Sensitivity analysis of erosion on the landward slope of an earthen flood defence located in southern France submitted to wave overtopping

Clément Lutringer[1,2], Adrien Poupardin[1], Philippe Sergent[3], Abdelkrim Bennabi[1], and Jena Jeong[1,2]

[1]ESTP Paris, 28 Avenue du Président Wilson, 94230 Cachan, France

[2]Gustave Eiffel University, Cité Descartes, 77200 Marne-la-Vallée, France

[3]CEREMA Risques Eau Mer, 134 rue de Beauvais, 60280 Margny-les-Compiègne, France

Correspondence: Clément Lutringer (clutringer@estp-paris.eu)

Abstract. The study aims to provide a complete analysis framework applied to an earthen dyke located in Camargue, France. This dyke is regularly submitted to erosion on the landward slope that needs to be repaired. Improving the resilience of the dyke calls for a reliable model of damage frequency. The developed system is a combination of copula theory, empirical wave propagation and overtopping equations as well as a global sensitivity analysis in order to provide the return period of erosion damage on a set dyke while also providing recommendations in order for the dyke to be reinforced as well the model to be self-improved. The global sensitivity analysis requires to calculate a high amount of return periods over random observations of the tested parameters. This gives a distribution of the return periods, providing a more general approach on the behavior of the dyke. The results show a return period peak around the two-year mark, close to reported observation. The distribution being skewed, the mean value is however higher and is thus less reliable as a measure of dyke safety. The sensitivity analysis shows that the geometrical characteristics of the dyke - slope angles and dyke height - are the ones carrying the highest amount of uncertainty into the system, showing that maintaining a homogeneous dyke is of great importance. Some empirical parameters intervening inside the propagation and overtopping process are also fairly uncertain and suggest that using more robust methods at their corresponding steps could improve the reliability of the framework. The obtained return periods have been confirmed by current *in situ* observations but the uncertainty increases for the most severe events due to the lack of long-term data.

1 Introduction

The site of the Salin-de-Giraud located in the Camargue area in southern France is a historically low-lying region and is thus frequently exposed to numerous storms. The latest Intergovernmental Panel on Climate Change report (Pörtner et al., 2022) points a general increase in variability of extreme events. Storm surges are expected to become more violent and the climate generally more uncertain, meaning that correctly designing structures to withstand rare events is becoming more difficult than ever. In fact, all the infrastructures on the site as well as the land itself must be maintained in order to ensure its exploitation and new methods must be applied in order to keep the maintenance cost at a reasonable level. An earthen dyke, named Quenin, has been constructed on the site in order to protect the salt marshes during storm surges. The structure is quite large, covering a few kilometers along the coastline. The dyke is approximately 2 meters high with large rocks on the seaward slope while the landward slope is only covered by sand. A picture is displayed in (fig. 1).

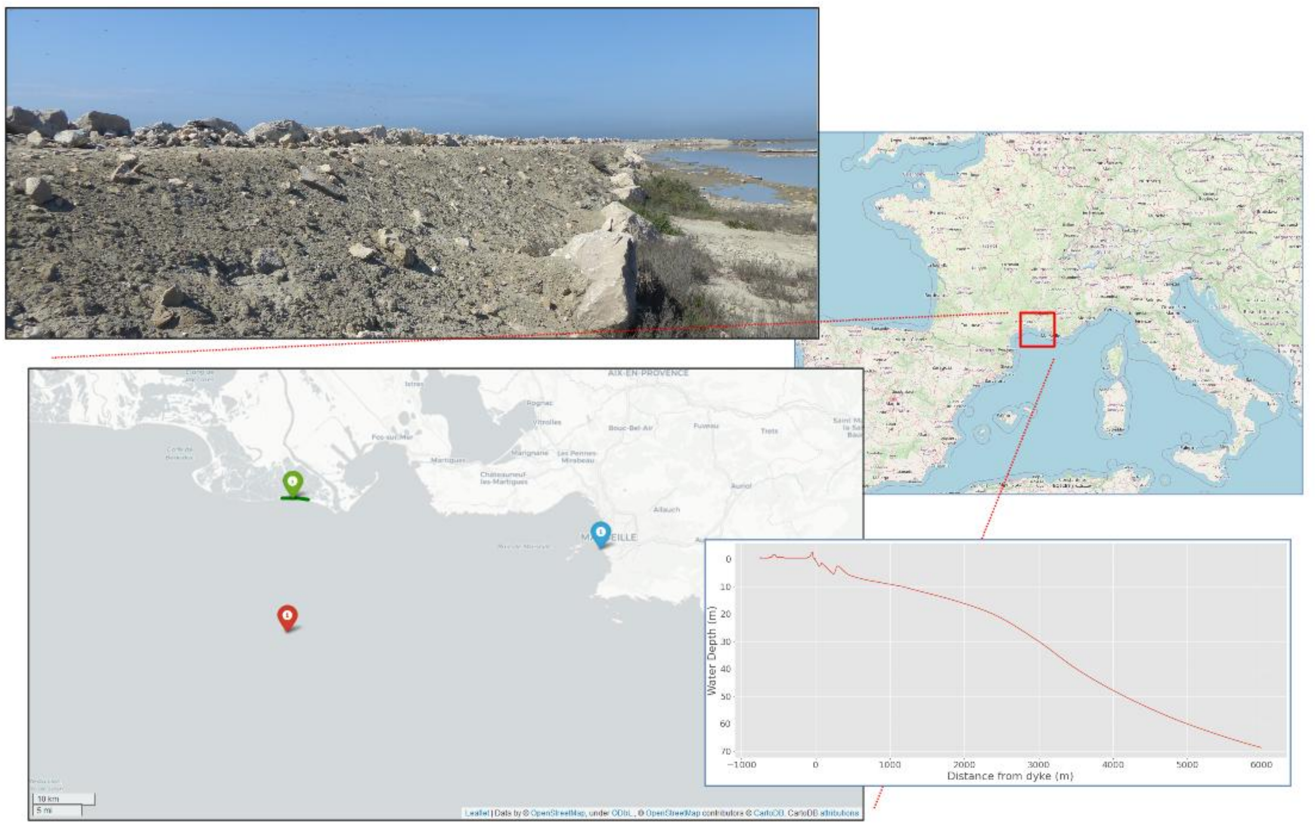

Figure 1. Map of the different locations used to collect data : The bathymetry (bottom right) is extracted at the location of the dyke (in green), The water elevation data comes from the Marseille historical gauge (blue) and the significant wave height is calculated offshore (red)

The erosion problem of the dyke is common in this area and therefore assessment of erosion is necessary. The semi-empirical approach based on the hydraulic loading has been well established and traditionally used. Wave propagation from deep water to the surf-zone has been well explored both analytically, numerically and experimentally in the literature. A large overview of theory surrounding random sea wave propagation theory was provided by Goda (2000) and brought advices on coastal protection. An evaluation of the different methods available on the subject has also been given by Liu and Han (2017). The complex nature of the overtopping phenomena makes it more difficult to model using only simple equations. Thus many large scale experiments have been conducted to deduce empirical laws as done by van der Meer (2011) as well as Hughes and Nadal (2008); Hughes et al. (2012) with use of the Wave Overtopping Simulator. Numerical simulations have also been explored by Li et al. (2003) using the Volume of Fluid method. More recently, the EurOtop manual (van der Meer et al., 2018) laid an extensive set of recommendations and experimentally based equations in order to functionally model the overtopping phenomenon. Bergeijk et al. (2019) also provided a more refined analytical model of overtopping using a set of coupled equations validated by numerical simulations and experiments.

Regarding the statistical tool to predict a higher risk, copula theory has been well accepted and used to calculate multivariate return periods of natural hazards. De Michele et al. (2007); Bernardara et al. (2014) wrote extensively on the subject with guidelines on using copulas to predict storm surges. More specifically, Kole et al. (2007) found that the Student's and Gumbel copulas are particularly interesting for risk management applications. Liu and Han (2017) deemed that the Clayton and Gumbel copulas are to be preferred for calculating multivariate joint return periods of natural hazards. Many methods are currently in use when estimating the probability of storm surges from sea states such as numerical models (SWAN or SWASH for example) as well as models based on wave energy. However, a more statistical approach based on bivariate copulas combining wave height and sea elevation are also widely used as they are reliable and computationally inexpensive, as seen in Salvadori and Michele (2007) but Orcel et al. (2020) expanded the method to trivariate copulas, allowing the method to yield the probability of structural failure. As indicated by many sources, we have a large choice of different copulas to link our different deep water conditions (Durante and Sempi, 2010, 2016; Tootoonchi et al., 2022). Among them, the Gumbel-Hougaard Copula is commonly used to link the Still Water Level to the wave height as done by (Wahl et al., 2010; Chini and Stansby, 2006). As mentioned by Orcel et al. (2020), this will lead to the calculation of an "and" return period, yielding the expected mean time between two events where all metrics overreach a certain level (as opposed to a "or" return period where only one metric needs to overreach). However, there are very few researches on assessment of erosion of dyke combining statistical and probability approach and theoretical and semi-empirical approach as well. Mehrabani and Chen (2015) worked on joint probabilistic approach for assessment of climate change effect on hydraulic loading. However, the authors constrained themselves to the frame of copula theory, assessing the risk to offshore conditions. That approach has not considered an interaction with a dyke nor propagation of deep water wave, but used a physical erosion criteria to put a threshold metric. In the present study, we used global sensitivity analysis to assess the most important parameters in the framework as the ones that contribute the most to the variance of the system in order to provide self-improvement to the framework as well as recommendations to improve the resiliency of the dyke. Combining different approaches, sensitivity analysis, a fully functional and modular overtopping framework and copula theory into a full stack has not been explored before and might provide use for the practitioner. Most works that laid the foundation for the last EurOtop manual (van der Meer et al., 2018) did not go further than predicting wave behavior up to overtopping but do not go further than this point. It makes sense as the focal point of such study is often led by damages on infrastructures laid behind the dyke. However, providing this extra step allows quantification of the erosion damages provoked on the dyke itself which is the main focus here as salt marshes do not bear costly infrastructures to protect. Also, erosion damages is often easier to observe and quantify than the overtopping phenomenon which is quick, volatile, and difficult to measure on-site. The second section will describe the data used in the study. The third section will be focused on the methodology of the article, the most important equations regarding both the physical wave process and the statistical processes.

Results are presented in the fourth section followed by discussions on the advantages and shortcomings of the study as well as future potential improvements in the fifth section.

2 Data

The statistical study of coastal events requires relatively large, well-documented, high-quality datasets. Such historical data is not easy to find even in an area containing a dense network of coastal sensors. As a unified database of all records regarding offshore and coastal characteristics does not exist, we used data coming from different bases which contained the measures of interest with correct time synchronicity. We present the data in this section.

2.1 Bathymetry

We have at our disposal the bathymetry of the dyke up to the deepwater point provided by the SHOM. The survey has a resolution of 3 meters. This is precise enough to calculate the mean slope of the beach in front of the dyke. The data is displayed in (fig. 1)

2.2 Sea level records

As there is no sensor that recorded the sea level in the immediate vicinity of the dyke, which would be highly sensitive to waves anyway, we had to resort to the nearest gauge that had large record of measures, which was located in Marseilles's harbour (fig. 1). The data of the gauge is maintained by the SHOM in the REFMAR database which is part of the Global Sea Level Observing System (GLOSS) and provides more than 100 years of hourly water elevation level. The acquisition is done using a permanent GNSS station. The place being located inside a port is protected from sea waves.

2.3 Significant Wave Height records

In situ data of the significant weight height provided at a hourly rate are difficult to find reliably over a long period of time (decades). This means that we have to resort to data provided by a numerical model. We use the data extracted from the ANEMOC-2 database currently maintained by the CEREMA[1], reproducing numerically the sea conditions over a long period of time (from 1980 to 2010). The significant wave height is estimated by calculating the mean value of the upper third of the recorded waves every hour. Thus, one value is given hourly at each chosen location. We have selected point 3667 (fig. 1) as it is both in front of dyke and located where the water depth is high enough to be considered offshore ($\approx 80m$).

2.4 Identifying storm surges and coupling the data sets

The time series data itself is not directly exploitable as the copula that we want to generate is based on the identification of extreme events implying a locally high value of both N and H_0 that we will call here "storms". We use the same protocol as in Kergadallan (2015) which is :

[1] https://www.cerema.fr/fr

- Search for high peak values V_i on data set A (the water level for example) with the peak-over threshold method, the height threshold is chosen using the methodology described in Bernardara et al. (2014);
- Associate each value V_i with its time of occurrence t_i;
- Define a time window Δt which would be the expected mean duration of a storm;
- For each peak, look for the maximum value W_i of data set B (the significant wave height for example) during the $[t_i - \Delta t/2; t_i + \Delta t/2]$;
- create the couple (V_i, W_i) as the characteristics for storm i.

An example of the method is given in (fig. 2). The peak is identified on the significant wave height sample as a local maximum. Then, the zone is defined around the peak and the local max is searched in the same time interval on the water level data.

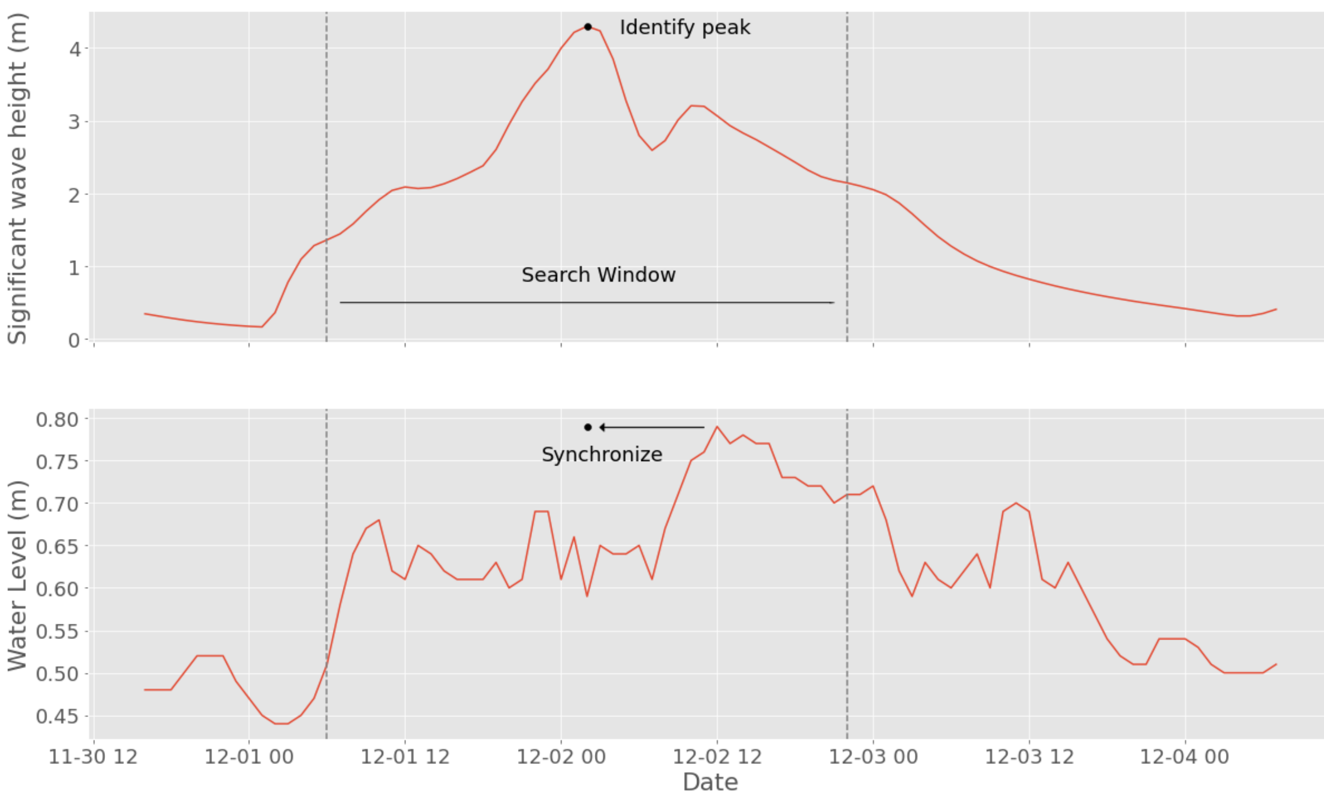

Figure 2. Illustration of the storm identification process in two steps. The peak is located on the significant wave height (black dot). A Search Window is defined around this point in which the maximum is found in (b). The Search Window used here is 38 hours.

One could wonder which data set should be used first in the protocol. The basic recommendation is to identify the peaks the physically most important contributor to the event that we are looking for. However, we are only interested in the final distribution of data sets and the choice of the data to use first did not seem to impact the distribution in our case. As a preliminary check, the correlation coefficient between the two datasets was calculated using Pearson's coefficient gave $C_{corr} = 0.6$ which indicates a moderate positive relationship. It is probable that a lot of the dependency structure linking the two metrics is non-linear and could be better encapsulated by a copula-based approach, motivating the following methods.

3 Methods

3.1 Multivariate statistical theory using copulas

The copula method is a popular approach for estimating return periods of extreme events in hydrology or finance and is commonly used as a tool of risk-management as univariate statistical analysis might not be enough to provide reliable probabilities with correlated variables as stated by Chebana and Ouarda (2011). One key advantage of the copula method over other methods is that it allows more flexible modeling of the dependence structure between variables. Other methods, such as the traditional joint probability method or the design storm method, assume a specific type of dependence structure (e.g. independence or fixed correlation), which may not accurately reflect the true relationship between the variables. In addition, the copula method can provide more accurate estimates of extreme events and their return periods, especially for events with very low probabilities of occurrence. This is because the copula method allows for more precise estimation of the tail dependence between variables, which is important for accurately estimating extreme events. As we are looking for rather rare events in this study, this represents a considerable advantage. It however appears that the practitioner has a large selection of copulas to choose from depending of the nature of the data. The choice of which copula to choose varies from the type of data as well as the physics of the setup and even so we are left with a rather large selection. Merging multiple copulas in order to combine their properties has also be explored by Hu (2006), complicating further the decision process. Wahl et al. (2010) suggested that the Gumbel-Hougaard copula was particularly adapted when combining water level and wave intensity, although they used the time integral of the wave height over a threshold instead of the significant wave height and the region of interest was the North Sea. Orcel et al. (2020) also recommended using the Gumbel-Hougaard or Clayton copulas for coastal waves on the Atlantic shores of France. An application of the Gumbel-Hougaard copula has also been explored on UK shores aiming in an aim to study extreme coastal waves by Chini and Stansby (2006). This motivates us to directly use the Gumbel-Hougaard copula as the most adapted choice. The formula is mentionned in (eq. 1).

$$F(u,v\,|\,\theta) = \exp\left[-\left[(-\log(u))^{\theta} + (-\log(v))^{\theta}\right]^{1/\theta}\right] \tag{1}$$

where u and v are the cumulative distribution functions of the histograms originated from the data sets. The copula parameter θ represents the interdependency of the data.

The value of this copula parameter is important, and can be calculated using a panel of different methods, ie. the Error method (see Appendix B for the equation as written by Capel (2020)) and Maximum-Likelihood method.

Once done, the copula can be calculated using equation 1, attributing a probability of occurrence of any event E with one of the variables having a value smaller or equal to the defined ones, noted $E(H \leq h||N \leq n)$. The logical inverse $E(H > h, N > n)$ can then be obtained by calculating the survival copula C_θ^{-1} defined as :

$$C_\theta^{-1} = C_\theta + u + v - 1 \quad (2)$$

Finally, we can associate to each value of C_θ^{-1} a return period using the formula provided in Salvadori and Michele (2007) :

$$RP = \frac{\mu}{C_\theta^{-1}} \quad (3)$$

where μ is the average interarrival time between two events of interest i.e., the storms. The offshore conditions have been determined by a couple (N, H), the water level and the significant wave height respectively, with an associated return period. This gives us the properties of an offshore wave.

3.2 Maximum-Likelihood Method

The principle of the maximum-likelihood method that we use is that we try to maximize the function L in (eq. 4) yielding the likelihood of generating the observed data for a set value of θ. It essentially means that given a set of data, a high value of L indicates that the function is highly likely to have been able to generate the data sample.

$$L(\theta) = \sum_{i=0}^{n} c_\theta(u(i), v(i)) \quad (4)$$

where c_θ is the copula density, which can be obtained by calculating the derivative of the copula function with respect to its cumulative density functions in (eq. 5):

$$c_\theta(u, v) = \frac{\partial^2 C_\theta(u, v)}{\partial u \partial v} \quad (5)$$

3.3 Wave theory : from offshore to the critical velocity

We are able to link a deep water state to a return period. However, this does not give us any information on the probability of occurrence of an event that would provoke erosion. Hence, we need to assess what kind of event provokes erosion using equations 6 - 11 to calculate the terminal velocity of the flow on the landward slope.

3.3.1 Propagation

The offshore significant wave height can be propagated up to the toe of the dyke. Among the numerous methods, the most convenient to use is the propagation formula written in equation 6 extracted from Goda (2000) allowing us to calculate the significant wave height at the toe of the dyke H_{m0}, which is the mean of the third of the highest wave height over a set period of time, as follows. This metric is important as it is bound to be used for future calculations of the overtopping characteristics. Note that refraction is neglected in our case :

$$H_{m0} = \begin{cases} K_s H_0 & \text{for } \frac{d}{L_0} > 0.2 \\ \min\left[\beta_0 H_0 + \beta_1 d; \beta_{max} H_0; K_s H_0\right] & \text{for } \frac{d}{L_0} < 0.2 \end{cases} \tag{6}$$

where H_0 is the offshore wave height. K_s is the shoaling coefficient, d is the water depth at the toe of the dyke and L_0 is the deepwater wavelength. The coefficients β_0, β_1 and β_{max} can be calculated as detailed in Goda (2000)[2] :

3.3.2 Overtopping Equations

Once the wave reaches the toe of the dyke, the wave will start interacting with the dyke in what is called the overtopping phase. This phenomenon is divided into 3 steps with equations detailed in van der Meer et al. (2018). We give a brief summary here of the used equations :

- **Run-up** : The wave reaches the dyke and flows up towards the crest. The run-up height reached by 2% of the incoming waves is calculated using equation 7

$$RU_{2\%} = \gamma_f \cdot \gamma_\beta \cdot \left(4 - \frac{1.5}{\sqrt{\gamma_b \cdot \xi}}\right) \cdot H_{m0} \tag{7}$$

where ξ is the Irribarren Number, H_{m0} the wave height at the toe of the dyke. The γ factors γ_b, γ_f and γ_β yield the contribution of the berm, the roughness and porosity of the seaward slope and the obliquity of the waves, respectively.

- **Crest flow** : The water flows on the crest up to the landward slope. We calculate the flow velocity and thickness at the beginning of the crest using equations 8 and 9, respectively:

$$v_{A,2\%} = c_{v2\%} \left(g(RU_{2\%} - z_A)\right)^{0.5} \tag{8}$$

$$h_{A,2\%} = c_{h2\%}(RU_{2\%} - z_A) \tag{9}$$

[2] This method is convenient and easy to use but can be imprecise, especially if the deepwater steepness is highly irregular and not constantly positive. The results can then be confirmed using numerical simulations using a wave propagator such as Tomawac. Sergent et al. (2015) gave an estimation of the reliability of the simplified Goda modal compared to numerical methods (BEACH and SWAN for instance), they obtained a reasonable concordance for a steepness inferior to 7%, which corresponds to our case study.

With $c_{v2\%}$ and $c_{h2\%}$ arbitrary coefficients that are used as fitting parameters. Z_A is the height of the dyke above the still water level and g the gravitational acceleration. These equations where compiled in van der Meer (2011); van der Meer et al. (2012) from the works led by Shüttrumpf and van Gent (2003) and Lorke et al. (2012).

The flow velocity will then decay along the crest following equation 10, which is a function of distance from the seaward side of the crest (x_c). Note that this formula is only valid for a few meters long crest as the formula becomes less precise for higher values of x_c.

$$\frac{v_{2\%}(x_c)}{v_{2\%}(x_c=0)} = exp(-1.4x_c/L_0) \tag{10}$$

With $L_0 = g \cdot T_0^2$ the deep water wavelength of the incoming waves.

According to van der Meer et al. (2012), the decrease of flow thickness upon reaching the crest is about one third and can be attributed to the change of direction of the flow and stays relatively constant along the crest.

- **Landward slope flow** : The water trickles down the landward slope, this is where erosion usually happens. Since the maximum velocity of the flow is quickly reached on a slope such as ours, we directly use (eq. 11) to compute the terminal velocity of the flow.

$$v_b = \sqrt[3]{\frac{2 \cdot g \cdot h_{b0} \cdot v_{b0} \cdot \sin\beta}{f}} \tag{11}$$

with h_{b0} and v_{b0} are the flow thickness and velocity at the entry of the slope, respectively. f is the friction coefficient, which is determined experimentally when possible and estimated otherwise, g the gravity acceleration and β the slope angle.

These equations rely on a large number of parameters that are detailed in table 1.

Variable Name	Description	Value in (fig. 7)	Source
H_{dyke}	Height of the dyke	2.2	*in situ* data
f	Friction coefficient	0.02	EurOtop (2018)
β	Landward slope	30^o	*in situ* data
α	Seaward slope	30^o	*in situ* data
γ_f	Influence of roughness and porosity	0.6	EurOtop (2018)
γ_b	Influence of berm	1.0	EurOtop (2018)
d	Water depth at the toe of the dyke	0.54	*in situ* data
C_{h2}	Arbitrary coefficient of equation 9	0.2	EurOtop (2018)
C_{v2}	Arbitrary coefficient of equation 8	1.4	EurOtop (2018)

Table 1. Main control parameters in the equation system of the framework.

Defining the value of these parameters is not easy and they may carry some amount of uncertainty that needs to be quantified. We use sensitivity analysis to resolve this problem.

3.4 Return period of soil erosion

We can now associate a terminal velocity to a set $\mathcal{S}_{vt} = \{(N, H_0), f(N, H_0) = v_t\}$ that is the set of couples (N, H_0) which are associated through the function f to a terminal velocity v_t.

By integrating the derivative of the copula with respect to H_0 along the isoline $\mathcal{S}_{vt}$, we can obtain the return period of event $E_{vt} = \{v_t* > v_t\}$ which is any event implying a terminal velocity equal or higher than v_t (see equations 12 to 14).

$$P(v_t^* > v_t) = \iint\limits_{C} \left(\frac{\partial^2 C_{N,H_0}}{\partial N \partial H_0} \right) dN dH_0 \tag{12}$$

$$P(v_t^* > v_t) = \int\limits_0^\infty \left[\frac{\partial C_{N,H_0}}{\partial H_0} \right]_{\mathcal{S}(H_0)}^\infty dH_0 \tag{13}$$

$$P(v_t^* > v_t) = -\int\limits_0^\infty \left(\frac{\partial C_{N,H_0}}{\partial H_0} (\mathcal{S}(H_0), H_0) \right) dH_0 \tag{14}$$

Where C is the surface of integration, which is the area above the velocity curve and $S(H_0)$ the velocity curve. This means that we can calculate the return period associated with a certain terminal velocity threshold for a defined dyke by fixing the parameters in Table 1. We give reference values to these parameters. They are obtained either experimentally from *in situ* data or extracted from the literature when observations are unavailable. The details of the values are explained in subsection 3.5.1.

3.5 Sensitivity analysis through Quasi-Monte-Carlo process

3.5.1 Uncertainty Parameters

The showcased system is indeed able to provide return periods associated to events leading to erosion or any dangerous event defined as a criteria on flow velocity. However, added to the deep water conditions used to generate the copula, are the characteristics associated to the dyke as well as many empirical parameters used to fit the laws allowing the calculations leading to the landward terminal velocity of the dyke. All of these parameters carry an intrinsic amount of uncertainty which has a non-negligible impact on the results. This calls for an accurate quantification on the whole potential range of variation of each parameter. Global sensitivity analysis through the computation of global sensitivity indices will be our tool of choice. A combination of the 1-st order and total effect sensitivity indices defined in equations (17 - 18) is a principled and classical approach that encapsulates a useful enough amount of information on the variation of system's characteristics.

We estimate the value of the indices using the Saltelli estimator defined in Saltelli et al. (2008). The number of dimensions being high, we accelerate the convergence of the estimator using a pseudo-random sampler, in our case the Sobol sequence, which generates a low discrepancy sample of points. The resulting distribution of the parameters is thus uniform, which is standard for the Monte-Carlo method. The performance comparison of the Monte-Carlo process against the improved Quasi-Monte-Carlo estimations has been extensively discussed, noticeably in Sobol' (1990, 1998); Sobol´ and Kucherenko (2005); Acworth et al. (1998). The improvement in performance is unanimously in favour of the Quasi-Monte-Carlo Method.

The first step is to define the parameters used in equations 1 to 11 that we are going to consider as relevant sources of uncertainty. They are compiled into Table 2 where we associate a potential range of variation that is deemed as reasonable with its source. Each parameter is further described in its associated description below.

Variable Name	Description	Range of variation	Source of interval
H_{dyke}	Height of the dyke	$[1.89, 2.47]$	*in situ* data
f	Friction coefficient	$[0.01, 0.03]$	EurOtop (2018)
β	Landward Slope	$[20^o, 50^o]$	*in situ* data
α	Seaward Slope	$[20^o, 50^o]$	*in situ* data
γ_f	Influence of roughness and porosity	$[0.4, 0.8]$	EurOtop (2018)
γ_b	Influence of berm	$[0.75, 1.0]$	EurOtop (2018)
C_{h2}	Arbitrary coefficient in equation 9	$[0.1, 0.4]$	Bosman (2007) + Schüttrumpf (2001,2005)
C_{v2}	Arbitrary coefficient in equation 8	$[0.7, 2.1]$	Bosman (2007) + Schüttrumpf (2001,2005)
θ	Interdependency parameter (copula)	$[1.45, 1.75]$	Numerical Estimator
v_c	Critical erosion velocity	$[1.0, 4.0]$	Hughes (2012)
d	Water depth at the toe of the dyke	$[0.47, 0.82]$	*in situ* data
b_0	First coefficient of equation A1	$[0.028, 0.052]$	Goda (2000)
b_1	First coefficient of equation A2	$[0.52, 0.63]$	Goda (2000)

Table 2. Characteristics of the parameters used during the sensitivity analysis.

We also provide a brief description of the parameters as well as the estimation technique.

- **The height of the dyke** H_{dyke} is defined as the vertical distance between the still water level in a calm sea condition and the culminating point of the dyke. Using *in situ* data from a Litto3D bathymetry map, we managed to obtain the distribution of the dyke height. We use the mean of the heights as the reference value for tab. 1 and give an interval of variation of approximately one standard deviation for sensitivity analysis. The same procedure is done for the geometrical parameters α, β and d.

- **The friction coefficient** f yields the resistance of contact between two materials, in our case between the landward slope of the dyke and water. A higher coefficient brings a slower flow velocity but also more shear stress. Different values can be used here. It is generally considered that for smooth surfaces and vegetation, a value close to 0.02 can be used. We assume that is it possible to use such value of small rocks with diameter of approximately 20 cm, which is what is currently implemented on the Quenin dyke.

- **The landward slope** β is defined from the end of the crest which is considered as flat. The steeper is the slope, the higher is the terminal velocity. It should be noted that a combination of high crest velocity and steep landward slope can provoke a flow separation at the end of the crest followed by an impact on the slope, resulting in added normal stresses. This behaviour may be significant and has been explored by Ponsioen et al. (2011).

- **The seaward slope** α is defined as the mean slope from the toe of the dyke to the beginning of the crest, assuming that the crest is flat. Its value is important as the behaviour of the up-rushing wave may change drastically for different values of α.

- **The influence of roughness and porosity** on the seaward slope γ_f is a factor with value in the range 0 to 1 scaling how much the run-up will be attenuated thanks to the slope surface characteristics (1 means no influence). This is difficult to estimate as it relies on *in situ* experiments. Evaluating this parameter is not easy. Hence, we chose a relatively large range of variation around the reference value as the rocks on the slope are expected to have an influence of the same order of magnitude as other structures described in the EurOtop.

- **The influence of the berm** γ_b with value between 0 and 1 indicating the attenuation of the wave due to the presence of a berm. This value can be estimated using the geometry of the dyke if it is simple. It is more uncertain for a more complicated geometry. We calculate this factor using equations given in the EurOtop. The dyke is heterogeneous through its length and its geometry is more complicated than what is used for the calculation as it is a natural berm. Thus we gave it some variability deciding that it could not result in more than 25% water height reduction, which is already dramatic.

- **The depth at the toe of the dyke** b is calculated *in situ* using the Litto3D map as previously cited. Its value is registered for every transversal cross-section of the dyke.

- **The scaling coefficients of the input crest velocity and thickness** C_{h2} and C_{v2}, respectively, are scaling factors on the equations calculating the velocity and thickness of the flow at the beginning of the crest from the run-up. The range is estimated as a variation of $+/-50\%$ from their suggested values in the EurOtop (2018).

3.5.2 Sobol indices

If we provide our framework inputs that are uncertain, it should be expected that the uncertainty will be carried through the system up to the outputs. We rely on sensitivity analysis to quantify such uncertainty by comparing the influence of each parameter on the variation of the outputs relative to their respective range of variation. Since there may be a lot of interaction between parameters and we need to assess the influence of the parameters over their whole range of variation, we use global sensitivity analysis.

Let $Y = f(X_1, ..., X_n)$ be a function of the X_i parameters with $i = 1, ..., n$. The uncertainty of the parameters X_i will carry over the uncertainty of the output Y. Therefore, it would be necessary to estimate the impact of parameters on the output Y.

In order to quantify the influence of a single parameter X_i on a complex system, a good starting point can be to fix this parameter to a defined value x_i. Logically, freezing a parameter, which is a potential source of variation, should reduce the variance $V(Y)$ of the output Y. Hence, a small value of variance $V_{X_{\sim i}}(Y|X_i = x_i)$ would imply a high influence of the parameter X_i. We can globalize the approach by calculating the average value of the variance over all valid values of x_i, preventing the dependence on x_i. This is written as :

$$E_{X_i}(V_{X_{\sim i}}(Y|X_i = x_i)) < V(Y) \quad (15)$$

The following relation is also useful in our case :

$$E_{X_i}(V_{X_{\sim i}}(Y|X_i = x_i)) + V_{X_i}(E_{X_{\sim i}}(Y|X_i = x_i)) = V(Y) \quad (16)$$

The conditional variance $V_{X_i}(E_{X_{\sim i}}(Y|X_i = x_i))$ is called the first-order effect of X_i on Y. We can then use the sensitivity measure called the sensitivity index or Sobol index (see Sobol (2001)) defined as :

$$S_i = \frac{V_{X_i}(E_{X_{\sim i}}(Y|X_i = x_i))}{V(Y)} \quad (17)$$

This gives the proportion of contribution of the parameter X_i alone on the total variance of the output Y relatively to the other parameters $X_{\sim i}$. The main drawback of this measure is that the interaction of the parameters between themselves is not taken into account. These measures are contained in higher-order indices. However, this may become quite time-consuming and impractical if the number of parameters is high as the total number of Sobol indices that could be calculated grows as $n!$ with n the number of parameters.

The Total Effect Sobol' Index , which measures the influence of a parameter i on the variance as well as its interaction with every other parameters is calculated following the same method but instead of freezing parameter i, we freeze every other parameter $j \neq i$. The formula is given in eq. 18.

$$S_{Ti} = 1 - \frac{V_{X_i}(E_{X_{\sim i}}(Y|X_i))}{V(Y)} = \frac{E_{X_i}(V_{X_{\sim i}}(Y|X_i))}{V(Y)} \quad (18)$$

Although concise, equation 17 and 18 are difficult to calculate analytically. We circumvent the problem by using the method developed by Sobol (2001) and further improved by Saltelli et al. (2008). The protocole can be summarized as follows :

- Define the input parameter space and the model output function.
- Generate a set of samples using Latin hypercube sampling (LHS) or another quasi-random sampling method (we used Sobol' sequence in our case).
- Compute the model output for each set of input parameters.
- Partition the output variance into components due to individual input variables and their interactions using an ANOVA-based decomposition.
- Calculate the first-order and total-effect Sobol indices, which measure the contribution of individual input variables and their interactions to the output variance, respectively.

A diagram displaying the full method is shown in (fig. 3).

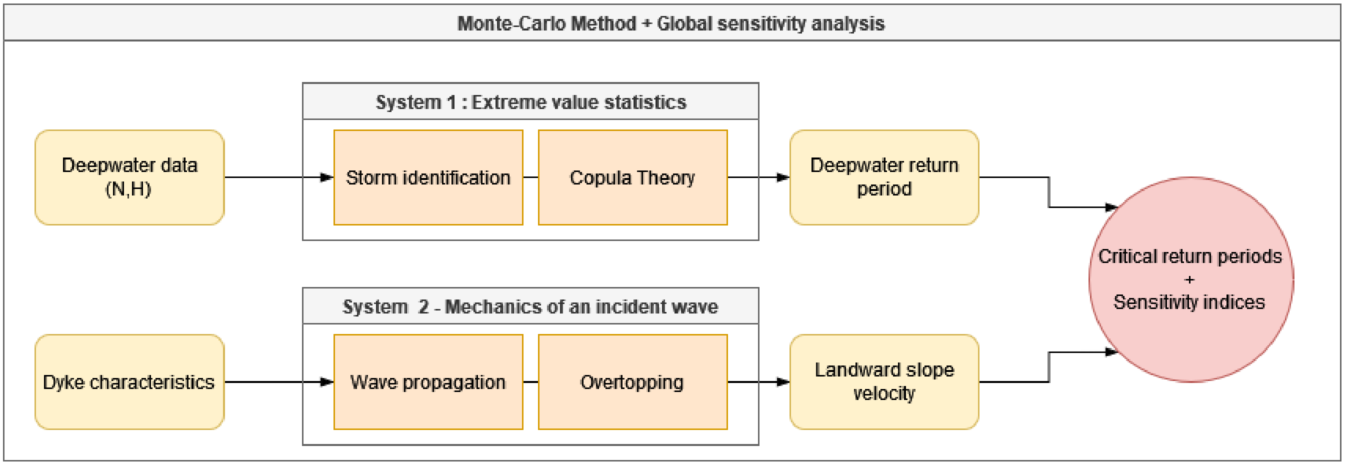

Figure 3. Diagram highlighting the main steps of the process as well as methods involved

4 Results

4.1 Return Periods Copula

We start by compiling the selected storm surge events into a histogram, giving the univariate probability densities of both datasets. However, since we only work with about 30 years of hourly data, we need to fit the cumulative histogram in order to create a cumulative distribution function that allows us to extrapolate to rarer events. We use the Generalized Extreme Value distribution which is used for the estimation of tail risks and is currently applied in hydrology for rainfalls and river discharges in the context of extreme events as in Muraleedharan et al. (2011).

This means that the events can then be sorted into an histogram for us to observe their respective distributions. In this case, the sample limits us to events that can happen up to once every 20 years since we have no data covering a larger period. In this case, we can obtain information about more extreme events by extrapolating the data using a fitted distribution. The Generalized Extreme Value distribution is particularly adapted for this kind of problem with cumulative distribution function formulated in (eq. 19).

$$[H]F(x) = \exp(t(x)) \qquad \text{with} \qquad t = \begin{cases} \left(1+\xi * \left(\frac{x-\mu *}{\sigma *}\right)\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ \exp\left(-\left(\frac{x-\mu *}{\sigma *}\right)\right) & \text{if } \xi = 0 \end{cases} \tag{19}$$

with (μ, σ, ξ) the location, scale and shape factor, respectively. The results are displayed in (fig. 4). The laws are fitted using the maximum-likelihood method. The fit gives a R^2 score higher than 0.999 for both data sets. Hence, we consider the fit almost perfect and will not be accounted for during future calculations of uncertainty as it is bound to be insignificant compared to other sources.

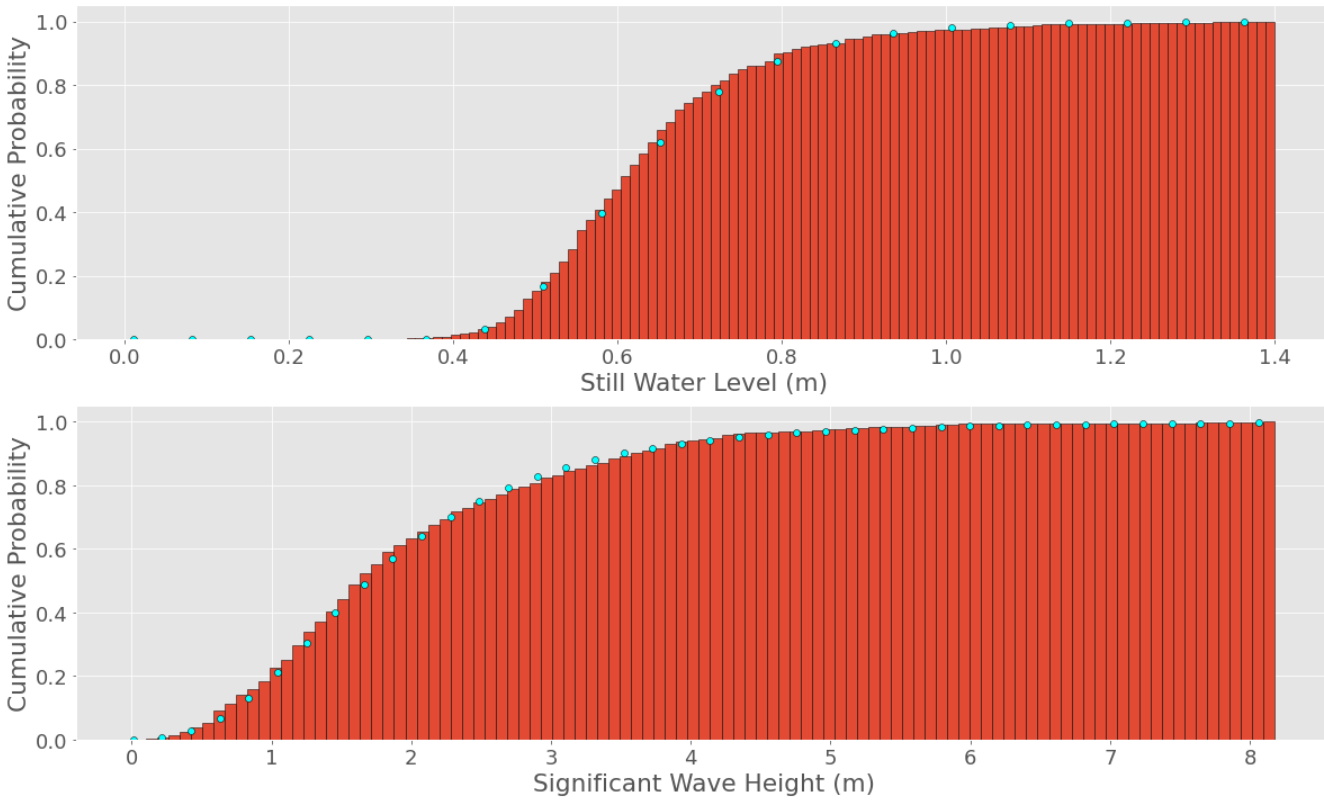

Figure 4. Cumulative Distribution Functions of the offshore significant wave height (a) and the still water level (b) displayed in a histogram. They are then fitted to a GEV distribution function using the maximum-likelihood method (blue dots)

We will then compute the derivative of the copula in (eq. 1) and maximize the value of $L(\theta)$ from (eq. 4). The variation of $L(\theta)$ is displayed in (fig. 5).

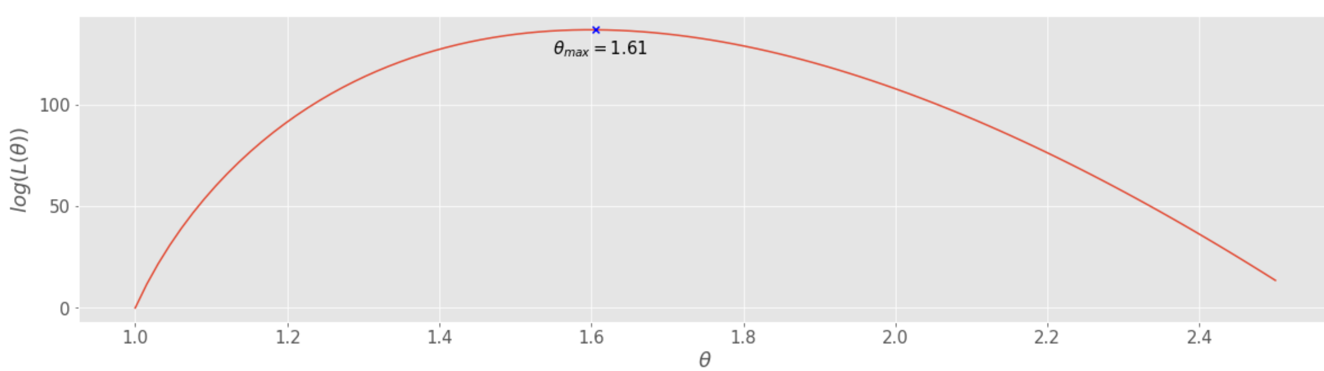

Figure 5. Value of the maximum-likelihood estimator with respect to the interdependence parameter θ.

The interdependency parameter can take values in the interval $[1, +\infty[$, where 1 is the independent copula and $+\infty$ means absolute correlation. A value of 1.6 means that there is a significant correlation between the two distributions. This can be seen visually on fig. 6 as the contour lines form an L-shape, indicating that a high value of water level is linked to a high value of significant wave height as increasing it does not affect much the return period. Hence, we can generate the copula using equation 1. The cumulative distribution function yields the probability of a value laying under a threshold. Hence, we use equation 2 to inverse the copula and obtain the survival copula (fig. 6). This allows us to evaluate the return period of any event E so that $E(N \leq x || H_0 \leq y)$ using (eq. 1-3).

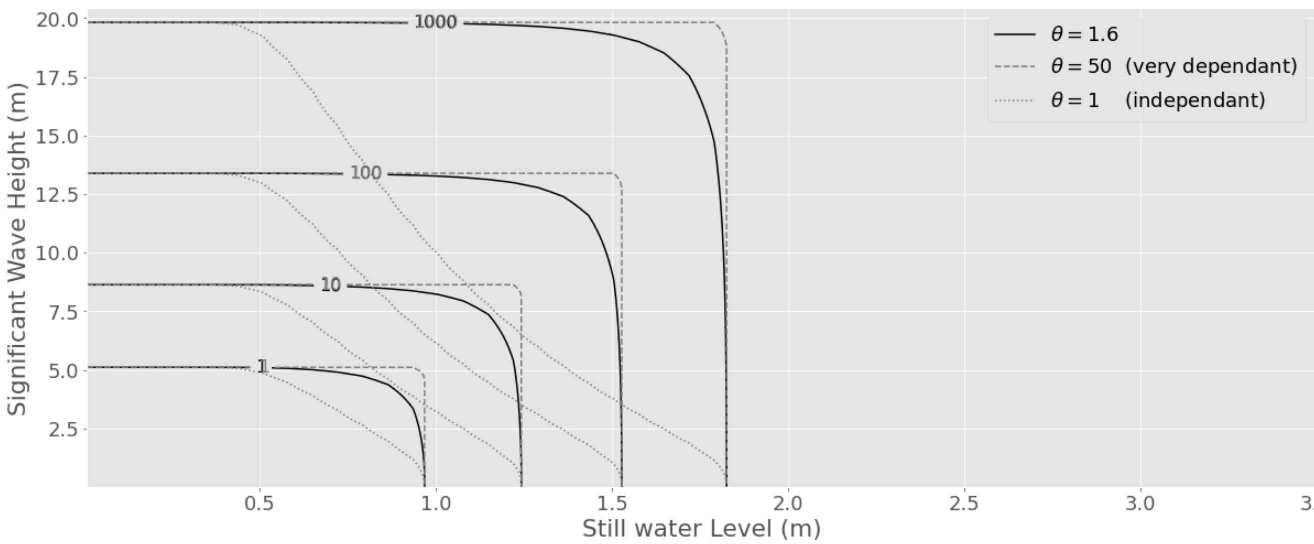

Figure 6. Return period of an event composed of a couple (N, H_0) or with a higher value of N or H_0, noted as $RP(E(N > x||H_0 > y))$ with the interdependence factor having a value of $\theta \approx 1.6$ (full line). The independant copula (dotted lines) and the high-dependence copula (dashed lines) are also displayed for reference.

The contour lines of the copula in (fig. 6) show that the data are coupled to some degree. Indeed, since the data are correlated, a both high value of the water level N and the significant wave height H should be more probable than if the data were uncorrelated, thus decreasing the return period and driving the contour lines towards the smaller values.

4.2 Computing the terminal velocity

We use the terminal velocity on the landward slope v_t as a criteria of erosion. Meaning that damage starts to occur when $v_t > v_c$ where v_c is the critical velocity which has to be determined using the literature. Using (eqs. 6 - 11), we can calculate it from any couple (N, H_0) of offshore water level and significant wave height, given that the mean slope of the bathymetry is known. The results are shown in (fig. 7).

Unsurprisingly, higher values of both N or H_0 induce higher values of terminal velocities. All values below the "0.0" line in (fig. 7) failed to produce overtopping and thus generate a null value while in fact there is no water flowing on the slope.

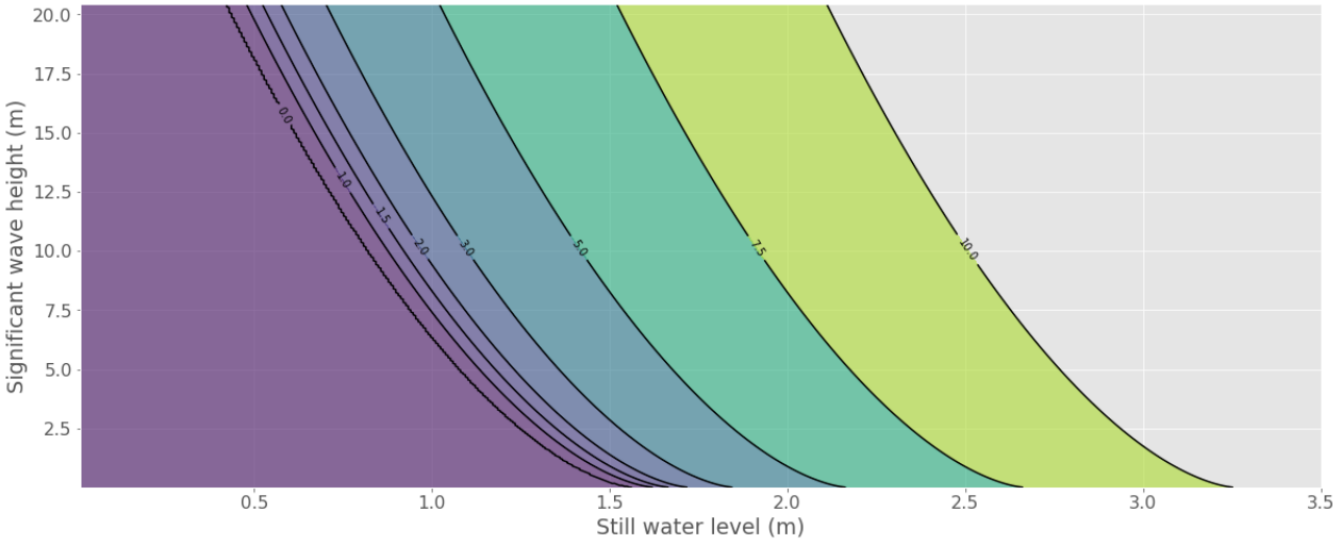

Figure 7. Terminal velocity (in $\mathrm{m/s}$)along the landward slope for any couple (N, H_0).

Typically, we observe that the Quenin dyke's landward slope is covered by rubble mounds which have an average diameter of 20 cm. Applying Peterka's formula (Peterka, 1958) (eq. 20) which is used by the U.S. Bureau of Reclamation, we can obtain the critical velocity of erosion on the dyke.

$$v^* = \sqrt{d_{50}/0.043} \tag{20}$$

where v^* is the critical erosion velocity and d_{50} is the median block parameter. For blocks with a diameter of 20 cm, we obtain a critical erosion velocity of approximately $2\ \mathrm{m/s}$. Our calculations estimate that such flow velocity will occur on average once every 5.86 years. This gives a higher value than what is reported by the Salins du Midi company, currently exploiting the dyke. The company reports significant damage that needs to be repaired approximately once every two years. This has been confirmed by its archives. This gap can be caused by uncertainty on the parameters which will be further estimated via sensitivity analysis.

4.3 Sensitivity indices

After generating a sample of parameter values, each set is computed through the framework, giving an associated return period from which we calculate the global sensitivity indices of both 1-st order and total effect. The results are compiled in (fig. 8).

The first observation is that some parameters contribute a lot more to the global variance of the system than others. Each parameter lies in four different categories to which we can attribute a degree of importance from the most important to the less important :

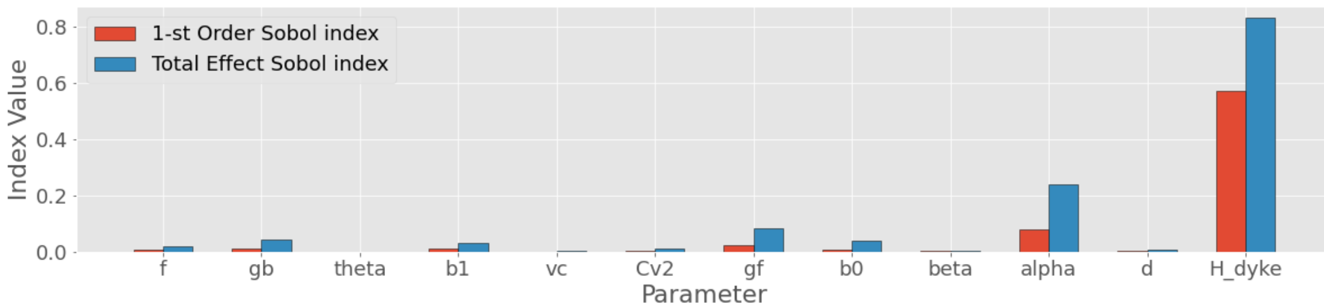

Figure 8. Value of the 1st-order sensitivity (in red) and total effect (in blue) indices for each tested parameter.

1. The parameters related to the geometrical features of the dyke (H_{dyke}, α, ...) seem to carry on average a lot of uncertainty and should be inspected thoroughly since some others don't look that important (β);

2. The parameters associated with the characteristics of the foreshore with parameters d and b_0 determine the initial behavior of the incoming wave. They are significant and should also be inspected;

3. The overtopping process relies on the intervention of many parameters which may have a some importance (γ_f, C_{v2}). Note that the chosen intervals focus on uncertainty, implementing a solution implying a high variation of a parameter outside of its considered range does not apply here;

4. The erosional process with parameter v_c however looks to be either well-defined or only mildly significant according to the values of the Sobol' indices.

Also, the values of the total effect indices seem to be much higher than for the 1-st order, which indicates that a great high amount of variance is hidden in higher-order indices, proving the presence of strong interactions between the parameters.

4.4 Return periods distribution

Launching such a high number of calculations allows us to compile the return periods into a histogram to evaluate the probability of the return periods taking into account uncertainties. The results are compiled in (fig. 9).

The results show that the distribution can be well fitted using a Generalized Extreme Value distribution which is right-skewed with a peak around the two years value and a long tail in the upper range of the return periods. The choice of the GEV is motivated by its common use to model extreme events such as floods, winds, and erosion. Furthermore, the root mean square error was lower on this fit than with other tested distribution such as the Lognormal or Pareto distribution. The mean value is close to ten years. This is high compared to what is expected from actual *in situ* records. However, the distribution is skewed. The median is of approximately 5 years, which is closer to records.

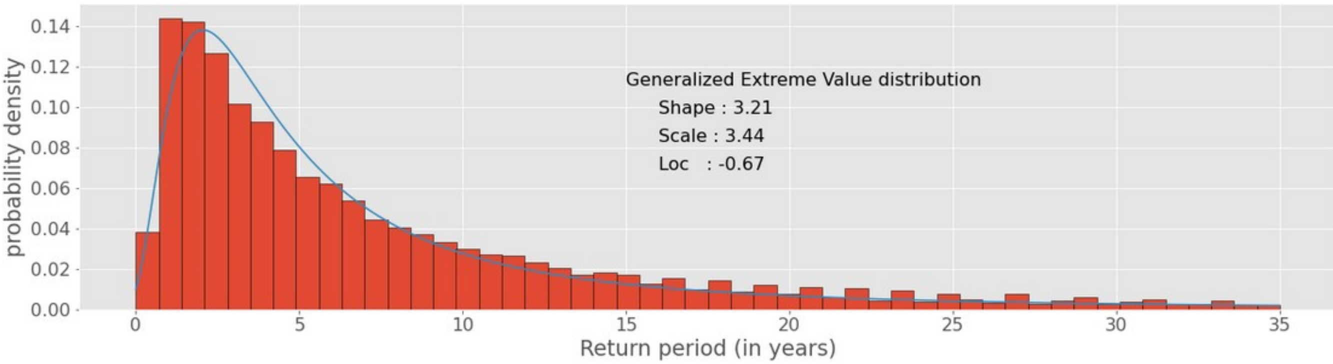

Figure 9. Distribution of the return periods of an event able to provoke some amount of erosion to landward slope at the dyke with random variation of the parameters in Table 2 according to their respective range of variation. The distribution is determined by three parameters : the localization is the beginning of the distribution at $y = 0$, the scale determines how the distribution stretches vertically and the shape controls the skewness of the distribution or how steep the decrease is after the initial peak.

The peak value is more representative than the mean as many of the extreme geometries represent weak points that are subject to the frequent erosion that are observed. Historical data gathered from the company monitoring the dyke seem to be in accordance with the choice of the peak value as the representative metrics of the distribution.

This asymmetry is expected since a negative return period would not make sense physically while it is not bounded by any high value.

5 Discussions

5.1 Results Validation

In order to make sure that the estimation of the sensitivity indices is accurate we need to ensure that the convergence of the estimator has been reached. We will do this by plotting the values of the indices and incrementally increasing the amount of points generated by the Sobol' sequence, this is called a validation curve. Note that the amount of plotting points is limited because the Sobol' sequence, being a non independent sample, is only valid for 2^n points. The results are displayed in (fig. 10).

Convergence has evidently been reached. It seems that we can safely use ≈ 15000 points which in our case is still fairly low as the computation of the terminal velocity is pretty fast. However, should the computation time increase by changing the methods of calculation, this could become a problem which would require more intensive optimizations.

5.2 Good practices and dyke improvements

Results from the Global Sensitivity Analysis give indications on how the dyke could be reinforced in order to increase the most the return periods. The recommendation would be to act upon the most significant parameters of the analysis, meaning the ones which yield the highest values of Sobol' indices. This indicates that the geometrical features of the dyke, the crest height

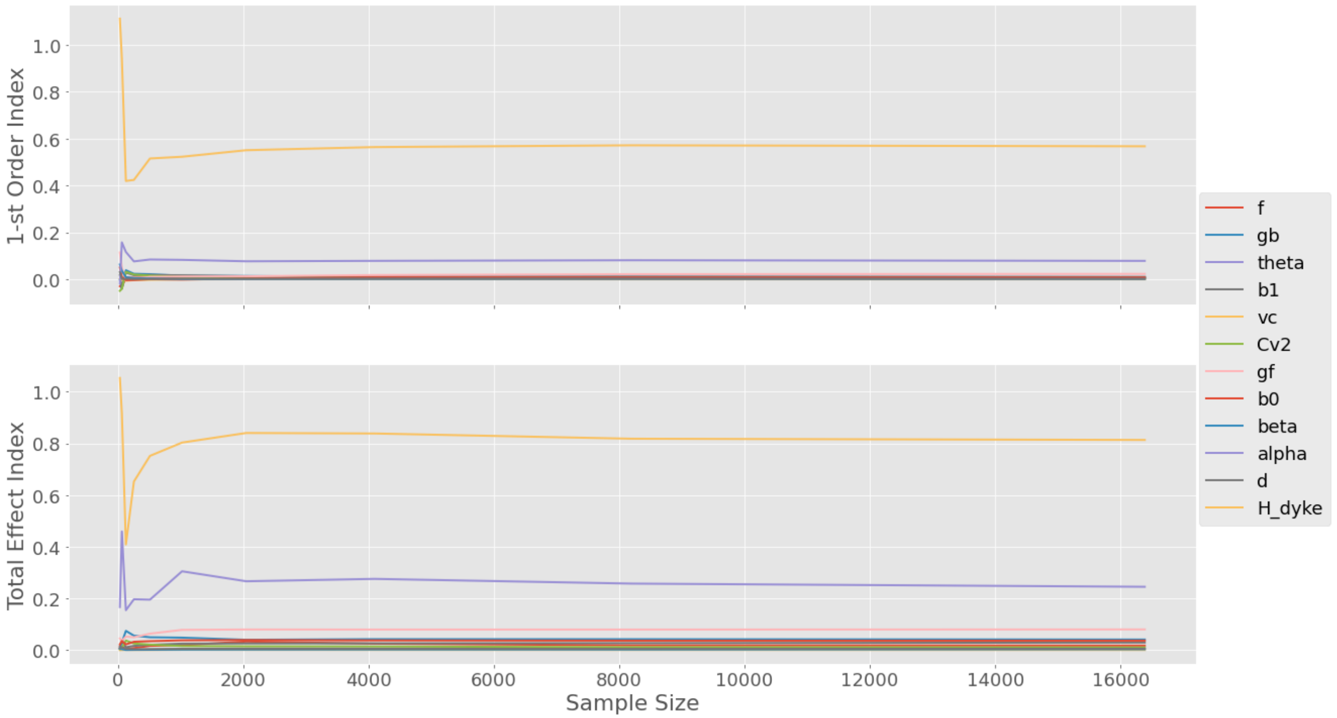

Figure 10. Evolution of the values of the 1-st order sensitiviy index (a) and total effect sensitivity index (b) with sample size.

as well as the slopes, should be acted upon first whenever is possible. Elevating the dyke or decreasing its seaward steepness should bring good results while altering the erosion properties of the landward slope does not look so promising. This focus on the geometrical features of the dyke is supported by Sibley et al. (2017). Generally, the recommendations of the USCE seem to focus mainly on geometrical features and secondly on erosion resistance when considering the design of levees. Approaching the problem using Sobol' indices in this partiular use case had not been done before and seems to provide similar results, confirming the value of the method. The recommendations stated here do not include however an analysis of cost effectiveness which should be one of the next milestones of the work that is presented here.

5.3 Limits of the study

The framework provides a rather complete approach but obviously suffers some limitations. Some of them are inherent to the system itself while others call for future improvements. Our main focus was to obtain an assessment of the risk of erosion on the landward slope of the dyke. A coastal protection is nonetheless submitted to many others damages such as erosion on other locations like the crest of the seaward slope. A more general criteria of security such as "any damage to the dyke" would require to broaden the calculations to take all possible damages into account. We have also limited our criteria of interest as a condition of whether or not the critical velocity has been overreached on the landward slope. The possibility of a breach or the amount of actual amount of eroded material is therefore not quantified. For practical reasons, we calculated return periods on an averaged profile of the dyke which as stated by the global sensitivity analysis can lead to return period different from the local profile. A location-wise study could bring reduced uncertainty and bring more relevant results. Finally, our problem focused on a rather fragile dyke designed with low return periods of dangerous events in mind. Some caution is advised for more resistant structures. Moreover, the features of the pilot site with a low breakwater along Mediterranean sea allowed us not to take into account a non-stationary climate as well as tidal variations. In other sites, these processes should be included.

6 Conclusion

We have been able to build a complete automated framework allowing the user to estimate the expected return periods of events leading to erosion on the rear side of the earthen dyke submitted to wave overtoppings, assuming the correctly assessed ranges of variation of the parameters are provided. The framework itself needs firstly meteocean data in order to create a reliable copula from wave and water level data, then a description of wave propagation to the toe of dyke and finally reliable laws representing wave overtopping process, run-off on the crest then on the landward slope and bottom erosion.

The return period for erosion on the Quenin dyke located in Salin-de-Giraud is firstly estimated from ~~average~~ reference parameters. This first estimate is equal to six years which is significantly higher than the value of two years written in reports from the operating company. The framework is then able to take parameters' uncertainty into account which provides a Generalized Extreme Value distribution of return periods which is right-skewed with a peak around the two years value and a long tail in the upper range of the return periods. This result shows that a statistical study is necessary to determine a return period of damages in accordance with observed damages. Damages on a long dyke are not observed on an average profile but on the weakest profile. That is why the peak of the statistical analysis is more representative than the first estimate based on average parameters. Sensitivity analysis is implemented into the framework and classifies the dyke's parameters in term of carried uncertainty. In the case of the Quenin dyke, the geometrical features of the dyke are the most important, followed in decreasing order by the foreshore conditions, the overtopping characteristics and finally the erosion process itself. The conclusions about sensitivity should only be used on this particular dyke as they are custom-made. This study case is indeed very specific with a very low return period for damages and large variations of the dyke crest. For any other dyke, the framework is applicable by providing the appropriate input values.

Finally, the results can be provided relatively quickly without an enormous amount of computing power. They can be validated indeed using only a small set of points for the Quasi-Monte-Carlo process (around fifteen thousand points at most).

Code and data availability. Freely available on demand to the corresponding author

Author contributions. C. Lutringer - Conceptualization, Methodology, Software, Investigation, Writing - Original Draft, Data Curation, Visualization

A. Poupardin - Supervision, Writing - Review and Editing, Methodology, Resources

P. Sergent - Conceptualization, Methodology, Validation, Surveilance, Project Administration, Funding acquisition

A. Bennabi - Conceptualization, Supervision, Project Administration

J. Jeong - Supervision, Writing - Review and Editing, Resources, Funding acquisition, Project Administration

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

Acknowledgements

We hereby thank the Salins du Midi company for financially supporting the work and especially Pierre-Henri Trapy for providing useful information on the site, guidance as well as access to the company's archives.

References

Acworth, P., Broadie, M., and Glasserman, P.: A Comparison of Some Monte Carlo and Quasi Monte Carlo Techniques for Option Pricing, Springer New York, 1998.

Bergeijk, V. V., Warmink, J., van Gent, M., and Hulscher, S.: An analytical model of wave overtopping flow velocities on dike crests and landward slopes, Coastal Engineering, 2019.

Bernardara, P., Mazas, F., Kergadallan, X., and Hamm, L.: A two-step framework for over-threshold modelling of environmental extremes, Natural Hazards and Earth Systems Sciences, 14, 635–647, 2014.

Capel, A.: Wave run-up and overtopping reduction by block revetments with enhanced roughness, Coastal Engineering, 104, 76–92, 2020.

Chebana, F. and Ouarda, T.: Multivariate quantiles in hydrological frequency analysis, Environmetrics, 22, 63–78, 2011.

Chini, N. and Stansby, P.: Extreme values of coastal wave overtopping accounting for climate change and sea level rise, Coastal Engineering, 65, 27–37, 2006.

De Michele, C., Salvadori, G., Passoni, G., and Vezzoli, R.: A multivariate model of sea storms using copulas, Coastal Engineering, 54, 734–751, 2007.

Durante, F. and Sempi, C.: Copula Theory: An Introduction, Lecture Notes in Statistics, 198, 3–31, 2010.

Durante, F. and Sempi, C.: Principles of Copula Theory, CRC Press, 2016.

Goda, Y.: Random Seas and Design of Maritime Structures, vol. 15, World Scientific Publishing Co., 2000.

Hu, L.: Dependence patterns across financial markets: a mixed copula approach, Applied Finance Economics, 16, 717–729, 2006.

Hughes, S. and Nadal, N.: Laboratory study of combined wave overtopping and storm surge overflow of a levee, Coastal Engineering, 56, 244–259, 2008.

Hughes, S., C.Thornton, der Meer, J. V., and Scholl, B.: Improvements in describing wave overtopping processes, Coastal Engineering, 2012.

Kergadallan, X.: Estimation des niveaux marins extrêmes avec et sans l'action des vagues le long du littoral métropolitain, 2015.

Kole, E., Koedijk, K., and Verbeek, M.: Selecting copulas for risk management, Journal of Banking and Finance, 31, 2405–2423, 2007.

Li, T., Troch, P., and Rouck, J. D.: Wave overtopping over a sea dike, Journal of Computational Physics, 198, 686–726, 2003.

Liu, S. and Han, J.: Energy efficient stochastic computing with Sobol sequences, Design, Automation & Test in Europe Conference & Exhibition, pp. 650–653, 2017.

Lorke, S., Borschein, A., Schüttrumpf, H., and Pohl, R.: Influence of wind and current on wave run-up and wave overtopping. Final report., FlowDike-D, 2012.

Mehrabani, M. and Chen, H.: Risk Assessment of Wave Overtopping of Sea Dykes Due to Changing Environments, Conference on Flood Risk Assessment, 2015.

Muraleedharan, G., Soares, C. G., and Lucas, C.: Characteristic and Moment Generating Functions of Generalised Extreme Value Distribution (GEV), Sea Level Rise, Coastal Engineering, Shorelines and Tides (Oceanography and Ocean Engineering), 14, 269–276, 2011.

Orcel, O., Sergent, P., and Ropert, F.: Trivariate copula to design coastal structures, Nat. Hazards Earth Syst. Sci., 21, 1–22, 2020.

Peterka, A.: Hydraulic design of stilling bassin and energy dissipators, Engineering Monograph, 25, 222, 1958.

Ponsioen, L., Damme, M. V., and Hofland, B.: Relating grass failure on the landside slope to wave overtopping induced excess normal stresses, Coastal Engineering, 14, 269–276, 2011.

Pörtner, H.-O., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., and Rama, B.: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2022.

Saltelli, A., Ratto, M., Terry, A., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis, The Primer, John Wiley and Sons, Ltd, 2008.

Salvadori, G. and Michele, C. D.: On the use of Copulas in Hydrology : Theory and Practice, Journal of Hydrologic Engineering, 12, 2007.

Sergent, P., Prevot, G., Mattarolo, G., Brossard, J., Morel, G., Mar, F., Benoit, M., Ropert, F., Kergadallan, X., Trichet, J., and Mallet, P.: Stratégies d'adaptation des ouvrages de protection marine ou des modes d'occupation du littoral vis-à-vis de la montée du niveau des mers et des océans, Ministère de l'écologie, du développement durable, du transport et du logement, 2015.

Shüttrumpf, H. and van Gent, M.: Wave overtopping at seadikes, Proc. Coastal Structures, pp. 431–443, 2003.

Sibley, H., Vroman, N., and Shewbridge, S.: Quantitative Risk-Informed Design of Levees, Geo-Risk, 2017.

Sobol', I.: Quasi-Monte Carlo methods, Progress in Nuclear Energy, 24, 55–61, 1990.

Sobol', I.: On quasi-Monte Carlo integrations, Mathematics and Computers in Simulation, 47, 103–112, 1998.

Sobol, I.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation, 55, 271–280, 2001.

Sobol ́, I. and Kucherenko, S.: On global sensitivity analysis of quasi-Monte Carlo algorithms, Monte Carlo Methods and Applications, 11, 83–92, 2005.

Tootoonchi, F., Sadegh, M., Haerter, J., Raty, O., Grabs, T., and Teutschbein, C.: Copulas for hydroclimatic analysis: A practice-oriented overview, WIREs Water, 9, 2022.

van der Meer, J.: The Wave Run-up Simulator. Idea, necessity, theoretical background and design, Van der Meer Consulting Report vdm11355, 2011.

van der Meer, J., Provoost, Y., and Steendam, G.: The wave run-up simulator, theory and first pilot test, Proc. ICCE, 2012.

van der Meer, J., Allsop, N., Bruce, T., de Rouck, J., Kortenhaus, A., Pullen, T., Schüttrumpf, H., Troch, P., and Zanuttigh, B.: EurOtop. Manual on wave overtopping of sea defences and related structures. An overtopping manual largely based on European research, but for worldwide application, www.overtopping-manual.com, 2018.

Wahl, T., Jensen, J., and Mudersbach, C.: A multivariate statistical model for advanced storm surge analyses in the North Sea, Proceedings of 32rd International Conference on Coastal Engineering, Shanghai, China, 2010.

Appendix A: Propagation equations from Goda (2000)

$$\beta_0 = b_0 \cdot \left(\frac{H_0}{L_0}\right)^{-0.38} * e^{20 \cdot m^{1.5}} \tag{A1}$$

$$\beta_1 = b_1 \cdot e^{4.2 \cdot \tan\theta_a} \tag{A2}$$

$$\beta_{\max} = \max[0.92, 0.32 \cdot (H_0/L_0)^{0.29} \cdot e^{2.4 \cdot \tan\theta_a} \tag{A3}$$

with m the average steepness of the seabed between the offshore point and the toe of the dyke, θ_a the angle of attack of the oblique waves and L_0 the deep water wavelength. b_0 and b_1 are coefficient determined empiracally from Goda (2000) who gives their values of 0.028 and 0.052, respectively.