further study is required to extend the results of the present study, including the influence of short-term heavy rainfall and equifinality among different parameters on the driving force of flash floods. However, the results of the present study provide a preliminary basis for the prevention and mitigation of flash floods in Shaanxi Province, and corresponding preventative measures can be implemented according to the drivers and characteristics of flash floods in different regions.

## ACKNOWLEDGEMENTS

The authors are grateful to the editors and the anonymous reviewers for their insightful comments and suggestions. We are also grateful for the data support provided by the Flash Flood Investigation and Evaluation Dataset of Shaanxi Province (FFIEDSP), the National Meteorological Information of China (CMA), the Geospatial Data Cloud (GDC), and the Resource and Environment Science and Data Center (RESDC).

## FUNDING

This study was supported by the Natural Science Basic Research Program of Shaanxi (Program No. 2019JLZ-15), the Water Science and Technology Program of Shaanxi (Program No. 2018slkj-4), and the Research Fund of the State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Xi'an University of Technology (Grant No. 2019KJCXTD-5).

## AUTHOR CONTRIBUTIONS

H.Z., J.G.L., and J.Y.W. conceptualized the study. H.Z. and J.Y.W. contributed to the methodology and prepared in writing the original draft. H.Z., J.G.L., and M.J.Y. wrote, reviewed, and edited the original manuscript. J.G.L. contributed to the funding acquisition.

## AVAILABILITY OF DATA AND CODE

The flash floods data of the Flash Flood Investigation and Evaluation Dataset of Shaanxi Province (FFIEDSP) and codes generated or used during the study are available from the corresponding author by request.

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## REFERENCES

- Adnan, M. S. G., Dewan, A., Zannat, K. E. & Abdullah, A. M. 2019 [The use of watershed geomorphic data in flash flood susceptibility zoning: a case study of the Karnaphuli and Sangu river basins of Bangladesh](#). *Natural Hazards* **99** (1), 425–448.
- Arora, A., Arabameri, A., Pandey, M., Siddiqui, M. A., Shukla, U. K., Bui, D. T., Mishra, V. N. & Bhardwaj, A. 2021 [Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India](#). *Science of the Total Environment* **750**.
- Chen, D. C., Xu, X. L., Sun, Z. Y., Liu, L., Qiao, Z. & Huang, T. 2020 [Assessment of urban heat risk in mountain environments: a case study of Chongqing Metropolitan Area, China](#). *Sustainability* **12** (1).
- Cortes, M., Turco, M., Llasat-Botija, M. & Llasat, M. C. 2018 [The relationship between precipitation and insurance data for floods in a Mediterranean region \(northeast Spain\)](#). *Natural Hazards and Earth System Sciences* **18** (3), 857–868.
- Esmailpour, M., Ghasemi, A. R., Khoramabadi, F. & Rashedi, S. 2021 [Spatiotemporal variability of trend in extreme precipitations using fuzzy clustering over Northwest Iran](#). *Earth Science Informatics*.
- Hou, J., Li, B., Tong, Y., Ma, L., Ball, J., Luo, H., Liang, Q. & Xia, J. 2020 [Cause analysis for a new type of devastating flash flood](#). *Hydrology Research* **51** (1), 1–16.
- Liu, Y. S., Yuan, X. M., Guo, L., Huang, Y. H. & Zhang, X. L. 2017 [Driving force analysis of the temporal and spatial distribution of flash floods in Sichuan Province](#). *Sustainability* **9** (9).
- Liu, Y. S., Yang, Z. S., Huang, Y. H. & Liu, C. J. 2018 [Spatiotemporal evolution and driving factors of China's flash flood disasters since 1949](#). *Science China-Earth Sciences* **61** (12), 1804–1817.
- Ngo, P. T. T., Pham, T. D., Nhu, V. H., Le, T. T., Tran, D. A., Phan, D. C., Hoa, P. V., Amaro-Mellado, J. L. & Bui, D. T. 2021 [A novel hybrid quantum-PSO and credal decision tree ensemble for tropical cyclone induced flash flood susceptibility mapping with geospatial data](#). *Journal of Hydrology* **596**.
- Nyikadzino, B., Chitakira, M. & Muchuru, S. 2020 [Rainfall and runoff trend analysis in the Limpopo river basin using the Mann Kendall statistic](#). *Physics and Chemistry of the Earth* **117**.
- Peng, J., Chen, S., Lu, H. L., Liu, Y. X. & Wu, J. S. 2016 [Spatiotemporal patterns of remotely sensed PM2.5 concentration in China from 1999 to 2011](#). *Remote Sensing of Environment* **174**, 109–121.
- Phuong, D. N. D., Tram, V. N. Q., Nhat, T. T., Ly, T. D. & Loi, N. K. 2020 [Hydro-meteorological trend analysis using the Mann-Kendall and innovative-Sen methodologies: a case study](#). *International Journal of Global Warming* **20** (2), 145–164.
- Seenu, P. Z. & Jayakumar, K. V. 2021 [Comparative study of innovative trend analysis technique with Mann-Kendall tests for extreme rainfall](#). *Arabian Journal of Geosciences* **14** (7).

- Shi, Y. S., Matsunaga, T., Yamaguchi, Y., Li, Z. Q., Gu, X. F. & Chen, X. H. 2018 Long-term trends and spatial patterns of satellite-retrieved PM<sub>2.5</sub> concentrations in South and Southeast Asia from 1999 to 2014. *Science of the Total Environment* **615**, 177–186.
- Shi, W. Z., Huang, S. Z., Liu, D. F., Huang, Q., Han, Z. M., Leng, G. Y., Wang, H., Liang, H., Li, P. & Wei, X. T. 2021 Drought-flood abrupt alternation dynamics and their potential driving forces in a changing environment. *Journal of Hydrology* **597**.
- Singh, P., Kumar, V., Thomas, T. & Arora, M. 2008 Changes in rainfall and relative humidity in river basins in northwest and central India. *Hydrological Processes* **22** (16), 2982–2992.
- Syed, A., Liu, X. P., Moniruzzaman, M., Roustia, I., Syed, W., Zhang, J. Q. & Olafsson, H. 2021 Assessment of climate variability among seasonal trends using In situ measurements: a case study of Punjab, Pakistan. *Atmosphere* **12** (8).
- Trigo, R. M., Ramos, C., Pereira, S. S., Ramos, A. M., Zezere, J. L. & Liberato, M. L. R. 2016 The deadliest storm of the 20th century striking Portugal: flood impacts and atmospheric circulation. *Journal of Hydrology* **541**, 597–610.
- Wang, J. F., Li, X. H., Christakos, G., Liao, Y. L., Zhang, T., Gu, X. & Zheng, X. Y. 2010 Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science* **24** (1), 107–127.
- Wang, J. F., Zhang, T. L. & Fu, B. J. 2016 A measure of spatial stratified heterogeneity. *Ecological Indicators* **67**, 250–256.
- Wang, X. H., Cong, P. T., Jin, Y. H., Jia, X. C., Wang, J. S. & Han, Y. X. 2021 Assessing the effects of land cover land Use change on precipitation dynamics in Guangdong-Hong Kong-Macao Greater Bay Area from 2001 to 2019. *Remote Sensing* **13** (6).
- Xiong, J. N., Pang, Q., Fan, C. K., Cheng, W. M., Ye, C. C., Zhao, Y. L., He, Y. R. & Cao, Y. F. 2020 Spatiotemporal characteristics and driving force analysis of flash floods in Fujian Province. *Isprs International Journal of Geo-Information* **9** (2).
- Yang, X. P., Jia, Y. T., Wang, Q. H., Li, C. M. & Zhang, S. X. 2021 Space-time evolution of the ecological security of regional urban tourism: the case of Hubei Province, China. *Environmental Monitoring and Assessment* **193** (9).
- Yu, W. H., Zhang, L. J., Zhang, H. W., Jiang, L. Q., Zhang, A. K. & Pan, T. 2020 Effect of farmland expansion on drought over the past century in Songnen Plain, Northeast China. *Journal of Geographical Sciences* **30** (3), 439–454.
- Zhang, W. Q., Jin, H. A., Shao, H. Y., Li, A. N., Li, S. Z. & Fan, W. J. 2021 Temporal and spatial variations in the leaf area index and Its response to topography in the Three-River Source Region, China from 2000 to 2017. *Isprs International Journal of Geo-Information* **10** (1).

First received 25 September 2021; accepted in revised form 25 November 2021. Available online 10 December 2021