Article

Stochastic Assessment of Scour Hazard

David Flores-Vidriales [1], Roberto Gómez [1] and Dante Tolentino [2],*

[1] Instituto de Ingeniería, Universidad Nacional Autónoma de Mexico, Mexico City 04510, Mexico; ing.dfv.1991@gmail.com (D.F.-V.); RGomezM@iingen.unam.mx (R.G.)
[2] Departamento de Materiales, Universidad Autónoma Metropolitana, Mexico City 02200, Mexico
* Correspondence: dantetl@azc.uam.mx

Abstract: Scour is the most frequent cause of bridge collapses in Mexico. Bridges located along the Mexican Pacific coast are exposed to extreme rainfall originating from tropical storms and hurricanes. Such environmental phenomena trigger sediment loss, which is known as scour. If maintenance actions are not taken after scouring events, the scour depth increases over time until the bridge collapses. A methodology to estimate the scour hazard considering both the scour–fill interaction and the Monte Carlo simulation method is proposed. The general extreme value probability distribution is used to characterize the intensity of the scouring events, the lognormal distribution is used to characterize the sedimentation process (fill), and a homogeneous Poisson process is used to forecast the occurrence of both types of events. Based on the above, several histories of scour–fill depths are made; such simulations are then used to develop time-dependent scour hazard curves. Different hazard curves associated with different time intervals are estimated for a bridge located in Oaxaca, Mexico.

Keywords: scour; stochastic analysis; forecasting; extreme events

1. Introduction

A bridge built over a river creates an obstruction to the water flow, changing the local flow field. The erosion around the obstruction is accelerated due to the following three variables: (a) a velocity surge, (b) a difference in pressure between upstream and downstream areas produced by the perpendicular action of the flow, and (c) the generation of vortices. Scour begins when a certain level of shear stress, known as critical Shields stress, is reached. The sediment at the river bed lifts due to the action of vortices and travels downstream; such vortices produce shear stress on the sediment particles lying on the river bed and contribute to the scour process. The critical Shields stress is a function of size, shape, and material of the sediment particles.

There are two current trends that study sediment transport: trend one uses computer fluid dynamics (CFD) to solve the Navier–Stokes equations; such an approach is useful for deterministic analyses and is the best option to precisely simulate the phenomenon under study, but it demands both a great amount of computer time and a high level of capabilities. Trend two consists of a stochastic approach that is not useful to accurately forecast the outcome of a single event, but it is efficient to forecast events in probabilistic terms, and it can be a good tool to study time-dependent problems such as the case of the thickness of the sediment layer on a river bed that changes over time. The sediment transport at any given point in a time lapse, t, is assumed to have two possible outcomes: erosion, which is also known as scour, and accretion, which is also known as fill. Regarding the latter, a mathematical model that consists of both a data generation process (DGP) and a set of suitable simplified sediment transport equations can be used to forecast the fluctuations of the sediment layer thickness within a reasonable error.

The mechanics of sediment transport is a complex issue to study from a stochastic point of view. Thus, assumptions need to be made in order to develop a simplified mathematical approach that is useful to a probabilistic approach. Research dealing with semi-empirical

equations for the sediment deposit is scarce. The rate at which sediment is deposited in the river bed is dependent on the availability of sediment and, thus, from the erosion of the bed and walls of the river from upstream. Variables of hydraulics, geometry, and geotechnical are fundamental for sediment transport; they also govern the behavior of secondary variables such as turbulence, flow velocity, and shear stress in the sediment particles [1]. Many studies related to the probabilistic assessment of scour hazard have been developed [2–13]. A simulation that allows forecasting the outcome of an event is a mathematical representation of events that have already happened; such simulation cannot add any new knowledge about the behavior of previous events, but it can provide useful information about possible future events.

In the last 20 years, Mexico has experienced an average of 2.5 bridge collapses per year due to scour effects [14], making scour the leading cause of bridge collapses in Mexico and most of the world [15]. The Mexican coastal highway system comprises geographic zones that are prone to extreme rainfall. For example, in 2005, hurricane Stan severely damaged highway infrastructure and caused the collapse of several bridges in the state of Chiapas. Thus, it is important to develop tools and methodologies to reduce bridge collapses due to scour.

This paper aims to propose a methodology to estimate scour hazard curves considering the following: (a) a preprocessing stage to fix missing data issues using data augmentation, (b) an information criterion for the probability distribution selection, and (c) the interaction of scour and fill events through time. The proposed methodology is illustrated in a bridge with potential scouring problems located in Mexico. Such an approach can deal with missing data, or problems such as outliers, and also has the novelty feature of addressing the accretion process that lowers maximum scour depth, making the mathematical model more complete. Moreover, the methodology is compared with the approach proposed by [5] in which the scour survival function is obtained by the SRICOS EFA method.

2. Data Processing

Most databases in Mexico are somewhat flawed: they either have missing data or have outliers. Such issues introduce an additional source of uncertainty that needs to be reduced in a preprocessing stage. First, outliers are defined as any value above or below 1.5 of the interquartile range, IQR. Such values are removed from the data set and then treated as missing at random, as per Rubin [16]. This mechanism considers data that are not systematically related to hypothetical values of known data. Thus, it is possible to use statistical inference to obtain probable values for the missing data without having to include any information about the missing data. Hence, a methodology that produces the least amount of bias is needed, such as data augmentation, K nearest neighbors, multiple imputation, etc. [17–20]. In this paper, the data augmentation technique [21] is used to address the missing data issue. Such a technique has two advantages: (a) it produces a low amount of bias, and (b) it is relatively simple when compared with complex optimization techniques used for missing data analysis.

2.1. Preprocessing

The available data that are recorded from the observation of a natural phenomenon such as discharge , Q, for a time interval $t \in Z$ can be used to form an observation set, F, in which its members are termed events, and denoted by $\omega \in F$. Let Ω be a set representing all the possible outcomes of the phenomenon. Then, the probability space of the phenomenon is $(\Omega, F, P_\theta \mid \theta \in I)$, where the parameter space, I, has two sources of statistical uncertainty: one is called stochastic uncertainty, which deals with the uncertainty about a fixed parameter θ (Figure 1); the other is called inductive uncertainty.
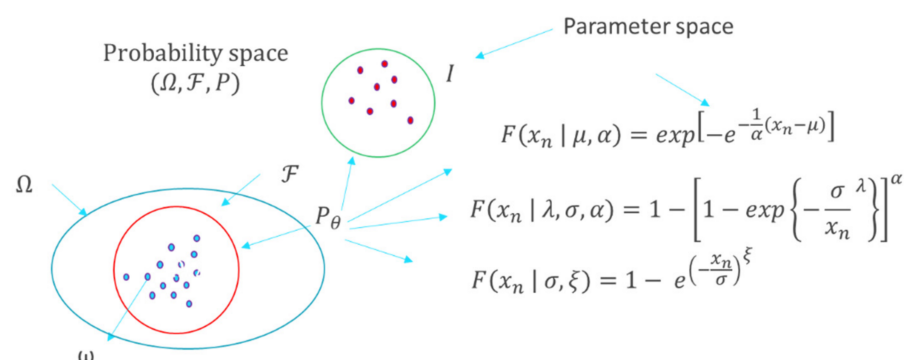
Figure 1. Probability space, parameter space, and probability distribution.

The uncertainty associated with each event of the subset *F* is represented by a probability function, *P*, which contains all possible outcomes of ω (if the sample is large enough). If the uncertainty of *P* is contained in the parameter space, then, for each *I* there is a possibility of having a different probability distribution, P_θ, $\theta \in I$, where θ is a set of parameters that govern the distribution function. Hence, stochastic problems are referred to as inductive, meaning that they provide predictions about events that have not yet happened using prior observations.

2.2. Statistical Inference of θ

The statistical estimation of the conditional distribution parameters of an observed data set $\theta \in I$ is known as inference. In order to infer the parameters for a given distribution $F = P_\theta[X < x \mid \theta \in I]$, an inference approach must be chosen first.

Frequentist inference is based on the idea of a limiting frequency. One of the issues of using this approach is the intrinsic epistemic uncertainty of the parameters θ. Such approach does not provide statements about θ, which is a fixed value. Thus, the uncertainty lies only in the observations ω (aleatory uncertainty). A frequentist approach usually has a predetermined sample size because *p* values are calculated over a sample space Ω [22]. In Bayesian inference, θ is considered a random variable. Some knowledge about the distribution of θ is assumed to be available, which is called a prior distribution. Then, after observing some data ω, the distribution of θ must be updated to match the knowledge obtained during the observation process. The Fisherian approach uses the likelihood function, *L*, to compare different values of θ with the probability of the observed data so that $L(\theta) = P_\theta$ [17]. Such likelihood provides a relative measure of the set of parameters θ, whose exact value is impossible to know. Hence, the parameter values that maximize the likelihood function $L(\theta)$ must be found. Based on the above, the Fisherian approach is used because maximizing the likelihood function $L(\theta)$ does not require previous knowledge of the phenomena, and θ does not have an underlying distribution to propose. The combined likelihood for a given distribution of *n* parameters can be expressed as

$$L(\theta|data) = \prod_{n=1} P_\theta(data \mid \theta) \tag{1}$$

The maximum of the likelihood function is obtained by two steps. The log-likelihood function is obtained, and then the score function, *S*, is calculated as follows:

$$S(\theta) \equiv \nabla loglog\ L(\theta) \tag{2}$$

Hence, we obtain the maximum likelihood estimators, MLE, and the solution of θ, $S(\theta) = 0$. MLE is not flawless, and for small data samples, the Simpson's paradox could occur. Moreover, most of the real score functions do not have a closed solution. Thus, a numerical method must be used to approach the solution. Maximum likelihood estimators

(MLE) are often the best way to deal with stochastic uncertainty because they require no prior knowledge about the phenomena.

2.3. Missing Data Inference

The second source of uncertainty is known as inductive uncertainty and deals with the lack of information of the observations available. Rubin [23] uses multiple imputation techniques to reduce the inductive uncertainty. When a small data set is used, stochastic uncertainty becomes relevant, and inductive uncertainty becomes critical [17]. Hence, it is important to rely on data sets with reasonable sizes. In order to reduce the inductive uncertainty, the missing at random mechanism (MAR) is used as follows [21]:

$$f(Q, M | \theta, \psi) = f(Q | \theta) \times f(M|Q, \psi) \tag{3}$$

where M is the missing data indicator with values between 0 and 1, which represent an observed value and no observed value, respectively; $f(\omega, M | \theta, \psi)$ is the joint probability distribution of (ω, M). The set of parameters that governs the distribution of the observed data in the subset F_{obs} is θ, whereas for the missing data F_{miss} is ψ. Hence, under the premise of MAR, given a set of observed data ω that is augmented by a quantity Z, the augmented posterior distribution $P(\theta|\omega, Z)$ can be calculated as

$$f(M | F_{obs}, F_{miss}, \psi) = f(M | F_{obs}, \psi) \times f(F_{obs} | \theta) \tag{4}$$

which means that the parameters from the data model θ are independent of the parameters in the missing data. Authors in [24,25] developed the following data augmentation algorithm:

$$F_{miss}^{(t+1)} \approx P_r(F_{miss}|F_{obs}, \theta^t) \tag{5}$$

$$\theta^{(t+1)} \approx P_r(F|Q_{obs}, F_{miss}^{(t+1)}) \tag{6}$$

where P_r is the probability associated to F_{miss}. Equation (5) is the imputation step, which generates imputed values for F_{miss} using the missing data probability distribution, observed data F_{obs}, and the set of parameters θ at the given iteration t. Equation (6) generates parameter values for the posterior distribution, using the observed data F_{obs} and the imputed values of F_{miss} at the iteration $t + 1$. This procedure is performed iteratively until convergence is reached, and a new "complete" data set F_{full} is generated.

2.4. Suitable Probability Distribution

The suitable distribution function of the data $F(\omega, t \in T)$ is the backbone of the DGP. In order to select the best fitting distribution to the data set, the Akaike information criterion [26], *AIC*, is used, which selects the best probability distribution based on the parsimony principle. *AIC* is obtained using the following equation:

$$AIC = -2n \, \underline{L}(\theta|F) + 2k \tag{7}$$

where n is the number of observations in F, $\underline{L}(\theta|F)$ is the log-likelihood function, and k is the number of parameters in the mathematical model.

2.5. Simulation of Events for the Proposed Approach

The block maxima approach (BMA) is used to estimate the intensity and occurrence of extreme river discharges associated with extreme rain falls. Such an approach divides the observations of the "complete" data set, F_{full}, into non-overlapping blocks that restrict the observation of the maximum on such periods. One-year blocks are used in this research since there is no seasonality in the maximum annual discharge [27]. The generalized

extreme value distribution (GEV) is used to describe the annual maximum discharge, whose CDF is

$$F(x \mid \sigma, \mu, \xi \neq 0) = exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \tag{8}$$

$$F(x \mid \sigma, \mu, \xi = 0) = exp\left\{-\frac{x-\mu}{\sigma}\right\} \tag{9}$$

where μ, σ, ξ are the location, scale, and shape parameters, respectively. The law of small numbers provided the basis for the selection of the Poisson stochastic process for the time of occurrence of events that trigger scour and fill. Using this law, the stochastic process of events that exceed a given threshold is

$$P(k \geq K \mid t = T) = e^{-\lambda T}\frac{(\lambda T)^k}{k!} \tag{10}$$

where

$$\lambda_e = \frac{\sum_{i=1}^{n} m \mid \forall\, m \geq \psi}{n} \tag{11}$$

and

$$\psi \cong 0.85IQR \tag{12}$$

Since there is only interest in the probability of occurrence of a single event [28], the cumulative distribution function (CDF) of the arrivals $F_a(t)$ in the Homogeneous Poisson Process (HPP) is obtained as follows:

$$F_a(t) = \int_0^\infty 1 - e^{-\lambda t}\, d_t \tag{13}$$

Having characterized the intensity of the events and its occurrence, the joint probability distribution (JPD) is

$$f_{xy}(x,y) = \sum_{y=1}^{n} \frac{\lambda^y e^{-\lambda}}{y!}\left(\int_{-\infty}^{\infty}\frac{1}{\sigma}exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}\left[1 + \xi_1\left(\frac{x-\mu}{\sigma}\right)\right]^{-1-\frac{1}{\xi}}dx\right) \tag{14}$$

where y is the number of events in a one-year block maxima.

Events that trigger fill have a random nature and have a somewhat fixed occurrence in a year. Such events occur both at the beginning and at the end of the rainy season, when the discharge is at its weakest intensity, but it is still sufficiently strong to transport sediment from the upstream river flow. Therefore, a stochastic model is needed to account for these observations that will trigger the accretion process. The intensity of the discharge, q_t, is considered to follow a lognormal distribution function whose CDF is

$$F(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^x \frac{1}{t}exp\left\{\frac{-\log(t-\mu)^2}{2\sigma^2}\right\}dt \tag{15}$$

where μ and σ are the shape and location parameters, respectively. The occurrence is obtained in a similar way as extreme events, with the difference in the threshold ψ_2 as follows:

$$\lambda_r = \frac{\sum_{i=1}^{n} m \mid \forall\, m \leq \psi_2}{n} \tag{16}$$

where ψ_2

$$\psi_2 \cong 1.2IQR \tag{17}$$

Thus, the JPD for the fill is

$$f_{xy}(y, t) = \sum_{y=1}^{n} \frac{\lambda^y e^{-\lambda}}{y!} \left(\frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{1}{t} exp \left\{ \frac{-\log(t-\mu)^2}{2\sigma^2} \right\} dt \right. \tag{18}$$

2.6. Simulation of Time Series for SRICOS-EFA Method

A time series process is a set of observations that are well defined and have regular intervals; thus, it is a set of random variables $\{X_t\}$ indexed by integers t (often associated by dates). A stochastic process is a parametrized collection of random variables $\{X_t\}_{t \in T}$ defined on a probability space $(\Omega, F, P_\theta | \theta \in I)$. Then, a time series can be understood as a realization of a stochastic process. Time series are often studied on the time domain using an autoregressive approach. According with the above, an autoregressive moving integrated average model (ARIMA) is used as follows:

$$y\prime_t = c + \phi_1 y\prime_{t-1} + \dots \phi_p y\prime_{t-p} + \theta_p \varepsilon_{t-1} + \dots \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{19}$$

where $y\prime_t$ is the difference series, p is the order of the autoregressive, d is the degree of the first differencing, and q is the order of the moving average. The ARIMA model that best fits the data is selected by the AIC criterion [26].

Once the model is selected, simulations are performed under a random walk approach for the extreme events data set (maximum annual discharge) and then used to estimate the scour depth using the SRICOS-EFA methodology [5]. These results are then compared to the proposed approach explained in Section 2.5.

3. Scour and Fill
3.1. Scour

Local scour depth, y_s around a bridge pier is obtained using the one-dimensional HEC 18 approach [19] as follows:

$$y_s = y_n K_4 \left[2K_1 K_2 K_3 \left(\frac{b}{y_n} \right)^{0.65} F^{0.43} \right] \tag{20}$$

where y_n is the height of the stream (derived from the Manning equation), K_4 is a bed material coefficient, K_1 is a correction factor for the pier nose shape, K_2 is a correction factor for the angle of attack of the flow, K_3 is a correction factor for bed condition, b is the pier width, and F is the Froude number. The local scour area, A_s, and wetted perimeter of the scour hole, P_{mh}, are calculated as

$$A_s \cong y_s^2 \tag{21}$$

$$P_{mh} \approx 2.33 y_s \tag{22}$$

When calculated, these values are added to the hydraulic area and wetted perimeter after each scouring event.

3.2. Fill

Haschenburguer [29] proposed a model to estimate the mean fill using both the Shields stress and the hydraulic radius, which are both considered discharge-dependent variables. The Haschenburguer fill model estimates the sediment deposit from a given intensity for small-scale events in stable channels. The model uses the Shields stress as the fundamental variable, and it does not limit the amount of accretion, a fact that could lead to unrealistic heights of the sediment layer. Such issue is also present in the HEC 18 equations for scour. Hence, it is considered to have a sufficient degree of approximation [30,31].

The scour–fill process is controlled by the transfer of material in the river bed. Thus, scour depths are time-dependent random variables. The underlying process which governs sediment transport has been a constant topic for research [32–34]. A Poisson distribution

coupled with a homogeneous Poisson process (HPP) is used, so that the occurrence of the phenomenon can be expressed with a semiempirical PDF as [35]

$$f\left(y_f\right) = \phi e^{-\phi y_{yf}} \tag{23}$$

where $f\left(y_f\right)$ is the proportion of stream bed fill for a given flow depth, y_{yf}, which is the flow depth obtained from the minimum discharge at the start of the rainy season.

$$\phi = 3.3 e^{-1.52\tau^*/\tau} \tag{24}$$

where the Shields mean stress, τ^*, is

$$\tau^* = \frac{\rho_w R_H S}{(\rho_s - \rho_w) D_{50}} \tag{25}$$

where ρ_w is the density of the sediment, ρ_s is the density of water at 20 °C, D_{50} is the mean size value of the particles, and R_H is the hydraulic ratio at the beginning and at the end of the rainy season. The hydraulic ratio is different from the one considered in the scour process, and S is the slope of the hydraulic grade line. A shear velocity criterion for incipient motion of sediment is used as follows [36]:

$$\tau \cong 0.215 + \frac{6.79}{D_{50}^{1.70}} - 0.75 \times exp\left(-2.61 \times 10^{-3} \times D_{50}\right) \tag{26}$$

The local fill area, A_{sf}, and the wetted perimeter, P_{mf}, are subtracted after each event and are calculated as

$$A_{sf} \cong f\left(y_f\right)^2 \tag{27}$$

$$P_{mf} \approx 2.33 f\left(y_f\right) \tag{28}$$

4. Scour Hazard Curves

4.1. Simulation of Random Events

Both the hydraulic radius (H_t) and the Froude number (F) change through time as scour and fill develop. Hence, an analytical solution to the stochastic process is difficult to obtain. A numerical approximation to the solution is used instead. Monte Carlo simulation (MCS) is a tool for the analysis of complex systems. The MCS approach used in this research simulates a large number of independent histories of occurrence–intensity, which are then used to estimate the time progression of scour. Each simulated history is one possible outcome in the probability space. Each history behaves differently because of the intrinsic stochastic nature of the state variables (scour depth, flow, Froude number, etc.). If a sufficient number of simulated histories is known, the probability function P of the phenomena can be approximated (law of large numbers).

Usually scour and fill in a bridge system (a collection of variables and components) are dealt with in a discrete time fashion, especially when a probabilistic approach is used. In this research, a discrete-event system is used, assuming that the state variables change instantly through discrete points in the simulation time. Since the form of the solution is unknown, a raw sampling approach was used.

4.2. Hazard Curves

Hazard curves show the intensity of scour depth and its corresponding probability of exceedance. The probabilistic scour assessment provides the scour intensity corresponding to a target hazard level (annual probability of exceedance). The proposed hazard curve for scour is site specific, which means that there is not a variable that deals with the source-to-site distance, which is a consequence of the fundamental approach to the phenomena. The location of the phenomenon does not have an influence on the model in the approach

proposed. Therefore, the effects of the phenomenon govern the model, and its location is negligible. The annual rate of exceedance for a given scour depth can be estimated as follows [37]:

$$P[y_s(te) > y] = \frac{\lambda_e}{\xi_0}[\xi_{yt}] \tag{29}$$

where ξ_0 is the mean number of events that exceed a given scour threshold, (ψ_s), during an interval t. $P[y_s(te) > y]$ is the probability that one or more scour depth events $y_s(te)$ are greater than a threshold (y), and ξ_{yt} is the exceedance rate given a time interval t and is estimated as

$$\xi_{yt} = 1 - (1 - P[y_s(te) > y])^t \tag{30}$$

Thus, the annual rate of exceedance can be obtained by

$$\xi_0 = \frac{\left[1 - \int_{-\infty}^{x} exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right]dx\right] * \xi_{yt}}{\lambda_e} \tag{31}$$

5. Case Study

The bridge is located in the state of Oaxaca in one of the highest scour-prone zones in Mexico. The bridge is a four-span, simply supported bridge; each span has a length of 56 m, as shown in Figure 2a. Bridge piers and foundation piles are built with reinforced concrete in a multi-column bent type with circular columns of 1.2 m in diameter, as seen in Figure 2a. A simplified trapezoidal cross-section is used, as shown in Figure 2b.

(**a**)   (**b**)
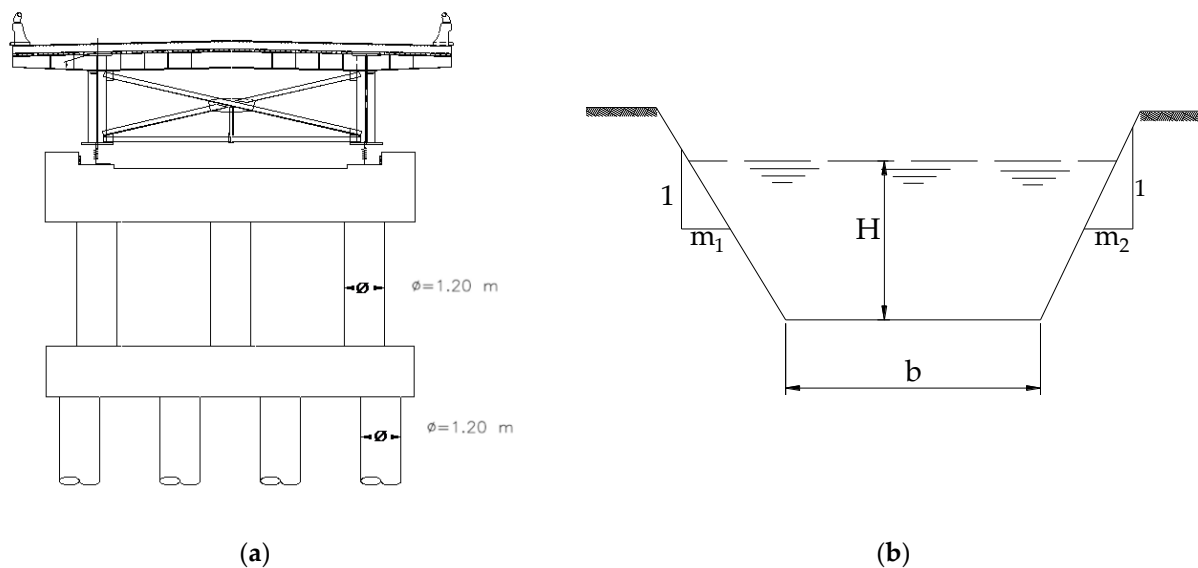
Figure 2. Study case: (**a**) typical pier of the bridge; (**b**) simplified cross-section of the river.

The properties of the simplified cross-section for a given value of H are as follows:

$$A_h = H\left\{b + \frac{H(m_1 + m_2)}{2}\right\} \tag{32}$$

$$P_m = b + H\left[\sqrt{1 + m_1^2} + \sqrt{1 + m_2^2}\right] \tag{33}$$

$$Q = \frac{1}{n}A_h R_h^{\frac{2}{3}} S_h^{\frac{1}{2}} \tag{34}$$

where A_h is the hydraulic area, P_m is the wetted perimeter, $R_h = A_{ht}/P_m$, and H is the height of the stream for a given discharge Q. Equation (34) can be written as

$$Q = \frac{1}{n}\left[\frac{\left(H\left\{b + \frac{H(m_1+m_2)}{2}\right\}\right)^{5/3}}{\left(b + H\left[\sqrt{1+m_1^2} + \sqrt{1+m_2^2}\right]\right)^{2/3}}\right]S_h^{\frac{1}{2}} \qquad (35)$$

H_t is the new value of y_s, and in a similar way τ^* is deducted from the river discharge for recurring events; thus, two different hydraulic ratios are used to find the values of y_s and τ^*. The values used in Equation (35) are taken from the Monte Carlo simulation. For this case study $b = 80\ m$, $m_1 = 0.8$, $m_2 = 1$, $s = 0.0068$, and $n = 0.025$.

5.1. Data Augmentation

Figure 3 shows the river discharge, Q, obtained from the National Superficial Water Database for this river. Both the maximum annual record and the minimum annual discharge are presented. Scour phenomenon is controlled by events that are infrequent in nature and have a high magnitude; these are known as "extreme events". Extreme events are represented in the maximum annual discharge data base, and they are used to forecast scour depths. Unlike scour, fill is controlled by recurring events at the beginning and end of the rainy season; then, the minimum discharge is used to forecast sediment deposit depths. From the data set shown in Figure 3, a "complete" data set is estimated using the data augmentation algorithm described in Section 2. Results are shown in Figure 4.
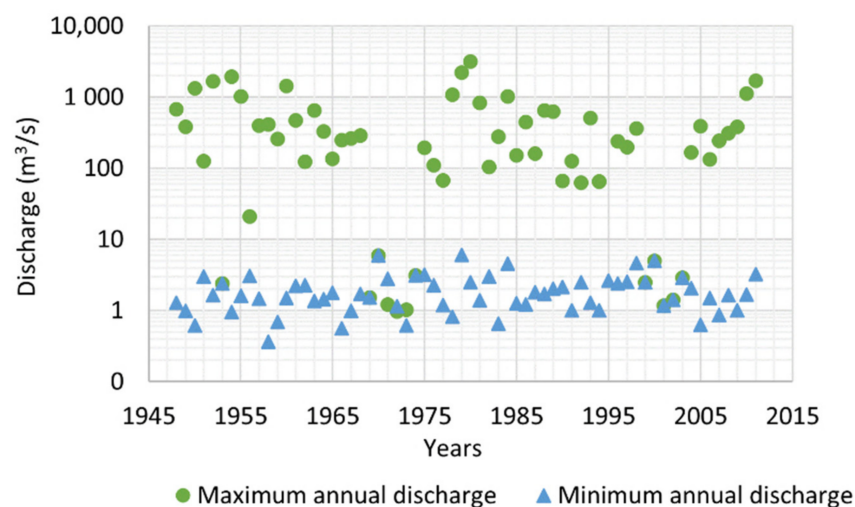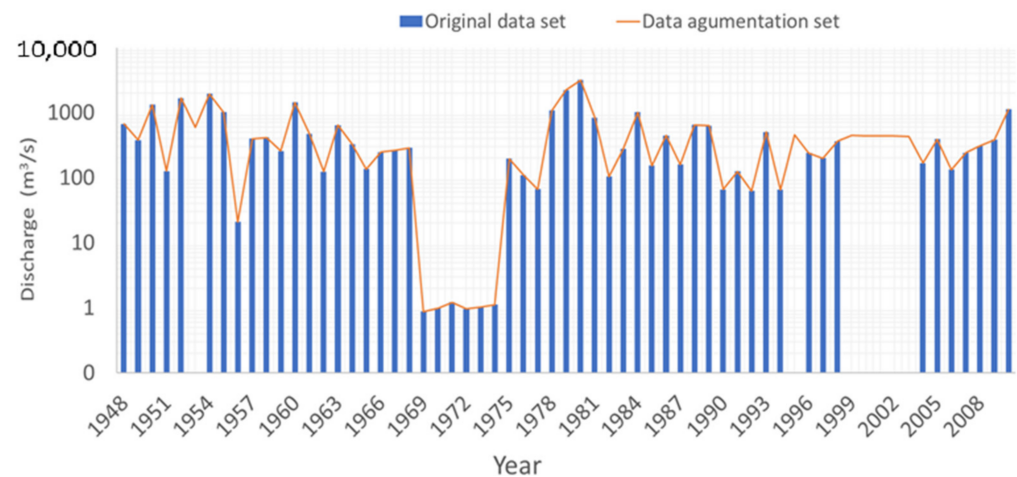

Figure 3. Discharge data.

Figure 4. Comparison between original data set and data augmentation set.

5.2. Fit to a Probability Distribution

Different probability distribution functions (PDFs) are tested to find the appropriate PDF that best characterizes the maximum annual discharge and, in a similar fashion, the minimum annual discharge. Therefore, the PDF with the lowest AIC is selected. Figure 5a shows the AIC for different PDFs such as generalized extreme value (GEV), Weibull 2P, normal, exponential, and logistic. It is noticed that the general extreme value distribution is the best fit for the maximum annual discharge. Figure 5b shows different PDFs to characterize the minimum annual discharge data at the beginning and end of the rainy season. Figure 5b indicates that the lognormal distribution presents the minimum AIC; thus, it is the best fit for the minimum flow.
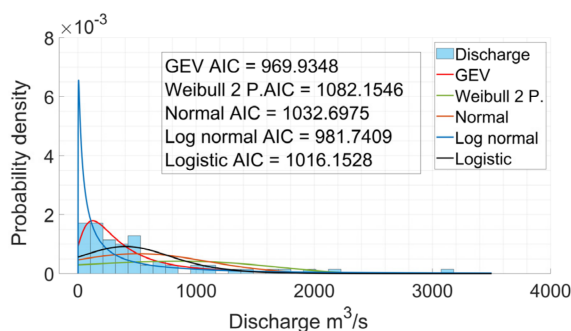

(**a**)
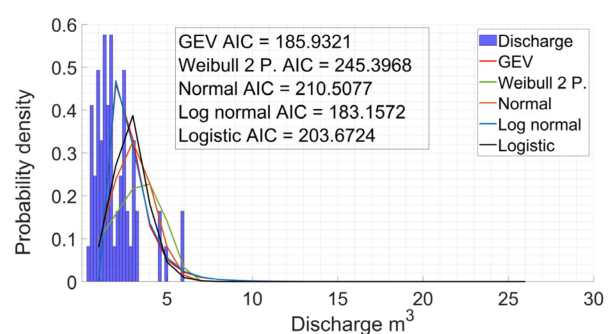
(**b**)

Figure 5. Different probability distribution functions: (**a**) data histogram of maximum annual flow; (**b**) histogram of minimum flow.

The selected PDFs that characterize the maximum and minimum annual discharge are represented in probability papers. The probability paper can be defined as a mathematical–graphic technique to verify if the data follow a PDF. Figure 6a shows the maximum annual discharge and the GEV PDF with a continuous line in probability paper. Figure 6b shows the probability paper of the minimum annual discharge with a lognormal PDF.
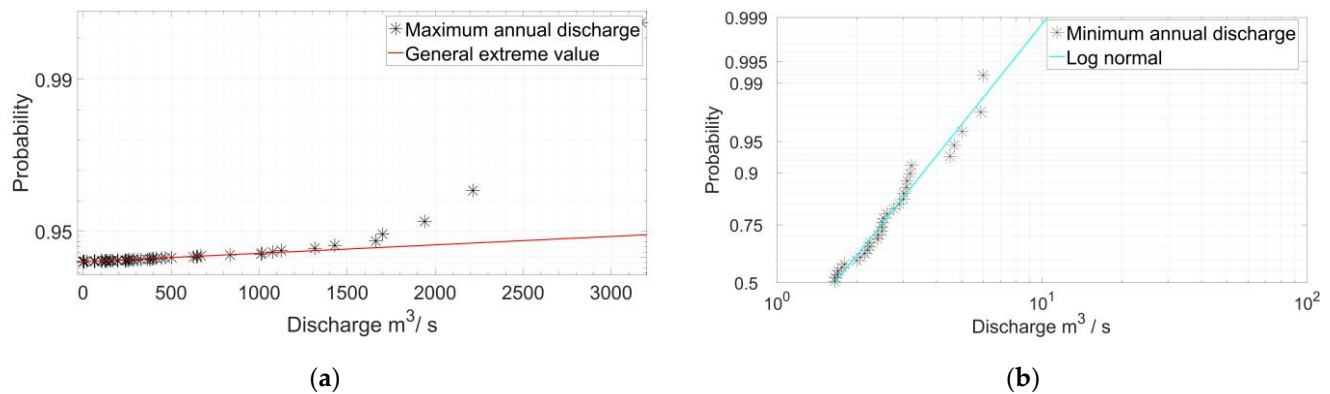
Figure 6. Probability papers for (**a**) maximum annual discharge; (**b**) minimum annual discharge.

Once the probability distribution for the intensity of events has been found, the mean number of events at a specific time interval is found, as described in Section 2.5, and then the joint probability distribution (JPD) is used with the simulations, as shown in Figure 7.
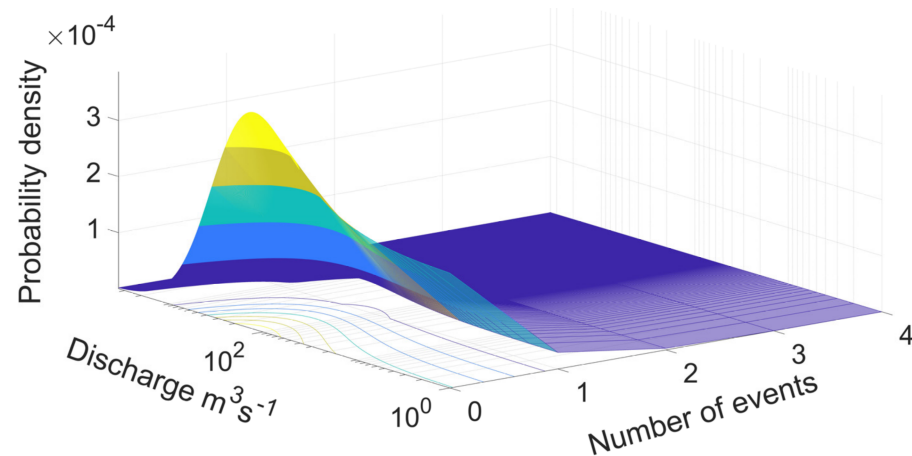
**Figure 7.** Joint PDF of maximum annual flow and number of events per year.

Table 1 shows all the constants and variables needed to carry out the simulations of scour and fill events; $k_1, k_2 = 1$ for the scour process.

Table 1. Variables and their distributions.

Variable	Description	Distribution	σ	μ	ξ	λ
ω_q	Maximum annual discharge	GEV/Poisson	230	209	0.512	$\cong 0.409$
q_t	Minimum annual discharge	Log normal/Poisson	1.74	1.73	-	$\cong 0.818$
d_{50}	Characteristic particle size (mm)		8	4		
k_3	Bed condition correction factor	Rand (Normal)	1.15			
d_a	Soil density kg/m^3	Normal	2000	250		

5.3. Simulation of Events

Monte Carlo simulations were carried out to generate random events that could trigger the scour and fill process. In this context, the necessary number of simulations to have a probability $\alpha \equiv 0.99$ providing a percentage error $\varepsilon \leq 5\,\%$, is 1060, which is the number of simulations carried out to solve the stochastic processes for each simulated time interval. Figure 8a,b shows the outcome of the simulations for a time interval equal to 170 years for the extreme and recurring events. The discharge extreme events are several orders of magnitude bigger than the recurring events. The simulations of occurrences and intensity are used for the estimation of scour and fill in a discrete time model.

(**a**)

(**b**)
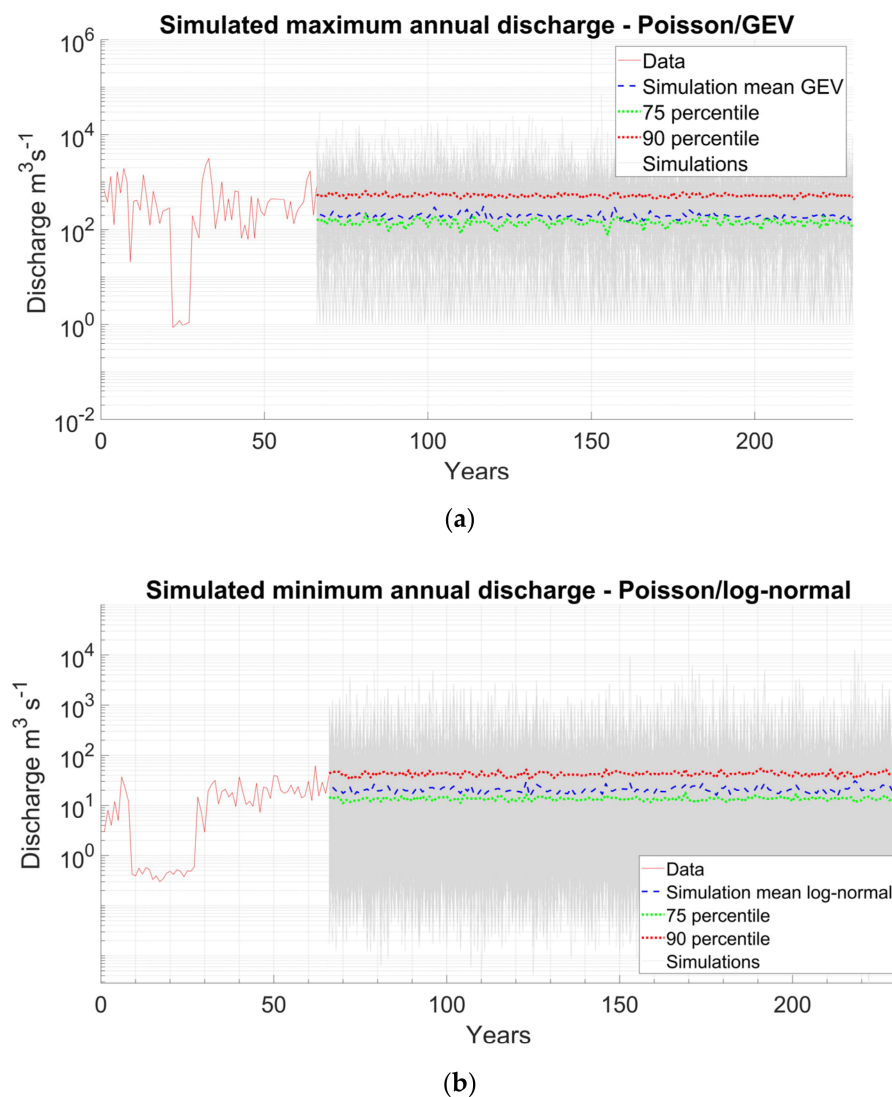
Figure 8. Simulation outcome of occurrences and intensity in accordance with both phenomena and its occurrence rates λ: (**a**) maximum annual streamflow; (**b**) minimum annual streamflow.

For comparison's sake, a set of 1024 simulations of 175 years of daily mean discharge is investigated using an ARIMA model as described in Section 2.6. The mean daily discharge data used are shown in Figure 9a. Figure 9b shows the outcome of the simulations. It can be noted that Figure 9a has outliers that have a negative impact on the forecasting precision. Since the data sets used for time series are generally much larger than that used to forecast extreme events (because of the BMA approach), outliers are more common and hard to deal with. This is a limitation on the use of time series such as ARIMA. In addition, when the data series is large and complex, advanced forecasting techniques such as wavelet multiresolution analysis should be used to improve accuracy.

5.4. Scour Survival Function

Considering that the bridges in Mexico are designed for a service life of 100 years, several simulated histories were developed for time intervals of 20, 30, 40, 60, 75, 90, 110, 130, 150, and 170 years, in order to track the degradation of thickness of the sediment layer. Figure 10a shows one simulation of the thickness of the sediment layer $y_{s(t)}$ in the river bed as a function of time; positive values denote scour, whereas negative values denote fill. Figure 10a shows the result of a single simulation for $t = 170$ years (simulation starts at the end of the recorded data year 64). Scouring events and fill events are added in the

simulation. Since scour is controlled by extreme events, its magnitude is bigger than fill. Figure 10b shows the outcome of 1064 simulations.

(**a**)   (**b**)
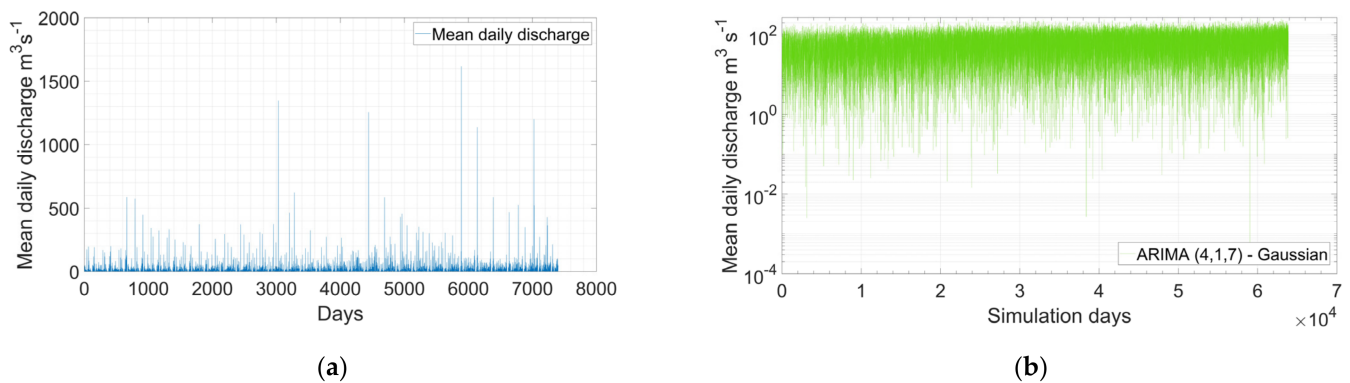
Figure 9. Mean daily discharge: (**a**) available readings of mean daily discharge at the bridge site; (**b**) 175 years of mean daily discharge simulation using ARIMA model with p:5, d:1, q:7 and a Gaussian innovation distribution.


(**a**)

(**b**)
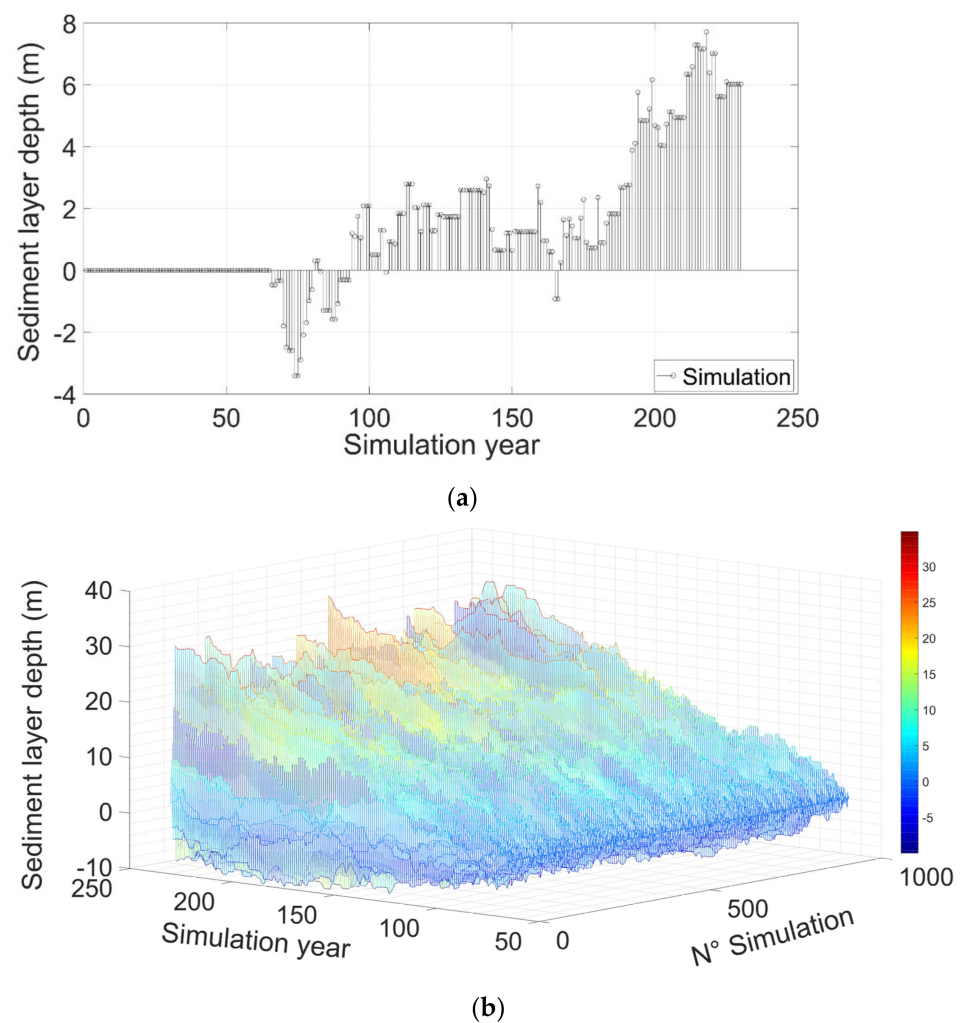
Figure 10. Simulation outcome for scour–fill depth over time (simulation starts at year 64): (**a**) one 170-year simulation; (**b**) 1064 simulations of 170 years.

The maximum scour depth value is obtained from each one of the 1064 simulations, and the best fit distribution is obtained as per 2.4 of this study. Once the probability function is found, the survival function is obtained as follows:

$$P[y_s(te) > y \mid t] = 1 - \int_{-\infty}^{t} exp\left[-\left(1 + \xi\frac{t-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right] dt \tag{36}$$

Figure 11a shows the survivor function of the GEV function for different time intervals, and Figure 11b shows the PDF of the maximum scour depth.
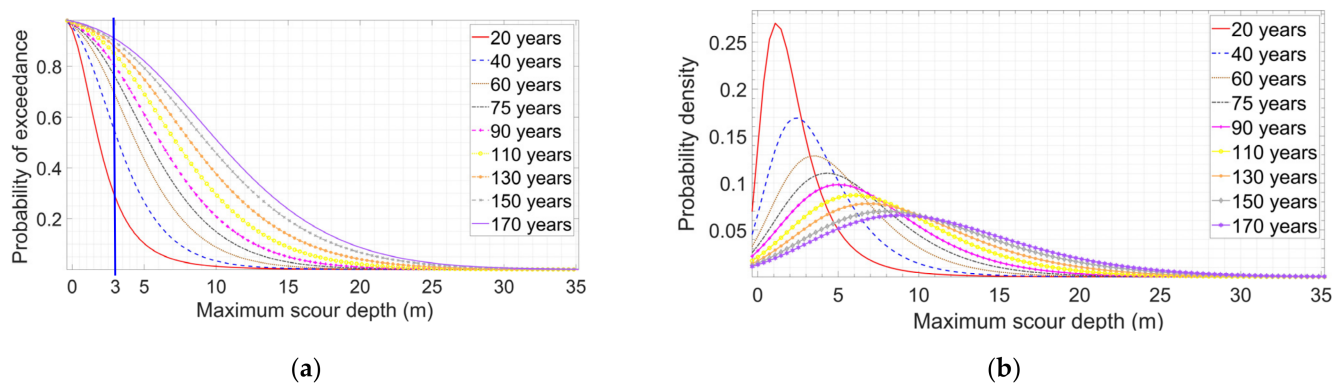
Figure 11. Outcome of the simulations for different time intervals of maximum scour depth: (**a**) GEV survivor function, 100-year return period scour depth is shown as a blue line; (**b**) GEV PDF.

Results show that, as expected, scour accumulates through time. Figure 11b shows the PDFs of several time intervals of the simulation; the distribution tends to zero skewness as time increases, and it has a negative skewness and positive kurtosis at lower time intervals. A scour depth of 2.67 m was computed using the HEC 18 methodology assuming a flood with a return period of 100 years, and it is used for comparison. Such a threshold has a notable probability of exceedance, as shown in Figure 11a. Multiple events of scour with low return periods can accumulate in time, and they can be larger than the design event when the bridge is not repaired. As time tends to infinity, the maximum scour depth increases because, with the approach proposed, scour depth is not limited by harder soils that are naturally found in deeper layers.

Figure 12a shows the scour survivor function of several time intervals using the SRICOS-EFA methodology [5]. Figure 12b shows a comparison between the proposed approach that takes in to account an estimation of sediment deposit, which lessens the severity of scour accumulation in time, and the SRICOS-EFA methodology.
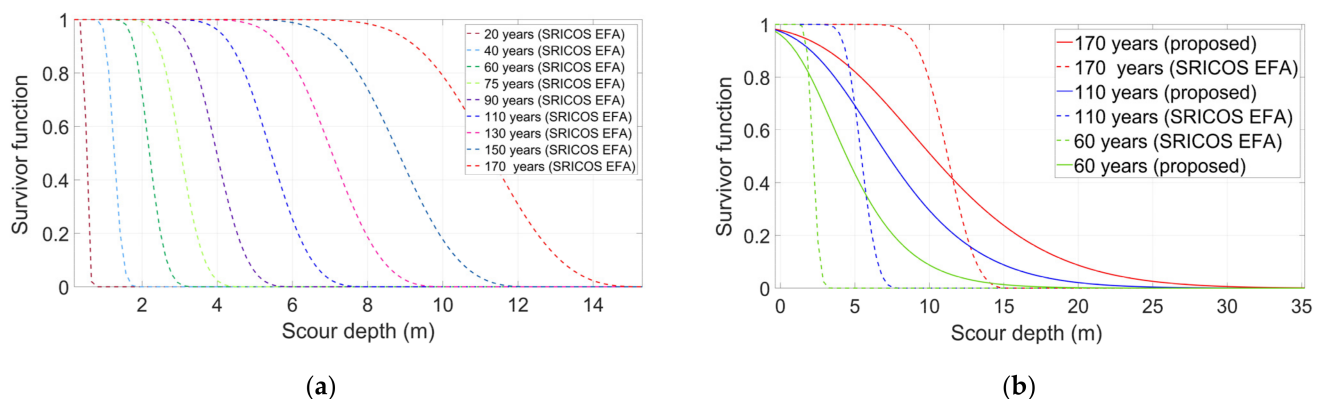
Figure 12. Simulations for different time intervals of maximum scour depth: (**a**) SRICOS EFA survivor function; (**b**) comparison between the proposed methodology and SRICOS EFA [5].

5.5. Discussion

The difference in results shown in Figure 12b are due to the differences in event simulation and in the methodology used to estimate the scour depths. Some notable differences are as follows: (**a**) SRICOS-EFA does not take in to account every event for the scour accumulation (flood has to be larger than flood to produce scour), and in a similar way a shear threshold is needed to produce scour. (**b**) It relies on average daily discharge produced by an ARIMA model, which has limitations in dealing with outliers. (**c**) It does not take in to account the accretion process (thus the sharp fall in the survivor function). (**d**) The SRICOS-EFA method uses equations obtained from flume test conducted at Texas A&M University [5], which are not equal to the HEC 18 equations for scour depth, and this leads to a natural spread on the comparison.

The proposed model has the following key aspects that differ from the SRICOS-EFA: (**a**) larger scour depths in the proposed methodology are likely due to the BMA approach on the maximum annual flow, and this approach leads to higher discharge and, thus, higher scour depths; (**b**) the addition of the accretion process leads to a heavy tail in the GEV distribution and a smoother fall in the survivor function; (**c**) in the proposed approach the hydraulic radius changes in every time step (unlike the SRICOS-EFA); (**d**)every discharge causes scour since the HEC 18 methodology does not establish a shear threshold for the scour development.

5.6. Scour Hazard

Scour hazard curves provide annual exceedance probabilities of scouring events having different intensities. Since scour is time-dependent, several time interval curves are needed because such curves are developed using all the available data of the events that could trigger the scour or fill. They reflect the behavior of the phenomena in a time-dependent fashion, and they can be used to evaluate both new and deteriorating bridges (Figure 13). The estimates of maximum scour are derived from simplified 1D equations. Such estimates are not exact and do not predict the future, but they provide a reasonable approximation to it.
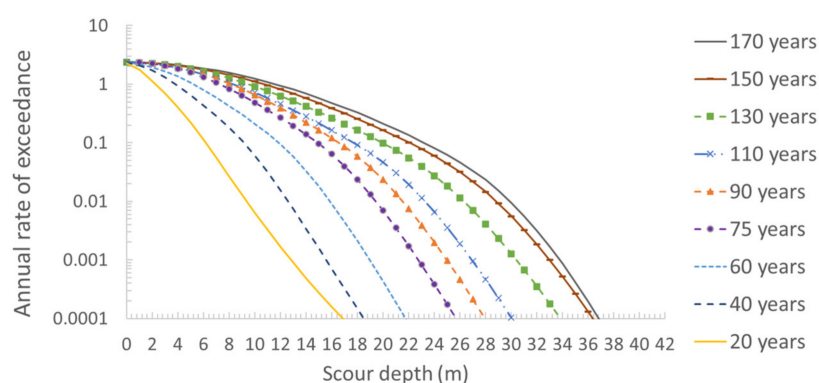
Figure 13. Scour hazard curves.

Figure 13 shows different exceedance rates due to scour or scour hazard curves for different time intervals. It is noticed that the scour depth increases as the time interval increases, which means that a certain scour depth takes different values for each time interval. Scour hazard curves may be improved if the next items are included in the model. First of all, a multi-layer approach is needed to improve the estimation of the scour process, and either a nonstationary stochastic process or an autoregressive integrated moving average with additive outliers in the ARIMA-AO model is needed to calculate the discharge forecast. In addition, it is important to consider climate change, which could impact the parameters related to discharge. It is also important to verify the computed scour depth with the actual bridge scour.

6. Concluding Remarks

A time-dependent scour hazard model that took into account missing data and the interaction between erosion and accretion was developed. The proposed model is straightforward, can be easily applied, and can be useful for the design and the re-design of new bridges prone to scour deterioration. In the case of existing bridge structures, the model can be used to estimate the time instant in which the structure could present an undesirable performance level caused by scour effects.

The proposed stationary stochastic model developed to study the scour and fill process shows that a single event with a high return period, such as $T_r = 100$ years, might not be adequate to represent the maximum scour that could be developed throughout the return period. Bridges subjected to continuous events of scour and fill without maintenance can acquire scour depths greater than the designed scour depth. This phenomenon makes bridges more vulnerable and partially explains why scour is the leading cause of bridge collapses in Mexico.

Scour hazard curves for different time intervals were estimated based on a reinforced concrete bridge located in one of the zones with high values of river discharge, such as Oaxaca. The scour hazard curves provide both an idea of an expected exceedance of scour associated with different time intervals and represent an important tool to estimate the reliability due to scour. Moreover, in zones with a high frequency of seismic occurrences, the scour hazard and seismic hazard can be treated from a multi-hazard point of view to estimate reliability indicators such as fragility curves, exceedance demand rates, mean annual rate of failure, or confidence factor. The above indicators lead engineers and decision-makers to re-design or repair the structural system in the case of new or existing bridges, respectively. The approach used to preprocess, characterize, and simulate events can be used for any phenomenon that has an "extreme" nature. In addition, the inclusion of missing data in the model improves the scour hazard estimation.

References

1. Pizarro, A.; Manfreda, S.; Tubaldi, E. The science behind scour at bridge foundations: A review. *Water* **2020**, *12*, 374. [CrossRef]
2. Mahalder, B.; Schwartz, J.S.; Palomino, A.M.; Zirkle, J. Scour hole development in natural cohesive bed sediment around cylinder shaped piers subjected to varying sequential flow events. *Water* **2021**, *13*, 3289. [CrossRef]
3. Manfreda, S.; Link, O.; Pizarro, A. A theoretically derived probability distribution of scour. *Water* **2018**, *10*, 1520. [CrossRef]
4. Van Noortwijk, J.; Kok, M.; Cooke, R. Optimal maintenance decisions for the sea-bed protection of the Eastern-Scheldt barrier. In *Engineering Probabilistic Design and Maintenance for Flood Protection*; Cooke, R., Mendel, M., Vrijling, H., Eds.; Springer: Boston, MA, USA, 1997; pp. 25–56. [CrossRef]
5. Brandimarte, L.; D'Odorico, P.; Montanari, A. A probabilistic approach to the analysis of contraction scour. *J. Hydraul. Res.* **2006**, *44*, 654–662. [CrossRef]

6. Tubaldi, E.; Macorini, L.; Izzuddin, B.A.; Manes, C.; Laio, F. A framework for probabilistic assessment of clear-water scour around bridge piers. *Struct. Saf.* **2017**, *69*, 11–22. [CrossRef]
7. Pizarro, A.; Tubaldi, E. Quantification of modelling uncertainties in bridge scour risk assessment under multiple flood events. *Geosciences* **2019**, *9*, 445. [CrossRef]
8. Briaud, J.-L.; Brandimarte, L.; Wang, J.; D'Odorico, P. Probability of scour depth exceedance owing to hydrologic uncertainty. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2007**, *1*, 77–88. [CrossRef]
9. Liao, K.-W.; Hoang, N.-D.; Gitomarsono, J. A probabilistic safety evaluation framework for multi-hazard assessment in a bridge using SO-MARS learning model. *KSCE J. Civ. Eng.* **2018**, *22*, 903–915. [CrossRef]
10. Johnson, P.A.; Dock, D.A. Probabilistic bridge scour estimates. *J. Hydraul. Eng.* **1998**, *124*, 750–754. [CrossRef]
11. Kallias, A.N.; Imam, B. Probabilistic assessment of local scour in bridge piers under changing environmental conditions. *Struct. Infrastruct. Eng.* **2016**, *12*, 1228–1241. [CrossRef]
12. Bolduc, L.C.; Gardoni, P.; Briaud, J.-L. Probability of exceedance estimates for scour depth around bridge piers. *J. Geotech. Geoenviron. Eng.* **2008**, *134*, 175–184. [CrossRef]
13. Contreras-Jara, M.; Echaveguren, T.; Chamorro, A.; Vargas-Baecheler, J. Estimation of exceedance probability of scour on bridges using reliability principles. *J. Hydrol. Eng.* **2021**, *26*, 04021029. [CrossRef]
14. Gómez, R.; Flores, D.; Arenas, M.; Flores, R. *Vulnerabilidad de Estructuras de Puentes en Zonas de Gran Influencia de Ciclones Tropicales*; Technical report; CENAPRED: Mexico City, MC, Mexico, 2017. (in Spanish)
15. Wardhana, K.; Hadipriono, F.C. Analysis of recent bridge failures in the United States. *J. Perform. Constr. Facil.* **2003**, *17*, 144–150. [CrossRef]
16. Rubin, D. Inference and missing data. *Biometrika* **1976**, *63*, 581–592. [CrossRef]
17. Pawitan, Y. *All Likelihood: Statistical Modelling and Inference Using Likelihood*; Oxford University Press Inc.: New York, NY, USA, 2001; pp. 336–351.
18. Schafer, J.L.; Graham, J.W. Missing data: Our view of the state of the art. *Psychol. Methods* **2002**, *7*, 147–177. [CrossRef] [PubMed]
19. Arneson, L.A.; Zevenbergen, L.W.; Lagasse, P.F.; Clopper, P.E. *Evaluating Scour at Bridges*, 5th ed.; Publication No. FHWA-HIF-12-003; FHWA: Springfield, VA, USA, 2012.
20. Folch-Fortuny, A.; Arteaga, F.; Ferrer, A. Missing data imputation toolbox for MATLAB. *Chemometr. Intell. Lab. Syst.* **2016**, *154*, 93–100. [CrossRef]
21. Tan, M.T.; Tian, G.-L.; Ng, K.W. *Bayesian Missing Data Problems: EM, Data Augmentation and Noniterative Computation*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2009; pp. 39–42.
22. Young, G.A.; Smith, R.L. *Essentials of Statistical Inference*; Cambridge University Press: Cambridge, UK, 2005; pp. 13–15.
23. Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*; Wiley & Sons. Inc.: New York, NY, USA, 1987; p. 258.
24. Takahashi, M. Statistical Inference in Missing Data by Mcmc and Non-Mcmc Multiple Imputation Algorithms: Assessing the Effects of between-Imputation Iterations. *Data Sci. J.* **2017**, *16*, 1–17. [CrossRef]
25. Schafer, J.L. *Analysis of Incomplete Multivariate Data*; Chapman and Hall/CRC: New York, NY, USA, 1997; pp. 55–59.
26. Akaike, H. Likelihood of a model and information criteria. *J. Econom.* **1981**, *16*, 3–14. [CrossRef]
27. Katz, R.W.; Parlange, M.B.; Naveau, P. Statistics of extremes in hydrology. *Adv. Water Resour.* **2002**, *25*, 1287–1304. [CrossRef]
28. Ferreira, A.; De Haan, L. On the block maxima method in extreme value theory: PWM estimators. *Ann. Stat.* **2015**, *43*, 276–298. [CrossRef]
29. Haschenburger, J.K. Scour and Fill in a Gravel-Bed Channel: Observations and Stochastic Models. PhD Thesis, University of British Columbia, Vancouver, BC, Canada, 1996.
30. Haschenburger, J.K. Channel scour and fill in coastal streams. In *Carnation Creek and Queen Charlotte Islands Fish/Forestry Workshop: Applying 20 Years of Coastal Research to Management Solutions*; BC Ministry of forest; Land management handbook No. 41; Hogan, D.L., Tschaplininski, P.J., Chatwin, S., Eds.; Crown Publications Inc.: Victoria, BC, CA, 1998; pp. 109–117.
31. Bigelow, P.E. Scour, fill, and salmon spawning in a California coastal stream. Master Thesis, Humboldt State University, Arcata, CA, USA, 2003.
32. Einstein, H.A. *The Bed-Load Function for Sediment Transportation in Open Channel Flows*; Technical report No. 1026; US Government Printing Office: Washington, DC, USA, 1950.
33. Todorovic, P.; Zelenhasic, E. A stochastic model for flood analysis. *Water Resour. Res.* **1970**, *6*, 1641–1648. [CrossRef]
34. Sirangelo, B.; Versace, P. Flood-induced bed changes in alluvial streams. *Hydrol. Sci. J.* **1984**, *29*, 389–398. [CrossRef]
35. Haschenburger, J.K. A probability model of scour and fill depths in gravel-bed channels. *Water Resour. Res.* **1999**, *35*, 2857–2869. [CrossRef]
36. Simões, F.J.M. Shear velocity criterion for incipient motion of sediment. *Water Sci. Eng.* **2014**, *7*, 183–193. [CrossRef]
37. Alamilla, J.L.; Campos, D.; Ortega, C.; Soriano, A.; Morales, J.L. Optimum selection of design parameters for transportation of offshore structures. *Ocean Eng.* **2009**, *36*, 330–338. [CrossRef]