



# Comparative study on flash flood hazard assessment for Nam Ou River Basin, Lao PDR

Jun Du<sup>1</sup> · Zhong-jie Fan<sup>1,2</sup> · Jian Pu<sup>1</sup>

Received: 20 May 2019 / Accepted: 23 April 2020 / Published online: 8 May 2020  
© Springer Nature B.V. 2020

## Abstract

Laos is a mountainous, rainy and less developed country in Southeast Asia. In Laos, floods represent a major constraint on social economic development, causing a large number of casualties and property losses each year, among which the impact from flash floods is also very prominent. Especially in recent decades, with the development of social economy and the intensification of hydropower development, the serious threat from flash floods is becoming more and more obvious. However, there is no fundamental defence system for flash floods been established yet in this country, and the basic knowledge on local flash flood development is also ignored. For filling this gap, taking the Nam Ou River Basin as an example, this paper tries to find out the most helpful hazard assessment method for current Laos based on comparative analysis from the flash flood potential index, the calculation of curve number–rainfall erosivity and the extrapolation method. The results show that the extrapolation method based on spatial lag model constructed by the data from Yunnan, China, presents the most reliable outcome compared with the other two methods, indicating the spatial autocorrelation model can also be useful for extrapolation, effectively. Besides, the scale effect of different potential impact elements on flash flood, i.e. the spatial correlation between each element and the distribution of flash flood events at different spatial statistic units, was also preliminary studied. It is found that the degree of correlation in spatial analysis depends on the short board effect, i.e. only the element that restricts the regional flash flood developmental system can be the key factor. The correlations of most watershed elements increase or decrease directly with the rise of watershed scale, and the values of coefficients tend to be stable at large watershed scales.

**Keywords** Flash flood · FFPI · Spatial autocorrelation · Hazard assessment · Scale effect

---

✉ Jun Du  
dxjx2006@126.com

<sup>1</sup> Research Center on Mountain Torrent and Geologic Disaster Prevention of Ministry of Water Resources, Changjiang River Scientific Research Institute, No. 23, Huangpu Road, Jiang-an District, Wuhan 430010, People's Republic of China

<sup>2</sup> Chengdu University of Technology, State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu 610059, China

## 1 Introduction

The flash flood or mountain torrent here, as a kind of dramatic surface runoff that features abrupt rise and fall on water level, often develops in the hilly or mountain areas and becomes flooding in low-lying areas (National Weather Service, NWS). It is usually caused by natural or human factors such as rainstorm, snow melting or dam failure from natural or artificial dams. The abrupt rise and fall on water level are the most significant feature in flash flood processes, lasting no longer than 6 h with an influential area no more than 400 km<sup>2</sup> in general (Georgakakos 1986; Marchi et al. 2010).

Since most flash floods are the result of concentrated release of water accumulated in a very short time, their peak discharges could be up to 500–2000 m<sup>3</sup>/s or higher even if the watershed area is only tens of square kilometres (Li et al. 1979; Tang and Tie 2009; Liu et al. 2011; Gan et al. 2012; Su et al. 2012). This is also one of the reasons why the flash flood is so powerful and could trigger landslide and debris flow easily. As the statistical data show, only in terms of fatality rate, flash flood is the second major meteorological disaster next to extreme heat in USA (Ashley and Ashley 2008) and also the second major flood disaster next to storm-surge flood in Europe (Borga et al. 2011). Likewise, the casualty and missing people proportion of flash flood/general flooding events in China have been over 60% for a long time and the rate even up to 87.6% in 2010 (Chinese Flood Control and Drought Relief Headquarters 2013). Thus, it is believed that the flash flood is one of the most important threats that cannot be ignored all over the world.

As a mountainous, rainy and undeveloped country in Southeast Asia, Laos suffers frequent flood disasters every year. The flooding from the Mekong River causing great threats to the lower reaches' residents (Plate 2007), and the losses caused by flash floods are also very prominent. According to the historical record, the flash flood events occurred in Oudomxay of northern Laos in 1986 caused 52 casualties (Laos Water Resource and Environment Administration 2008). Especially in recent years, with the accelerated development of hydropower, the potential risk of flash flood disasters increases sharply. Only in 2018, the flash flood happened due to the dam failure in Attapeu of southern Laos caused more than 100 people dead or missing. However, in facing such a severe threat, Laos has not yet established a fundamental risk management system for flash floods prevention from the central government to local administrations, even in conceptual. Therefore, the hazard assessment, as an important premise of risk management, is the first gap should be filled in current status.

The concept of hazard assessment came from the theory of risk on natural disasters, in which the risk was defined as a mathematical expectation of losses in a disaster-prone process (Alexander 1993). Generally, it is believed that the concept of natural disaster risk and its basic formula, proposed by the Department of Humanitarian Affairs of the United Nations in 1992, are the theoretical basis for corresponding assessment work. In this formula, the natural disaster risk is expressed by the product of the hazard degree of natural process and the vulnerability of the corresponding disaster-bearing body. Obviously, as a representation of the occurrence probability of potential disaster process, the hazard assessment is also a very fundamental and important work; it can provide a basis for the planning of villages system and the layout of prevention facilities.

Hazard assessment on flash flood disasters can be roughly categorized into two types, i.e. the process analysis method and the cause analysis method, according to their different analytical perspective. The process analysis method is mainly based on rainfall–runoff simulation and flood routing model, and the output could be specific on inundation range

and velocity distribution of the floods at different recurrences. For instance, Zaharia et al. (2015) estimated the possible affected area, buildings and roads under 1%, 5% and 10% probabilities of flood discharge in a catchment of 427 km<sup>2</sup> of Romania based on the HEC-RAS model; Zhang et al. (2016) simulated a typical mountain torrent event in Yunnan, China, by using the Floodarea model in order to obtain the critical rainfalls under different inundation depths. Compared with the normal models for flash flood forecasting, the process analysis method can provide more details on physical mechanism theoretically because of no limitation on lead time.

Based on runoff hydrograph, some scholars proposed indicators to reflect the characteristics of a specific catchment on hydrological responses to the potential causative rainstorm, such as the rising curve gradient, flood magnitude ratio and flood response time by Bhaskar et al. (2000) and the flashiness by Saharia et al. (2017). These indices usually express a single feature like rising limb, time to peak, magnitude of peak or the comprehensive part of them, and they do reflect some key features of flash flood processes in many cases. However, hazards from flash floods are not only related to some part of the hydrograph, but also to the uncertainties from runoff yield, flood routing, bridge and culvert siltation, formation and collapse of barrier body and so on. The idea can be treated as a special or simplified process analysis method.

Basically, the process analysis method relies heavily on the reliability of data on rainstorm distribution, runoff yield, channel roughness and cross-section morphology. Moreover, some key parameters, such as the infiltration rate in a rainfall–runoff model, need the field work of calibration to achieve higher output accuracy. Therefore, the process analysis method is poorly reliable in completely ungauged areas, especially in current Laos, and it is also hard to be widely promoted.

In cause analysis method, hazard attributes are calculated based on the empirical relationships between the flash flood disasters and their main natural elements generally, and the GIS method of layer algebraic overlay is used commonly to synthesize the layers of different elements, in order to obtain the hazard pattern of assessment objects. The building of indices system and relevant weights design are the two core works in this method. Researchers usually establish an indices system by selecting the data from rainstorm, topography, soil texture, lithology, fracture, vegetation, land use, historical disasters or even some flood process indices, based on the qualitative and quantitative knowledge of the causes of flash flood disasters.

If the historical disaster events are recorded and supported sufficiently, the contribution of each index to the historical events can be estimated by the methods of regression, information model, sensitivity analysis, grey incidence model or other models, so as to assist the weights design (Luo et al. 2011; Du et al. 2015; Costache and Zaharia 2017). Otherwise, the subjective method, objective methods or the combined methods, such as the analytic hierarchy process (AHP), the entropy method and the scoring method are mostly used for scoring. For instance, Youssef et al. (2011) scored the risk-related indices from 11 sub-basins in Egypt based on the relationships between the bifurcation ratio and the drainage density or drainage frequency; Zeleňáková et al. (2015) determined the importance of flash flood causal factors in a Slovakia basin by using AHP; Zeng et al. (2016) used a mixed method of the AHP and the information entropy theory for generating a FFPI map in Yunnan, China.

Besides, the machine learning related methods, such as the logistic model, support vector machine, decision tree or random forest, can also be used to obtain the output directly based on an appropriate setting of the input/output parameters, rules or classification logic. Wang et al. (2015) evaluated the flood hazard in Dongjiang River, Southern China, based

on 5000 samples by using the random forest model and found that the maximum 3-day precipitation, runoff depth, typhoon frequency, DEM, and topographic wetness index are the most important five indices. Costache (2019a, b) compared different machine learning models and the related hybrid algorithms on flash flood potential assessment in Romania, and several quantitative methods, such as the ROC Curve model, were used to validate the outputs. The results indicated that the hybrid models performed excellent (AUC values > 0.87) and presented significant advantages with the subjective weighting method listed in the study. It is believed that the machine learning technology would play a more and more important role on natural hazard assessment in the future, with the accumulation of historical data and the optimization of various classification algorithms.

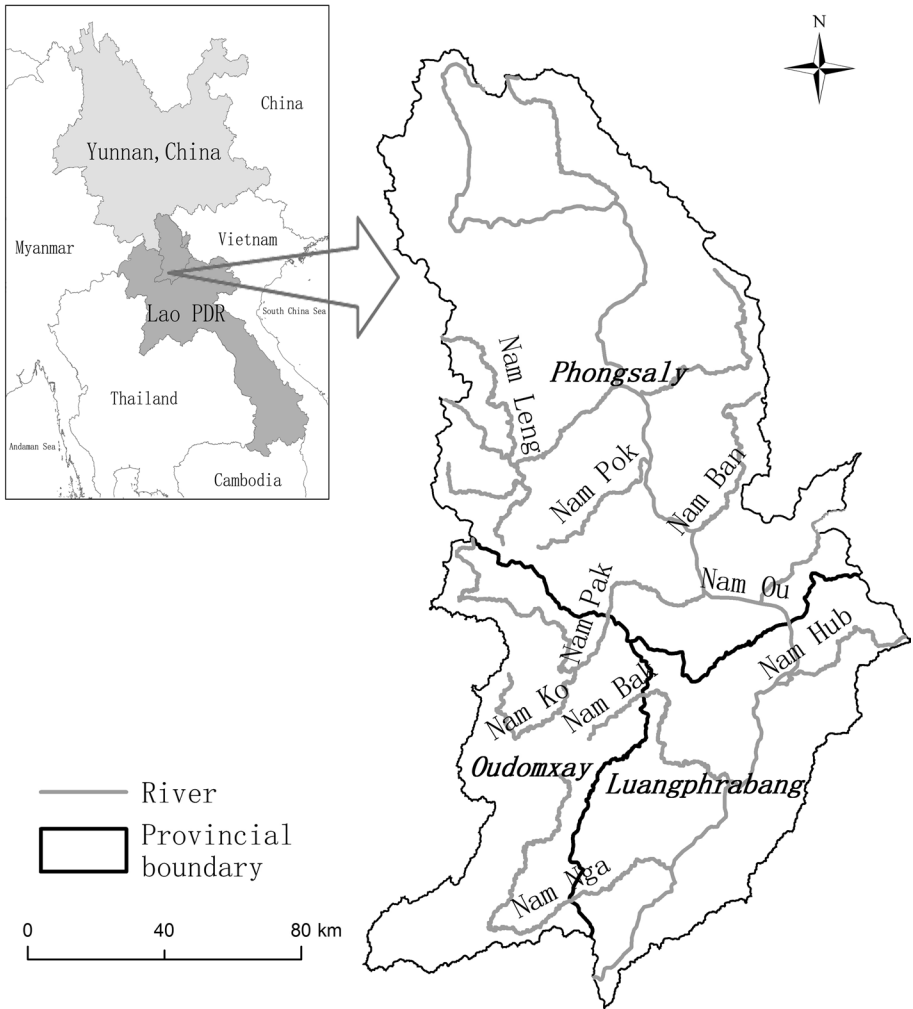
Compared with the process analysis method, the cause analysis method is more mature on theoretical framework, more flexible on application and lower requirements for data accuracy. Although the specific range of flooding cannot be provided, the output of regional hazard pattern can still be very useful for preliminary risk management, and it is easier to be verified. Therefore, for current Laos where lacks the basic prevention and event record system on flash flood disasters in most area, it is believed that the cause analysis method is more suitable for local promotion and application.

The Nam Ou River, one of the most important tributaries of the lower Mekong in northern Laos, is rich on hydropower and ecological resources. In addition, its population and navigation capacity also rank the first class among the main rivers in Laos. Because of the high relief on topography and the intensified and frequent rainstorms, the Nam Ou River Basin is always highly susceptible to flash floods and threatens the local residents' production and life heavily. Accordingly, a comparative study focused on flash flood hazard assessment of the Nam Ou River Basin is implemented here based on the methods such as GIS, flash flood potential index (FFPI), regression and curve number calculation in this paper, in order to (1) obtain a reliable map on hazard pattern for local flash flood disaster prevention; (2) compare and analyse the applicability of the assessment methods used in this study for the area with little fundamental work.

## 2 Methodology and data preparation

### 2.1 Study area

The Nam Ou River originates from the mountain area at the frontier between the Jiangcheng County, Yunnan, China, and the Phongsaly, Laos (Fig. 1). It is one of the most important tributaries in the lower Mekong River in Laos, drainages an area of 24,600 km<sup>2</sup> and has a length with 475 km in mainstream, ranking first among the 12 biggest rivers except the Mekong mainstream in Laos, respectively. Besides, the regional population, i.e. 404,000 persons in 2015, and the ecological and navigation scales are also ranking the first class among the above rivers. The average annual discharge of the Nam Ou River is about 610 m<sup>3</sup>/s, contributing 3% discharge to the whole Mekong River only. However, its hydropower potential is up to 1670 MW, accounting for 11.3% of the technically available hydropower resources of the Mekong Basin. In recent years, with the confirmation of the "corridor country" economic strategy in Laos national level, i.e. promoting cross-border trade from inland areas, the Nam Ou River, as a border river among China, Laos and Vietnam, would play a more and more important role in future regional trade.



**Fig. 1** Nam Ou River Basin, Lao PDR

Although the development prospects of the Nam Ou River Basin are highly expected, the current situation on flood control is not optimistic. The regional topography is dominated by mountainous area with typical features of tectonic denudation, erosion accumulation and karst landform. Slope gradient more than 30% accounts for about 60% of the total area of the basin, while only 2.2% of the area inner basin is flat. On the other hand, influenced by warm and wet airflow from the Bay of Bengal and the Beibu Gulf, tropical monsoon climate prevails in the basin, with abundant rainfall and distinct dry and wet season characteristics. The average annual rainfall in the basin can reach 1610 mm, which decreases gradually from north to south. The wet season rainfall from June to November can account for 80% of the whole year. In addition, the regional surface is occupied widely by Acrisol, a kind of thin, sticky and low-fertility soil, which is not conducive to the traditional cultivation and tillage. Thus, the shifting cultivation, mainly slash-and-burn method, was popular in this area naturally, which in turn intensifies the destruction of forest resources, leading to runoff increased locally. Based

on these backgrounds, the regional floods, especially the flash floods, occur frequently in the Nam Ou River Basin, causing a great threat to local residents. However, due to the low level of social productivity and the lack of financial and technical support, there is still a long way to go on flash floods prevention for local administration, currently.

## 2.2 Methodology

The hazard pattern of regional flash floods can be quantitatively described by the methods such as FFPI, regression or other runoff yield methods. Considering the current data support, the convenience of promotion and the reliability of methods, herein the FFPI, the calculation of curve number–rainfall erosivity and the so-called extrapolation method are used to make a comparative analysis of the hazard distribution of flash floods in the Nam Ou River Basin. Among these three methods, the one with the results closer to the actual disaster situation would be the preferred plan.

### 2.2.1 FFPI method

The FFPI is an assessment method of flash flood occurrence based on factor empirical analysis, which was developed in 2003–2004 by the Colorado Basin River Regional Forecast Center, NWS. The original intention of developing this method is to participate in forecasting and pre-warning of flash floods as a supplement to LFFG, i.e. Lumped Flash flood Guide system. However, since its principles and considerations are of indicative significance to describing the threat level of regional flash floods, it can also be used for hazard assessment of flash flood disasters (Stanimir et al. 2017). The principle of FFPI is to quantize and score the effect of each element by 10 degrees based on physical and empirical relationships between the four elements, i.e. the soil type, land use, slope and vegetation cover, and the runoff generation (Smith 2010); then, with the support of GIS technology, the distribution map of flash flood occurrence potential can be calculated through the algebraic overlay of layers of the four elements. The formula for FFPI is as follows:

$$I_{\text{FFPI}} = (k_1 \cdot M + k_2 \cdot L + k_3 \cdot S + k_4 \cdot V) / N \quad (1)$$

where  $I_{\text{FFPI}}$  is the specific value of FFPI with a theoretical range of 0–10: the higher the value is, the greater the occurrence potential of flash floods would be;  $M$ ,  $L$ ,  $S$  and  $V$  indicate the slope gradient, land use, soil property and forest cover, respectively, and the numerical range of these indicators is also 0–10;  $k_1$ ,  $k_2$ ,  $k_3$  and  $k_4$  are the respective weight coefficients, generally set as equal weight; considering that the slope gradient plays a more important role on flash flood development commonly, the  $k_1$  is usually given a higher value;  $N = k_1 + k_2 + k_3 + k_4$ .

Since the release of FFPI, the model and its related derivatives have been applied in many countries or regions including the USA, China, Vietnam, Japan and Serbia. Currently, the FFPI may be one of the most influential cause analytical methods of flash flood occurrence worldwide.

### 2.2.2 Curve number–rainfall erosivity method (C–R method)

The famous NRCS runoff curve number is an indicator proposed in 1972 by the US Natural Resources Conservation Service, i.e. Soil Conservation Service formerly, to reflect the

regional potential runoff yield. Its principle is to propose an empirical parameter reflecting the infiltration or runoff yield performance of the underlying surface according to the hydrological relationship between the soil, land use type and the runoff yield. Theoretically, combined with the data of real-time rainfall and antecedent soil moisture (Table 1), the parameter can be used further to calculate the specific amount of runoff yield from each rainfall event.

In this study, only the general scenario is calculated, i.e.  $CN$  value of AMCH ( $CN_2$ ), which can be obtained with reference to the  $SCS$  handbook ( $SCS$  1972). Since the effect of slope is not considered in traditional  $CN$  method, the Sharpley and Williams (1990) method (formula 2) is used to obtain the improved  $CN_2$  value, i.e.  $CN_{2\alpha}$

$$CN_{2\alpha} = CN_2 \frac{322.79 + 15.63(\alpha)}{\alpha + 323.52} \tag{2}$$

where  $\alpha$  indicates the slope gradient.

After the  $CN_{2\alpha}$  is obtained, the amount of runoff yield in each rainfall event can be calculated by the formulas (3)–(5)

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \tag{3}$$

$$I_a = 0.2S \tag{4}$$

$$S = \frac{25400}{CN} - 254 \tag{5}$$

where  $Q$  indicates the amount of runoff yield in each rainfall event;  $P$  indicates the amount of each event;  $I_a$  indicates the initial interception; and  $S$  indicates the potential maximum moisture capacity.

Unfortunately, only the monthly precipitation data in recent decades were supported by the Laos government; thus, the specific runoff depth cannot be calculated in this study. For fully used these data, the Wischmeier’s empirical formula on rainfall erosivity (formula 6), in which the storm-related rainfall intensity is also considered, is used instead to reflect the rainfall power on flash flood development.

$$R = \sum_{i=1}^{12} \left( 1.735 \times 10^{\left( 1.5 \log \frac{p_i^2}{P} - 0.8188 \right)} \right) \tag{6}$$

**Table 1** Antecedent soil moisture condition (AMC) used to determine  $CN$  value. *Source:* Chow et al. (1988)

Grade AMC	Antecedent precipitation in 5 days (cm)	
	Non-monsoon period	Monsoon period
I. Optimum soil conditions from plastic limit to wilting point	Lower than 1.25	Lower than 3.5
II. Average annual flood	1.25–2.75	3.5–5.25
III. High or low precipitation and low temperature 5 days before the rainstorm	Higher than 2.75	Higher than 5.25

where the  $P$  indicates the mean annual precipitation (mm) and the  $P_i$  indicates the monthly precipitation (mm). The  $R$  is the rainfall erosivity in the unit of  $100 \text{ ft t in ac}^{-1} \text{ h}^{-1} \text{ a}^{-1}$ .

Once the layers of  $CN_{2\alpha}$  and  $R$  are prepared, the hazard outcome based on the C–R method can be obtained by the overlay of the two layers in GIS.

### 2.2.3 Extrapolation method

This method is based on the relationships between flash flood disasters and their impact factors in the Yunnan Province, China, where borders the Nam Ou River Basin and shares a similar geographical environment in general. In this method, the stable empirical relationships between the disaster events and the development elements should be established first by regression or other mathematical tools, and then, the relationships would be extrapolated into the target area, i.e. Nam Ou River Basin here, with the inputs of corresponding local data of impact factors.

As one of the most commonly used models in the world, the general linear regression is selected as a basic model of extrapolation method. Recently, the logistic model is proved to be very useful on flash flood assessment (Costache 2019a; Xiong et al. 2019); thus, the model is also chosen for comparing. Besides, considering that the flash floods usually develop in clusters widely in mountain area (Du et al. 2016; Xiong et al. 2019), presenting significant autocorrelation characteristic and in line with the first law of geography (Tobler 1970), the spatial regression model, i.e. the spatial lag model or the spatial error model in this study (Anselin 1988), was also used as an important statistical tool in addition to the ordinary multiple regression and the logistic model. Since the latter two models are well known in academic world, the following will focus on the fundamental principles of spatial regression model and its general parameter settings of this study.

The classical linear regression model (formula 7) consists of dependent variable  $Y$ , constant term  $a$ , independent variable  $X$  and its coefficient  $b$ , as well as error term  $e$ . In this model, the dependent variable  $Y$  is only affected by the independent variable  $X$ , and the error terms  $e$  are completely independent and subject to normal distribution. However, for the phenomena of spatial autocorrelation, the use of a classical regression model may lead to fitting failure or inadequate goodness of fit, resulting from two situations: firstly, the dependent variable  $Y$  is not only affected by independent variable  $X$ , but also affected by the adjacent dependent variable  $Y'$ , i.e. spatial autocorrelation exists between dependent variables; secondly, although spatial autocorrelation of the dependent variable itself is not significant, there is a spatial autocorrelation between error terms  $e$  for some reasons. For the first situation, the spatial lag model (SLM) can be used for fitting (formula 8), which has a spatial weight matrix  $\rho WY$  of dependent variable added on the basis of classical regression model; for the second situation, the spatial error model (SEM) can be used for fitting (formula 9), which has a spatial weight matrix  $\lambda W\epsilon$  of error term added on the basis of classical regression model

$$Y = a + bX + e \quad (7)$$

$$Y = a + \rho WY + bX + e \quad (8)$$

$$Y = a + bX + \lambda W\epsilon + e \quad (9)$$

The spatial weight matrix  $W$  is an important concept in spatial autoregressive simulation. It is a set of matrices built to describe the spatial adjacent relation of an object or unit.



Suppose that there are  $n$  neighbouring units around a central unit, each neighbouring unit gives the central unit an influential weight of  $1/n$ ; when  $n=0$ , it means that the spatial autocorrelation is not necessary to be considered for the central unit. There are three forms of adjacent relation: The Queen contiguity, i.e. units contact by point and edge; the Rook contiguity, i.e. units contact only by edge; the Distance contiguity, i.e. units contact according to a certain distance threshold. It is believed that the selection of adjacent relation should follow a principle, which is the less neighbouring units, the better (Florax and Rey 1995; Griffith 1996; Meng et al. 2007). Accordingly, the Rook contiguity is used here to determine the number of adjacent units.

## 2.3 Data sources and layers calculation

### 2.3.1 Rainfall erosivity

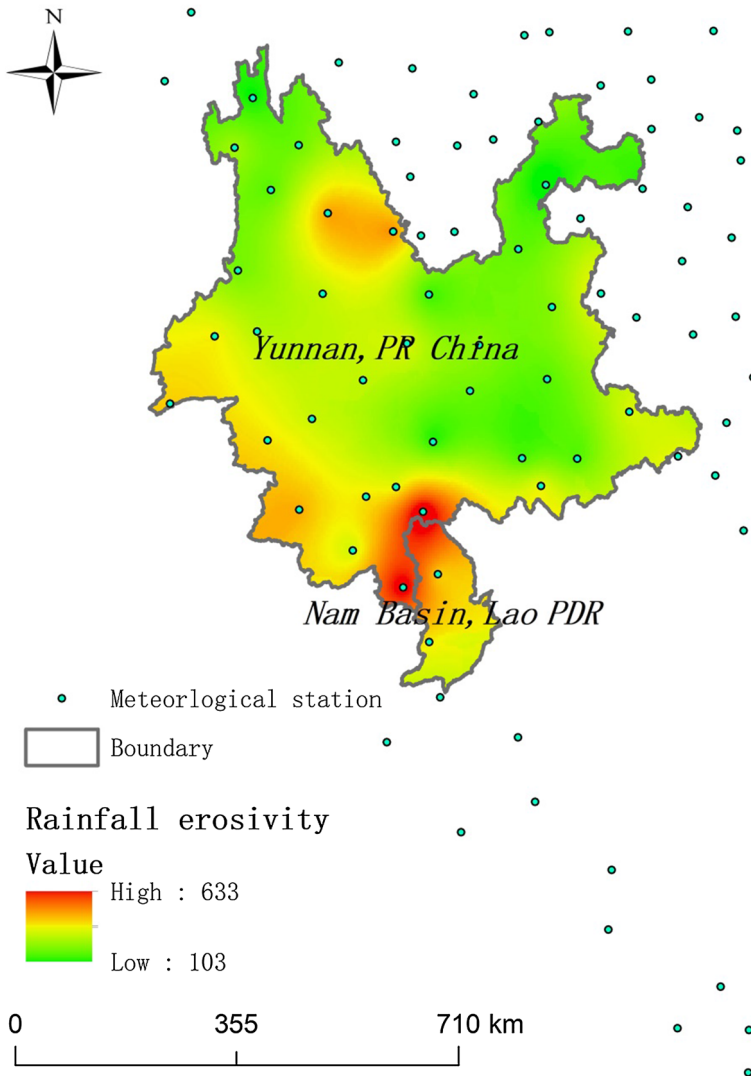
Based on the monthly rainfall data, i.e. 13 stations during 1994–2009 in Laos and 68 stations during 1951–2002 inner and around Yunnan, provided by the Ministry of Natural Resources and Environmental Protection of Laos and the China Meteorological Administration, respectively, the rainfall erosivity of each station was calculated by the empirical formula (formula 6) of Wischmeier and interpolated by the Kriging method in ArcMap so as to obtain the raster layers (Fig. 2).

### 2.3.2 Soil and land use indices

The effect of soil on runoff yield is mainly reflected by infiltration performance, which is closely related to the soil texture. Studies show that the soil with better pore development usually presents higher infiltration level, and the clay, silt and sand grains support the pore development differently due to their different mechanical compositions and organic contents (Helalia 1993; Shan et al. 1998). It is believed that the clay presents low-level infiltration because the soil pores can be hardly developed for its smaller grain size trending to form crust easily when exposed to water (Helalia et al. 1988), while the sandy soil is on the contrary. Accordingly, the soil potential of flash flood development in different regions can be defined based on the soil texture.

FFPI quantifies the soil potential of flash flood development into 10 levels based on the mechanical components of different types of soil (Table 2). In this study, the vector layer of soil texture distribution in the Nam Ou River Basin and Yunnan was generated by GIS based on the data from the World Soil Properties Database and the official literature (Laos Department of Water Resources 2016), and then the scoring process was implemented according to the FFPI standard on soil (Fig. 3a).

The land use types also affect runoff yield theoretically. On one hand, the urban and industrial land presents more runoff yield than that of other land use types due to the hardening of the earth surface. On the other hand, unreasonable development activities may lead to a regional reduction in vegetation coverage and aggregation of soil erosion, which increases the runoff yield to a certain extent. The FFPI on land use sets the quantitative levels according to the empirical effects of different land use types on runoff yield (Lin et al. 2016). In this study, the land use types of Yunnan in China and Nam Ou River Basin in Laos are assigned with values based on the standards (Table 3, Fig. 3b). Relevant data derive from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences and Laos Ministry of Natural Resources and Environmental Protection.

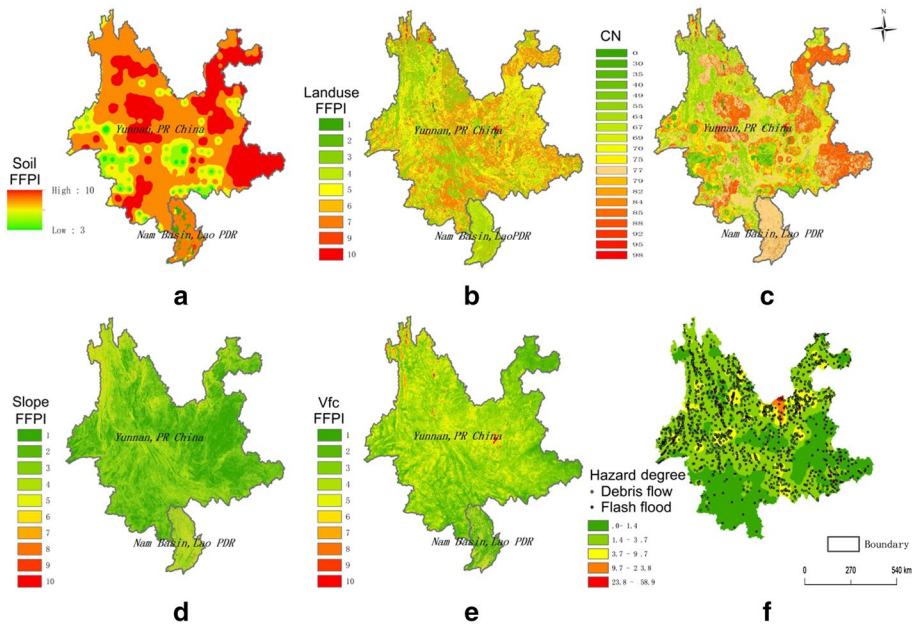


**Fig. 2** Distribution map of rainfall erosivity in Nam Ou River Basin and Yunnan

The *CN* value reflects the potential runoff generation on different kinds of underlying surfaces. Specifically, it is an indicator reflecting the maximum retention capacity of precipitation (Smith 2010), including the interception by vegetation and soil and the infiltration of underlying surface. The higher values on *CN* mean lower infiltration rates and higher amount on runoff yield. The soil characteristic and land use are two key aspects of *CN* values estimation. By checking the hydrological soil data (Table 4) determined by the US Department of Agriculture in 1986 based on the soil infiltration rate from high to low, the hydrological soil groups of the study area can be determined. Meanwhile, the original *CN* assignment table and the corresponding vector layer can be

**Table 2** Scoring basis of soil FFPI value. *Source:* Smith (2010)

FFPI index	Sand%	Silt%	Clay%
1	92	5	3
2	83	11	6
3	74	17	9
4	65	24	12
5	56	30	15
6	47	32	21
7	37	36	27
8	28	27	45
9	19	27	55
10	10	30	60



**Fig. 3** Layers of underlying surface elements and hazard degree in Nam Ou River Basin and Yunnan

obtained (Fig. 3c) by referring to historical research (Hadley 2002; Ebrahimian et al. 2009).

### 2.3.3 Slope

The steep mountain causes various topographic rains, especially the night storm which causing disaster easily. Besides, it can also accelerate the processes of runoff generation and confluence. Furthermore, the mountainous area usually leads to large-scale landslides and collapses, providing source and path for the occurring of debris flows. Hence, the topography affects the development of flash flood significantly and directly. The FFPI

**Table 3** Assignment of CN and FFPI values based on land use types and hydrological soil properties

Land use types in China	HSGs				Land use FFPI	Land use types in Laos	HSGs		Land use FFPI
	A	B	C	D			A	D	
Towns or industrial zones	98	98	98	98	10	Cities or building areas	98	98	10
Bare rocks	98	98	98	98	10	Other agricultural land	64	85	6
Bare land	89	94	92	95	9	Bamboo forest	30	77	4
Shrub land	67	85	77	88	7	Shrub land	67	88	7
Dry farmland	64	82	75	85	6	Agricultural plantation	–	85	6
Paddy field	64	82	75	85	6	Paddy field	64	85	6
Grassland	49	79	69	84	5	Grassland	49	84	5
Orchard	64	82	75	85	6	Mixed fallen leaves	30	77	4
Open woodland	30	70	55	77	4	Open woodland	30	77	4
Woodland	30	70	55	77	3	Non-forest land	30	77	3
Wetland	20	40	35	40	2	Dried dipterocarp	30	77	3
Swamp	20	40	35	40	2	Low-dry evergreen forest	30	77	3
Open water area	0	0	0	0	1	Open water area	0	0	1
Perennial ice and snow	0	0	0	0	1				

**Table 4** Hydrological soil groups of disturbed soil profiles. *Source*: NRCS (1986)

HSG	Soil texture
Group A	Sand, loam, or sandy loam
Group B	Silt loam or loam
Group C	Sandy clay loam
Group D	Clay loam, silty clay loam, sandy clay, silty clay or clay

system divides the gradient of the slope into 10 grades at equal intervals, in which Grade-1 is between 0 and 10% and Grade-10 is 90% and above (Smith 2010). Based on the revised ASTER 30 m DEM, this study divides the gradient of Yunnan and Nam Ou River Basin into different FFPI grades and generates the vector layer in GIS (Fig. 3d).

### 2.3.4 Forest coverage

The growth of forest can retain the rainfall and runoff effectively by canopy interception, water storage from fallen leaves and dry branches as well as soil structure improvement by the root system. Though these effects could be weakened or even reversed due to the vegetation's water conservation under continuous storm floods, it is believed that the forest vegetation is generally effective in inhibiting the formation of flash floods. Similar to the slope grades, FFPI system divides the forest coverage rate into 10 levels at equal intervals, of which Gradel-1 is between 91 and 100% and Grade-10 between 0 and 10%.

With the monthly synthetic product from MODIS-500 m-NDVI during 2000–2002, the forest coverage rate in Yunnan and Nam Ou River Basin was mapped and scored based on relevant standards of the FFPI system (Fig. 3e).

### 2.3.5 Shape and area of watershed units

The shape and area of the watershed are important hydrological parameters to the development of flash flood. In this study, the shape coefficient  $K_e$ , i.e. the ratio of actual perimeter of the watershed to the hypothetical perimeter of a circle with the same area, is used to reflect the hydrological influence from the watershed shape. When the shape of watershed becomes more and more like a circle, the  $K_e$  value is closer to 1, and the duration of confluence is shorter, which is more favourable to the formation of flash floods. Once the scales of watershed units been decided, the  $K_e$  value and the area can be directly calculated in GIS with the raster calculator tool.

### 2.3.6 Flash flood hazard pattern in Yunnan based on historical disaster records

According to the historical disaster events reported by the administration of Yunnan in the project of National Planning for Flash Floods Disasters Prevention and Control in China, 738 flash flood points (952 disaster events) and 1481 debris flow points (1532 disaster events) caused by flash floods were collected and recorded during 1949–2002. Unfortunately, no sufficient messages, such as the losses, the estimated frequency or scale of each flood or rainfall, the involved areas, were provided in this database except the locations. In addition, the historical disasters represent only where the events have happened but no clear indication of the disaster-prone range.

For presenting the disaster-prone or hazard areas of flash floods in Yunnan more closely (Fig. 3f), the method of spatial interpolation was used to complement necessary information based on the assumption of that area adjacent to disaster points share the similar disaster-prone environment. The assumption is rational according to the spatial autocorrelation of flash flood development.

The specific steps are as follows: (1) a concrete hazard value is set for each flash flood point according to its frequency. For instance, a point where the flash flood occurred for  $n$  times shall have the hazard value of  $n$ ; (2) considering that the debris flow is more complex and causes greater threats to local residents commonly, the hazard value for each debris flow point is three times of the frequency; (3) the areas have no events recorded definitely would be interpolated points with the hazard value of 0, the points should be distributed in space as evenly as possible; (4) the interpolation is performed with the inverse distance weighting (IDW) method, and the best parameter setting should be the minimum value of root-mean-square error ( $RMSE$ ) of the result.

## 3 Results and discussion

### 3.1 Scale effect of the watershed unit

It is common to explore the correlation between two elements by their geographical relevance in spatial analysis. However, the relevance might be changed at different spatial statistic scales, i.e. there might be a scale effect between the elements on spatial

correlation. The previous studies have paid great attention on the scale effect of single geographical phenomena (Lam and Quattrochi 1992; Xu and Yan 2005; Shi 2008), instead of the changes of spatial relationships between elements with different scales. Besides, the proper statistic scale, i.e. the most representative area scale of watershed units, in Yunnan also needs to be clarified in the extrapolation method of this study. Accordingly, a scatter analysis was made by using the tool of hydrology in GIS. In this tool, various statistic scales of watershed units can be generated based on different flow accumulation thresholds. Once the scales of watershed units were determined, the geo-statistic characteristics of each element, including the potential impact factors and the average hazard degree based on historical events recorded of Yunnan, could be obtained and the Pearson correlation analysis between each factor and the hazard degree would be processed; then, the changes in the correlation coefficients at different scales could be presented.

Theoretically, potential impact factors, such as land use and slope, play important roles in flash flood development. However, in spatial analysis whether the element becomes the key or 'bottleneck' factor depends on the so-called short board effect, i.e. only the element that really restricts the entire development system can be the key factor in a specific area. For example, it is believed that the rainstorm is an important trigger factor for a specific flash flood event, but to an area that basically presents the same rich level on rainfall characteristics, the role of rainstorm can hardly be detected in a spatial analysis because the flash flood events are not sensitive to rainstorm distribution in macroscopical sense. In other word, if the role of rainstorm in commonsense could not be detected in a specific area, it does not mean the rainstorm has no impact on flash flood development; only indicates that it is not the 'bottleneck' factor in this region.

In this case, the indices of land use FFPI and the rainfall erosivity in Yunnan present negative correlations with the hazard degree, i.e. the greater the rainfall erosivity or land use FFPI is, the smaller the frequency of events will be (Table 5), which is contrary to the common recognition. Obviously, it indicates that these indices are definitely not the key factors on flash flood development in Yunnan. Similarly, the shape coefficient of  $Ke$  also presents a false correlation in this area. Generally, the larger the  $Ke$  is, the shape of watershed tends to be narrower and longer, and the capacity of runoff yield and confluence becomes weaker. However, the correlation analysis shows that the coefficients between  $Ke$  and the hazard degree are positive at most statistic scales.

Except for the obvious false correlation of land use FFPI, rainfall erosivity and  $Ke$ , the rest indices of slope FFPI, soil FFPI, vegetation FFPI as well as the watershed area at different statistic scales were connected with the hazard degree to form a scatter diagram on correlation coefficients (Fig. 4). It seems that except the soil FFPI, the correlation between the other factors and hazard degree basically increases with the larger of the statistic scale, of which the watershed area has the highest correlation level at 555 km<sup>2</sup> average watershed area and after that remains stable, and the vegetation FFPI varies similarly, while the correlation of the slope FFPI reaches the highest level at the largest statistic scale.

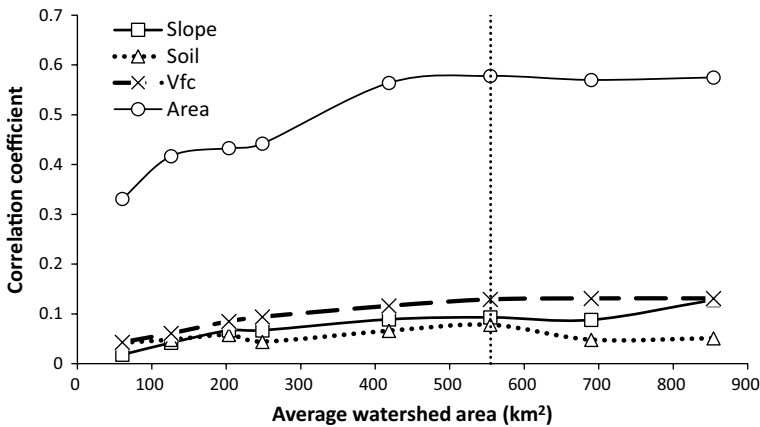
Although the water amount of flash flood comes from the mountainous parts of a watershed mostly, the disasters occur mainly in low-lying parts where human activities are relatively frequent. Hence, the effects from natural elements such as landform and vegetation cover on flash flood events are easily disturbed by human activities at small spatial scales, showing low correlation levels. However, with the increase in spatial scales, the effects from natural characteristics of watershed become more and more obvious, and the corresponding correlations increase accordingly. Since the soil properties are not necessarily related to the formation of a

**Table 5** Correlation coefficients between potential impact factors and the flash flood hazard based on historical data on different scales in Yunnan

FAT	AWA	Landuse	Slope	Soil	Vfc	Re	Area	Ke
5000	61	-0.036	0.018	0.042	0.043	0.043	0.331	-0.028 <sup>a</sup>
10,000	126	-0.065	0.042	0.048	0.061	-0.059	0.417	-0.021 <sup>a</sup>
15,000	204	-0.095	0.067	0.058	0.085	-0.079	0.433	-0.017 <sup>a</sup>
20,000	249	-0.118	0.067	0.044 <sup>a</sup>	0.094	-0.084	0.442	0.024 <sup>a</sup>
30,000	418	-0.123	0.089	0.066 <sup>a</sup>	0.116	-0.102	0.564	0.133
40,000	555	-0.162	0.093	0.078 <sup>a</sup>	0.129	-0.094	0.578	0.223
50,000	690	-0.175	0.088	0.048 <sup>a</sup>	0.131	-0.119	0.570	0.246
60,000	854	-0.18	0.128	0.051 <sup>a</sup>	0.131	-0.123	0.575	0.261

FAT flow accumulation threshold, accumulation units in GIS, AWA average watershed area, km<sup>2</sup>, Landuse land use FFPI, Slope slope FFPI, Soil soil FFPI, Vfc vegetation fractional cover, Re rainfall erosivity, Area watershed area, Ke coefficient of basin shape

<sup>a</sup>Means the value is insignificant statistically



**Fig. 4** Scatter plot of correlation coefficients between main potential impact factors and the flash flood hazard based on historical data in Yunnan during different scales. Note: the vertical dashed line indicates the selected watershed scale in the following study

watershed, and the spatial heterogeneity of soil FFPI is also very high in Yunnan (Fig. 3a), the correlation between soil FFPI and hazard degree presents more complex trend (Table 5).

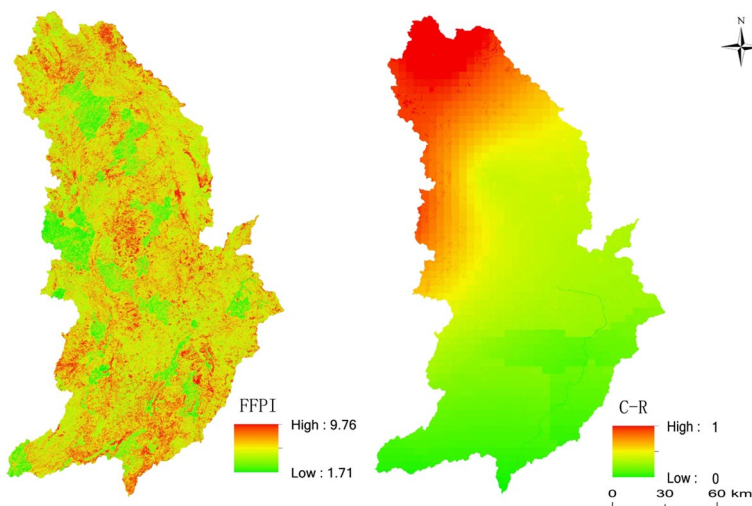
As a summary, most of the indices, i.e. watershed area, vegetation FFPI and slope FFPI, have the highest or relatively stable correlation level with the development of flash flood events at the scale of average watershed area of 555 km<sup>2</sup>. Accordingly, the corresponding flow accumulation threshold, i.e. 40,000 here, would be set as the basis for watershed division in the subsequent extrapolation study.

### 3.2 Hazard maps based on the three methods

Combined with the data layers of various elements prepared previously and the method of algebraic overlay in GIS, the hazard pattern of flash floods in the Nam Ou River Basin based on FFPI algorithm can be obtained by the formula (1) (Fig. 5, left). Similarly, by the product overlay of the layers of revised *CN* and rainfall erosivity, the hazard map based on C–R algorithm can also be achieved (Fig. 5, right). The two outcomes present the same direction of value, i.e. the larger the hazard value is, the higher the regional occurrence probability or hazard degree of the flash flood is. The data presented in FFPI outcome remain unchanged, while that in C–R outcome is normalized in 0–1.

Normally, the FFPI outcome reflects the comprehensive effects of the underlying surface on flash flood development. In the Nam Ou River Basin, the influence of soil infiltration properties is highlighted. The high content of gravel and sand in cambisol weakens the runoff yield; thus, the local occurrence rate of flash floods is significantly lower than that of the area no cambisol developed (Fig. 5, left). Meanwhile, the topography also plays an important role, causing the high-relief mountain areas in the northern, central and southern parts of the basin at relatively high hazard degrees. In comparison, the hazard outcome based on C–R method has a more obvious feature on spatial trend: the hazard degree gradually decreases from northwest to southeast, reflecting the direct influence of rainfall pattern on flash flood development (Fig. 2).

The analysis in part 3.1 indicated that the FFPI indices of soil, *vfc*, slope and the area index in Yunnan were correlated with the development of flash flood effectively. Therefore, in the method of extrapolation, the four indices as well as the *CN* value were used as the independent variables to build the regression models, i.e. the spatial lag model, spatial error model, general regression model and logistic model, for understanding the fundamental empirical relationships between the potential impact factors and the flash flood disasters in this area, and the basic statistic units are set in 555 km<sup>2</sup> of average watershed area accordingly.



**Fig. 5** Hazard maps based on FFPI and C–R method



The flash flood hazard degree in Yunnan was divided into three levels based on an objective classification method, i.e. the Jenks Natural Breaks, which can provide the best arrangement of values by reducing the variance within classes and maximize the variance between classes (Esri 2010). Taking the three-level hazard characteristics as the dependent variables (Table 6), the multinomial logistic model (MLM) was built for hazard assessment firstly. Although the model was significant statistically, it performed bad on identifying the watershed units with medium and high hazard level in test (Table 8). Except for the large amount of noises of the samples, another reason for this is the independent variables are all continuous or quasi-continuous and can hardly be categorized with rational basis. Hence, the units with medium and high hazard levels were combined into one category for building the binary logistic model (BLM) in order to obtain better prediction performance, and the model is acceptable after some noise data were eliminated (Tables 6, 7, 8). However, when the model was extrapolated into the Nam Ou River Basin, all the watersheds were identified as the units with medium and high hazard level because of the high value on regional slope, which means all of the watersheds are prone to develop flash floods. The

**Table 6** The characteristics of dependent variables in logistic models

MLM	Samples number	Percentage (%)	BLM	Samples number	Percentage (%)
Hazard 3			Hazard 2		
1	320	59.5	1	273	62.2
2	179	33.3	2	166	37.8
3	39	7.2			
Sum	538	100.0	Sum	439	100.0

**Table 7** Results of independent variables in logistic models

MLM	-2 log likelihood	Chi square	Sig.	BLM	$\beta$	SE	Exp ( $\beta$ )	Sig.
CN	876.201	16.511	0.000	CN	0.229	0.029	1.258	0.000
Slope	891.199	31.509	0.000	Slope	3.710	0.387	40.849	0.000
Vfc	878.686	18.995	0.000	Vfc	2.651	0.324	14.168	0.000
Area	874.549	14.858	0.001	Area	0.000	0.000	1.000	0.141
Intercept	907.893	48.203	0.000	C	-35.809	3.821	0.000	0.000

$\beta$ , SE and Exp ( $\beta$ ) denote the partial regression coefficient, standard error and odds ratio, respectively

**Table 8** Test results of the logistic models

MLM					BLM				
Observed	Predicted			Correct%	Observed	Predicted		Correct%	
	1	2	3			1	2		
1	287	33	0	89.7	1	238	35	87.2	
2	134	44	1	24.6	2	52	114	68.7	
3	29	9	1	2.6					
Overall				61.7	Overall			80.2	

accuracy of this result is hard to judge, but it cannot provide the hazard distribution pattern in the target area at least.

As for the other three regression models, according to the three statistic indicators proposed by Anselin (1988), i.e. the log-likelihood value, the Akaike info criterion and the Schwarz criterion, as well as the corresponding criteria for model judgement, the larger the logarithm likelihood value is or the smaller the statistic value of Akaike and Schwarz criterion are, the better the fitting performance of the model will be. The logarithmic likelihood value of the general regression model is  $-2124.98$ , much smaller than that of the spatial regression models (SLM,  $-765.20$ ; SEM,  $-768.40$ ), while its statistic value of Akaike and Schwarz criterion is much larger than that of the corresponding spatial regression models. These indicate that in this study the general regression model is indeed worse than the spatial regression model on fitting performance due to the insufficient consideration on spatial autocorrelation. Similarly, the indicators of spatial lag model are slightly better than that of the spatial error model, meaning the SLM is the preferred plan among the three regression models (Table 9).

Based on the results of statistic indicators of each factor in SLM, the spatial autocorrelation factor of dependent variable presents the most significant level among the independent variables, followed by slope FFPI and Vfc FFPI, while the *CN* value that reflecting the comprehensive properties of land use and soil presents the lowest significance. It seems that besides the strongly influenced by the spatial autocorrelation of processes themselves, the flash flood disasters in Yunnan are mainly affected by the elements of landform and forest cover. Assuming that the Nam Ou River Basin and Yunnan share the similar natural environment on flash flood development, the corresponding data of independent variables can be substituted into the SLM model to output the hazard map on flash floods in the Nam Ou River Basin.

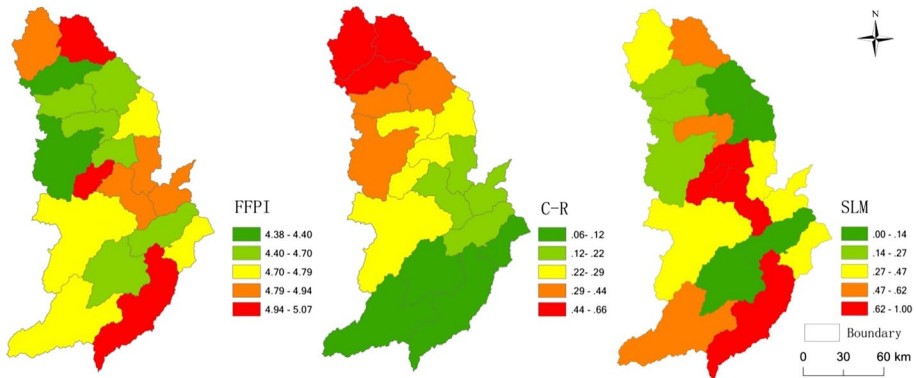
### 3.3 Comparison and validation

For comparison, the Jenks Natural Breaks is used to classify the three hazard maps based on the methods of FFPI, C–R and SLM, respectively (Fig. 6). It seems that the fundamental distributed characteristic in the maps of FFPI and SLM is similar: the low-hazard areas are all distributed in the northern part of the basin, while the high-hazard areas are concentrated in the central and southern parts. Differences between the two maps are only focused on some local changes of hazard degrees. By contrast, the hazard map based on C–R method presents more differences on spatial trend, i.e. the hazard levels of watershed units gradually decrease from northwest to southeast. In principle, the FFPI and SLM maps mainly present the effects of an underlying surface on flash flood development, while the C–R map basically reflects the role of rainfall pattern. Therefore, methodologically speaking, the three methods have different underlying assumptions, and the difference in the results shows that either the underlying assumption of SLM and FFPI, or the C–R map assumptions may not be valid. However, since the annual precipitations in the whole areas of Nam Ou River Basin are very rich and concentrated, the rainfall can hardly be the element that restricts the regional development of flash floods; thus, the map based on C–R plan may not represent the hazard pattern accurately.

For further comparing the reliability of the hazard maps based on different plans, the flood disaster events in recent period were collected and summarized by the field investigation, literature research (Laos Department of Water Resources 2016; Changjiang Institute of Survey, Planning, Design and Research 2018) and interviews with local officials in

**Table 9** Comparison of results in models of OLS, SLM and SEM

Model	$R^2$	Log likelihood	Akaike info criterion	Schwara criterion	Independent variable	Coefficient	z value/t value	P value
SLM	0.77	-765.20	1540.40	1561.90	W	0.90	51.01	0.00
					C	-1.30	-2.70	0.01
					Vfc	0.12	2.30	0.02
					CN	0.01	2.02	0.04
					Slope	0.14	2.59	0.01
SEM	0.76	-768.40	1544.90	1562.02	C	0.53	0.60	0.55
					Vfc	0.18	1.58	0.11
					CN	0.01	0.94	0.35
					Slope	0.21	2.07	0.04
					LAMBDA	0.91	53.10	0.00
OLS	0.10	-2124.98	4257.97	4275.12	C	-4.10	-4.30	0.00
					Vfc	0.62	5.80	0.00
					CN	0.04	4.10	0.00
					Slope	0.49	4.50	0.00



**Fig. 6** Hazard maps based on FFPI, C–R and SLM in watershed units

Phongsaly, Oudomxay and Luangprabang. By using the statistic tools in GIS, the average hazard values at district scale were extracted from the three hazard maps, and the scatter diagrams were drawn with the density of historical events in each district (Fig. 7). Besides, although the final result from the binary logistic model is not distinguishable in the target area, its intermediate result, i.e. the probability, is quantified. Thus, this outcome can still be used to plot the scatter diagram with the density of historical events, so as to further understand the effect of the model.

Because there are no statistical data specifically for flash flood disasters in Laos officially at present, the historical events data collected here are the total flood disaster events, and the statistical scope is general and broad; thus, the linear fitting relationships of the scatter diagrams are generally poor. Relatively speaking, the plan based on SLM model presents the best fitting performance, i.e. the higher the hazard degree, the greater the disaster density; thus, it is indeed the best plan among the six models. On the contrary, the fitting performance of the plan based on C–R method is very poor and illogical because it seems that the higher the hazard degree, the lower the disaster density. As mentioned above, the hazard map from C–R method mainly reflects the rainfall distribution, and the rainfall is always been an important trigger factor for a specific flash flood event. However, hazard assessment here is not designed for a specific flash flood forecast but for understanding the overall flash flood development pattern. The rich rainfall is not the restrict factor for flash flood distribution in the Nam Ou River Basin; thus, it is possible for a watershed which has relatively lower rainfall level but suffering greater disasters, because other factors ignored by the C–R method may play the key impacts.

The fitting performance of the binary logistic model ranks second in the four models listed (Fig. 7). Considering the parameters of this model has been calibrated based on the extrapolation method framework, this goodness of fit is unacceptable. The logistic model is essentially a linear regression model; its fitting performance mainly depends on the linear relationship of the samples. However, since the flash flood disasters are very complicated and dangerous, it is difficult to record the first hand data and define the flood types, especially in undeveloped region. Hence, the collected historical disaster data from Yunnan and Laos contain noises, inevitably. Obviously, these noises seriously interfere with the fitting of general linear model (Table 9), and the logistic regression model can hardly achieve an ideal result, either. On the other hand, the logistic model is not good at distinguishing the non-balance samples. For example, the proportion of the first, the second and the

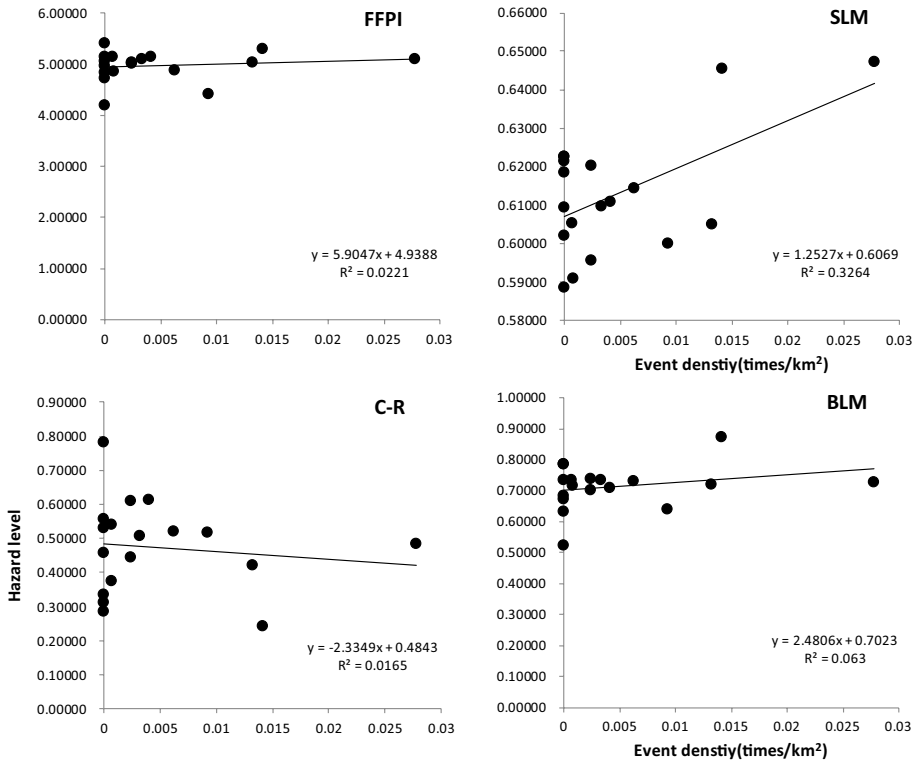


Fig. 7 Scatter diagrams of average hazard level-events density of each plan in Nam Ou River Basin

third categories of samples in the multinomial logistic model of this study is close to 6:3:1 (Table 6), which means the final loss function can still be acceptable even if the samples from the second and the third categories are all identified as the first one.

### 3.4 Rational hazard assessment plan of flash flood for current Laos

In previous studies, the spatial autocorrelation models, especially the spatial lag and error models, were used mostly in the correlation analysis of element layers rather than extrapolation in hydrological geography. Because the results from these models are usually very sensitive to the spatial zoning, while different zoning may change the neighbouring relationships among the statistic units dramatically. However, in this study the fundamental statistic unit is determined by the natural water system, and the watershed scale was also decided based on the analysis of scale effect. Thus, the statistic relationships from SLM can be extrapolated into the other region with the similar environment and watershed scale, theoretically. The optimal fitting performance in the Nam Ou River Basin indicated that the SLM may be the best hazard assessment plan of flash floods for current Laos among the three methods. The successful practice means the spatial autocorrelation models can also be used for extrapolation, effectively.

Nevertheless, considering that the extrapolation method relies heavily on historical recorded, the SLM in this study may lead to misestimating to some extent if it be promoted

to the whole land of Laos and the data used to build the model has not been updated in time, especially in the southern Laos where the situation of precipitation, population, vegetation cover and landform are quite different from that of northern mountainous area. In this case, it is believed that the FFPI can still be useful as an important auxiliary tool due to its comprehensive consideration on the main formed elements of flash floods. Besides, the FFPI is easier to promote for its simpler in the calculation and no validation process. Therefore, we believed that the rational hazard assessment plan of flash flood for current Laos is a mixed method, i.e. SLM proposed here can be implemented as the main assessment tool for the whole country while the FFPI map could also be made for reference in southern Laos.

The performance of logistic model in this study is not as good as that from the reports of Xiong et al. (2019) and Costache (2019a). One important reason for this is the quality of samples here is not good enough for the linear-relationship-based model, which also indicates that the spatial regression model is more adaptable to the rough samples of this study because of the spatial autocorrelation of the flash flood disasters. Optimistically, with the improvement on historical recorded system of flash flood in Laos, the new SLM updated by local input of data would become more and more reliable, and the revised FFPI based on logistic model or other machine learning methods could also be more useful in future study.

## 4 Conclusions

Hazard assessment for the ungauged area is always been a knotty problem on flash flood disaster prevention. As a mountainous, rainy and less developed country in Southeast Asia, Laos suffers a lot from flash floods for poorly preparation on flood monitoring, event record and prevention system construction. In this paper, taking the Nam Ou River Basin in Laos as an example, three basic methods that represent three different solving ideas, including six models totally, were presented and compared, in order to provide a relatively reliable and easy-to-promote plan for current Laos on flash flood hazard assessment.

The FFPI and C–R here are the two methods that do not have to be calibrated, in which the FFPI focuses on representing the influence of underlying surface on the development of flash floods, while the C–R prefers to reflect the runoff yield and rainfall distribution characteristic. At the stage of validation, the fitting performance of the two methods is relatively poor, of which the FFPI method performs better because the development of flash flood in the Nam Ou River Basin seems to be affected greatly by the underlying surface elements.

The model parameters of extrapolation method need to be calibrated by the data from the adjacent region, i.e. Yunnan, China here, firstly. Hence, the fitting performance is better than FFPI and C–R in general. However, the four models from the extrapolation method present different effects at the stages of calibration and validation. Since the spatial autocorrelation of flash flood disasters is considered, the two spatial regression models, i.e. the SLM and SEM, are much better than that of the multiple linear regression model in terms of the statistic indicators for model judgement, and the goodness of fit of the two spatial models can reach more than 0.7, while the corresponding indicator is about 0.1 for the linear regression model.

Although the binary logistic model (BLM) can achieve a comprehensive accuracy rate of recognition at 80.2% in test, the rate on identifying the high and medium hazard

watershed units is only 68.7%, which made the model fail to provide the hazard distribution pattern in the Nam Ou River Basin, and its goodness of fit is also far from satisfactory at validation stage. Too much noises in the samples maybe the main reason for this result.

In the end the SLM, as well as the FFPI was selected as a mixed plan for current Laos' flash flood hazard assessment, in which the SLM performed as the main assessment tool because of the best fitting performance, and the FFPI applied as an auxiliary tool for reference.

Since the adjacent relationships of units could be changed dramatically in different spatial scales, the spatial autocorrelation models, especially the SLM and SEM, have rarely been used for extrapolation in previous studies. This work indicated that the models can also be used for extrapolation effectively, even if the samples are not good enough for general linear and logistic models.

Besides, the scale effect of different potential impact elements on flash flood, i.e. the spatial correlation between each element and the distribution of flash flood events at different spatial statistic units, was also preliminary studied. It is found that the degree of correlation in spatial analysis depends on the short board effect, i.e. only the element that restricts the regional flash flood developmental system can be the key factor. The correlations of most watershed elements, such as the average slope, vegetation cover and watershed area, increase or decrease directly with the rise of watershed scale, and the values of coefficients tend to be stable at large watershed scales.

**Acknowledgements** This research has been sponsored by the National Key R&D Program of China (Grant No. 2018YFC1505201), the program of National Natural Science Foundation of China (Grant No. 41501109, 41902300) and the Opening Fund of State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology (Grant No. SKLGP2019K001). We are grateful to Silvan Ragettli, a researcher from hydrosolutions ltd, Zurich, Switzerland, for suggestions and guidance of this paper.

## References

- Alexander DE (1993) Natural disasters. UCL Press Limited, London, pp 1–632
- Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dodrecht
- Ashley S, Ashley W (2008) Flood fatalities in the United States. *J Appl Meteorol Climatol* 47:806–818
- Bhaskar NR, French MN, Kyiamah GK (2000) Characterization of flash floods in Eastern Kentucky. *J Hydrol Eng* 5(3):327–331
- Borga M et al (2011) Flash flood forecasting, warning and risk management: the HYDRATE project. *Environ Sci Policy* 14:834–844
- Changjiang Institute of Survey, Planning, Design and Research (2018) Comprehensive planning of the Nam Ou River Basin in Laos, report for the Ministry of Natural Resources and Environmental Protection, Lao PDR
- Chinese Flood Control and Drought Relief Headquarters (2013) Bulletin on flood and drought disasters in China in 2013. China Water and Power Press, Beijing
- Chow VT, Maidment DR, Mays LW (1988) Applied hydrology. McGraw-Hill, New York
- Costache R (2019a) Flash-flood potential assessment in the upper and middle sector of Prahova river catchment (Romania). A comparative approach between four hybrid models. *Sci Total Environ* 659:1115–1134
- Costache R (2019b) Flash-flood Potential Index mapping using weights of evidence, decision Trees models and their novel hybrid integration. *Stoch Env Res Risk Assess* 33:1375–1402
- Costache R, Zaharia L (2017) Flash-flood potential assessment and mapping by integrating the weights-of-evidence and frequency ratio statistical methods in GIS environment—case study: Basca Chiojdului River catchment (Romania). *J Earth Syst Sci* 126:59. <https://doi.org/10.1007/s12040-017-0828-9>
- Department of Water Resources (2016) Nam Ou River Basin profile, Vientiane, Lao PDR

- Du J, Xiao X, Cai DM et al (2015) Hazard assessment of mountain torrent debris flow in Wenchuan seismic area. *J Yangtze River Sci Res Inst* 32(3):77–83
- Du J, Ren HY, Zhang PC et al (2016) Comparative study of the hazard assessment of mountain torrent disasters in macro scale. *J Catastrophol* 31(03):66–72
- Ebrahimian M, See LF, Ismail MH et al (2009) Application of natural resources conservation service–curve number method for runoff estimation with GIS in the Kardeh watershed, Iran. *Eur J Sci Res* 34:575–590
- Esri (1995–2010) Arcgis 10 help, USA
- Florax RJ, Rey SJ (1995) The impacts of misspecified spatial interaction in linear regression models. In: Anselin L, Florax RJGM (eds) *New directions in spatial econometrics*. Springer, Heidelberg
- Gan JJ, Sun HY, Huang RQ et al (2012) Study on mechanism of formation and river blocking of Hongchuanou Giant Debris Flow at Yingxiu of Wenchuan County. *J Catastrophol* 27(1):5–16
- Georgakakos KP (1986) On the design of national, real-time warning systems with capability for site-specific, flash flood forecasts. *Bull Am Meteor Soc* 67(10):1233–1239
- Griffith DA (1996) Some guidelines for specifying the geographic weights matrix contained in spatial statistical models. In: Arlinghaus SL (ed) *Practical handbook of spatial statistics*. CRC Press, Boca Raton
- Hadley J (2002) Report for 2002TX58B: real-time distributed runoff estimation using NEXRAD precipitation data. Retrieved 2003. <http://twri.tamu.edu/docs/funding/usgs/2002-03/hadley.pdf>
- Helalia AM (1993) The relation between soil infiltration and effective porosity in different soils. *Agric Water Manag* 24(8):39–47
- Helalia AM, Letey J, Graham RC (1988) Crust formation and clay migration effects on infiltration rate. *Soil Sci Soc Am* 52:251–255
- Lam SN, Quattrochi DA (1992) On the issues of scale, resolution, and fractal analysis in the mapping sciences. *Prof Geogr* 44(1):88–98
- Li J, Chen QD, Kang ZC (1979) A study on the mudflow of the Jiangjia Ravine, Dongchuan District, Yunnan Province. *Acta Geogr Sin* 34(2):156–167
- Lin DJ, Zhen ZC, Zhang XZ et al (2016) Characteristic and influencing factors of soil infiltration of different land use patterns. *J Soil Water Conserv* 24(1):33–36
- Liu CZ, Miao TB, Chen HQ et al (2011) Basic feature and origin of the “8·8” mountain torrent-debris flow disaster happened in Zhouqu County, Gansu, China, Aug. 8, 2010. *Geol Bull China* 30(1):141–150
- Luo ZF, Pu DC, Xie HB et al (2011) Landslide hazard evaluation in debris flow catchment area based on GIS and information method. *China Saf Sci J* 21(11):144–150
- Marchi L et al (2010) Characterisation of selected extreme flash floods in Europe and implications for flood risk management. *J Hydrol* 394(1–2):118–133
- Meng Q, Cieszewski CJ, Madden M et al (2007) K nearest neighbor methods for forest inventory using remote sensing data. *GISci Remote Sens* 44(2):149–165
- National Weather Service. NWS glossary. <http://w1.weather.gov/glossary/index.php>
- Natural Resources Conservation Service (1986) *Urban hydrology for small watersheds*, TR-55. US Department of Agriculture
- Plate EJ (2007) Early warning and flood forecasting for large rivers with the lower Mekong as example. *J Hydroenviron Res* 1(2):80–94
- Saharia M, Kirstetter PE, Vergara H et al (2017) Mapping flash flood severity in the United States. *J Hydro-meteorol* 18(2):397–411
- Shan XZ, Wei YQ, Yan HJ et al (1998) Influence of organic matter content on soil hydrodynamic parameters. *Acta Pedol Sin* 35(1):1–9
- Sharpley AN, Williams JR (1990) EPIC-erosion/productivity impact calculator: 1. Model documentation. US Department of Agriculture Technical Bulletin No. 1768, US Government Printing Office, Washington, DC
- Shi CX (2008) Scaling effects on sediment yield in the upper Yangtze River. *Geogr Res* 4:800–810
- Smith GE (2010) Development of a flash flood potential index using physiographic data sets within a geographic information system. *Dissertations & Theses-Gradworks*
- Soil Conservation Service (1972) *SCS national engineering handbook*, section 4 hydrology. Soil Conservation Service, US Department of Agriculture, Washington, DC
- Stanimir K, Slavoljub D, Tomislav S et al (2017) Torrential flood prevention in the Kolubara river basin. *J Mt Sci* 14(11):2230–2245
- Su PC, Wei FQ, Xie T et al (2012) Causal analysis of debris flow on August 18, 2010 in Gongshan, Yunnan Province and its harm to the development and utilization of mineral resources. *Resour Sci* 34(7):1248–1256
- Tang C, Tie YB (2009) Reconnaissance and analysis on the rainstorm induced debris flow in Weijiagou Valley of Beichuan City after the Wenchuan Earthquake. *Mt Res* 27(5):625–630



- Tobler W (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46(Supplement):234–240
- Wang Z, Lai C, Chen X et al (2015) Flood hazard risk assessment model based on random forest. *J Hydrol* 527:1130–1141
- Water Resource and Environment Administration (2008) National water profile, Vientiane, Lao PDR
- Xiong JN, Li J, Chen WM et al (2019) Spatial–temporal distribution and the influencing factors of mountain flood disaster in southwest China. *Acta Geogr Sin* 74(7):1374–1391
- Xu JX, Yan YX (2005) Scale effects on specific sediment yield in the Yellow River basin and geomorphological explanations. *J Hydrol* 307:219–2321
- Youssef AM, Pradhan B, Hassan AM (2011) Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environ Earth Sci* 62(3):611–623
- Zaharia L, Costache R, Prăvălie R et al (2015) Assessment and mapping of flood potential in the Slănic catchment in Romania. *J Earth Syst Sci* 124(6):1311–1324
- Zelenáková M, Gaňová L, Purcz P et al (2015) Methodology of flood risk assessment from flash floods based on hazard and vulnerability of the river basin. *Nat Hazards* 79(3):2055–2071
- Zeng Z, Tang G, Long D et al (2016) A cascading flash flood guidance system: development and application in Yunnan Province, China. *Nat Hazards* 84(3):2071–2093
- Zhang MD, Li M, Dai CR et al (2016) Simulation study of flood in Yunnan based on the flood area model. *J Catastrophol* 31(1):78–82

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.