

Synthetic events for flood risk calculation by using a nested Copula structure

Tim Franken^{1,a}, Joris Blanckaert¹ and Gert Leyssen¹

¹International Marine and Dredging Consultants, Antwerp, Belgium

Risk analysis requires considering the entire frequency domain of flood consequences. Synthetic events were generated for the entire river system of the Scheldt estuary. This estuary contains multiple navigable waterways and is situated in Belgium and the Netherlands. Extreme water levels are influenced by rainfall-runoff discharges, tiding, storm surges, and wind speed and direction. For the generation of hydraulic boundary conditions for flood risk assessment, these influences and their mutual dependencies and correlations are taken into account by means of a nested extreme value copula structure. The variation in time is taken into account by standardized profiles, computed by normalizing all recorded extreme events and fitting a probability distribution to the variation of the standardized events, yielding 5 profile classes through another stratification. Eventually this resulted in a total of 1920 sets of synthetic events. All events were run through the hydrodynamic model of the river system. The frequency distribution of the resulting water levels are calculated by accumulation of the corresponding probabilities of occurrence of the synthetic events at each location. The methodology has the advantage that it determines a statistical distribution of consequences, rather than assigning frequencies to hydrodynamic boundary conditions.

1 Introduction

Flood risk is calculated from the probability distribution of flood consequences, hence the probability distribution of inundation depths must be determined. This requires considering the entire frequency domain since severe consequences can compensate for extremely low frequencies and vice versa.

As part of the Flemish implementation of the European Flood Directive, a probabilistic approach has been developed and applied to assess flood risk in Flemish non-navigable waterway catchments [1]. This methodology uses synthetic hydrographs, generated by stratified sampling. The advantage of stratified sampling lies in the fact that it allows for determining a statistical distribution of the consequences, rather than assigning statistical frequencies to the loads, in this case the hydrodynamic boundary conditions.

This methodology has been further developed and applied to the entire river system of the Scheldt estuary. The estuary contains multiple navigable waterways situated in Belgium and the Netherlands. Chapter 2 provides a more detailed description of the estuary, the available data and the multiple interactions in the system. Chapter 3 gives a brief overview of the synthetic event methodology. Chapter 4 focuses on the stratification process to discretise the multivariate loading domain. The generation of synthetic events through standardized profiles is discussed in Chapter 5. The overall frequency

of occurrence of each event is determined in Chapter 6. After hydraulic simulation of the synthetic events an empirical frequency distribution of inundation depths is obtained. This is discussed in Chapter 7. Eventually Chapter 8 summarizes the main conclusions of this study.

2 The Scheldt estuary and boundary conditions

The Scheldt estuary has a total drainage area of 21 860 km² and contains multiple navigable waterways situated in Belgium and the Netherlands. Extreme water levels are influenced by rainfall-runoff discharges, tiding, storm surges and wind set-up. The water level in the rivers and inundation depths in the flood plains can be simulated using a 1D hydrodynamic model developed in Mike11 (Figure 1).

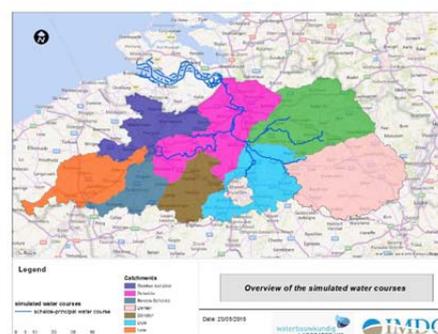


Figure 1. Overview of the Scheldt estuary

^a Corresponding author: tim.franken@imdc.be

This model has a downstream water level boundary in Vlissingen and 42 upstream discharge boundaries. The influence of wind speed and direction is also included in the model. For each boundary condition of the hydrodynamic model a long-term time series was available.

Lumped hydrological models were used to generate time series of rainfall-runoff for 40 upstream boundaries. The hydrological models were fed by catchment averaged rainfall, calculated using the Thiessen polygon method. Evaporation was calculated using the Penman-Monteith method. The length of calculated time series of rainfall runoff discharge was 28 years.

The two remaining upstream discharge boundaries were derived from long-term simulations with 1D hydrodynamic models of the upstream water courses. These models were fed with similar rainfall data as described for the hydrological models.

Measured water level data at Vlissingen was provided by Flanders Hydraulics at the downstream boundary of the model. The water level time series has been divided in a time series of astronomical tide and a time series of skewed storm surge, i.e. the difference between measured high water level and the corresponding astronomical high water level, using a harmonic analysis (Figure 2).

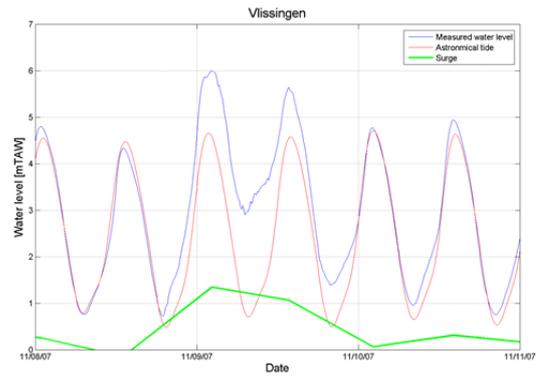


Figure 2. Measured water level time series filtered into an astronomical tiding time series and a skewed surge time series by harmonic analysis

Time series of wind data were made available by Flanders Hydraulics.

3 Synthetic events

Risk is by definition calculated by integrating consequences with their probabilities. This implies that extreme events with low frequencies but severe consequences as well as less extreme events with higher frequencies and less severe consequences must be considered:

$$R = \int S * dP \quad (1)$$

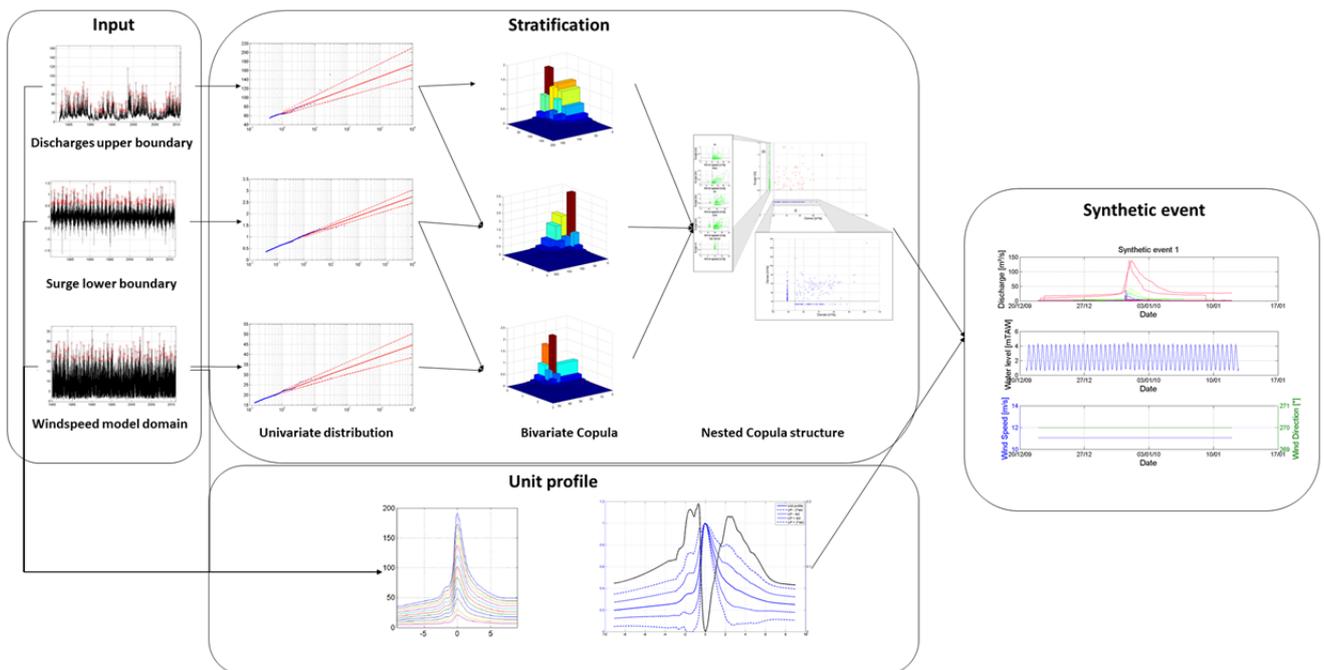


Figure 3. Schematic overview of the different steps to generate synthetic events for the Scheldt estuary

R usually converges asymptotically to a maximum, since flood risk consequences usually don't increase significantly for the extremely low probabilities. The formula above can be discretized by a stratification of the multivariate probability domain in k hypercubes:

$$R = \sum_{i=1}^k \sum_{\text{flood area}} \text{Consequence}_i * \Delta p_i \quad (2)$$

Hence a synthetic event has to be designed for each single hypercube or class Δp . A synthetic event consists of a set of synthetic boundary conditions which are generated by multiplying the time profiles of normalized unit profiles with the representative values of that particular class. The probability or frequency of occurrence of each class is determined by taking the multivariate frequency of each of the boundary conditions into account. Figure 3 gives a schematic overview of the different steps to generate synthetic events for the Scheldt estuary and compute the associated frequency of occurrence. Each of these steps will be discussed below.

4 Stratification

4.1 Peak over threshold

In this study Peak Over Threshold (POT) values are selected from the long-term time series of discharge, wind speed and storm surge. POT values are independent extremes exceeding a set threshold. The independency of the events is obtained by use of a maximum inter-event level and a minimum time lag between two successive events. The initial choice of the threshold has to be sufficiently low in order to include an average number of events of about 12 per year [2]. In a later step a higher optimized threshold will be determined.

4.2 Threshold excess models

The tail behaviour of extreme POT values can be modeled by the Generalized Pareto Distribution (GPD), a generalized distribution representing the tail behaviour of different threshold excess models. A model calibration method suggested by Beirlant [3] and Coles [4] is adopted in this study. By means of the mean excess function it is possible to get an estimate of the GPD shape parameter ξ , which determines the appropriate conditional distribution.

The mean excess is the mean of the excess values of all POT values. Each conditional distribution has a theoretical mean excess function which returns the mean excess as a function of the threshold. The shape of the empirical mean excess function of the POT values can be compared with theoretical mean excess functions of the different threshold models in order to select the appropriate conditional distribution. If the empirical mean excess function has an increasing trend, the

corresponding distribution will belong to the GPD ($\xi > 0$), the Pareto distribution or the Conditional Weibull distribution ($\tau < 1$). In case of a horizontal mean excess function the data will follow GPD ($\xi = 0$), or the Exponential Distribution. If the mean excess function is decreasing the model is GPD ($\xi < 0$), or the Conditional Weibull Distribution ($\tau > 1$) (Figure 4). After selecting a conditional threshold excess model, the parameter estimation of the threshold models is based on the choice for an optimal threshold, as shown in Figure 5.

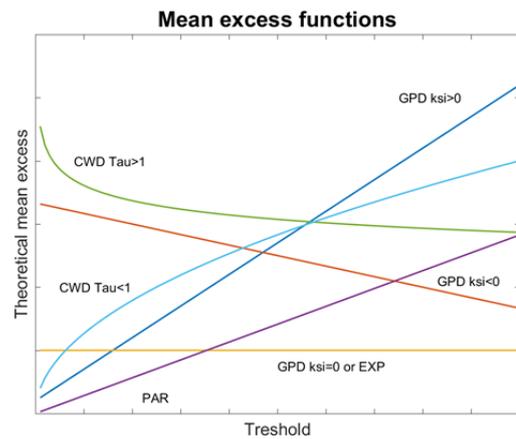


Figure 4. Mean excess functions

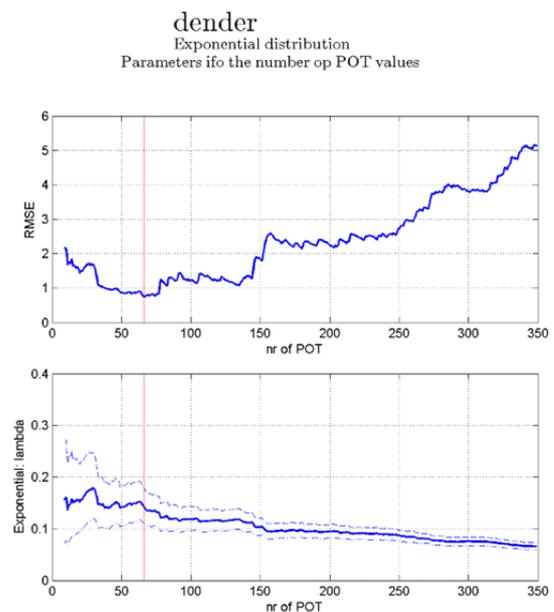


Figure 5. Selection of optimized threshold for a GPD with $\xi = 0$ (Exponential Distribution, flow discharge Dender).

In this example the mean excess comparison suggests an exponential distribution. Hence an exponential distribution is repeatedly fitted (Maximum Likelihood fit) to the POT values, taking into account a decreasing number of POT-values, corresponding to an increasing threshold. An optimized selection of the threshold is obtained if the RMSE of the fit is at a local minimum with the calculated scale parameter located in a stable

zone. With the accordingly computed scale and shape parameter, the threshold model can be drawn, as shown in Figure 6.

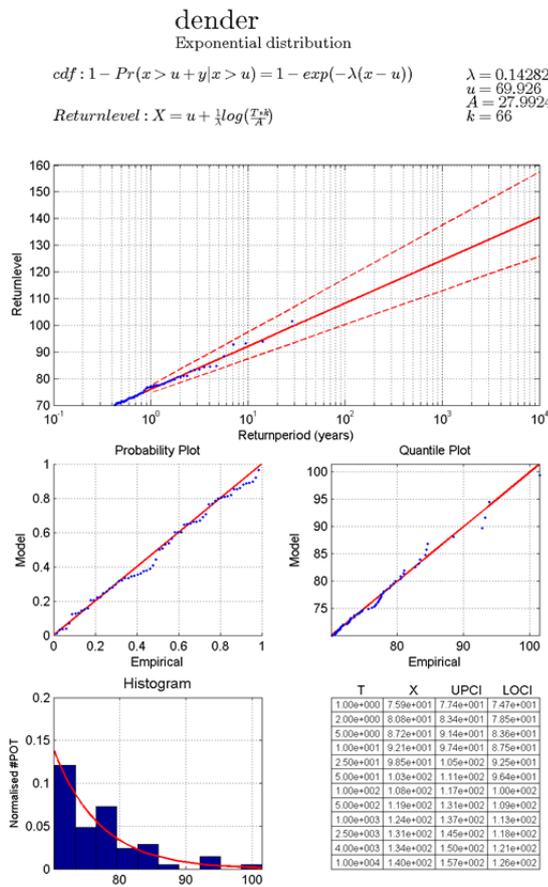


Figure 6. Univariate extreme value distribution for flow discharge Dender (Exponential Distribution)

4.3 Correlation analysis

The flood consequences along the Scheldt estuary are determined by multiple variables which can mutually influence each other. When multiple boundary conditions are involved two different approaches are available to generate the parameters for the synthetic events. The multivariate approach, in which a joint distribution for all variables is determined followed by a stratification of the multivariate domain, or the conditional approach, in which the different variables are related e.g. by means of a regression model. With a high number of variables, as is the case in the Scheldt estuary, the multivariate approach will result in a for computational reasons unrealistic number of events. The conditional approach has the advantage not to increase the number of synthetic events, yet it is only applicable for highly correlated variables.

In order to assess the most suitable method it is important to understand the influence and the mutual dependencies and correlations of all involved variables.

4.3.1 Correlation between discharge boundaries

The POT values of the upstream discharge boundaries have been coupled using a time window of two days around the peak discharge. The correlation between the coupled upstream discharge boundaries was studied using a correlation analysis. Three types of correlation (Spearman, Kendall and Pearson) have been calculated to assess for non-parametric and nonlinear correlation. Table 1 provides an overview of the corresponding parameters and the p-values for three of the largest catchments in the basin. Considering these values the hypothesis of independence can be rejected at a significance level of $\alpha = 0.05$. The underlying meteorological events are correlated.

		Demer - Dender	Dender - LBS	Demer - LBS
Pearson	ρ	0.40	0.64	0.54
	p-test	2.08E-04	3.06E-12	5.81E-08
Kendall	τ	0.24	0.37	0.28
	p-test	1.73E-03	1.27E-07	1.70E-04
Spearman	ρ	0.33	0.51	0.40
	p-test	2.36E-03	1.39E-07	1.36E-04

Table 1. Parameters ρ and τ as a result of the correlation analysis between discharge boundaries

4.3.2 Correlation between downstream water level and wind

Wind can be separated in its two constituents wind speed and wind direction. The downstream water level has been separated in a storm surge component and an astronomical tide component based on harmonic analysis.

The POT values of wind speed and storm surge have been coupled using a time window of two days around the peak value. The correlation between the downstream boundaries was studied using a correlation analysis. Table 2 gives an overview of the corresponding parameters and the p-values. Considering these values the hypothesis of independence can be rejected at a significance level of $\alpha = 0.05$. The associated meteorological events are correlated.

	ρ and τ	p-test
Pearson	0.60	2.35E-35
Kendall	0.46	1.05E-36
Spearman	0.65	1.82E-42

Table 2. Parameters ρ and τ as a result of the correlation analysis between downstream water level and wind speed

Figure 7 shows all coupled and uncoupled extreme values for both wind speed and storm surge. The correlation between the coupled extreme values is clearly visible, but

both variables also contain a considerable amount of uncoupled values. Hence, since both variables strengthen each other's effect on the water levels, relating wind speed and storm surge by means of a regression approach would imply an overestimation of the water levels.

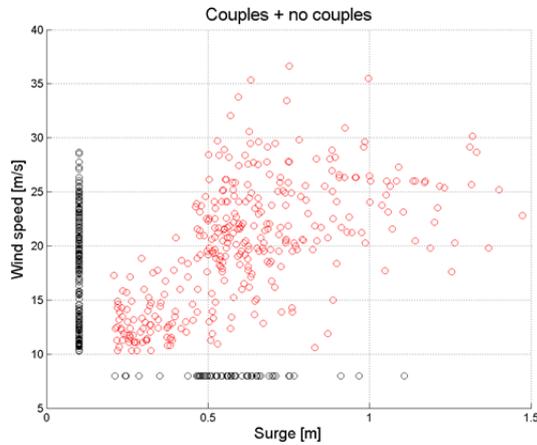


Figure 7. Overview of coupled and uncoupled extremes of storm surge and wind speed

Wind direction is another important component that might have a considerable influence on the extreme water levels. Figure 8 summarizes the results of a sensitivity analysis in which wind direction was varied while keeping the other variables constant. Based on the sensitivity analysis 4 dominant wind directions (N, NW, W and SW) have been identified with a significant effect on the water levels.

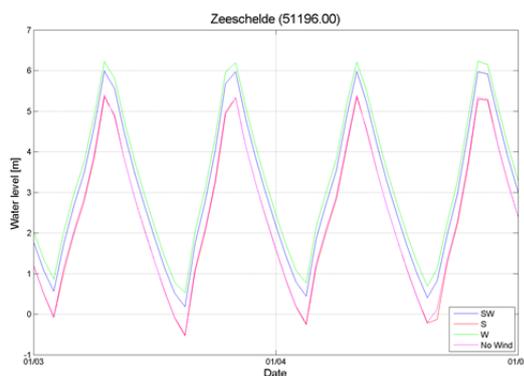


Figure 8. Sensitivity analysis for the effect of wind direction on simulated water levels.

4.3.3 Correlation between upstream discharge and downstream water level

The correlation between the upstream discharge boundaries and the downstream water level was investigated by means of a correlation analysis between the runoff of one of the largest catchments (Demer) and the storm surge in Vlissingen. The POT values of both variables have been coupled using a time window of two days around the peak value. Table 3 provides an

overview of the parameters and the p-values from the correlation analysis on coupled values. Considering these values the hypothesis of independence cannot be rejected at a significance level of $\alpha = 0.05$.

	ρ and τ	p-test
Pearson	0.06	0.64
Kendall	0.03	0.76
Spearman	0.05	0.69

Table 3. Parameters ρ and τ as a result of the correlation analysis between downstream water level and upstream discharge

The results of the correlation analysis do not exclude the possibility that a downstream extreme storm surge coincides with an upstream extreme discharge event. Such an event has a potential high impact in the Scheldt estuary. Figure 9 shows the coupled and uncoupled values of storm surge and upstream discharge. Although the values of the extremes cannot be correlated, the frequency of the bivariate domain is not to be neglected considering the potential impact of such events.

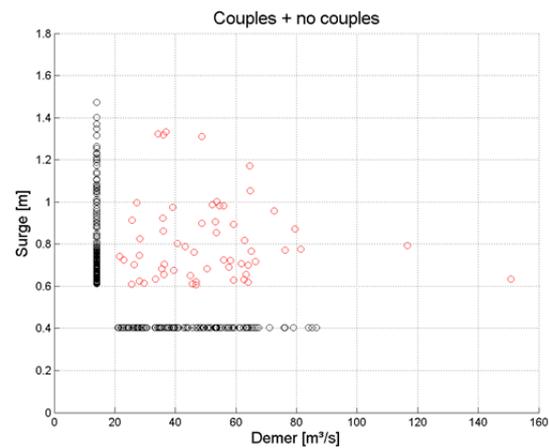


Figure 9. Overview of coupled and uncoupled extreme values for storm surge and discharge in Demer

4.4 Copula

4.4.1 Multivariate approach

Based on the correlation analysis between the different variables it can be concluded that the conditional approach, with all variables correlated, might result in an oversimplified representation of reality, resulting in an overestimation of the water levels.

Hence the multivariate approach is recommended. Considering the amount of variables involved, a multivariate approach with a joint distribution for all variables however will result in an impractical number of events. Therefore a compromise approach has been designed that uses a nested Copula structure that involves

conditional Gumbel Copulas and an independent Copula. A Copula is a function that joins or couples one-dimensional marginal distribution functions ([5] and [6]). The Gumbel Copula was first discussed by Gumbel [7] and is given by:

$$C(u, v) = \exp(-[(-\ln(u))^\alpha + (-\ln(v))^\alpha]^{1/\alpha}) \quad (3)$$

With C the Copula of variables u and v and $\alpha \in [1, \infty]$.

The independent Copula implies that the variables u and v are considered independent and α equals 1. The resulting Copula is given by:

$$C(u, v) = u * v \quad (4)$$

4.4.2 Stratification of a Copula

A bivariate Copula combines 2 univariate distributions into a multivariate distribution. To estimate the dependency, couples of simultaneous events are needed. Because univariate threshold excess models are only valid above their threshold, only couples of extremes that simultaneously exceed their corresponding thresholds can be used. This will leave some extremes without a corresponding extreme of the other variable. The greater the correlation between the two investigated variables, the smaller this number of uncoupled POT. Figure 10a gives the domains in which the couples and the uncoupled POT's of variable 1 and 2 occur. Domains 2 and 3 in Figure 10a contain events of, respectively, variable 1 and 2 with an associated value for the second variable below the optimal threshold.

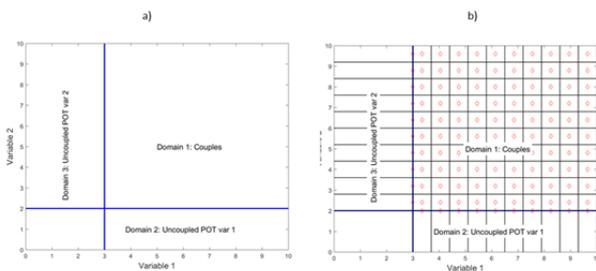


Figure 10. Domain of occurrence of couples and uncoupled POT. b) Stratified domain of occurrence of couples and uncoupled POT

The exceedance probability of the joint occurrences in a certain point of domain 1 can be calculated by the Copula. If both variable 1 and variable 2 are subdivided into 10 classes, 121 cells and 121 (11^2) exceedance probabilities will be obtained. The probability of occurrence per year of each class is:

$$P(i, j) = \frac{k_1}{A_1} (C(i + 1, j + 1) - C(i, j + 1) - C(i + 1, j) + C(i, j)) \quad (5)$$

Where C is the exceedance probability of the Copula, k is the number of couples and A is the amount of years of data. The value of variable 1 and 2 in the middle of each class will be representative to that class (red dots in Figure 10b). This yields a stratified two dimensional space with known frequencies.

4.4.3 Copula between upstream discharge and downstream storm surge

Based on the correlation analysis between upstream discharge and the downstream storm surge it was concluded that the values of both variables are independent but that the frequency of the bivariate domain cannot be neglected due to the potentially high impact of events in this domain. This can be modelled by means of an independent Copula which implies that we consider the possibility of coincidental joint occurrences, without supposed dependency between the variables.

This Copula is considered the central Copula in the nested structure and contains three domains. Domain 1 is the bivariate domain. Stratification of this domain results in events with a known discharge and downstream storm surge (Figure 11). All other variables (wind speed and direction, other upstream discharge values) are determined by relating them to the main variables.

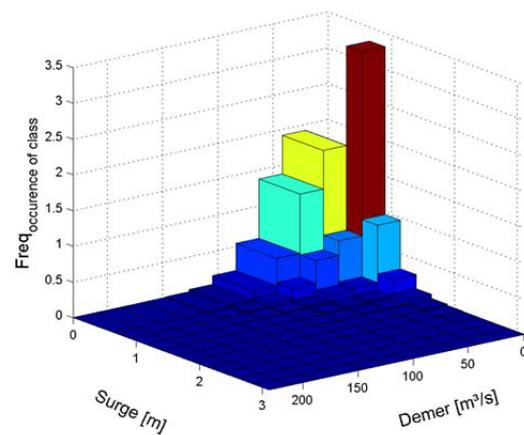


Figure 11. Stratification of Copula between discharge Demer and downstream storm surge

4.4.4 Copula between upstream discharges

Domain 2 of the main Copula (Figure 13) is the univariate domain of extreme discharges. Correlation analysis revealed that the upstream discharge values are not independent. In order to determine the joint probabilities of all variables we could therefore assume full dependency. This might however result in an overestimation of the resulting water levels considering the extent of the Scheldt estuary. Local rainfall patterns imply e.g. that a peak discharge in the river Dender (western part of the basin) does not always coincide with

peak discharges on the river Demer (eastern part of the basin). Determining the joint probabilities by a partial dependency is a more accurate representation of reality.

Therefore it was decided to further refine domain 2 of the main Copula by calculating the joint probabilities of extreme Dender discharge and extreme Demer discharge with a Gumbel Copula (Figure 12). The joint probability of the 40 other upstream discharges are considered fully dependent to the extreme discharge of the Dender or the Demer and the associated univariate extreme value distributions are stratified by a dependant stratification with the most correlated discharge of both.

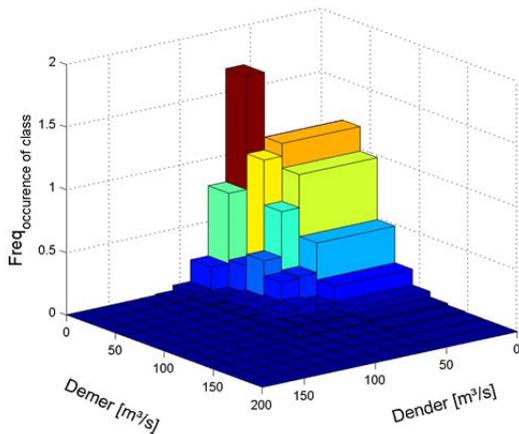


Figure 12. Stratification of Copula between discharge Demer and discharge Dender

4.4.5 Copulas between downstream storm surge and wind speed

Domain 3 of the main Copula is the univariate domain of extreme storm surges. Correlation analysis revealed that the downstream storm surge is not independent from the wind speed. But full dependency would again result in an overestimation of the water levels corresponding to a certain probability of exceedance. In addition it was concluded that wind direction plays an important role.

Therefore it was decided to refine domain 3 of the main Copula by modelling four independent Gumbel Copulas between wind speed and storm surge and a univariate distribution of storm surge. The four Gumbel Copulas correspond to the four dominant wind directions that influence water levels in the Scheldt estuary (N, NW, W and SW). The extreme values for storm surge and wind speed for each of the directions have been derived by means of a directional extreme value analysis. The univariate distribution of the storm surge corresponds to the extreme values that are not related to any of the dominant wind directions.

4.4.6 Nested Copula structure

The eventual nested Copula structure is shown schematically in Figure 13. From the central Copula only the events in the bivariate domain (domain 1) will be calculated. The events in the univariate domains (domain 2 and domain 3) can be replaced by the more detailed synthetic events and associated probabilities corresponding to the Gumbel Copulas between, respectively, the upstream discharges and between the upstream discharge and the downstream storm surge.

The calculated frequency of occurrence for the Gumbel Copula between the upstream discharges (Formula 3) has to be rescaled in order to comply with the nested Copula structure. This adjustment is necessary since the frequencies of occurrence from the Gumbel Copula are calculated using the entire domain of the univariate distribution of Demer. A subdomain of this is however taken into account in domain 1 of the central Copula. Without adjustments a part of the probability domain of Demer will be counted twice. The frequency of occurrence of domain 1 and domain 3 of the Copula between the upstream boundaries is accordingly recalculated using the following formula:

$$Frq_h(i) = Frq(i) * \left(\frac{Frq_{domain2,us-ds}}{Frq_{domain1,ds} + Frq_{domain3,us}} \right) \quad (6)$$

With $Frq_h(i)$ the rescaled frequency of occurrence for event i , $Frq(i)$ the frequency of occurrence based on the bivariate Copula domain, $Frq_{domain2,us-ds}$ the sum of the frequencies from domain 2 of the central Copula, and $Frq_{domain1,us}$ and $Frq_{domain3,ds}$ the sum of frequencies from respectively domain 1 and domain 3 of the Copula between the upstream boundaries. The frequencies of domain 2 of the Copula between the upstream boundaries are not rescaled since these events are not coupled with the univariate domain of Demer.

In a similar way the frequencies for the Gumbel Copula's between downstream storm surge and wind speed are rescaled using the following formula:

$$Frq_h(i) = Frq(i) * \left(\frac{Frq_{domain3,us-ds}}{Frq_{ds,Copulas} + Frq_{ds,Univariat}} \right) \quad (7)$$

With $Frq_h(i)$ the rescaled frequency of occurrence for event i , $Frq(i)$ the frequency of occurrence based on the bivariate Copula domain, $Frq_{domain3,us-ds}$ the sum of the frequencies from domain 3 of the central Copula, $Frq_{ds,Copulas}$ the sum of frequencies from the Copulas between storm surge and wind speed for the four dominant wind directions and $Frq_{ds,Univariat}$ the sum of frequencies from the univariate distribution of storm surge for the remaining wind directions.

The frequency of occurrence for the bivariate domain in the central Copula can be determined using formula 4. Calculating synthetic events in the central domain of the

Copula implies the assumption of full dependency between all upstream discharge values and between wind speed and storm surge. This is a necessary assumption to keep the number of events feasible with respect to the hydrodynamic computations, but has also implications for the calculated frequency of occurrence. The calculated frequencies for the Copula has been determined using the entire domain of the univariate distribution of the discharge in the river Demer. As a consequence of the dependent stratification between the discharges in the rivers Demer en Dender only events from the bivariate domain from the Copula between these discharges have been calculated. Likewise, wind speed for these events is calculated by means of a regression formula. Since not the entire domains of the univariate distributions of discharge in the river Demer and storm surge in Vlissingen is calculated, the related frequencies of occurrence will be overestimated without rescaling. Therefore the frequencies of domain 1 from the central Copula have to be rescaled using the following formula:

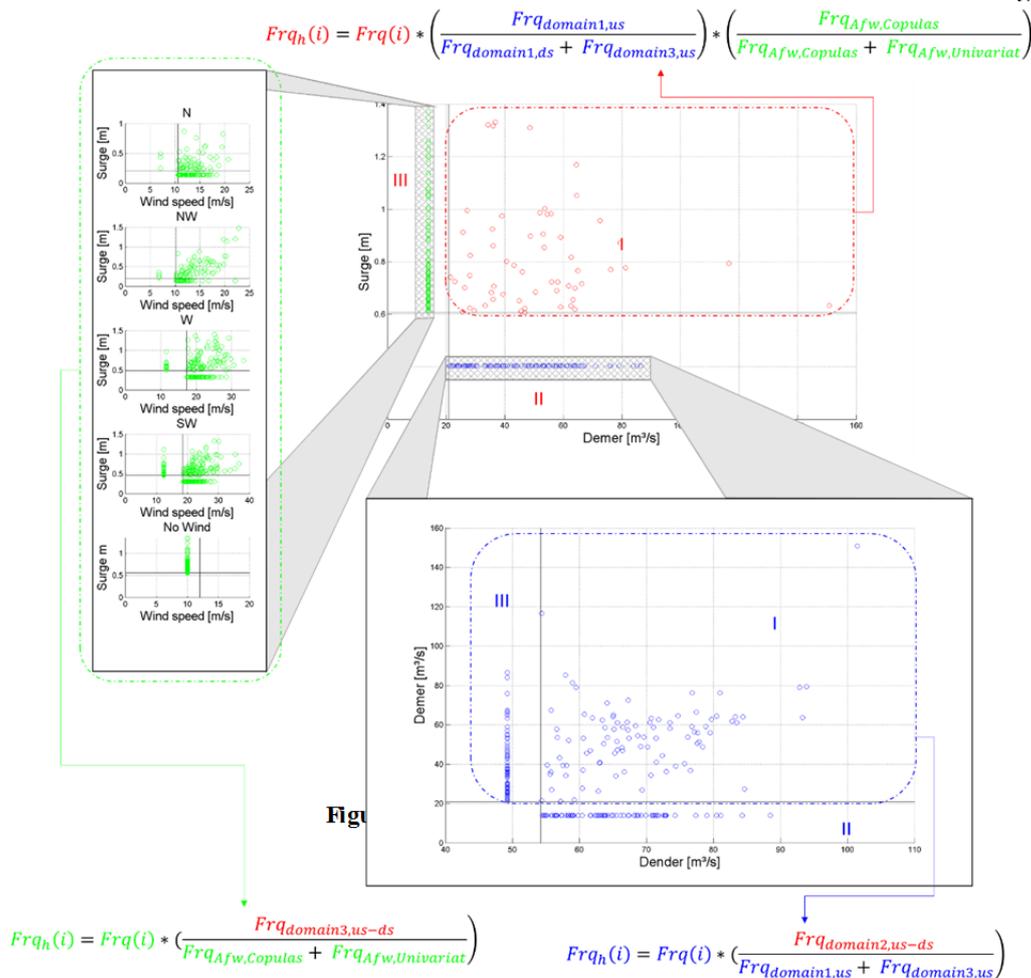
$$Frq_h(i) = Frq(i) * \left(\frac{Frq_{domain1,us}}{Frq_{domain1,us} + Frq_{domain3,us}} \right) * \left(\frac{Frq_{ds,Copulas}}{Frq_{ds,Copulas} + Frq_{ds,Univariat}} \right) \quad (8)$$

With $Frq_h(i)$ the rescaled frequency of occurrence for event i , $Frq(i)$ the frequency of occurrence based on the bivariate Copula domain, $Frq_{domain1,us}$ and $Frq_{domain3,us}$ the sum of frequencies from, respectively, domain 1 and domain 3 of the Copula between the upstream boundaries, $Frq_{ds,Copulas}$ the sum of frequencies from the Copulas between storm surge and wind speed for the 4 dominant wind directions, and $Frq_{ds,Univariat}$ the sum of frequencies from the univariate distribution of storm surge for the remaining wind directions.

5. Synthetic events

5.1 Normalized storm profile

Synthetic events are derived from the combination of normalized unit profiles and the representative values of the different Copula domain classes. These normalized profiles are computed by scaling the most extreme recorded historical events between 0 and 1. A normalized unit profile is then computed as the mean profile of these scaled profiles. The number of extreme events is a compromise between a focus on extreme behaviour, by decreasing the number of extreme events, and reducing the influence of random behaviour of a single event, thus



increasing the number of extreme events.

This approach has been repeated at each model boundary and for each different hydrometeorological variable.

5.2 Stratification of normalized profile

The normalized unit profile was computed as a mean time profile $m(t)$ and a standard deviation $s(t)$ of the scaled events. To account for variations of the synthetic events, 5 hydrographs are generated in each discharge class:

- $m(t) - 2s(t)$: normalized unit profile minus 2 times standard deviation
- $m(t) - s(t)$: normalized unit profile minus standard deviation
- $m(t)$: normalized unit profile
- $m(t) + s(t)$: normalized unit profile plus standard deviation
- $m(t) + 2s(t)$: normalized unit profile plus 2 times standard deviation

This yields another stratification, with 5 profile classes. Corresponding probabilities are computed under the assumption of a normal distribution, with first moment $\mu = m(t)$ and second moment $\sigma = s(t)$ for each t , as explained in Figure 14. The assumption of normality is reported to be conservative [8].

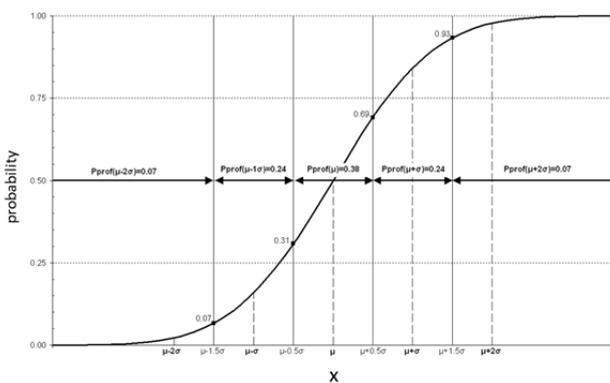


Figure 14. Stratification of hydrograph profile variation

6. Synthetic event frequencies

6.1 Synthetic event frequencies

For each class i 5 different synthetic events were generated according to the profile variation, by stratification of the normal variation into 5 classes k . Hence the probability of a single synthetic event yields:

$$P_{synth_{i,k}} = P\Delta Q_i * Pprof_k \quad (9)$$

With $P\Delta Q_i$ computed as shown in Figure 13, and $Pprof_k$ computed as shown in Figure 14.

6.2 Statistical uncertainty

As part of the extreme value analysis, confidence bounds are computed by means of the bootstrap resampling technique. This interval represents the statistical uncertainties of the GPD parameter estimation. The statistical uncertainty can be implemented in the stratification by recalculating the class probabilities $P\Delta Q_i$ for the upper and lower limit of the 95% confidence interval of the GPD. Eventually this yields a 95% upper and lower confidence bound of the empirical frequency distribution of inundation depths. Since the statistical uncertainty only affects the class probabilities, leaving the synthetic time profiles unchanged, no additional hydrodynamic runs are required to calculate statistical uncertainties for the empirical distribution function of the flood consequences.

7. Impact and validation

Eventually a set of boundary conditions is obtained with a known frequency of occurrence. Each event is a set of time series for every model boundary (Figure 15). The amount of events is determined by the nested Copula structure, the stratification resolution of each domain and the variation of the time unit profile. In this case 1920 synthetic events are generated.

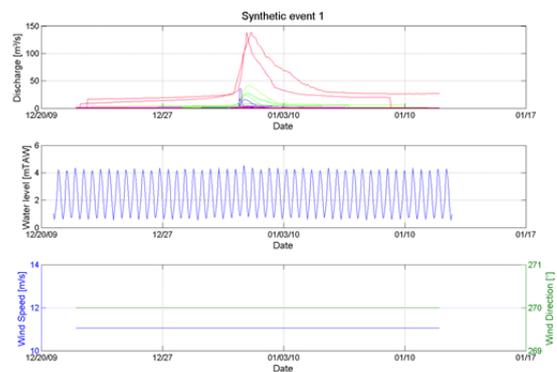


Figure 15. Set of boundary conditions for a synthetic event

The impact of each event is calculated using a hydrodynamic 1D model of the Scheldt estuary. Hence an empirical frequency distributions of the impact, i.e. water levels in the flood plains and calculation nodes, can be drawn by sorting the maximum inundation depths of the synthetic events in descending order and accumulating the corresponding probabilities $P_{synth_{i,k}}$.

In order to verify the reliability of the methodology, a long term hydrodynamic simulation of 28 years is used as a reference. This simulation results in the empirical frequency distribution of inundation depths which can be

compared with the frequency distribution of the synthetic event runs.

The comparison between the resulting frequency distributions of the synthetic events and the reference events was repeated for several locations along the river. Examples are shown in Figure 16 and Figure 17 for an upstream and a downstream location, revealing a satisfactory correspondence between the empirical distribution function of the synthetic events, with 95% confidence interval, and the empirical frequencies of the reference run. Similar results were obtained for the verification of the empirical distribution of accumulated storage along the river, making the methodology appropriate for e.g. design purposes of retention zones.

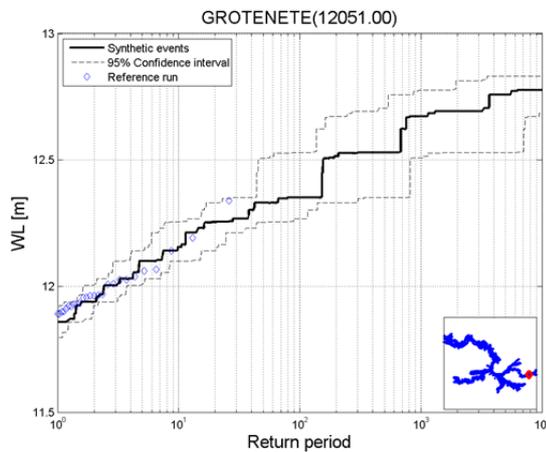


Figure 16. Frequency distribution of water levels in an upstream node of the model. Comparison between synthetic events and the reference run

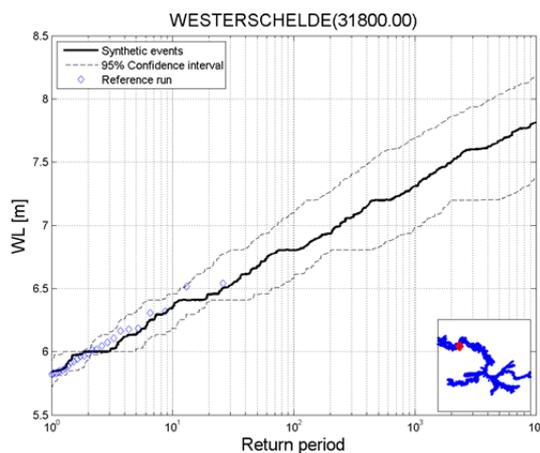


Figure 17. Frequency distribution of water levels in a downstream node of the model. Comparison between synthetic events and the reference run

8. Conclusions

In this paper a methodology has been presented for probabilistic assessment of flood risk in the Scheldt

estuary. This includes the generation of synthetic events for hydrodynamic simulation with multiple boundary conditions which can influence each other.

Based on a correlation analysis between the different upstream and downstream boundaries, a nested Copula structure was developed and stratified to sample synthetic events. The nested structure of bivariate domains is a trade-off between a full multivariate approach and a conditional approach. The nested Copula structure ensures that the combination between upstream and downstream events is considered without compromising on the level of detail of the description of upstream or downstream events, while keeping the number of simulations acceptable with respect to impact model calculations. The resulting methodology appears to be an efficient technique to cover the complete frequency domain of interest, reducing the number of samples by taking into account the probabilities of the classes. Furthermore, statistical uncertainties can be taken into account without additional hydrodynamic computations.

The methodology has the advantage that it aims at determining a statistical distribution of consequences. Hence, this makes the methodology particularly appropriate for flood risk assessment and related applications.

9. References

1. Blanckaert, J., Swings, J., Franken, T. and Gullentops, C. (2012). Stratified Sampled Synthetic Hydrographs for Flood Risk Calculations.
2. Willems, P. (2013). Waterloopmodellering : algemene methodologie voor numerieke modellering van waterlopen.
3. Beirlant, J., Goegebeur, Y. and Segers, J. (2004). Statistics of Extremes: Theory and Applications
4. Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values.
5. Li, D. X. (2000). On Default Correlation: A Copula Function Approach. *Springer*.
6. Nelsen, B. R. (2006). An Introduction to Copulas. *Springer New York*.
7. Gumbel, É. J. (1960). Distributions des Valeurs Extremes en Plusieurs Dimensions. *ResearchGate 9*
8. Singh, V. P. and Chowdhury, P. K. (1985). On Fitting Gamma Distribution to Synthetic Runoff Hydrographs. *ResearchGate 16*