Coastal Engineering 118 (2016) 21–34

Contents lists available at ScienceDirect

Coastal Engineering

journal homepage: www.elsevier.com/locate/coastaleng

Predicting coastal hazards for sandy coasts with a Bayesian Network

Laurens Poelhekke [a,b,c,*], Wiebke S. Jäger [b], Ap van Dongeren [a], Theocharis A. Plomaritis [d], Robert McCall [a], Óscar Ferreira [d]

[a] Deltares, Department of Applied Morphodynamics, Delft, The Netherlands
[b] Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Hydraulic Engineering, Delft, The Netherlands
[c] CDR International, Amersfoort, The Netherlands
[d] University of the Algarve, CIMA, Faro, Portugal

ARTICLE INFO

Article history:
Received 11 April 2016
Received in revised form 16 August 2016
Accepted 23 August 2016
Available online xxxx

Keywords:
Early Warning System
Bayesian Network
Sandy coasts
XBeach
Probabilistic
Hazards

ABSTRACT

Low frequency, high impact storm events can have large impacts on sandy coasts. The physical processes governing these impacts are complex because of the feedback between the hydrodynamics of surges and waves, sediment transport and morphological change. Predicting these coastal changes using a numerical model requires a large amount of computational time, which in the case of an operational prediction for the purpose of Early Warning is not available. For this reason morphodynamic predictions are not commonly included in Early Warning Systems (EWSs). However, omitting these physical processes in an EWS may lead to potential under or over estimation of the impact of a storm event.

To solve this problem, a method has been developed to construct a probabilistic Bayesian Network (BN). This BN connects three elements: offshore hydraulic boundary conditions, characteristics of the coastal zone, and onshore hazards, such as erosion and overwash depths and velocities. The hydraulic boundary conditions are derived at a water depth of approximately 20 m from a statistical analysis of observed data using copulas, and site characteristics are obtained from measurements. This BN is trained using output data from many pre-computed process-based model simulations, which connect the three elements. Once trained, the response of the BN is instantaneous and can be used as a surrogate for a process-based model in an EWS in which the BN can be updated with an observation of the hydraulic boundary conditions to give a prediction for onshore hazards.

The method was applied to Praia de Faro, Portugal, a low-lying urbanised barrier island, which is subject to frequent flooding. Using a copula-based statistical analysis, which preserves the natural variability of the observations, a synthetic dataset containing 100 events was created, based on 20 years of observations, but extended to return periods of significant wave height of up to 50 years. These events were transformed from offshore to onshore using a 2D XBeach (Roelvink et al., 2009) model. Three BN configurations were constructed, of which the best performing one was able to predict onshore hazards as computed by the model with an accuracy ranging from 81% to 88% and predict events with no significant onshore hazards with an accuracy ranging from 90% to 95%. Two examples are presented on the use of a BN in operational predictions or as an analysis tool.

The added value of this method is that it can be applied to many coastal sites: (1) limited observations of offshore hydrodynamic parameters can be extended using the copula method which retains the original observations' natural variability, (2) the transformation from offshore observations to onshore hazards can be computed with any preferred coastal model and (3) a BN can be adjusted to fit any relevant connections between offshore hydraulic boundary conditions and onshore hazards. Furthermore, a BN can be continuously updated with new information and expanded to include different morphological conditions or risk reduction measures. As such, it is a promising extension of existing EWSs and as a planning tool for coastal managers.

1. Introduction

Over the past decades a number of large storm events have demonstrated the vulnerability of the coastal zone in Europe, such as to the North Sea Flood of 1953 in the Netherlands, Belgium and the United Kingdom (Gerritsen, 2005), Xynthia (2010) affecting the entire coast of south-western Europe (Bertin et al., 2012) and Hercules (2014) causing severe coastal erosion and flooding in parts of France and the United Kingdom (Masselink et al., 2015a) among many others. Larger and more extreme events such as the hurricanes in the USA, e.g. Katrina in 2005 (Knabb et al., 2006) and Sandy in 2012 (Blake et al., 2013), and Typhoons in Asia, e.g. Haiyan in 2013 and Nargis in 2008, have also shown the devastating effects of these low-frequency, high impact flood events.

* Corresponding author at: CDR International, Amersfoort, The Netherlands.
E-mail address: l.poelhekke@cdr-international.nl (L. Poelhekke).

http://dx.doi.org/10.1016/j.coastaleng.2016.08.011
0378-3839/

Sandy coasts are especially vulnerable as they exhibit large responses to these low-frequency high impact events, such as extensive beach and dune erosion (Castelle et al., 2015; Masselink et al., 2015b), and even breaching (Roelvink et al., 2009; McCall et al., 2010). Buildings and infrastructure built on these sandy coasts are not only vulnerable to flooding but also to damage caused by overwash and coastal erosion (Smallegan et al., 2015). Furthermore, extreme overwash and breaching can alter the barrier and lagoon morphology and possibly affect the water level inside the lagoon, changing the back barrier flooding hazard.

Coastal hazard prediction has long been the focus of the scientific and coastal management community. Early approaches were mainly focused on the classification of offshore hazard based on wave power (Dolan and Davis, 1992) or associated coastal response (like erosion) to offshore forcing (Miller and Livermont, 2008; Mungar and Kraus, 2010). The most commonly-used approaches to derive coastal hazards use simple relations (Kriebel and Dean, 1993; Larson et al., 2004; Mendoza and Jiménez, 2006; Stockdon et al., 2007). Using hindcasted or operational predictions of wave and surge levels with or without flood propagation, coastal impacts such as flooding over an invariant topography are computed using the "bathtub" or flooded valley approach (Leatherman, 1990; Carrasco et al., 2012). Overwash is routinely computed using empirical equations (e.g. (Rodrigues et al., 2012)), and beach erosion using static models (e.g. (Ferreira et al., 2006)). Due to the complexity of the storm processes Stockton et al. (2007) classified storm impact based for the different regimes proposed by (Sallenger, 2000). The above consider each hazard separately and do not include the morphodynamic response of a coast to high impact events and usually do not include feedback mechanisms between waves, currents, sediment transport and morphological change. For sandy shores with beaches and dunes the coastal morphological response is non-negligible and influences the pathway of a coastal hazard to the hinterland, changing for instance the flood duration, the extent and the depth fields (McCall et al., 2010; Cañizares and Irish, 2008). Current models can therefore under or over predict the hazard intensities and impact of coastal hazards on sandy coasts, both of which are detrimental for planning or evacuation purposes.

Morphodynamic process-based models such as XBeach (Roelvink et al., 2009) are capable of simultaneously computing wind- and wave-induced water levels and velocities as well as the associated morphological response. However, these models are complex, which comes at the expense of increased computational cost. This poses a problem for the use of such models in an Early Warning System (EWS), where the computational window is limited to the short period between successive, updated meteorological forecasts. Early attempts to incorporate morphodynamic models in EWS (Plomaritis et al., 2012; Vousdoukas et al., 2012a) resulted in simplification of the morphology to several 1D profiles in order to increase the models operability.

To solve this problem, a solution is proposed in which a probabilistic model based on a Bayesian Network (BN) is utilized as a surrogate for a process-based model. A BN is in essence a probabilistic graphical model, which consists of random variables and conditional dependencies between said variables. The random variables are the hydraulic boundary conditions (defined at the 20 m depth contour), such as the surge and wave parameters, and the onshore hazard intensities, such as erosion, overwash depths and flow velocities. The conditional dependencies between the random variables can be determined by training the BN using output data from many pre-computed process-based model simulations, as well as from observations. Once trained, the response of the BN is instantaneous. It can be included in an operational EWS in which the BN can be conditioned with predicted waves and water levels from offshore hydrodynamic models to produce a prediction for onshore coastal impacts.

BNs have been proven useful in a number of coastal applications. Hapke & Plant (2010) applied them to predict cliff erosion by connecting the forcing variables (e.g. wave conditions) and initial conditions (e.g. cliff geometry). Dune erosion volumes due to storm impact, as predicted by an empirical model, have also been reproduced by a BN (Den Heijer et al., 2012). A BN has been used to predict coastal vulnerability to sea level rise (Gutierrez et al., 2011), and to assess the interactions between barrier island geomorphic variables (Gutierrez et al., 2015). Van Verseveld et al. (2015) applied it to relate the onshore hazard intensities to observed building damage for the case of the impact of hurricane Sandy on a part of New York. In this paper, we will build on this previous work and apply the BN to relate offshore hydraulic boundary conditions to onshore hazards through a transformation using process-based model simulations.

The key point of this paper is the development of a method in which a BN is a surrogate for a process-based model within an EWS. The main focus of the paper is on the development of a generic method with an application on a low-lying barrier coast which serves as an example. In this paper, we will focus on the method and application, and provide references to background literature on the individual elements.

In Section 2, a method is developed to construct a BN, which is a surrogate for a complex process-based model and can be implemented in an EWS for urbanised sandy coasts. The method is applied in the case study site of Praia de Faro (Algarve, Portugal) in Section 3. The discussion is in Section 4 and a summary and conclusions are presented in Section 5.

2. Methods

A probabilistic model, a BN, will be constructed which can be used as a surrogate for a process-based coastal morphodynamic model in an EWS. A BN consists of nodes and arcs, in which the nodes represent the variables of interest and the arcs indicate the conditional dependencies between them (Pearl and Russel, 1988). When the nodes and arcs are set up the network can be trained with a dataset, which may consist of observations or of synthetic numerical model results. In the present application the BN will be trained with the results of a process-based numerical model which is able to transform a range of offshore hydraulic boundary conditions (water levels, wave heights, wave periods, etc.) in deeper water, typically at 20 m depth, to onshore hydrodynamic (water depths and currents) and morphological (erosion) hazards on the coast. Hydraulic boundary conditions will be derived from a statistical analysis of observational data. The combination of the hydraulic boundary conditions and onshore hazards simulated in the process-based model forms the dataset that trains the BN. Once the BN is trained it can be conditioned with an observation, or prediction, of the hydraulic boundary conditions to give a prediction of onshore hazards.

The method involves five steps: (1) a dataset from which conditions with large return periods can be derived is synthesised if long-term observations are not available (which is usually the case), (2) the time variation of the offshore hydrodynamic parameters during an event is schematised, (3) a process-based model is constructed to transform the offshore hydraulic boundary conditions to onshore hazards (4) a BN is set up that is sufficiently complex (in terms of nodes and arcs) to represent the relations between variables, yet as simple as possible to minimise the required amount of training data and (5) metrics are determined to assess the predictive value of the BN.

2.1. Synthesis of extreme events

The first step is to obtain a large enough dataset from which extreme conditions can be derived to train the BN. If long-term observations are not available, this dataset needs to be synthesised, for which a method using copulas (Sklar, 1959) is applied. Copulas are mathematical tools that can be used to construct multivariate distributions. An example are storm events which can be characterised by a set of random variables such as the significant wave height, surge level, storm duration and peak wave period. The interrelations between these variables may be characterised by a multivariate probability distribution. Classical parametric families of multivariate distributions, e.g. Gaussian,

student-t or extreme value distributions force the marginal behaviour of the random variables to belong to the same family, which is not the case for the hydrodynamic variables under consideration. Copulas, however, are able to describe the interrelation between several random variables without this restriction (Genest and Favre, 2007; Schmidt, 2007). Just like marginal distributions can be fit to a random variable, a copula can be 'fit' to describe the dependence between two or more random variables. Several copula types exist which can be used to emulate a dataset and be tested on their goodness of fit. Once an appropriate copula model is found it can be used to create a synthetic dataset that mimics the characteristics of the original data.

The method followed in this research is a combination of the work of Wahl et al. (2012), Corbella and Stretch (2013) and Jäger and Morales Nápoles (2014). Observed events in the area of interest are used to create bivariate distributions between pairs of variables. These bivariate distributions resemble the pattern of the observed data and can be used to sample new data points that form the set of synthetic events that will be used to train the BN.

There are four steps that have to be taken to produce the synthetic dataset and are elaborated in the application Section 3.2: (1) obtain and constrain a dataset of observations; (2) fit marginal distributions to the measured data points; (3) fit different copulas to each pair of variables and (4) use the copulas in combination with the marginal distributions to sample from the bivariate distributions to obtain the synthetic data.

2.2. Storm event schematisation

Significant wave heights, periods and water levels show a time variation over the course of a storm event. This variation is unknown a priori and therefore need to be schematised. To mimic the typical behaviour of a storm increasing and then decreasing in strength, its time evolution is schematised by a symmetric linear increase, until the peak of the storm is reached, and subsequent decrease of the wave height and surge level (Fig. 1a). Assuming constant wave steepness, the peak period is related directly to the wave height.

To drive the process-based model that transforms the hydraulic boundary conditions to onshore hazards, time series of sea states of waves and total water levels are imposed at the seaward boundary of the coastal process-based model. The total water levels consist of surges and tides. Since tidal amplitudes vary over a spring-neap cycle, and are statically uncorrelated to the storm surge, several tidal signals need to be considered for each modelled event. To this end, several representative tidal amplitudes are extracted from observations. The tide is then imposed on the model starting in a random phase (schematized in Fig. 1b). Because the total water level is the sum of the surge and the tidal elevation, the maximum does not necessarily occur at the surge peak of the event.

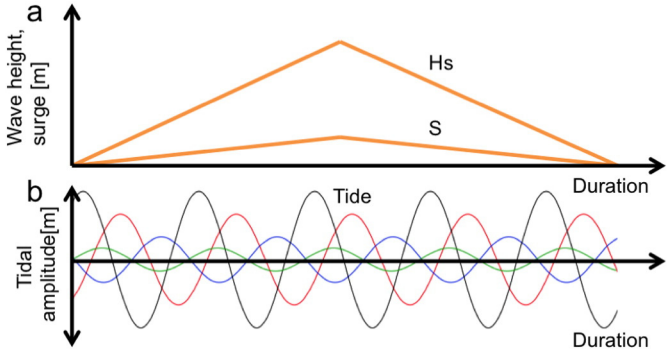

Fig. 1. Modelling of an event. a) The wave height is linearly increased to its maximum value after which it linearly decreases. The surge level follows the same pattern. b) The tidal signal is independent of the wave height and surge levels; therefore, several signals are modelled for each storm.

2.3. Construction of a process-based model

A process-based model is used to transform hydraulic boundary conditions to onshore hazards such as erosion, flooding and overwash velocities. In the current method, XBeach is selected for this purpose. "XBeach solves the 2D horizontal nonlinear shallow water equations with time-varying forcing obtained from wave action equations. It thus resolves motions at the time scale of wave groups (infragravity waves), which have been shown to be of importance in the dune erosion process (Van Thiel de Vries, 2009). It is capable of seamlessly modelling all four dune impact regimes as defined by (Sallenger, 2000), and model skill has been demonstrated on barrier islands (McCall et al., 2010; Lindemer et al., 2010) and urbanised coasts (van Verseveld et al., 2015; Nederhoff et al., 2014) among others. XBeach is chosen as the most appropriate numerical model to use, because it has been extensively validated for simulating morphological change over complex 2D bathymetry. Also, coastal structures can be represented as hard, non-erodible layers" (paragraph adopted from Smallegan et al. (2015)).

A complex model such as XBeach 2D has a large computational demand. Thus, an evaluation of the required accuracy versus computational time must be performed in order to achieve an optimum. In our application, we achieved a reduction in computation time by using the morphological acceleration ("morfac") factor (Ranasinghe et al., 2011) and by reducing the alongshore grid resolution as the coast in our example is relatively long, straight and regular.

The model performance is assessed using the Brier Skill Score (BSS) on observed and computed morphological change in the active coastal zone (van Rijn et al., 2003; Sutherland et al., 2004). A BSS score of 1 indicates a perfect model prediction, and a value of 0 means that the model's bed level prediction is no better than the assumption of a static bed.

2.4. Development of a Bayesian Network

In the present case a BN is used to describe hazards in the coastal zone, which are determined by the hydraulic boundary conditions and the local characteristics of a section of the coast, as shown in Fig. 2. A BN is a computational tool to describe a system in a probabilistic way using conditional probability tables (CPTs). The relation between the variables in a system comes from prior knowledge about the system. A prediction is made by updating the BN with an observation, such as the significant wave height associated with an event. The BN then uses Bayes' rule to update the likeliness of the other variables in the network that the observation is linked to, according to:

$$p(F_i|O_j) = \frac{p(O_j|F_i)p(F_i)}{p(O_j)} \quad (1)$$

In which, O_j is a subset of observations and F_i is a forecast variable. $p(F_i|O_j)$ is the conditional probability of the forecast F_i, given O_j. In other words, $p(F_i|O_j)$ is the probability that F_i is true given O_j. $p(F_i)$ and $p(O_j)$ represents the prior marginal probabilities of the forecast variable and the observation.

Setting up a BN consists of three steps: (1) setting up the structure by identifying the variables that need to be included in the network and the dependencies between the variables, (2) discretising each node into bins, and (3) training the network with a dataset. In this paper, three different configurations are explored by varying the discretisation and the number of variables. The software used to create

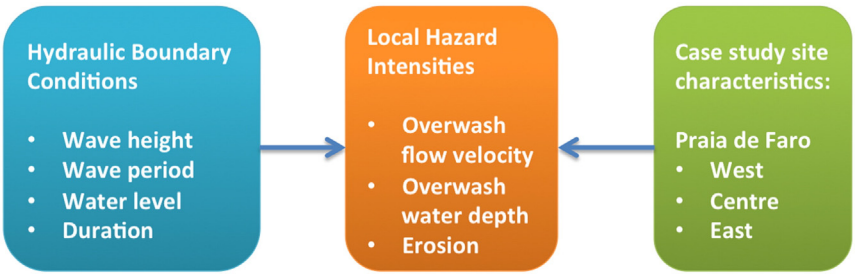

Fig. 2. Relationship between the hydraulic boundary conditions, local hazard intensities and the case study site characteristics. From left to right: The nodes that make up the hydraulic boundary conditions, local hazard intensities and the case study site characteristics.

the BN is Netica version 5.12, obtained from Norsys (http://www.norsys.com).

The dataset consists of a number of cases, in which one case is a single data entry into the BN and consists of one value at each individual node. In this application one case consists of a data entry of the boundary conditions of a single event and the local hazard intensities associated with that event for a specific area, and represents the result of one simulation with the coastal process-based model.

The nodes of the BN are discretised into bins that cover the full range of the input data. The discretisation is important because it influences the predictive capabilities of the BN. This is best illustrated by the number of entries in a CPT and the way a BN learns from data. The number of entries in a CPT is defined by:

$$CPT_{Entries} = \#Bins_{Child} \times \prod_{N=1} \#Bins_{Parent,N} \quad (2)$$

In this case a child node is a hazard and a parent node is a hydraulic boundary condition or a location at a given study area. This means that more bins and more parents create larger CPTs. In Netica the learning of the BN is based on a simple *counting-learning* algorithm (Norsys, 2016). Initially assuming uniform distributions within the CPT, this algorithm requires each of its entries to be satisfied by multiple cases such that it can gain "experience". If the CPT is too large for the dataset the experience values will be low and the probabilities close to uniform and hence uninformative. A well-trained BN has high experience values and will provide predictions with small standard deviations and large occurrence probabilities.

2.5. Assessment of the predictive value

The predictive value of the BN is assessed by comparing its predictive skill to one of a competing model. As a competing model the prior probabilities of the BN may be considered. These are the probabilities of the hazard nodes in the BN after it has been trained with a dataset, but before the BN is conditioned with an observation of the boundary conditions. The comparison is made using the log likelihood ratio (LLR) as has been done in Plant & Holland (2011). This ratio is positive if the likelihood of a prediction increases as compared to the prior probability, likewise the ratio is negative if the predictions likelihood decreases as result of the updated prediction:

$$LLR = \sum_{j=1}^{n} \log\left\{p(F_i|O_j)_{F_i=O_j}\right\} - \log\left\{p(F_i)_{F_i=O_j}\right\} \quad (3)$$

In which, F_i is the forecast variable, O_j is a subset of observations that are inserted into the BN to make a prediction and O'_j is the observation that is withheld from the prediction. A positive result indicates that the BN has predictive skill since the updated prediction is then better than the outcome based on the prior distributions of the network. The BN is trained with 90% of the available dataset and the remaining 10% of the data is used to test its predictive skill, following van Verseveld et al. (2015)). This is done for ten randomly chosen subsets of the dataset such that all data are used in the testing. The LLR of the BN is then obtained by averaging over the LLRs of the BNs created with the 10 subsets of the data.

To compare the performance of two BNs trained with the same dataset, a similar approach can be used. Instead of comparing the prior probability of a BN with its updated probability, the updated probabilities of two BNs can be compared:

$$LLR = \sum_{j=1}^{n} log\left\{p(F_{BM,i}|O_j)_{F_{BM,i}=O_j}\right\} - \log\left\{p(F_{CM,i}|O_j)_{F_{CM,i}=O_j}\right\} \quad (4)$$

In which $F_{BM,i}$ is the updated probability of the Base Model and $F_{CM,i}$ is the updated probability of the competing model.

3. Application to Praia de Faro

3.1. Site description

Praia de Faro (Faro Beach) is the beach community of Faro, which is the regional capital of the Algarve (Portugal; Fig. 3a). Praia de Faro is located at the Ancão Peninsula which is part of the Ria Formosa barrier island system, consisting of a lagoon sheltered from the Atlantic Ocean by five barrier islands and two peninsulas. The coast is NW-SE oriented and the central part of the peninsula is urbanised which has changed its natural configuration, namely by dune lowering and replacement by buildings, a car park, roads and a camping site. An aerial view shows the barrier (Fig. 3b). Praia de Faro has an access road connecting the barrier to the mainland. At the end of the access road lies a parking lot that extends from the back of the barrier to the berm crest. Directly east of the access road, buildings have been built on top of the dunes and facing the beach berm. Immediately west of the access road, buildings are slightly recessed from the beach and are protected by a single row of frontal dunes. At the western end of Praia de Faro, members of the local fishing community have dwellings directly on top of the dunes. The fishing community also occupies the eastern end of Praia de Faro, where houses are scattered mostly along the back barrier. For the purpose of the performed analysis only the densely-occupied area near the parking lot is considered where mostly second homes are located along with restaurants and bars and the access road. This area of Praia de Faro is impacted by annually recurring overwash events (Almeida et al., 2012) and has experienced damage to houses and infrastructure due to severe erosion caused by storm events.

The beach face at Praia de Faro has a mean grain size about 0.5 mm (Achab et al., 2014) is steep, typically above 10%, and ranging from 6% to 15% (Achab et al., 2014; Vousdoukas et al., 2012b), categorising the beach as intermediate to reflective (Wright and Short, 1984). Ria Formosa is exposed to storms of which approximately 70% originate from

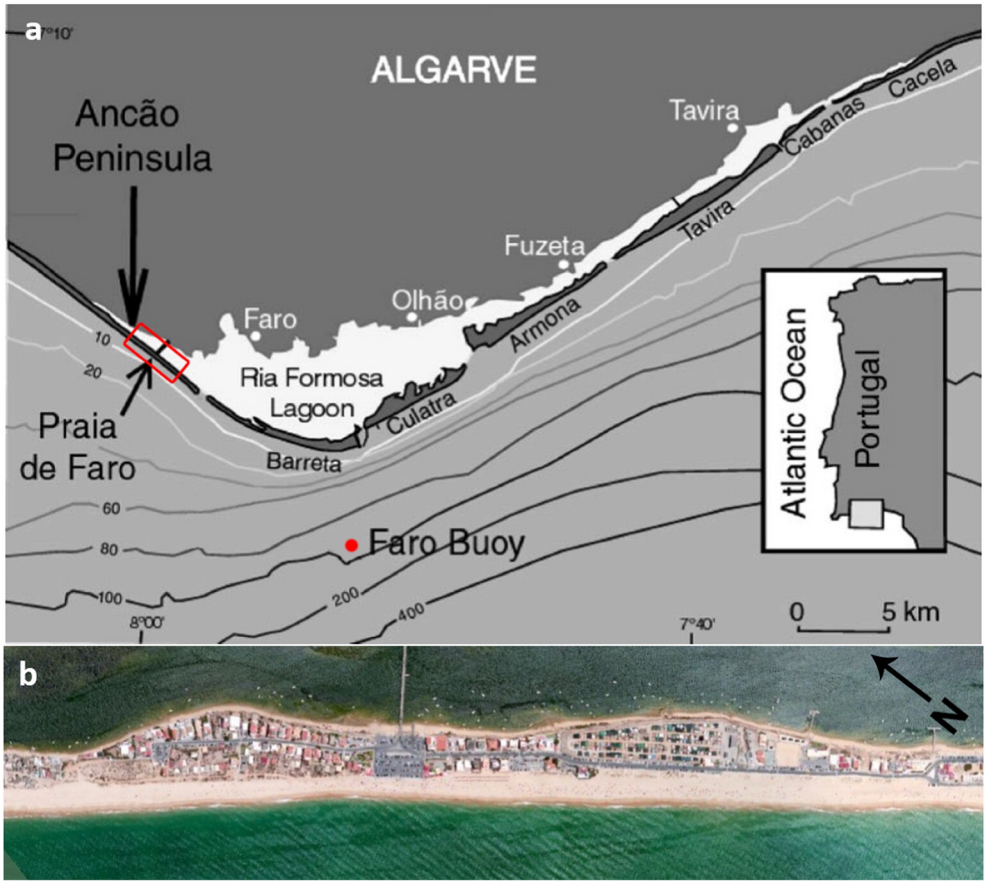

Fig. 3. a): Location of the case study site in the Algarve (red box), Portugal (adapted from Ferreira et al. (2006). b): Aerial photograph of the central part of Praia de Faro, showing the parking lot and the access road in the middle. Source: iGEO. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the west-southwest and 30% from the east-southeast (Almeida et al., 2011). Due to the natural triangular shape of Ria Formosa, Praia de Faro is relatively protected from the latter direction. Moreover, due to a narrow continental shelf the storm surge levels are relatively low (<1 m) (Fortunato et al., 2016). The semi-diurnal tide with an average range of 2.8 m for spring-tides (maximum 3.5 m) and 1.3 m during neap tides is more important in determining the maximum storm water level than surge (Fortunato et al., 2016). The return periods for the significant wave height (from W-SW), associated peak period and storm surge levels are summarized in Table 1.

3.2. Synthetic dataset

A synthetic dataset of hydraulic boundary conditions has been created for the case study site following the four steps described in Section 2.1: (1) obtain and constrain a dataset of observations; (2) fit marginal distributions to the measured data points; (3) fit different copulas to each pair of variables and (4) use the copulas in combination with the marginal distributions to sample from the bivariate distributions to obtain the synthetic data.

This is done for a range of events with a significant wave height of at least 3.0 m to a maximum of 8.1 m. These limits are based on Almeida et al. (2012), who showed that storms with a significant wave height <3.0 m do not lead to appreciable coastal impact. Storms with a significant wave height of 8.1 m have a return period of 50 years, which is commonly used for risk assessment in the coastal management plans for the study area.

In the first step, wave data are obtained from a directional wave-rider "Faro Buoy", placed at 93 m depth near Praia de Faro (Fig. 3a), which has 20 years of data between January 1993 and December 2013. A peak over threshold analysis with a threshold of 2.5 m was used to identify individual events. Surge levels are available for the period of June 1997 until June 2007, obtained from a nearby Spanish network of tide gauges, some 80 km east of the study area (Rodrigues et al., 2012). Surge levels are provided and are calculated by elimination of the tidal signal using local tidal constituents derived from the total time-series.

Only events originating from the south-western quadrant (180°–270°) have been included in the dataset, because of the exposure of the case study site. The observed maximum significant wave heights versus storm duration, surge and wave peak period associated with the significant wave height are shown in Fig. 4 (black dots). The dependency between the significant wave height and the storm duration and surge levels are significant, with r-squared values of respectively 0.54 and 0.41. The dependency between the significant wave height and the associated peak period show a large scatter with an r-squared value of 0.04. These characteristics are transferred to the synthetic dataset by using copulas, as described below.

Table 1
Return periods for the significant wave height, peak period and surge levels for Praia de Faro, Portugal. * (Pires, 1998), ** (Rodrigues et al., 2012).

Return period (yr)	Hs (m)*	Tp (s)**	Surge (m)**
5	5.7	11.3	0.46
10	6.4	11.9	0.54
25	7.4	12.7	0.65
50	8.1	13.3	0.72
100	8.8	13.9	0.80

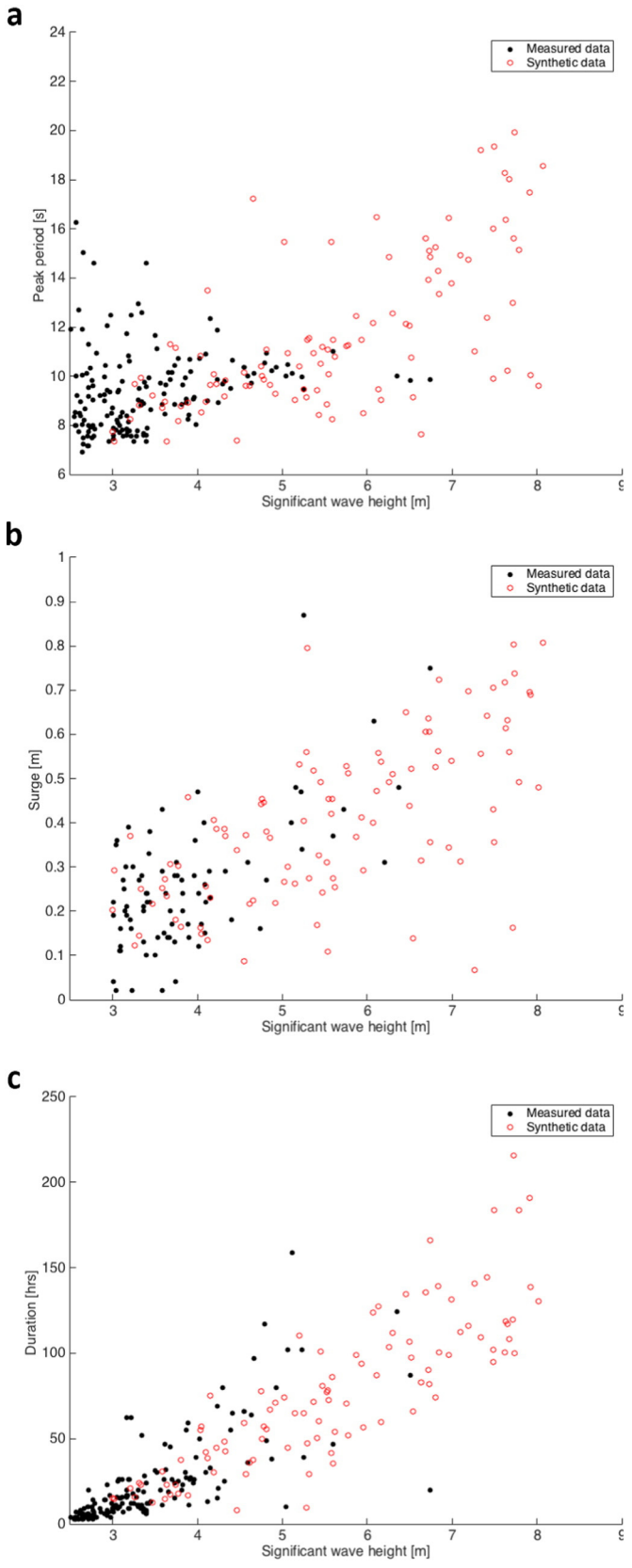

Fig. 4. Scatter plots of the significant wave height versus a) the peak period, b) surge levels and c) storm duration. The observed data is given in black and the synthetic events produced with the copulas are given in red dots. The synthetic events are produced uniformly on the interval [3 m, 8.1 m] H_S to be used for the creation of a BN. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Best marginal distribution fits for the individual variables of the storm dataset according to the Akaike information criterion.

Variable	Best fit distribution
Hs	Generalized Pareto
Tp	Generalized extreme value
Duration	Exponential
Surge	Rayleigh

In the second step, the individual variables are considered separately: the significant wave height, the peak wave period, the storm duration and the surge level. For each variable several distributions have been fit and ranked according to the Akaike Information Criterion (AIC), following the approach in Corbella & Stretch (2013). The top-ranked distributions (shown in Table 2) have been used for the copula fitting.

The third step is to fit the copulas. Similar to marginal distributions, which can be fit to a random variable, a copula can be 'fit' to describe the dependence between two or more random variables. A general introduction to copulas can be found at Schmidt (2007), and a specific application for hydrological phenomena at Genest & Favre (2007). The method followed in this research is a combination of the work of Corbella & Stretch (2013) and Jäger and Morales Nápoles (2014) who both use copulas to describe sea conditions.

Several copula types exist and can be used to emulate a dataset and tested on their goodness of fit. Once a copula is selected it can be used, together with the marginal distributions, to create a synthetic dataset that mimics the characteristics of the original data. Three variable pairs have been identified for which a copula have been fit: (1) the significant wave height and the storm duration, $H_s - D$, (2) the significant wave height and the storm surge level, $H_s - S$ and (3) the significant wave height and the peak period, $H_s - T_p$. The choice for the pairs of bivariate copulas was driven by the availability of data; one dataset contained wave heights, periods and storm duration and a second contained wave height and storm surge. Even though other variable pairs are not coupled directly they are dependent through the wave height. For the first and second variable pair four possible copula types have been applied: Gaussian, Clayton, Frank and Gumbel. An additional copula has been fit for the third pair, the gamma factor copula (Joes, 2015). This copula allows for skewness in the dataset, which is due to the phenomena of wave breaking (not shown). For simplicity only a general description of the fitting procedure is presented. More elaborate explanations are provided in the above recommended literature.

In copula modelling the dependence structure between two variables is described in isolation from their marginal behaviour. Therefore, the marginal distributions of the variables are transformed to uniform distributions by using the respective cumulative density functions of their marginal parametric distributions. The copulas are then fit using Spearmans's rank correlation coefficient ρ.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{5}$$

Table 3
Cramer von Mises test results for each copula and variable pair. The test scores can only be compared within a variable pair. A lower test score indicates a better fit.

Copula	Hs - Duration	Hs – Surge	Hs – Tp
Clayton	0.4427	1.4861	0.7977
Frank	0.4616	1.1665	0.7756
Gaussian	**0.3249**	1.0484	0.7764
Gumbel	0.4071	**0.7737**	0.8170
Gamma factor	–	–	**0.6123**

Table 4
Best fitting copulas and parameters for three variable pairs based on the CvM statistic and a visual check.

Variable pair	Copula	Parameters
$H_s - D$	Gumbel	$\rho = 0.8814$
$H_s - S$	Gaussian	$\alpha = 1.4621$
$H_s - T_p$	Gamma factor	$\theta_0 = 0.3487$, $\theta_{vec} = [0.1029,\ 1.2286]$

For which $d_i = x_i - y_i$ and x and y are the ranked variables and n the size of the sample.

The goodness of fit of a copula is judged with two methods following Jäger and Morales Nápoles (2014). Firstly, the test statistic of the Cramer von Mises (CvM) test (e.g. Genest and Favre (2007)) is applied, which uses least-squares to indicate the difference between the data and the parametric copula model. Thus, a lower test score represents a better fit. Note that this test is relative, meaning that the scores can only be compared within each variable pair. Secondly, empirical and parametric conditional probabilities have been visually compared. For the visual test, a dataset was generated from the copulas and compared to the original dataset. The scatterplots of the ranks of the data pairs as well as conditional probabilities in each quadrant were compared qualitatively. Based on the CvM test statistc (Table 3) and the visual test the Gaussian copula was chosen for the $H_s - D$ variable pair, the Gumbel copula for the $H_s - S$variable pair and the gamma factor copula for

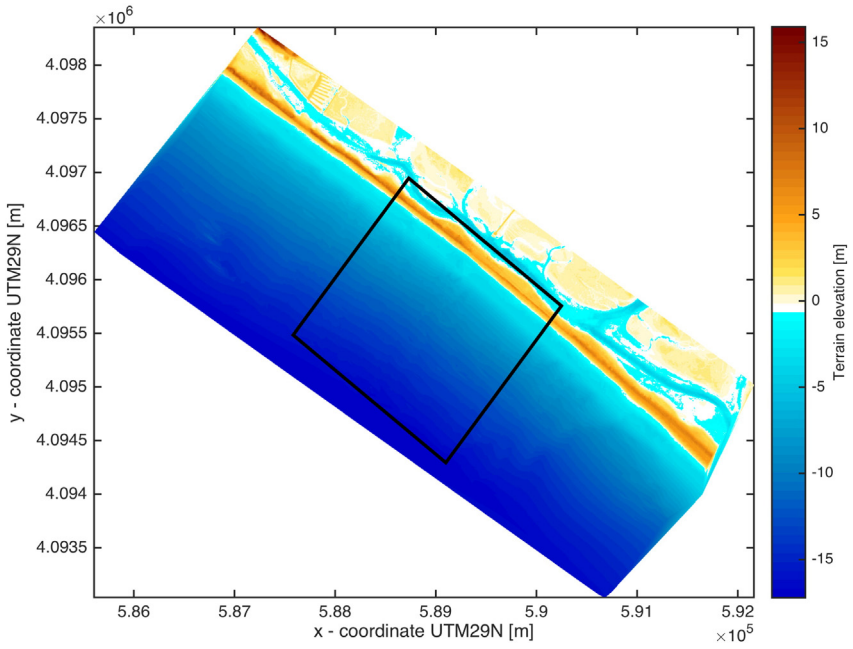

Fig. 5. Composite topographic and bathymetric digital terrain model and location of the research area (black rectangle). Elevations are relative to MSL.

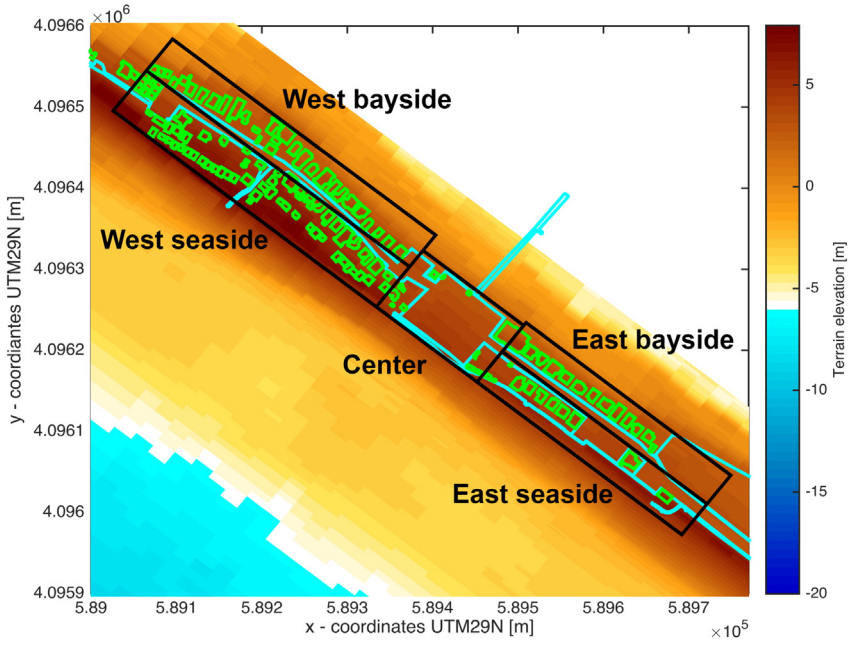

Fig. 6. Locations specified for the Bayesian Network are enclosed in the black boxes. Building footprints are indicated in green and infrastructure in light blue. Elevations are relative to MSL. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the$H_s - T_p$variable pair. The parameters of the copulas are given in Table 4.

A large synthetic dataset of one million samples has been generated with each copula. These points are marginally uniformly distributed on the interval [0,1]and are transformed back by applying the inverse CDFs. Since the BN needs to be trained evenly over the whole range of possible events, in the final step a uniformly distributed set of one hundred events has been resampled on the interval [3.0 m, 8.1 m] of significant wave height, from the one million data points. The results are shown in Fig. 4, where the black dots are the observations and the red dots the synthetic data set, which will be used to force the XBeach model.

3.3. Configuration of the XBeach model

In this study, XBeach version 1.22.4867 has been used (http://www.xbeach.org). The model inputs consist of bathymetric and topographic information and waves and water level time series at the offshore boundary.

Topographic and bathymetric information from the summer of 2011 is available for the Ancão Peninsula (Fig. 5). It consists of bathymetric cross sections of the wet areas and a LIDAR of the dry area with a resolution of 2 by 2 m. As a case study an area of interest with a total alongshore length of 1.7 km was selected, indicated by the box in Fig. 6. It is centred on the entrance road/parking lot of Praia de Faro, which is considered the most vulnerable location (Almeida et al., 2012). A 2DH model was setup in which the grid cell sizes vary in the cross shore direction from 20 m at the offshore boundary to 1.4 m at the coast and in the alongshore direction from 25 m in the centre to 40 m at the lateral boundaries. The model was calibrated against measured pre- and post-storm upper beach profiles obtained during a number of consecutive storms from December 2009 to January 2010 (Vousdoukas et al., 2012b), which had return periods of approximately 2 years. Although the return periods were small, the conditions were relatively energetic and responsible for strong morphological changes. The calibration yielded a BSS of 0.8. The parameter values after calibration are given in Appendix A.

3.4. Bayesian Network

3.4.1. Structure

The structure of the BN follows Fig. 2, in which the hydraulic boundary condition nodes are identified as: (1) the maxima of the offshore wave height, (2) the peak wave period, (3) the water level that occurs during the peak of the storm (as defined in Section 2.2) and (4) the storm duration. The site is characterised by five distinct areas in terms of the dune height, barrier island elevation and distance between buildings and infrastructure and the shoreline (Fig. 6). The hydraulic boundary conditions and the case study characteristics nodes are the parent nodes. The child nodes are the coastal hazard nodes. These are subdivided over infrastructure and buildings and are identified as (1) erosion depth, (2) overwash flow velocity and (3) overwash water depth. Erosion is defined as the positive vertical distance at a grid point between the pre storm bed level and the lowest bed level measured during an event. The overwash depth at each grid point is defined as the water depth between the instantaneous bed elevation and water level.

3.4.2. Training data

The training dataset for the BN comprises 2000 cases: 100 events were selected from the synthetic dataset and run in combination with four tidal signals with a random phase. In each XBeach simulation aggregated data have been extracted for the five areas discussed above. The maximum offshore wave height, peak period and storm duration were obtained for each case from the synthetic dataset. The maximum water level is defined within the peak interval of the event, defined as the 25% of the total duration in which the highest waves occur. Model output of erosion, overwash depth and flow velocity are interpolated to the corner of each building and infrastructure polygon (Fig. 6), and the maximum hazard values computed for each of the five areas are subsequently stored in the BN coastal hazard nodes. An example of XBeach model results of the overwash water depth, overwash flow velocity and erosion for a high-energy event is shown in Fig. 7.

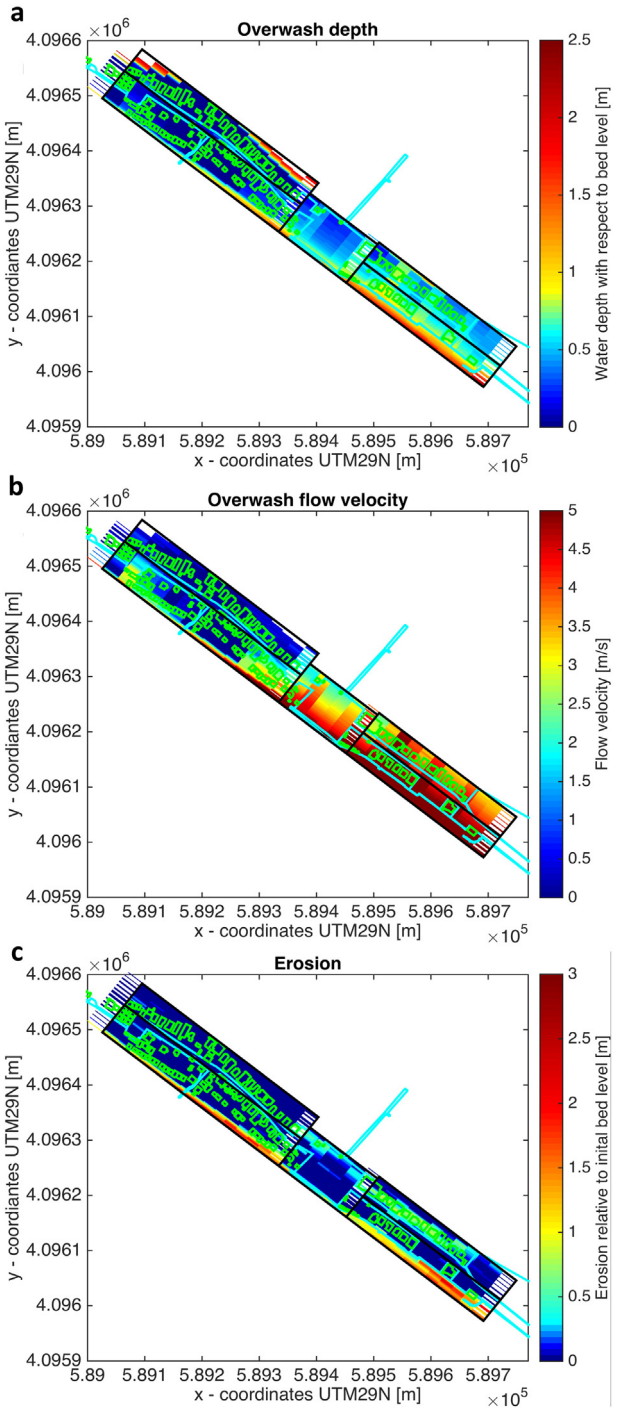

Fig. 7. XBeach model results for the synthetic storm with a significant wave height of 8.1 m, peak wave period of 18.5 s, duration of 306 h and surge level of 0.80 m. a): maximum values of the overwash water depth. b): Maximum overwash flow velocities. c): Maximum bed erosion during the event. Sedimentation is not shown as it is not considered a hazard.

3.4.3. BN configurations

Since the nodal structure and discretisation of the nodes of a BN influences its predictive skill, three BN configurations have been set up with a varying number of nodes and bins (Figs. 8, 9 and 10) of which the predictive skills are assessed in Section 3.4.4.

All configurations show the hydraulic boundary condition nodes (water level during peak, maximum significant wave height, peak period and storm duration) on the left. The case study site location node is shown on the right and the local hazard intensity nodes (overwash depth at buildings and infrastructure, overwash flow velocity at buildings and at infrastructure, erosion at buildings and infrastructure) are in the middle column.

In the first BN (Fig. 8), defined as Configuration 1, each node is discretized in four bins. The BN shows histograms of the prior probability of occurrence of a quantity in each bin. The mean value and the standard deviation are shown at the bottom of each node. Because of the narrow first bin for overwash depths of 0 to 0.1 m and a first bin for erosion of 0 to 0.5 m, which effectively signify "no overwash", and "little erosion", this BN configuration gives insight into whether or not overwash or erosion will occur. When overwash or erosion does occur, the three other bins provide a quantification of the overwash and erosion extent. Configuration 2 is a *binary* or green light/red light version in which the boundary conditions are discretised in four bins and the hazards in two (Fig. 9). This BN gives a prediction whether or not overwash or erosion will take place above a certain threshold, as specified by the bin sizes. It does not however, give quantitative insight into the hazards as the first configuration did. Configuration 3 (Fig. 10) is similar to Configuration 2 but has fewer boundary condition nodes.

The CPTs of the hazard nodes of the BN differ because of the different structures (Eq. (3)). Configuration 1 will have the largest CPTs and therefore the lowest experience values. The CPTs for Configuration 2 are smaller and will therefore have higher experience values. The cost of these higher experience values is that the magnitude of the hazard is not resolved. Configuration 3 has the smallest CPTs and therefore the highest experience values. The cost for this BN is that fewer boundary conditions can be considered. This increases the spread in the updated probability histograms of the hazards meaning that it is not necessarily a better performing BN (not shown).

3.4.4. Bayesian Network performance

The three BN configurations have been tested according to the validation method described in section 2.5. LLRs have been calculated for the individual hazard nodes using Eq. (3) and by conditioning on the hydraulic boundary conditions and areas (Table 5). The LLR scores are positive for all three configurations, indicating predictive skill; on average the updated probabilities for the bin in which the realisation falls for the three BNs are larger than their prior probabilities (Eq. (3)). These numbers can however not be used for an inter-comparison because their prior probabilities are not the same.

Using Eq. (4), LLRs are calculated to compare the three configurations (Table 6). Configurations 2 and 3 score positively against Configuration 1, which means the latter one is the lowest performing configuration. This is in line with expectations since this BN has the

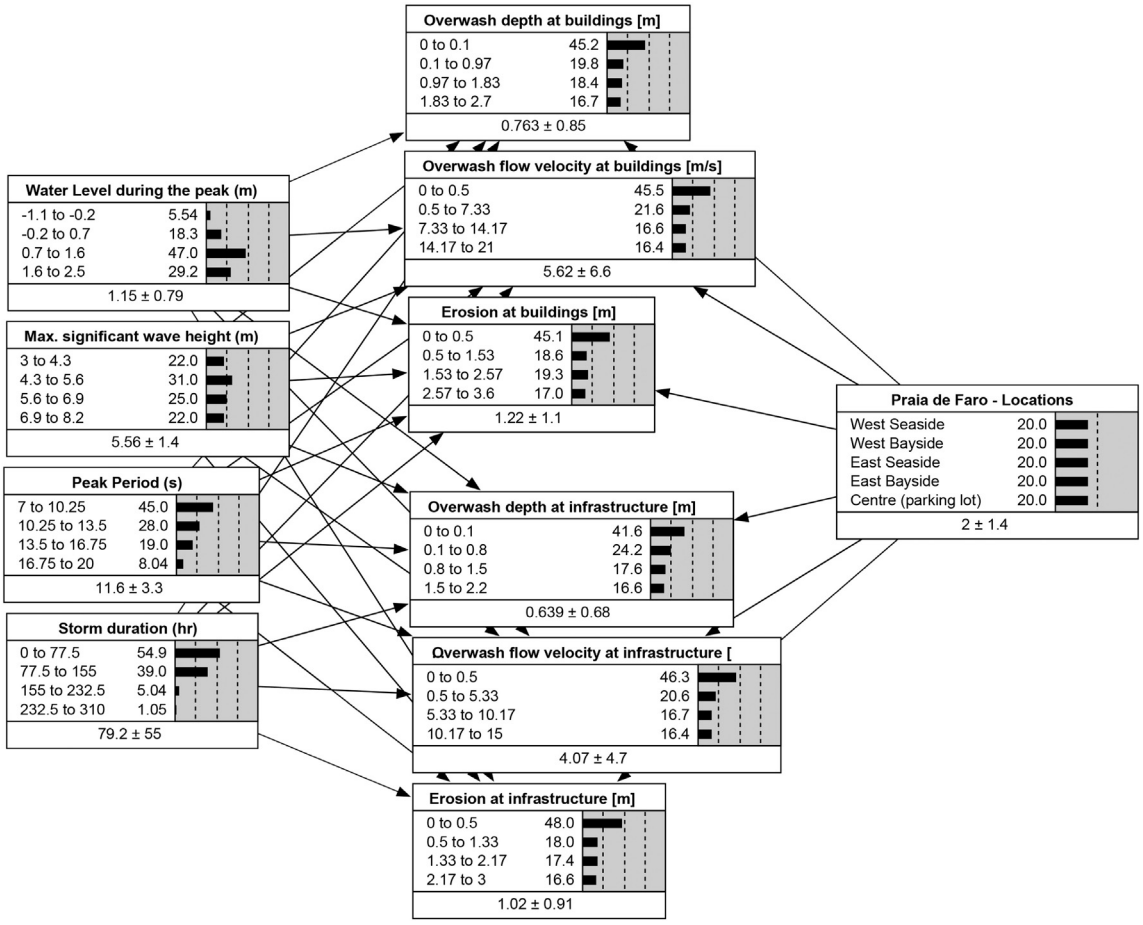

Fig. 8. Configuration 1 of the Bayesian Network (BN) with each node discretized into four bins. Four hydraulic boundary conditions are considered: the water level during the peak of the storm, the maximum significant wave height, the peak period and the storm duration. As hazard intensities the overwash depth and flow velocity and the erosion are considered for buildings as well as for infrastructure. A prediction can be requested for five output locations as specified in Fig. 7. The BN shows the prior probabilities.

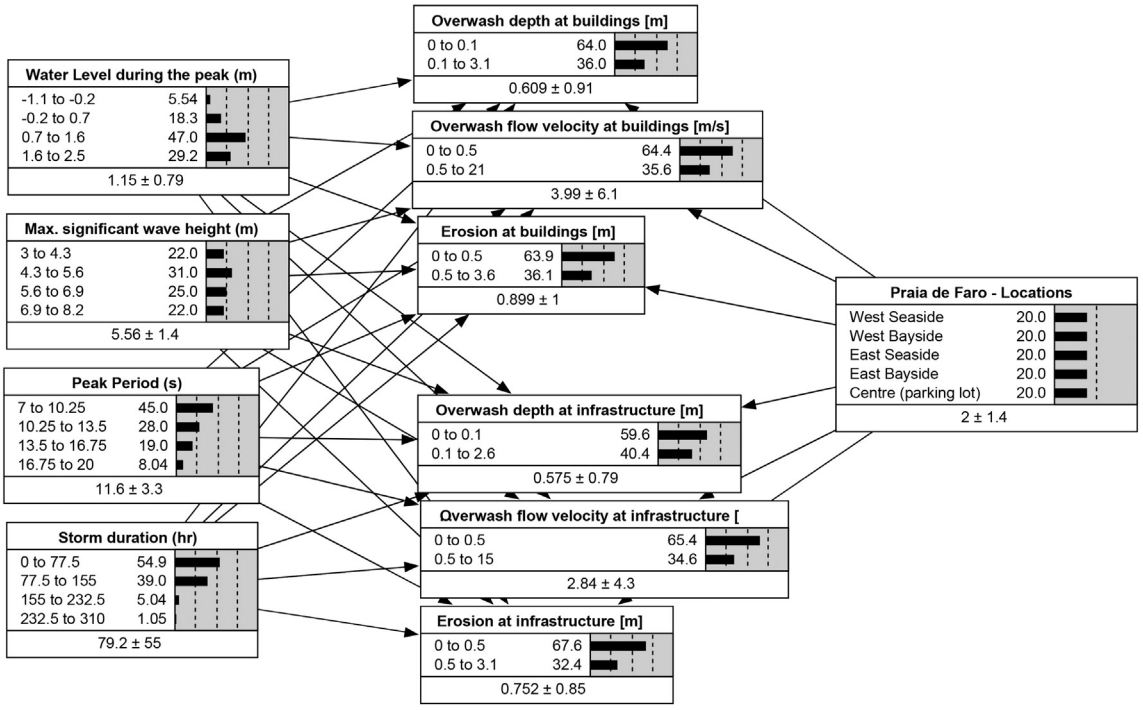

Fig. 9. Configuration 2 of the Bayesian Network (BN) with the boundary conditions discretized into four bins and the hazard intensities discretized into two bins. Four boundary conditions are considered: the water level during the peak of the storm, the maximum significant wave height, the peak period and the storm duration. As hazard intensities the overwash depth and flow velocity and the erosion are considered for buildings as well as for infrastructure. A prediction can be requested for five output locations as specified in Fig. 7. The BN shows the prior probabilities.

largest CPTs and therefore the lowest experience values. The two simpler BN have very similar performance, as indicated by the small LLR values in their inter-comparison (Table 6, bottom row). This is expected since the only difference is the reduced number of boundary condition nodes. Less intuitive is that Configuration 2, while more complex, shows better performance than Configuration 3. This can be explained by the spread in the data; fewer boundary conditions can be selected meaning that there is a larger spread in the data for the simpler BN.

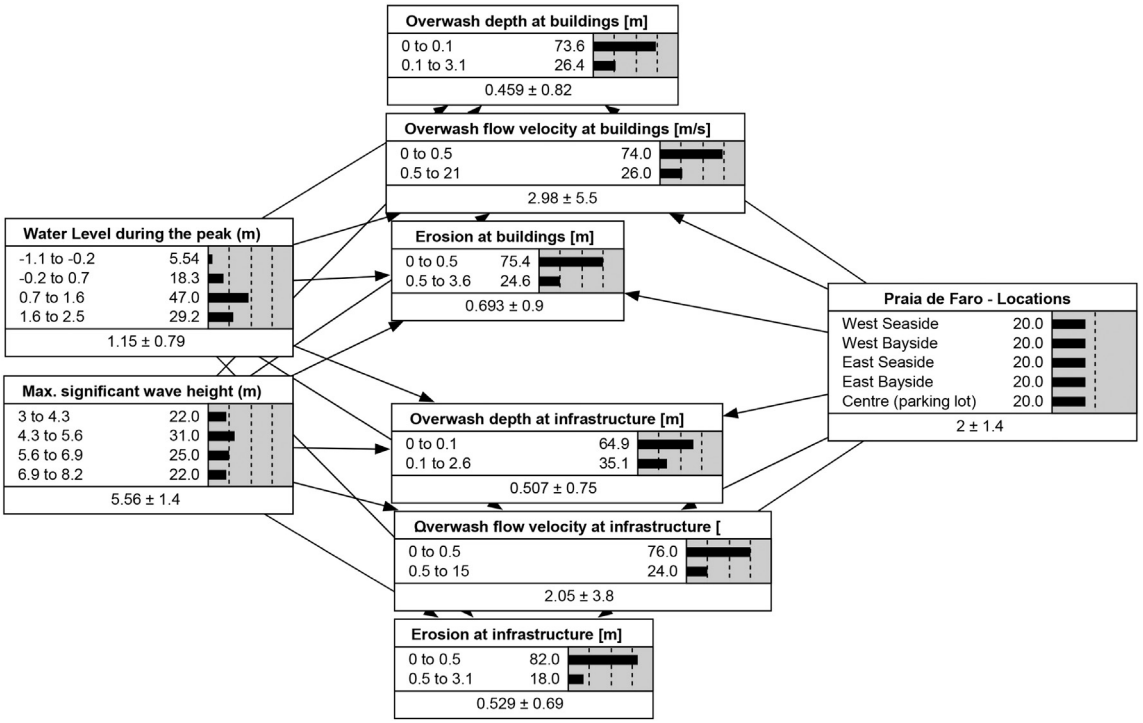

Fig. 10. Configuration 3 of the Bayesian Network (BN) with only two hydraulic boundary conditions, discretized into four bins, and the hazard intensities discretized into two bins. Two boundary conditions are considered: the water level during the peak of the storm and the maximum significant wave height. As hazard intensities the overwash depth and flow velocity and the erosion are considered for buildings as well as for infrastructure. A prediction can be requested for five output locations as specified in Fig. 7. The BN shows the prior probabilities.

Table 5
Log-likelihood ratio (LLR) test scores for three different BN setups. The LLR scores are the average of all tested datasets and given per node. A positive score means that on average the updated probabilities of the BN were larger than its prior probabilities and thus indicates predictive skill.

BN Configurations	Overwash depth buildings	Overwash velocity buildings	Erosion buildings	Overwash depth infrastructure	Overwash velocity infrastructure	Erosion Infrastructure
1	38.49	41.56	38.29	40.13	39.57	33.12
2	28.69	27.89	30.03	28.89	27.00	22.54
3	22.27	20.96	25.57	24.67	18.75	17.14

Table 6
Log-Likelihood Ratio (LLR) test scores comparing the three BN configurations against each other. The LLR scores are the average of all tested datasets and given per node. A positive score means that the updated probabilities of the base model were larger on average than the updated probabilities of the competing model. A positive score therefore indicates that the base model has more predictive skill than the competing model.

Base model Configuration	Competing model	Overwash depth buildings	Overwash vel. Buildings	Erosion buildings	Overwash depth infrastructure	Overwash vel. Infrastructure	Erosion infrastructure
2	1	27.8	21.4	28.9	28.3	22.5	23.8
3	1	23.4	16.2	27.8	24.1	16.6	237
2	3	4.4	5.2	1.1	4.1	5.8	0.1

This larger spread is reflected by the BN in that it is less certain of a prediction than when it can be conditioned on more boundary conditions, as is the case in Configuration 2. In terms of physical processes this result is more intuitive, because it shows the importance of the wave period and the storm duration on the hazards: wave period is a very important factor in determining overwash and storm duration is an important factor for erosion.

The performance of the best scoring Configuration 2 is additionally assessed with a confusion matrix, based on the same principle of training the network with 90% of the cases and testing it with the remaining 10% of cases, for 10 randomly chosen subsets of the data. The confusion matrix displays the percentage of time that the state with the highest belief of a node was also the state in which the XBeach model results fell and the percentage of time that the state with the highest belief was not the state in which XBeach model result fell (Table 7). Overall, the BN predicts hazards with an accuracy ranging from 81% to 88% and predicts cases with no significant onshore hazards with an accuracy ranging from 90% to 95%. This indicates that the BN would perform well as a surrogate for an operational process-based model simulation.

Ideally, the BN should be validated against hazard data obtained during extreme events. However, at our case study site, there is insufficient and at best only sparse and qualitative data on these rare events. A complete validation dataset would need to include pre- and post-event data (topography, bathymetry, overwash depths, velocities, damages) as well as hydrodynamic forcing data during such extreme events.

Table 7
Confusion matrices for the hazard nodes of the *binary* BN. The confusion matrix gives the percentage of times that the state with the highest belief of a node was also the state in which the actual observation fell and the percentage of times that the state with the highest belief was not the state in which the actual observation fell.

		Predicted	
	Actual	0 to 0.1	0.1 to 3.1
Overwash depth buildings	0 to 0.1	93%	14%
	0.1 to 3.1	7%	86%
		0 to 0.5	0.5 to 21
Overwash velocity buildings	0 to 0.5	92%	15%
	0.5 to 21	8%	85%
		0 to 0.5	0.5 to 3.6
Erosion at buildings	0 to 0.5	95%	12%
	0.5 to 3.6	5%	88%
		0 to 0.1	0.1 to 2.6
Overwash depth at infrastructure	0 to 0.1	90%	17%
	0.1 to 2.6	10%	83%
		0 to 0.5	0.5 to 15
Overwash velocity at infrastructure	0 to 0.5	93%	14%
	0.5 to 15	7%	86%
		0 to 0.5	0.5 to 3.1
Erosion at infrastructure	0 to 0.5	95%	19%
	0.5 to 3.1	5%	81%

4. Discussion

This research describes a method to construct a BN, which is a surrogate for a complex process-based model and can be implemented in an EWS for urbanised sandy coasts. The method involves five steps: 1) synthesis of a dataset, (2) schematisation of the storm event, (3) construction of a process-based model, (4) setting up a BN and (5) determining assessment metrics. In this discussion we address the first and third points, and show two examples of application of the BN.

In the first step, the limited availability of offshore hydrodynamic observations has been addressed by creating a synthetic dataset following a statistical approach using copulas. This approach has the advantage over simpler linear regression methods of incorporating the natural variability in the synthetic dataset and thus in the BN. If this natural variability is not incorporated the system may end up in operational mode facing a storm that the BN is not trained for. A large variability in the synthetic data, however, also leads to a larger variability in the predictions of the BN, ultimately leading to a higher demand for training data. In other words, including the natural variability by using copulas means the model is a better representation of reality but ensuring sufficient training data is more difficult. The number of synthetic events needed for the application of this method will therefore depend for a large part on the scatter in the observations.

Another advantage of using copulas over classical bivariate distributions is that the marginal distributions of the variables do not have to be of the same family, such that the marginal behaviour is better described. This advantage was utilized in the application to Praia de Faro since the marginal behaviour of the variables was best described by different distributions (Table 2). In the method presented here, bivariate copulas are used for the creation of the synthetic storm dataset, linking the significant wave height to the peak period, surge level and storm duration. In principle, the approach can be extended to higher dimensional copulas, which combine more than two variables. Whether this has

benefits for the application to hydrodynamic data is yet to be determined. Although more advanced than more commonly-used methods, the use of copulas and marginal distributions in the generation of synthetic datasets still has its limitations as illustrated by outliers in the significant wave height versus peak wave period plot (Fig. 5b). Wave heights of 7 m with a period of 8 s are very steep and would in fact break. Furthermore, the limitation that the existence of unknown asymptotes due to unknown physical limits is not addressed.

In the third step, a single measured topography was used as input, which does not reflect the site's natural dynamics of quick recovery and erosion periods (Vousdoukas et al., 2012b). This affects the result of a prediction since the berm width used in the model may be different than the actual one. Furthermore, the effect of storm sequences (Ferreira, 2005; Callaghan et al., 2008; Splinter et al., 2014; Karunarathna et al., 2014) in between which the beach does not have time to fully recover, is not considered in the current approach. Both aspects could be included by adding another node with different representative beach types.

A strength of the BN is its versatility of its application. It can both be used in a forecast mode as part of an EWS and as an analysis tool for coastal management. Two examples are shown. In Fig. 11, Configuration 2 has been conditioned on the hydraulic boundary condition nodes and on the *eastern seaside* location. If the BN is used as part of an EWS, the nodes would be conditioned with observations, which may come from either an operational wave-surge model or from real-time buoy observations. In the case of an analysis tool, design conditions may be used to assess the vulnerability of each area. In either way, the conditioning constrains the probability distribution of the hydraulic boundary condition nodes, but also provides an update of the probability of the hazards. The BN predicts overwash and erosion above the given thresholds with a large certainty for buildings. For infrastructure it also gives a large probability for overwash above the threshold, but indicates that erosion will most likely stay below the threshold.

In the second example the BN is used in a reverse way (Fig. 12). It is provided with a hypothetical onshore hazard scenario in which all intensities are above the specified thresholds. The BN has now updated the probability histograms for the hydraulic boundary conditions, which gives an indication what forcing conditions would be responsible for causing such a response. The BN has also updated the histogram for the locations, indicating that *seaside* and the *centre* are the most likely areas for which the hazard intensities are largest, since the *bayside* is further from the sea and in the central area the dunes are lowest. Furthermore, the updated histograms for the boundary conditions indicate which states of the hydraulic boundary conditions are most likely to cause these hazards.

The method is not limited to Praia de Faro but can be applied to other coasts which experience limited observations of offshore hydrodynamic parameters or for which the computational time of a coastal response model is too long for its use in an operational EWS.

The method is not specific to the process-based model that is presently used. The quality of the results of the BN is as good as the underlying process-based or empirical model, which needs to capture the essential physics. Any model that has better physics or is calibrated on better data would yield a better EWS.

Furthermore, the BN can be adjusted to fit any relevant connections between hydraulic boundary conditions and onshore hazards. Furthermore, a BN can be continuously updated with new information and expanded to include different morphological conditions or risk reduction measures. Finally, if the computational demand of a single process-based model simulation is not a concern, a BN still has the advantage that it is a probabilistic approach rather than deterministic; the BN returns a prediction in the form of a probability density function that is based on more than one numerical simulation, and thus gives an indication of the uncertainties. Therefore, BNs are a promising extension of existing EWSs and a valuable planning tool for coastal managers.

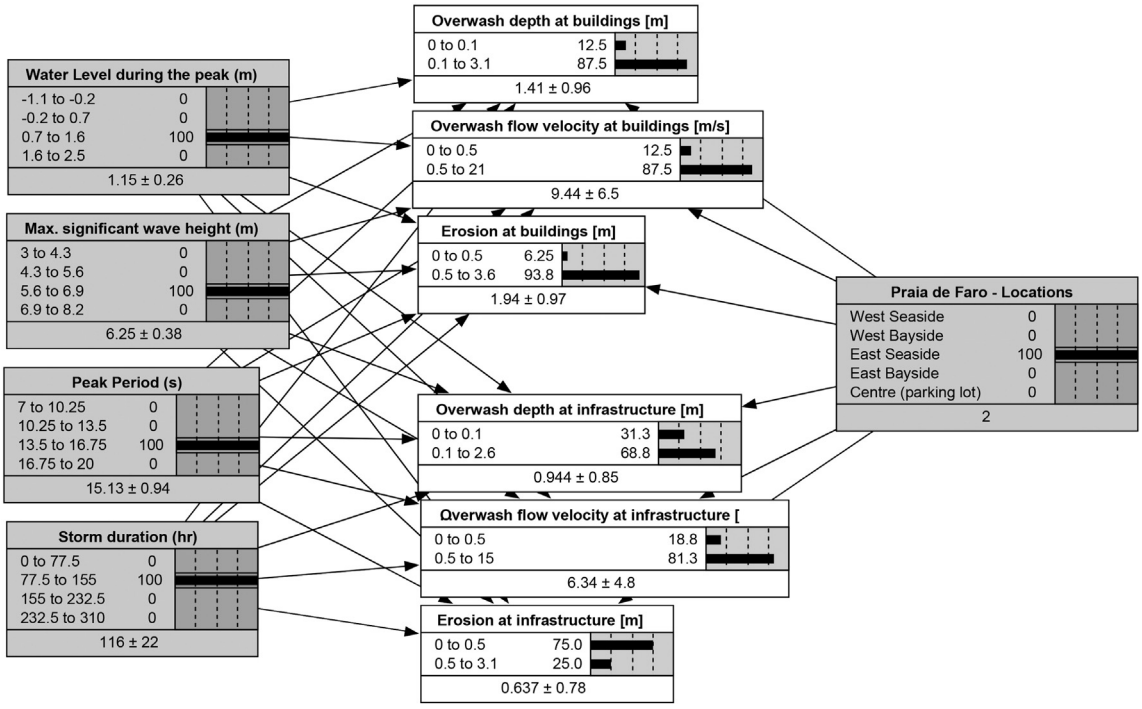

Fig. 11. Example of the application of Configuration 2 conditioned on the hydraulic boundary conditions of an event at the location *East Seaside*. Based on the observations the BN has updated the histograms of the hazard nodes. The BN indicates large probabilities of overwash (>0.1 m and 0.5 m/s) and erosion (>0.5 m) for buildings and a large probability for overwash at infrastructure (>0.1 m and 0.5 m/s). It also indicates that the probability of erosion at the infrastructure is low (<0.5 m).

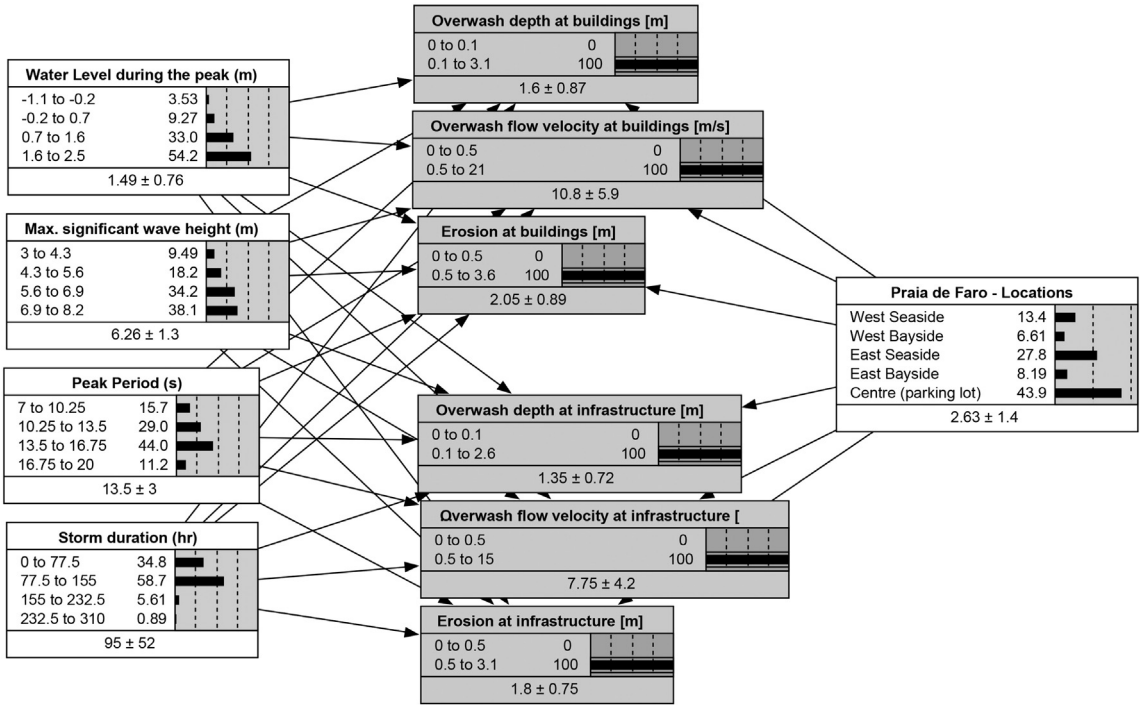

Fig. 12. Example of the application of Configuration 2 conditioned with observations of the hazard nodes. Based on the observations the BN has updated the histograms of the hydraulic boundary condition nodes and the location node. The BN indicates which areas are most vulnerable to the hazards and which state of the boundary conditions are most likely to be the cause.

5. Summary and conclusions

A method has been developed to construct a probabilistic Bayesian Network (BN), which acts as a surrogate for a process-based model, and can be used as part of an Early Warning System (EWS) for sandy coasts.

The BN connects three elements: hydraulic boundary conditions at the 20 m depth contour, characteristics of the coastal zone, and onshore hazards. Hydraulic boundary conditions were derived from a statistical analysis of observed data using copulas, and site characteristics were obtained from measurements. This BN was trained using output data from many pre-computed process-based model simulations, which connect the three elements. Once trained, the response of the BN is instantaneous and can be applied in operational mode as a surrogate for the process-based model. As part of an EWS, the BN can be updated with an observation of the hydraulic boundary conditions to give a prediction for onshore hazards such as erosion, overwash depth and velocities. As an analysis tool, the BN can be used to assess the effect of constraining the probability of one or more variables on the rest of the network.

The method was applied to Praia de Faro, Portugal, a low-lying urbanised barrier island, which is subject to frequent flooding. Using a copula-based statistical analysis, which preserves the natural variability of the observations, a synthetic dataset containing 100 events was created, based on 20 years of observations, but extended to return periods of significant wave height of up to 50 years. These events, characterised by the significant wave height, peak wave period, maximum water level and the storm duration, were transformed from offshore to onshore using a 2D XBeach (Roelvink et al., 2009) model. The onshore hazard intensities which are predicted are erosion, overwash depth and flow velocity.

Three BNs configurations were constructed, of which the best performing one was able to predict onshore hazards as computed by the model with an accuracy ranging from 81% to 88% and predict no significant onshore hazards with an accuracy ranging from 90% to 95%. Two examples were presented on the use of a BN in operational predictions or as an analysis tool.

The added value of this method is that it generic enough to be applied to other coastal sites. Thus, BNs are a promising extension of existing EWSs and a valuable planning tool for coastal managers.

Acknowledgements

All authors received funding through the RISC-KIT (Resilience-increasing Strategies for Coasts—Toolkit) project, supported by the European Community's Seventh Framework Programme through contract no. 603458. Ap van Dongeren and Robert McCall were additionally supported through Deltares Strategic Funding on “Hydro- and Morphodynamics during Extreme Events” (project number 1230002). We thank the Portuguese Hydrografic Office and the Spanish Port Authorities for supplying wave and surge data, respectively.

Appendix A

Table A. 1
XBeach calibration settings. All other parameter settings are defaults.

Parameter	Value	Units	Description
D50	0.005	m	D50 grain size per grain type
D90	0.002	m	D90 grain size per grain type
bedfriction	Manning	–	Bed friction formulation
bedfriccoef	0.02 to 0.04	s/m1/3	Bed friction coefficient
delta	0.2	–	Fraction of wave height to add to water depth
gammax	2.364	–	Maximum ratio wave height to water depth
beta	0.138	–	Breaker slope coefficient in roller model
alpha	1.262	–	Wave dissipation coefficient in Roelvink formulation
facSK	0.2	–	Calibration factor time averaged flows due to wave skewness
facAs	0.4	–	Calibration factor time averaged flows due to wave asymmetry
gamma	0.541	–	Breaker parameter in Baldock or Roelvink formulation
wetslp	0.26	–	Critical avalanching slope under water.
hswitch	0	m	Water level at which is switched from wetslp to dryslp.
Gamma_js	3.3	–	JONSWAP (Hasselmann et al., 1980) peak enhancement factor
s	20	–	directional spreading coefficient, characterising a mix of sea and swell waves (Goda, 1985)

References

Achab, M., Ferreira, O., Alverinho Dias, J.M., 2014. Evaluation of sedimentological and morphological changes induced by the rehabilitation of sandy beaches from the Ria Formosa barrier island system (South Portugal). Thalassas 30, 21–31.

Almeida, L.P., Ferreira, Ó., Vousdoukas, M.I., Dodet, G., 2011. Historical variation and trends in storminess along the Portuguese South Coast. Nat. Hazards Earth Syst. Sci. 11, 2407–2417.

Almeida, L.P., Vousdoukas, M.V., Ferreira, Ó., Rodrigues, B.A., Matias, A., 2012. Thresholds for storm impacts on an exposed sandy coastal area in southern Portugal. Geomorphology 143–144, 3–12. http://dx.doi.org/10.1016/j.geomorph.2011.04.047.

Bertin, X., Bruneau, N., Breilh, J.F., Fortunato, A.B., Karpytchev, M., 2012. Importance of wave age and resonance in storm surges: the case Xynthia, Bay of Biscay. Ocean Model 42, 16–30. http://dx.doi.org/10.1016/j.ocemod.2011.11.001.

Blake, E.S., Kimberlain, T.B., Berg, R.J., Cangialosi, J., Beven II, J.L., 2013. Tropical Cyclone Report: Hurricane Sandy. http://dx.doi.org/10.1017/CBO9781107415324.004.

Callaghan, D.P., Nielsen, P., Short, A., Ranasinghe, R., 2008. Statistical simulation of wave climate and extreme beach erosion. Coast. Eng. 55, 375–390.

Cañizares, R., Irish, J.L., 2008. Simulation of storm-induced barrier island morphodynamics and flooding. Coast. Eng. 55, 1089–1101.

Carrasco, A.R., Ferreira, Ó., Matias, A., Freire, P., 2012. Flood hazard assessment and management of fetch-limited coastal environments. Ocean Coast. Manag. 65, 15–25. http://dx.doi.org/10.1016/j.ocecoaman.2012.04.016.

Castelle, B., Marieu, V., Bujan, S., Splinter, K.D., Robinet, A., Sénéchal, N., Ferreira, S., 2015. Impact of the winter 2013–2014 series of severe Western Europe storms on a double-barred sandy coast: beach and dune erosion and megacusp embayments. Geomorphology 238, 135–148. http://dx.doi.org/10.1016/j.geomorph.2015.03.006.

Corbella, S., Stretch, D.D., 2013. Simulating a multivariate sea storm using Archimedean copulas. Coast. Eng. 76, 68–78. http://dx.doi.org/10.1016/j.coastaleng.2013.01.011.

Den Heijer, C., Knipping, D.T.J.A., Plant, N.G., Van Thiel de Vries, J.S.M., Baart, F., Van Gelder, P.H.A.J.M., 2012. Impact assessment of extreme storm events using a Bayesian network. Proc. Coast. Eng. Conf. (http://www.scopus.com/inward/record.url?eid=2-s2.0-84884938068&partnerID=tZOtx3y1).

Dolan, R., Davis, R.E., 1992. An intensity scale for Atlantic Coast northeast storms. J. Coast. Res. 8, 840–853.

Ferreira, Ó., 2005. Storm groups versus extreme single storms: predicted erosion and management consequences. J. Coast. Res. 221–227 (<Go to ISI>://WOS:000230629600025).

Ferreira, Ó., Garcia, T., Matias, A., Taborda, R., Dias, J.M.A., 2006. An integrated method for the determination of set-back lines for coastal erosion hazards on sandy shores. Cont. Shelf Res. 26, 1030–1044. http://dx.doi.org/10.1016/j.csr.2005.12.016.

Fortunato, A.B., Li, K., Bertin, X., Rodrigues, M., Martín Miguez, B., 2016. Determination of extreme sea levels along the Iberian Atlantic coast. Ocean Eng. 111, 471–482. http://dx.doi.org/10.1016/j.oceaneng.2015.11.031.

Genest, C., Favre, A., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 347–368 http://dx.doi.org/10.1061/(ASCE)1084–0699(2007)12:4(347) (accessed July 21, 2015).

Gerritsen, H., 2005. What happened in 1953? The big flood in the Netherlands in retrospect. Philos. Transact. A. Math. Phys. Eng. Sci. 363, 1271–1291. http://dx.doi.org/10.1098/rsta.2005.1568.

Goda, Y., 1985. Random Seas and Design of Maritime Structures. University of Tokyo Press.

Gutierrez, B.T., Plant, N.G., Thieler, E.R., 2011. A Bayesian network to predict coastal vulnerability to sea level rise. J. Geophys. Res. 116. http://dx.doi.org/10.1029/2010JF001891.

Gutierrez, B.T., Plant, N.G., Thieler, E.R., Turecek, A., 2015. Using a Bayesian network to predict barrier island geomorphologic characteristics. J. Geophys. Res. 120, 2452–2475. http://dx.doi.org/10.1002/ 2015JF003671.

Hapke, C., Plant, N., 2010. Predicting coastal cliff erosion using a Bayesian probabilistic model. Mar. Geol. 278, 140–149. http://dx.doi.org/10.1016/j.margeo.2010.10.001.

Hasselmann, D.E., Dunckel, M., Ewing, J.A., 1980. Directional wave spectra observed during JONSWAP 1973. J. Phys. Oceanogr. 10, 1264–1280. http://dx.doi.org/10.1175/1520-0485(1980)010%3C1264:DWSODJ%3E2.0.CO;2.

Jäger, W.S., Morales Nápoles, O., 2014. Sampling Joint Time Series of Significant Wave Heights and Periods in the North Sea.

Joes, H., 2015. Dependence Modeling with Copulas. 1st ed. Chapman and Hall/CRC.

Karunarathna, H., Pender, D., Ranasinghe, R., Short, A.D., Reeve, D.E., 2014. The effects of storm clustering on beach profile variability. Mar. Geol. 348, 103–112.

Knabb, R.D., Rhome, J.R., Brown, D.P., 2006. Tropical Cyclone Report: Hurricane Katrina. Natl. Hurric. Cent. http://dx.doi.org/10.1017/CBO9781107415324.004.

Kriebel, D.L., Dean, R.G., 1993. Convolution method for time dependent beach profile response. J. Waterw. Port Coast. Ocean Eng. 119, 204–226. http://dx.doi.org/10.1061/(ASCE)0733-950X(1993)119:2(204).

Larson, M., Erikson, L., Hanson, H., 2004. An analytical model to predict dune erosion due to wave impact. Coast. Eng. 51, 675–696. http://dx.doi.org/10.1016/j.coastaleng.2004.07.003.

Leatherman, S.P., 1990. Modelling shore response to sea-level rise on sedimentary coasts. Prog. Phys. Geogr. 14, 447–464.

Lindemer, C.A., Plant, N.G., Puleo, J.A., Thompson, D.M., Wamsley, T.V., 2010. Numerical simulation of a low-lying barrier island's morphological response to Hurricane Katrina. Coast. Eng. 57, 985–995. http://dx.doi.org/10.1016/j.coastaleng.2010.06.004.

Masselink, G., Castelle, B., Scott, T., Dodet, G., Suanez, S., Derek, J., Floc'h, F., 2015a. Extreme wave activity during 2013/14 winter and impacts along the Atlantic coast of Europe. Nat. Geosci. (submitted).

Masselink, G., Scott, T., Poate, T., Russell, P., Davidson, M., Conley, D., 2015b. The extreme 2013/14 winter storms: hydrodynamic forcing and coastal response along the southwest coast of England. Earth Surf. Process. Landf. 41, 378–391. http://dx.doi.org/10.1002/esp.3836.

McCall, R.T., Van Thiel de Vries, J.S.M., Plant, N.G., Van Dongeren, A.R., Roelvink, J.A., Thompson, D.M., Reniers, A.J.H.M., 2010. Two-dimensional time dependent hurricane overwash and erosion modeling at Santa Rosa Island. Coast. Eng. 57, 668–683. http://dx.doi.org/10.1016/j.coastaleng.2010.02.006.

Mendoza, E.T., Jiménez, J.A., 2006. Storm-induced erosion potential on the Catalonian coast. J. Coast. Res. 81–88. http://dx.doi.org/10.1061/40855(214)98.

Miller, J.K., Livermont, E., 2008. An index for predicting storm erosion due to increased waves and water levels. Disasters. ASCE , pp. 561–572http://link.aip.org/link/?ASCECP/312/51/1.

Mungar, S., Kraus, N.C., 2010. Frequency of Extreme Storms Based on Beach Erosion at Northern Assateague Island, Maryland, Shore & Beach. 78 pp. 3–11.

Nederhoff, C.M., Lodder, Q.J., Boers, M., den Bieman, J.P., Miller, J.K., 2014. Modeling the Effects of Hard Structures on Dune Erosion and Overwash: A Case Study of the Impact of Hurricane Sandy on the New Jersey Coast. Coast. Sediments 2015, San Diego, USA.

Norsys, 2016. Netica help. https://www.norsys.com/WebHelp/NETICA.htm (accessed March 17, 2016).

Pearl, J., Russel, S., 1988. Bayesian Networks. University of California http://dx.doi.org/10.1145/1859204.1859227.

Pires, H.O., 1998. Project INDIA. Preliminary report on wave climate at Faro, Lisboa, Portugal. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Project+INDIA.+Preliminary+report+on+wave+climate+at+Faro#0 (accessed March 18, 2015).

Plant, N.G., Holland, K.T., 2011. Prediction and assimilation of surf-zone processes using a Bayesian network. Coast. Eng. 58, 119–130. http://dx.doi.org/10.1016/j.coastaleng.2010.09.003.

Plomaritis, T., Benaventa, J., Del Ria, L., Reyes, E., Dastis, C., Gómez, M., Bruno, M., 2012. Storm Early Warning System as a Last Plug-in of a Regional Operational Oceanography System: The Case of the Gulf of Cadiz. Coast. Eng. Proceedings; no 33, Santander, Spain.

Ranasinghe, R., Swinkels, C., Luijendijk, A., Roelvink, D., Bosboom, J., Stive, M., Walstra, D., 2011. Morphodynamic upscaling with the MORFAC approach: dependencies and sensitivities. Coast. Eng. 58, 806–811.

Rodrigues, B.A., Matias, A., Ferreira, Ó., 2012. Overwash hazard assessment. Gelologica Acta. 10, 427–437. http://dx.doi.org/10.1344/105.000001743.

Roelvink, J.A., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., Lescinski, J., 2009. Modelling storm impacts on beaches, dunes and barrier islands. Coast. Eng. 56, 1133–1152. http://dx.doi.org/10.1016/j.coastaleng.2009.08.006.

Sallenger, A.H.J., 2000. Storm impact scale for barrier islands. J. Coast. Res. 16, 890–895http://www.jstor.org/stable/4300099 (accessed March 17, 2015).

Schmidt, T., 2007. Coping with copulas. In: Rank, J. (Ed.), Copulas From Theory to Appl. Financ. Risk Books, London, pp. 3–34.

Sklar, A., 1959. Fonctions de répartition à N Dimensions et Leurs Marges. 8. Publ. l'Institut Stat. l'Université Paris, pp. 229–231.

Smallegan, S.M., Irish, J.L., van Dongeren, A.R., den Bieman, J.P., 2015. Morphological response of a sandy barrier island with a buried seawall during Hurricane Sandy. Coast. Eng.

Splinter, K.D., Carley, J.T., Golshani, A., Tomlinson, R., 2014. A relationship to describe the cumulative impact of storm clusters on beach erosion. Coast. Eng. 83, 49–55.

Stockdon, H.F., Sallenger, A.H., Holman, R.A., Howd, P.A., 2007. A simple model for the spatially-variable coastal response to hurricanes. Mar. Geol. 238, 1–20. http://dx.doi.org/10.1016/j.margeo.2006.11.004.

Sutherland, J., Peet, A.H., Soulsby, R.L., 2004. Evaluating the performance of morphological models. Coast. Eng. 51, 917–939.

van Rijn, L., Walstra, D.J., Grasmeijer, B., Sutherland, J., Pan, S., Sierra, J., 2003. The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based profile models. Coast. Eng. 47, 295–327. http://dx.doi.org/10.1016/S0378-3839(02)00120-5.

Van Thiel de Vries, J.S.M., 2009. Dune Erosion during Storm Surges. Delft Univesity of Technology.

van Verseveld, H.C.W., van Dongeren, A.R., Plant, N.G., Jäger, W.S., den Heijer, C., 2015. Modelling multi-hazard hurricane damages on an urbanized coast with a Bayesian Network approach. Coast. Eng. 103, 1–14. http://dx.doi.org/10.1016/j.coastaleng.2015.05.006.

Vousdoukas, M.I., Ferreira, Ó., Almeida, L.P., Pacheco, A., 2012a. Toward reliable storm-hazard forecasts: XBeach calibration and its potential application in an operational early-warning system. Ocean Dyn. 62, 1001–1015. http://dx.doi.org/10.1007/s10236-012-0544-6.

Vousdoukas, M.I., Almeida, L.P.M., Ferreira, Ó., 2012b. Beach erosion and recovery during consecutive storms at a steep-sloping, meso-tidal beach. Earth Surf. Process. Landf. 37, 583–593. http://dx.doi.org/10.1002/esp.2264.

Wahl, T., Mudersbach, C., Jensen, J., 2012. Assessing the hydrodynamic boundary conditions for risk analyses in coastal areas: a multivariate statistical approach based on Copula functions. Nat. Hazards Earth Syst. Sci. 12, 495–510.

Wright, L.D., Short, A.D., 1984. Morphodynamic variability of surf zones and beaches: a synthesis. Mar. Geol. 56, 93–118http://www.sciencedirect.com/science/article/pii/0025322784900082 (accessed March 17, 2015).