Hydrol. Earth Syst. Sci., 21, 1491–1514, 2017
www.hydrol-earth-syst-sci.net/21/1491/2017/
doi:10.5194/hess-21-1491-2017

Hydrology and Earth System Sciences
Open Access

An integrated probabilistic assessment to analyse stochasticity of soil erosion in different restoration vegetation types

Ji Zhou[1,2], Bojie Fu[1,2], Guangyao Gao[1], Yihe Lü[1], and Shuai Wang[1]

[1]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Science, Chinese Academy of Science, Beijing 100085, People's Republic of China
[2]University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

Correspondence to: Bojie Fu (bfu@rcees.ac.cn)

Received: 28 July 2016 – Discussion started: 1 August 2016
Revised: 24 December 2016 – Accepted: 16 February 2017 – Published: 10 March 2017

Abstract. The stochasticity of soil erosion reflects the variability of soil hydrological response to precipitation in a complex environment. Assessing this stochasticity is important for the conservation of soil and water resources; however, the stochasticity of erosion event in restoration vegetation types in water-limited environment has been little investigated. In this study, we constructed an event-driven framework to quantify the stochasticity of runoff and sediment generation in three typical restoration vegetation types (*Armeniaca sibirica* (T1), *Spiraea pubescens* (T2) and *Artemisia copria* (T3)) in closed runoff plots over five rainy seasons in the Loess Plateau of China. The results indicate that, under the same rainfall condition, the average probabilities of runoff and sediment in T1 (3.8 and 1.6 %) and T3 (5.6 and 4.4 %) were lowest and highest, respectively. The binomial and Poisson probabilistic model are two effective ways to simulate the frequency distributions of times of erosion events occurring in all restoration vegetation types. The Bayes model indicated that relatively longer-duration and stronger-intensity rainfall events respectively become the main probabilistic contributors to the stochasticity of an erosion event occurring in T1 and T3. Logistic regression modelling highlighted that the higher-grade rainfall intensity and canopy structure were the two most important factors to respectively improve and restrain the probability of stochastic erosion generation in all restoration vegetation types. The Bayes, binomial, Poisson and logistic regression models constituted an integrated probabilistic assessment to systematically simulate and evaluate soil erosion stochasticity. This should prove to be an innovative and important complement in understanding soil erosion from the stochasticity viewpoint, and also provide an alternative to assess the efficacy of ecological restoration in conserving soil and water resources in a semi-arid environment.

1 Introduction

Soil erosion is a global environmental problem. In recent centuries, the erosion rate worldwide has been accelerating due to climate change and anthropogenic activities, causing soil deterioration and terrestrial ecosystem degradation (Jiao et al., 1999; Marques et al., 2008; Fu et al., 2011; Portenga and Bierman, 2011). The uncertainty and intensity of soil erosion are major features of the erosion phenomenon. Although many studies have concentrated on the intensity of erosion at different spatiotemporal scales (Cantón et al., 2011; Puigdefábregas et al., 1999), the uncertainty of soil erosion generation is a further challenge for researchers working to improve the accuracy of erosion prediction. The stochasticity of environment and spatiotemporal heterogeneity of soil loss is the main influence on the randomness of runoff production and sediment transportation in natural conditions (Kim et al., 2016). But the complex mechanism of erosion generation also increases the uncertainty and variation of erosion processes (Sidorchuk, 2005, 2009). Therefore, how to effectively describe erosion stochasticity and to reasonably assess its impacting factors is necessary and important for understating soil erosion science from the perspective of randomness.

Published by Copernicus Publications on behalf of the European Geosciences Union.

First, combinations of various probabilistic, conceptual and physical models have been reported as different integrated approaches to describe the stochasticity of soil erosion intensity (see Table 1). As one form of erosion intensity, the runoff has been shown as a stochastic process by different mathematic simulation models. Some studies (Moore, 2007; Janzen and McDonnell, 2015) have also simulated the stochasticity, and further quantified the randomness of runoff production and its connectivity dynamics in hillslope and catchment scales by using different probabilistic distribution functions and conceptual models. In these studies, the theory-driven conceptual models simplified the main hydrological behaviours related to runoff production, highlighting the stochastic effects of infiltration and precipitation on runoff processes. Based on the above precondition, the data-driven probabilistic models further simulated the stochastic runoff production by mapping or calibrating the difference between observed and predicted probabilistic values. As a result, the stochastic-conceptual approaches have formed an effective framework to model rainfall–runoff processes (Freeze, 1980), as well as to assess flood forecasting (Yazdi et al., 2014).

The stochasticity of soil erosion rate which is another pattern of erosion intensity has been investigated by probabilistic and physical models in some studies. The theory-driven physical models in these studies (Sidorchuk, 2005) integrated hydrological and mechanical mechanisms of overflow and soil structure with sediment transpiration processes, stressing the stochastic effect of physical principles on erosion rate in different spatial scales (Table 1). Sidorchuk (2005) introduced stochastic variables and parameters into probabilistic models by randomizing the physical properties of overflow and soil structure. This approach developed the understanding of uncertainty of sediment transpiration processes, causing the randomness simulation to better fit the reality of stochastic erosion rate (Sidorchuk, 2009). Additionally, the stochasticity of soil erosion rate also reflected the erosion risk which was assessed by the integration of a theory-driven empirical model with geostatistics (Jiang et al., 2012; Wang et al., 2002; Kim et al., 2016). Erosion risk analysis has generally concentrated on the uncertainty or variability of soil erosion rate at catchment and regional scales, highlighting the impact of the spatiotemporal heterogeneous rainfall and other environment conditions on the stochastic erosion rate. In summary, these probabilistic and physical models constituted a systematical analysis framework closely related to the principle of water balance and basic hydrological assumptions. This effectively described the randomness of soil erosion rate affected by complex hydrological processes (Bhunya et al., 2007). However, few studies have been made to analyse the stochasticity of soil erosion events. In particular, there has been little effort to systematically investigate how the signal of stochastic rainfall is transmitted to erosion events occurring in different restoration vegetation types based on observational data rather than on other model assumptions. Yet such event-based investigation deriving from specific experiment results may be more practically meaningful for understanding the stochastic interaction between rainfall and erosion events.

Secondly, the probabilistic approaches have also been reported as a crucial tool to describe the stochasticity of factors affecting soil erosion rate (Table 1). The randomness of soil water content (Ridolfi et al., 2003), antecedent soil moisture (Castillo et al., 2003), infiltration rate (Wang and Tartakovsky, 2011) and soil erodibility (Wang et al., 2001) in heterogeneous soil types have all been modelled by different probability distribution functions. The stochasticity of soil hydrological characteristics related to erosion rate mainly impacted in various ways the spatiotemporal distribution of erosion rate, especially at regional or larger spatial scales. Meanwhile, as the main driving force of soil erosion generation, the uncertainty of rainfall event to some extent represents the environment stochasticity (Andrés-Doménech et al., 2010). Eagleson in 1978 applied probabilistic-trait models to characterise the stochasticity of rainfall event by using Poisson and Gamma probability distribution functions. The stochastic rainfall distribution in different spatiotemporal scales has also been applied to examine the effect of runoff and sediment yield (Lopes, 1996), to calibrate the runoff–flood hydrological model (Haberlandt and Radtke, 2014), as well as to evaluate sewer overflow in urban catchment (Andrés-Doménech et al., 2010).

The role of spatial distribution of vegetation in controlling the soil erosion rate under different spatiotemporal scales has been well recognized (Wischmeier and Smith, 1978; Puigdefábregas, 2005). How the plants reduce soil erosion rate has also been illuminated and interpreted by various physical and empirical models (Liu, 2001; Mallick et al., 2014; Prasannakumar et al., 2011). In theory, Puigdefábregas (2005) proposed vegetation-driven spatial heterogeneity (VDSH) to explain the relationship between vegetation patterns and erosion fluxes, which improves understanding of the hydrological function of plants in erosion processes. The trigger–transfer–reserve–pulse (TTRP) framework proposed by Ludwig in 2005 systematically explored the responses and feedback between vegetation patches and runoff erosion during ecohydrological processes. Theoretically, the stochastic signals of different rainfall events could also be disturbed by the hydrological function of plants, finally affecting the randomness of runoff and sediment events occurring in various vegetation types. However, little effort has been made to explore the effect of different vegetation types on the stochasticity of soil erosion events. In particular, few approaches have been used to analyse how the properties of rainfall, soil and vegetation impact on the stochastic erosion events through event-based investigation deriving from observational data rather than via theory-based models. Actually, logistic regression modelling (LRM) probably deals with the above problems. LRM evaluates the causal effects of categorical variables on dependent variables, and quantifies the probabilistic contri-

Table 1. Summary of the research on the stochasticity of soil erosion rate and the stochasticity of factors affecting soil erosion rate.

[a]Stochasticity (uncertainty)	[b]Approach or method	[c]Driven types	Main hydrological behaviours	Main influencing factors	Spatiotemporal scale	Reference
Stochasticity of soil erosion rate						
Runoff connectivity	Probabilistic model Conceptual model	1. Data-mapping 2. Theory	Infiltration processes Precipitation	Topography Soil depth	Hillslope scale in the USA	Janzen and McDonnell (2015)
Runoff processes	Probabilistic model Conceptual model	1. Simulation 2. Theory	Infiltration processes Precipitation	Topography		Janzen and McDonnell (2015)
Runoff production	Probabilistic model Conceptual model	1. Theory 2. Simulation	Runoff absorption Water storage Infiltration capacity	Soil moisture Evaporation recharge	Point and basin scale	Moore (2007)
Flood prediction and runoff	Probabilistic model Multivariate analysis	1. Simulation 2. Data calibration	Stochastic rainfall process	Parameters in rainfall–runoff model	Multiple catchment scales in Iran	Yazdi et al. (2014)
Rainfall and runoff processes	Probabilistic model hydrological mechanism	1. Simulation 2. Random event 3. Theory	Soil storage	Given climate regime hydraulic conductivity landform development	Hillslope scale	Freeze (1980)
Erosion rate	Probabilistic model Mechanical mechanism	1. Data calculation 2. Stochastic forcing		Bed shear stress Critical shear stress	Laboratory scales in the Netherlands	Prooijen and Winterwerp (2010)
Erosion rate	Physical model Probabilistic model Conceptual model	1. Theory 2. Simulation	Simulated near-bed flow	Soil structure Oscillating flow		Sidorchuk (2005)
Erosion risk	Empirical model Geostatistics	1. Data mapping	Erosive precipitation	Factors in RUSLE	Annual and regional scales in China	Jiang et al. (2012)
Uncertainty of soil loss	Empirical model Geostatistics Error analysis	1. Simulation 2. Data calibration	Erosive precipitation Runoff and sediment	Spatiotemporal rainfall erosivity distribution	Annual time and catchment scale in the USA	Wang et al., 2002
Uncertainty and variability of erosion rate	Empirical model	1. Hypotheses 2. Data calculation	Total rainfall volume and 30 min rainfall intensity	Stochastic environment conditions Scale effect		Kim et al. (2016)
Stochasticity of factors affecting soil erosion rate						
Soil moisture related to soil erosion	Probabilistic model Physical model	1. Hypotheses, 2. Simulation 3. Theory	Precipitation Evapotranspiration	Temporal patterns of rainfall property	Daily time and hillslope scale in	Ridolfi et al. (2003)
Antecedent soil moisture related to soil erosion	Probabilistic model Physical model	1. Data mapping 2. Theory	Runoff response Infiltration processes		Daily time and multiple catchment scales in Spain	Castillo et al. (2003)
Stochastic rainfall related to flood and runoff	Probabilistic model Conceptual model	1. Data calibration 2. Random event 3. Hypothesis	Stochastic storm Runoff and flood	Parameters in Peak flow models	Hourly–daily time and multiple catchment scales in Germany	Haberlandt and Radtke (2014)
Stochastic rainfall related to runoff and erosion	Physical model Empirical model	1. Simulation 2. Data calibration	Overland/channel flow Erosion transport Precipitation	Spatiotemporal rainfall distribution	Seasonal and annual time catchment scale in the USA	Lopes (1996)
Uncertainty of soil erodibility	Empirical model Geostatistics	1. Simulation 2. Data mapping		Spatiotemporal soil types, depth and parent material	Regional scales in the USA	Wang et al. (2001)
Stochastic rainfall related to runoff	Probabilistic model Conceptual model Physical model	1. Data calibration 2. Theory	Sewer overflows	Rainfall depth and duration, climate conditions	Seasonal and annual time catchment scales in Spain	Andrés-Doménech et al. (2010)

[a] The main contents of different studies focusing on the stochasticity (uncertainty) of soil erosion and its influencing factors. [b] The main statistical methods or different types of mathematical and physical models employed to describe and analyse the stochasticity of soil erosion. [c] The main properties of analysis framework in the different studies and the characteristics of data application on the evaluation of stochasticity of soil erosion.

bution of influencing factors on the randomness of responsive random events in terms of an odds ratio (Hosmer et al., 2013). This can be seen as another probabilistic model to explore the probability attribution of influencing factors. However, little literature is available on LRM being used to explore the probabilistic attribution of stochastic erosion events under complex environmental conditions.

In this study, we have hypothesized that the uncertainty of erosive events was also an important property of the soil erosion phenomenon, and monitored erosion events occurring in three typical restoration vegetation types at runoff plot scale over five consecutive years' rainy seasons. We aim to (1) comprehensively describe the stochasticity of runoff and sediment events in detail by using probability theory, and (2) systematically evaluate the effect of the properties of soil, plant and rainfall on the stochastic erosion events by employing the LRM approach. The probabilistic description attribution approach constitutes an integrated probabilistic assessment based on event-driven probability theory and data-driven experimental observation. The investigation of stochastic soil erosion events by integrated assessment is an innovative and important complement in understanding soil erosion from the stochasticity viewpoint, and could also provide an alternative way to assess the efficacy of ecological restoration for conserving soil and water resources in a semi-arid environment.

2 Method

2.1 Definition and classification of random events

Each observed stochastic weather condition with different durations in the field monitoring period was defined as a random experiment. All possible outcomes of a random experiment constituted a sample space (Ω) defined as a random observational event (O event, for short). Two mutually exclusive random event types – random rainfall event (I event, for short) and random non-rainfall event (C event, for short) – constituted the O event. Precipitation is a necessary condition of runoff generation, and the random runoff production event (R event, for short) is a subset of the I event. Similarly, R event is also a necessary condition of random sediment migration event (S event, for short), which leads to S event being a subset of R event. As a result, O, C, I, R and

S events constituted a random events framework (OCIRS) to reflect the stochasticity of the environment in which soil erosion events occur.

The random event duration in OCIRS is an important property of stochastic weather conditions. In particular, the duration property of I events was closely related to the transmission of stochastic signals of rainfall into the R and S events. According to the rainfall duration patterns in China (Wei et al., 2007), the time interval between two adjacent I events is set to be more than 6 h, forming the criterion for individual rainfall classification. Meanwhile, based on the observation of random events over five consecutive rainy seasons, we summarised the duration property of all I events and further classified them into four mutually exclusive I event types: a random extremely long rainfall event type (Ie event, for short), a random general long-duration rainfall event type (Il event, for short), a random spanning rainfall event type (Is event, for short) whose duration spans two consecutive days, and a random within rainfall event type (Iw event, for short) occurring in a day. The C event can also be divided into two types at the daily scale: the random non-rainfall event type lasting a whole day (Cd event, for short) and the random non-rainfall event type whose duration is less than 24 h (Ch event, for short) which is interrupted by an I event.

Table 2 shows the physical, probabilistic properties and implications of all random event types in OCIRS. The classification process of all random event types is illustrated in Fig. 1a, and a Venn diagram of all random event types in OCIRS is shown in Fig. 1c. Considering the observed longest duration of an Ie event approximating 72 h, in Fig. 1b, we have summarised a series of random event sequences in terms of different combination patterns of I and C events in every 3 consecutive days during the whole monitoring period.

2.2 Probabilistic description of erosion event

2.2.1 Conditional probability of erosion event

In the sample space Ω, for any random event type E in OCIRS, we defined $P(\mathrm{E})$ as the proportion of time that E occurs in terms of relative frequency:

$$P(\mathrm{E}) = \lim_{n\infty} \frac{n(\mathrm{E})}{n} = p_{\mathrm{E}},\ p_{\mathrm{E}} \in [0, 1]. \tag{1}$$

Theoretically, $n(\mathrm{E})$ is the number of times in n outcomes of observed random experiment that the event E occurs. According to the law of total probability (Robert et al., 2013), the probability of R event is defined as

$$P(\mathrm{R}) = P(\mathrm{RI}) = P\left(\mathrm{R}|\cup_{m=1}^{4}\mathrm{I}_m\right) P\left(\cup_{m=1}^{4}\mathrm{I}_m\right) = \sum_{m=1}^{4} P(\mathrm{R}|\mathrm{I}_m)\, P(I_m) = p_{\mathrm{R}}. \tag{2}$$

I_m, $m = 1$, 2, 3 and 4 represent the Ie, Il, Is, and Iw respectively, and $P(\mathrm{R}|\mathrm{I}_m)$ represents the conditional probability that R event occurs given that the mth I event type has occurred. Similarly, the probability of S event is defined as

$$\mathrm{P}(\mathrm{S}) = \mathrm{P}(\mathrm{SI}) = \mathrm{P}\left(\mathrm{S}|\cup_{m=1}^{4}\mathrm{I}_m\right) \mathrm{P}\left(\cup_{m=1}^{4}\mathrm{I}_m\right) = \sum\nolimits_{m=1}^{4} \mathrm{P}(\mathrm{S}|\mathrm{I}_m)\,\mathrm{P}(\mathrm{I}_m) = p_{\mathrm{S}}. \tag{3}$$

Equations (2) and (3) quantify the stochastic soil erosion events by using conditional probability.

2.2.2 Probability distribution functions of erosion event

We define X and Y as two discrete random variables, representing two real-valued functions defined on the sample space (Ω). Let X and Y denote the numbers of times of R and S event occurrence respectively, and assign the sample space Ω to another random variable Z. $X(\mathrm{R}) = x, Y(\mathrm{S}) = yZ(\Omega) = z, y \leq x \leq z$. The x, y, and z are integers. The ranges of X and Y are $\mathrm{R}_X = \{\text{all } x : x = X(\mathrm{R}), \text{all } \mathrm{R} \in \Omega\}$ and $\mathrm{R}_Y = \{\text{all } y : y = Y(\mathrm{S}), \text{allS} \in \Omega\}$. The probability of x_i or y_j numbers of times of R or S events can be quantified by probability mass function (PMF) as follows:

$$\mathrm{PMF}_X(x_i) = \mathrm{P}[\{\mathrm{R}_i : X(\mathrm{R}_i) = x_i, x_i \in \mathrm{R}_X\}] \tag{4}$$

$$\mathrm{PMF}_Y(y_j) = \mathrm{P}\left[\left\{\mathrm{S}_j : Y\left(\mathrm{S}_j\right) = y_j, y_j \in \mathrm{R}_Y\right\}\right] \text{for } i \geq j. \tag{5}$$

The PMF in Eqs. (4) and (5) describe the general expression of probability distribution of all possible numbers of times of R or S events.

The random variables X and Y obey the binomial distribution with n independent Bernoulli experiments (Robert et al., 2013). Therefore, the PMF of X and Y can be defined as follows:

$$\mathrm{PMF}_{X\mathrm{bin}}(x) = \mathrm{P}_{X\mathrm{bin}}(X = x) = \begin{cases} \binom{n}{x} p_{\mathrm{R}}^{x}(1 - p_{\mathrm{R}})^{n-x} & x = 0, 1, 2, \ldots, n \\ 0 & \text{elsewhere} \end{cases} \tag{6}$$

$$\mathrm{PMF}_{Y\mathrm{bin}}(y) = \mathrm{P}_{y\mathrm{bin}}(Y = y) = \begin{cases} \binom{n}{y} p_{\mathrm{S}}^{y}(1 - p_{\mathrm{S}})^{n-y} & y = 0, 1, 2, \ldots, n \\ 0 & \text{elsewhere,} \end{cases} \tag{7}$$

where x and y indicate all possible numbers of times of R and S occurring over n I events. However, when the Bernoulli experiment is performed infinite independent times ($n \to \infty$), the binomial PMF can be transformed into the Poisson PMF (proved in Appendix A), and finally expressed as follows:

$$\mathrm{PMF}_{X\mathrm{poi}}(x) = \mathrm{P}_{X\mathrm{poi}}(X = x) = \begin{cases} \frac{\lambda_{\mathrm{R}}^{x} e^{-\lambda_{\mathrm{R}}}}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{elsewhere} \end{cases} \tag{8}$$

$$\mathrm{PMF}_{Y\mathrm{poi}}(y) = \mathrm{P}_{Y\mathrm{poi}}(Y = y) = \begin{cases} \frac{\lambda_{\mathrm{S}}^{y} e^{-\lambda_{\mathrm{S}}}}{y!} & y = 0, 1, 2, \ldots \\ 0 & \text{elsewhere,} \end{cases} \tag{9}$$

where the parameter $\lambda_{\mathrm{R}} \approx n p_{\mathrm{R}}, \lambda_{\mathrm{S}} \approx n p_{\mathrm{S}}$. Equations (6)–(9) reflect two PMF models to simulate the probability distribution of R or S events.

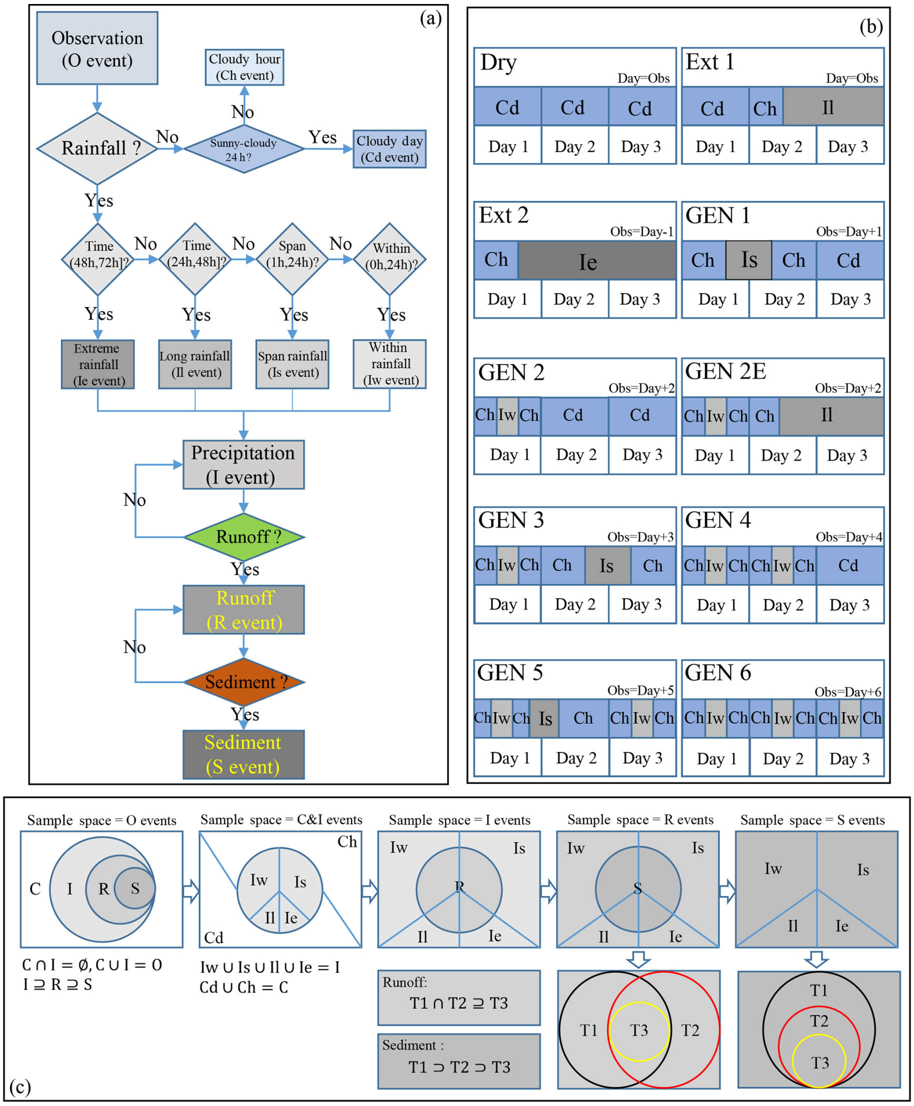

Figure 1. The OCIRS system: **(a)** a flow chart to determine random event types in the OCIRS framework; **(b)** the different combination patterns of rainfall and non-rainfall events in 3 consecutive days to form ten observed random event sequences in five rainy seasons; **(c)** Venn diagram showing the relationship between random event types in the OCIRS framework.

2.3 Probabilistic attribution of erosion events

2.3.1 Bayes model

Based on the Bayes formula theory (Sheldon, 2014), if we want to evaluate how much the probabilistic contributions of kth type of random rainfall event on one stochastic runoff or sediment event which has been generated and observed, the Bayes model can quantify the results as follows:

$$\mathrm{P}(\mathrm{I}_k|\mathrm{R}) = \frac{\mathrm{P}(\mathrm{I}_k\mathrm{R})}{\mathrm{P}(\mathrm{R})} = \frac{\mathrm{P}(\mathrm{R}|\mathrm{I}_k)\,\mathrm{P}(\mathrm{I}_k)}{\sum_{m=1}^{4}\mathrm{P}(\mathrm{R}|\mathrm{I}_m)\,\mathrm{P}(\mathrm{I}_m)} \tag{10}$$

$$\mathrm{P}(\mathrm{I}_k|\mathrm{S}) = \frac{\mathrm{P}(\mathrm{I}_k\mathrm{S})}{\mathrm{P}(\mathrm{S})} = \frac{\mathrm{P}(\mathrm{S}|\mathrm{I}_k)\,\mathrm{P}(\mathrm{I}_k)}{\sum_{m=1}^{4}\mathrm{P}(\mathrm{S}|\mathrm{I}_m)\,\mathrm{P}(\mathrm{I}_m)}. \tag{11}$$

In fact, the Bayes model provides an important explanation that how the a priori stochastic information ($\mathrm{P}(\mathrm{I}_k)$) was modified by the posterior stochastic information ($\mathrm{P}(\mathrm{R})$ or $\mathrm{P}(\mathrm{S})$). The application of Bayes model in Eqs. (10)–(11) reflects the feedback of random erosion events on the stochastic rainfall events. It could also be regarded as one pattern of probabilistic attribution to assess the effect of different random rainfall events on the uncertainty of soil erosion events without considering the diversity of restoration vegetation.

2.3.2 Logistic regression model

Firstly, we constructed an event-driven logistic function, and defined Y_R and Y_S as two dichotomous dependent variables. When we denote Y_R and Y_S as 1 or 0, it means that a R and

Table 2. Definition and explanation of all random events in OCIRS.

Symbol	Physical meaning of random event types	Probabilistic meaning of random event types	Influencing factors and implication
O	observation events with time step ranging from 0 to 72 h, including non-rainfall and rainfall events	random events composing the sample space of OCIRS system. The probability P(O) = 1	indicating the general stochastic weather conditions over rainy seasons
C	non-rainfall events with time step ranging from 0 to 24 h, including sunny or cloudy weather condition at hour or day scales	random events, the probability of C events is the ratio of numbers of C events to O events $C \subset O$, $0 \leq P(C) \leq P(O) = 1$	implying the extent of evaporation or potential evapotranspiration in weather condition
Cd	non-rainfall events with time step being 24 h, including observed sunny or cloudy at day scale	random events composing the subset of C events, $Cd \subseteq C$, $0 \leq P(Cd) \leq P(C)$	implying the duration of evaporation or evapotranspiration at day scale
Ch	non-rainfall events with time step being less than 24 h, including observed sunny or cloudy at hour scales which intercepted by rainfall events within a day	random events composing the subset of C events, the intersection of Ch and Cd is null, $Ch \subseteq C$, $Cd \cup Ch = C$, $Cd \cap Ch = \varnothing$, $0 \leq P(Ch) \leq P(C)$	influenced by the frequency of rainfall events generation, and implying the alternation of sunny and rainy in a day
I	an individual rainfall event with different precipitation, intensity and duration ranging from 0 to 72 h, the time interval between two I events is more than 6 h	random events, the probability of I event is ratio of numbers of I events to O events over observation $I \subset O$, $I \cup C = O$, $I \cap C = \varnothing$, $0 \leq P(I) \leq P(O) = 1$	a driven force of soil erosion, which could be intercepted by vegetation and transformed into throughfall
Ie	an extreme longest individual rainfall event whose average precipitation, intensity and duration were 96.6 mm, 0.022 mm min^{-1}, and 73 h, respectively.	random events composing the subset of I events, $Ie \subseteq I$, $0 \leq P(Ie) \leq P(I)$	rainfall events with low intensity and longest duration, leading to infiltration–excess runoff generation
Il	second longest individual rainfall event type whose average precipitation, intensity and duration were 47.3 mm, 0.027 mm min^{-1}, and 30 h, respectively.	random events composing the subset of I events, the intersection of Il and Ie is null, $Il \subseteq I$, $Il \cap Ie = \varnothing$, $0 \leq P(Il) \leq P(I)$	rainfall events with low intensity and long duration, leading to infiltration–excess runoff generation
Is	rainfall event type spanning 2 days whose average precipitation, intensity and duration were 22.7 mm, 0.042 mm min^{-1}, and 10 h, respectively	random events composing the subset of I events, $Is \subseteq I$, $Is \cap Il \cap Ie = \varnothing$, $0 \leq P(Il) \leq P(I)$	rainfall events with strongest rainfall intensity in middle duration, leading to runoff and sediment generation
Iw	rainfall event type occurring within a day whose average precipitation, intensity and duration were 9.8 mm, 0.045 mm min^{-1}, and 5 h, respectively. It usually occurs several times within 1 day.	random events composing the subset of I events, $Iw \subseteq I$, $Iw \cap Is \cap Il \cap Ie = \varnothing$, $Iw \cup Is \cup Il \cup Ie = I$, $0 \leq P(Iw) \leq P(I)$	rainfall events with least and shortest precipitation and duration, which is difficult to trigger soil erosion
R	runoff event type occurring on vegetation land type; it occurs on rainfall processes, and its duration is negligible	random events responding to I events, $R \subset I$, $R \cap C = \varnothing$, $0 \leq P(R) < P(I)$	influenced by rainfall and vegetation properties
S	sediment event occurring on vegetation land types, it occurs on runoff processes, and its duration is negligible	random events responding to R events, $S \subset R \subset I$, $S \cap C = \varnothing$, $0 \leq P(S) \leq P(R) < P(I)$	driven by R events, and affected by rainfall and vegetation properties

S event has occurred or not occurred. Given that Y_R is a dichotomous dependent variable of R event in the linear probability model in can be expressed as follows:

$$Y_{R_i} = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_n x_{ni} + \xi_i = \alpha + \sum_{n=1}^{n} \beta_n x_{ni} + \xi_i. \quad (12)$$

We then further transform Eq. (12) into the conditional probability of R event which has occurred in the ith observation time as follows:

$$P\left(Y_{R_i} = 1 | \cap_{n=1}^{n} x_{ni}\right) = P\left[\left(\alpha + \sum_{n=1}^{n} \beta_n x_{ni} + \xi_i\right) \geq 0\right] = P\left[\xi_i \leq \left(\alpha + \sum_{n=1}^{n} \beta_n x_{ni}\right)\right] = F\left(\alpha + \sum_{n=1}^{n} \beta_n x_{ni}\right), \quad (13)$$

where $\alpha\beta$ are constants and $F\left(\alpha + \sum_{n=1}^{n} \beta_n x_{ni}\right)$ is the cumulative distribution function of ξ_i when $\xi_i = \alpha + \sum_{n=1}^{n} \beta_n x_{ni}$. Equations (12) and (13) quantify the stochasticity of Y_{R_i} depending on the linear combination of n influencing factors x_n and measurement error ξ under i observation times of stochastic runoff generation.

Secondly, assuming that the probabilistic distribution of ξ_i satisfies logistic distribution and $P\left(Y_{R_i} = 1 | \cap_{n=1}^{n} x_{ni}\right) = p_i$, then the LRM expression of $Y_{R_i} = 1$ is deduced as follows:

$$p_i = F\left(\alpha + \sum_{n=1}^{n} \beta_n x_{ni}\right) = \frac{1}{1 + e^{-\left(\alpha + \sum_{n=1}^{n} \beta_n x_{ni}\right)}} = \frac{e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}}{1 + e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}}. \quad (14)$$

Correspondingly, the LRM of $Y_{R_i} = 0$ can be expressed as

$$P\left(Y_{R_i} = 0 | \cap_{n=1}^{n} x_{ni}\right) = 1 - p_i = \frac{1}{1 + e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}}. \quad (15)$$

The ratio of Eq. (14) to (15) is defined as the odds of the R event:

$$\text{Odds} = \frac{p_i}{1 - p_i} = \frac{\frac{e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}}{1 + e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}}}{\frac{1}{1 + e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}}} = e^{\alpha + \sum_{n=1}^{n} \beta_n x_{ni}}, \quad \text{odds} \in [0, 1]. \quad (16)$$

In this study, odds in Eq. (16) is a probabilistic attribution index to quantify how much the n influencing factors affect the generation of the ith stochastic runoff event. Specifically, when the odds of an influencing factor is greater (less) than 1, it means that the corresponding influencing factor exerts positive (negative) effects on the probability of R generation.

Finally, taking the natural logarithms of both sides of Eq. (16), we transform the odds of stochastic runoff event into the linear Eq. (17) reflecting the standard expression of LRM:

$$\ln\left[\frac{P\left(Y_{R_i} = 1 | \cap_{n=1}^{n} x_{ni}\right)}{P\left(Y_{R_i} = 0 | \cap_{n=1}^{n} x_{ni}\right)}\right] = \ln\left(\frac{p_i}{1 - p_i}\right) = \alpha + \sum_{n=1}^{n} \beta_n x_{ni}. \quad (17)$$

LRM can be regarded as another probabilistic attribution pattern to evaluate the effect of multiple impacting factors – such as soil, vegetation and rainfall – on the randomness of soil erosion events occurring in different restoration vegetation types.

3 Experimental design and data analysis

3.1 Study area

The study was implemented in the Yangjuangou Catchment (36°42′ N, 109°31′ E; 2.02 km^2), which is located in the typical hilly-gully region of the Loess Plateau in China (Fig. 2a). A semi-arid climate in this area is mainly affected by the North China monsoon. Annual average precipitation reaches approximately 533 mm, and the rainy season here spans from June to September (Liu et al., 2012). The Calcaric Cambisol soil type (FAO-UNESCO, 1974) with weak structure and higher erodibility in the Loess Plateau is vulnerable to water erosion. For these reasons, soil and water loss was one of the most serious environmental problems to degrade the ecosystem in the Loess Plateau before the 1980s (Miao et al., 2010; Wang et al., 2015). After that, as a crucial soil and water resource protection project, the Grain-for-Green Project was widely implemented in the Loess Plateau. A large number of steeply sloped croplands were abandoned, restored or naturally recovered by local shrub and herbaceous plants (Cao et al., 2009; Jiao et al., 1999). In the Yangjuangou Catchment, the main restoration vegetation distributed on hillslopes includes *Robinia pseudoacacia* Linn., *Lespedeza davurica*, *Aspicilia fruticosa*, *Armeniaca sibirica*, *Spiraea pubescens* and *Artemisia copria*. All the restoration vegetation was planted over 20 years ago.

3.2 Design and monitoring

In the Yangjuangou Catchment, systematic long-term field experiments have been conducted, including the monitoring of soil erosion (Liu et al., 2012; Zhou et al., 2016), observation of soil moisture dynamic (Wang et al., 2013; Zhou et al., 2015) and assessment of soil controlling service in this typical water-restricted environment (Fu et al., 2011).

In this study, we first monitored the soil erosion events occurring in three typical restoration vegetation types (*Armeniaca sibirica* (T1), *Spiraea pubescens* (T2) and *Artemisia copria* (T3)) from the rainy seasons of 2008–2012 (Fig. 2b). Each restoration vegetation type was designated by three 3 m by 10 m closed-runoff plots located on southwest-facing hillslopes with 26.8 % aspect. The boundaries of each runoff plot were perpendicularly fenced with impervious polyvinyl chloride (PVC) sheet of 50 cm depth. Collection troughs and storage buckets were installed at the bottom boundary to collect the runoff and sediment (Zhou et al., 2016). Under natural precipitation condition, we recorded the number of times that stochastic runoff and sediment events occurred in each runoff plot over the five rainy seasons. Also, we collected runoff and sediment, separated them and, after settling for 24 h, the samples were dried at 105° over 8 h and weighed.

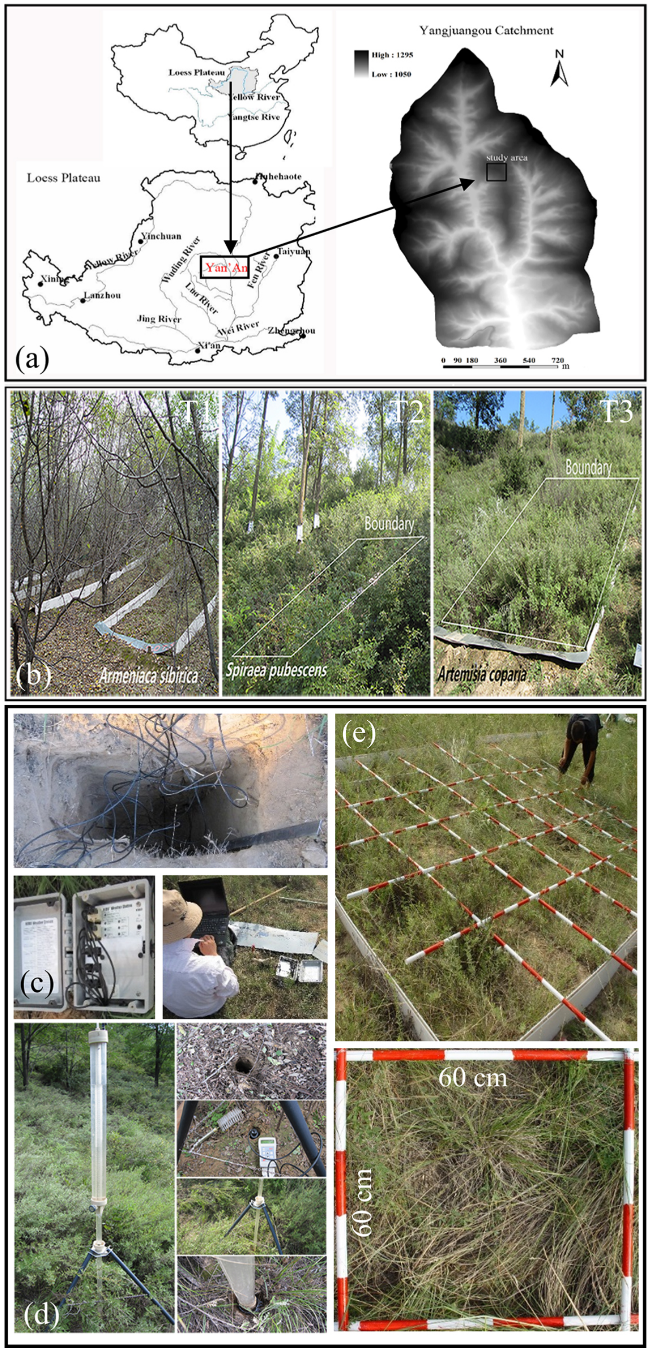

Figure 2. Study area and experimental design: **(a)** location of the Yangjuangou Catchment; **(b)** three restoration vegetation types including *Armeniaca sibirica* (T1), *Spiraea pubescens* (T2) and *Artemisia copria* (T3); **(c)** the dynamic measurement of soil moisture and data collection to provide the information about average antecedent soil moisture; **(d)** the measurement of field-saturated hydraulic conductivity to determine the average infiltration capability; **(e)** investigation of the morphological properties of restoration vegetation by setting quadrats.

Secondly, we systematically monitored the hydrological properties of soil in different restoration vegetation types. In the rainy season of 2010, we began to measure the dynamics of soil moisture in the study region (Wang et al., 2013). The real-time dynamic data of soil water content at intervals of 10 min were recorded by S-SMC-M005 soil moisture probes (Decagon Devices Inc., Pullman, WA, USA), and were collected by HOBO weather station logger (Fig. 2c). These data provided the information about average antecedent soil moisture (ASM) before every rainfall event occurring in the two rainy seasons between 2010 and 2012. We further measured the field-saturated hydraulic conductivity (SHC) in all vegetation types by a model 2800 K1 Guelph permeameter (Soilmoisture Equipment Corp., Santa Barbara, CA, USA) to determine the average infiltration capability of the soil matrix (Fig. 2d).

Thirdly, we investigated the morphological properties of different vegetation types in each runoff-plot for 2–3 times over different periods of rainy season. We measured the average crown width, height and the thickness of litter layer in three restoration vegetation types by setting 60×60 cm quadrats in each runoff plot (Bonham, 1989) (Fig. 2e).

Finally, two tipping bucket rain gauges were installed outside the runoff plots to automatically record the rainfall processes over the five rainy seasons with an accuracy of 0.2 mm precipitation. Table 3 summarises the properties of the four types of random rainfall event, and the basic characteristic of soil and vegetation is shown in Table 4.

3.3 Statistics

We employed nonparametric statistical tests – one-way ANOVA and post hoc LSD – to determine the significant difference of soil, vegetation and erosive properties in the three restoration vegetation types. The maximum likelihood estimator (MLE) and uniformly minimum variance unbiased estimator (UMVUE) (Robert et al., 2013) were explored to compare the suitability of the binomial PMF and Poisson PMF for predicting the uncertainty of runoff and sediment generation over the long term.

4 Results

4.1 Environmental stochasticity in different rainy seasons

The probabilistic distribution of random rainfall events (I events) and random non-rainfall events (C events) forms the environmental stochasticity which is the background of stochastic soil erosion generation. Within the OCIRS framework, the stochastic environment at monthly and seasonal scales over five rainy seasons is described in Fig. 3. For the rainy seasons of 2008 to 2012, the probability of I event generation first increased with later monitoring period and then decreased in the last two rainy seasons. In the rainy season of 2008, the average probability of I event was lower than the other four rainy seasons, being less than 15 %. However, the I event type was most complex in 2008. The random extremely long rainfall event (Ie event) only appeared in this rainy season, with the probability reaching 2.5 %. On the other hand, the average probability of I event was the highest in the rainy season of 2010, being larger than 18 %. But there were only two types of I events (Iw and Is events) in this rainy season. Over the five rainy seasons, the average probability of Iw (12.3 %) and Ie (0.8 %) event occurrence was the highest and lowest, respectively. The average probability of Is (1.7 %) and Il (1.3 %) events ranged between Iw and Ie. The probability of Cd event was higher than Ch in each month of rainy season, with average probability being 54.4 and 29.4 %, respectively. As seen in Table 3, the difference in average precipitation and duration of the four types of I events was significant. But the average rainfall intensity of Iw and Is events was nearly twice that of Il and Ie events.

Table 3. Main characteristics of the four types of random rainfall event over five rainy seasons.

Rainy season	Rainfall event types	Average precipitation (mm)	Average intensity ($mm\,min^{-1}$)	Average duration (h)
2008	Iw	16.7	0.122	2.3
	Is	19.2	0.066	4.8
	Il	53.2	0.032	27.7
	Ie	96.6	0.022	73.2
2009	Iw	9.0	0.027	5.6
	Is	35.4	0.059	10.0
	Il	47.9	0.032	24.9
	Ie	×	×	×
2010	Iw	9.0	0.018	8.3
	Is	7.6	0.012	10.6
	Il	×	×	×
	Ie	×	×	×
2011	Iw	3.3	0.031	1.8
	Is	21.5	0.040	9.0
	Il	42.5	0.020	35.4
	Ie	×	×	×
2012	Iw	10.8	0.028	6.4
	Is	30.0	0.031	16.1
	Il	45.5	0.023	33.0
	Ie	×	×	×
Average	Iw	9.8	0.045	4.9
	Is	22.7	0.042	10.1
	Il	47.3	0.027	30.3
	Ie	96.6	0.022	73.2

Table 4. Basic properties of soil, vegetation and erosion in different restoration vegetation types.

Basic properties of different vegetation type	[h] N	Restoration vegetation types		
		Armeniaca sibirica Type 1 (T1)	*Spiraea pubescens* Type 2 (T2)	*Artemisia copria* Type 3 (T3)
Topography property				
Slope aspect	9	Southwest	Southwest	Southwest
Slope gradation (%)	9	≈ 26.8	≈ 26.8	≈ 26.8
Slope size for each (m)	9	3 × 10	3 × 10	3 × 10
Soil property				
[a]DBD ($g\,cm^3$)	30	1.28 ± 0.08	1.16 ± 0.12	1.23 ± 0.10
Clay (%)	30	11.07 ± 2.43	11.98 ± 3.05	9.54 ± 1.48
Silt (%)	30	26.11 ± 1.50	25.24 ± 3.84	26.72 ± 2.87
Sand (%)	30	62.82 ± 0.94	62.78 ± 4.51	63.74 ± 3.24
[b]Texture type		Sandy loam	Sandy loam	Sandy loam
[c]SHC ($cm\,min^{-1}$)	20	0.46 ± 0.82(a)	2.22 ± 0.66(b)	0.50 ± 0.60(a)
[d]SOM (%)	30	1.28 ± 0.63(a)	0.98 ± 0.15(b)	0.90 ± 0.09(b)
Vegetation property				
Restoration years	9	20	20	20
Crown diameters (cm)	27	211.6 ± 15.4(c)	80.5 ± 4.5(b)	64.1 ± 6.3(a)
Litter layer (cm)	30	1.2 ± 0.3(a)	3.4 ± 1.8(b)	1.8 ± 0.5(a)
Height (cm)	27	256.3 ± 11.1(c)	128.3 ± 8.3(b)	61.8 ± 1.1(a)
LAI	27	×	2.31	1.78
[e]Ave. coverage (%)	27	85	90	90
Rainfall/erosion property				
Times of rainfall events			130	
Times of runoff events		30/30/30	45/45/45	45/45/45
Times of sediment events		13/13/13	19/19/19	32/32/32
[f]Ave. runoff depth (cm)		0.012(a)	0.014(a)	0.083(b)
[g]Ave. sediment amount (g)		5.8(a)	6.8(a)	25.7(b)

[a] Dry bulk density. [b] Texture type is determined by textural triangle method based on USDA. [c] Field saturated hydraulic conductivity, and all the values with same letter in each row indicates non-significant difference at $\alpha = 0.05$ which is the same as follow rows. [d] Soil organic matter. [e] Average coverage of three restoration vegetation types over five rainy seasons. [f] Average runoff depth in restoration types over rainy seasons. [g] Average sediment yield in restoration types over rainy seasons. [h] Sample number.

4.2 Stochasticity of soil erosion events

4.2.1 Probability of erosion events in vegetation types

The stochasticity of erosion events was quantified by the probability of runoff and sediment generation in three restoration vegetation types (T1, T2 and T3) at monthly and rainy season scales (Fig. 4). Over the five rainy seasons, the probability of soil erosion occurring in all vegetation types generally decreased with later monitoring period, and then increased in 2012. At the early period of erosion monitoring (2008), the randomness of erosion events is similar, and the probability of R and S events ranged from 6 to 13 % and from 3 to 13 % respectively. After that, from the rainy seasons of 2009 to 2011, the highest probabilities of erosion events in each vegetation type all concentrated in the July and August of each season. Regarding runoff production, the average probability of R event in T1 (3.78 %) was less than that for T2 (5.60 %) and T3 (5.58 %) under the same precipitation condition. With respect to sediment yield, the average probability of S event in T1 (1.65 %) was also the lowest in all restoration vegetation types. In particular, in the last two rainy seasons, there was no S event occurring in T1, but the average probability of S event in T2 and T3 reached 1.83 and 3.36 % respectively in the corresponding rainy seasons. Consequently, affected by the same stochastic signal of rainfall events, T1 and T3 have the lowest and highest probability of erosion event generation over the five rainy seasons respectively.

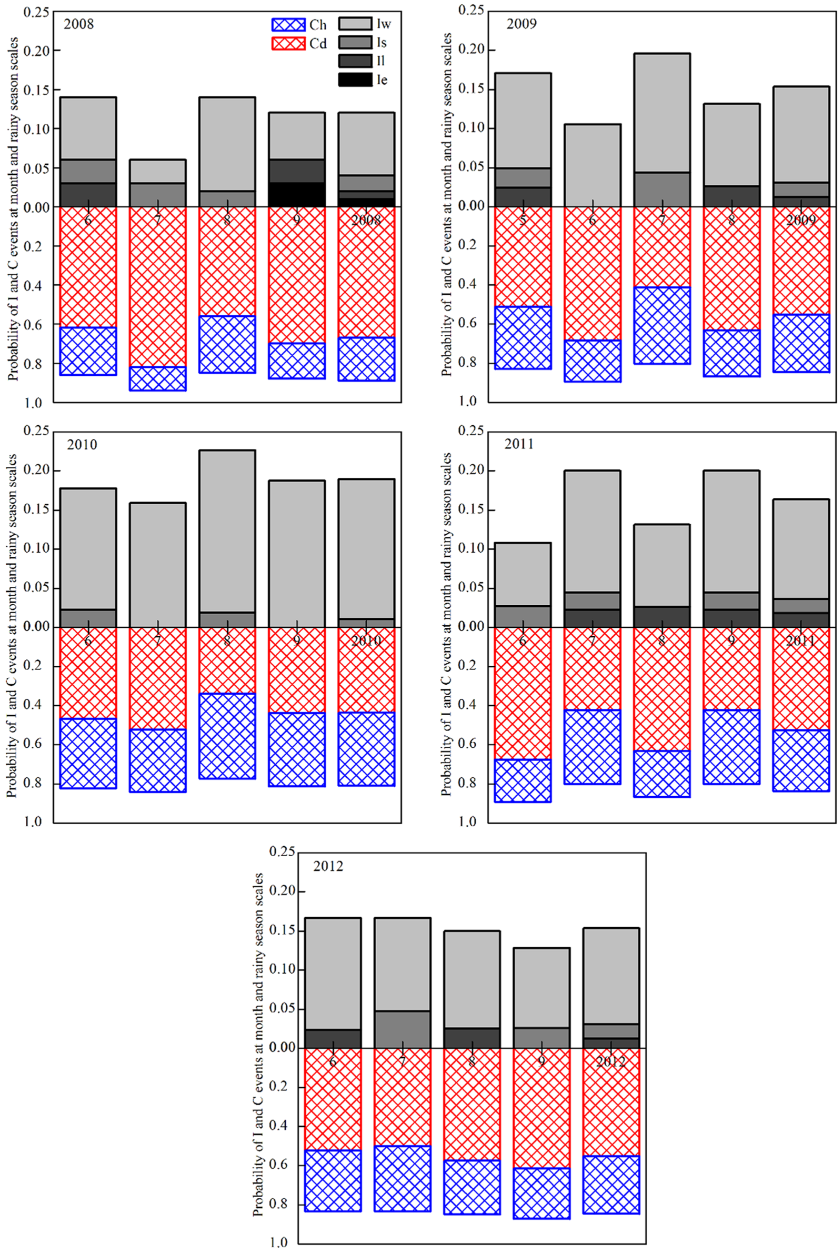

Figure 3. The probability distribution of different random rainfall event types (Iw, Is, Il and Ie) and random non-rainfall event types (Ch and Cd) at monthly and seasonal scales from the rainy seasons of 2008 to 2012.

4.2.2 Probabilistic distribution of erosion events in vegetation types

More detailed stochastic information of erosion events in different vegetation types was simulated by binomial and Poisson PMFs at monthly scale. We also compared the frequency distributions of different numbers of observed erosion events with the corresponding simulated results by the two PMFs in Fig. 5. Firstly, as to the detailed stochastic information of R events, the two PMFs generally provided a better simulation to the observations in T1 than in T2 and T3. When comparing the simulated and observed values, the binomial PMF supplied better simulation to the observed numbers of time of R events with larger frequency (such as 2–4 times) than did the Poisson PMF. However, the Poisson PMF simulated the observed numbers of time of R events with lower frequency (such as 6–8 times) better than the binomial PMF. Secondly, in relation to the detailed stochastic information of S event, the two PMFs provided better simulation to the observations in T3 than in T1 and T2. In particular, when the number of times of S event generation reaches two in T1 and T2, the corresponding simulated probability values were all nearly

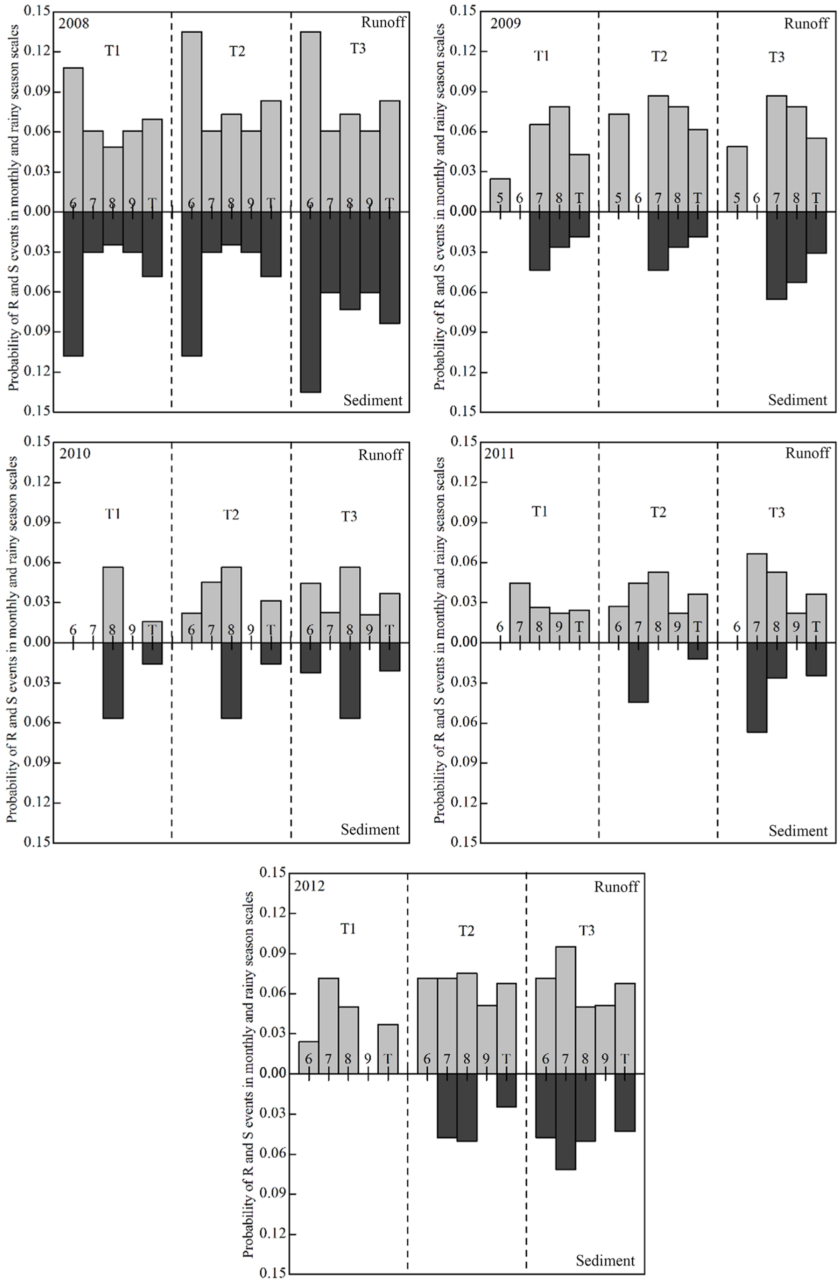

Figure 4. The probability distribution of random runoff and sediment events occurring in three restoration vegetation types at monthly and seasonal scales from the rainy seasons of 2008 to 2012; the Arabic numbers and letter "T" on the abscissa indicate the month and season respectively (also in the following figures).

two times larger than the observed frequencies, reflecting the greatest simulation error of the two PMFs. Moreover, with the restoration vegetation types changing from T1 to T3, both the simulated and observed numbers of time of R and S events with largest probability or frequency increased in consistently. In summary, comparing the observed frequency of numbers of erosion events, both PMFs showed good simulation ability to detail the stochasticity of runoff and sediment events at the monthly scale.

4.3 Stochastic attribution of soil erosion events

4.3.1 Effect evaluation of stochastic erosion events by Bayes model

The Bayes model was applied to analyse the effect of random rainfall events (including Iw, Is, Il and Ie) on stochastic erosion events in different restoration vegetation types. Specifically, the Bayes model evaluated the different probabilistic

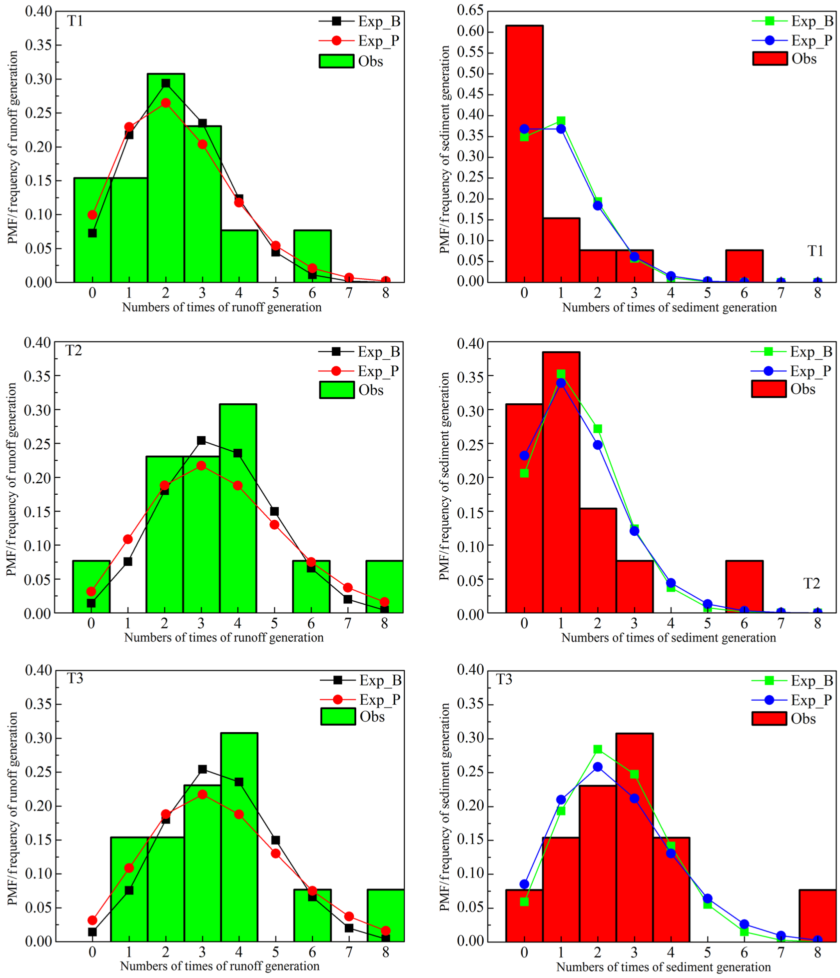

Figure 5. The comparison between simulation of stochasticity of runoff and sediment events by binomial and Poisson PMFs and the observed frequencies of numbers of times of soil erosion events in three restoration vegetation types; Exp_B and Exp_P indicate the simulated values in binomial and Poisson PMF respectively; the histogram shows the observed values.

contributions of four types of I events on one observed erosion event stochastically generated in specific vegetation type at monthly and rainy seasonal scales (Fig. 6). In the rainy season of 2008, the types of I events driving one stochastic erosion event was more complex than in the other rainy seasons. In contrast, only one stochastic soil erosion occurrence in three vegetation types was attributed to Iw and Is events in the rainy season of 2010. In the other three rainy seasons, when one R or S event stochastically generated in T1, the main contributing I event types concentrated on Is and Il events, which have relatively higher precipitation and longer duration, respectively. On the other hand, if one R or S event occurred in T2 or T3 randomly, the main contributing I event type was the Iw event which, however, had no contribution to S event occurring in T1.

In general, over five rainy seasons, the composition of I event driving one R event was more complex than that driving one S event. The relatively longer-duration rainfall events (Il and Ie) became the main probabilistic contributors of one stochastic erosion event occurring in T1, and the relatively stronger-intensity rainfall events (Iw and Is) mainly caused one random erosion event occurring in T2 and T3.

4.3.2 Effect evaluation of stochastic erosion events by LRM

According to the results of significant difference analysis in Table 4, we defined the properties of soil and plant as ordinal variables, and classified them into four grades (Table 5). Meanwhile, based on previous studies (Liu et al., 2012; Wei et al., 2007) and rainfall properties in this study area, we further subdivided all precipitation and rainfall intensity into four grades with different scores.

First, the intensity of positive and negative effects of a single influencing factor on the probability of runoff and sediment generation in all restoration vegetation types was quantified in terms of odds ratio of erosion events by LRM (Ta-

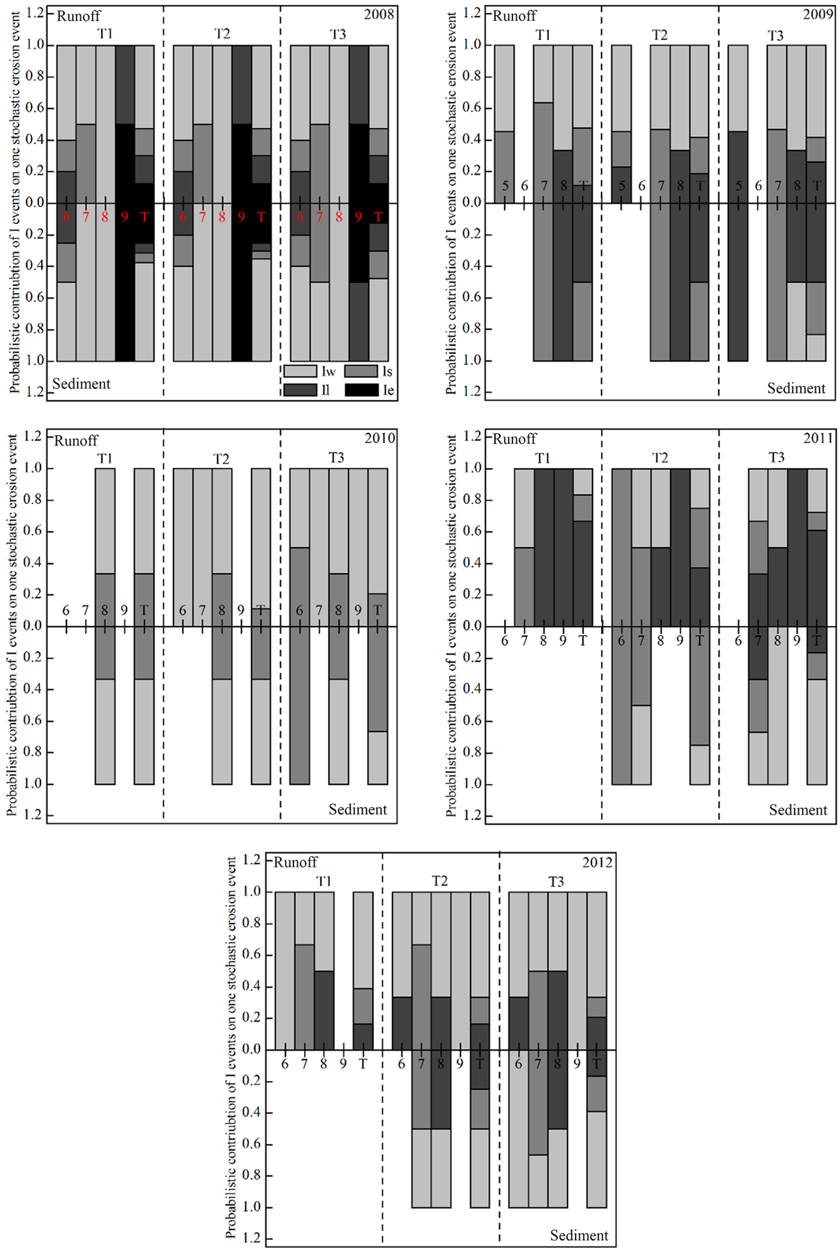

Figure 6. The distribution of probabilistic contribution of four random rainfall event types on any one runoff or sediment event stochastically occurring in three restoration vegetation types at monthly and seasonal scales from rainy season of 2008 to 2012.

ble 6). In the LRM, the highest and lowest odd ratios appeared in rainfall intensity ordinal variable (INT) and average crown width ordinal variable (CRO). An increasing INT and CRO (from middle to extreme grade) significantly increased and decreased the odds ratio of erosion events, respectively. This means that INT and CRO have two of the most important roles in improving and restraining the probability of stochastic erosion generation in all restoration vegetation types. Additionally, the increasing of antecedent soil moisture ordinal variable (ASM) and the SHC ordinal variable (from middle to high grade) in the LRM also significantly increased and decreased the odds ratio of R and S events, respectively. However, the average thickness of litter layers (TLL) ordinal variable did not have significant effect on the odds ratio of erosion events. Tables S1 and S2 in the Supplement systematically describe the processes of LRM to evaluate the effect of single factors on the odds ratio of erosion event.

Secondly, we applied LRM to evaluate the interactive effects of multiple influencing factors on the odds ratio of R and S events in all restoration vegetation types (Table 7). Regarding the interactive effect of two soil hydrological properties, the interaction between low grade of SHC and increasing grade of ASM significantly raised the odds ratio of erosion events – the odds ratio of R and S events affected by the interactive effects of low-grade SHC and extreme-grade ASM were respectively 7.02 and 1.82 times larger than the interactive effects of low-grade SHC and low-grade ASM. Similarly, regarding the effect of two vegetation properties, the interactive effect of low-grade CRO and increasing-grade TLL reduces the odds ratio of erosion events – the odds ratio of R and S events influenced by the interaction between low-grade CRO and high-grade TLL were respectively only 0.12 and 0.33 times larger than the interactive effects of low-grade CRO and low-grade TLL. Additionally, with respect to the interaction between soil and plant properties, the interactive effect of low-grade CRO and increasing-grade ASM properties also significantly raised the odds ratio of erosion events. The processes of LRM used to evaluate the interactive effect of multiple factors on odds ratio of erosion event are detailed in Supplement Tables S3–S5.

5 Discussion

5.1 The integrated probabilistic assessment of erosion stochasticity

The probabilistic attribution and description of stochastic erosion events constituted the framework of integrated probabilistic assessment (IPA).

First, as one pattern of probabilistic attribution in the IPA, the Bayes model supplies a supplementary view and algorithm about how to evaluate the feedback of a result which had stochastically occurred on all possible reasons (Wei and Zhang, 2013). Under the conditions of insufficient information about an occurred result, the Bayes model can determine which reasons have relatively greater probability to trigger the occurrence of the result through some prior information. Specific to this study, the Bayes model was used to evaluate the probabilistic contribution of four types of I events on one stochastic R ($P(I_k|R)$) and S ($P(I_k|S)$) event generated in each restoration vegetation type. Although there was no further specific information about a stochastic soil erosion event, the prior information ($P(R|I_m)P(S|I_m)P(I_m)$) can provide assistance for us to assess the feedback of the stochasticity of soil erosion on different random rainfall events by the Bayes model. Meanwhile, ($P(I_k|R)$) and ($P(I_k|S)$) also reflect the different probability threshold values of four rainfall event types triggering soil erosion. The Bayes model integrated with total probability theory to systematically quantify the interactive relationship between the stochasticity of precipitation and soil erosion, forming a relatively simple and practical risk assessment of soil erosion event occurring in complex restoration vegetation conditions.

Secondly, as a pattern of probabilistic description in the IPA, the binomial and Poisson PMFs are two crucial prob-

Table 5. The definition and classification of properties of rainfall soil and plant ordinal variables.

Ordinal variable	Physical meaning of classified influencing factors	Standard of influencing factor classification			
		Low (L)	Middle (M)	High (H)	Extreme (E)
PREC	classified precipitation variable of a single random rainfall event	0–15 mm	15–30 mm	30–60 mm	>60 mm
INT	classified intensity variable of a single random rainfall event	0–0.025 mm min^{-1}	0.025–0.05 mm min^{-1}	0.05–0.1 mm min^{-1}	>0.1 mm min^{-1}
ASM	classified variable of the antecedent soil moisture	0–5 %	5–10 %	10–20 %	>20 %
SHC	classified variable of the filed saturated hydraulic conductivity	0–1 cm min^{-1}	×	>1 cm min^{-1}	×
CRO	classified variable of the average crown width in vegetation type	0–60 cm	60–80 cm	>80 cm	×
TLL	classified variable of the average thickness of litter layers	0–2 cm	×	>2 cm	×
Y_R	dichotomous dependent variable to indicate whether a random runoff event has occurred or not	If $Y_R = 1$, it means that a random runoff event has occurred; If $Y_R = 0$, it means that a random runoff event has not occurred			
Y_S	dichotomous dependent variable to indicate whether a random sediment event has occurred or not	If $Y_S = 1$, it means that a random sediment event has occurred; If $Y_S = 0$, it means that a random sediment event has not occurred			

Table 6. Logistic regression model to analyse the single effect of rainfall, plant and soil ordinal variable on the erosion events presence/absence in all restoration vegetation types.

Grade levels	PREC (low)	INT (low)	ASM (low)	SHC (low)	CRO (low)	TLL (low)
			Odds ratio of all random runoff events			
Extreme	[a]$\times^{NS}$	[b]90.91***	[c]2.19*	Null	Null	Null
High	$\times^{NS}$	32.26***	2.01*	[d]0.85*	[e]$7.53\times10^{-3**}$	[f]$\times^{NS}$
Middle	$\times^{NS}$	2.09*	1.59*	Null	$7.17\times10^{-2**}$	Null
			Odds ratio of all random sediment events			
Extreme	142.85***	166.67***	15.40*	Null	Null	Null
High	16.95**	125.00***	13.79**	0.78*	$6.27\times10^{-3**}$	$\times^{NS}$
Middle	6.09**	34.48***	6.36*	Null	$2.55\times10^{-2**}$	Null

[a] Taking the low grade of PREC ordinal variable as reference, the odds ratio of all random runoff event in extreme grade of PREC is not significantly larger than that of low grade of PREC. [b] Taking the low grade of INT ordinal variable as reference, the odds ratio of all random runoff events in extreme grade of INT is 90.91 times significantly larger than that of low grade of INT, under the controlled PREC condition with $P\le0.001$. [c] Taking the low grade of ASM ordinal variable as reference, the odds ratio of all random runoff events in extreme grade of ASM is 2.19 times significantly larger than that of low grade of ASM, under the controlled PREC and INT condition with $P\le0.1$. [d] Taking the low grade of SHC ordinal variable as reference, the odds ratio of all random runoff events in high grade of SHC is 0.85 times significantly larger than that of low grade of SHC, under the controlled PREC, INT and ASM condition with $P\le0.1$. [e] Taking the low grade of CRO ordinal variable as reference, the odds ratio of all random runoff events in high grade of CRO is 7.53×10^{-3} larger than that of low grade of CRO, under the controlled PREC, INT, ASM and SHC condition with $P\le0.01$. [f] Taking the low grade of TLL ordinal variable as reference, the odds ratio of all random runoff events in high grade of TLL is not significantly larger than that of low grade of TLL, under the controlled PREC, INT, ASM, SHC and CRO condition. (the Wald test statistic is applied to test the significance of odds ratio: $^{***}P\le0.001$, $^{**}P\le0.01$, $^{*}P\le0.1$; NS: not significant; $\times^{NS}$: the nonsignificant value cannot be estimated).

Table 7. Logistic regression model to analyse the interactive effect of rainfall, plant and soil ordinal variables on the erosion events presence/absence in all restoration vegetation types.

Grade levels	Reference of grade levels	Soil_ASM				Plant_TLL	
		ASM (low)	ASM (middle)	ASM (high)	ASM (extreme)	TLL (low)	TLL (high)
				Odds ratio of all random runoff events			
Soil_SHC	SHC (low)	Ref.	[a]2.23^{NS}	3.19^{NS}	7.02*	Null	Null
Plant_TLL	TLL (Low)	Ref.	2.23^{NS}	3.19^{NS}	7.02*	Null	Null
Plant_CRO	CRO (low)	Ref.	[b]64.34*	70.77*	486.43**	Ref.	[c]0.12***
	CRO (middle)	Ref.	$\times^{NS}$	2.32^{NS}	22.49*	Null	Null
	CRO (high)	Ref	Null	Null	Null	Null	Null
				Odds ratio of all sediment runoff events			
Soil_SHC	SHC (low)	Ref.	$\times^{NS}$	1.22^{NS}	1.82^{NS}	Null	Null
Plant_TLL	TLL (low)	Ref.	$\times^{NS}$	1.22^{NS}	1.82^{NS}	Null	Null
Plant_CRO	CRO (low)	Ref.	$\times^{NS}$	$\times^{NS}$	$\times^{NS}$	Ref.	0.33**
	CRO (middle)	Ref.	$\times^{NS}$	$\times^{NS}$	$\times^{NS}$	Null	Null
	CRO (high)	Ref	Null	Null	Null	Null	Null

[a] Taking the interactive effect of low grade of SHC and low grade of ASM as reference, the odds ratio of all random runoff events affected by the interactive effect of low grade of SHC and middle grade of ASM is 2.23 times larger than the interactive effect of low-grade SHC and low-grade ASM under controlled rainfall conditions. [b] Taking the interactive effect of low grade of CRO and low grade of ASM as reference, the odds ratio of all random runoff events affected by the interactive effect of low-grade CRO and middle-grade ASM is 64.34 times significantly larger than that interactive effect of low grade of CRO and low grade of ASM under controlled rainfall conditions, with $P\le0.1$. [c] Taking the interactive effect of low grade of CRO and low grade of TLL as reference, the odds ratio of all random runoff events affected by the interactive effect of low grade of CRO and high grade of TLL is 0.12 times significantly larger than that interactive effect of low grade of CRO and low grade of TLL, with $P\le0.001$. (the Wald test statistic is applied to test the significance of odds ratio: $^{***}P\le0.001$, $^{**}P\le0.01$, $^{*}P\le0.1$; NS: not significant, $\times^{NS}$: the nonsignificant value cannot be estimated).

abilistic functions to characterise many random hydrological phenomena and to model their ecohydrological effects in natural condition (Eagleson, 1978; Rodriguez-Iturbe et al., 1999, 2001). In this study, the two PMFs were found to give good simulations of the frequency of times of soil erosion events in three restoration vegetation types. However, it is necessary and meaningful for the reliability and accuracy of the IPA to assume whether the two PMFs can both stably and reasonably simulate the erosion stochasticity at closed-runoff plot over a longer monitoring period. Therefore, based on the above assumption, two important point estimation methods – the maximum likelihood estimator (MLE) and uniformly minimum variance unbiased estimator (UMVUE) (Robert et al., 2013) – were applied to evaluate the stability of erosion stochasticity estimation by means of analysing the unbiasedness and consistency of p_{R}, p_{S}, λ_{R} and λ_{S}. Taking parameter analysis of random runoff event for example, we defined X_i as the number of times R event occurred in some specific restoration vegetation type in the ith rainy season ($i = 1$, 2, 3, 4 and 5). The five independent and identical (iid) random variables satisfy the same and mutually independent binomial or Poisson PMFs as follows:

$$X_1, X_2, \ldots, X_5 \overset{\text{iid}}{\longrightarrow} \text{binomial}\,(p_{\mathrm{R}}) \text{ or } X_1, X_2, \ldots, X_5 \overset{\text{iid}}{\longrightarrow} \text{Poisson}\,(\lambda_{\mathrm{R}})\,. \tag{18}$$

Considering longer monitoring periods, we supposed that the numbers of corresponding I events (n) and rainy seasons (i) would approach infinity (n, $i \to \infty$), and Eq. (18) can be transformed as follows:

$$X_1, X_2, \ldots, X_i \overset{\text{iid}}{\longrightarrow} \text{binomial}(p) \text{ or } X_1, X_2, \ldots, X_i \overset{\text{iid}}{\longrightarrow} \text{Poisson}\,(\lambda)\,. \tag{19}$$

We take MLE and UMVUE methods to search for the best reasonable population estimators $\hat{p}$ and $\hat{\lambda}$ to approximate the unknown p and λ in Eq. (19), and finally obtain more comprehensive stochastic information about the randomness of R event over i rainy seasons. Appendix B shows that the best estimator $\hat{p}$ in binomial PMF is the unbiasedness and consistency of the MLE of p. However, as shown in Appendix C, the best estimator $\hat{\lambda}$ in the Poisson PMF has more reliability as it has not only the unbiasedness and consistency of the MLE of λ, but also the UMVUE of MLE. The UMVUE in the Poisson PMF implies that the lowest variance unbiased estimator can cause the Poisson PMF to more steadily and accurately simulate the stochasticity of soil erosion events over long-term observations than the binomial PMF.

Thirdly, besides having better simulation of the stochastic soil erosion events at larger temporal scale, the Poisson PMF is also more suitable for simulating the randomness of S event in the closed-design plot system than the binomial PMF.

Following the hypothesis of Boix-Fayos et al. (2006), the closed runoff-plot design forms an obstruction to prevent the transportable material from entering the close monitoring system, which, in particular, leads the transport-limited erosion pattern to gradually transform into a detachment-limited pattern in the closed plot over time (Boix-Fayos et al., 2007; Cammerraat, 2002). Consequently, with the extension of monitoring period, this closed-runoff plot design would make it more and more difficult for the sediment to migrate out of the plot, which also reduces the probability of observed S events under the same precipitation condition. In fact, the effect of closed-runoff plot on stochastic sediment event is also implied by the algorithm of the Poisson PMF. Specifically, in order to satisfy that $\lambda = np$ in the Poisson PMF is an unknown constant, the extension of monitoring period could lead the numbers of times of I events (n) to approach infinity, and then the probability (p) of R or S event generation would have to approach zero. The above inference coincides with the assumption about the decreasing of sediment generation in the closed-plot system, and further shows that the Poisson PMF is more reliable to simulate the stochastic erosion events at longer temporal scale.

5.2 The effect of influencing factors on erosion stochasticity

The effects of rainfall, soil and vegetation properties on erosion stochasticity in different restoration vegetation types were evaluated by LRM. This integrated stochastic rainfall events with their precipitation and intensity grades, and connected the ecohydrological functions of soil and plant with their classified hydrological and morphological features.

Just as in previous studies (Verheyen and Hermy, 2001a, b; Verheyen et al., 2003), LRM in this study explored the relative importance of morphological features disturbing the transmission of stochastic signal of I events into R and S events in different restoration vegetation types. These disturbances are closed related to the complex hydrological functions owned by different morphological structures, which finally affect the whole processes of runoff production and sediment yield (Bautista et al., 2007; Puigdefábregas, 2005).

First, many previous field experiments and mechanism models have shown that canopy structure has capacity for intercepting precipitation. This specific hydrological function can prevent rainfall from directly forming overland flow or splashing soil surface particles (Liu, 2001; Mohammad and Adam, 2010; Morgan, 2001; Wang et al., 2012). The precipitation retention by canopy structure has been regarded as an indispensable positive factor to reduce the soil erosion rate. Meanwhile, as a crucial complement to understanding the hydrological function of canopy structure, the result of LRM in this study indicated that the higher-grade canopy structure was a most important morphological feature to reduce the odds ratio of random soil events in all restoration vegetation types. This result suggests that larger canopy diameter would have relatively stronger capacity for disturbing the transmission processes of stochastic signal of rainfall on the soil surface than other morphological properties. From the perspective of erosion stochasticity, the higher-grade canopy structure could finally be attributed to the lower probability of R and S event generation. Therefore, the diversity of canopy structures in different vegetation types could play a key role in reducing both the intensity and probability of soil erosion generation.

Secondly, many studies have also discovered that denser root system distribution in the soil matrix improves the overland reinfiltration (Gyssels et al., 2005). This reinfiltration process is an effective way to recharge soil water stores when the overland flow starts to occur in hillslopes, which is also an indispensable contributing factor to reduce the unit area runoff production (Moreno-de las Heras et al., 2009, 2010). In this study, the potential reinfiltration capacity of the soil matrix could be positively affected by the saturated hydraulic soil conductivity (SHC) index. Figure 7 indicates the distribution patterns of root system in three restoration vegetation types. Meanwhile, the result of LRM also implied that the grade of SHC could negatively affect the odds ratio of stochastic erosion event, which improved the understanding of the hydrological function of plant root distribution from the viewpoint of erosion randomness. This suggests that the denser root system creates more macropores in the subsurface to provide more probability of reinfiltration of overland flow. This disturbance of overland flow by SHC can reduce the probability of erosion event generation.

Thirdly, the litter layer was shown to play multiple roles in conserving the rainfall, by improving infiltration of throughfall, as well as cushioning the splashing of raindrops (Gyssels et al., 2005; Munoz-Robles et al., 2011; Geißler et al., 2012). Therefore, the thicker litter layer in T2 (Fig. 7) probably has stronger capacity for conserving and infiltrating throughfall, as well as inhibiting splash erosion than that of other restoration vegetation types (Woods and Balfour, 2010). Although the result of LRM indicated that there was no significant correlation between the TLL and the odds ratio of soil erosion (Table 6), the interactive effect of TLL and CRO significantly affects the odds ratio of stochastic erosion events (Table 7). The interaction result implied that, under the relative low-grade CRO condition, the higher-grade TLL could have stronger disturbance on the transmission of stochastic signals of rainfall to improve the throughfall absorption to reduce the probability of splash or sheet erosion occurrence.

Additionally, Table 7 explored more interactive effects of the soil and plant properties on the odds ratio of random runoff and sediment event. These explorations suggested that the interactions between soil and vegetation properties formed more complex hydrological functions to affect the stochastic soil erosion event during ecohydrological processes in semi-arid environment (Ludwig et al., 2005).

Although the hydrological traits of vegetation played core roles in reducing the soil erosion depending on the mechanical properties of their morphological structures (Zhu et al., 2015), the LRM analysis in this study further illuminated that these hydrological-trait morphological structures of vegetation may also play an important role in affecting the stochasticity of soil erosion. Actually, the different stochasticity of soil erosion in three restoration vegetation types reflected the different extent of disturbance of vegetation type on the transmission of stochastic signals of rainfall into soil–plant systems. Therefore, the relatively smaller canopy structure, thinner litter layer and shallower root system in T3 have relatively weaker capacity to disturb the stochastic signal of rainfall than that of T1 and T2 with obvious hydrological-trait morphological structures (Fig. 7). The effect of diverse morphological structures on stochasticity of soil erosion was a meaningful complement to studying the hydrological functions of restoration vegetation types in semi-arid environment.

5.3 The implication of integrated probabilistic assessment

The IPA is an important complement to expand on the understanding of hydrological function existing in vegetation types. The hydrological-trait of morphological structures owned by different plants is closely related to the function of vegetation-driven spatial heterogeneity (VDSH) in affecting the intensity of erosion events. The VDSH theory (Puigdefábregas, 2005) can be regarded as a clear and concise summary to emphasise the dominant role of vegetation in restructuring soil erosion processes. It reflects the effect of spatial distribution patterns of vegetation on their corresponding hydrological functions in controlling erosion rate in patch, stand and even at regional . Therefore, VDSH theory has provided an innovative view to investigating the soil erosion and other ecohydrological phenomena affected by vegetation (Sanchez and Puigdefábregas, 1994; Puigdefábregas, 1998; Boer and Puigdefábregas, 2005). In the study, depending on the long-term experimental data and fundamental probability theories, the IPA concentrated on the hydrological function of VDSH in affecting the randomness of erosion events rather than the erosion rate. This can enrich the comprehension of the hydrological function of vegetation morphological structure in soil erosion phenomena, and also be an effective complement to the application of VDSH theory in interpreting stochastic erosion events.

Additionally, in our study, the IPA also provides a new framework for practitioners to develop restoration strategies focused on controlling the risk of erosion generation rather than only on reducing erosion rate. The framework contains three stages: construction of stochastic environment, description of random erosion events, and evaluation of probabilistic attribution (Fig. 8).

The first stage in the framework aims to build a unified platform to describe the stochasticity of different hydrological phenomena closely related to the erosion event. This stage generally investigates the stochastic background under which soil erosion occurs, which is also an indispensable pre-

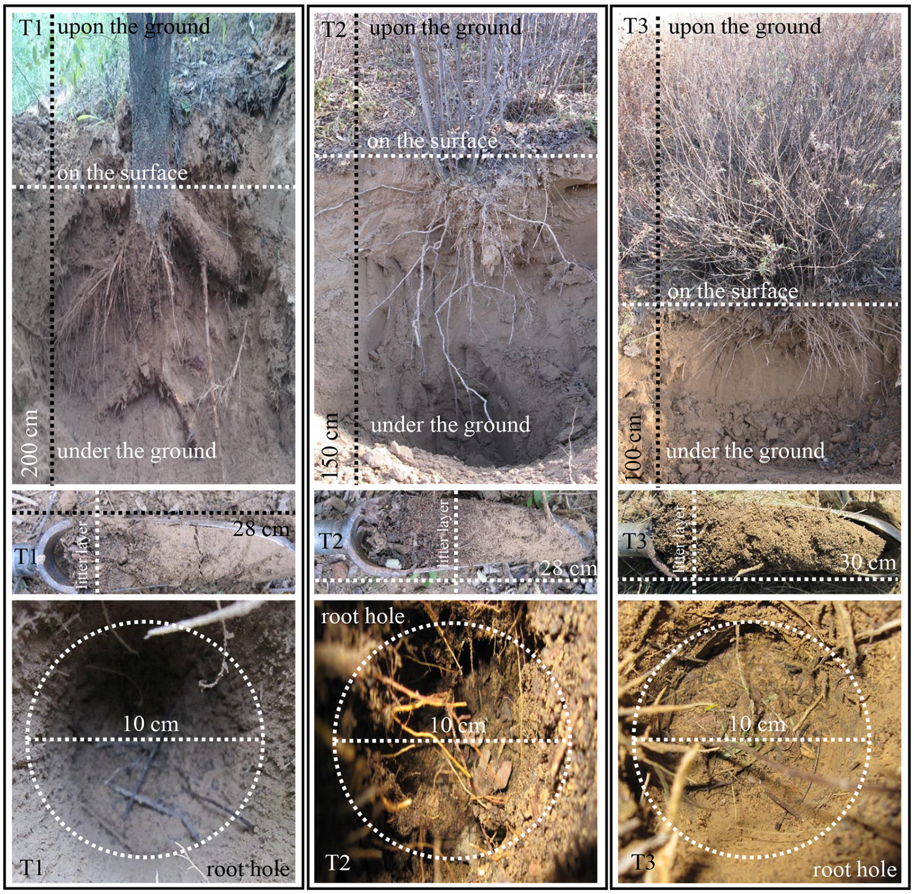

Figure 7. Morphological properties of three restoration vegetation types including the thickness of litter layer and the distribution of root system. The dashed lines indicate the diameter and depth of soil samples, approximately 10 and 30 cm respectively.

condition for quantifying the probability of R and S in stage II. The second stage is designed to construct a phased adjustment of monitoring processes based on the principle of Bayes theory as well as on the parameter analysis of binomial and Poisson models. In this phased-adjustment monitoring, the Bayes, binomial and Poisson models were applied to simulate the randomness of erosion events in short-term, mid-term and long-term monitoring stages, respectively. This model-driven monitoring approach can be regarded as a more reasonable method to explore the complexity of stochastic erosion events in larger temporal scales, but also provide a new perspective for researchers to more effectively evaluate the stochasticity of erosion events in stage III. The objective of stage III is to assess the probabilistic attribution of rainfall, soil and vegetation properties on erosion event generation. This probabilistic attribution evaluation by LRM could develop the restoration strategies for more effectively selecting vegetation types with stronger capacity for reducing the erosion risk, and finally improve the management of soil and water conservation in a semi-arid environment.

As a result, this stochasticity-based restoration strategy was developed by a combination of experimental data with multiple probabilistic theories to deal with the soil erosion randomness under complex stochastic environment. It is different from the trait-based restoration scheme derived from the functional diversity of vegetation community to reduce the soil erosion rate (Zhu et al., 2015; Baetas et al., 2009). Meanwhile, with increased monitoring duration, more stochastic information of erosion events could be added into the IPA framework. This addition could finally fulfil the self-renewal and self-adjustment of the IPA to improve the restoration strategy for selecting more reasonable vegetation types with stronger capacity for controlling erosion risk in the long term. Therefore, the IPA framework containing three stages could translate the event-driven erosion stochasticity into restoration strategies concentrating on erosion randomness, which may be a helpful complement for restoration management in a semi-arid environment.

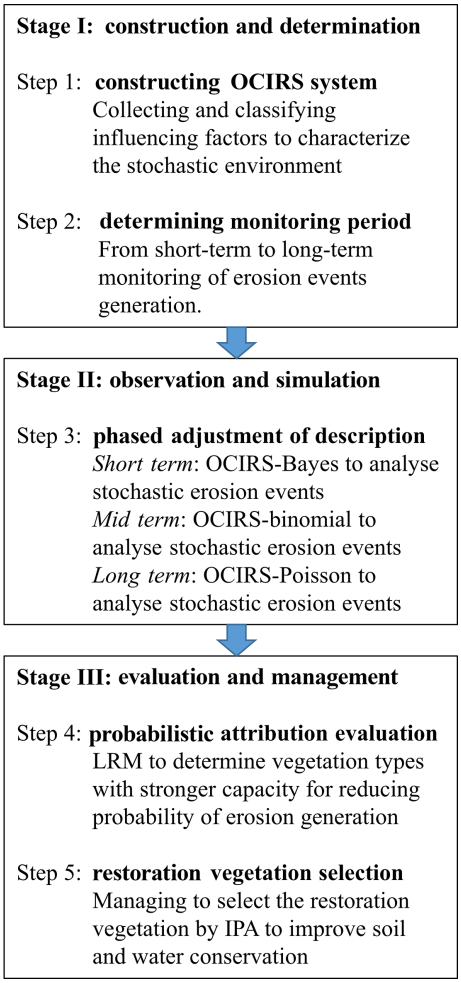

Figure 8. The framework of integrated probabilistic assessment for soil erosion monitoring and restoration strategies.

6 Conclusion

In this study, we applied an integrated probabilistic assessment (IPA) to describe, simulate and evaluate the stochasticity of soil erosion in three restoration vegetation types in the Loess Plateau of China, and draw the following conclusions.

In the IPA, the OCIRS was an innovative event-driven system to standardise the definition of hydrological random events, which is also a foundation for quantifying the stochasticity of soil erosion events under complex environmental conditions.

Both binomial and Poisson PMFs in the IPA can simulate the probability distribution of the numbers of runoff and sediment events in all restoration vegetation types. However, the Poisson PMF more effectively simulated the stochasticity of soil erosion at larger temporal scales.

The difference of morphological structures in restoration vegetation types is the main source of different stochasticity of soil erosion from T1 to T3 under the same rainfall condition. Larger canopy, thicker litter layer and denser root distribution could more effectively affect the transmission of stochastic signal of rainfall into soil erosion.

The IPA is an important complement to developing restoration strategies to improve the understanding of stochasticity of erosion generation rather than only of the intensity of erosion event. It could also be meaningful to researchers and practitioners to evaluate the efficacy of soil control practices in a semi-arid environment.

Data availability. All the data used in this study are available on request, and they can be accessed by contacting the corresponding author.

Appendix A: The transformation from binomial to Poisson PMF

Let $p = \frac{\lambda}{n}$, then

$$\mathrm{PMF}_{X\mathrm{bin}}(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1-\frac{\lambda}{n}\right)^{n-x}$$
$$= \frac{\lambda!}{x!} \cdot \frac{n(n-1)(n-2)\ldots 1}{(n-x)(n-x-1)\ldots 1} \cdot \frac{1}{n^x} \cdot \left(1-\frac{\lambda}{n}\right)^{n-x}$$
$$= \frac{\lambda!}{x!} \cdot 1 \cdot \left(1-\frac{1}{n}\right) \cdot \left(1-\frac{2}{n}\right) \ldots \left(1-\frac{x-1}{n}\right) \cdot \left(1+\frac{-\lambda}{n}\right)^n \cdot \left(1-\frac{\lambda}{n}\right)^{-x} \quad \text{(A1)}$$

In Eq. (A1), when $n \to \infty$, and x, λ is finite and constant, then

$$\lim_{n\infty}\left(1-\frac{1}{n}\right) = \ldots = \lim_{n\infty}\left(1-\frac{x-1}{n}\right) = \lim_{n\infty}\left(1-\frac{\lambda}{n}\right)^{-x} = 1 \quad \text{(A2)}$$

and

$$\lim_{n\infty}\left(1+\frac{-\lambda}{n}\right)^n = e^{-\lambda} \quad \text{(A3)}$$

and in accordance with Eqs. (A2) and (A3), Eq. (A1) can be transformed as

$$\lim_{n\infty}\left[\frac{n!}{x!(n-x)!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1-\frac{\lambda}{n}\right)^{n-x}\right] = \frac{\lambda^x e^{-\lambda}}{x!} x = 0, 1, 2, \ldots \quad \text{(A4)}$$

or

$$\mathrm{PMF}_{X\mathrm{bin}}(x) \xrightarrow{n\to\infty} \frac{\lambda^x e^{-\lambda}}{x!} = \mathrm{PMF}_{X\mathrm{poi}}(x). \quad \text{(A5)}$$

Appendix B: Parameter estimation of p in Poisson PMF

B1 Derivatization of the MLE $\hat{p}$

Let the random sample $X_1, X_2, \ldots, X_i \xrightarrow{\text{iid}} \mathrm{PMF}_{X\mathrm{bin}}(p)$ and assume the binomial distribution as

$$\mathrm{P}(X = x_i) = \binom{m}{x_i} p^{x_i} (1-p)^{m-x_i}. \quad \text{(B1)}$$

The likelihood function $L(p)$ is a joint binomial PDF with parameter p as follows:

$$L(p) = f_X(X_1, \ldots, X_n, p) = \prod_{i=1}^{n} \binom{m}{x_i} p^{\sum_{i=1}^{n} X_i}$$
$$(1-p)^{(mn - \sum_{i=1}^{n} X_i)}. \quad \text{(B2)}$$

By taking logs on both sides of Eq. (B2),

$$\ln L(p) = \ln\left(\prod_{i=1}^{n}\binom{m}{x_i}\right)$$
$$+ \sum_{i=1}^{n} X_i \ln p + \left(mn - \sum_{i=1}^{n} X_i\right) \ln(1-p) \quad \text{(B3)}$$

and differentiating with respect to p in $\ln L(P)$ and letting the result be zero,

$$\frac{\partial \ln L(p)}{\partial p} = \frac{\sum_{i=1}^{n} X_i}{p} - \frac{\left(mn - \sum_{i=1}^{n} X_i\right)}{(1-p)} = 0. \quad \text{(B4)}$$

For solution $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{mn}$, let $m = n \Longrightarrow \hat{p} = \frac{\overline{X}}{n}$.

Therefore, $\hat{p} = \frac{\overline{X}}{n}$ is the MLE of population parameter p in the binomial PMF model.

B2 Discussion of the unbiasedness and consistency of $\hat{p}$

Let $E_p(\hat{p})$ be the expectation of MLE $\hat{p}$ when population parameter p is true in random sample which is X_1, X_2, …, $X_i \xrightarrow{\text{iid}} \mathrm{PMF}_{X\mathrm{bin}}(p)$, then

$$E_p(\hat{p}) = E_{\mathrm{P}}\left(\overline{X}/n\right) = \frac{1}{n^2}\sum_{i=1}^{n} E_{\mathrm{P}}(X_i) = \frac{1}{n^2} n^2 p = p \quad \text{(B5)}$$

which shows that MLE $\hat{p} = \frac{\overline{X}}{n}$ is an unbiased estimator for p. Furthermore, let $\mathrm{Var}_p(\hat{p})$ be the variance of $\hat{p}$ when population p is true:

$$\mathrm{Var}_p(\hat{p}) = \mathrm{Var}_p\left(\sum_{i=1}^{n} X_i/n^2\right) = \frac{1}{n^4}\sum_{i=1}^{n}$$
$$\mathrm{Var}_p(X_i) = \frac{p(1-p)}{n^2}. \quad \text{(B6)}$$

As the n approaches infinity,

$$\lim_{n\infty} \mathrm{Var}_p(\hat{p}) = \left(\frac{p(1-p)}{n^2}.\right) = 0 \quad \text{(B7)}$$

Equations (B5)–(B7) satisfy the theme of the weak law of larger number, which leads the $\hat{p} = \frac{\overline{X}}{n}$ to probabilistic convergence to population parameter p:

$$\lim_{n\infty} P\left(\left|\hat{p} - p\right| \geq \varepsilon\right) = 0, \text{ for all } \varepsilon > 0. \quad \text{(B8)}$$

Consequently, the unbiased MLE $\hat{p} = \frac{\overline{X}}{n}$ is consistent for p.

Appendix C: Parameter estimation of λ in Poisson PMF

C1 Derivatization of the MLE $\hat{\lambda}$

Let the random sample X_1, X_2, …, $X_i \xrightarrow{\text{iid}} \mathrm{PMF}_{X\mathrm{poi}}(\lambda)$, and assume the Poisson distribution as

$$\mathrm{PMF}_{X\mathrm{poi}}(x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \quad \text{(C1)}$$

The likelihood function $L(\lambda)$ is joint PDF with parameter λ as follows:

$$L(\lambda) = f_X(X_1, \ldots, X_n, \lambda) = f(X_1, \lambda) \times \ldots \times f(X_n, \lambda)$$
$$= \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}. \quad \text{(C2)}$$

Taking logs on $L(\lambda)$ in Eq. (B4) and differentiating logarithm function with respect to λ:

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = \frac{\partial(\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!})}{\partial \lambda} = -n \frac{\lambda^{\sum_{i=1}^{n} X_i}}{(x_1 x_2 \ldots x_n)!} e^{-n\lambda} + \frac{\sum_{i=1}^{n} X_i \lambda^{(-1+\sum_{i=1}^{n} X_i)}}{(x_1 x_2 \ldots x_n)!}. \tag{C3}$$

Letting Eq. (C3) equal zero has solution

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}. \tag{C4}$$

Therefore, $\hat{\lambda} = \overline{X}$ is the MLE of population parameter λ in the Poisson PMF model.

C2 Discussion of the unbiasedness and consistency of $\hat{\lambda}$

Let $E_\lambda\left(\hat{\lambda}\right)$ be the expectation of MLE $\hat{\lambda}$ when population parameter λ is true in random sample $X_1, X_2, \ldots, X_i \xrightarrow{\text{iid}} \mathrm{PMF}_{X\mathrm{poi}}(\lambda)$, then

$$E_\lambda\left(\hat{\lambda}\right) = E_\lambda\left(\overline{X}\right) = \frac{1}{n^2} \sum_{i=1}^{n} E_\lambda(X_i) = \frac{1}{n} n\lambda = \lambda, \tag{C5}$$

which shows that MLE $\hat{\lambda} = \overline{X}$ is an unbiased estimator for λ. Meanwhile, let $\mathrm{Var}_\lambda\left(\hat{\lambda}\right)$ be the variance of MLE $\hat{\lambda}$ when population parameter λ is true:

$$\mathrm{Var}_\lambda\left(\hat{\lambda}\right) = \mathrm{Var}_\lambda\left(X\right) = \mathrm{Var}_\lambda\left(\sum_{i=1}^{n} X_i / n^2\right) = \frac{1}{n^4} \sum_{i=1}^{n} \mathrm{Var}_\lambda\left(X_i\right) = \frac{\lambda}{n} \tag{C6}$$

and

$$\lim_{n\infty} \mathrm{Var}_\lambda\left(\hat{\lambda}\right) = \left(\frac{\lambda}{n}\right) = 0. \tag{C7}$$

According to the weak law of large numbers (Eqs. B7, B8, C1), the unbiased MLE $\hat{\lambda} = \overline{X}$ probabilistically converges to λ:

$$\lim_{n\infty} \mathrm{P}\left(\left|\hat{\lambda} - \lambda\right| \geq \varepsilon\right) = 0, \text{ for all } \varepsilon > 0. \tag{C8}$$

Therefore, MLE $\hat{\lambda} = \overline{X}$ is consistent for population parameter λ.

C3 Determination of UMVUE $\hat{\lambda}$ of population parameter

Firstly, MLE $\hat{\lambda} = \overline{X}$ is an unbiased estimator of parameter λ which is the precondition of UMVUE determination. Secondly, by using Cramer–Rao lower bound to check whether the unbiased MLE was UMVUE or not. Then we have:

$$\ln f_X\left(X, \lambda\right) = -\ln x! + x \ln \lambda - \lambda \tag{C9}$$

$$\frac{\partial(\ln f_X\left(X, \lambda\right))}{\partial \lambda} = \frac{x}{\lambda} - 1 \tag{C10}$$

and

$$\frac{\partial^2 \ln f_X(X, \lambda)}{\partial \lambda^2} = \frac{\partial\left(\frac{x}{\lambda} - 1\right)}{\lambda} = -\frac{x}{\lambda^2} \tag{C11}$$

Accordingly the expectation of Eq. (C11) when the population parameter λ is true:

$$E_\lambda\left[\frac{\partial^2 \ln f_X(X, \lambda)}{\partial \lambda^2}\right] = E_\lambda\left(-\frac{X}{\lambda^2}\right) = -\frac{1}{\lambda^2} E_\lambda\left(X\right) = -\frac{\lambda}{\lambda^2} = -\frac{1}{\lambda} \tag{C12}$$

So the Cramer–Rao lower bound (CRLB) is

$$\mathrm{CRLB} = \frac{1}{-nE_\lambda\left[\frac{\partial^2 \ln f_X(X,\lambda)}{\partial \lambda^2}\right]} = \frac{1}{-n \cdot (-\frac{1}{\lambda})} = \frac{\lambda}{n} = \mathrm{Var}_\lambda\left(\hat{\lambda}\right) = \mathrm{Var}_\lambda\left(\overline{X}\right) \tag{C13}$$

Consequently, MLE $\hat{\lambda} = \overline{X}$ is UMVUE of population parameter λ.

The Supplement related to this article is available online at doi:10.5194/hess-21-1491-2017-supplement.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was funded by the National Natural Science Foundation of China (no. 41390464) and the National Key Research and Development Program (no. 2016YFC0501602). We are especially grateful to the associate editor and the two reviewers, whose suggestions and advice improved the quality of this study. We also thank Chen Lin-An with National Chiao Tung University (NCTU) for his great help on the mathematical statistical inference in this paper, as well as Liu Yu, Liu Jianbo and Wang Jian for their support for soil erosion monitoring.

Edited by: L. Wang
Reviewed by: J. P. Puigdefábregas, C. Miao, and one anonymous referee

References

Andrés-Doménech, I., Montanari, A., and Marco, J. B.: Stochastic rainfall analysis for storm tank performance evaluation, Hydrol. Earth Syst. Sci., 14, 1221–1232, doi:10.5194/hess-14-1221-2010, 2010.

Baetas, S. D., Poesen, J., Reubens, B., Muys, B., Baerdemaeker, D., and Meersmans, J.: Methodological framework to select plant species for controlling rill and gully erosion: application to a Mediterranean ecosystem, Earth Surf. Proc. Land., 34, 1374–1392, 2009.

Bautista, S., Mayor, A. G., Bourakhouadar, J., and Bellot, J.: Plant spatial pattern predicts hillslope runoff and erosion in a semiarid Mediterranean landscape, Ecosystems, 10, 987–998, 2007.

Bhunya, P. K., Berndtsson, R., Ojha, C. S. P., and Mishra, S. K.: Suitability of Gamma, Chi-square, Weibull, and Beta distributions as synthetic unit hydrographs, J. Hydrology, 334, 28–38, 2007.

Boer, M. and Puigdefábregas, J.: Effects of spatially structured vegetation patterns on hillslope erosion in a semiarid Mediterranean environment: a simulation study, Earth Surf. Proc. Land., 30, 149–167, 2005.

Boix-Fayos, C., Martinez-Mena, M., Arnau-Rosalen, E., Calvo-Cases, A., Castillo, V., and Albaladejo, J.: Measuring soil erosion by field plots: Understanding the sources of variation., Earth-Sci. Rev., 78, 267–285, 2006.

Boix-Fayos, C., Martinez-Mena, M., Calvo-Cases, A., Arnau-Rosalen, E., Albaladejo, J., and Castillo, V.: Causes and underlying processes of measurement variability in field erosion plots in Mediterranean conditions, Earth Surf. Proc. Land., 32, 85–101, doi:10.1002/esp.1382, 2007.

Bonham, C.: Measurements for Terrestrial Vegetation, second Edn., John Wiley & Sons. Ltd, UK, 1989.

Cammerraat, L. H.: A review of two strongly contrasting geomorphological systems within the context of scale, Earth Surf. Proc. Land., 27, 1201–1222, 2002.

Cantón, Y., Solé-Benet, A., de Vente, J., Boix-Fayos, C., Calvo-Cases, A., Asensio, C., and Puigdefábregas, J.: A review of runoff generation and soil erosion across scales in semiarid south-eastern Spain, J. Arid Environ., 75, 1254–1261, 2011.

Cao, S. X., Chen, L., and Yu, X. X.: Impact of China's Grain for Green Project on the landscape of vulnerable arid and semi-arid agricultural regions: a case study in northern Shaanxi Province, J. Appl. Ecol., 46, 536–543, 2009.

Castillo, V., Gomezplaza, A., and Martinezmena, M.: The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach, J. Hydrol., 284, 114–130, 2003.

Eagleson, P. S.: Climate, soil and vegetation 2.The distribution of annual precipitation derived from observed storm sequences, Water Resour. Res., 14, 713–721, 1978.

FAO-UNESCO: Soil map of the world (1 : 5000000), Food and agriculture organiation of the Unite Nations, UNESCO, Paris, 1974.

Freeze, R. A.: A Stochastic-Conceptual Analysis of Rainfall-Runoff Processes on a Hillslope, Water Resour. Res., 16, 391–408, 1980.

Fu, B., Liu, Y., He, C., Zeng, Y., and Wu, B.: Assessing the soil erosion control service of ecosystems change in the Loess Plateau of China, Ecol. Complex., 8, 284–293, 2011.

Geißler, C., Kühn, P., Böhnke, M., Bruelheide, H., Shi, X., and Scholten, T.: Splash erosion potential under tree canopies in subtropical SE China, Catena, 91, 85–93, 2012.

Gyssels, G., Poesen, J., Bochet, E., and Li, Y.: Impact of plant roots on the resistance of soils to erosion by water: a review, Prog. Phys. Geog., 29, 189–217, 2005.

Haberlandt, U. and Radtke, I.: Hydrological model calibration for derived flood frequency analysis using stochastic rainfall and probability distributions of peak flows, Hydrol. Earth Syst. Sci., 18, 353–365, doi:10.5194/hess-18-353-2014, 2014.

Hosmer, W., Lemeshow, S., and Sturdivant, R.: Applied Logistic Regression, Third Edn., John Willey & Sons. Inc., USA, 2013.

Janzen, D. and McDonnell, J. J.: A stochastic approach to modelling and understanding hillslope runoff connectivity dynamics, Ecol. Model., 298, 64–74, 2015.

Jiang, Z., Su, S., Jing, C., Lin, S., Fei, X., and Wu, J.: Spatiotemporal dynamics of soil erosion risk for Anji County, China, Stoch. Env. Res. Risk A., 26, 751–763, 2012.

Jiao, J., Wang, W., and Hao, X.: Precipitation and erosion characteristics of rain-storm in different pattern on Loess Plateau, Journal of Arid Land Resources and Environment, 13, 34–42, 1999.

Kim, J., Ivanov. V, and Fatichi, S.: Enviromental stochasticity controls soil erosion variabilty, Scientific Report, 6, 22065, doi:10.1038/srep22065, 2016.

Liu, S.: Evaluation of the Liu model for predicting rainfall interception in forest world-wide, Hydrol. Process., 15, 2341–2360, 2001.

Liu, Y., Fu, B., Lü, Y., Wang, Z., and Gao, G.: Hydrological responses and soil erosion potential of abandoned cropland in the Loess Plateau, China, Geomorphology, 138, 404–414, 2012.

Lopes, V. L.: On the effect of uncertainty in spatial distribution of rainfall on catchment modelling, Catena, 28, 107–119, 1996.

Ludwig, J. A., Wilcox, B. P., Breshears, D. D., Tongeay, D. J., and Imeson, A. C.: Vegetation pathces and runoff-erosion as interacting ecohydrological processes in semarid landscape, Ecology, 86, 288–297, 2005.

Mallick, J., Alashker, Y., Mohammad, S. A.-D., Ahmed, M., and Abul Hasan, M.: Risk assessment of soil erosion in semi-arid mountainous watershed in Saudi Arabia by RUSLE model coupled with remote sensing and GIS, Geocarto International, 29, 915–940, 2014.

Marques, M. J., Bienes, R., Perez-Rodriguez, R., and Jiménez, L.: Soil degradation in central Spain due to sheet water erosion by low-intensity rainfall events, Earth Surf. Proc. Land., 33, 414–423, 2008.

Miao, C. Y., Ni, J., and Borthwick, A. G.: Recent changes in water discharge and sediment load of the Yellow River basin, China, Prog. Phys. Geog., 34, 541–561, 2010.

Mohammad, A. G. and Adam, M. A.: The impact of vegetative cover type on runoff and soil erosion under different land uses, Catena, 81, 97–103, 2010.

Moore, R. J.: The PDM rainfall-runoff model, Hydrol. Earth Syst. Sci., 11, 483–499, doi:10.5194/hess-11-483-2007, 2007.

Moreno-de las Heras, M., Merino-Martin, L., and Nicolau, J. M.: Effect of vegetation cover on the hydrology of reclaimed mining soils under Mediterranean-continental climate, Catena, 77, 39–47, 2009.

Moreno-de las Heras, M., Nicolau, J. M., Merino-Martin, L., and Wilcox, B. P.: Plot-scale effects on runoff and erosion along a slope degradation gradient, Water Resour. Res., 46, W04503, doi:10.1029/2009WR007875, 2010.

Morgan, R. P. C.: A simple approach to soil loss prediction: a revised Morgan-Finney model, Catena, 44, 305–322, 2001.

Munoz-Robles, C., Reid, N., Tighe, M., Biriggs, S. V., and Wilson, B.: Soil hydrological and erosional responses in patches and inter-patches in vegetation states in semi-arid Australia, Geoderma, 160, 524–534, 2011.

Portenga, E. W. and Bierman, P. R.: Understanding earth's eroding surface with ^{10}Be, GSA Today, 21, 4–10, 2011.

Prasannakumar, V., Shiny, R., Geetha, N., and Vijith, H.: Spatial prediction of soil erosion risk by remote sensing, GIS and RUSLE approach: a case study of Siruvani river watershed in Attapady valley, Kerala, India, Environmental Earth Sciences, 64, 965–972, 2011.

Prooijen, B. C. and Winterwerp, J. C.: A stochstic formulation for erosion of cohesive sediments, J. Geophys. Res., 115, C01005, doi:10.1029/2008JC005189, 2010.

Puigdefábregas, J.: The role of vegetation patterns in structuring runoff and sediment fluxes in drylands, Earth Surf. Proc. Land., 30, 133–147, 2005.

Puigdefábregas, J.: Ecological impacts of global change on drylands and their implications for desertification, Land Degrad. Dev., 9, 393–406, 1998.

Puigdefábregas, J., Solé-Benet, A., Gutierrez, L., Barrio, G., and Boer, M.: Scales and processes of water and sediment redistribution in drylands: results from the Rambla Honda field site in Southeast Spain, Earth-Sci. Rev., 48, 39–70, 1999.

Ridolfi, L., D'Odorico, P., Porporato, A., and Rodríguez-Iturbe, I.: Stochastic soil moisture dynamics along a hillslope, J. Hydrol., 272, 264–275, 2003.

Robert, V. H., Joseph, W. M., and Allen, T. C.: Introdcution to Mathematical Statistics, Pearson Education, Inc., USA, 2013.

Rodriguez-Iturbe, I., Porporato, A., Laio, F., and Roidolfi.L: Plants in water-controlled ecosystems:active role in hydrologic processes and response to water stress I. Scope and general outline, Adv. Water Resour., 24, 695–705, 2001.

Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V., and Cox, D. R.: Probabilstic modeling of water balance at a point: the role of climate, soil and vegetation, P. Roy. Soc Lond. A Mat., 455 3789–3805, 1999.

Sanchez, G. and Puigdefábregas, J.: Interactions of plant growth and sediment movement on slopes in a semi-arid environment, Geomorphology, 9, 243–360, 1994.

Sheldon, R.: A first course in probability, Ninth Edn., Pearson Education Limited, USA, 2014.

Sidorchuk, A.: Stochastic modelling of erosion and deposition in cohesive soil, Hydrol. Process., 19, 1399–1417, 2005.

Sidorchuk, A.: A third generation erosion model: the combination of probailistic and determinstic components, Geomorphology, 2009, 2–10, 2009.

Verheyen, K. and Hermy, M.: The relative improtance of disperal limitation of vascular plants in secondary forest succession in Muizen Forest, Belgium, J. Ecol., 89, 829–840, 2001a.

Verheyen, K. and Hermy, M.: An integrated analysis of the spatio-temporal colonization patterns of forest plant species in a mixed deciduous forest, J. Veg. Sci., 12, 567–578, 2001b.

Verheyen, K., Guntenspergen, G., Biesbrouck, B., and Hermy, M.: An integrated analysis of the effects of past land use on forest herb colonization at the landscape scale, J. Ecol., 91, 731–742, 2003.

Wang, S., Fu, B., Piao, S., Lu, Y., Ciais, P., Ciais, P., and Wang, Y.: Reduced sediment transport in the Yellow River due to anthropogenic changes, Nat. Geosci., 9, 38–41, 2015.

Wang, G., Gertner, G., Liu, X., and Anderson, A.: Uncertainty assessment of soil erodibility factor for revised univeral soil loss equation, Catena, 46, 1–14, 2001.

Wang, G., Gertner, G., Singh, V., Shinkareva, S., Parysow, P., and Anderson, A.: Spatial and temporal prediction and uncertainty of soil loss using the revised universal soil loss equation: a case study of the rainfall–runoff erosivity R factor, Ecol. Model., 153, 143–155, 2002.

Wang, P. and Tartakovsky, D. M.: Reduced complexity models for probabilistic forecasting of infiltration rates, Adv. Water Resour., 34, 375–382, 2011.

Wang, S., Fu, B. J., Gao, G. Y., Liu, Y., and Zhou, J.: Responses of soil moisture in different land cover types to rainfall events in a re-vegetation catchment area of the Loess Plateau, China, Catena, 101, 122–128, 2013.

Wang, X., Zhang, Y., Hu, R., Pan, Y., and Berndtsson, R.: Canopy storage capacity of xerophytic shrubs in Northwestern China, J. Hydrol., 454, 152–159, 2012.

Wei, L. and Zhang, W.: Bayes Analysis, Press of University of Science and Technology of China, Beijing, 2013.

Wei, W., L., C., Fu, B., Huang, Z., Wu, D., and Gui, L.: The effect of land uses and rainfall regimes on runoff and soil erosion in the semi-arid loess hilly area, China, J. Hydrol., 335, 247–258, 2007.

Wischmeier, W. H. and Smith, D. D.: Predicting rainfall erosion losses: A guide to conservation planning, Agriculture handbook Number 537, United States Department of Agriculture, 1978.

Woods, S. W. and Balfour, V. N.: The effects of soil texture and ash thickness on the post-fire hydrological response from ash-covered soils, J. Hydrol., 393, 274–286, 2010.

Yazdi, J., Salehi Neyshabouri, S. A. A., and Golian, S.: A stochastic framework to assess the performance of flood warning systems based on rainfall-runoff modeling, Hydrol. Process., 28, 4718–4731, 2014.

Zhou, J., Fu, B., Gao, G., Lü, N., Lü, Y., and Wang, S.: Temporal stability of surface soil moisture of different vegetation types in the Loess Plateau of China, Catena, 128, 1–15, 2015.

Zhou, J., Fu, B., Gao, G., Lü, Y., Liu, Y., Lü, N., and Wang, S.: Effects of precipitation and restoration vegetation on soil erosion in a semi-arid environment in the Loess Plateau, China, Catena, 137, 1–11, 2016.

Zhu, H., Fu, B., Wang, S., Zhu, L., Zhang, L., and Jiao, L.: Reducing soil erosion by improving community functional diversity in semi-arid grassland, J. Appl. Ecol., 52, 1063–1072, doi:10.1111/1365-2664.12442, 2015.