Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR030559

Failure Probability Analysis of Levees Affected by Mammal Bioerosion

Matteo Balistrocchi[1], Giovanni Moretti[1], Roberto Ranzi[2], and Stefano Orlandini[1]

[1]Department of Engineering Enzo Ferrari, University of Modena and Reggio Emilia, Modena, Italy, [2]Department of Civil Environmental Architectural Engineering and Mathematics, University of Brescia, Brescia, Italy

Key Points:

- Fully bivariate analysis of peak flow discharge and flood duration is used to describe the hydrologic forcing to levees
- Vorogushyn et al.'s unsteady seepage flow model is extended to derive the failure probability of variably disturbed levees
- Return period of levee failure due to excessive seepage reduces from 100 to 9 years when the mammal den extends for 84% of levee thickness

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. Balistrocchi,
matteo.balistrocchi@unimore.it

Citation:

Balistrocchi, M., Moretti, G., Ranzi, R., & Orlandini, S. (2021). Failure probability analysis of levees affected by mammal bioerosion. *Water Resources Research*, *57*, e2021WR030559. https://doi.org/10.1029/2021WR030559

Received 7 JUN 2021
Accepted 9 NOV 2021

Author Contributions:

Data curation: Matteo Balistrocchi, Giovanni Moretti, Roberto Ranzi, Stefano Orlandini
Funding acquisition: Stefano Orlandini
Investigation: Matteo Balistrocchi, Giovanni Moretti, Stefano Orlandini
Methodology: Matteo Balistrocchi, Giovanni Moretti, Roberto Ranzi, Stefano Orlandini
Software: Matteo Balistrocchi, Giovanni Moretti
Supervision: Roberto Ranzi, Stefano Orlandini
Visualization: Matteo Balistrocchi, Giovanni Moretti

Abstract Mammal bioerosion is an emergent threat to the functionality of levees. In the present paper, the problem of assessing the failure probability of levees affected by mammal bioerosion is addressed. A fully bivariate description of peak flow discharge and flood duration is combined with a deterministic unsteady seepage flow model to obtain a suitable model of variably disturbed levee response to the observed natural variability of floods. Monte Carlo analysis is also implemented to evaluate the epistemic uncertainty connected to the description of the river system. The obtained model is tested with respect to a real-world levee located along the Secchia River in northern Italy, which underwent a disastrous failure caused by mammal bioerosion in 2014. The convex linear combination of two Archimedean copulas is found to fit the empirical dependence structure between peak flow discharge and flood duration. The reliability of the unsteady seepage flow model is tested against detailed numerical simulations of the seepage occurring through the levee body. A limit state function is obtained by comparing the maximum extent of the seepage front to the distance between the den end and the riverside levee slope, and the corresponding levee safety and failure regions are delimited. Results obtained from the developed model reveal a significant impact of mammal dens located near the levee crest in terms of failure probability and related return period. This impact is consistent with failures observed in the study area.

Plain Language Summary Burrowing mammals often find in levees a suitable habitat. Unfortunately, mammal dens can significantly compromise the functionality of levees by creating preferential flow paths for flood water seeping through the levee bodies, and by causing ultimately levee failures due to excessive seepage and internal erosion. In fact, many levee failures have been connected to the levee weakening caused by mammal dens. Mammal bioerosion significantly increases the failure probability of levees and the related flood risk in densely populated floodplains. Estimating the failure probability of levees affected by mammal bioerosion is therefore a relevant societal need. In the present study levee, safety and failure conditions are estimated by combining a fully bivariate statistical description of peak flow discharge and flood duration with a computationally efficient unsteady seepage flow model. The resulting modeling framework incorporates the natural variability of floods and the essential hydraulic properties of disturbed/undisturbed levees. Model results reveal that the return period of levee failure due to excessive seepage reduces from 100 to 9 years, namely of −91%, when the mammal den extends for 84% of levee thickness. These results can be used to inform levee design and maintenance programs for the safety of societies living in floodplains worldwide.

1. Introduction

Bioerosion is a global and ubiquitous ecological process transforming habitats, increasing biodiversity and shaping landscapes (Davidson et al., 2018). Unfortunately, it can also have strong adverse impacts on human assets, from the economic, social and ecological points of view (Bayoumi & Meguid, 2011; Orlandini et al., 2015). The interaction between human infrastructures and wildlife has been known by engineers for centuries and is presently becoming an increasingly important research topic (Harvey et al., 2019; Haubrock et al., 2019). Special attention is focused on burrowing mammals, such as the crested porcupine (Hystrix cristata), the European badger (Meles meles), the red fox (Vulpes vulpes) and the nutria (Myocastor coypus), affecting levees devoted to flood control. These mammals burrow levees to construct their dens, thus obtaining shelter from predators and thermal extremes, and a suitable environment for reproduction. However, mammal dens compromise both the structural integrity and the hydraulic properties of levees by ultimately affecting their functionality. Among levee failure mechanisms, seepage and internal erosion induced by dens have recently attracted much attention, since

Writing – original draft: Matteo Balistrocchi, Giovanni Moretti, Stefano Orlandini
Writing – review & editing: Matteo Balistrocchi, Roberto Ranzi, Stefano Orlandini

these processes have been found to explain an increasing number of levee breaches causing disastrous inundations (Orlandini et al., 2015).

The awareness of researchers toward the issues that mammal dens in levees pose to flood prone communities is demonstrated by the rapidly growing number of papers dedicated to this topic (Calamak et al., 2020; Dassanayake & Mousa, 2020; Li et al., 2020; Palladino et al., 2020). In spite of the problem relevance, burrowing of earthen structures was previously discussed only by a limited number of publications such as failure reports, maintenance recommendations and restoration guidelines of authorities involved in flood control or newspaper articles. A broad review of this gray literature was provided by Bayoumi and Meguid (2011), who first raised the problem within the scientific community. On a global scale, 22 animal species were acknowledged to have a detrimental impact on the structural integrity and on the vulnerability to seepage of earthen dams and levees, including human vandals (Federal Emergency Management Agency, 2005; Woodward & Mayfield, 1999). Nevertheless, Bayoumi and Meguid (2011) underlined the disproportionately low attention given to burrowing with respect to its socio-economic impact.

The development of the research interest toward burrowing can also be justified by the increase in the invasive activity of animals against earthen structures, that has recently been noticed. Several reasons can be advocated to explain this evidence: (a) the establishment of nature reserves along river courses and the promulgation of laws protecting endangered species (Council of Europe, 1979, 1992), (b) some species take advantage from habitat fragmentation due to the urbanization sprawl and from deforestation (Alexandre et al., 2020; McMahon et al., 2019), (c) in the past decades the release of some allochthonous species in habitats where they did not find natural competitors or predators yielded their widespread proliferation (Carter & Leonard, 2002; Harvey et al., 2019), (*iv*) the decrease in the frequency of riverbanks maintenance practices during the last two decades favors mammal denning (Mori & Assandri, 2019).

The den structure depends on several factors, including the kind of species, the community size, the den age and the soil characteristics. Although strongly variable on a global scale, dens of semi-fossorial mammals are basically relevant to levees along the secondary river network where levee thicknesses are limited. For instance, nutria dens can extent 5 m in depth (Harvey et al., 2019). An increase in the overall failure probability of levees is therefore expected since these mammal species usually burrow their dens between the toe and the crest. During floods, the den tunnels can be directly inundated, if the entrance is placed on the riverside slope. Conversely, if the entrance is placed on the landside slope, the den tunnels provide a pipe network for the seepage waters to be conveyed, when the saturated front reaches the den end. In both situations, the actual soil thickness that the saturated front of seepage must cross to trigger the internal soil erosion can be significantly shorter than the design levee thickness. As the fine fraction of the soil is progressively removed by hydrodynamic forces, the seepage process auto-increases, until a pressurized pipe flow is established between the riverside slope and the landside slope. Then, soil removal continues due to shear stress erosion, and the pipe cross section enlarges. Finally, the levee portion located above the den subsides, yielding the levee overflow and then a breach (Orlandini et al., 2015).

When dealing with an extensive levee system affected by mammal bioerosion, it is important to develop practical methodologies for assessing the increase in levee failure probability, by using simple geometric characteristics and soil hydraulic properties. First, estimating levees vulnerability to bioerosion would allow authorities devoted to levee maintenance to better address survey campaigns. Second, classifying surveyed dens in terms of hazard would improve the effectiveness of restoring programs. Levee failure mechanisms induced by excessive seepage, as well as other levee failure mechanisms related to overflowing, internal erosion and mechanical instabilities, depends on both riverflow stages and their duration. Bearing in mind the natural variability of floods, this delineates a multivariate problem in which at least two hydrograph characteristics must be accounted for. In addition, the geometrical characterization of dens is affected by high uncertainty, related to the difficulties involved in the indirect techniques used for their surveys (Borgatti et al., 2017). Thus, accounting for the epistemic uncertainty is essential. At the state-of-the-art, two approaches are available to evaluate the response of levees to the seepage of flood waters: (a) numerical simulations and (b) fragility curves.

Numerical simulations are capable to comprehensively represent the seepage processes by solving the 3D Richards equation in variably saturated soil conditions (e.g., Butera et al., 2020; Calamak et al., 2020; Orlandini et al., 2015). However, they are computationally intensive and require a large amount of data that are not always accessible with the appropriate accuracy. Therefore, their practical application is normally limited to the

simulation of a few individual floods, which can be observed events or synthetic design events, often derived from flood reduction curves (Bacchi et al., 1992). In both cases, accounting for the epistemic uncertainty, especially that related to the den geometry, is a nontrivial task. A stochastic simulation, based on the probabilistic description of hydroclimatic forcing, a simple deterministic modeling of levee response and the statistical description of uncertain parameters, is therefore a useful alternative. Fragility curves use schematic models or analytical solutions derived for simple cases (e.g., Marchi, 1961), that can be implemented in stochastic generation procedures to account for the epistemic uncertainty (e.g., ; Camici et al., 2017; Hall et al., 2003, 2005; Vorogushyn et al., 2009). However, fragility curves express the failure probability variability for given flood durations and must be regarded as conditioned distribution functions. By using fixed durations of floods, the actual dependence structure of flood characteristics is lost (Balistrocchi et al., 2017). In addition, the peak flow discharge is often assumed to be constant over the chosen duration. Albeit conservative, this hypothesis is far from being realistic, so that dealing with a real-world flood, it must be relaxed by approximating the flood hydrograph through a step function.

As first suggested by Balistrocchi et al. (2019) for the levee failure due to the overflowing, levee failure probability can be estimated in a conceptually correct manner by combining a fully bivariate statistical description of the hydrologic forcing to the levee with a simplified schematization of the failure mechanism requiring a reasonable computational burden. To this aim, bivariate distributions or copulas have largely been used in literature (De Michele et al., 2005; Goel et al., 1998; Grimaldi & Serinaldi, 2006; Salvadori et al., 2007; Yue, 2000). Recent advances have been achieved in the estimation of the frequency of occurrence of multivariate events (Balistrocchi & Bacchi, 2017; Balistrocchi et al., 2019; Balistrocchi & Grossi, 2020; Requena et al., 2013; Salvadori et al., 2015; Serinaldi, 2015, 2016; Volpi & Fiori, 2014). If a functional relationship exists between a multiple random variable representing the hydrological forcing and a univariate random variable representing the hydrologic load to the levee, the derived distribution theory implying the equality of exceedance probabilities of corresponding events can be used (e.g., Kunstmann & Kastens, 2006). If the transformation function represents the failure mechanism, it can be implemented in a limit state function that discriminates the hydrologic events in two dichotomous classes: (a) those that trigger the failure mechanism and (b) those that do not. Hence, the multivariate population of the hydrologic event variables can be split into a safety region and a failure region, making it possible to estimate the failure probability. This modeling strategy has never been used to investigate the levee failure due to excessive seepage and, more specifically, to determine the impact of burrowing mammals on levee failure probability.

In the present paper, a new stochastic model is specifically developed for estimating the increase in failure probability of levees due to mammal bioerosion. This model combines a fully bivariate statistical description of peak flow discharge and flood duration, a deterministic unsteady seepage flow model, and a Monte Carlo analysis of epistemic uncertainty in the characterization of the river system. Although the developed model is of general interest as it may be used for any disturbed and undisturbed levees, it is illustrated on the real case of the Secchia River in Italy, where bioerosion caused a remarkable flood damage in 2014. The paper is organized as follows. Methods are described in Section 2. The case study, the available data set, the calibration and the verification of model components are illustrated in Section 3. The discussion on results is reported in Section 4. Conclusions are drawn in Section 5.

2. Methods

2.1. Bivariate Modeling of Flood Events

The natural variability of flood events is modeled by using a joint distribution function (JDF) of peak flow discharge q_p and flood duration d. Peak flow discharge is related to the maximum riverflow stage through the uniform flow rating curve. River flow stage can directly be used in preference to flow discharge and uniform flow rating curve when these data are available. In addition, the combined use of peak flow discharge and flood duration is informative about the flood volume. Copulas are used to derive such a bivariate distribution (Joe, 1997, 2015; Nelsen, 2006). According to this approach, the JDF F_{Q_pD} expressing the overall mutual dependence of the marginal variables can be decomposed as given by

$$F_{Q_pD}\left(q_p, d\right) = H_\rho\left[F_{Q_p}\left(q_p\right), F_D\left(d\right)\right] \tag{1}$$

where F_{Q_p} is the cumulative distribution function (CDF) of q_p, F_D is the CDF of d and $H_{\boldsymbol{\rho}}$ is the underlying copula function, completely defined by the parameter vector $\boldsymbol{\rho}$ (Sklar, 1959). The bivariate function $H_{\boldsymbol{\rho}}$ varies in the domain $[0,1]^2$ of the marginal uniform variables u and v, defined as the non-exceedance probabilities of the natural marginal variables, as given by

$$H_{\boldsymbol{\rho}}(u,v):[0,1]^2 \rightarrow [0,1] \text{ with } u = F_{Q_p}(q_p) \text{ and } v = F_D(d). \tag{2}$$

Consequently, $H_{\boldsymbol{\rho}}$ is independent of the marginal distributions and it only expresses the dependence structure. To fit the JDF given by 1, individual independent events must be sampled from an extended time series of observed riverflow discharges. To this aim, a peak-over-threshold (POT) criterion based on a threshold discharge q_t and a minimum interevent time t_d is adopted. The first sampling parameter allows us to separate the continuous series into partial duration series, including only those hydrograph portions which exceed q_t (Lang et al., 1999; Todorovic, 1978). These hydrograph portions are identified as individual flood events, separated by the others by interevent periods. The second sampling parameter aims at ensuring the independence between two subsequent flood events (Brunner et al., 2017). Hence, two hydrographs are considered to be independent if they are separated by an interevent period greater than t_d. Otherwise, they are aggregated into a unique flood event, whose partial duration spans from the beginning of the first hydrograph to the end of the second one.

Sampling parameters significantly affects the statistical properties of the flood sample, and thus the overall model reliability. The method suggested by Balistrocchi et al. (2017) provides, however, a sound strategy to select these values. In fact, q_t must yield floods significant to the system behavior, while t_d must be long enough for the system to be restored to the initial condition, at the beginning of the subsequent flood. Once independent flood events are isolated, the observation sample of peak flow discharge $\hat{q}_{pi}$ and flood duration $\hat{d}_i$ couples are derived along with the average annual number of flood events n_y. According to copula formalism in Equation 1, one can split the inference of a parametric model for F_{Q_pD} into the separate inferences of marginal distributions and copula, enabling for greater flexibility in the choice of these components and making the fitting procedure easier. This allows us to investigate the dependence structure and the marginal distributions in a more straightforward manner, and to construct a JDF that closely suits the observed joint variability.

The model of the theoretical copula $H_{\boldsymbol{\rho}}$ is obtained in the present study by applying a convex linear combination of two copulas as given by

$$H_{\boldsymbol{\rho}}(u,v) = \omega C_{\phi}(u,v) + (1-\omega) G_{\psi}(u,v) \text{ with } \boldsymbol{\rho} = \{\omega,\phi,\psi\} \text{ and } 0 \le \omega \le 1 \tag{3}$$

where C_{ϕ} is a Clayton copula having parameter ϕ, G_{ψ} is a Gumbel-Hougaard copula having parameter ψ, and ω is the convex linear combination coefficient. Both copulas are mono-parametric and Archimedean (Salvadori et al., 2007). The bivariate CDFs of the Clayton and Gumbel-Hougaard copulas are given by

$$C_{\phi}(u,v) = \left[\max\left\{ u^{-\phi} + v^{-\phi} - 1, 0 \right\} \right]^{-1/\phi} \text{ with } \phi \ge -1 \tag{4}$$

and

$$G_{\psi}(u,v) = \exp\left\{ -\left[(-\ln u)^{\psi} + (-\ln v)^{\psi} \right]^{1/\psi} \right\} \text{ with } \psi \ge 1 \tag{5}$$

respectively. The interest for combining these functions mainly arises from the different modeling of the tails. The Clayton copula features a strong concordance in the lower tail, but not in the upper tail. The contrary occurs for the Gumbel-Hougaard copula. Hence, their convex linear combination has tail dependence coefficients given by the same combination of the corresponding coefficients of the single copulas, as given by

$$\lambda_L = \omega 2^{-1/\phi} \text{ with } \phi \ge 0 \tag{6}$$

for the lower tail coefficient λ_L, and by

$$\lambda_U = (1-\omega)\left(2 - 2^{1/\psi}\right) \text{ with } \psi \ge 1 \tag{7}$$

for the upper tail coefficient λ_U. The theoretical copula given by 3 can be fit to data leaving apart the marginal distributions by deriving the pseudo-observations $\left(\hat{u}_i, \hat{v}_i\right)$ from the observations $\left(\hat{q}_{pi}, \hat{d}_i\right)$. Although the most popular method for fitting multi-parameter copulas is based on the maximization of the pseudo-log-likelihood estimator (Genest & Favre, 2007), in the present study the method based on the minimization of the Cramer-von Mises statistics is used to increase computational efficiency (Genest et al., 2009; Salvadori et al., 2007). To complete the JDF given by 1, the CDFs belonging to the generalized Pareto model, given by

$$F_{Q_p}\left(q_p\right) = 1 - \left[1 + \frac{\kappa}{\xi}\left(q_p - q_t\right)\right]^{-\frac{1}{\kappa}} \text{ with } q_p \geq q_t \tag{8}$$

and the Weibull model, given by

$$F_D\left(d\right) = 1 - \exp\left[-\left(\frac{d}{\zeta}\right)^{\delta}\right] \text{with } d \geq 0 \tag{9}$$

are selected to represent the marginal distribution of q_p and d, respectively. In Equation 8, only the shape parameter κ and the scale parameter ξ must be estimated, since the location parameter for the peak flow discharge must be set equal to the threshold flow discharge q_t. In Equation 9, only a shape parameter δ and a scale parameter ζ are needed, as the flood duration lower limit is zero. Owing to the copula approach, such CDFs can be individually fit and tested according to commonly used criteria for univariate distributions (Kottegoda & Rosso, 2008).

2.2. Failure Hazard Estimate

In multivariate models, the estimate of the hazard referred to a planning period, or alternatively of the return period, suffers from the absence of a total order relation of the events. Hence, various approaches have been proposed in order to split the multivariate populations into sub critical regions and super critical regions: "AND," "OR" (Goel et al., 1998), Kendall (Salvadori & De Michele, 2010), and structural (Balistrocchi & Bacchi, 2017; Balistrocchi et al., 2017; Requena et al., 2013; Salvadori et al., 2015; Volpi & Fiori, 2014). Owing to its reliability, the structural approach has recently attracted great interest and appears to be appealing in failure probability analyses of levees (Balistrocchi et al., 2019). Dealing with the estimate of failure probabilities, the multivariate events must be classified in two classes, including those that trigger the failure mechanism and those that do not. Thus, the failure mechanism itself induces a dichotomous partition of the population, that is split into the failure region and the safety region. The boundary separating these regions is referred to as limit state, which is a (k-1)-dimensional space in a k-dimensional population.

In general, the population partition can be obtained by using the limit state function, or reliability function, as given by

$$Z = R - L \tag{10}$$

where R is the structural resistance and L is the external load. The difference Z between R and L is positive in the safety region, equal to zero in the limit state and negative in the failure region (Apel et al., 2004; Vorogushyn et al., 2009). Equation 10 can incorporate the natural variability of floods along with the epistemic uncertainty of the structural characteristics. In our study, the load L is derived from the bivariate JDF given by 1 through a deterministic transformation function estimating the maximum extent of the saturation front inside the levee body. Although the transformation function is deterministic, its parameters that can be randomly varied by using a Monte Carlo simulation technique in order to take into consideration the limited knowledge and the variability of the structural characteristics, as well as the errors due to model assumptions. In undisturbed levees, the resistance R is given by the seepage length needed for the saturation front to reach the landside levee slope. In the disturbed levees addressed in the present study, the resistance R is given by the seepage length needed for the saturation front to reach the den system.

2.2.1. Seepage Modeling

Seepage in variably disturbed levees is sketched in Figure 1. The flood hydrograph forcing the levee is sketched in Figure 1a. A picture showing a crested porcupine at the entrance of its den is reported in Figure 1b. The

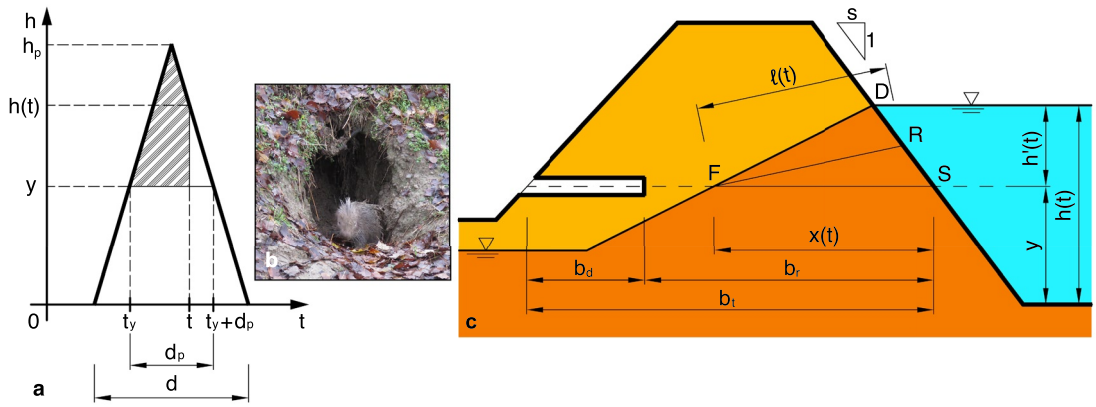

Figure 1. (a) Shape of the flood hydrograph adopted in this study. (b) Crested porcupine photographed on the entrance of its den along a river system. (c) Sketch of the seepage at a given time t in the ponding period starting at time t_y and ending at time $t_y + d_p$ at elevation y.

conceptual scheme proposed by Vorogushyn et al. (2009) for the levee failure mechanism due to excessive seepage is extended as shown in Figure 1c. The saturation front length ℓ at time t is given by

$$\ell(t) = \sqrt{2\frac{K_c}{\theta_a}\int_0^t h'(\tau)d\tau} \tag{11}$$

where K_c is the saturated hydraulic conductivity, θ_a is the air-filled porosity, and $h' = h - y$ is the hydraulic head with respect to the den elevation y. The integral $\int_0^t h'(\tau)d\tau$ is represented by the shaded area in Figure 1a. In order to account for the unsteady flow conditions determined by the gradual rise and recession of the flow stage during a flood, a time dependent water head $h' = h'(t)$ is considered. Equation 11 makes it possible to overcome the usual hypothesis of rectangular shape for the flood hydrograph, which is needed to integrate analytically the filtration equations under unsteady flow conditions (Michelazzo et al., 2018; Palladino et al., 2020). To make this hypothesis conservative, a constant flow discharge equal to the flood peak is set throughout the flood duration. In contrast, by using Equation 11, any shape for the flood hydrograph can be adopted to derive the maximum length of the seepage front. In the present study the triangular hydrograph shape first suggested by Wycoff and Singh (1976) is used (Figure 1a).

The maximum horizontal advance x_m in Figure 1c can be computed by assuming ℓ as the median RF of triangle DSF as given by

$$x_m = \begin{cases} \sqrt{2\frac{K_c}{\theta_a}\left(|\mathrm{SMP}| + \bar{h}\right)d_p - \left(\frac{\bar{h}}{2}\right)^2} + \frac{s\,\bar{h}}{2} & \text{if } 2\frac{K_c}{\theta_a}\left(|\mathrm{SMP}| + \bar{h}\right)d_p - \left(\frac{\bar{h}}{2}\right)^2 > 0 \\ s\sqrt{2\frac{K_c}{\theta_a}\left(|\mathrm{SMP}| + \bar{h}\right)d_p} & \text{elsewhere} \end{cases} \tag{12}$$

where SMP is the soil matric potential, $\bar{h}$ is the average hydraulic head with respect to the den elevation y as is given by

$$\bar{h} = \frac{1}{d_p}\int_{t_y}^{t_y+d_p}\left[\, h(\tau) - y \,\right]d\tau \tag{13}$$

d_p is the flood impoundment duration, t_y is the time at which impoundment at the level y begins, and s is the levee riverside slope as shown in Figure 1c. The gentler the slope s, the larger the infiltration area on the riverside is and the longer the seepage front advance x_m is. The maximum length of the seepage front ℓ_m and the maximum horizontal advance x_m occurring during the flood event are obtained when the integral given by Equation 13 covers the total duration of the flood impoundment d_p. Indeed, these values are related to the maximum possible value of the hydrograph integral in Equation 11. It is remarked that the variable x_m can hardly be observed directly and the statistics for this quantity needs therefore to be derived from observed flood variables.

In Equation 12, the soil matric potential SMP corresponding to the initial levee saturation is also considered as driving force in addition to the gravitational forces. This contribution could be significant in presence of soils with fine granulometry, such as those that are commonly used in the construction of levees. Owing to the nature of the methodology proposed in the present study, the whole population of flood events, including those with short duration and high peak, must be investigated even though their probability of occurrence is low. These events were not considered in the study by Vorogushyn et al. (2009) and they may cause the radical of the first case of Equation 12 to become negative. To deal with this problem, the solution given by the second case of Equation 12 is introduced in the present study. This solution assumes a seepage length ℓ_m equal to the half of the average water stage and allows the continuity of the limit state boundary to be preserved. The cross-sectional river geometry sketched in Figure 1c is consistent with the conceptual model described by Equation 12 and it is similar to the cross-sectional river geometry considered by Vorogushyn et al. (2009) or, more recently, by Palladino et al. (2020). In this conceptual model the gently sloping floodplain located between the levee riverside and the active channel are not considered.

To estimate the hydrograph integral in Equation 13, the triangular shape shown in Figure 1a appears to be a reasonable approximation. Although flood hydrographs are generally flatter on the recession limb, an isosceles triangle hydrograph is considered in Figure 1a as $\bar{h}$ used in Equation 12 does not depend on the slopes of the rising and recession limbs. The base-time directly corresponds to the random variable d, and the maximum water stage h_p can be assessed from the random variable q_p by using an estimated uniform flow or looped rating curve. As illustrated in Figure 1a, the impoundment duration d_p at a level y and the corresponding average water head $\bar{h}$ are easily computed as

$$d_p = d\frac{h_p - y}{h_p} \tag{14}$$

and

$$\bar{h} = \frac{1}{2}\left(h_p - y\right) \tag{15}$$

respectively. According to the derived distribution theory, Equations 12–15 configure as a transformation function, suitable to relate the bivariate space of the flood variables q_p and d to the univariate space of x_m. In addition, this derivation procedure is consistent with the structural based approach for the estimation of the failure hazard or the failure return period, and makes it possible the exploitation of the total order of the univariate random variable x_m to estimate the return period of a bivariate flood event defined by q_p and d.

2.2.2. Mammal Den Modeling

Under undisturbed conditions, it is routinely assumed that the failure mechanism triggers immediately when the saturated front reaches the landside levee toe. Accordingly, in this analysis, a den with an entrance located on the landside slope is idealized as a cylindric horizontal pipe, which extends toward the riverside slope for a length b_d. As illustrated in Figure 1c, given the total thickness of the levee b_t at the elevation y, the residual levee thickness b_r represents the levee thickness at elevation y from the river thalweg, that the saturated front must cross to trigger the failure mechanism. This schematization can also stand for a den with the entrance located on the riverside slope if an instantaneous filling of the den volume is assumed. In the limiting case of a den crossing the levee from the riverside to the landside (b_r equal to zero), the failure mechanism can be considered triggered by any flood whose water stages reach the den. A similar schematization is adopted by Saghaee et al. (2016) and Palladino et al. (2020), to analyze the impact of dens on levee stability by using physical modeling and finite element analysis, respectively.

2.2.3. Hazard Estimate

The seepage failure probability associated with an individual independent flood is defined as the integral of the probability density function of its constitutive variables over the failure region. According to the derived distribution theory, this integral can be computed in the probability space by using the copula density function $\eta = \partial^2 H_{\rho} / \left(\partial u\, \partial v\right)$. Due to the complexity of the copula functions, the integration must be performed by

numerical techniques. However, the semi-infinite regions of the space of natural variables are transformed in definite regions, included in the unitary square $[0,1]^2$, greatly simplifying the integral computation. In this case, the general form of the limit state function in Equation 10 takes the form

$$Z\left(q_p, d\right) = b_r - x_m\left(q_p, d\right) \tag{16}$$

where the levee residual thickness b_r represents the resistance R and the maximum advance of the saturated front x_m the load L. Hence, Equation 16 can explicitly incorporate the mutual variability of flood variables, along with the epistemic uncertainty through the randomization of its parameters according to Monte Carlo techniques. The state variable Z is thus function of the natural random variables q_p and d, which can be expressed in terms of quantiles of the corresponding uniform variables u and v by inverting the marginal CDFs given by 8 and 9, respectively. The failure region Λ in the probability space can thus be delimitated as indicated by

$$\Lambda = \left\{\left(u,v\right)\epsilon\left[0,1\right]^2 \mid Z\left[\ F_{Q_p}^{-1}\left(u\right), F_D^{-1}\left(v\right)\ \right] < 0\right\}. \tag{17}$$

The hazard related to a structural failure in a planning period of N years can finally be computed as

$$H_N = 1 - \left(1 - n_y \int_\Lambda \eta\left(u,v\right) du\, dv\right)^N \tag{18}$$

In the past, this quantity was referred to as long-term risk, or inherent hydrologic risk (US Army Corps of Engineers, 1996). In this formulation, the annual hazard is computed by multiplying the event failure probability for the average annual event number n_y. The hazard H_N is thus the probability that a load L exceeding the resistance R occurs in a planning period of N years (Akan & Houghtalen, 2003; Chow et al., 1988). The return period T_r is therefore related to the hazard H_N as given by

$$T_r = \frac{1}{1 - \left(1 - H_N\right)^{\frac{1}{N}}} \tag{19}$$

2.3. Implementation of the Failure Probability Estimation Model

The stochastic model estimating the failure probability of levees affected by bioerosion is implemented by performing the following steps:

1. The copula function H_ρ given by Equation 3 expressing the natural mutual variability of peak flow discharge and flood duration is fitted to pseudo-observations by minimizing the sum of the squared residuals between the empirical and the theoretical copula, as described in Section 2.1.
2. The soil hydraulic parameters affecting the load factor x_m given by Equation 12 and the parameters of the uniform flow rating curve are randomly generated by using the Monte Carlo procedure described in the following Section 3.2.3.
3. The failure region Λ defined by Equation 17 is delimited in the probability space $[0,1]^2$ for chosen den elevation y and levee residual thickness b_r, through the limit state function Z given by Equation 16, by varying the uniform variables u and v in [0,1] according to a sufficiently small interval:
   3.1. Uniform variables u and v are fixed in [0,1].
   3.2. Uniform variables are transformed in the corresponding natural variables q_p and d through the CDFs given by 8 and 9, respectively.
   3.3. The maximum water stage h_p is obtained from the peak flow discharge q_p by using the uniform flow rating curve, as illustrated in the following Section 3.2.3.
   3.4. The average water head $\bar{h}$ and the ponding duration d_p at elevation y are evaluated by using Equations 14 and 15.
   3.5. The maximum horizontal advance x_m, namely the load factor L in 10, is computed from Equation 12 and subtracted to the residual thickness b_r, namely the resistance factor R in 10, as given by Equation 16, to obtain the limit state function Z introduced in Equation 10.

3.6. The point (u,v) is attributed to failure region Λ if the value of the state function Z computed by using Equation 16 is negative.

4. The hazard H_N is estimated by integrating numerically the copula density function η, derived from the copula H_ρ given by Equation 3, over the failure region Λ as given by Equation 18.
5. Steps from 2 to 4 are repeated for a sufficiently large number of runs.
6. The sample of hazards H_N is statistically described by sample median and quartiles.

It is specified here that this stochastic modeling framework is composed of deterministic and statistical components. The deterministic component is the seepage flow model given by Equation 12. The statistical components are the bivariate distribution 1 of peak flow discharge and flood duration describing the natural variability of floods and the statistical distributions of uncertain parameters in Equation 12 and in the uniform flow rating curve described in Section 3.2.3. The probability of occurrence of the hydrologic load x_m is directly estimated by using JDF 3 and no further estimations are needed. An equivalent result can be achieved by statistically characterizing a sample of x_m occurrences, derived from JDF 3 by means of copula simulation techniques. In theory, the estimate method used in this study should lead to a more precise estimate of the failure probability, since it relies on the whole population of flood events and not on a limited sample of flood occurrences. In the present study, major sources of uncertainty are taken into consideration. Nevertheless, the uncertainties related to probabilistic model estimate could also be accounted for, through the randomization of the parameters of JDF 3. In addition, the different weight associated with the single sources of uncertainty could be quantified through an analysis of variance approach as shown, for instance, by Qi et al. (2016).

Annual levee failure probability can interchangeably be expressed as $1 / T_r$, where T_r is the levee failure return period, or hazard H_N over the planning period N, which is related to T_r by Equation 19. Although the model outlined in the present section can mathematically yield estimates of levee failure return period having unlimited magnitude, the statistical reliability of results must be qualified in regard to the flow record length. Computed events having return periods that do not exceed by a factor of two the flow record length are commonly considered to be statistically reliable (Benson, 1962, 1963; Hu et al., 2020; Natural Environment Research Council, 1975; UK Center for Ecology & Hydrology, 1999). In addition, more recent studies indicate a greater reliability of quantiles obtained from POT series with respect to those obtained from the corresponding annual maximum series (Todorovic, 1978). In the case study reported in the following Section 3, where flow discharge is observed over a period of 42 years and a POT series of 188 flood events is used, all model results are illustrated and events having return period that do not exceed 100 years are assumed to be statistically reliable.

3. Case Study

3.1. Data

The failure probability model herein developed is applied to and tested on an existing levee located in northern Italy along the left bank of the Secchia River (Figure 2). As shown in Figure 2a, the Secchia River is a main right bank tributary of the Po River in the Tuscan-Emilian Apennines originating at Alpe di Succiso at an elevation of about 1,450 m asl. It heads north, touches Modena on the right bank and runs into the Po River south of Mantua. The course is 172 km long and the drainage basin has a total area of 2,292 km^2. The riverflow has been monitored by a river station located in Ponte Bacchello, 73.5 km upstream of the river mouth, since 1923 by the Italian Hydrographic Service (Servizio Idrografico Italiano). As shown in Figure 2, the Ponte Bacchello river station lies on the plain and drains a drainage basin having area of 1,292 km^2 and time of concentration of about 15 hr. The annual rainfall depth averagely amounts to 1,270 mm. The runoff coefficient is estimated to be on average equal to 0.47 and the mean annual discharge is equal to 22.4 m^3 s^{-1}. The riverflow regime is mostly driven by rainfall, which is characterized by two maxima and two minima. The main maximum occurs in November and the main minimum in July. Nevertheless, snowmelt supplies an appreciable contribution to the riverflow in March, when the average monthly runoff coefficient is close to one. Owing to high temperatures, scarce rainfalls and the absence of glacier contributions, low riverflow is frequently observed in August. Conversely, the most severe floods usually occur in November, after the drainage basin moisture is restored by the stratiform rainfall events.

The river reach on the plain is bounded by an extensive levee system that underwent a severe failure yielding a disastrous inundation of a densely populated floodplain on January 19, 2014 (Figure 2a). The breach occurred in the right bank levee at San Matteo, a few kilometers upstream of the Ponte Bacchello section. The event featured

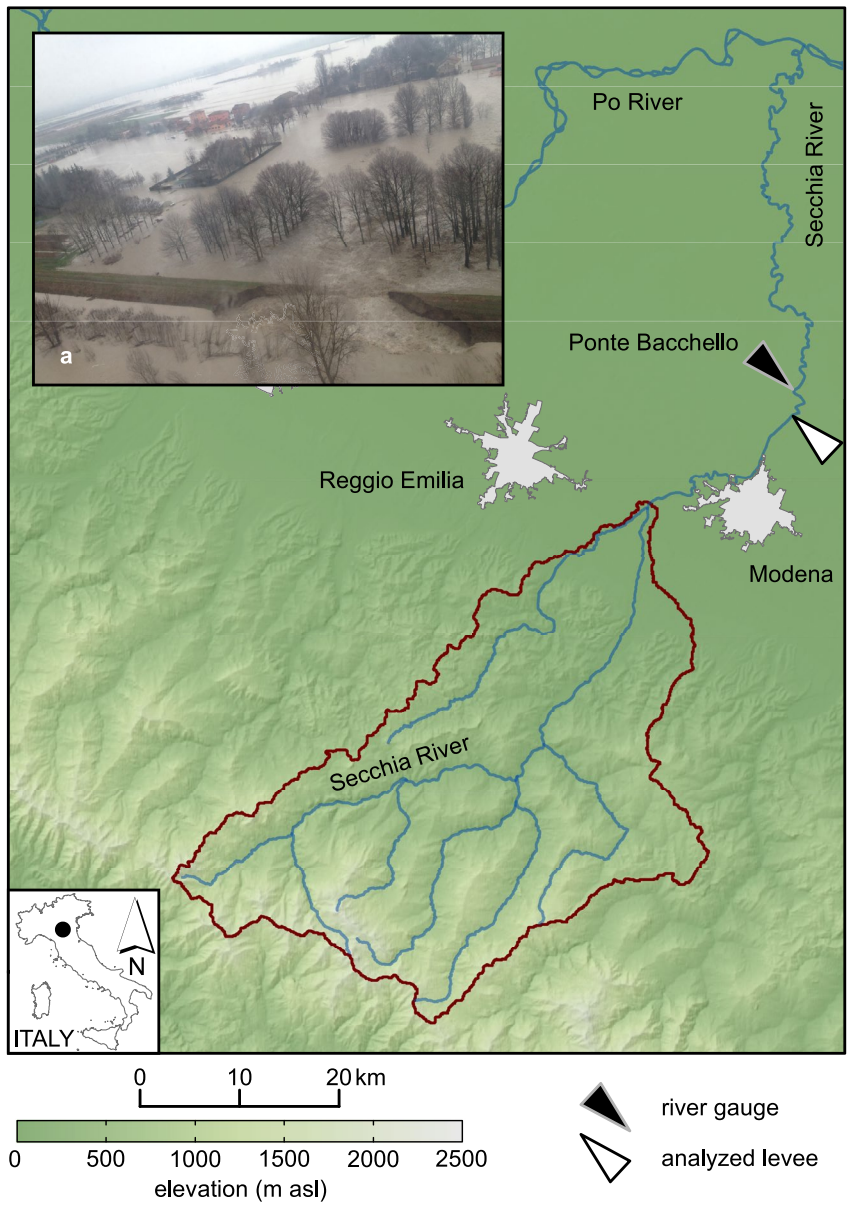

Figure 2. Secchia River drainage basin and (inset a) aerial photo of the levee breach occurred on January 19, 2014 near San Matteo (Modena, Italy).

a flooding volume of about 36×10^6 m^3, affecting an area of 52 km^2 and yielding a damage in excess of $500 million. The failure event was extensively analyzed by Orlandini et al. (2015) and a mammal den was found to be the most plausible cause for the levee failure. Therefore, a detailed knowledge of the hydraulic characteristics of the levees in this area was achieved. In addition, a 42-year long series of hourly flow discharges recorded at the Ponte Bacchello river gage is available, making it possible to statistically characterize the flood variability in this river reach. The analyzed period spans from 1923 to 1933 and from 1951 to 1981. After this period, the national hydrographic agency was dismissed, and regional hydrographic services were established. As a result, the observation series was interrupted for several years, during which the Secchia River underwent important hydraulic works, including the construction of an extensive flood control reservoir. Therefore, flood variability has been deeply changed and flow discharges presently recorded cannot be considered homogeneous with respect to previously recorded data. The data series used in the present study represents the longest available time series describing the flood variability of the Secchia River.

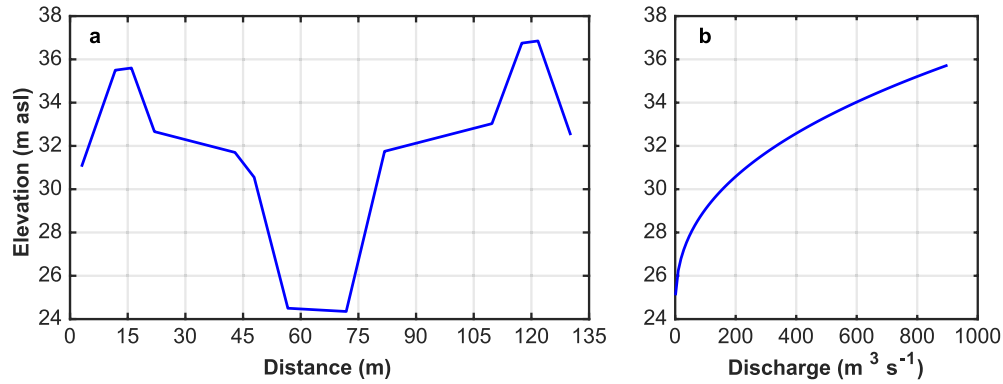

Figure 3. (a) Analyzed river cross section and (b) corresponding uniform flow rating curve.

The analyzed levee location is shown in Figure 2. The river cross section and the corresponding uniform flow rating curve are reported in Figures 3a and 3b, respectively. The river cross section geometry is determined from a 1-m digital elevation model generated from a lidar survey featuring horizontal errors less than 30 cm and elevation errors less than 15 cm. The uniform flow rating curve was set by separating the compound section in elementary subsections. Different values are adopted for the Gauckler-Strickler conductance coefficients k_S, which are set equal to 32 and 25 $m^{1/3}$ s^{-1} for the main channel and the floodplains, respectively. Such values are in accordance with the data reported in literature for a plain straight riverbed and banks covered by grass and light brush (Chow et al., 1988). The obtained uniform flow rating curve is consistent with the maximum of the observed flow discharge series, which was estimated to be 817.4 m^3 s^{-1}. The maximum flow discharge is therefore conveyed through this section without overflowing the shortest levee on the left bank. In fact, during the observation period overflowing events were not recorded. It is worth to underline that, during the last two decades, the overall conveyance capacity of this river reach has greatly reduced due to a decrease in maintenance. The derived uniform flow rating curve in Figure 3b must therefore be considered as representative of the river condition during the observation period. This does not compromise the generality of the performed analysis since its main objective is to illustrate the potential and the reliability of the hazard estimate procedure herein developed, and not to provide specific results for the reported case study.

3.2. Calibration and Test of Model Components

The assessments and the reliability verifications of the probabilistic and the deterministic model components of the Monte Carlo simulation technique used in the present study to estimate the failure hazard given by 18, as well as its uncertainty, are described in the following subsections. The probabilistic component is portioned into the natural variability of flood variables in Section 3.2.1. The deterministic component is given by the transformation function reported in Equation 16 as described in Section 3.2.2. The epistemic uncertainty is finally illustrated in Section 3.2.3.

3.2.1. Natural Variability of Floods

In consideration of the hydrologic characteristics of the river section and its upstream contributing area, a threshold flow discharge q_t equal to 180 m^3 s^{-1} and a minimum interevent time t_d equal to 24 hr are used to sample the independent floods from the discharge series. According to the uniform flow rating curve reported in Figure 3b, the threshold flow discharge is needed for the flow to begin the wetting of the left bank. An interevent period equal to about the double of drainage basin time of concentration is considered to ensure the independence of two subsequent floods. The selected parameters lead to a sample size n of 188 individual independent floods, with an average annual number of flood events n_y equal to 4.48. A scatterplot of the derived pseudo-observations is supplied in Figure 4a, visually showing a moderate-strong concordance for the association of peak flow discharge and flood duration. The Kendall rank correlation coefficient is estimated to be 0.68, while the independence hypothesis can be rejected for a test significance level less than 0.1%. This kind of moderate-strong concordant association is consistent with those observed in other flood frequency studies (Balistrocchi et al., 2017; Serinaldi & Kilsby, 2013).

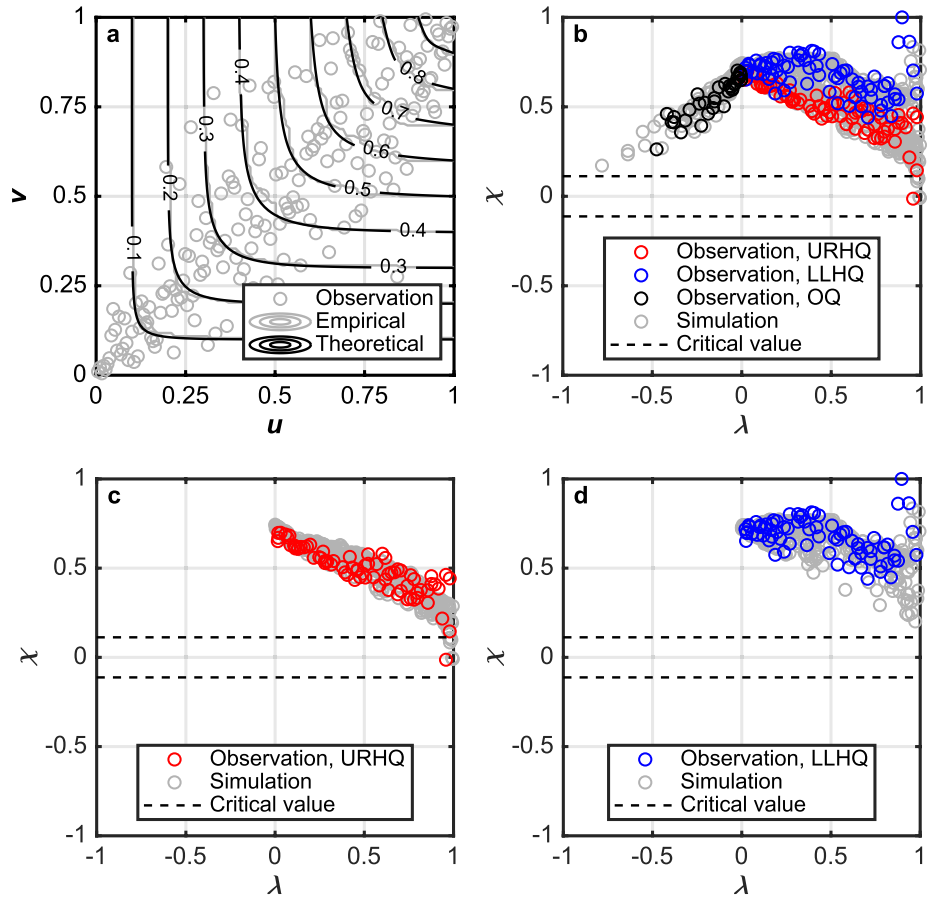

Figure 4. Global goodness-of-fit displayed by the copula H_ρ with respect to pseudo-observations: (a) copula CDF contour lines, (b) χ-plot of pseudo-observations derived from data and simulated by using the theoretical copula H_ρ given by Equation 1, (c) χ-plot of the pseudo-observations falling in the upper right-hand quadrant (URHQ), and (d) χ-plot of the pseudo-observations falling in the lower left-hand quadrant (LLHQ).

The parameter vector $\boldsymbol{\rho}$ which minimizes the sum of squared residuals between the theoretical copula given by 3 and the empirical copula (Cramer-von Mises statistics) was estimated by only considering the value ranges of concordant associations ($\phi > 0$ and $\psi > 1$) and assuming a dominant contribution of the Clayton copula ($\omega > 0.5$). This choice is intended to combine the tail dependence properties of the copulas to better suit the empirical evidence. The result is $\boldsymbol{\rho} = \{\omega = 0.52,\ \phi = 3.2,\ \psi = 4.7\}$, corresponding to a Cramer-von Mises statistics equal to 0.0094, which is far less than those achievable by the single copulas (0.0458 for C_ϕ and 0.0884 for G_ψ). The theoretical upper tail coefficient λ_U (6) and the lower tail coefficient λ_L (7) are estimated to be 0.40 and 0.42, respectively. The goodness-of-fit of the theoretical copula is illustrated in Figure 4 by comparing the contour lines of the cumulative function H_ρ to those of the empirical copula (Figure 4a). An alternative representation of the global dependence structure is given by the χ-plot reported in Figure 4b. A χ-plot is a scatterplot relating the measure λ of the distance of the pseudo-observations $(\hat{u}_i, \hat{v}_i)$ from the bivariate median (0.5,0.5), that is the center of the data set, to χ values, representing the departure from independence. Details for the computation of λ, χ and the critical values for testing independence can be found in Fisher and Switzer (2001). In Figure 4 critical values under the null hypothesis of independence refer to a significance level α of 10%. As highlighted by Joe (2015), empirical measures of tail dependences do not really exist because of the limit inherent in their definition. Graphical inspections by using scatterplots of pseudo-observations limited to suitable subsets of $[0,1]^2$ can however be conducted. According to Abberger (2005), in this paper a focus on the dependence structure featuring the tails was obtained by χ-plots reporting pseudo-observations belonging only to specific quadrants with respect to the bivariate median, namely the upper-right hand quadrant (URHQ) for the upper tail, the lower-left hand quadrant

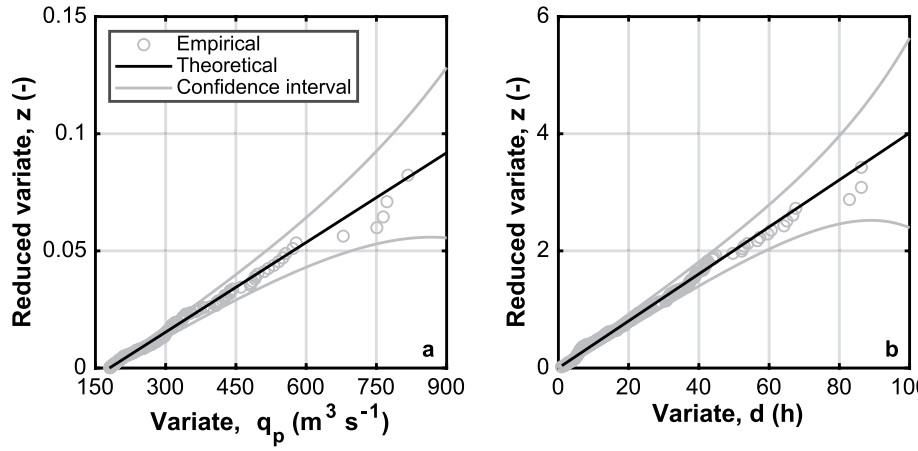

Figure 5. Probability plots showing the confidence interval for the significance level α of 10% for the marginal distributions of (a) peak flow discharge ($z = \kappa\left(q_p - q_t\right) / \xi$), and (b) flood duration ($z = d / \zeta$).

(LLHQ) for the lower tail. The remaining pseudo-observations are referred to other quadrants (OQ). Such χ-plots are reported in Figures 4c and 4d, respectively.

The χ-plot reported in Figure 4b indicates a moderate to strong concordant association which appears to be statistically significant as almost all the χ values are positive and larger than the critical values for the independence hypothesis not to be rejected. When pseudo-observations belonging to specific quadrants are considered, the greatest χ values are found to belong to the LLHQ (χ values varying between 0.5 and 1.0). A less marked concordance, but still statistically significant, is shown in the URHQ (χ values varying between 0.0 and 0.70). Therefore, a theoretical dependence structure featuring asymmetric tail dependence properties, in which the lower tail dependence coefficient is greater than the upper tail dependence coefficient, appears to fit the observed dependence structure. Actually, χ-plots based on 1,000 simulated pseudo-observations are visually compared to those derived for actual pseudo-observations, revealing a satisfactory global agreement (Figure 4b) and satisfactory agreements both in the URHQ (Figure 4c) and in LLHQ (Figure 4d). Details on copula simulation methods can be found in Mai and Scherer (2012) and Joe (2015).

The satisfactory goodness-of-fit visually shown in Figure 4 was quantitatively investigated by using test statistics. In particular, the blanket test proposed by Genest et al. (2009) and based on Cramer-von Mises statistics is adopted. By using a bootstrap procedure, an empirical estimate of the *p-value* that is the test significance level for which the null hypothesis cannot be rejected is obtained. In this case, the tested null hypothesis was that the underlying copula is the theoretical function given by 3. After 10,000 copula simulation runs, the *p-value* is found to be 64.5%. The obtained *p-value* is far larger than the usual 10% significance level routinely adopted in test statistics and it supports the suitability of the selected theoretical function.

The marginal CDFs for q_p and d given by 8 and 9, respectively, are fitted by using the maximum likelihood criterion obtaining κ equal to 0.015 and δ equal to 1.345, for the shape parameters, ξ equal to 118.24 m^3 s^{-1} and ζ equal to 24.9 hr, for the scale parameters. The overall goodness-of-fit was first evaluated by using the probability plots shown in Figure 5 that evidence a satisfactory agreement between the theoretical and the empirical functions. The confidence limits shown in Figure 5 are derived as suggested by Kite (1977) for a significance level α equal to 10%. Since the empirical distributions completely lie inside the 80% pointwise confidence intervals, the null hypotheses that the observation samples are distributed according to theoretical CDFs (8) and (9) cannot be rejected.

Additional evaluations were conducted by using the Kolmogorov-Smirnov (K-S) test and the Anderson-Darling (A-D) test (Kottegoda & Rosso, 2008). The first is useful to highlight the maximum residual between the theoretical and the empirical functions, while the second focuses more specifically on the fit of the tails. The K-S and the A-D tests were conducted by implementing a bootstrap procedure. The results are reported in Table 1, where the test statists D_{max} and A^2 along with the *p-values* obtained after 100,000 simulations are listed. All *p-values* are greater than 10% test significance level adopted in the present study, so that none of the null hypotheses should be

Table 1
Results of K-S and A-D Tests Reporting Test Statistics (D_{max} and A^2), Critical Values for a Significance Level α of 10% ($D_{n\alpha}$ and A^2_α) and the Empirical Estimates of the p-values

	K-S			A-D		
Marginal	D_{max}	$D_{n\alpha}$	*p-value*	A^2	A^2_α	*p-value*
F_{Q_p} given by 8	0.0523	0.0890	59.4%	0.6913	0.9160	57.6%
F_D given by 9	0.0624	0.0890	47.7%	0.6367	0.6370	10.4%

rejected. For the sake of completeness, Table 1 also reports the critical values $D_{n\alpha}$ and A^2_α, corresponding to 10% significance levels. In all the cases, test statistics are less than the corresponding critical values, and thus consistent with the results of the bootstrap procedure.

3.2.2. Transformation Function

The soil unsaturated hydraulic properties of the levee required to calibrate Equation 12 have been determined experimentally by using 10 undisturbed samples collected near the location of the levee failure occurred on January 19, 2014 (Figures S1 and S2 and Table S1 in Supporting Information S1). Levee soil at the site can be classified as a loam. Soil retention experiments and hydraulic conductivity tests were performed as described in Romano and Santini (1999) and results are summarized in Table S1 in Supporting Information S1. Additional information on foundation soils can be found in D'Alpaos et al. (2014). The constitutive equation relating the volumetric soil water content θ to the soil matric potential SMP is given by the van Genuchten equation $\Theta = \left\{1 / \left[1 + \left(a_{vg}\left|\mathrm{SMP}\right|\right)^{n_{vg}}\right]\right\}^{m_{vg}}$, where $\Theta = \left(\theta - \theta_r\right) / \left(\theta_a - \theta_r\right)$ is the degree of saturation, θ_a is the air-filled porosity, θ_r is the residual soil water content, SMP is the soil matric potential, a_{vg}, n_{vg}, and m_{vg} are parameters. The constitutive equation relating the unsaturated hydraulic conductivity K to the volumetric soil water content θ was defined by the Mualem model as $K = K_c\, K_r$, where K_c is the soil saturated hydraulic conductivity and the soil relative hydraulic conductivity is given by $K_r = \Theta^{1/2}\left\{1 - \left[1 - \Theta^{1/m_{vg}}\right]^{m_{vg}}\right\}^2$. Soil parameters for the analyzed levee are $\theta_r = 0.079$, $\theta_a = 0.404$, $n_{vg} = 1.495$, $m_{vg} = 1\text{–}1/n_{vg} = 0.331$, $a_{vg} = 1.543\ \mathrm{m}^{-1}$, and $K_c = 1.881 \times 10^{-6}$ m s^{-1}. In particular, the air-filled porosity θ_a and the saturated hydraulic conductivity K_c in Equation 12 are assessed as the arithmetic mean and the geometric mean of the sample values, respectively.

To test the simplified seepage flow model given by Equation 12, detailed numerical simulations based on the Richards equation are performed by using the soil hydraulic properties reported above as shown in Figure 6. The levee geometry incorporated in the simplified seepage flow model is shown by the gray dashed line. In the simplified seepage flow model, the soil domain extending between the levee body and the active channel is not entirely represented. This is equivalent to assuming that the omitted soil portion saturates instantaneously during the flood event. The soil matric potential SMP in Equation 12 is estimated to be −1.0 m. Numerical simulations are performed by using the FEFLOW model (Diersch, 2014). The levee geometry in the numerical model is based on a 1-m resolution lidar survey. As shown in Figure 6a, the computational soil domain has a planar extension along the river cross-section of about 450 m and a soil thickness of about 30 m. Boundary conditions assigned to the land surface are Dirichlet-type (assigned potential) conditions, representing a levee boundary wetted by riverflow. The head is assigned to the mesh nodes only when the river stage is greater than the node elevation.

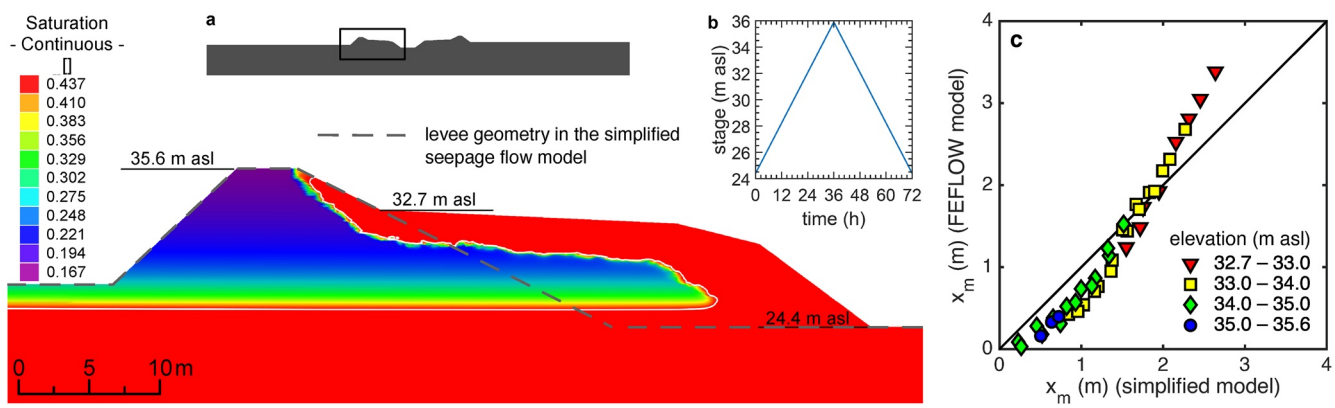

Figure 6. Detailed numerical modeling of seepage flow performed by the FEFLOW model to test the simplified seepage flow model given by Equation 12. Panel (a) shows the volumetric soil water content at time $t = 37$ hr when the levee is forced by the stage hydrograph reported in panel (b). The term "saturation" reported in the legend of panel (a) is used in the FEFLOW model to denote the volumetric soil water content. The phreatic line is reported in white. Predictions of maximum horizontal advance x_m obtained from the simplified model and from the FEFLOW model are compared in panel (c).

For all the mesh nodes belonging to the land surface outside the wetted perimeter of the river cross section, a Dirichlet-type (assigned potential) conditions constraint by the flux is assigned. When flow is directed outward the domain, the head equal to the elevation of the node is assigned. This makes it possible to correctly simulate the possible return to the land surface of flood water seeping across the levee. A no-flux Neumann-type boundary condition is assigned to the bottom of the domain, whereas a Dirichlet-type was assigned to the lateral boundaries (far from the levee) to represent a water table depth of about 1.50 m. The mesh is refined to obtain nodes along the levee riverside with a spatial resolution in elevation of at least 0.10 m and a 10-min time step for the definition of boundary conditions is used.

As exemplified in Figure 6b, the seepage process is simulated in response to four triangular flood hydrographs by combining two flood durations d of 48 and 72 hr, with two peak water stages of 11.5 and 10.5 m, respectively (35.9 and 34.9 m asl, respectively). These values are selected to represent relevant hydroclimatic forcing to the analyzed levee. The triangular shape is used to ensure consistency between the compared models, namely Equation 12 and FEFLOW. The maximum horizontal advances x_m are then obtained at different elevations y varying between the levee toe on the riverside and the levee crest (35.6 and 32.7 m asl, respectively) with resolution of 25 cm. A sample of 41 simulated values was obtained and compared to the corresponding estimates obtained by Equation 12. The visual comparison of such estimates is reported in Figure 6c. By assuming that numerical simulations are benchmarking, the estimates of the Nash-Sutcliffe efficiency and of the linear correlation coefficient are 0.86 and 0.98, respectively. Therefore, the model comparison is satisfactory and supports the implementation of the simplified model given by Equation 12. It is finally remarked that according to these simulations, the time for the levee soil to be completely drained after a flood event is assessed in no more than 2–3 days, indicating a negligible memory effect in multiple flood events in these systems.

Examples of the bivariate population splits between the safety region and the failure region that can be achieved by using Equation 17, with soil hydraulic properties previously reported and uniform flow rating curve $h = 0.749\,q^{0.401}$ (Figure 3b), are plotted in Figure 7 for left bank levee and for different den elevations and residual thicknesses when $K_c = 1.88 \times 10^{-6}$ m s^{-1}; $\theta_a = 0.404$; SMP $= -1.0$ m. The flow discharge of 400 m^3 s^{-1} corresponds to the value needed for the levee toe to be pounded. The dashed line indicates the crest overflow discharge, that is the threshold flow discharge over which the additional failure mechanism of overflowing must be taken into consideration.

3.2.3. Epistemic Uncertainty

Epistemic uncertainty deals with the imperfect knowledge of the analyzed system as a result of the limited ability to measure and describe natural phenomena. Limited extension and quality of hydrologic data is another relevant source of uncertainty. In hydrological applications such as levee breach, the epistemic uncertainty is remarkable and deserves to be suitably accounted for to provide meaningful estimates of the variables of interest (Vorogushyn et al., 2011). As highlighted by Apel et al. (2004), in a Monte Carlo analysis the epistemic uncertainty can be modeled by randomizing the parameters of the deterministic model components. Sources of uncertainty are assumed to be mutually independent, so that random values are generated separately by using univariate techniques. The sources of uncertainty considered in the present study are listed below, along with the characteristics of the CDFs implemented in the Monte Carlo analysis to estimate the uncertainty in the obtained estimates of the levee failure hazard. These sources are:

- Soil saturated hydraulic conductivity K_c. The uncertainty is modeled by using the conductivity values estimated for the 10 soil samples through a triangular distribution. The density function was assumed to vary between the minimum estimate of 8.333×10^{-8} m s^{-1} and the maximum estimate of 1.301×10^{-5} m s^{-1}, while the geometric mean of 1.881×10^{-6} m s^{-1} is assumed to be the mode. The triangular distribution is used in preference of the lognormal distribution as the triangular distribution makes it possible to equal the pdf mode to the sample geometric mean while also keeping a positive skewness (Figure S3 in Supporting Information S1). It is remarked here that the triangular distribution of soil saturated hydraulic conductivity observed at a given point is used in the Monte Carlo analysis to determine the value at any point inside the levee body. The probability that this value is less than the minimum observed value or greater than the maximum observed value is assumed to be insignificant.
- Soil air-filled porosity θ_a. The uncertainty is modeled by using the porosity values estimated for the 10 soil samples through a normal distribution, characterized by a mean of 0.404 and a standard deviation of 0.047.

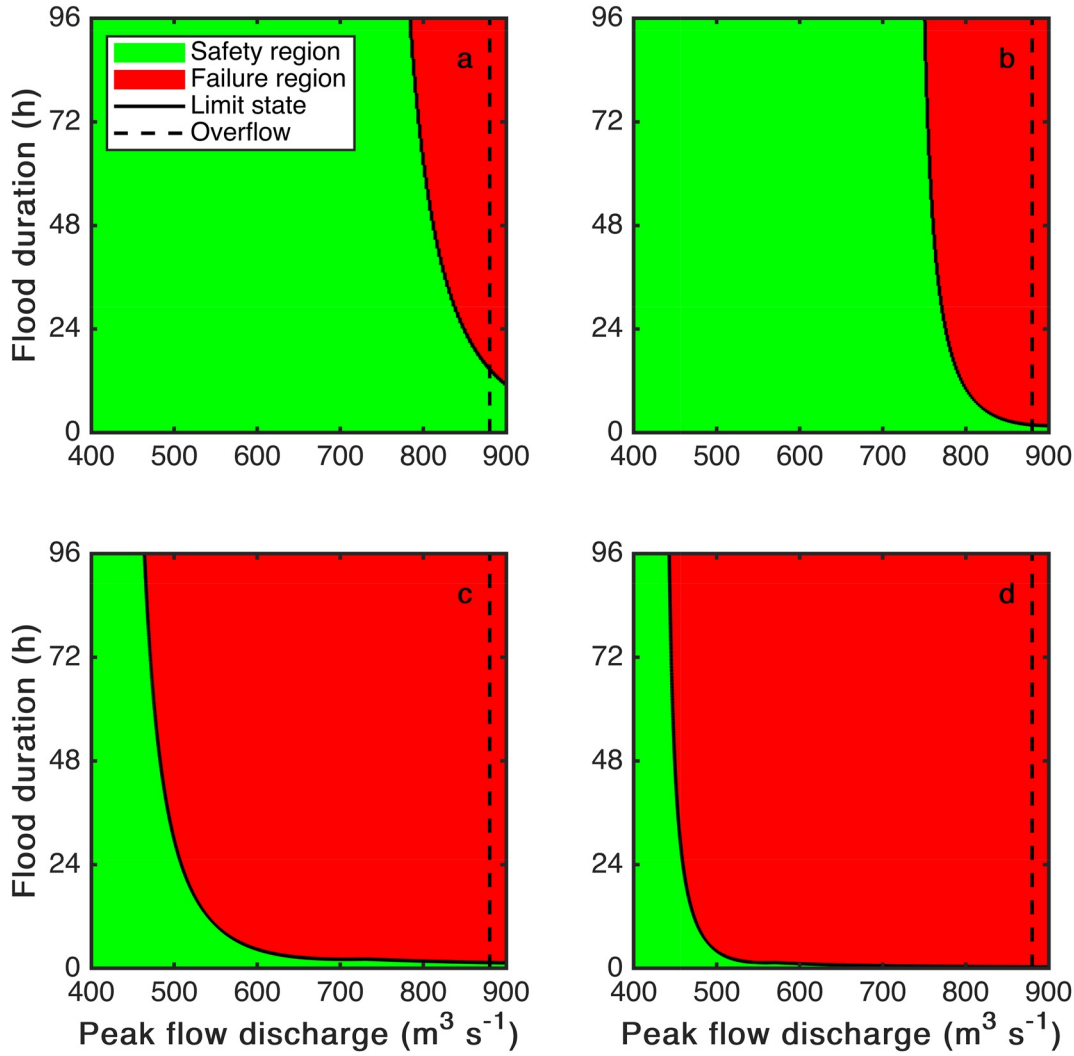

Figure 7. Examples of the bivariate population splits for (a) den elevation $y = 10.5$ m (about 0.7 m below the levee crest) and residual thickness $b_r = 1.0$ m, (b) $y = 10.5$ m and $b_r = 10.5$ m, (c) $y = 8.5$ m (about 2.7 m below the levee crest) and $b_r = 1.0$ m, and (d) $y = 8.5$ m and $b_r = 0.5$ m. Vertical dashed lines indicate the value of the threshold flow discharge of 880 $m^3 s^{-1}$ above which overflow occurs.

- Soil matric potential SMP in Equation 12. In consideration of its high uncertainty, a uniform density function was adopted for the soil matric potential. A minimum value of −1.20 m and a maximum value of −0.80 m are considered.
- Uniform flow rating curve. Owing to the precision and the reliability of the survey that yielded the cross-section geometry reported in Figure 3a, the uncertainty is mainly related to the Gauckler-Strickler conductance coefficients k_S of the main channel and the floodplains. This parameter is also variable due to the season and the maintenance status. With respect to the mean parameter values reported in Section 3.1, a variability of ±5% and of ±10% are therefore assumed for the main channel and the floodplains, respectively. If the uniform flow rating curves are approximated by simple power functions of the kind $h = a\,q^b$, its variability can be expressed in terms of the variability of parameters a and b. Hence, normal distributions are assumed for both the coefficient a (mean equal to 0.749, standard deviation equal to 0.003) and the exponent b (mean equal to 0.401, standard deviation equal to 0.002). The mean values delineate the uniform flow rating curve in Figure 3b, whereas the standard deviation is set by assuming that the maximum and the minimum uniform flow rating curves bound a variability range equal to six times the standard deviation.

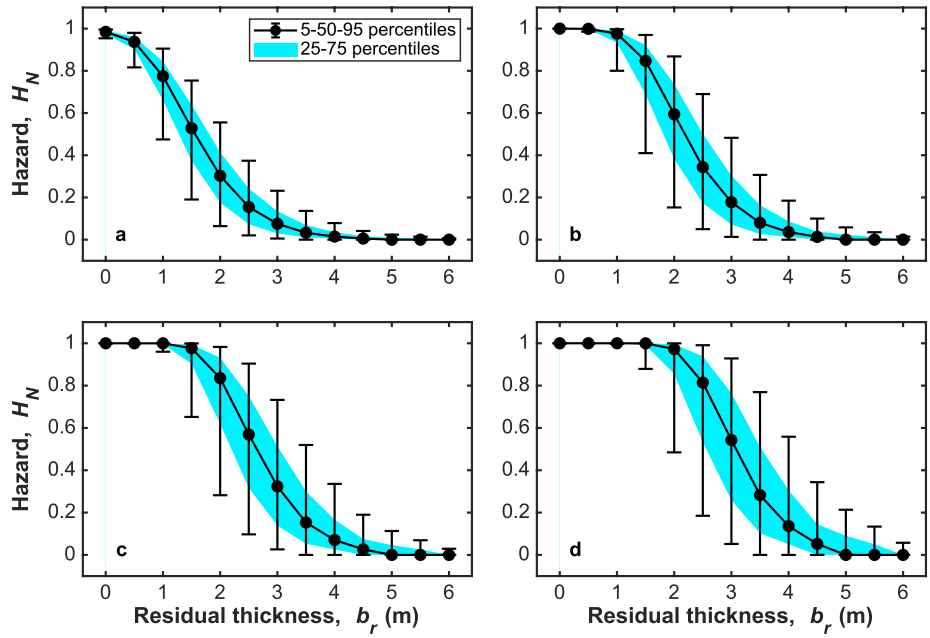

Figure 8. Hazard trends and related uncertainties as obtained by varying the residual thickness b_r at a constant elevation y = 9.5 m (about 2.0 m below the levee crest) and for increasing planning period N: (a) 20 years, (b) 50 years, (c) 100 years, and (d) 200 years.

3.3. Hazard Estimation Results

The hazard estimates conducted by using Equation 18 for different planning periods N can be expressed in terms of functional boxplots as reported in Figure 8. The reference value is given by the median, and the uncertainty band is delimitated by the first and the third quartiles. For the sake of completeness, the 5% and the 95% quantiles are also reported in Figure 8. The parameters of the limit state function are randomly varied according to the CDFs discussed in Section 3.2.3. To achieve a satisfactory stability of results while also maintaining computational efficiency, 1,000 simulation runs are conducted. Under the considered model assumptions, a variation of only 2%–3% in the percentile estimates of the event failure probability is obtained by increasing the number of simulation runs up to 10,000. In this derivation, a fixed elevation of 9.5 m (about 2.0 m below the levee crest) is chosen to reproduce a condition similar to that of the levee failure occurred on January 19 2014 along this river reach. Different plots can be obtained by varying the den geometrical parameters. For instance, Figure 9 shows the functional boxplots of the hazard increasing trends with respect to the planning period N, estimated for increasing values of the den elevation y when a constant residual thickness b_r equal to 1.5 m is considered. Conversely, Figure 10 shows the same trends estimated when the den length b_d is set equal to 5.6 m and the residual thickness b_r is varied.

4. Discussion

The convex linear combination of Clayton and Gumbel-Hougaard copulas defined in Equations 3–5 is found to provide a suitable model closely fitting the empirical distribution (Section 3.2.1). As shown in Figure 4, asymmetric and statistically significant tails were suitably reproduced by combining the different tail properties of the individual components. Despite the increase in probabilistic model complexity, the suggested linear combination is selected in preference to other copula families commonly adopted in flood frequency analysis such as the elliptical copulas. In fact, the computational advantages of the Archimedean copulas related to the existence of the explicit analytical expression of the copula cumulative function are maintained. It is acknowledged that investigating tail properties is always a nontrivial task, due to the conceptual issues of deriving estimators for the empirical distributions (Joe, 2015). However, the empirical copula characteristics outlined in Figure 4 are often observed in bivariate flood frequency analysis and fueled the debate on the use of copula featuring only lower tail dependence (Balistrocchi et al., 2017, 2019; Chowdhary et al., 2011), only upper tail dependence (Karmakar

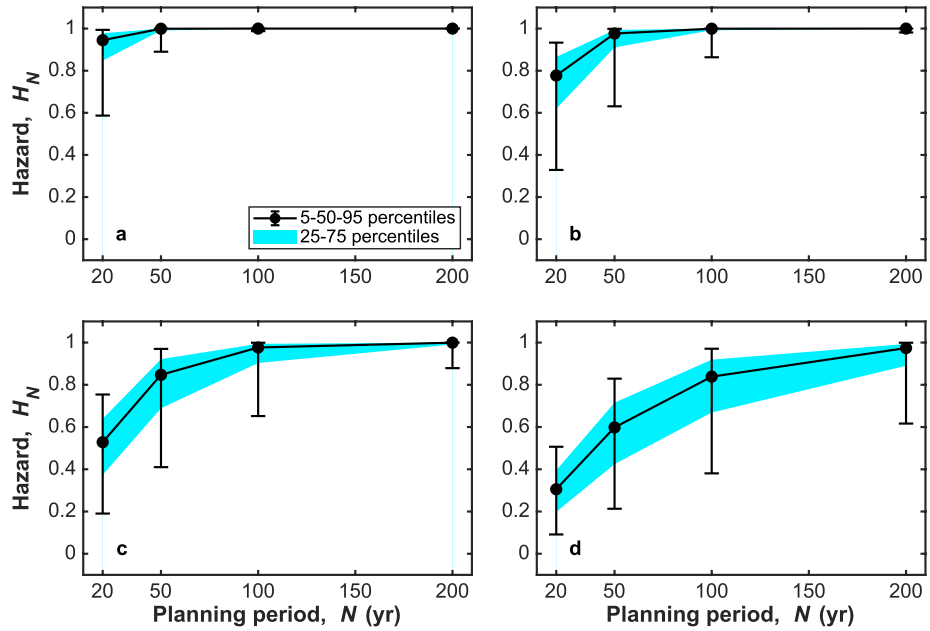

Figure 9. Hazard trends and related uncertainties as obtained by varying the planning period N for a constant residual thickness b_r = 1.5 m and for increasing elevation y: (a) 8.5 m, (b) 9.0 m, (c) 9.5 m, and (d) 10.0 m.

& Simonovic, 2009; Requena et al., 2013; Zhang & Singh, 2007), or both tail dependences (Ganguli & Reddy, 2013). Therefore, despite the estimation of tail dependences in empirical samples is theoretically challenging, these aspects of the joint variability do deserve investigation as made in Figure 4.

Numerical experiments conducted through the FEFLOW model to investigate seepage dynamics in the analyzed levee system demonstrate that the time needed for the seepage water to completely drain after a significant flood is in the order of 2–3 days (Section 3.2.2). This period is far less than the average interevent time between two

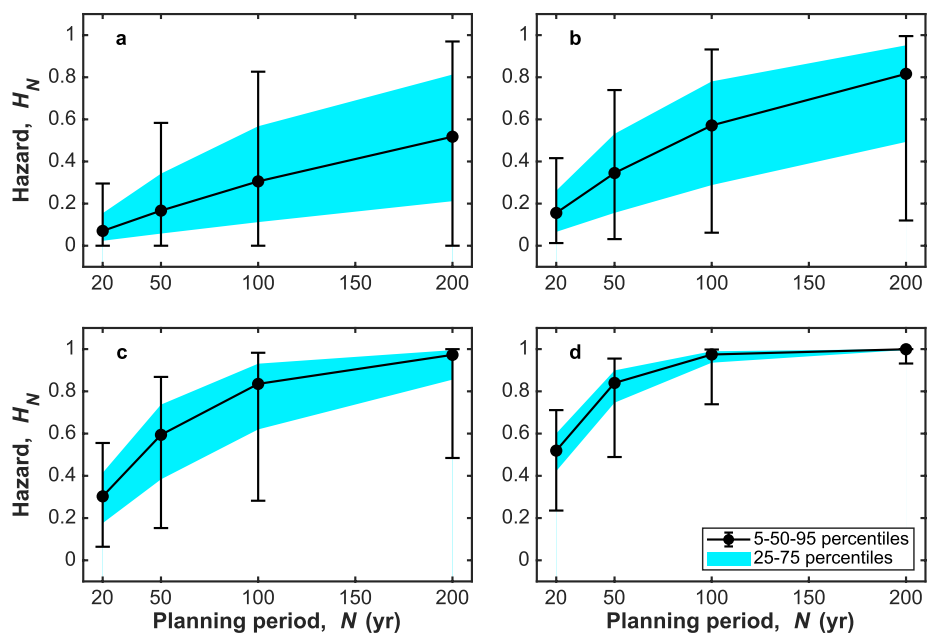

Figure 10. Hazard trends and related uncertainties as obtained by varying the planning period N for a constant den length b_d = 5.6 m and for increasing elevation y: (a) 8.5 m (b_r = 4.0 m), (b) 9.0 m (b_r = 3.0 m), (c) 9.5 m (b_r = 2.0 m) and (d) 10.0 m (b_r = 1.0 m).

subsequent floods of about 80 days, that corresponds to an average annual number of flood events n_y equal to 4.48 (Section 3.2.1). Thus, the levee moisture content can be assumed to be restored to the initial condition, or at least to an unsaturated soil condition, when an independent flood occurs. This supports the hypothesis of statistical independence of the sample flood events, from the physical point of view concerning seepage flow dynamics.

The complete model proposed by Vorogushyn et al. (2009), and herein modified as given by Equation 12 to be suitably implemented in the limit state function given by 16, was found to provide results comparable to those obtained from detailed numerical simulations (Figure 6). As shown in Figure 6c, deviations between the simplified and detailed model results are found to be acceptable in consideration of the uncertainties related to the description of seepage processes. The scatterplot reported in Figure 6c indicates that there is a satisfactory overall agreement between the values of the maximum horizontal advance x_m obtained from the simplified seepage flow model and by the numerical FEFLOW model (Nash-Sutcliffe efficiency equal to 0.86, linear correlation coefficient equal to 0.98). The simplified seepage flow model underestimates x_m for values of x_m greater than about 2 m, which are especially obtained at lowest elevations, and overestimates x_m for values of x_m less than about 2 m, which are especially obtained at highest elevations. The obtained trend is likely to be connected to the high nonlinearity of soil hydraulic conductivity with soil water content incorporated in the numerical FEFLOW model and to the longer impoundment of the low points of the levee with respect to those located at a higher elevation. In addition, the computational burden decreases dramatically when Equation 16 is used in preference to detailed numerical models, making it possible to account for the epistemic uncertainty in the whole modeling framework developed.

Other studies used a simplified version of Vorogushing et al.'s model, but they did not account for the contribution to seepage flow due to the infiltration from the riverside slope (Michelazzo et al., 2018; Oliver et al., 2018). More precisely, in previous studies the seepage length ℓ_m estimated in Equation 11 is set equal to the seepage horizontal advance x_m (Figure 1) and the impact of soil matric potential SMP is disregarded. These assumptions have also been tested in the present study. The model given by Equation 11 is found to yield significant underestimations of the maximum saturated front extents when compared to numerical simulations. It is concluded that the terms neglected in Equation 11 are important and in the results obtained from previous studies the effect of these terms is compensated, at least in part, by the use of a rectangular hydrograph in preference to a triangular hydrograph. In accordance with Vorogushyn et al. (2009), Equation 11 can be used to describe the substratum piping processes, whereas Equation 12 and triangular hydrographs are needed to describe the seepage process through the levee body.

The limit state function obtained from Equation 16 is conceptually meaningful. As shown in Figure 7, a shorter flood duration is needed for the failure mechanism to be triggered as the peak flow discharge increases. This configuration of the limit state function appears to be realistic since the higher the peak flow discharge, the steeper the seepage flow hydraulic gradient is. In addition, the failure region expands when the residual thickness b_r decreases for a fixed elevation (compare Figure 7a to Figure 7b, and Figure 7c to Figure 7d), and when the den elevation y decreases for a fixed residual thickness (compare Figures 7a–7c, and Figures 7b–7d). The first behavior is expected as the residual thickness is the path length that must be crossed by the seepage to trigger the failure mechanism. Conversely, the second behavior can be explained by considering that, at a lower elevation, the den is subject to a longer pounding duration and to a higher hydraulic load. As shown in Figure 7, levee overflowing under noncritical seepage flow conditions occur for very short flood durations, which may not be critical for levee crest erosion and breach formation. In any case, levee failure due to excessive seepage appears to be dominant when dens exist. It is finally remarked that the limit state function defined by Equation 16 and illustrated in Figure 7 identifies infinite flood events expressed in terms of peak flow discharge q_p and flood duration d that correspond to a unique value of maximum horizontal advance x_m inside the levee, and thus to a unique value of return period.

The method developed in Section 2.2.1 takes into consideration the potential presence of a mammal den inside the levee body, making it possible to carry out assessments for undisturbed and disturbed levees. As shown in Figure 8, a decreasing trend in the hazard H_N with respect to the residual thickness b_r is obtained. Hazard is found to be negligible when b_r is larger than 5.0 m, even for a planning period of 200 years. Since the levee thickness is 6.0 m at this elevation, the hazard related to seepage is almost null in the absence of mammal dens and the seepage failure mechanism could be disregarded in the estimate of the overall hazard (Figure 8d). In contrast, the hazard is unacceptably high for a planning period of 20 years when b_r is less than 2.5 m (Figure 8a). In this

situation, the hazard is actually estimated to be 15.5%. When the den completely crosses the total levee thickness (that is when b_r is equal to zero) the hazard equals the probability that the den is pounded during the planning period. This event is almost certain for the 20 years planning period as the median hazard is estimated at 98.5%. The return period of a levee failure due to the excessive seepage mechanism only can be estimated by Equation 19. In the absence of dens, T_r is greater than 200 years and decreases down to 9 years when bioerosion extends for 84% of the levee thickness and 5 years when the den completely crosses the levee. Major uncertainties are related to a residual thickness varying between 1.5 and 3.0 m, as the other values lead to hazard estimates very close to one or zero, respectively. With regard to the planning periods of 20 and 200 years, the maximum widths of the uncertainty bands are 17.9% and 40.4%, respectively. This analysis was repeated at various elevations, revealing that the maximum distance for the hazard to be negligible is always shorter than the levee total thickness b_t for return periods far greater than 200 years. Therefore, the analyzed levee can be considered globally verified, with regard to the flood variability previously featuring the plain river reach. The results obtained for undisturbed and disturbed levees are consistent with the events observed in the study area.

A strong impact of mammal dens on the levee structural integrity is clearly revealed. As shown in Figure 8, the hazard referred to the excessive seepage failure mechanism is negligible in undisturbed conditions, whereas it increases dramatically when the levee body is affected by bioerosion. Indeed, based on obtained model results, the levee analyzed in the present study is verified with respect to excessive seepage induced failure for flood return periods greater than 200 years as required for levees belonging to the main hydrographic network. It is, however, remarked here that the statistical analysis reported in the present study is based on a flow record having length of 42 years, from which a POT series of 188 flood events is extracted and used, and only events having return period that do not exceed 100 years are conservatively considered to be statistically reliable. The functional boxplots obtained for the hazard H_N by varying the planning period N (Figures 9 and 10) show an increasing trend. When a fixed residual thickness and different elevations are considered, the hazard trends shown in Figure 9 are obtained. Since the levee thickness increases as the elevation decreases, a constant residual thickness involves more extended dens near the levee toe. Under this assumption, the worst situation is represented by a den located near the levee toe (y equal to 8.5 m, den length b_d equal to 8.0 m), for which the hazard H_N is estimated to be 94.5% for N equal to 20 years (Figure 9a). However, near the levee crest (y equal to 10.0 m, den length b_d equal to 5.0 m), the estimate of the hazard H_N is 30.6% for N equal to 20 years, which increases up to 97.4% for N equal to 200 years (Figure 9d). Hence, the return period T_r of a failure event only due to the excessive seepage mechanism increases from 7 to 55 years, when the elevation y increases from 8.5 to 10.0 m and the residual thickness is 1.5 m. Such a behavior can be explained by considering that the lower the den elevation, the longer the pounding durations are and the higher the hydraulic heads are. Owing to the hazard values close to the upper limit of 100%, uncertainty bands are quite narrow for low elevations (Figures 9a and 9b). The maximum uncertainty is observed for a den elevation y of 10.0 m and a planning period N ranging between 50 and 100 years where the band width is assessed to range from 25% to 29%.

A different behavior with respect to the variation of the den elevation is nevertheless assessed if the den length b_d is kept constant instead of the residual levee thickness b_r. In the example shown in Figure 10, the worst situation is given by a den placed near the levee crest. The hazard for a den elevation of 8.5 m varies between 7.0% and 51.8%, as the planning period increases from 20 yr to 200 years (Figure 10a). In contrast, the corresponding variation for a den elevation of 10.0 m is estimated to range between 51.8% and 99.9% (Figure 10d). The return period of the failure event thus decreases from 270 to 30 years, as the den elevation increases. This can be explained by admitting a major contribution of the increase in the resistance factor (the residual levee thickness b_r) with respect to the worsening of the load factor (pounding duration d_p and hydraulic head $\bar{h}$) in the limit state function given by 16. The uncertainty bands are large for a den elevation of 8.5 m (Figure 10a) and of 9.0 m (Figure 10b), where the maximum width is estimated to range from 49% to 60% for the planning periods ranging between 100 and 200 years. The greatest uncertainties of the hazard estimates at the lowest elevations can be explained by considering the great contribution of the uncertainties of the soil hydraulic parameters when large soil thicknesses are involved. The assumption of constant den length (Figure 10) appears to be more realistic than assuming constant residual levee thickness (Figure 9). Although the den geometry shows significant variability on a global scale, such a variability greatly diminishes on regional and local scales, where the population number is in balance with respect the resources of the surrounding habitat and similar soil types are used to construct the levee system. The analysis reported in the present paper confirms therefore that the major threat to levee stability is due to mammal

species burrowing dens close to the levee crest, such as crested porcupines and badgers, as reported in Orlandini et al. (2015).

Aiming at making the illustration of the developed methodology more effective, the geometrical characteristics of the den are not implemented in the uncertainty analysis. Thus, the residual levee thickness b_r, that is the actual length that seepage must cross to reach the landside slope, and the den elevation y were considered to be known (Section 3.3). Nevertheless, these parameters can be assessed through on-site surveys by using indirect methods or remote sensing. At the state of the art, den surveys are affected by uncertainties (*e.g.*, Borgatti et al., 2017). The Monte Carlo procedure herein developed makes it possible to implement this additional uncertainty sources in the hazard uncertainty estimate, by randomizing the den elevation y and the residual thickness b_r according to the specific uncertainties related to the survey methodology. This is however left to future more comprehensive analyses. Among the uncertainty sources considered in this study (Section 3.2.3), a major role is played by the estimate of the soil hydraulic conductivity K_c. Owing to the width of the uncertainty bands shown in Figure 10, it must be remarked that a precise estimate of this soil parameter is crucial to obtain meaningful estimations of the hazard when the den is located near the levee toe, where the residual thickness is basically large. As shown in Figures 10a and 10b, when the residual thickness b_r is greater than 3 m, uncertainty bands are significantly wider than those obtained for b_r less than 2 m and reported in Figures 10c and 10d. Clearly, when the uncertainty bands are wide, caution must be exercised in the use of estimated median value of the hazard for the longest planning periods.

The results reported in the present study must be qualified by recognizing limitations in the description of real levees. Mammal dens can form a complex tridimensional tunnel network often including chambers and multiple entrances. Complexity of these dens depends on the size and the age of mammal community, along with the mechanical characteristics of the soil and the potential for the mammal community to grow up. The essential parameters are the den elevation y with respect to the thalweg and the corresponding minimum residual thickness b_r. These are critical parameters affecting the hydrologic forcing and the excessive soil saturation in the specific locations where the levee is weakened by a den. Critical points where the residual thickness is small can exist at different elevations. The impact on levee failure of multiple critical points can be accounted for by applying the total probability theorem, by estimating the probability that excessive saturation occurs at any critical point. Given the joint distribution function of peak flow discharge and flood duration, such a probability can be estimated by integrating the probability density function over the failure region obtained as the union of individual failure regions related to all critical points. Alternatively, a detailed numerical model can be used to incorporate complex den geometries in the modeling framework described in the present study at the expense of higher computational cost. Only the levee failure mechanism due to excessive seepage is considered in the present paper. Levee failure due to overflowing is analyzed in Balistrocchi et al. (2019). The total failure probability for a levee can however be computed as the probability of the union of events caused by all possible failure mechanisms. These events include those due to internal erosion, caused by excessive seepage, those due to crest erosion, caused by overflowing, and those due to mechanical instability, caused by soil saturation. These events are not mutually exclusive as an individual flood event can trigger one or more failure mechanisms.

It is finally acknowledged that real levees may display soil heterogeneity across the levee body and the levee foundation, and that seepage cutoff walls may be another relevant source of heterogeneity. Based on the experimental analysis of undisturbed soil samples reported in Section 3.3.2, the assumption of homogeneity for the levee body considered in the present study appears to be acceptable. No undisturbed soil samples were collected in the levee foundations as the problem addressed in the present study was observed to affect the levee body only. However, geotechnical surveys indicated that no relevant soil heterogeneities are present in the levee foundation (D'Alpaos et al., 2014). Seepage cutoff walls are a possible means for reducing the impact of borrowing mammals on levees. Seepage cutoff walls for levees may have limited durability and may not entirely solve the problem addressed in the present study. Theoretical and experimental work is, however, needed to appropriately test this possible structural measure. In order to address structural heterogeneity of levees and related foundations, in the modeling framework proposed in the present study, a detailed numerical model has to be used in preference to the simplified seepage flow model described in Section 2.2.1 at the expense of higher computational cost. In any case, the modeling framework proposed in the present study can be applied to any levee with the appropriate characterization of soil properties and the appropriate selection of the deterministic component describing seepage flow.

5. Conclusions

In the present study, a new model for estimating the failure probability of levees affected by bioerosion was developed (Section 2). The statistical component of the model implements a fully bivariate approach to represent the natural variability of peak flow discharge and flood duration (Section 2.1), whereas the deterministic component was developed by using an extended version of the Vorogushyn et al.'s (2009) analytical model of seepage flow in the levee body (Section 2.2). This model was found to be reliable, computationally efficient and able to incorporate any flow hydrograph shape. Comparison with FEFLOW numerical model, revealed its capability to represent the maximum extent of saturated fronts inside the levee body under unsteady unsaturated flow conditions (Section 3.2.2). The bivariate nature of flood events was summarized in a state variable through the deterministic component that operates as a transformation function (Section 2). Owing to the derived distribution theory, the hazard estimated by using the model herein developed is related to a fixed period of time, such as a planning period or a return period, in a conceptually meaningful manner. From a theoretical point of view, this assessment was found to be more appropriate than those obtained by using detailed numerical simulations of individual synthetic floods or conditional approaches relying on arbitrary assumptions on flood durations to understand and predict the long-term behavior of levee systems.

The increase in failure probability due to mammal bioerosion revealed by the developed bivariate approach was found to be dramatic, for commonly observed locations and extents of the dens (i.e., den extending for most of the levee width leaving a residual thickness of undisturbed soil of 1 m on the riverside). The analyzed levee located along the Secchia River was tested against seepage flow under undisturbed conditions and the analyzed flood regime, for a planning period of 200 years or a return period greater than 100 years (Figure 8d). In contrast, when a den reproducing the essential structure observed in the real-world is incorporated in the model, the levee failure is almost certain during a planning period of 5 years and the failure return period is less than 10 years (Figure 8a and Section 3.3). These results are consistent with the events observed in the study area. Future work is needed to investigate the impact of den complexity and levee structural heterogeneity on levee failure probability. Additional research is also needed to assess the impact of riverflow regulation on levee instability due to the combined occurrence of overflowing, excessive seepage, and mechanical instability.

Data Availability Statement

Data used in this study are available at the repository https://doi.pangaea.de/10.1594/PANGAEA.931759.

Acknowledgments
The research reported in the present paper was supported by Fondazione di Modena through the grant 2018-0093 and by University of Modena and Reggio Emilia through the grant FAR 2020 Mission Oriented. The authors are grateful to Alberto Montanari, Francesco Serinaldi, and two anonymous reviewers for comments that led to improvements in the manuscript.

References

Abberger, K. (2005). A simple graphical method to explore tail-dependence in stock-return pairs. *Applied Financial Economics*, *15*(1), 43–51. https://doi.org/10.1080/0960310042000280429

Akan, A. O., & Houghtalen, R. J. (2003). *Urban hydrology, hydraulics and stormwater quality*. Hoboken, NJ: Wiley & Sons.

Alexandre, M., Hipólito, D., Ferreira, E., Fonseca, C., & Rosalino, L. M. (2020). Humans do matter: Determinants of red fox (vulpes vulpes) presence in a western Mediterranean landscape. *Mammal Research*, *65*(2), 203–214. https://doi.org/10.1007/s13364-019-00449-y

Apel, H., Thieken, A. H., Merz, B., & Blösch, G. (2004). Flood risk assessment and associated uncertainty. *Natural Hazards and Earth System Sciences*, *4*(2), 295–308. https://doi.org/10.5194/nhess-4-295-2004

Bacchi, B., Brath, A., & Kottegoda, N. T. (1992). Analysis of the relationships between flood peaks and flood volumes based on crossing properties of river flow processes. *Water Resources Research*, *28*(10), 2773–2782. https://doi.org/10.1029/92WR01135

Balistrocchi, M., & Bacchi, B. (2017). Derivation of flood frequency curves through a bivariate rainfall distribution based on copula functions: Application to an urban catchment in northern Italy's climate. *Hydrology Research*, *48*(3), 749–762. https://doi.org/10.2166/nh.2017.109

Balistrocchi, M., & Grossi, G. (2020). Predicting the impact of climate change on urban drainage systems in northwestern Italy by a copula-based approach. *Journal of Hydrology: Regional Studies*, *28*, 100670. https://doi.org/10.1016/j.ejrh.2020.100670

Balistrocchi, M., Moretti, G., Orlandini, S., & Ranzi, R. (2019). Copula-based modeling of earthen levee breach due to overtopping. *Advances in Water Resources*, *134*, 103433. https://doi.org/10.1016/j.advwatres.2019.103433

Balistrocchi, M., Orlandini, S., Ranzi, R., & Bacchi, B. (2017). Copula-based modeling of flood control reservoir. *Water Resources Research*, *53*(11), 9883–9900. https://doi.org/10.1002/2017wr021345

Bayoumi, A., & Meguid, M. A. (2011). Wildlife and safety of earthen structures: A review. *Journal of Failure Analysis and Prevention*, *11*(4), 295–319. https://doi.org/10.1007/s11668-011-9439-y

Benson, M. A. (1962). Evolution of methods for evaluating the occurrence of floods. In *Geological survey water-supply*. Washington, DC: United States Geological Survey. https://doi.org/10.3133/wsp1580A

Benson, M. A. (1963). Factors influencing the occurrence of floods in a humid region of diverse terrain. In *Geological survey water-supply*. Washington, DC: United States Geological Survey. https://doi.org/10.3133/wsp1580B

Bezak, N., Brilly, M., & Šraj, M. (2014). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, *59*(5), 959–977. https://doi.org/10.1080/02626667.2013.831174

Borgatti, L., Forte, E., Mocnik, A., Zambrini, R., Cervi, F., Martinucci, D., et al. (2017). Detection and characterization of animal burrows within river embankments by means of coupled remote sensing and geophysical techniques: Lessons from river Panaro (northern Italy). *Engineering Geology*, *226*, 277–289. https://doi.org/10.1016/j.enggeo.2017.06.017

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A.-C., & Seibert, J. (2017). Flood type specific construction of synthetic design hydrographs. *Water Resources Research*, *53*(2), 1390–1406. https://doi.org/10.1002/2016WR019535

Butera, I., Climaci, M., & Tanda, M. G. (2020). Numerical analysis of phreatic levels in river embankments due to flood events. *Journal of Hydrology*, *590*, 125382. https://doi.org/10.1016/j.jhydrol.2020.125382

Calamak, M., Larocque, L. A., & Chaudhry, M. H. (2020). Numerical modelling of seepage through earthen dams with animal burrows: A case study. *Journal of Hydraulic Research*, *59*, 488–499. https://doi.org/10.1080/00221686.2020.1780502

Camici, S., Barbetta, S., & Moramarco, T. (2017). Levee body vulnerability to seepage: The case study of the levee failure along the Foenna stream on 1 January 2006 (Central Italy). *Journal of Flood Risk Management*, *10*(3), 314–325. https://doi.org/10.1111/jfr3.12137

Carter, J., & Leonard, B. P. (2002). A review of the literature on the worldwide distribution, spread of, and efforts to eradicate the coypu (Myocastor coypus). *Wildlife Society Bulletin*, *30*(1), 162–175.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). *Applied hydrology*. Singapore, SG: McGraw-Hill.

Chowdhary, H., Escobar, L. A., & Singh, V. P. (2011). Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrology Research*, *42*(2–3), 193–216. https://doi.org/10.2166/nh.2011.065

Council of Europe (1979). *Convention on the Conservation of European Wildlife and Natural Habitats* (Vol. 104). European Treaty Series No. Bern.

Council of Europe (1992). *Directive 92/43/CEE of the Council of 21 May 1992 on the conservation of natural habitats and of natural fauna and flora*. Official Journal of European Communities.

D'Alpaos, L., Brath, A., Fioravante, V., Gottardi, G., Mignosa, P., & Orlandini, S. (2014). Causes of the levee failure occurred on the Secchia River at San Matteo on January 19, 2014. In *Final Report, Reg. Emilia-Romagna Comm*.

Dassanayake, S. M., & Mousa, A. (2020). Probabilistic stability evaluation for wildlife-damaged earth dams: A Bayesian approach. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *14*(1), 41–55. https://doi.org/10.1080/17499518.2018.1542499

Davidson, T. M., Altieri, A. H., Ruiz, G. M., & Torchin, M. E. (2018). Bioerosion in a changing world: A conceptual framework. *Ecology Letters*, *21*(3), 422–438. https://doi.org/10.1111/ele.12899

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., & Rosso, R. (2005). Bivariate statistical approach to check adequacy of dam spillway. *Journal of Hydrologic Engineering*, *10*(1), 50–57. https://doi.org/10.1061/(asce)1084-0699(2005)10:1(50)

Diersch, H. -J. G. (2014). *FEFLOW: Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media* (p. 996). Springer.

Federal Emergency Management Agency (2005). *Technical Manual for Dam Owners: Impacts of Animals on Earthen Dams* (Vol. 473). Washington, DC: FEMA.

Fisher, N. I., & Switzer, P. (2001). Graphical assessment of dependence: Is a picture worth 100 tests? *The American Statistician*, *55*(3), 233–239. https://doi.org/10.1198/000313001317098248

Ganguli, P., & Reddy, M. J. (2013). Probabilistic assessment of flood risks using trivariate copulas. *Theoretical and Applied Climatology*, *111*(1–2), 341–360. https://doi.org/10.1007/s00704-012-0664-4

Genest, C., & Favre, A. C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, *12*(4), 347–368. https://doi.org/10.1061/(asce)1084-0699(2007)12:4(347)

Genest, C., Remillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, *44*(2), 199–213. https://doi.org/10.1016/j.insmatheco.2007.10.005

Goel, N. K., Seth, S. M., & Chandra, S. (1998). Multivariate modeling of flood flows. *Journal of Hydraulic Engineering*, *124*(2), 146–155. https://doi.org/10.1061/(asce)0733-9429(1998)124:2(146)

Grimaldi, S., & Serinaldi, F. (2006). Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*, *29*(8), 1115–1167. https://doi.org/10.1016/j.advwatres.2005.09.005

Hall, J. W., Dawson, R. J., Sayers, P. B., Rosu, C., Chatterton, J. B., & Deakin, R. (2003). A methodology for national-scale flood risk assessment. *Proceedings of the ICE - Water and Maritime Engineering*, *156*(3), 235–247. https://doi.org/10.1680/wame.2003.156.3.235

Hall, J. W., Sayers, P. B., & Dawson, R. J. (2005). National-scale assessment of current and future flood risk in England and Wales. *Natural Hazards*, *35*(1–2), 147–167. https://doi.org/10.1007/s11069-004-4546-7

Harvey, G. L., Henshaw, A. J., Brasington, J., & England, J. (2019). Burrowing invasive species: An unquantified erosion risk at the aquatic-terrestrial interface. *Reviews of Geophysics*, *57*(3), 1018–1036. https://doi.org/10.1029/2018RG000635

Haubrock, P. J., Inghilesi, A. F., Mazza, G., Bendoni, M., Solari, L., & Tricarico, E. (2019). Burrowing activity of Procambarus clarkii on levees: Analysing behaviour and burrow structure. *Wetlands Ecology and Management*, *27*(4), 497–511. https://doi.org/10.1007/s11273-019-09674-3

Hu, L., Nikolopoulos, E. I., Marra, F., & Anagnostou, E. N. (2020). Sensitivity of flood frequency analysis to data record, statistical model, and parameter estimation methods: An evaluation over the contiguous united states. *Journal of Flood Risk Management*, *13*(1), e12580. https://doi.org/10.1111/jfr3.12580

Joe, H. (1997). *Multivariate models and dependence concepts*. Chapman and Hall.

Joe, H. (2015). *Dependence modeling with copulas*. Boca Raton, FL: CRC Press.

Karmakar, S., & Simonovic, S. P. (2009). Bivariate flood frequency analysis part 2: A copula-based approach with mixed marginal distributions. *Journal of Flood Risk Management*, *2*(1), 32–44. https://doi.org/10.1111/j.1753-318X.2009.01020.x

Kite, G. W. (1977). *Frequency and risk analysis in hydrology*. Littleton, CO: Water Resources Publications, LLC.

Kottegoda, N. T., & Rosso, R. (2008). *Applied statistics for civil and environmental engineers* (2nd ed.). Blackwell Publishing.

Kunstmann, H., & Kastens, M. (2006). Direct propagation of probability density functions in hydrological equations. *Journal of Hydrology*, *325*(1–4), 82–95. https://doi.org/10.1016/j.jhydrol.2005.10.009

Lang, M., Ouarda, T. B. M. J., & Bobée, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, *225*(3–4), 103–118. https://doi.org/10.1007/978-1-59259-261-6_9

Li, Y., Dong, Z.-Y., Pan, D.-Z., & Pan, C.-H. (2020). Effects of subterranean termite nest architectures on earth embankment seepage and stability. *Paddy and Water Environment*, *18*(2), 367–384. https://doi.org/10.1007/s10333-019-00788-1

Mai, J. -F., & Scherer, M. (2012). *Simulating Copulas. Vol. 4 of series in quantitative finance*. Imperial College Press.

Marchi, E. (1961). Sulla filtrazione attraverso gli argini fluviali. In *Proceedings of the VII Convegno di Idraulica e Costruzioni Idrauliche* (p. 8).Palermo, IT.

McMahon, B. J., Doyle, S., Gray, A., Kelly, S. B. A., & Redpath, S. M. (2019). European bird declines: Do we need to rethink approaches to the management of abundant generalist predators? *Journal of Applied Ecology*, *57*(10), 1885–1890. https://doi.org/10.1111/1365-2664.13695

Michelazzo, G., Paris, E., & Solari, L. (2018). On the vulnerability of river levees induced by seepage. *Journal of Flood Risk Management*, *11*(S2), S677–S686. https://doi.org/10.1111/jfr3.12261

Mori, E., & Assandri, G. (2019). Coming back home: Recolonisation of abandoned dens by crested porcupines hystrix cristata and european badgers meles meles after wood-cutting and riparian vegetation mowing events. *Hystrix—The Italian Journal of Mammalogy*, *30*(1), 39–43. https://doi.org/10.4404/hystrix-00176-2019

Natural Environment Research Council. (1975). *Flood studies report*. Natural Environment Research Council.

Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). New York, NY: Springer.

Oliver, J., Qin, X. S., Larsen, O., Meadows, M., & Fielding, M. (2018). Probabilistic flood risk analysis considering morphological dynamics and dike failure. *Natural Hazards*, *91*(1), 287–307. https://doi.org/10.1007/s11069-017-3126-6

Orlandini, S., Moretti, G., & Albertson, J. D. (2015). Evidence of an emerging levee failure mechanism causing disastrous floods in Italy. *Water Resources Research*, *51*(10), 7995–8011. https://doi.org/10.1002/2015WR017426

Palladino, M. R., Barbetta, S., Camici, S., Claps, P., & Moramarco, T. (2020). Impact of animal burrows on earthen levee body vulnerability to seepage. *Journal of Flood Risk Management*, *13*(S1). https://doi.org/10.1111/jfr3.12559

Qi, W., Zhang, C., Fu, G., & Zhou, H. (2016). Imprecise probabilistic estimation of design floods with epistemic uncertainties. *Water Resources Research*, *52*(6), 4823–4844. https://doi.org/10.1002/2015WR017663

Requena, A. I., Mediero, L., & Garrote, L. (2013). A bivariate return period based on copulas for hydrologic dam design: Accounting for reservoir routing in risk estimation. *Hydrology and Earth System Sciences*, *17*(8), 3023–3038. https://doi.org/10.5194/hess-17-3023-2013

Romano, N., & Santini, A. (1999). Determining soil hydraulic functions from evaporation experiments by a parameter estimation approach: Experimental verifications and numerical studies. *Water Resources Research*, *35*(11), 3343–3359. https://doi.org/10.1029/1999wr900155

Saghaee, G., Mousa, A. A., & Meguid, M. A. (2016). Experimental evaluation of the performance of earth levees deteriorated by wildlife activities. *Acta Geotechnica*, *11*(1), 83–93. https://doi.org/10.1007/s11440-015-0373-0

Salvadori, G., & De Michele, C. (2010). Multivariate multiparameter extreme value models and return period: A copula approach. *Water Resources Research*, *46*, W10501. https://doi.org/10.1029/2009wr009040

Salvadori, G., De Michele, C., Kottegoda, N. T., & Rosso, R. (2007). *Extremes in nature: An approach using copulas*. Dordrecht, The Netherlands: Springer.

Salvadori, G., Durante, F., Tomasicchio, G. R., & D'Alessandro, F. (2015). Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering. *Coastal Engineering*, *95*(6), 77–83. https://doi.org/10.1016/j.coastaleng.2014.09.007

Serinaldi, F. (2015). Dismissing return periods. *Stochastic Environmental Research and Risk Assessment*, *29*(4), 1179–1189. https://doi.org/10.1007/s00477-014-0916-1

Serinaldi, F. (2016). Can we tell more than we can know? The limits of bivariate drought analyses in the United States. *Stochastic Environmental Research and Risk Assessment*, *30*(6), 1691–1704. https://doi.org/10.1007/s00477-015-1124-3

Serinaldi, F., & Kilsby, C. G. (2013). The intrinsic dependence structure of peak, volume, duration, and average intensity of hyetographs and hydrographs. *Water Resources Research*, *49*(6), 3423–3442. https://doi.org/10.1002/wrcr.20221

Sklar, A. (1959). *Fonctions de Répartition à n Dimensions et Leures Marges* (Vol. 8, pp. 229–231). Publications de l'Institut Statistique de l'Université de Paris.

Todorovic, P. (1978). Stochastic models of floods. *Water Resources Research*, *14*(2), 345–356. https://doi.org/10.1029/wr014i002p00345

UK Centre for Ecology & Hydrology. (1999). Statistical procedures for flood frequency estimation. In *Flood estimation handbook*. (Vol. 3). Wallingford, UK: UK Centre for Ecology & Hydrology.

US Army Corps of Engineers. (1996). Risk-based analysis for flood damage reduction studies. In *Engineer manual 1110-2-1619*. Washington, DC: US Army Corp of Engineers.

Volpi, E., & Fiori, A. (2014). Hydraulic structures subject to bivariate hydrological loads: Return period, design, and risk assessment. *Water Resources Research*, *50*, 885–897. https://doi.org/10.1002/2013wr014214

Vorogushyn, S., Apel, H., & Merz, B. (2011). The impact of the uncertainty of dike breach development time on flood hazard. *Physics and Chemistry of the Earth*, *36*(7–8), 319–323. https://doi.org/10.1016/j.pce.2011.01.005

Vorogushyn, S., Merz, B., & Apel, H. (2009). Development of dike fragility curves for piping and micro-instability breach mechanisms. *Natural Hazards and Earth System Sciences*, *9*(4), 1383–1401. https://doi.org/10.5194/nhess-9-1383-2009

Woodward, D. K., & Mayfield, S. M. (1999). A survey of ASDSO/ICODS representatives on animal damage to Earth filled dams and appurtenances. In *ASDSO/FEMA Specialty Workshop on ''Plant and Animal Penetrations of Earth Filled Dams''*. Knoxville, TN: University of Tennessee Conference Center.

Wycoff, R. L., & Singh, U. P. (1976). Preliminary hydrologic design of small flood detention reservoirs. *Water Resources Bulletin*, *12*(2), 337–349. https://doi.org/10.1111/j.1752-1688.1976.tb02683.x

Yue, S. (2000). The bivariate lognormal distribution to model a multivariate flood episode. *Hydrological Processes*, *14*(14), 2575–2588. https://doi.org/10.1002/1099-1085(20001015)14:14<2575::aid-hyp115>3.0.co;2-1

Zhang, L., & Singh, V. P. (2007). Trivariate flood frequency analysis using the Gumbel-Hougaard copula. *Journal of Hydrologic Engineering*, *12*(4), 431–439. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(431)